Free-form estimation of the unsaturated soil hydraulic properties by inverse modeling using global optimization

S. C. Iden and W. Durner

Received 21 December 2006; revised 27 March 2007; accepted 19 April 2007; published 28 July 2007.

Inverse modeling is a powerful technique for identifying the hydraulic properties of unsaturated porous media. However, the selection of an appropriate parameterization of the soil water retention and hydraulic conductivity function remains a challenge. In this article, we present an improved algorithm for estimating these two relationships without assigning an a priori shape to them. The approach uses cubic Hermite interpolation and a global optimization strategy. A multilevel routine identifies the adequate number of degrees of freedom by balancing model performance, the statistical interaction of the estimated model parameters, and their number. A first-order uncertainty analysis provides a quantitative measure of how well the soil hydraulic properties can be identified in different ranges of pressure head. This offers great potential for designing optimal experimental procedures for identifying the hydraulic properties of porous media. We demonstrate the effectiveness of the algorithm for the evaluation of multistep outflow experiments by investigating synthetic data sets and real measurements. The free-form approach yields optimal model parameters that show only moderate correlation, indicating well-posed inverse problems. Since parameterization errors are almost completely avoided, the algorithm is well suited to identifying other error sources in unsaturated flow problems, e.g., limitations in the applicability of the Richards equation or problems caused by spatial heterogeneity.


1. Introduction

An accurate knowledge of the soil hydraulic properties (SHP) is crucial for solving many basic and applied problems in agricultural and environmental research [Vanclooster et al., 2004]. Models used to predict the transport and fate of pollutants in the subsurface rely on a proper soil hydraulic characterization [Śmigielski, 2005], because the SHP control the dynamics of water flow in the vadose zone.

The standard approach to simulate unsaturated water flow in porous media is the Richards equation [van Dam et al., 2004]. Solving Richards’ equation requires appropriate initial and boundary conditions and knowledge of two constitutive relationships, namely the soil water retention curve and the unsaturated hydraulic conductivity curve. A variety of methods has been developed to determine the SHP in the laboratory or in the field. These methods range from relatively simple, static or steady state flow experiments to experiments under transient conditions [Dane and Topp, 2002; Durner and Lipsius, 2005]. The latter are advantageous over the traditional methods with regard to time requirements and the use of nonrestrictive boundary conditions [Hopmans et al., 2002; Vrugt and Dane, 2005].

An exact evaluation of experiments under transient flow conditions is only possible by modeling the flow process in combination with parameter estimation. Thus considerable effort has been spent on improving the computational tools for solving inverse problems of unsaturated flow. The trend toward more powerful and reliable methods includes the adaptation and application of modern global optimization techniques [Abbaspour et al., 1997; Pan and Wu, 1998; Takeshita, 1999; Abbaspour et al., 2001; Lambot et al., 2002] and reliable methods of uncertainty analysis [Vrugt et al., 2003]. Most of the studies on inverse parameter estimation in soil physics are restricted to the laboratory and the lysimeter scale. However, as the vadose zone scientific community is striving to bridge the gap between small and regional-scale simulation studies, the application to larger scales receives more attention, for instance to derive effective properties [Harter and Hopmans, 2004; Vrugt et al., 2004].

Any mathematical expression that describes the SHP can be referred to as a soil hydraulic property function (SHPF). The majority of SHPF consist of closed-form expressions with a relatively small number of parameters. These functions relate water content and hydraulic conductivity to the water potential over the entire pressure head domain (see Leij et al. [1997] for an extensive survey). The most frequently used expressions for the retention curve are those of van Genuchten [1980] and Brooks and Corey [1964]. For soils with heterogeneous pore systems, com-
posite functions have been developed [Durner, 1994; Othmer et al., 1991; Ross and Smettem, 1993] to overcome the inability of the classic unimodal retention functions to correctly describe measured soil water retention data. Parameterizations of the hydraulic conductivity function are often coupled to the retention curve by either empirical equations or capillary models. Although physically motivated, such approaches may give rise to practical problems, since (1) the capillary model itself may be only approximately valid for the porous medium under study and (2) any error of the employed model of the retention function propagates into errors in the predicted conductivity function. This is particularly the case if the model of the retention function systematically misfits the true relationship close to saturation [Durner, 1994]. Consequently, it was suggested to decouple the two functions [Nielsen and Luckner, 1992]. As an alternative to the classic parametric approach, some authors suggest to apply spline interpolation methods [Kastanek and Nielsen, 2001; Prunty and Casey, 2002; Bitterlich et al., 2004]. This offers greater flexibility in the SHP as no global shape of the function is assumed a priori, and therefore limitations of the closed-form expressions may be overcome.

In a recent review on inverse modeling of soil hydraulic properties, Vrugt and Dane [2005] have classified the ongoing research efforts into five categories. Among these, this article refers primarily to the development and selection of suitable models of the hydraulic properties and the quantification of the uncertainty with which hydraulic properties can be determined from multistep outflow experiments. Our approach is closely related to the ones published by Bitterlich et al. [2004] who used a spline interpolation for identifying hydraulic properties in an inverse modeling framework and the work of Knabner et al. [2005] who applied linear interpolation in a multilevel framework for describing equilibrium and nonequilibrium sorption of solutes in column breakthrough experiments. In this article, we present a new algorithm using cubic Hermite interpolation for the description of the SHP. Different from Bitterlich et al. [2004], we use a global optimization strategy, the Shuffled Complex Evolution (SCE-UA) algorithm [Duan et al., 1992] which we adapted to parameter estimation problems with nonequality constraints. The new algorithm is embedded in an improved multilevel framework for identifying the adequate number of degrees of freedom (df). A balance between the flexibility in the SHP and the identifiability of the spline coefficients is achieved by calculating a parameter collinearity index introduced by Brun et al. [2001]. The uncertainty of the estimated model parameters is assessed using a classic first-order approximation and uncertainty bounds for the SHP are inferred. Since the spline approach assigns a local influence on the SHP to the estimated parameters, the uncertainty limits get a local meaning as well. This enables an uncertainty assessment for different ranges of pressure head and an answer to the question in which pressure head domain a given experimental design provides sufficient information for determining the SHP.

