Experimental design for outflow experiments based on a multi-level identification method for material laws

To cite this article: Sandro Bitterlich and Peter Knabner 2003 Inverse Problems 19 1011

Related content
- Direct analytic model of the L-curve
  P J Mc Carthy
- Iterative regularization of parameter identification problems by SQP methods
  Martin Burger and Wolfram Mühruber
- Efficient methods for L-hypersurface
  Murat Belge, Misha E Kilmer and Eric L Miller

Recent citations
- Nonlinear ill-posed problem analysis in model-based parameter estimation and experimental design
  C. Diana C. López et al
- Optimal experimental design for nonlinear ill-posed problems applied to gravity dams
  Tom Lahmer
- Optimal experimental design of ill-posed problems: The METER approach
  Andrés and Bardow
Experimental design for outflow experiments based on a multi-level identification method for material laws

Sandro Bitterlich and Peter Knabner

Institute for Applied Mathematics, Friedrich-Alexander University of Erlangen-Nuremberg, Martensstr. 3, 91058 Erlangen, Germany

E-mail: bitterli@am.uni-erlangen.de and knabner@am.uni-erlangen.de

Received 5 March 2003, in final form 30 June 2003
Published 31 July 2003
Online at stacks.iop.org/IP/19/1011

Abstract
A multi-level method for the determination of nonlinear coefficient functions of parabolic partial differential equations is presented as an example of the reconstruction of the hydraulic properties of porous media from soil column outflow experiments. A linear stability and error analysis is used to characterize the ill-posedness of the inverse problem. Founded on this analysis, optimal experimental design problems can be formulated. Numerical examples are presented which show that the method described leads to a stable solution of the inverse problem.

1. Introduction

The knowledge of characteristic material laws is an essential condition for the understanding of transport and reaction processes in porous media. Only if these material laws are well known can a physical process be simulated by the numerical solution of an appropriate mathematical model. Often, it is not possible to measure the material properties of porous media directly. Therefore, it is common to perform a suitable laboratory experiment, providing an indirect measurement by means of the medium’s reaction to a specific excitation. Mathematically, this leads to an inverse problem in which unknown solution-dependent coefficient functions are determined from experimental (boundary) observations of the solution of a corresponding direct problem in the form of a nonlinear parabolic partial differential equation.

On the one hand, the experiment must ensure the identifiability of the material properties and, on the other hand, a bounded amplification of small errors of the observations is desirable, i.e. an inverse problem which is as little ill-posed as possible. If more than one type of experiment is feasible, the experiment which minimizes the ill-posedness of the inverse problem in an appropriately defined sense is chosen for the determination of the material properties.

To obtain stable solutions of the inverse problem a regularization method must be applied. We will use a projection method defined by an interpolation with \( B \) splines for the regularization.
of the identification of nonlinear coefficient functions of a parabolic partial differential equation. Further stabilization can be achieved by restricting the resulting parameter space with linear inequality constraints which are derived from monotonicity conditions for the coefficient functions which have to be determined.

In this paper we consider a specific example: the identification of the unsaturated hydraulic properties of porous media from soil column outflow experiments. The identification problem is transformed into the minimization of a least squares error functional. The ill-posedness will be characterized by a linear error analysis which is derived from a first-order expansion of the forward operator similar to Chavent et al. (1994). The identification is embedded in a multi-level algorithm which was developed by Igler (1998), Igler and Knabner (2000) for the determination of sorption characteristics by soil column breakthrough experiments. The principle of the identification method is also described by Bitterlich and Knabner (2001, 2002). There we turn our attention to efficient algorithms for computing the sensitivity matrix.

This paper is organized as follows. In section 2, we formulate the inverse problem and the corresponding direct problem. In section 3, we will describe the identification method and the stability and error analysis. The identification will also be considered from the statistical point of view. We will demonstrate the identification by means of numerical examples. In section 4, special experimental design problems for the identification of hydraulic properties are considered. Furthermore, examples are presented which illustrate the importance of the experimental conditions to the ill-posedness of the inverse problem. Finally, a heuristic algorithm for solving the design problem is presented.

