

## Estimating Uncertain Flow and Transport Parameters Using a Sequential Uncertainty Fitting Procedure

K. C. Abbaspour,\* C. A. Johnson, and M. Th. van Genuchten

### ABSTRACT

Inversely obtained hydrologic parameters are always uncertain (nonunique) because of errors associated with the measurements and the invoked conceptual model, among other factors. Quantification of this uncertainty in multidimensional parameter space is often difficult because of complexities in the structure of the objective function. In this study we describe parameter uncertainties using uniform distributions and fit these distributions iteratively within larger absolute intervals such that two criteria are met: (i) bracketing most of the measured data (>90%) within the 95% prediction uncertainty (95PPU) and (ii) obtaining a small ratio (<1) of the average difference between the upper and lower 95PPU and the standard deviation of the measured data. We define a model as calibrated if, upon reaching these two criteria, a significant  $R^2$  exists between the observed and simulated results. A program, SUFI-2, was developed and tested for the calibration of two bottom ash landfills. SUFI-2 performs a combined optimization and uncertainty analysis using a global search procedure and can deal with a large number of parameters through Latin hypercube sampling. We explain the above concepts using an example in which two municipal solid waste incinerator bottom ash monofills were successfully calibrated and tested for flow, and one monofill also for transport. Because of high levels of heavy metals in the leachate, monitoring and modeling of such landfills is critical from environmental points of view.

**I**NVERSE MODELING is becoming increasingly popular in many branches of the earth and environmental sciences. One major reason to apply inverse modeling is to estimate parameters that cannot be measured directly because of a variety of reasons, including scale issues. What makes inverse modeling attractive is that various output variables are often much easier to measure than some of the required input parameters. Numerous applications of inverse modeling now exist in the literature. These include estimation of soil hydraulic parameters (e.g., Pan and Wu, 1998; Simunek et al., 1999; Abbaspour et al., 2001; Young et al., 2002; Wang et al., 2003), hydrologic parameters (e.g., Duan et al., 1992; Beven and Binley, 1992; Lahkim et al., 1999; Savenkoff et al., 2001), and transport parameters (e.g., Schmied et al., 2000; von Gunten and Furrer, 2000; McKenna et al., 2001; Bell et al., 2002), as well as parameters in many other applications such as flood discharge prediction (Sulzer et al., 2002), barotropic ocean tide simulations (Egbert

and Erofeeva, 2002), and seawater intrusion (Iribar et al., 1997).

Inverse modeling has brought new opportunities, but also new challenges. Some of the advantages are savings in time and cost of laboratory and/or field experiments generally needed to obtain the unknown parameters and attainment of a better fit with available data. Another advantage is the usefulness of inverse modeling in the analysis of model structure (the invoked conceptual model), boundary conditions, and prevailing subsurface flow and contaminant transport processes. One of the limitations of inverse modeling is that the fitted parameters are conditioned on the experimental setup and its set of measured variables, which is usually limited in both time and space. Other conditioning factors include the choice of the parameter estimation routine, the objective function, and the weights associated with the different components or parts of the form of objective function.

The main challenge of inverse modeling is to find that "unique" set of effective parameters that best describe the data and envisioned processes. However, many stochastic inverse methods are based on the philosophy that there is no such unique set of effective parameter values. Beven and Binley (1992), among others, explain that since the invoked hydrological model structure usually contains some error, and since all observations and measurements on which inverse modeling is based are also subject to error, no reason exists why any one set of parameter values will represent the sought after effective parameters in a unique way. Hence, the nonuniqueness problem is an inherent characteristic of inverse modeling regardless of the calibration procedure or experimental setup.

Furthermore, with field-scale time series data, as opposed to relatively simple well-defined laboratory solute breakthrough curves for example (e.g., van Genuchten, 1981), definition of a good simulation can be very subjective because visually different simulations with very different parameter combinations may yield identical values of the objective function (Boyle et al., 2000). In such cases, parameter fitting leads to parameter regions with similar values of the objective function. These parameter regions comprise the uncertainty in the parameters. Quantifying the degree of uncertainty in the parameters leads to more reliable models (Cooley, 1993). Uncertainty regions of potential candidates for effective parameter values can be used to produce prediction uncertainties (Konikow, 1986, p. 183). Or vice versa, as is done in our study, if a desired prediction uncertainty

K.C. Abbaspour and C.A. Johnson, Swiss Federal Inst. for Environmental Science and Technology (EAWAG), Ueberlandstrasse 133, P.O. Box 611, CH-8600 Dübendorf, Switzerland; M. Th. van Genuchten, USDA-ARS, George E. Brown, Jr. Salinity Lab., 450 West Big Springs Road, Riverside, CA 92507. Received 12 Dec. 2003. Original Research Paper. \*Corresponding author (abbaspour@eawag.ch).

Published in Vadose Zone Journal 3:1340–1352 (2004).

© Soil Science Society of America

677 S. Segoe Rd., Madison, WI 53711 USA

**Abbreviations:** EC, electrical conductivity; GLUE, Generalized Likelihood Uncertainty Estimation; MSWI, municipal solid waste incinerator; SSQ, sum of square errors; SUFI, Sequential Uncertainty Fitting; 95PPU, 95% prediction uncertainty.

level is obtained, all parameter distributions producing this uncertainty are potential solutions. In this paper, we outline a procedure that uses a stopping rule based on a certain definition of the prediction uncertainty.

Inverse modeling usually reduces to an optimization problem in which a vector of unknown parameters is estimated by minimizing an assumed goal or objective function. Many procedures have been developed to solve the resulting nonlinear minimization problem (Brad, 1974; Beck and Arnold, 1977; Yeh, 1986; Kool et al., 1987; Duan et al., 1992; Yapo et al., 1998; Vrugt et al., 2003a,b). Among soil scientists and vadose zone hydrologists the gradient approach of Levenberg-Marquardt (Marquardt, 1963) method has become the method of choice (e.g., van Genuchten 1981; Kool et al., 1985, 1987), while modelers of larger-scale problems prefer using more global approaches and multiobjective optimization functions (e.g., Beven and Binley, 1992; Duan et al., 1992, 2003; Yapo et al., 1998; Gupta et al., 1998).

A common problem with gradient methods is that they are very sensitive to the initial values of the optimized parameters and often lead to a local minimum. Also, their application to problems having a large number of parameters is sometimes problematic. Moreover, gradient methods do not provide a reliable estimate of parameter uncertainty and, hence, prediction uncertainty. Global procedures such as the shuffled complex (Duan et al., 1992) and genetic algorithms (Wang, 1997) generally have the drawback of being inefficient and requiring too many iterations. As an example, Gupta et al. (2003) reported that a complete single-criterion optimization problem using the shuffle complex model as reported in Sorooshian et al. (1993) would require as many as 10 000 function calls, while 100 Pareto solutions will require approximately one million function calls. In a multiobjective formulation, a Pareto solution is a solution that cannot be improved without disadvantaging at least one of the objectives. To overcome some of these problems we developed a procedure that consists of a sequence of steps in which the initial (large) uncertainty in the model parameters is progressively reduced until certain calibration criteria for prediction uncertainty are met. The program uses an efficient sampling procedure (Latin hypercube), along with a global search algorithm that examines the behavior of an objective function by analyzing the Jacobian and Hessian matrices. The current algorithm is henceforth referred to as SUFI-2, as it is the second version of a previously developed Sequential Uncertainty FITting (SUFI) program (Abbaspour et al., 1997).

In this study, the SUFI-2 procedure was applied to two municipal solid waste incinerator (MSWI) bottom ash monofills. Landfills of this type have recently been investigated by, among others, Hjelmar (1996), Belevi et al. (1992), Johnson et al. (1998), and Hartmann et al. (2001). Bottom ash in these landfills is typically composed of equal amounts of fine ash material and melted components (of which one-half have crystallized) and small quantities of metallic components, ceramics, and stones (Kirby and Rimstidt, 1993). The Lostorf landfill, a MSWI bottom ash monofill in Switzerland, was stud-

ied in detail by Johnson et al. (1999). Chemical analyses of leachate from this landfill at discrete time intervals between 1994 and 1996 showed average total concentrations of 44.5, 47.1, 11.8, 0.63, 8.2, and 12.4 mM for Na, Cl, K, Mg, Ca, and  $\text{SO}_4$ , respectively. Many other metals, such as Cu, Zn, Sb, Cr, Cd, Mo, V, Mn, and Pb, were also detected. While the leachate composition was found to be relatively constant during dry periods, considerable dilutions occurred during rain events. The relatively good reproducibility of the experimental observations in response to rain events motivated us to use transport modeling to predict heavy metal concentrations.

Different models have been used to simulate flow through MSWI bottom ash landfills (Guyonnet et al., 1998; Hartmann et al., 2001; Johnson et al., 2001). In the study of Johnson et al. (2001), flow through the Lostorf landfill was modeled using the programs MACRO (Jarvis, 1994), HYDRUS5 (Vogel et al., 1996), a neural network approach (Schaap and Bouten, 1996), and a linear storage model adapted from Huwe et al. (1994). They found that flow was dominated by preferential paths. The best simulation results were obtained with the variably saturated dual-permeability model MACRO of Jarvis (1994).

We extend the work of Johnson et al. (2001) by simulating both flow and transport at the Lostorf landfill site. Solute transport was simulated in terms of the electrical conductivity (EC). A second bottom ash landfill, Seckenberg, was also calibrated for flow. This was done by linking SUFI-2 with MACRO. Our objectives were to test the SUFI-2 procedure by calibrating and testing the MACRO model using hourly discharge and EC data.