2. Theory

Water flow in rigid, unsaturated porous media under isothermal conditions can be simulated on the macroscopic scale using Richards’ equation. For one-dimensional flow without sinks and sources, Richards’ equation in the pressure head form reads

$$ C(h) \frac{\partial h}{\partial t} = \frac{\partial}{\partial z} \left[ K(h) \frac{\partial h}{\partial z} - K(h) \right] $$

(1)

where \( t [T] \) is time, \( z [L] \) is depth, positive downward, \( h [L] \) is the pressure head, \( C(h) [L^{-1}] \) is the soil water capacity function, and \( K(h) [L T^{-1}] \) is the unsaturated hydraulic conductivity function. The latter can be written as

$$ K(h) = K_s K_r(h) $$

(2)

where \( K_r(h) [-] \) is the relative hydraulic conductivity function and \( K_s [L T^{-1}] \) is the saturated hydraulic conductivity. The soil water capacity function \( C(h) \) is the partial derivative of the soil water retention function \( \theta(h) \) with respect to pressure head, where the latter relates volumetric water content \([L^3 L^{-3}]\) to pressure head \( h \).

In order to solve the nonlinear partial differential equation (1), appropriate initial and boundary conditions have to be specified. For a typical outflow experiment, these can be specified as [Hopmans et al., 2002]

$$ h(0,z) = z \quad 0 \leq z \leq L_c $$

$$ \frac{\partial h}{\partial z}(t,0) = 1 \quad 0 \leq t \leq t_{max} $$

$$ h(t,L_c) = h_{LB}(t) \quad 0 \leq t \leq t_{max} $$

where \( L_c \) denotes the length of the soil column and \( t_{max} \) denotes the termination time of the experiment.

2.1. Free-Form Functions

Spline interpolation methods are nowadays widely applied for deriving smooth functions over a given interval [Press et al., 1992]. In order to describe the SHP by spline interpolation, the pressure head relevant for the unsaturated flow problem of interest has to be partitioned into subintervals. Interpolating between nodal values of \( \theta \) and \( \log K \) yields smooth functional relationships of the SHP. In the article of Bitterlich et al. [2004] this was referred to as “free-form approach” indicating that no a priori shape of the hydraulic properties is assumed except for their monotonicity.

In order to appropriately partition the pressure head axis, one needs to bear in mind that outflow experiments provide information about the SHP only in the range of potentials occurring in the column during the experiment. For the initial and boundary conditions given by equation (3), the pressure head range inside the column is restricted by the pressure applied at the bottom of the column and the length of the column in [DuChateau, 1997]. Instead of partitioning the relevant pressure head interval on a linear scale, as done by Bitterlich et al. [2004], we found it preferable to partition the pressure head range on the log scale. This is conveniently achieved by expressing the pressure head \( h \) by the decimal logarithm of its absolute value: \( pF \equiv \log|h| \), where \( h \) is given in units of cm.
two hydraulic functions are kept at their saturated values \( \theta_s \) and \( K_s \). This effectively introduces an air entry value of \(-1 \) cm corresponding to an effective maximum pore size diameter of 3 mm (assuming full wettability). This is physically justified since flow in pores wider than this cannot be adequately described by Darcy’s law. Moreover, introducing a distinct air entry value leads to a better numerical performance of algorithms solving equation (1) [Vogel et al., 2001]. We allow the two functions to stay either constant or decrease on the pF interval \( \Gamma = (0, pF_{\max}) \) without any further assumption on their shape except the ones imposed by the selected interpolation technique.

[12] The new free-form algorithm adopts the basic idea presented by Bitterlich et al. [2004], as illustrated in Figure 1. A number of \( r \) nodes is distributed on the interval \( \Gamma \) and the hydraulic properties are obtained by piecewise cubic Hermite interpolation [Fritsch and Carlson, 1980] of the \( (r + 1) \) nodal values of \( \theta \) and \( \log K \), respectively. The number of estimated model parameters sums up to \( 2r \) (\( r \) water contents and \( r \) log hydraulic conductivities). In order to ensure the monotonicity of the properties (both functions must decrease or stay constant with increasing pF), the following inequality constraints are imposed:

\[
\begin{align*}
\theta_j & \geq \theta_1 \\
\theta_j & \geq \theta_{j+1} \quad 1 \leq j \leq r - 1 \\
K_j & \geq K_1 \\
K_j & \geq K_{j+1} \quad 1 \leq j \leq r - 1
\end{align*}
\]

[13] Although the unsaturated hydraulic conductivity function is fixed at its saturated value for \( pF \leq 0 \), it is noteworthy that \( K_s \) does not scale the relative hydraulic conductivity function \( K_r(h) \) as it is frequently done by capillary bundle models. In fact, the free-form approach determines \( K(h) \) instead of estimating \( K_r(h) \). As long as the selected value of \( K_s \) is not smaller than any other value of the true function \( K(h) \), this does not lead to problems to correctly estimate the unsaturated hydraulic conductivity function.

2.2. Inverse Modeling

[14] In a weighted least squares framework, the objective function used for estimating the unknown soil hydraulic parameter vector \( p \) from measurements of outflow data and records of state variables from within the soil column (like tensiometer readings and TDR-data) is given by

\[
O(p) = \sum_{i=1}^{n} w_i (y_i - \hat{y}_i(p))^2
\]

where \( y_i \) denotes a measured value and \( \hat{y}_i \) denotes a corresponding model prediction dependent on the parameter vector \( p \). The summation is carried out over the total number of observations \( n \) (comprising different data types). The weights \( w_i \) should be assigned to the residuals \( r_i = y_i - \hat{y}_i \) such that those data points which are considered less reliable than others carry a lower weight. It is common practice to scale the residuals \( r_i \) by the inverse of the variance of the measurement error \( \sigma_i^2 \) of each record:

\[
w_i = \sigma_i^{-2}
\]

which is in agreement with the maximum-likelihood principle [Omlin and Reichert, 1999]. Other weighting schemes and objective functions can be applied on the basis of statistical reasoning [Finsterle, 2004; Finsterle and Najita, 1998]. However, in cases where a model error contributes significantly to deviations between observed and predicted values, a sound statistical basis for deriving an objective function may not exist.

[15] In our work, the inverse problem (minimizing \( O(p) \)) with respect to \( p \), subject to the constraints given by equation (4) was solved by combining a modified version of the Hydrus1D software code for numerically solving the Richards equation [Simunek et al., 1998] with the shuffled-complex-evolution algorithm (SCE-UA) presented by Duan et al. [1992]. The SCE-UA algorithm is a global optimization strategy which has shown its efficiency in many studies dealing with the calibration of conceptual rainfall runoff models [Gan and Biftu, 1996; Thyer et al., 1999]. Vrugt and Bouten [2002] are among the first to have used the SCE-UA method to estimate hydraulic properties from multistep outflow experiments. The method combines the strength of the local simplex procedure of Nelder and Mead [1965] with concepts like competitive evolution [Holland, 2001].
and complex shuffling [Duan et al., 1992]. The algorithm relies solely on function evaluations whereas many classic, gradient-based optimization strategies require first derivatives of the objective function with respect to the estimated parameters [Finsterle, 2004]. For the purpose of our study, the original SCE-UA algorithm was modified. The feasible parameter domain was redefined to account for the non-equivalence constraints. The change of this affects both to the triggering of the mutation step and the offspring generation when a mutation occurs. The application of a global optimization algorithm is one of the improvements over the algorithm presented by Biterlich et al. [2004] who used a classic nonlinear programming method accounting for non-equality constraints.