2. Model statement

2.1. The Richards equation

The description of a non-stationary fluid transport through porous media is founded on the Richards equation. In our work we use the Richards equation in its pressure head form which consists of a volume balance equation:

$$\frac{\partial}{\partial t} \Theta(\psi) + \nabla \cdot q = 0$$

(1)

and Darcy’s law:

$$q = -K(\psi) \nabla (\psi + z).$$

(2)

Here, $\psi = \psi(x, t)$ (length) denotes the pressure head, $q = q(x, t)$ (length/time) is the volumetric flow rate per surface area and $z$ (length) is the height against gravitational direction. The nonlinear coefficient functions of the partial differential equation characterize the hydraulic properties of the porous medium. In detail these are the hydraulic conductivity $K(\psi)$ (length/time) and the fluid retention curve $\Theta(\psi)$ which describes the relation between the pressure head and the fluid content. The coefficients are assumed to be continuous functions of $\psi$.

With the Richards equation and appropriate initial and boundary conditions we can simulate soil column outflow experiments as they have been realized by Durner and Zurmühl (1998), Durner et al (1998, 1999a, 1999b). The principle of such an experiment is the following. A homogeneous vertically oriented and initially saturated soil column of length $L$ is drained by decreasing the pressure head at the bottom of the column. Fluid flow across the top of the column is not allowed. This set-up is modelled by: a Dirichlet condition for the outflow boundary and a homogeneous flux boundary condition at the top of the column.

The physical properties of the experiment allow us to describe the fluid flow in the column by the Richards equation in one spatial dimension, such that the resulting model of the direct problem is as follows:
\[ \frac{\partial}{\partial t} \Theta(\psi) + \vartheta \psi \frac{\partial}{\partial x} q = 0 \]
\[ q = -K(\psi)(\vartheta \psi - 1) \quad (x, t) \in \Omega \times (0, T) \]
\[ \psi(x, 0) = \psi_0(x) \quad x \in \Omega \]
\[ \psi_0 \quad \xi \in \Omega \quad \xi \in \Omega \times (0, T), \]
where \( \Omega = (0, L) \), \( \psi_0(x) \geq 0 \) and \( h(t) \) is a continuous suction head applied at the bottom of the column with \( h(0) = \psi_0(L) \) and \( h'(t) \leq 0 \).

2.2. Formulation of the inverse problem

In an outflow experiment, it is possible to measure the pressure head at a fixed space coordinate \( x_d \in [0, L] \) and the cumulative outflow at the bottom of the column for prescribed time points \( 0 \leq t_1 < \cdots < t_n \leq T \). If the coefficients \( \Theta \) and \( K \) are given, the outflow experiment may be simulated by solving (3). The measured data then correspond to the following computed values:
\[ \omega_i^{\psi} := \omega_i^{\psi}(x_d; t^i) \quad \omega_i^q := \int_0^{t^i} q(L, \tau) \, d\tau, \quad i = 1, \ldots, n. \]

The dependence of the output vectors \( \omega_{\psi} := (\omega_1^{\psi}, \ldots, \omega_n^{\psi}) \in \mathbb{R}^n \) and \( \omega_q := (\omega_1^q, \ldots, \omega_n^q) \in \mathbb{R}^n \) (note that \( q(L, t) \geq 0 \), see, e.g., DuChateau (1997)) on the coefficients \( \Theta \) and \( K \) will be indicated by the notation
\[ (\omega_{\psi}, \omega_q) = B(x_d; \Theta, K). \]

We can now formulate our identification problem.

Identification problem. Find coefficients \( \Theta \) and \( K \) such that the pair \((\Theta, K)\) reproduce the given measured data \( \omega_{\psi}^\star \) and \( \omega_q^\star \) which are obtained from an outflow experiment, i.e. the equation
\[ (\omega_{\psi}^\star, \omega_q^\star) = B(x_d; \Theta, K) \]
holds.