## THEORY

### Simulation Program MACRO

MACRO is a dual-permeability model based on the concept of having two different domains: a micropore or matrix domain, in which flow is governed by the Richards equation, and a macropore domain that drains by gravity flow. The Richards equation for one-dimensional isothermal flow in a variably saturated rigid porous medium is given by

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z} \left[ K \left( \frac{\partial h}{\partial z} - 1 \right) \right] \pm S_w \quad [1]$$

where  $\theta$  is the volumetric water content ( $\text{L}^3 \text{L}^{-3}$ ),  $K$  is the hydraulic conductivity ( $\text{L T}^{-1}$ ),  $h$  is the pressure head (L),  $S_w$  is a term signifying a source or a sink ( $\text{T}^{-1}$ ), and  $z$  is depth (L) positive downward. For the macropore region, a simplified approach is used in which fluxes are predicted by assuming no change in  $h$  with depth (i.e.,  $\partial h / \partial z = 0$ ), thus reflecting laminar flow under gravity.

The source-sink term  $S_w$  in Eq. [1] is used here to account for water exchange between the micropores and macropores. The flux between the macropores to the micropores is given by

$$S_w = \left( \frac{3D_w \gamma_w}{d^2} \right) (\theta_b - \theta_{mi}) \quad [2]$$

where  $d$  is the effective diffusion path length (L),  $\theta_b$  is the saturated water content ( $\text{L}^3 \text{L}^{-3}$ ),  $\theta_{mi}$  is the actual water content in the micropores ( $\text{L}^3 \text{L}^{-3}$ ),  $D_w$  is the effective water diffusivity ( $\text{L}^2 \text{T}^{-1}$ ), and  $\gamma_w$  is a scaling factor introduced to match the approximate and exact solutions of the diffusion problem. The

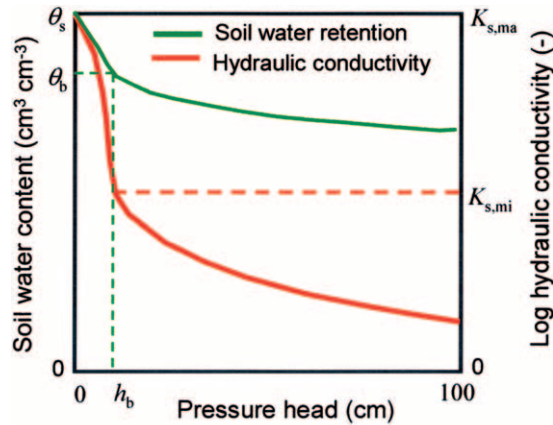


Fig. 1. Water retention and hydraulic conductivity curves as used in MACRO.

scaling factor varies somewhat with the initial water content and hydraulic properties, but not strongly (Gerke and van Genuchten, 1993). This parameter is set to an average value of 0.8 in MACRO. The model assumes that water is immediately routed from the micropores to the macropores when water content exceeds saturation,  $\theta_b$ . Dispersion in the macropores is neglected since advection is assumed to dominate solute transport.

Different functions are used in MACRO to describe the water retention characteristics of the micropore and macropore regions, as illustrated in Fig. 1. For the micropores, the water retention function,  $h_{mi}(\theta)$ , is given by the modified Brooks and Corey (1964) function:

$$h_{mi}(\theta) = h_b S_{mi}^{-1/\lambda} \quad \theta_r \leq \theta \leq \theta_b \quad [3]$$

where

$$S_{mi} = \frac{\theta - \theta_r}{\theta_b - \theta_r} \quad \theta_r \leq \theta \leq \theta_b \quad [4]$$

and  $h_b$  is the critical boundary head (L) at which micropores drain (also known as the air-entry value),  $\theta_r$  is the residual water content, and  $\lambda$  is the pore size distribution index. The hydraulic conductivity function for the micropore region,  $K_{mi}$ , is derived from the retention curve using Mualem's (1976) model, yielding

$$K_{mi} = K_{s,mi} S_{mi}^{n+2+2/\gamma} \quad [5]$$

where  $K_{s,mi}$  is the hydraulic conductivity ( $L T^{-1}$ ) at the critical boundary head  $h_b$ , and  $n$  is the tortuosity factor of the micropore region.

Because flow in the macropore domain is assumed to be driven by gravity only, no retention curve is required. The hydraulic conductivity of the macropores,  $K_{ma}$ , is given by a simple power function of the degree of saturation in the macropores ( $S_{ma}$ ):

$$K_{ma} = K_{s,ma} S_{ma}^{n^*} \quad [6]$$

where  $\theta_s$  is the saturated water content,  $n^*$  is an empirical parameter (unit-less) describing the tortuosity in macropores, and  $S_{ma}$  is defined as

$$S_{ma} = \frac{\theta_{ma}}{\theta_s - \theta_b} \quad [7]$$

where  $\theta_{ma}$  is the water content in macropores.

Solute transport in MACRO is described with the standard advection-dispersion equation for nonsorbing tracers:

$$\frac{\partial \theta c}{\partial t} = \frac{\partial}{\partial z} \left( D \theta \frac{\partial c}{\partial z} - qc \right) \pm U \quad [8]$$

where  $c$  is the solute concentration of the liquid phase ( $M L^{-3}$ ),  $q$  is the Darcian water flux ( $L T^{-1}$ ),  $U$  is a source-sink term ( $M L^{-3} T^{-1}$ ), and  $D$  is the dispersion coefficient ( $L^2 T^{-1}$ ) given by

$$D = D_v v + D_0 f^* \quad [9]$$

where  $D_0$  is the diffusion coefficient in free water ( $L^2 T^{-1}$ ),  $f^*$  is a constant impedance factor,  $D_v$  is the dispersivity (L), and  $v$  is the pore water velocity ( $L T^{-1}$ ). Dispersion in the macropores is set to zero since solute transport is assumed to be dominated by advection. The source-sink term  $U$  accounts for diffusive and advective mass transfer between the two flow domains as follows:

$$U = \left( \frac{3D_e \theta_{mi}}{d^2} \right) (c_{ma} - c_{mi}) + S_w c' \quad [10]$$

where the prime indicates solute concentration in either the macropores or micropores depending upon the direction of flow (the sign of  $S_w$ ), and  $D_e$  is an effective diffusion coefficient ( $L^2 T^{-1}$ ) given by

$$D_e = D_0 f^* S_{ma} \quad [11]$$

Infiltration from the soil surface into the micropores is assumed to be limited by the saturated conductivity,  $K_{s,mi}$ , of the micropore region. The portion of rain or irrigation exceeding  $K_{s,mi}$  flows directly into the macropores. We calculated the solute concentration of water infiltrating into the macropores,  $c_{ma}^*$ , by assuming complete mixing of the infiltrating water with water stored in a shallow "mixing depth,"  $z_d$ :

$$c_{ma}^* = \frac{Q_d(t+\Delta t) + R c_r}{R + z_d \theta_1} \quad [12]$$

where  $Q_d$  is the amount of solute stored in the mixing depth ( $M L^{-2}$ ),  $R$  is the amount of rain reaching the surface within time interval  $\Delta t$  (L),  $c_r$  is the concentration in the rain water ( $M L^{-3}$ ), and  $\theta_1$  is the water content in the layer  $z_d$ . The amount of solute added or removed from the micropores in the top layer within time interval  $\Delta t$ ,  $Q_{mi(l)}$ , is calculated as the difference:

$$Q_{mi(l)} = R c_r - I_{ma} c_{ma}^* \quad [13]$$

where the  $I_{ma}$  denotes infiltration in macropores (L).

## Inverse Procedure SUFI-2

We assume that inversely obtained parameters are always uncertain and that an inverse optimization approach should seek the smallest possible parameter uncertainty or parameter range that satisfies a desired prediction uncertainty. The SUFI-2 procedure described below uses elements of Generalized Likelihood Uncertainty Estimation (GLUE) (Beven, 1989, Beven and Binley, 1992) with elements of a gradient approach (Kool and Parker, 1988), modified to enable a global search. Similar to GLUE, the procedure is applied to parameter sets, rather than to individual parameter values, so that any interactions between parameters are taken into account explicitly.

Also similar to GLUE, our final objective is not necessarily to find a set of best-fit parameters, even though such a best-fit parameter set is calculated. Unlike GLUE, however, the key output of SUFI-2 is a "best range" for each parameter. Parameter combinations within the parameter ranges are ensured to produce high quality simulations because of the two calibration criteria discussed below.



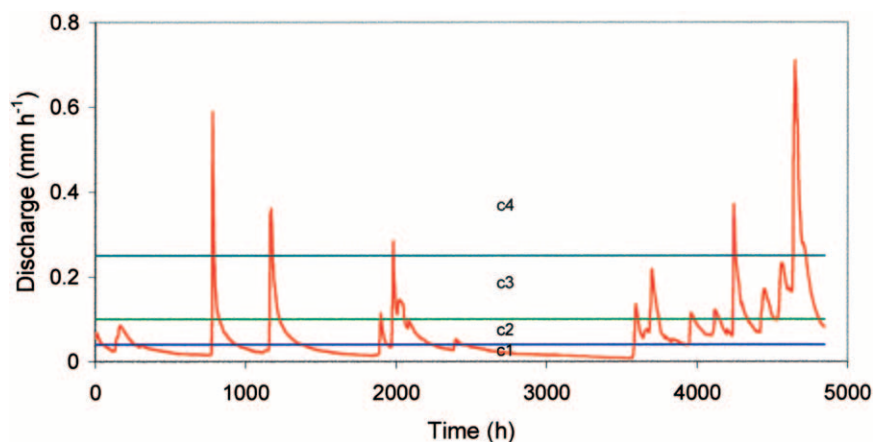


Fig. 2. An example showing the partitioning of a response signal into different compartments: base flow (c1), the recession or rise (c2), intermediate peaks (c3), and large peaks (c4). The weights are calculated such that each compartment contributes equally to the objective function.

Starting with some initial parameter value ranges, SUFI-2 is iterated until (i) the 95% prediction uncertainty (95PPU) between the 2.5th and 97.5th percentiles bracket more than 90% of the measured data and (ii) the average distance between the 2.5th and 97.5th percentiles is smaller than the standard deviation of the measured data. A model is considered calibrated if, upon reaching the above two criteria, a significant  $R^2$  exists between the best simulation and calibration data. Further testing of the calibrated parameter ranges, as performed in this study, should increase model confidence. If calibration cannot be attained with the above criteria, the model structure should be reexamined.