### 2.3. Improved Multilevel Procedure

[16] As stated by Biterlich et al. [2004], the inverse modeling results obtained by the free-form approach depend very much on the number of df (i.e., number of nodes) used for parameterizing the hydraulic functions. A higher number of nodes in general leads to an increased flexibility in the hydraulic properties which is appreciated in cases where the pore size distribution differs significantly from an approximately lognormal shape. However, when increasing the number of estimated parameters, the statistical independence of the model parameters increases in parallel, which finally may lead to nonidentifiability of single model parameters [Brun et al., 2001].

[17] Biterlich et al. [2004] presented a cascading approach for introducing new points into a given vector of pressure heads which form the basis for cubic Hermite interpolation. The procedure consisted of inserting new nodal values of the pressure head at the center between all existing nodes. The disadvantage of this technique is twofold: (1) the nodes are always equidistantly spaced on the given pressure head interval irrespective of the local shape of the hydraulic functions and (2) the number of nodes increases very quickly at higher levels of r (2, 3, 5, 9, 17, 33, ... nodes). As the number of nodes is almost doubled when new nodes are added, the estimation algorithm may become inefficient, since more parameters than necessary may be estimated when moving through the cascade.

[18] In our new multilevel algorithm, only one new node is inserted per step of the cascade, and the positions of the original nodes are shifted. The algorithm consists of the following steps:

1. Determine the best fit parameter vector \( \mathbf{p}_0 \) consisting of \( r \) water content values and \( r \) log conductivity values by applying the SCE-UA algorithm. This is based on \( r \) nodes located on the pF axis.
2. Map the optimum retention curve on the basis of the last fit (using \( r \) nodes) to a transformed plane spanned by a normalized pF axis and a normalized saturation axis.
3. Add a new node and redistribute the resulting \( r + 1 \) points on the retention curve in the transformed plane such that they exhibit equal Euclidian distances.
4. Transform the resulting vector of normalized pF values back to the regular pF range.
5. Use the resulting pF table for the subsequent optimization with \( 2(r + 1) \) df. The starting nodal values of \( \theta \) and \( \log K \) are assigned to the new pF table by cubic Hermite interpolation using the results from the antecedent optimization.

### 2.4. Diagnostic Variables

[24] In order to find a balance between flexibility in the properties and identifiability of the model parameters, we applied a number of diagnostic variables for identifying the adequate number of df. These variables can be categorized into those characterizing the goodness of fit and those characterizing the degree of interaction between the estimated model parameters. In practice, one has to find a balance between model performance, parameter interaction and the number of estimated model parameters to identify an optimal number of nodes \( r \).

[25] For evaluating the goodness of fit, we compared the minimum value of the objective function for the different number of df applied. If the weights in the definition of \( O(\mathbf{p}) \) are chosen according to equation (6), the value of the objective function at the minimum follows a \( \chi^2 \) distribution with \( n - n_p \) degrees of freedom [Press et al., 1992] where \( n \) denotes the number of observations and \( n_p \) denotes the number of estimated model parameters, respectively. On this basis, a probability of model adequacy \( p_{adeq} \) can be calculated from the \( \chi^2 \) cumulative distribution function [Press et al., 1992; Hollenbeck and Jensen, 1998]. The model adequacy is a measure of how well the model fits the observations given the experimental error [Vrugt et al., 2003].

[26] Additional measures for assessing the goodness of fit or model performance used in this work were the root-mean-square error (RMSE) and the number of runs of the residuals \( n_r \). A run is defined as a sequence of residuals with equal sign [Bard, 1974; Draper and Smith, 1998]. Since it only uses the structure of the sign of the residuals, the number of runs is independent of a priori assumptions on the magnitude of measurement error and therefore complementary to the RMSE and the probability of model adequacy. On the basis of the calculated number of runs \( n_r \), a probability of complete randomness \( p_{ran} \) in the residuals can be calculated from the standard normal distribution function. However, when judging the results of the runs test, one has to take into consideration that autocorrelated measurement error can lead to a failure of the runs test [Bard, 1974]. We calculated both RMSE and \( n_r \) separately for different data groups in order to individually assess the quality of fit for different data types contained in the objective function.

[27] In addition to assessing the goodness of fit, we address the issue of nonidentifiability of model parameters by calculating the collinearity index \( \gamma \) of Brun et al. [2001]. Consider the \( n \times n_p \) sensitivity or derivative matrix of the model output with respect to the model parameters evaluated at the optimal parameter set \( \mathbf{p}_0 \):

\[
\mathbf{V} = \frac{\partial \mathbf{y}}{\partial \mathbf{p}} \Bigg|_{\mathbf{p} = \mathbf{p}_0}
\]

[28] This matrix is composed of \( n_p \) column vectors \( \mathbf{v}_j \) each of which represents the change in the model output with respect to model parameter \( p_j \). Because of its simplicity and generality, the most widely used approach for calculating \( \mathbf{V} \) is the finite difference or perturbation method [Hill,
1998). We calculated $V$ by applying a central difference scheme because this yields a more accurate approximation than a forward or backward finite difference [Doherty, 2004]. For the free-form hydraulic functions, the nodal values of $\theta$ and $\log K$ where increased and decreased by 1% whenever this did not violate the monotonicity constraints imposed by equation (4). In cases where a perturbation lead to a violation of the constraints given by equation (4), the level of perturbation was decreased until the corresponding equality in equation (4) holds. In the case where a nodal value of $\theta$ or $\log K$ equals the values of the two adjacent nodes, the parameter is regarded as not being estimated since its value is determined by the monotonicity constraints.