3. Solving the identification problem

3.1. Regularization technique

Following a common method for solving an inverse problem, see, for example, Tarantola (1987) and Isakov (1998), we transform the above identification problem into a minimization problem of the least squares error functional:
\[ J(\Theta, K) = \sum_{i=1}^n (\alpha_{\psi}^{i} (\omega_{\psi}^i - \omega_{\psi}^\star)^2 + \alpha_q^{i} (\omega_q^i - \omega_q^\star)^2) \]
with positive weighting factors \( \alpha_{\psi}^{i}, \alpha_q^{i} \in \mathbb{R}_+ \).

Minimization problem. Find coefficients \( \Theta_{\text{opt}} \) and \( K_{\text{opt}} \) with
\[ J(\Theta_{\text{opt}}, K_{\text{opt}}) = \inf_{\Theta, K} J(\Theta, K) \]
for the measured data \( \omega_{\psi}^\star \) and \( \omega_q^\star \). Equation (8) describes an optimization problem in an appropriate space of continuous functions, i.e. an infinite-dimensional problem has to be solved.
The experiment explores only a bounded pressure head interval $[\psi_\ast, 0]$ which is determined by the suction head $h(t)$: $\psi_\ast = -L + h(T)$. This property was proved by DuChateau (1997). Thus the experiment determines the unknown coefficients $\Theta(\psi)$ and $K(\psi)$ on a finite interval, only: if we change the coefficients $\Theta(\psi)$ and $K(\psi)$ for $\psi < \psi_\ast$, we obtain the same observations $\omega_\psi$ and $\omega_q$ as before, i.e. the coefficients are not identifiable outside the above interval. Therefore, we consider the space $C[\psi_\ast, 0]$. By means of an interpolation in a finite-dimensional subspace of $C[\psi_\ast, 0]$ with basis functions $\Phi_j$ ($j = 1, \ldots, r$), we can define a projection of $C[\psi_\ast, 0]$ functions onto this subspace:

$$\Pi_r f := \sum_{j=1}^r p_j \Phi_j \quad \forall f \in C[\psi_\ast, 0].$$

(9)

The values $p_1, \ldots, p_r$ are defined by the interpolation. In the case of a piecewise linear interpolation we use the hat functions corresponding to a given partition $\psi_\ast = \psi_1 < \cdots < \psi_r = 0$ as basis functions and the parameters $p_j$ ($j = 1, \ldots, r$) are given by $p_j := f(\psi_j)$. Analogously we can define an interpolation for smoother basis functions. In particular we will use quadratic $B$ splines as basis functions.

By equation (9) we may generate a sequence of projections which satisfy the following properties:

$$\lim_{r \to \infty} \|\Pi_r f - f\|_\infty = 0 \quad \forall f \in C[\psi_\ast, 0]$$

(10)

and

$$\Pi_r C[\psi_\ast, 0] \subset \Pi_{r'} C[\psi_\ast, 0] \quad \text{for } r \leq r'.$$

(11)

Because a function of the space $\Pi_r C[\psi_\ast, 0]$ is uniquely defined by parameters $p_1, \ldots, p_r$, we identify the projection space $\Pi_r C[\psi_\ast, 0]$ with the corresponding parameter space $P_r \subset \mathbb{R}^r$. The dimension $r$ of the parameter space $P_r$ is called the number of degrees of freedom.

Now, we replace the minimization problem equation (8) by a finite-dimensional parameter minimizing problem.

Parameter minimizing problem. Find parameter vectors $p^\Theta_{r, \text{opt}}, p^K_{r, \text{opt}} \in P_r$ with

$$J(p^\Theta_{r, \text{opt}}, p^K_{r, \text{opt}}) = \inf_{p^\Theta, p^K \in P_r} J(p^\Theta, p^K),$$