## Step-by-Step Application of SUFI-2

### Step 1

In this first step an objective function is defined. The literature shows many different ways of formulating an objective function (e.g., Legates and McCabe, 1999; Gupta et al., 1998). Each formulation may lead to a different result; hence, the final parameters and parameter ranges are always conditioned on the form of the objective function. To overcome this problem, some studies (e.g., Yapo et al., 1998) combine different types of statistics together (e.g., RMSE, absolute difference, logarithm of differences,  $R^2$ ,  $\chi^2$ ) to yield a “multicriteria” formulation. The use of multiobjective formulation (Duan et al., 2003; Gupta et al., 1998), where different variables, including different variables at different locations, are embedded in the objective function, is also important to reduce the non-uniqueness.

Since there is no unique formulation of the objective function that can be applied universally to all situations, the choice of an objective function must correspond to the goal of the project. In this study, we required a good overall simulation of the peaks as well as the base flow; hence, we used the formulation of Boyle et al. (2000) and partitioned the response function into four compartments, as illustrated in Fig. 2. This formulation, henceforth referred to as “multicompartment” formulation, depicts base flow (c1 in Fig. 2), rise and recession (c2), intermediate peaks (c3), and large peaks (c4). For each compartment, the RMSE is calculated and weighted so they have equal contributions to the objective function. This formulation has proven to provide an overall good fit if the model has no conceptual limitation to simulating each compartment of the response function (i.e., not being able to simulate large peaks due to the absence of a macropore flow component). If simulation of the large peaks is important, the peak region

could be given a larger weight. If two variables such as discharge and concentration are considered simultaneously, then a multiobjective, multicompartment formulation with eight RMSE values is defined, in which case the objective function becomes

$$g(\mathbf{b}) = \sum_{i=1}^I w_i \text{RMSE}_i \quad [14]$$

where  $\mathbf{b}$  is the parameter vector,  $I$  (here equal to 8) is the total number of compartments, and  $w_i$  is calculated here as

$$w_i = \frac{\text{avg}(\text{RMSE})_1}{\text{avg}(\text{RMSE})_i} \quad i = 1, \dots, I \quad [15]$$

where the average (avg) is computed for all simulations within a Latin hypercube sampling round (see Step 5),  $w_i$  is always equal to 1, and

$$\text{RMSE}_i = \sqrt{\frac{1}{k} \sum_{j=1}^k (q^o - q^s)^2_j} \quad [16]$$

where  $q$  is a measured variable,  $k$  is the number of observations in the  $i$ th compartment, and the superscripts “o” and “s” refer to observed and simulated values, respectively. Equation [15] implies that the weights are different for each sampling round.

### Step 2

The second step establishes physically meaningful absolute minimum and maximum ranges for the optimized parameters. There is no theoretical basis for excluding one specific distribution. However, because of lack of information, we assume that all parameters were uniformly distributed within the region bounded by minimum and maximum values. Because the parameter ranges play a constraining role, they should be as large as possible, yet physically meaningful:

$$b_j: b_{j,\text{abs\_min}} \leq b_j \leq b_{j,\text{abs\_max}} \quad j = 1 \dots m \quad [17]$$

where  $b_j$  is the  $j$ th parameter and  $m$  is the number of parameters to be estimated.

### Step 3

This step involves an optional, yet highly recommended sensitivity analysis for all parameters. We pose that no automated optimization routine can replace the insight from physical understanding and knowledge of the effects of parameters on system response. The sensitivity analysis is carried out by

keeping all parameters constant to realistic values, while varying each parameter within the range assigned in Step 1. For each parameter, about 10 simulations are conducted by simply dividing the absolute ranges in equal intervals and allowing the midpoint of each interval to represent that interval. Plotting results of these simulations along with the observations on the same graph gives insight into the effects of the parameters on observed signals.

#### Step 4

Initial uncertainty ranges are assigned to parameters for the first round of Latin hypercube sampling; that is,

$$b_j: [b_{j,\min} \leq b_j \leq b_{j,\max}] \quad j = 1, m \quad [18]$$

In general, the above ranges are smaller than the absolute ranges, are subjective, and depend on experience. The sensitivity analysis in Step 3 can provide a valuable guide for selecting appropriate ranges. Although important, these initial estimates are not crucial since they are updated and allowed to change within the absolute ranges.

#### Step 5

A Latin hypercube (McKay et al., 1979) sampling results in  $n$  parameter combinations, where  $n$  is the number of desired simulations. This number should be relatively large ( $\approx 1000$ – $2000$ ). The simulation program is run  $n$  times and the simulated output variable(s) of interest, corresponding to the measurements, are saved.

#### Step 6

As a first step in assessing the simulations, RMSEs are calculated followed by calculation of the weights and the objective function for each simulation.

#### Step 7

In this step a series of measures is calculated to evaluate each sampling round. First, the sensitivity matrix or Jacobian of  $g(\mathbf{b})$  is computed using

$$J_{ij} = \frac{\Delta g_i}{\Delta b_j} \quad i = 1, \dots, C_2^n, \quad j = 1, \dots, m \quad [19]$$

where  $C_2^n$  is the number of rows in the Jacobian (equal to all possible combinations of two simulations), and  $j$  is the number of columns (number of parameters). Following the Gauss–Newton method and neglecting the higher-order derivatives of the Hessian matrix,  $\mathbf{H}$ , the Hessian matrix of  $g(\mathbf{b})$  is calculated as

$$\mathbf{H} = \mathbf{J}^T \mathbf{J} \quad [20]$$

According to the Cramer–Rao theorem (Press et al., 1992), an estimate of the lower bound of the parameter covariance matrix,  $\mathbf{C}$ , is calculated from

$$\mathbf{C} = s_g^2 (\mathbf{J}^T \mathbf{J})^{-1} \quad [21]$$

where  $s_g^2$  is the variance of the objective function. The estimated standard deviation and 95% confidence interval of a parameter  $b_j$  is calculated from the diagonal elements of  $\mathbf{C}$  (Press et al., 1992) from

$$s_j = \sqrt{C_{jj}} \quad [22]$$

$$b_{j,\text{lower}} = b_j^* - t_{v,0.025} s_j \quad [23]$$

$$b_{j,\text{upper}} = b_j^* + t_{v,0.025} s_j \quad [24]$$

where  $b_j^*$  is the current best estimate of parameter  $b$  (i.e., that parameter value that produces the smallest value of the objective function), and  $v$  is the degrees of freedom ( $n - m$ ). Parameter correlations can be assessed using the diagonal and off-diagonal terms of the covariance matrix as follows:

$$A_{ij} = \frac{C_{ij}}{\sqrt{C_{ii} C_{jj}}} \quad [25]$$

It is important to note that the correlation matrix  $\mathbf{A}$  quantifies the change in the objective function as a result of a change of parameter  $i$ , relative to changes of the other parameters  $j$ .

Parameter sensitivities,  $S$ , are calculated next by averaging the columns of the Jacobian matrix as expressed by

$$S_j = \bar{b}_j \frac{1}{C_2^n} \sum_{i=1}^{C_2^n} \left| \frac{\Delta g_i}{\Delta b_j} \right| \quad j = 1, \dots, m \quad [26]$$

where  $\bar{b}_j$  is the average value of the  $j$ th parameter. We emphasize that the measures of sensitivity given by Eq. [26] are different from the sensitivities calculated in Step 3. The sensitivities given by Eq. [26] are estimates of the average changes in the objective function resulting from changes in each parameter, while all other parameters are changing. Therefore, the sensitivities in Eq. [26] are relative sensitivities. By contrast, the sensitivities in Step 3 define the absolute sensitivity of a parameter that can change when other parameters take on different optimized values.

#### Step 8

In this step the 95PPUs of all predicted variables are computed using the  $n$  Latin hypercube simulations. Distribution of the predicted variables is not always Gaussian, and could be highly skewed. Thus, the usual calculation of uncertainty limits as a function of the variance of the predicted values is not applicable (Beven and Binley, 1992). In this study we represent the prediction uncertainty as the 2.5th ( $q_L$ ) and 97.5th ( $q_U$ ) percentiles of the cumulative distribution of every simulated point.

The goodness of fit is calculated for each variable  $q$  from the percentage of measured data that fall within the 95PPU region (prediction uncertainty), the  $R^2$  between the optimized and observed data, and the average distances  $\bar{d}$  between the upper and lower 95PPU as determined from

$$\bar{d} = \frac{1}{K} \sum_{l=1}^K (q_U - q_L)_l \quad [27]$$

in which  $l$  is a counter, and  $K$  is the total number of observations for variable  $q$ . The best outcome is that 100% of the measurements fall within the 95PPU,  $R^2$  is close to 1, and the  $\bar{d}$  value is close to zero. However, because of measurement errors and model uncertainties, generally this will not be the case.

We consider a model satisfactorily calibrated if approximately 90% of the measured data are within the 95PPU when  $\bar{d}$  is smaller than the standard deviation of the measured data and when  $R^2 > 0.8$ .