[29] It is common practice to scale the entries of matrix $V$, for instance by the experimental error of each observation [Romano and Santini, 1999]. This leads to a scaled sensitivity matrix $S$. Dimension-free sensitivity information is obtained by normalizing the column vectors of $S$ by their individual $L_2$ norm yielding the scaled sensitivity matrix $U$. Brun et al. [2001] proposed to estimate the joint influence of the estimated parameters on the model output by assessing the collinearity of the column vectors of $U$. For this purpose, a collinearity index $\gamma$ was calculated which is defined as [Brun et al., 2001]

$$\gamma = \frac{1}{\sqrt{\lambda_{\text{min}}}}$$

where $\lambda_{\text{min}}$ is the smallest eigenvalue of $U^TU$. Brun et al. [2001] state that critical values of $\gamma$ lie in a range of 5–20. For each number of nodes $r$ we calculated three different collinearity indices: $\gamma_\theta$ relates to the full matrix $U$, whereas $\gamma_\theta$ and $\gamma_K$ relate only to the columns of $U$ corresponding to water contents and log hydraulic conductivities, respectively. By this, we quantify the parameter interactions of the full parameter set and the two subsets defining the two hydraulic functions.

[30] For identifying the optimal number of nodes $r$, we introduce a test variable which aggregates the information contained in the quality of fit, the values of the collinearity index $\gamma_r$ and the number of estimated model parameters $2r$. In most cases, an increase in the number of nodes $r$ leads to a better fit to the measured data thereby reducing the final value of the objective function $O_{\text{min}}$. Conversely, estimating more model parameters leads to stronger parameter interactions thereby increasing the value of the collinearity index $\gamma_r$. An optimal model structure yields a good model performance with acceptable values of $\gamma_r$. Accounting for the principle of parsimony requires to introduce a penalty on the number of df, since among those models performing equally well, the one with minimum $r$ should be preferred.

[31] In order to account for all three aspects that are relevant for model selection, we define a penalized misfit function $P^*$ as the sum of (1) the final value of the objective function normalized by the number of measurements, (2) the value of $\gamma_r$ normalized by 5 (the lower limit of its critical range as reported by Brun et al. [2001]) and (3) the logarithm of the number of estimated parameters $2r$ normalized by the logarithm of 12. We selected 12 as empirical benchmark for the number of parameters because estimating more parameters than this sequentially increases the likelihood of obtaining hydraulic functions of rugged shape. In cases where the model is adequate (defined in the sense of $P_{\text{adap}}$, i.e., adequate relative to the assumed measurement error), the first contributing part to $P^*$ is around unity. The normalized collinearity index has values greater than unity if $\gamma$ is greater than five and the penalty exerted by the number of df is greater than unity for $r > 6$. The minimum value of $P^*$ yields the optimal value of $r$. $P^*$ has some similarity to statistical model selection criteria which penalize the log likelihood criterion by the number of estimated parameters and the norm of the Fisher information matrix (see Carrera and Neumann [1986] for a brief review), but lacks a strict statistical basis. However, it can be heuristically applied to the problem of model selection without relying on the validity of statistical assumptions.

### 2.5. Uncertainty Analysis for Hydraulic Functions

[32] Assessing the uncertainty in the identified soil hydraulic functions is of crucial interest for any soil hydrological study. Although many studies have addressed the uncertainty of single estimated soil hydraulic parameters, e.g., van Genuchten $\alpha$ or $n$, only a few authors have addressed the issue of uncertainty of the functions $\theta(h)$ and $K(h)$ (see Vrugt et al. [2003] for a Bayesian approach). From our point of view, the uncertainty and identifiability of single soil hydraulic parameters must be accompanied by an assessment of the uncertainty in the hydraulic properties. While the identifiability of single parameters is crucial for the well posedness of the inverse problem, it should be kept in mind that the accurate determination of the SHP is the objective of most of the inverse modeling studies in soil hydrology.

[33] In order to assess the uncertainty in $\theta(h)$ and $K(h)$, we employed a first-order approximation [Vrugt and Bouten, 2002] and inferred the parameter covariance matrix $C_p$ from the scaled sensitivity matrix $S$ as

$$C_p = \frac{O_{\text{min}}}{n-p} (S^T S)^{-1}$$

where $O_{\text{min}}$ is the minimum value of the objective function defined by (5). Note that the use of $S$ in this context is restricted to cases where the weights in the objective functions are assigned according to equation (6).

[34] Confidence intervals of single model parameters for a given level of significance were calculated from the diagonal elements of $C_p$, which are the variances of the estimated model parameters. The confidence intervals relate directly to the soil hydraulic properties as the estimated parameters are nodal values of $\theta$ and $\log K$. Consequently, piecewise cubic Hermite interpolation between the parameter bounds yields the uncertainty of the SHPF. To ensure the monotonicity of the uncertainty bounds, the nodal bounds were adjusted prior to interpolation so that they meet the nonequality constraints given by equation (4).

### 3. Materials and Methods

#### 3.1. Inverse Modeling Using Synthetic Data

[35] In order to test the algorithm, we generated three synthetic data sets and performed numerical inversions with the free-form global optimization algorithm. The data sets were selected on the basis of the article by Zurmühl and
Table 1. Parameters of the MVGMD Model Used for Generation of Synthetic Data Sets

<table>
<thead>
<tr>
<th>Scenario</th>
<th>( \theta_s )</th>
<th>( \theta_r )</th>
<th>( \alpha_1 )</th>
<th>( n_1 )</th>
<th>( w_1 )</th>
<th>( \alpha_2 )</th>
<th>( n_2 )</th>
<th>( K_s )</th>
<th>( \tau )</th>
<th>( h_s )</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>0.5</td>
<td>0.03</td>
<td>0.1</td>
<td>2.7</td>
<td>0.3</td>
<td>0.01</td>
<td>1.9</td>
<td>10</td>
<td>0.5</td>
<td>2</td>
</tr>
<tr>
<td>S2</td>
<td>0.5</td>
<td>0.03</td>
<td>0.15</td>
<td>1.8</td>
<td>0.4</td>
<td>0.02</td>
<td>1.3</td>
<td>10</td>
<td>0.5</td>
<td>2</td>
</tr>
<tr>
<td>S3</td>
<td>0.5</td>
<td>0.03</td>
<td>0.5</td>
<td>4.0</td>
<td>0.1</td>
<td>0.01</td>
<td>1.8</td>
<td>10</td>
<td>0.5</td>
<td>2</td>
</tr>
</tbody>
</table>

\( \theta_s \) is the saturated water content, \( \theta_r \) is the residual water content, \( \alpha_1 \), \( n_1 \), \( w_1 \), \( \alpha_2 \), \( n_2 \), \( K_s \), \( \tau \), and \( h_s \) are parameters controlling the SHPF. Both the saturated water content \( \theta_s \) and the saturated hydraulic conductivity \( K_s \) are known values. Keeping these parameters fixed is common practice when evaluating multistep outflow data since the experiment is quite insensitive to conductivity estimates close to saturation [Durner et al., 1999], and the outflow experiment provides information on changes in water content, but not on absolute values of \( \theta \) if no additional water content measurements are available. For the uncertainty analysis, we regarded the uncertainty of the independently determined \( \theta_s \) and \( K_s \) as zero.