(12)

where $J(p^\Theta, p^K) = J(\Pi_r \Theta, \Pi_r K)$ with

$$\Pi_r \Theta = \sum_{j=1}^r p^\Theta_j \Phi_j \quad \text{and} \quad \Pi_r K = \sum_{j=1}^r p^K_j \Phi_j$$

for a given $r \in \mathbb{N}_+$. Analogously to Louis (1989) we have defined a regularization strategy which is given by a sequence of appropriate projections, where $r$ is the regularization parameter. Note that problem (12) is not well defined in general, as the direct problem (3) requires certain conditions on $\Theta$ and $K$ to exhibit a unique solution. Sufficient conditions are that the coefficients are non-negative and monotone increasing. These are enforced (approximately) by linear constraints based on a partition $\psi_\ast = \psi_1 < \cdots < \psi_s = 0$ of $[\psi_\ast, 0]$. Furthermore, the unsaturated fluid content and the saturated conductivity can be measured independently from the outflow experiment such that we may assume knowledge of

$$\Theta(0) = \Theta_{\text{sat}} \quad \text{and} \quad K(0) = K_{\text{sat}}.$$

Thus the problem to deal with is as follows.
Constrained minimization problem. Find a pair of parameter vectors \((p_{r,\text{opt}}^{\Theta}, p_{r,\text{opt}}^{K})\) ∈ \(\mathbb{R}^r \times \mathbb{R}^c\) with

\[
J(p_{r,\text{opt}}^{\Theta}, p_{r,\text{opt}}^{K}) = \min_{p_{r,\text{opt}}^{\Theta}, p_{r,\text{opt}}^{K} \in \tilde{P}_r} J(p_{r,\text{opt}}^{\Theta}, p_{r,\text{opt}}^{K}).
\] (13)

The restricted parameter set \(\tilde{P}_r \subset P_r\) is defined by the constraints
\[
\begin{align*}
\Pi_r \Theta(\psi_r) &= \Theta_{\text{sat}}, \\
\Pi_r \Theta(\psi_1) &\geq 0, \\
\Pi_r K(\psi_r) &= K_{\text{sat}}, \\
\Pi_r K(\psi_1) &\geq 0
\end{align*}
\]

and
\[
\begin{align*}
\Pi_r \Theta(\psi_{k+1}) &\geq \Pi_r \Theta(\psi_k), \\
\Pi_r K(\psi_{k+1}) &\geq \Pi_r K(\psi_k) \quad \text{for } k = 1, \ldots, s - 1.
\end{align*}
\]

For the specific shape functions \(J\) (without regard to the parameter restrictions). This leads to
\[
\begin{align*}
\text{min} & \quad J(p_{r,\text{opt}}^{\Theta}, p_{r,\text{opt}}^{K}) \\
\text{s.t.} & \quad \psi \in \Omega, \quad \omega \in \Omega_r, \quad \kappa \in \Omega_k,
\end{align*}
\]

with
\[
\begin{align*}
\omega &\in \Omega_r, \\
\kappa &\in \Omega_k
\end{align*}
\]

We solve the above constrained optimization problem by a sequential quadratic programming method which, for instance, is described by Schittkowski (1988). Instead of solving the constrained minimization problem it is possible to minimize a penalty function, e.g.

\[
J(p_{r,\text{opt}}^{\Theta}, p_{r,\text{opt}}^{K}) + v_1 |\Pi_r \Theta(\psi_r) - \Theta_{\text{sat}}|^2 + v_2 |\Pi_r K(\psi_r) - K_{\text{sat}}|^2
\]

\[
+ \mu_1 g(\Pi_r \Theta(\psi_1)) + \sum_{k=1}^{r-1} \mu_{k+1} g(\Pi_r \Theta(\psi_{k+1}) - \Pi_r \Theta(\psi_k))
\]

\[
+ \mu_{s+1} g(\Pi_r K(\psi_1)) + \sum_{k=1}^{s-1} \mu_{s+k+1} g(\Pi_r K(\psi_{k+1}) - \Pi_r K(\psi_k))
\] (14)

for penalty parameters \(v_1, v_2 > 0\) and \(\mu_i > 0, i = 1, \ldots, 2s\), where
\[
g(x) = \begin{cases} |x|^2 & \text{for } x < 0, \\ 0 & \text{otherwise} \end{cases}
\]