#### Step 9

Because in general parameter uncertainties are large initially, the value of  $\bar{d}$  tends to be quite large in the first sampling round when about 100% of the data are within the 95PPU. Hence, further sampling rounds are needed with updated parameter ranges calculated from:

**Table 1. Data periods for calibration and testing.**

Landfill	Calibration <sup>†</sup>	Testing
	(yr-mo-d-h)	
Lostorf	1996-04-22-13 to 1996-11-10-20	1995-01-02-00 to 1995-03-31-13
Seckenberg	1998-04-01-00 to 1998-12-03-23	1999-01-01-00 to 1999-05-26-20

<sup>†</sup> The year with larger amount of data was chosen for calibration.

$$b'_{j,\min} = b_{j,\text{lower}} - \text{Max} \left[ \frac{(b_{j,\text{lower}} - b_{j,\min})}{2}, \frac{(b_{j,\max} - b_{j,\text{upper}})}{2} \right]$$

$$b'_{j,\max} = b_{j,\text{upper}} + \text{Max} \left[ \frac{(b_{j,\text{lower}} - b_{j,\min})}{2}, \frac{(b_{j,\max} - b_{j,\text{upper}})}{2} \right] \quad [28]$$

where  $b'$  indicates updated values. The above criteria, while producing smaller parameter ranges, ensure that the updated parameter ranges are always centered on the current best estimates. If the best estimates are close to their limits, parameter ranges are increased while not exceeding the absolute boundaries. In the final step, parameters are ranked according to their sensitivities, and highly correlated parameters are identified. Of the highly correlated parameters, those with the smaller sensitivities should be fixed to their best estimates and removed from additional sampling rounds.

This procedure is continued until the two stopping rules for model prediction uncertainty are satisfied. The resulting parameter ranges are considered the best-fit parameters for the problem being investigated. If the stopping rules cannot be met (i.e., a small value of  $\bar{d}$ , but too many observed data are excluded from the 95PPU), the model cannot be calibrated and the structure of the model and its boundary conditions need to be reexamined.

## APPLICATION

### Sites and Data

The SUFI-2 sequential uncertainty fitting procedure was applied to the two MSWI landfills in Switzerland. The Lostorf landfill, a MSWI bottom ash monofill near Buchs AG, Switzerland, is situated in an unused gravel pit. The Seckenberg landfill is situated on sandstone and dolomite formations on a hillside above Frick, Switzerland. For the Lostorf landfill, measured discharge and EC data were available for several months in 1995 and 1996. For the Seckenberg landfill only flow data were available for a few months in 1997 and 1998. We refer the reader to Johnson et al. (1999, 2001) for more complete descriptions of the two landfills.

Hourly values of rainfall, global radiation, relative humidity, air pressure, wind velocity at 2 m, and temperature were obtained from the Swiss Meteorological Institute at the Buchs/Suhr and Frick/Rheinfelden sites for application to Lostorf and Seckenberg, respectively. Potential evapotranspiration rates from landfill surfaces were estimated using the hourly form of the FAO Penman–Monteith equation (Allen et al., 1994), but modified to obtain actual evaporation rates for both landfill systems according to procedures described by Johnson et al. (2001).

**Table 2. Estimated flow and transport parameters for the Lostorf and Seckenberg landfills**

Parameter <sup>†</sup>	Pseudo value	Initial range	Optimized range	Optimized value	Sensitivity ranking
$c_1$ -1, mS mm <sup>-1</sup>	0.767	0.1–2	0.57–0.67	0.634	31
$c_1$ -2	0.974	0.5–1.5	0.83–1.09	1.06	35
$c_1$ -3	1.39	1.4–1.5	1.38–1.40	1.39	1
$\theta_1$ -1, cm <sup>3</sup> cm <sup>-3</sup>	14.23	17–23	14.7–15.3	14.86	3
$\theta_1$ -2	12.14	17–23	13.1–15.4	14.4	16
$\theta_1$ -3	16.56	19–25	16.4–17.6	16.63	8
$c_r$ , mS mm <sup>-1</sup>	0.5	0.2–1	0.47–0.54	0.5	29
$D_v$ , mm	97.4	10–100	95.4–102.3	95.7	14
$d$ -1, mm	111.55	20–100	106.4–114.9	113.5	18
$d$ -2	126.52	20–100	124.7–134.2	129.6	17
$d$ -3	3.70	20–100	2.7–3.9	3.3	37
$h_0$ -1, cm	18.92	5–25	18.1–19.2	18.69	11
$h_0$ -2	7.13	5–25	6.0–7.0	6.31	30
$h_0$ -3	15.28	5–25	14.4–15.5	15.42	19
$K_s$ -1, mm h <sup>-1</sup>	276.20	150–500	266.7–300.2	280.44	25
$K_s$ -2	476.43	150–500	432.3–468.9	468.7	20
$K_s$ -3	142.31	150–500	138.8–147.6	140.4	12
$K_{sm}$ -1	1.075	0–2	0.99–1.16	1.019	33
$K_{sm}$ -2	0.017	0–1	0.014–0.02	0.019	36
$K_{sm}$ -3	0.132	0–1	0.12–0.154	0.151	34
$\theta_1$ -1, cm <sup>3</sup> cm <sup>-3</sup>	14.23	12–16.5	12.0–14.0	13.2	32
$\theta_1$ -2	12.14	12–16.5	12.1–12.6	12.2	6
$\theta_1$ -3	16.56	12–18.5	16.7–17.8	17.2	13
$\theta_1$ -1	30.34	27–30	29.9–31.3	30.2	7
$\theta_1$ -2	27.92	27–30	27.5–29.9	27.7	21
$\theta_1$ -3	27.66	26–30	27.5–28.3	27.9	2
$\theta_1$ -1, cm <sup>3</sup> cm <sup>-3</sup>	25.84	23–26.9	25.5–26.3	25.62	4
$\theta_1$ -2	22.27	23–26.9	22.1–23.4	22.90	10
$\theta_1$ -3	24.65	23–25.9	24.1–24.9	24.69	5
$\lambda$ -1	0.183	0.1–0.4	0.17–0.20	0.181	26
$\lambda$ -2	0.174	0.1–0.4	0.17–0.20	0.192	28
$\lambda$ -3	0.426	0.1–0.4	0.47–0.51	0.492	15
$n$ -1	0.071	0.01–0.15	0–0.04	0.024	38
$n$ -2	0.152	0.01–0.15	0.15–0.164	0.160	24
$n$ -3	0.097	0.01–0.15	0.084–0.095	0.093	27
$n^*$ -1	3.691	1–4	3.5–3.9	3.77	22
$n^*$ -2	2.767	1–4	2.7–2.84	2.83	9
$n^*$ -3	0.682	1–4	0.62–0.69	0.67	23

<sup>†</sup> Numbers 1, 2, and 3 refer to soil layers.

**Table 3. Estimated flow and transport parameters for the Lostorf and Seckenberg landfills.**

Parameters†	Optimized calibration range	
	Lostorf	Seckenberg
$c_1$ -1, mS mm <sup>-1</sup>	0.6–1.2	—
$c_1$ -2	0.6–1.0	—
$c_1$ -3	1.3–1.4	—
$\theta_1$ -1, cm <sup>3</sup> cm <sup>-3</sup>	0.18–0.20	0.24–0.26
$\theta_1$ -2	0.16–0.19	0.22–0.25
$\theta_1$ -3	0.20–0.22	0.22–0.24
$c_2$ , mS mm <sup>-1</sup>	0.22–0.52	—
$D_1$ , mm	57–108	—
$d$ -1, mm	84–112	66–74
$d$ -2	94–137	38–50
$d$ -3	2–27	33–49
$h_1$ -1, cm	14.8–20.3	19.9–21.8
$h_1$ -2	2.9–8.1	10.6–15.3
$h_1$ -3	13.6–19.1	10.1–13.8
$K_1$ -1, mm h <sup>-1</sup>	248–409	318–352
$K_1$ -2	400–510	183–238
$K_1$ -3	135–265	421–494
$K_{s,mi}$ -1	0.88–1.47	1.1–1.3
$K_{s,mi}$ -2	0.0–0.16	0.0–0.003
$K_{s,mi}$ -3, mm h <sup>-1</sup>	0.08–0.26	0.11–0.26
$\theta_1$ -1, cm <sup>3</sup> cm <sup>-3</sup>	0.13–0.15	0.13–0.15
$\theta_1$ -2	0.12–0.13	0.18–0.19
$\theta_1$ -3	0.15–0.18	0.16–0.18
$\theta_2$ -1	0.29–0.31	0.28–0.29
$\theta_2$ -2	0.28–0.29	0.28–0.30
$\theta_2$ -3	0.27–0.28	0.27–0.28
$\theta_3$ -1	0.26–0.27	0.25–0.26
$\theta_3$ -2	0.22–0.24	0.23–0.25
$\theta_3$ -3	0.24–0.25	0.23–0.24
$\lambda$ -1	0.15–0.25	0.11–0.13
$\lambda$ -2	0.10–0.26	0.29–0.39
$\lambda$ -3	0.34–0.45	0.30–0.36
$n$ -1	0.04–0.07	0.07–0.08
$n$ -2	0.15–0.22	0.037–0.59
$n$ -3	0.08–0.17	0.08–0.09
$n^{*}$ -1	2.4–3.7	3.1–3.5
$n^{*}$ -2	2.6–3.2	2.4–2.6
$n^{*}$ -3	0.46–1.3	1.1–1.25

† Numbers 1, 2, and 3 refer to soil layers.