3.2. Inverse Modeling Using Real Data

To illustrate the performance of the algorithm on real data, we used a data set from a multistep outflow experiment conducted on an undisturbed soil column (sandy loam) sampled from a plough horizon close to the city of Bayreuth, Germany. The column was 14.4 cm high and 9.4 cm in diameter. It was mounted on top of a ceramic porous plate with a saturated hydraulic conductivity of \( K_s = 4.9 \text{ cm/hr} \) (value obtained from independent measurement). Ten pressure steps were applied to the sample by decreasing the pressure at the bottom of the porous plate, starting from full saturation. Cumulative outflow at the bottom and the pressure head at 3.5 and 9.0 cm distance from the top of the column was measured. Both the saturated water content \( \theta_s = 0.298 \) and the saturated hydraulic conductivity \( K_s = 4.0 \text{ cm/hr} \) of the sample were measured independently. As with the synthetic data, the uncertainty analysis of the SHPF was carried out by assuming these two parameters to be free of error. The parameters controlling the SCE-UA optimization were set as discussed above.

In order to assess the performance of the free-form algorithm, we compared the free-form results with those obtained by fitting classic parametric hydraulic functions of varying flexibility. The five examined models comprised (1) the bimodal van Genuchten–Mualem model according to Durner [1994] with 6 df, (2) the same model with seven df, (3) the modified bimodal model (MVGMD) with 6 df, (4) the MVGMD with 7 df, (5) and the MVGMD model with 8 df. Whenever the air entry pressure \( h_a \) was treated as free parameter, it was allowed to vary between \(-5 \) and \(-0.5 \) cm. Otherwise it was set to \(-1 \) cm in agreement with the free-form approach. Parameter estimation was carried out using the SCE-UA algorithm with the same convergence criteria as outlined above. The number of complexes was set to twice the number of estimated parameters \( n_p \) and the number of points per complex was set to \( 2n_p + 1 \). The weights in the objective function were chosen according to equation (6) with the standard deviation of the cumulative outflow and pressure head data assumed to be \( 0.01 \) cm and \( 0.25 \) cm, respectively.

4. Results and Discussion

4.1. Synthetic Experiments

The inverse modeling of the outflow scenarios S1 to S3 lead in each case to excellent fits between simulated and measured data.
“observed” data. This is summarized in Table 2, where values of selected diagnostic variables are listed as function of the number of nodes \( r \). For data set S1 the minimum objective function value decreases almost monotonically with the number of nodes \( r \). The probability of model adequacy \( p_{\text{adeq}} \) reaches its maximum value of 0.35 for \( r = 13 \). For \( r = 14 \) and \( r = 15 \) the free-form approach fits the data almost equally well as for \( r = 13 \). However, this is achieved with a higher number of df leading to a decrease in \( p_{\text{adeq}} \).

The collinearity index for the entire parameter set \( g \) used to diagnose the parameter interdependence increases monotonically with \( r \). This reflects the fact that the more parameters are estimated from the multistep outflow data set, the stronger become the interactions between the parameters. The value of 6.62 obtained for \( r = 13 \) (highest probability of model adequacy) seems acceptable given the critical range of 5–20 for \( g \) [Brun et al., 2001]. As \( g \) amounts to only 7.04 for \( r = 15 \), the information content of the data supports the identification of at least up to \( g \) parameters.

### Table 2. Results for the Three Synthetic Data Sets S1, S2, and S3

<table>
<thead>
<tr>
<th>( r )</th>
<th>( r_{\text{min}} )</th>
<th>( p_{\text{adeq}} )</th>
<th>( \gamma_{\theta} )</th>
<th>( \gamma_{K} )</th>
<th>( \gamma_{\theta} )</th>
<th>( \gamma_{K} )</th>
<th>( \gamma_{\theta} )</th>
<th>( \gamma_{K} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>265623</td>
<td>0.00</td>
<td>1.25</td>
<td>ND</td>
<td>ND</td>
<td>15927</td>
<td>0.00</td>
<td>1.31</td>
</tr>
<tr>
<td>2</td>
<td>25962</td>
<td>0.00</td>
<td>2.12</td>
<td>1.02</td>
<td>1.81</td>
<td>18800</td>
<td>0.00</td>
<td>3.17</td>
</tr>
<tr>
<td>3</td>
<td>17636</td>
<td>0.00</td>
<td>3.10</td>
<td>1.38</td>
<td>1.77</td>
<td>3627</td>
<td>0.00</td>
<td>5.67</td>
</tr>
<tr>
<td>4</td>
<td>3317</td>
<td>0.00</td>
<td>3.51</td>
<td>1.59</td>
<td>1.80</td>
<td>1913</td>
<td>0.03</td>
<td>6.61</td>
</tr>
<tr>
<td>5</td>
<td>3197</td>
<td>0.00</td>
<td>4.09</td>
<td>1.84</td>
<td>1.79</td>
<td>1833</td>
<td>0.26</td>
<td>7.05</td>
</tr>
<tr>
<td>6</td>
<td>2183</td>
<td>0.00</td>
<td>4.60</td>
<td>1.98</td>
<td>1.82</td>
<td>1820</td>
<td>0.33</td>
<td>7.15</td>
</tr>
<tr>
<td>7</td>
<td>2215</td>
<td>0.00</td>
<td>4.89</td>
<td>2.08</td>
<td>1.81</td>
<td>1806</td>
<td>0.40</td>
<td>7.36</td>
</tr>
<tr>
<td>8</td>
<td>1995</td>
<td>0.00</td>
<td>5.05</td>
<td>2.20</td>
<td>1.80</td>
<td>1802</td>
<td>0.42</td>
<td>7.59</td>
</tr>
<tr>
<td>9</td>
<td>1887</td>
<td>0.05</td>
<td>5.07</td>
<td>2.42</td>
<td>1.79</td>
<td>1800</td>
<td>0.41</td>
<td>7.73</td>
</tr>
<tr>
<td>10</td>
<td>1866</td>
<td>0.09</td>
<td>5.07</td>
<td>3.24</td>
<td>1.78</td>
<td>1799</td>
<td>0.41</td>
<td>7.82</td>
</tr>
<tr>
<td>11</td>
<td>1829</td>
<td>0.22</td>
<td>5.17</td>
<td>3.76</td>
<td>1.96</td>
<td>1796</td>
<td>0.41</td>
<td>7.92</td>
</tr>
<tr>
<td>12</td>
<td>1820</td>
<td>0.26</td>
<td>5.33</td>
<td>3.72</td>
<td>1.99</td>
<td>1796</td>
<td>0.40</td>
<td>7.95</td>
</tr>
<tr>
<td>13</td>
<td>1803</td>
<td>0.35</td>
<td>6.62</td>
<td>4.26</td>
<td>2.34</td>
<td>1797</td>
<td>0.38</td>
<td>14.04</td>
</tr>
<tr>
<td>14</td>
<td>1802</td>
<td>0.34</td>
<td>6.67</td>
<td>4.53</td>
<td>2.57</td>
<td>1793</td>
<td>0.39</td>
<td>8.92</td>
</tr>
<tr>
<td>15</td>
<td>1803</td>
<td>0.32</td>
<td>7.04</td>
<td>5.27</td>
<td>2.94</td>
<td>1792</td>
<td>0.39</td>
<td>10.27</td>
</tr>
</tbody>
</table>