3.2. Stability and error analysis

In this subsection, we provide a possibility for the characterization of the ill-posedness of our parameter identification problem. Similar to Chavent et al (1994) and Igler and Knabner (2000) we consider data \(\omega^*_y, \alpha^*_y\) with the minimum \(p_{r,\text{opt}}^{\Theta}, p_{r,\text{opt}}^{K}\) and a variation \(\omega^*_y + \delta \omega^*_y, \alpha^*_y + \delta \alpha^*_y\). To approximate the corresponding variations \(\delta p_{r,\text{opt}}^{\Theta}, \delta p_{r,\text{opt}}^{K}\) in the parameter vectors \(p_{r,\text{opt}}^{\Theta}, p_{r,\text{opt}}^{K}\), we apply a Gauss–Newton step in consideration of the optimality of \(p_{r,\text{opt}}^{\Theta}, p_{r,\text{opt}}^{K}\) (without regard to the parameter restrictions). This leads to

\[
J(\delta p_{r,\text{opt}}^{\Theta}, \delta p_{r,\text{opt}}^{K}) = \sum_{i=1}^{n} \left( \frac{d\left(\sqrt{\alpha^*_y \omega^*_y, \text{opt}}\right)}{d(p_{r,\text{opt}}^{\Theta}, p_{r,\text{opt}}^{K})} \delta(p_{r,\text{opt}}^{\Theta}, p_{r,\text{opt}}^{K}) - \sqrt{\alpha^*_y \delta \omega^*_y}\right)^2 \\
+ \left( \frac{d\left(\sqrt{\alpha^*_y \omega^*_y, \text{opt}}\right)}{d(p_{r,\text{opt}}^{\Theta}, p_{r,\text{opt}}^{K})} \delta(p_{r,\text{opt}}^{\Theta}, p_{r,\text{opt}}^{K}) - \sqrt{\alpha^*_y \delta \alpha^*_y}\right)^2 \rightarrow \min_{\delta p_{r,\text{opt}}^{\Theta}, \delta p_{r,\text{opt}}^{K} \in \tilde{P}_r}
\] (15)

where the data \(\omega^*_y, \text{opt}\) and \(\alpha^*_y, \text{opt}\) correspond to the solution vectors \(p_{r,\text{opt}}^{\Theta}, p_{r,\text{opt}}^{K}\) of (13). Thus these deviations are coupled by the (weighted) sensitivity matrix:

\[
S_r = \frac{d\left(\sqrt{\alpha^*_y \omega^*_y, \text{opt}}\right)}{d(p_{r,\text{opt}}^{\Theta}, p_{r,\text{opt}}^{K})} \in \mathbb{R}^{2n \times 2r}.
\] (16)
A solution of (15), i.e. of
\[ S_r(\delta p^\Theta_r, \delta p^K_r) = \delta \omega^* \]
with
\[ \delta \omega^* = \left( \sqrt{\alpha^*_i \omega^*_i} \right)_{i=1,...,n} \in \mathbb{R}^{2n}, \]
in the least squares sense of (15) is given by means of a singular value decomposition of the sensitivity matrix \( S_r \). We obtain for \( n \geq r \), and assuming a full rank \( 2r \) for \( S_r \), the following identification-measurement-error relation:
\[ (\delta p^\Theta_r, \delta p^K_r) = \sum_{j=1}^{2r} \sigma_j^{-1} \langle \delta \omega^*, u_j \rangle v_j. \]
(18)

Here \( \{\sigma_j; u_j, v_j\}_{j=1,...,2r} \) denotes a singular system of the sensitivity matrix. Relation (18) shows that the transfer of potential measurement errors to corresponding identification errors is mainly influenced by the singular values \( \sigma_j \) of the sensitivity matrix.

According to Louis (1989) we characterize the ill-posedness of the identification problem by the expansion behaviour of the singular values. For this we define the (generalized) condition number of the sensitivity matrix \( S_r \):
\[ \mu_{\text{cond}}(S_r) := \max_{i,j=1,...,2r} \frac{\sigma_i}{\sigma_j}, \]
and the maximum error amplification:
\[ \mu_{\text{max}}(S_r) := \max_{j=1,...,2r} \frac{1}{\sigma_j}. \]

These two characteristic numbers can be used to determine the most appropriate number of degrees of freedom for the parametrization of the unknown coefficient functions. If the values of \( \mu_{\text{cond}}(S_r) \) and \( \mu_{\text{max}}(S_r) \) are small, we expect stable solutions of the identification problem.