### Calibration and Testing of the Flow and Transport Model

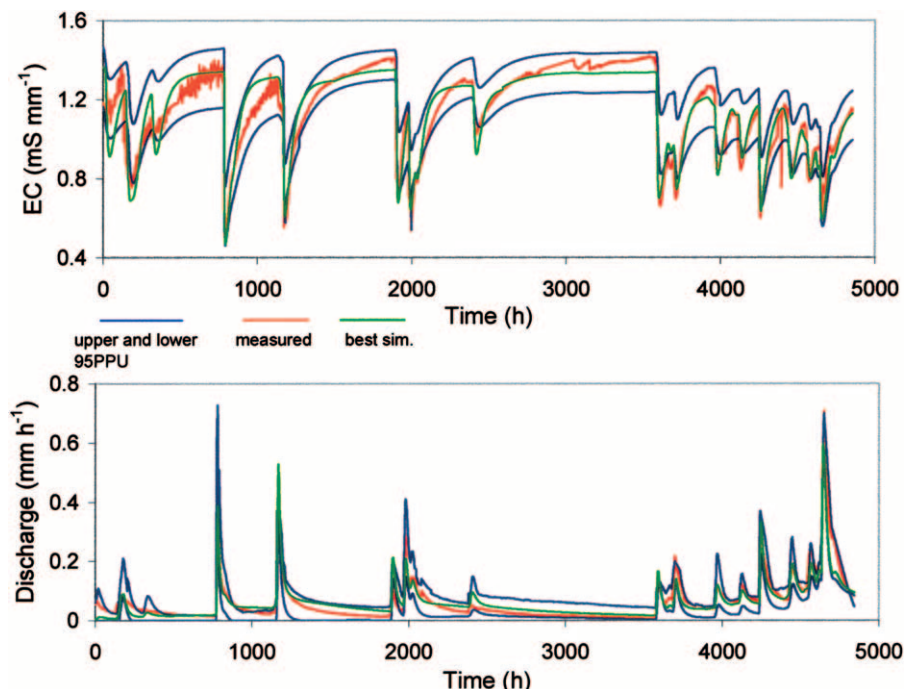
The Lostorf landfill was calibrated using both discharge and EC in the objective function given by Eq. [14] and [15]. The objective function for the Seckenberg landfill included only discharge data. Time periods used for calibration and model testing are given in Table 1.

Based on the filling history of the landfills, a three-layer profile was modeled by treating the parameters listed in Table 2 as unknowns. A common problem of historic landfill databases is that the initial conditions for both the flow and transport simulations are poorly defined. To determine the initial water content and electrical conductivity profiles, we examined two options: (i) simulating the flow regime by letting the landfills drain from a nearly saturated condition until the simulated discharge equaled the observed discharge at time zero (the water content and concentration profiles at this time were assumed the correct initial water content and solute distributions) and (ii) fitting the initial conditions along with the other parameters. Although both options have problems, we opted for the second option since it was more straightforward and produced better results. The main problem with the first option was that flow parameters defining flow rates were still unknown. To reduce the effect of the initial condition, we started the simulations a few months before time zero. Including the initial water content and concentration distributions produced a total of 33 unknown parameters when only flow is considered, and 38 parameters when transport is modeled (Table 2). Transport simulations were performed by using electrical conductivity as a tracer.

The bottom boundary conditions were assumed to have constant hydraulic gradients, while the surface boundaries were assumed to be atmospheric. Rain or irrigation exceeding  $K_{s,mi}$  was allowed to flow directly into the macropores.

### RESULTS AND DISCUSSION

The absolute parameter ranges and the starting initial ranges were selected based on our experience with the



**Fig. 3. Calibration results for flow and transport at Lostorf. The blue curves represent the 95% prediction uncertainty (95PPU), the red curve represents the measured data, and the green curve shows the best fit of optimization.**



**Table 4. Statistics of measured variables and the corresponding modeling results.**

Variable	Number of data points	Average and SD of measured variables	Percent of measured data in the 95PPU <sup>†</sup>	$R^2$ <sup>‡</sup>	Average width of 95PPU and its ratio over SD of measured data
Lostorf discharge calib. test	4855	0.062 (0.077)	90	0.86	0.06 (0.78)
	1600	0.15 (0.14)	91	0.85	0.16 (1.1)
Lostorf EC calib. test	1600	1.15 (0.21)	90	0.88	0.23 (1.1)
	2125	1.02 (0.24)	90	0.82	0.40 (1.7)
Seckenberg discharge calib. test	5927	0.105 (0.16)	93	0.94	0.07 (0.44)
	3500	0.19 (0.33)	89	0.88	0.12 (0.36)

<sup>†</sup> 95PPU, 95% prediction uncertainty.

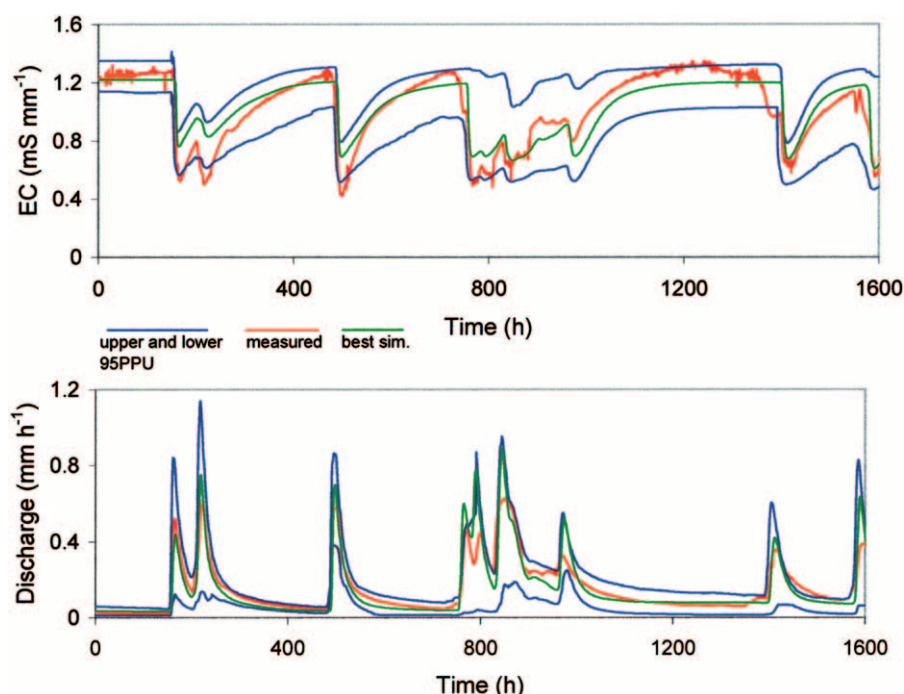
<sup>‡</sup> All regressions were significant at  $P < 0.0001$ .

parameters, literature review (e.g., Dubus and Brown, 2002), and personal communication with the author of MACRO (Jarvis, 1994). We conducted the sensitivity analysis of Step 3 after the first sampling round. We concluded that (i) some parameters affected the early stages of the discharge and concentration (e.g., initial water contents and concentrations), while others affected the later stages (e.g., dispersivity,  $D_v$ ; effective diffusion path length,  $d$ ), (ii) some parameters had a strong effect on the peak discharge (e.g., hydraulic conductivities  $K_{s,mi}$  and  $K_{s,ma}$ ), while others affected mainly base flow (especially the diffusion path length,  $d$ ), and (iii) some parameters affected the recession or timing of the peak discharges (e.g., the pore size distribution index,  $\lambda$ , and the macropore tortuosity index  $n^*$ ), while others affected all compartments as defined in Eq. [15] (especially the saturated hydraulic conductivity,  $K_{s,mi}$ ). We used this sensitivity analysis to obtain better estimates of the initial parameter ranges and to improve our understanding of the effect of all parameters on discharge and EC.

On the basis of some preliminary runs we selected a set of pseudo parameters for the Lostorf landfill and generated a synthetic set of discharge and EC data. We

used these output data in SUFI-2 to examine how well we could recover the pseudo parameters. Table 2 summarizes the list of pseudo parameters, the initial parameter uncertainties, the optimized parameter uncertainties, the optimized best parameter values, and the relative sensitivity ranking. In the fourth Latin hypercube sampling round, 99% of the measured EC and discharge data were bracketed by the 95PPU, while the ratio of average difference between the upper and the lower 95PPU over the standard deviation of generated measured data was 0.83 for EC and 0.53 for discharge, indicating excellent fits. As shown in Table 2, however, some of the insensitive pseudo parameters (i.e.,  $c_i-1$ ) were outside the optimized range, indicating that parameters insensitive to the objective function could not be fitted.

Table 3 summarizes the final optimized parameter uncertainties for the Lostorf landfill, while Fig. 3 presents calibration results for flow and transport at Lostorf. Plots in Fig. 3 show the best simulation results and the upper and lower limits of the 95PPU of the calibrated parameter ranges. Results were obtained after five sampling rounds for a total of 5000 simulations. The 95PPU contained 90% of the measured discharge and EC data (Table 4), while the ratio of the average



**Fig. 4. Test results of flow and transport at Lostorf. The blue curves represent the 95% prediction uncertainty (95PPU), the red curve represents the measured data, and the green curve shows the best fit of optimization.**



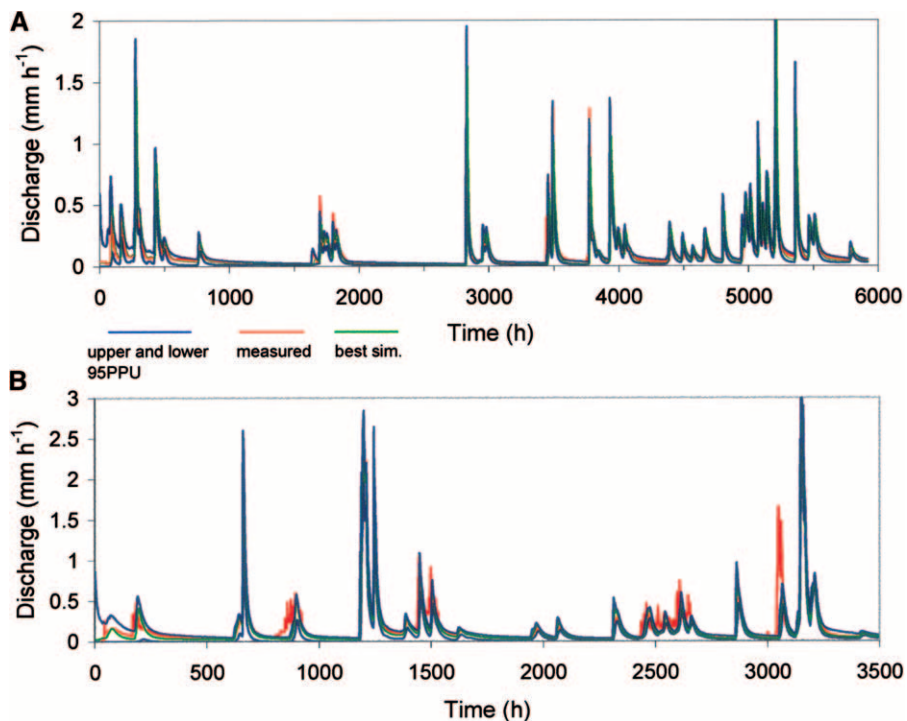


Fig. 5. (A) Calibration and (B) test results for flow at Seckenberg. The blue curves represent the 95% prediction uncertainty (95PPU), the red curve represents the measured data, and the green curve shows the best fit of optimization.

difference between the upper and the lower 95PPU over the standard deviation of the measured data was  $<1$  for discharge and only slightly above 1 for EC. With  $R^2$  values for discharge and EC of 0.86 and 0.88, respectively, almost all calibration requirements for the calibration data set of Lostorf were satisfied.