\( ^{a} \)Shown are the minimum value of the objective function \( O_{\text{min}} \), the probability of model adequacy \( p_{\text{adeq}} \), and the three-parameter collinearity indices \( \gamma \) as functions of the number of nodes \( r \). The overall number of measurements in the objective function is 1806. ND means not defined.

Figure 2. Original (dotted blue line) and identified (solid black line) hydraulic properties for the three synthetic data sets S1, S2, and S3 for 13, 8, and 14 nodes, respectively. The original properties were generated by a modified van Genuchten–Mualem–Durner model (bimodal with an air entry value \( h_{e} \) of \(-2 \) cm). The identified properties result from the free-form algorithm with the nodal values of \( \theta \) and \( \log K \) denoted by the black circles. The shaded areas illustrate the 95% confidence region of the estimated properties as inferred from the first-order statistical analysis. The optimal properties shown maximize the probability of model adequacy \( p_{\text{adeq}} \).
Figure 3. Development of the penalized misfit function $P^*$ for the three synthetic data sets S1, S2, and S3 and the real data set Bayreuth sandy loam. The identified optimal number of nodes is shown by the darker columns.

30 parameters given the initial and boundary conditions and the shape of the SHPF analyzed, even more parameters may be identifiable, but we stopped the algorithm at $r = 15$.

[44] In addition to quantifying the interactions between all estimated model parameters, i.e., water contents and $r$ log $K$ values, the interactions of both parameter groups are assessed independently by examining the values of $\gamma_q$ and $\gamma_k$. Both diagnostic variables have values smaller than five for almost all number of df indicating that the specific parameter interdependence among the estimated water contents and log conductivity values poses no challenge for solving the inverse problem. The distinctly higher values of $\gamma_r$ as compared to $\gamma_q$ and $\gamma_k$ can be attributed to the interaction between the soil water retention and unsaturated hydraulic conductivity function. Since these two properties are physically related, the correlation between the two is inevitable.

[45] The results for scenarios S2 and S3 are comparable to the ones of S1. For data set S2, the model adequacy peaks for $r = 8$ with a corresponding value of $\gamma_r$ of 7.59. Again, the interaction parameters for the two functions $\gamma_q$ and $\gamma_k$ are significantly lower. For data set S3, $P_{adeq}$ reaches its maximum value of 0.39 for $r = 14$. The corresponding value of $\gamma_r$ is 4.95, indicating a well-posed inverse problem. For $r = 10$ a relatively high value of $\gamma_q$ obviously results from a rather strong interaction between the estimated log $K$ values, since $\gamma_q$ lies within the usual order of magnitude whereas $\gamma_k$ is 13.33. However, this problem is overcome when increasing $r$ to 11. This illustrates an interesting aspect of the modified multilevel algorithm: since the positions of the nodes are shifted along the $pF$ axis when new nodes are introduced, increasing the number of nodes does not necessarily lead to a better fit, nor does it necessarily increase the value of the collinearity index.

[46] In addition to the excellent fit to all three synthetic data sets, the free-form algorithm yields optimum hydraulic properties that excellently match those used for generating the synthetic data. This is illustrated in Figure 2, which shows the functions corresponding to maximum values of $P_{adeq}$. In all three cases the identified SHPF show virtually no difference to the original function for $pF > 1$. Furthermore, the uncertainties of the estimated functions are amazingly small. In the range close to saturation, some deviations occur, which goes along with an increased uncertainty in $\theta(h)$ and $K(h)$. This reflects the insufficient information content of the multistep outflow experiment in that moisture range, a shortcoming of the multistep outflow method that has been discussed in other publications [Butters and Duchateau, 2002; Durner et al., 1999] and does not adversely affect the applicability of the algorithm. From the evidence provided by analyzing the synthetic data sets we conclude that the algorithm is well suited to solve the inverse unsaturated flow problem even for SHP with a relatively complex shape.

[47] An optimum number of nodes for the three test cases is identified by comparing the values of the penalized misfit function $P^*$ for different values of $r$. Figure 3 shows that optimal values of $r$ are 6, 4, and 8 for scenarios S1, S2, and S3, respectively, indicated by minimum values of $P^*$. The corresponding values of the objective function, as shown in Table 2, are already very close to the total number of measurements $n = 1806$ for scenarios S2 and S3. For scenario S1, the minimum value of $O_p(p)$ is only about 21% higher than $n$ indicating an excellent fit to the data with parameters estimated at only 6 nodes. In the case of S2,

The data set consists of three data types with 144 observations of cumulative outflow ($Q$) and the pressure head in two depths ($H_1$ and $H_2$).