It is also possible to disregard the weighting factors in equation (15). The reason for including the weighting factors in the sensitivity matrix is explained in the following subsection.

3.3. Maximum likelihood estimators

The parameters which minimize the least squares error functional can be interpreted as maximum likelihood estimations for the unknown coefficient functions. The applied experimental measurement methods justify the following additional assumption.

**Assumption.** Let \( \omega_\psi \) and \( \omega_q \) be the error-free observations for the coefficient functions that have to be determined. The measurement data \( \omega^*_\psi \) and \( \omega^*_q \), which are obtained from the soil column outflow experiment, are assumed to be uncorrelated and Gaussian distributed random vectors with mean values \( \omega_\psi \) and \( \omega_q \), respectively.

With this assumption the minimization of the least squares functional (15) with \( \delta \omega^*_\psi = \omega^*_\psi - \omega_\psi \) and \( \delta \omega^*_q = \omega^*_q - \omega_q \) is equivalent to the maximum likelihood method for point estimations of \( \delta p^\Theta_r, \delta p^K_r \) if the weighting factors \( \alpha'_i, \alpha'_q \) for \( i = 1, ..., n \) are the inverse values of the variances of the measurements (see, e.g., Cramér 1966). By this we know that the estimated values of \( \delta p^\Theta_r, \delta p^K_r \) and so the estimated parameter vectors \( p^\Theta_{r,\text{opt}}, p^K_{r,\text{opt}} \) are asymptotically unbiased and Gaussian distributed with covariance matrix
\[ \text{cov}(p^\Theta_{r,\text{opt}}, p^K_{r,\text{opt}}) \approx (S^T_r S_r)^{-1}. \]
(19)
Thus the inverse squares of the singular values of the sensitivity matrix simplistically approximate the eigenvalues of the covariance matrix of the maximum likelihood estimators. Note that a linearized model was used for the derivation of the covariance matrix. Therefore the approximation error of the covariance matrix increases for increasing measurement errors.

3.4. Multi-level algorithm and implementation

Now, we have put together the instruments which we need to formulate the numerical identification algorithm. We start the identification for a parametrization with the least possible number of degrees of freedom. After solving the constrained minimization problem we interpolate the resulting fitted coefficients to a parametrization with one or more degrees of freedom added. This sequential procedure will be continued until a stopping criterion is satisfied.

The crucial point is to define a reasonable stopping criterion. The objective of the stopping criterion is the \textit{a posteriori} determination of the optimal regularization parameter.

Isakov (1998) quotes the discrepancy principle, which selects the regularization parameter that is consistent with the accuracy of the data. The discrepancy principle is generalized by Hanke \textit{et al} (1995) and used by Scherzer (1998) in a multi-level method in conjunction with Landweber’s iteration for solving nonlinear ill-posed problems. But the convergence results of Scherzer (1998) need assumptions for the nonlinear forward operator, which are not satisfied for our problem.

The \textit{L}-curve method is another approach frequently used to find the optimal regularization parameter, see, for example, Hansen (2001) and Vogel (1996). If the regularization is done by Tikhonov’s method, the \textit{L} curve is a plot, on a log–log scale, of the norm of the regularized solution versus the corresponding residual norm. The \textit{L}-curve criterion is used to pick the regularization parameter which corresponds to the ‘corner’ of the curve. This method has, besides its advantages, some limitations. The method, for example, fails to converge for a certain class of problems, see Vogel (1996).

The stopping criterion used in this paper is explained by the stability analysis. According to it, we terminate the multi-level identification if

\[ \varepsilon \cdot \mu_r > \mu_{tol} \]

is satisfied for the first time (or the number of degrees of freedom \( r \) exceeds a given maximum value \( r_{\text{max}} \)). \( \mu_r \) denotes one of the two characteristic numbers \( \mu_{\text{cond}}(S_r) \) or \( \mu_{\text{max}}(S_r) \), \( \mu_{\text{tol}} > 0 \) is a predefined tolerance parameter and \( \varepsilon > 0 \) again indicates the measurement error level.