We tested the above calibration parameters using data from a previous year; results are shown in Fig. 4. The statistics in Table 4 indicate that the three calibra-

tion requirements were again satisfied for discharge, but that the ratio requirement was not met for the EC data. As will be discussed below, this indicates that the calibrated parameter ranges produced several relatively poor simulations for EC. This is a strong indication that a higher degree of multiobjective formulation is needed in the objective function to limit the number of bad simulations.

To test the SUFI-2 procedure against previous results

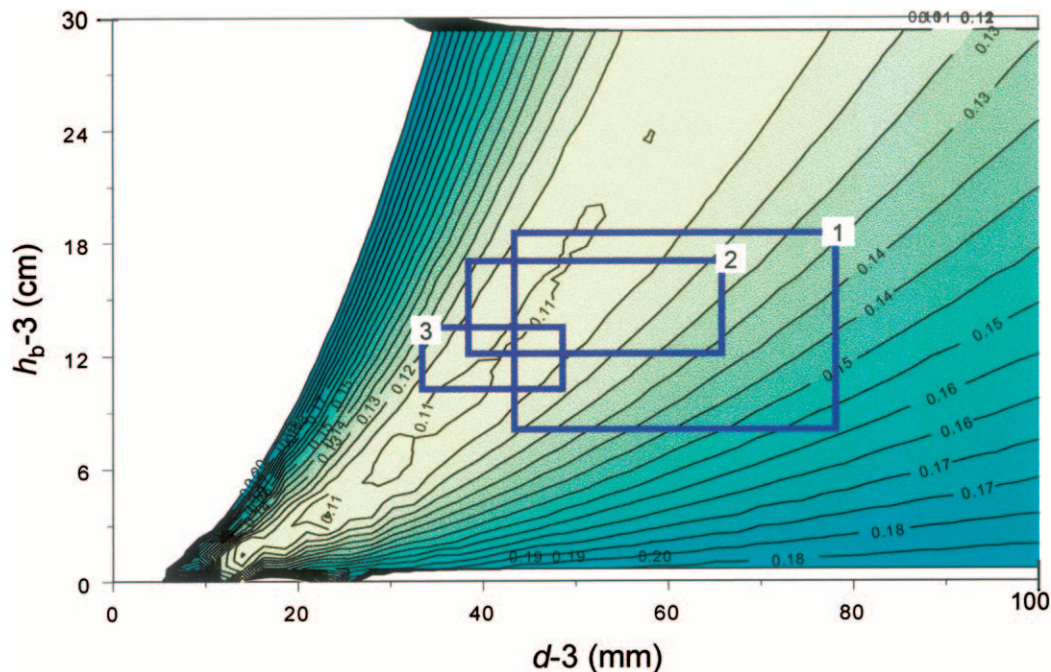


Fig. 6. The  $d-3-h_b-3$  response surface obtained after 2000 simulations. All other parameters were kept constant.

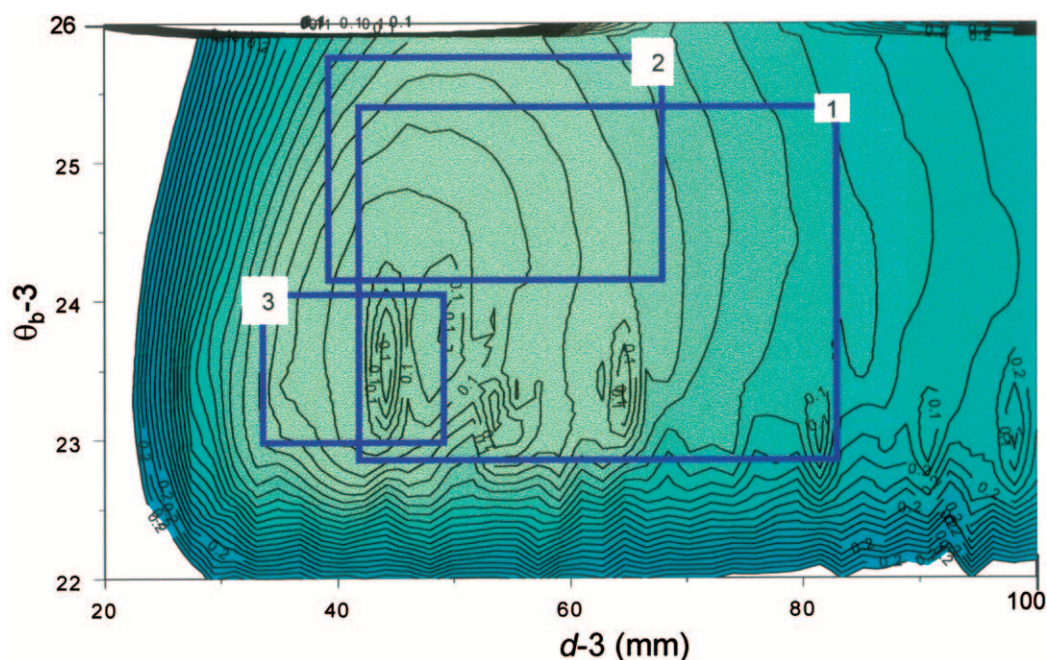


Fig. 7. The  $d-3-\theta_{v-3}$  response surface obtained after 2000 simulations. All other parameters were kept constant.

by Johnson et al. (2001) for this same landfill, we compared the sum of square errors (SSQ) between the measured and simulated results for both calibration and test cases. A direct comparison is not possible because the objective function previously did not include a transport component. The previously obtained SSQs for discharge calibration and test results were 10.1 and 41.1, respectively. The values obtained with SUFI-2 were 4.1 and 7.5, respectively. This shows a significant improvement in best-fit simulations obtained with our new sequential uncertainty fitting procedure.

Figure 5 presents the calibration (A) and test (B)

results for flow through the Seckenberg landfill. The plots show both the best simulation results and the 95PPU based on the calibrated parameter ranges of Table 3. The results were obtained after four sampling rounds. The 95PPU contained 93 and 89% of the measured discharge data for calibration and testing, respectively (Table 4). The ratios of the average difference between the upper and the lower 95PPU over the standard deviation of the measured data were substantially less than one for both the calibration and test data sets. The  $R^2$  requirement was also met with highly significant values. One reason for the better calibration results for the Secken-

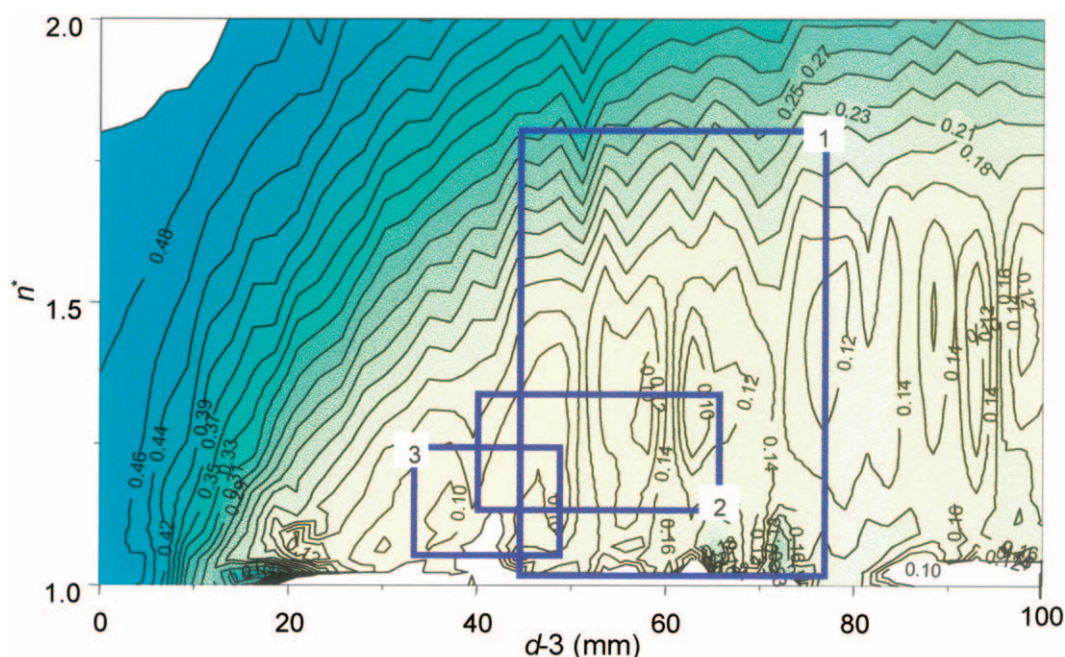


Fig. 8. The  $d-3-n^*$  response surface obtained after 2000 simulations. All other parameters were kept constant.