Table 3. Results for the Data Set Bayreuth Sandy Loam Using the Free-Form Approach$^a$

<table>
<thead>
<tr>
<th>$r$</th>
<th>$O_{min}$</th>
<th>RMSE$Q$</th>
<th>RMSE$H_1$</th>
<th>RMSE$H_2$</th>
<th>$n_c$</th>
<th>$n_c$</th>
<th>$n_c$</th>
<th>$\gamma_r$</th>
<th>$\gamma_q$</th>
<th>$\gamma_k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>61711</td>
<td>0.194</td>
<td>1.489</td>
<td>1.050</td>
<td>2</td>
<td>7</td>
<td>9</td>
<td>1.05</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>2</td>
<td>13212</td>
<td>0.071</td>
<td>1.240</td>
<td>1.026</td>
<td>3</td>
<td>9</td>
<td>9</td>
<td>3.21</td>
<td>1.43</td>
<td>3.15</td>
</tr>
<tr>
<td>3</td>
<td>3886</td>
<td>0.026</td>
<td>0.706</td>
<td>0.880</td>
<td>5</td>
<td>11</td>
<td>19</td>
<td>2.79</td>
<td>1.21</td>
<td>2.24</td>
</tr>
<tr>
<td>4</td>
<td>2379</td>
<td>0.019</td>
<td>0.470</td>
<td>0.773</td>
<td>12</td>
<td>17</td>
<td>24</td>
<td>4.40</td>
<td>1.49</td>
<td>2.19</td>
</tr>
<tr>
<td>5</td>
<td>1666</td>
<td>0.015</td>
<td>0.434</td>
<td>0.633</td>
<td>27</td>
<td>19</td>
<td>20</td>
<td>5.01</td>
<td>1.47</td>
<td>2.61</td>
</tr>
<tr>
<td>6</td>
<td>1557</td>
<td>0.015</td>
<td>0.417</td>
<td>0.607</td>
<td>15</td>
<td>19</td>
<td>20</td>
<td>6.47</td>
<td>1.90</td>
<td>2.83</td>
</tr>
<tr>
<td>7</td>
<td>1505</td>
<td>0.014</td>
<td>0.417</td>
<td>0.603</td>
<td>19</td>
<td>21</td>
<td>20</td>
<td>7.09</td>
<td>2.07</td>
<td>2.64</td>
</tr>
<tr>
<td>8</td>
<td>1505</td>
<td>0.014</td>
<td>0.417</td>
<td>0.601</td>
<td>23</td>
<td>21</td>
<td>20</td>
<td>11.08</td>
<td>2.22</td>
<td>3.63</td>
</tr>
<tr>
<td>9</td>
<td>1469</td>
<td>0.014</td>
<td>0.422</td>
<td>0.587</td>
<td>19</td>
<td>21</td>
<td>20</td>
<td>7.22</td>
<td>2.63</td>
<td>2.55</td>
</tr>
<tr>
<td>10</td>
<td>1420</td>
<td>0.013</td>
<td>0.415</td>
<td>0.580</td>
<td>21</td>
<td>21</td>
<td>20</td>
<td>7.92</td>
<td>2.47</td>
<td>2.89</td>
</tr>
<tr>
<td>11</td>
<td>1416</td>
<td>0.013</td>
<td>0.408</td>
<td>0.582</td>
<td>29</td>
<td>21</td>
<td>20</td>
<td>8.60</td>
<td>2.80</td>
<td>3.05</td>
</tr>
<tr>
<td>12</td>
<td>1402</td>
<td>0.013</td>
<td>0.413</td>
<td>0.570</td>
<td>20</td>
<td>19</td>
<td>22</td>
<td>10.89</td>
<td>2.84</td>
<td>3.03</td>
</tr>
<tr>
<td>13</td>
<td>1292</td>
<td>0.012</td>
<td>0.389</td>
<td>0.559</td>
<td>28</td>
<td>17</td>
<td>24</td>
<td>8.74</td>
<td>2.87</td>
<td>3.00</td>
</tr>
<tr>
<td>14</td>
<td>1284</td>
<td>0.013</td>
<td>0.391</td>
<td>0.552</td>
<td>24</td>
<td>19</td>
<td>26</td>
<td>14.07</td>
<td>3.28</td>
<td>3.37</td>
</tr>
<tr>
<td>15</td>
<td>1286</td>
<td>0.013</td>
<td>0.389</td>
<td>0.550</td>
<td>19</td>
<td>21</td>
<td>26</td>
<td>8.08</td>
<td>3.31</td>
<td>3.37</td>
</tr>
</tbody>
</table>

$^a$Shown are the minimum value of the objective function $O_{min}$, the root-mean-square error (RMSE) for each data set in the objective function, the number of runs $n_c$ for each data set, and the three-parameter collinearity indices $\gamma$ for different numbers of nodes $r$. The data set consists of three data types with 144 observations of cumulative outflow ($Q$) and the pressure head in two depths ($H_1$ and $H_2$).
estimates at only 4 nodes are required because the SHP have a less complex shape than S1 and S3 (see Figure 2).

4.2. Application to Measured Data

The results of applying the free-form approach and different parametric models to the Bayreuth sandy loam data set is summarized in Tables 3 and 4. Among the tested parametric models, the modified bimodal van Genuchten–Mualem model with fitted \( \tau \) and fitted \( h_s \), representing 8 df, performs best yielding a minimum objective function value of 3118. In comparison to this, the free-form approach yields a distinctively better fit with the same number of df (\( r = 4 \)). As can be seen in Figure 3, the penalized misfit function indicates that \( r = 5 \) is the optimal choice for the number of df. All performance measures given in Table 3, i.e., the RMSE and the number of runs for all three data types in the objective function, show that the free-form algorithm with \( r = 5 \) yields a superior fit to the observations if compared to any of the classic models. Furthermore, the parameter interaction is significantly lower for the free-form approach, with \( \gamma \) reaching a value of only 5.01 as opposed to 46.74 for the best parametric model. Again, the values of the individual collinearity indices \( \gamma_0 \) and \( \gamma_K \) are relatively low, indicating moderate correlation among the parameters defining each curve. Obviously, the model structure defined by the free-form approach is easier to invert and the corresponding inverse problem is better posed.

The superior fit of the free-form algorithm has two causes, namely (1) an enhanced flexibility in the SHPF and (2) the complete decoupling of the hydraulic conductivity curve from the soil water retention curve. This is illustrated in Figure 4, which shows the resulting hydraulic functions for the best MVGMD model and for three free-form fits with selected numbers of nodes. The retention curves for the relevant pF range (pF greater than unity because of the limited information content of the multistep outflow design) are very similar. However, differences become apparent for the \( K(h) \) function, in particular in the \( K(h) \) plot. For \( r = 15 \) the SHPF show oscillations indicating that the algorithm starts to fit measurement errors. The oscillations become apparent in the \( K(h) \) plot for \( r = 10 \) as well, although less pronounced.

Figure 5 presents the model predicted cumulative outflow and pressure head data obtained by fitting the MVGMD model with 8 df and the free-form functions with \( r = 5 \). Obviously both approaches have problems in describing the outflow and pressure head data obtained by fitting the MVGMD model with 8 df (including \( h_s \) and \( \tau \), dash-dotted gray line) and those obtained from the free-form algorithm for different numbers of nodes \( r \).