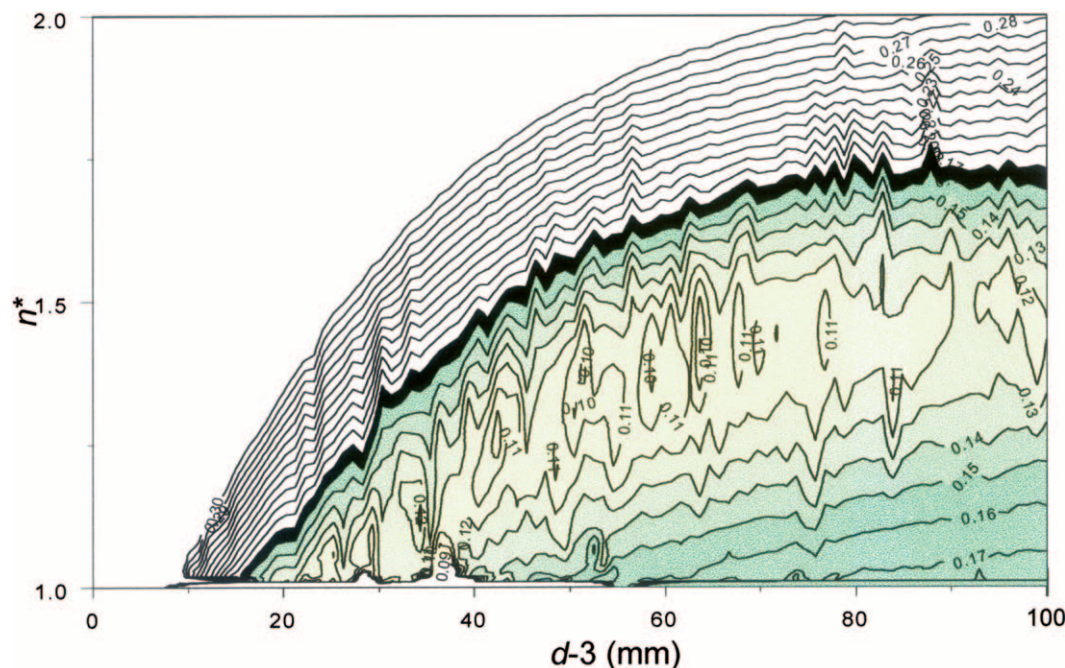


Fig. 9. The  $d$ -3- $n^*$  response surface obtained after 10 000 simulations. All other parameters were kept constant.

berg case is that the objective function contained only discharge data.

A comparison of the parameters for both landfills (Table 3) shows substantial overlaps, with the parameters for Seckenberg generally having smaller ranges. Applying the Lostorf parameter ranges to Seckenberg (except for the initial values which were those of the Seckenberg calibration) and vice versa produced relatively good simulation results, even though not all three calibration criteria were met. This suggests that it may be possible to use identical parameters for similar applications.

To investigate the progression toward convergence and limitations of the calibrated parameter ranges, we plotted the response surfaces for some of the most sensitive parameters. In doing so, all parameters of the Seckenberg example were fixed at their final best values. In the first example diffusion path length,  $d$ , and the boundary soil water tension  $h_b$  of the third layer were varied by drawing 2000 samples using the Latin hypercube procedure. The response surface of the objective function is plotted in Fig. 6. Superimposed on the response surface are the parameter ranges from the last three sampling rounds, delineated by three boxes num-

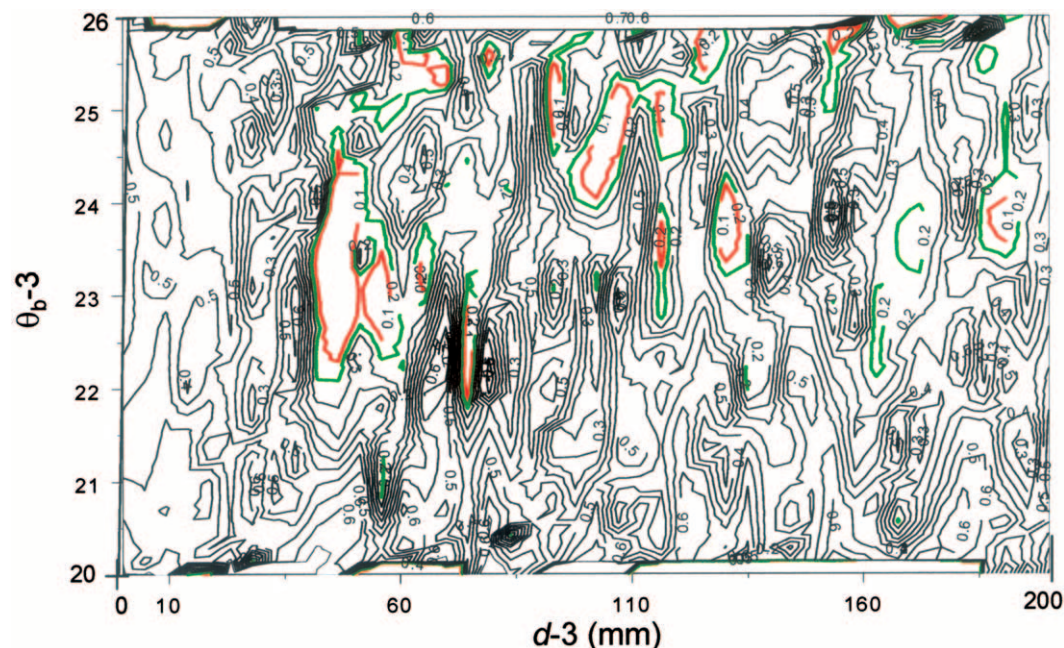


Fig. 10. The  $d$ -3- $\theta_b$  response surface obtained after 10 000 simulations. Parameters  $d$ -3,  $\theta_b$ -3,  $h_b$ -3, and  $n^*$  were varied simultaneously, while all others were kept constant.

bered 1, 2, and 3. The two-parameter  $d$ - $h_b$  response surface shows a clear valley of small objective function values.

The results in Fig. 6 reflect an apparent limitation of expressing the parameters in terms of ranges. The two-parameter surface plot suggests that the uncertainty ranges of both parameters could be much larger, albeit located along a narrow valley. However, if the size of the parameter ranges is increased, the prediction uncertainty may include many poor simulations as was the case with the EC of Lostorf, leading to a high value of the second criterion. For this reason, several (e.g., Yapo et al., 1998) have used the concept of Pareto optimality, in which the simulations are divided into sets of relatively good and poor simulations. The final solution is then not specified by means of a set of parameter distributions, but in terms of parameter combinations that produce good simulations. Still, this approach also does not guarantee that all good simulations can be captured by means of a limited number of function calls.

Additional calculations showed that the two-parameter response surface may be somewhat misleading when assessing the region with minimum values in a multidimensional parameter system. Figure 7 shows the response surface when only  $d$  and  $\theta_b$  are allowed to vary. Clearly, the region with minimum values of the objective function now has a different shape, where many disconnected minima islands exist. Another example with  $d$ - $n^*$  (Fig. 8) shows an even more complicated surface, with numerous local minima.

To verify that 2000 samples will give an accurate response surface, we repeated the  $d$ - $n^*$  simulations with 10 000 samples. Figure 9 shows a plot similar to that of Fig. 8 for the 2000 simulations. Notice that the response surface again contains many minima. If more than two parameters are allowed to vary simultaneously, plots of any combination of two-parameter surfaces will be even more complicated than those in Fig. 8 or 9. Figure 10 shows plots of the surface of the  $d$ - $\theta_b$  response function when  $d$ ,  $n^*$ ,  $\theta_b$ , and  $h_b$  were allowed to vary simultaneously. This graph again illustrates the difficulty of using procedures based on gradient methods to solve optimization problems.

## CONCLUSIONS

The SUFI-2 procedure we developed for inverse modeling uses a sequence of steps in which the initial (large) uncertainties in the model parameters are progressively reduced until a certain calibration requirement based on the prediction uncertainty is reached. From a series of two-parameter response surfaces we observed that also the SUFI-2 program is limited to finding one of many regions with local minima in a multidimensional parameter space. An encouraging observation, however, was that none of the detailed runs produced a region with smaller values of the objective function. The advantages of SUFI-2 are that it combines optimization with uncertainty analysis and can handle a large number of parameters. Moreover, the procedure is quite efficient

in terms of the number of function calls, and it is easy to implement.

Finally, we recall that one of the main purposes of inverse modeling is to use easily measurable model outputs to obtain parameters that are much more difficult and time-consuming to measure directly. In other words, we prefer to feed our inverse optimization procedure with data of relatively low value to obtain data of higher value. Unfortunately, the notion of getting much for little is inherently problematic. Despite the attractions, nature does not allow for such easy transformations. For this reason one must accept that our SUFI-2 procedure, or similar Bayesian inverse methods, will only produce a certain solution space, rather than a single unique solution. Hence, one may very well have to live with the idea that inverse methods will lead to many solutions regardless of the type of the objective function used in the optimization. To limit the nonuniqueness, however, a step forward would be to make the description of the objective function as constraining to parameters as possible by using a combination of multicriteria, multiobjective, and multicompartiment formulation.

## ACKNOWLEDGMENTS

The authors thank the Gemeindeverband Abfallbeseitigung Oberes Fricktal, and in particular Erwin Roth, for their active cooperation, Gholamabbas Sayyad for data processing and initial phases of modeling, Dr. Mark Borsuk for reviewing a preliminary version of the paper, and Dr. Peter Reichert for many fruitful discussions.

## 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.T. van Genuchten. 2001. Estimating unsaturated soil hydraulic parameters using ant colony optimization. *Adv. Water Resour.* 24:827–841.
- Allen, R.G., M. Smith, L.S. Pereira, and A. Perrier. 1994. An update for the calculation of reference evapotranspiration. *ICID Bull.* 43: 35–92.
- Beck, J.V., and K.J. Arnold. 1977. *Parameter estimation in engineering and science*. John Wiley and Sons, New York.
- Belevi, H., D.M. Stämpfli, and P. Baccini. 1992. Chemical behaviour of municipal solid waste incinerator bottom ash in monofills. *Waste Manage. Res.* 10:153–167.
- Bell, L.S.J., P.J. Binning, G. Kuczera, and P.M.H. Kau. 2002. Rigorous uncertainty assessment in contaminant transport inverse modeling: A case study of fluoride diffusion through clay liners. *J. Contam. Hydrol.* 57:1–20.
- Beven, K.J. 1989. Interflow. p. 191–219. *In* H.J. Morel-Seytoux (ed.) *Unsaturated flow in hydrologic modeling*. D. Reidel, Norwell, MA.
- Beven, K., and A. Binley. 1992. The future of distributed models: Model calibration and uncertainty prediction. *Hydrol. Processes* 6:279–298.
- Boyle, D.P., H.V. Gupta, and S. Sorooshian. 2000. Toward improved calibration of hydrological models: Combining the strengths of manual and automated methods. *Water Resour. Res.* 36:3663–3674.
- Brad, Y. 1974. *Nonlinear parameter estimation*. Academic Press, New York.
- Brooks, R.H., and A.T. Corey. 1964. Hydraulic properties of porous media. *Hydrology Paper 3*. Colorado State University, Fort Collins.
- Cooley, R.L. 1993. Exact Scheffe-type confidence interval for output from groundwater flow models. 1. Use of hydrogeologic information. *Water Resour. Res.* 29:17–33.
- Duan, Q., S. Sorooshian, and V. Gupta. 1992. Effective and efficient



- global optimization for conceptual rainfall-runoff models. *Water Resour. Res.* 28:1015–1031.
- Duan, Q., S. Sorooshian, V. Gupta, A.N. Rousseau, and R. Turcotte. 2003. Advances in calibration of watershed models. AGU, Washington, DC.
- Dubus, I.G., and C.D. Brown. 2002. Sensitivity and first-step uncertainty analyses for the preferential flow model MACRO. *J. Environ. Qual.* 31:227–240.
- Egbert, G.D., and S.Y. Erofeeva. 2002. Efficient inverse modeling of barotropic ocean tides. *J. Atmos. Ocean Technol.* 19:183–204.
- Gerke, H.H., and M.Th. van Genuchten. 1993. Evaluation of a first-order water transfer term for variably-saturated dual-porosity flow models. *Water Resour. Res.* 29:1225–1238.
- Gupta, H.V., L.A. Bastidas, J.A. Vrugt, and S. Sorooshian. 2003. Multiple criteria global optimization for watershed model calibration. p. 125–132. *In* Q. Duan et al. (ed.) *Calibration of watershed models*. Water Science and Application 6. AGU, Washington, DC.
- Gupta, H.V., S. Sorooshian, and P.O. Yapo. 1998. Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. *Water Resour. Res.* 34:751–763.
- Guyonnet, D., B. Didier-Guelorget, G. Provost, and C. Feuillet. 1998. Accounting for water storage effects in landfill leachate modeling. *Waste Manage. Res.* 16:285–295.
- Hartmann, F., H.P. Bader, R. Scheidegger, and P. Baccini. 2001. Water transport in a bottom ash landfill from a municipal solid waste ("MSW") incinerator. *J. Solid Waste Technol. Manage.* 27:76–81.
- Hjelmar, O. 1996. Disposal strategies for municipal solid waste incineration residues. *J. Hazard. Mater.* 47:345–368.
- Huwe, B., H. Goelz-Huwe, and J. Eberhardt. 1994. Parameterschaetzung und Modellrechnungen zum Gebietswasserhaushalt kleiner, heterogener Einzugsgebiete mit einfachen Modellkonzepten. *Mitteiln. Dtsch. Bodenkundl. Gesellsch.* 74:273–276.
- Iribar, V., J. Carrera, E. Custodio, and A. Medina. 1997. Inverse modeling of seawater intrusion in the Llobregat delta deep aquifer. *J. Hydrol. (Amsterdam)* 198:226–244.
- Jarvis, N.J. 1994. The Macro model (version 4.0). Technical description and sample simulations. Swedish University of Agriculture Sciences, Department of Soil Science, Uppsala.
- Johnson, C.A., M. Kaeppli, S. Brandenberger, A. Ulrich, and W. Baumann. 1999. Hydrological and geochemical factors affecting leachate composition in municipal solid waste incinerator bottom ash. Part II: The geochemistry of leachate from Landfill Lostorf, Switzerland. *J. Contam. Hydrol.* 40:239–259.
- Johnson, C.A., G.A. Richner, T. Vitvar, N. Schittli, and M. Eberhard. 1998. Hydrological and geochemical factors affecting leachate composition in municipal solid waste incinerator bottom ash. Part I: The hydrology of landfill Lostorf, Switzerland. *J. Contam. Hydrol.* 33:361–376.
- Johnson, C.A., M.G. Schaap, and K.C. Abbaspour. 2001. Model comparison of flow through a municipal solid waste incinerator ash landfill. *J. Hydrol. (Amsterdam)* 243:55–72.
- Kirby, C.S., and J.D. Rimstidt. 1993. Mineralogy and surface properties of municipal solid waste ash. *Environ. Sci. Technol.* 27:652–660.
- Konikow, L.F. 1986. Predictive accuracy of a ground-water model—Lessons from a post audit. *Ground Water* 24:173–184.
- 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. ONESTEP: A nonlinear parameter estimation program for evaluating soil hydraulic properties from one-step outflow experiments. *Bull. 85-3*. Virginia Agric. Exp. Stn., Blacksburg.
- Kool, J.B., J.C. Parker, and M.Th. van Genuchten. 1987. Parameter estimation for unsaturated flow and transport models—A review. *J. Hydrol. (Amsterdam)* 91:255–293.
- Lahkim, M.B., L.A. Garcia, and J.R. Nuckols. 1999. Stochastic modeling of exposure and risk in a contaminated heterogeneous aquifer. 2: Application of Latin hypercube sampling. *Environ. Eng. Sci.* 16: 329–343.
- Legates, D.R., and G.J. McCabe. 1999. Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resour. Res.* 35:233–241.
- Marquardt, D.W. 1963. An algorithm for least-squares estimation of non-linear parameters. *SIAM J. App. Math.* 11:431–441.
- McKay, M.D., R.J. Beckman, and W.J. Conover. 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 21:239–245.
- McKenna, S.A., L.C. Meigs, and R. Haggerty. 2001. Tracer tests in a fractured dolomite. 3. Double-porosity, multiple-rate mass transfer processes in convergent flow tracer tests. *Water Resour. Res.* 37: 1143–1154.
- Mualem, Y. 1976. A new model for predicting the hydraulic conductivity of unsaturated porous media. *Water Resour. Res.* 12:513–522.
- 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.
- Press, W.H., B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling. 1992. *Numerical recipe, the art of scientific computation*. 2nd ed. Cambridge University Press, Cambridge, UK.
- Savenkoff, C., A.F. Vezina, P.C. Smith, and G. Han. 2001. Summer transports of nutrients in the Gulf of St. Lawrence estimated by inverse modeling. *Estuarine Coastal Shelf Sci.* 52:565–587.
- Schaap, M.G., and W. Bouten. 1996. Modeling water retention curves of sandy soils using neural networks. *Water Resour. Res.* 32:3033–3040.
- Schmied, B., K.C. Abbaspour, and R. Schulin. 2000. Inverse estimation of parameters in a nitrogen model using field data. *Soil Sci. Soc. Am. J.* 64:533–542.
- Simunek, J., O. Wendroth, and M.Th. van Genuchten. 1999. Estimating unsaturated soil hydraulic properties from laboratory tension disc infiltrometer experiments. *Water Resour. Res.* 35:2965–2979.
- Sorooshian, S., Q. Duan, and V.K. Gupta. 1993. Calibration of rainfall runoff models: Application of global optimization to the Sacramento soil moisture accounting. *Water Resour. Res.* 29:1185–1194.
- Sulzer, S., R. Rutschmann, and W. Kinzelbach. 2002. Flood discharge prediction using two-dimensional inverse modeling. *J. Hydrol. Eng. ASCE*. 128:46–54.
- van Genuchten, M.Th. 1981. Non-equilibrium transport parameters from miscible displacement experiments. *Res. Rep. 119*. U.S. Salinity Laboratory, USDA-ARS, Riverside, CA.
- Vogel, T., K. Huang, R. Zhang, and M.Th. van Genuchten. 1996. The HYDRUS 5 code for simulating one-dimensional water flow, solute transport, and heat movement in variably-saturated media. Version 5. U.S. Salinity Laboratory, USDA-ARS, Riverside, CA.
- von Gunten, U., and G. Furrer. 2000. Steady-state modeling of biogeochemical processes in columns with aquifer material: 2. Dynamics of iron-sulfur interactions. *Chem. Geol.* 167:271–284.
- Vrugt, J.A., W. Bouten, H. V. Gupta, J. W. Hopmans. 2003a. Toward improved identifiability of soil hydraulic parameters: On the selection of a suitable parametric model. Available at [www.vadosezonejournal.org](http://www.vadosezonejournal.org). *Vadose Zone J.* 2:98–113.
- Vrugt, J.A., H.V. Gupta, W. Bouten, and S. Sorooshian. 2003b. 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.
- Wang, Q.J. 1997. Using genetic algorithms to optimise model parameters. *Environ. Model. Software* 12:27–34.
- Wang, W., S.P. Neuman, T. Yao, and P.J. Wierenga. 2003. Simulation of large-scale field infiltration experiments using a hierarchy of models based on public, generic, and site data. Available at [www.vadosezonejournal.org](http://www.vadosezonejournal.org). *Vadose Zone J.* 2:297–312.
- Yapo, P.O., H.V. Gupta, and S. Sorooshian. 1998. Multi-objective global optimization for hydrologic models. *J. Hydrol.* 204:83–97.
- Yeh, W.W.-G. 1986. Review of the parameter identification procedures in groundwater hydrology: The inverse problem. *Water Resour. Res.* 22:95–108.
- Young, M.H., A. Karagunduz, J. Simunek, and K.D. Pennell. 2002. A modified upward infiltration method for characterizing soil hydraulic properties. *Soil Sci. Soc. Am. J.* 66:57–64.