Table 4. As in Table 3 But for Different Parametric Functions Applied to the Bayreuth Sandy Loam Column

<table>
<thead>
<tr>
<th>Model</th>
<th>Estimated Parameters</th>
<th>( O_{\text{min}} )</th>
<th>RMSE( Q )</th>
<th>RMSE( H_1 )</th>
<th>RMSE( H_2 )</th>
<th>( n_{s,Q} )</th>
<th>( n_{s,H_1} )</th>
<th>( n_{s,H_2} )</th>
<th>( \gamma_0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>VGMD</td>
<td>( \theta, a_1, a_2 )</td>
<td>4194</td>
<td>0.029</td>
<td>0.874</td>
<td>0.734</td>
<td>12</td>
<td>13</td>
<td>15</td>
<td>15.96</td>
</tr>
<tr>
<td>VGMD</td>
<td>( \theta, a_1, a_2 )</td>
<td>4096</td>
<td>0.028</td>
<td>0.863</td>
<td>0.738</td>
<td>10</td>
<td>11</td>
<td>15</td>
<td>40.16</td>
</tr>
<tr>
<td>MVGMD</td>
<td>( \theta, a_1, a_2 )</td>
<td>4331</td>
<td>0.039</td>
<td>0.672</td>
<td>0.698</td>
<td>8</td>
<td>9</td>
<td>19</td>
<td>108.50</td>
</tr>
<tr>
<td>MVGMD</td>
<td>( \theta, a_1, a_2 )</td>
<td>3184</td>
<td>0.025</td>
<td>0.652</td>
<td>0.756</td>
<td>8</td>
<td>9</td>
<td>17</td>
<td>57.28</td>
</tr>
<tr>
<td>MVGMD</td>
<td>( \theta, a_1, a_2 )</td>
<td>3118</td>
<td>0.024</td>
<td>0.681</td>
<td>0.720</td>
<td>10</td>
<td>9</td>
<td>17</td>
<td>46.74</td>
</tr>
</tbody>
</table>

*For the bimodal van Genuchten–Mualem model (VGMD) and its modified form (MVGMD), \( \theta \) and \( K_s \) were set to their independently determined values, and Mualem parameter \( \tau \) was either fixed to a value of 0.5 or treated as a free parameter in the optimization. In the case of the MVGMD, the air entry pressure \( h_s \) was either fixed to a value of \(-1 \) cm or was fitted by the SCE-UA algorithm.
systematic measurement error, and limitations of the applicability of the Richards equation around the air entry point as caused, for instance, by dynamic effects [Schultze et al., 1999; Hassanizadeh et al., 2002]. Because of the high flexibility of the free-form approach, the result of zero model adequacy can hardly be attributed to errors in the SHP. In fact, the free-form approach can be used to identify other error sources in unsaturated flow problems, an important improvement over the use of classic parametric functions.

5. Conclusions

[51] We have presented a new algorithm to identify the SHPF using inverse modeling of transient flow experiments. The algorithm is a further development of the free-form approach introduced by Bitterlich et al. [2004]. Our algorithm uses cubic Hermite interpolation to describe the SHPF and is embedded into a new multilevel framework that enables one to identify the adequate number of nodes used for the interpolation. The optimization algorithm used is the SCE-UA algorithm of Duan et al. [1992] which has shown its global convergence behavior in many studies in different fields of research. In addition to yielding optimal SHPF without a priori assuming a global shape of the properties, our algorithm also quantifies the uncertainty of the estimated SHPF. Since the estimated model parameters exert a local influence on the shape of the SHPF, the obtained uncertainty bounds get a local meaning as well. This enables to assess in which range of pressure head a given experimental design provides enough information content for identifying the SHPF. The estimation of a local SHPF accuracy will help to improve experimental designs for estimating soil hydraulic properties.

[52] The algorithm has shown its ability to correctly identify the SHPF using synthetic data from the commonly applied multistep outflow technique. Because of the improved multilevel algorithm which distributes the nodes evenly on a normalized retention curve and inserts only one node at a time, the adequate number of nodes necessary to describe the experimental data was only moderately high. This means that in the majority of cases between 8 and 12 df will be sufficient to adequately describe measured data obtained from experiments under controlled boundary conditions. In addition to yielding the correct shape of the SHPF, the estimated model parameters, i.e., the nodal values of $\theta$ and $\log K$, show an acceptable low degree of correlation yielding a well-posed inverse problem for the synthetic data sets tested in this work.

[53] The application to real data proved the superiority of the free-form algorithm over approaches using classic parametric functions for an undisturbed sandy loam sampled from the plough horizon at a site close to Bayreuth. In comparison to the modified van Genuchten–Mualem–Durner model (MVGMD), for which we estimated 8 model parameters by applying the SCE-UA algorithm, the free-form algorithm yielded a significantly better fit to the experimental data. Moreover, the free-form approach yielded optimal SHPF showing significantly less interaction between the estimated model parameters than the MVGMD with 8 df. This shows that the model structure defined by the free-form approach is easier to invert given the initial

Figure 5. Measured and model-predicted values for cumulative outflow and the pressure head at 3.5 and 9.0 cm distance from the top of the column for Bayreuth sandy loam. Shown are the results of the free-form algorithm with $r = 5$ and the parametric fit, obtained from fitting the MVGMD model with 8 df (including $\tau$ and $h_s$).

Figure 6. Identified hydraulic properties and corresponding 95% confidence intervals (shaded areas) for $\theta(h)$ and $K(h)$ for the Bayreuth sandy loam column ($r = 5$).
and boundary conditions of the multistep outflow experiment. This should be viewed against the background that the use of parametric approaches for the SHPF is often claimed to regularize the inverse problem because the number of unknowns can be significantly reduced. In fact, restricting the number of unknowns is only helpful if the resulting model parameters interact only weakly.

[54] Our algorithm should be very helpful to the scientific community in addressing some of the most interesting and stimulating problems in unsaturated flow research. Because of its very high flexibility, we believe that any flow process that can be adequately described by the Richards equation can be successfully simulated by the free-form approach. This is not necessarily the case for the usual parametric approaches. Therefore limitations in the applicability of the Richards equation [see, e.g., Bayer et al., 2004] should be identifiable by the free-form approach, as inadequacies in the SHP model can effectively be excluded. This appears a fruitful idea for studying phenomena like dynamic effects in porous media on the laboratory scale. Other potential applications comprise the derivation of effective properties on larger scales and the phenomenon of hysteresis in unsaturated flow.

References


W. Durner and S. C. Iden, Institut für Geökologie, Technische Universität Carolo-Wilhelmina zu Braunschweig, Braunschweig D-38106, Germany. (s.iden@tu-bs.de)