A summary of the multi-level algorithm is given in table 1. The step size \( \Delta r \) and the maximum number of degrees of freedom have to be chosen appropriately.

By the interpolation of the optimization result for \( r \) degrees of freedom a good initial value is available for solving the constraint minimization problem with \( r + \Delta r \) degrees of freedom. Because of that the optimization by the sequential quadratic programming algorithm is accelerated which is, in general, slowly convergent for a high number of degrees of freedom. Moreover, we have to choose an initial parameter guess only in the first step of the multi-level algorithm. Test runs show that the identification results are nearly independent of the choice of this initial value, provided that sufficiently many iterations in the sequential quadratic programming algorithm are performed in each step of the multi-level algorithm such that the optimality condition is satisfied. These are the advantages of the multi-level algorithm.

The solution of the constrained minimization problem includes the solution of the direct problem. Numerically, the solution of the direct problem is approximated by a hybrid mixed finite element method developed by Schneid (2000). Efficient methods for the computation
of the sensitivity matrix are described by Bitterlich and Knabner (2002). The gradient of the least squares error functional, which is needed for solving the minimization problem, can be computed by means of the sensitivity matrix.

3.5. Numerical examples

To examine the behaviour of the multi-level identification algorithm, we consider soil columns of length $L = 15$ cm with known hydraulic properties $\Theta$ and $K$ described by the van Genuchten–Mualem model (cf van Genuchten 1980):

$$
\Theta(\psi) = \theta_{\text{res}} + (\theta_{\text{sat}} - \theta_{\text{res}}) \left( \frac{1}{1 + (\alpha \psi)^m} \right)^{(m-1)/m}
$$

$$
K(\psi) = K_{\text{sat}} \frac{(1 - (-\alpha \psi)m^{-1})(1 + (-\alpha \psi)m)^{(1-m)/m})^2}{(1 + (\alpha \psi)^m)(m-1/2m)}
$$

for $\psi < 0$ with the parameter values $\theta_{\text{sat}} = 0.52$, $\theta_{\text{res}} = 0.218$, $K_{\text{sat}} = 1.3167$, $\alpha = 0.0115$ and $m = 7.0$ for example 1, and $\theta_{\text{sat}} = 0.45$, $\theta_{\text{res}} = 0.0.05$, $K_{\text{sat}} = 2.1$, $\alpha = 0.003$ and $m = 2.1$ for example 2. The exact solution of the direct problem is approximated by the solution of the discrete direct problem for 3000 elements and $\Delta t = 0.05$. The pressure head at the bottom of the column decreases linearly in $T = 40$ h from 15 to $-120$ cm (example 1) and in $T = 100$ h from 15 to $-500$ cm (example 2), respectively. The observation of the pressure head takes place at the top of the column ($x_d = 0$ cm). The artificial measurement data are obtained by adding a 5% Gaussian noise. The coefficient spaces are given as spaces of piecewise quadratic functions. We start the identification with three degrees of freedom for every coefficient and refine the parametrization by successive equidistant bisections of the $\psi$ partition of the interval $[-140, 0]$ (example 1) and $[-520, 0]$ (example 2), respectively. In example 1 we have parametrized the logarithmic hydraulic conductivity $\log(K)$. During the identification, the experiment is simulated by a discretization with 100 elements. The time discretization is determined by the measurement time points with at least one intermediate step. Following the recommendation of Gribb (1996), we set the weighting factors of the least squares error functional to be the inverse squares of the measurement data:

$$
\alpha_{\psi}^i = \left( \frac{1}{n} \sum_{k=1}^{n} \alpha_{\psi}^k \right)^{-2}
$$

and

$$
\alpha_{\theta}^i = \left( \frac{1}{n} \sum_{k=1}^{n} \alpha_{\theta}^k \right)^{-2}
$$

for $i = 1, \ldots, n$.  

---

**Table 1. Multi-level algorithm.**

| Choose initial parameter vectors $p_{\text{res}}^{\Theta}, p_{\text{res}}^{K}, p_{\text{res}}^{\Theta}$, $p_{\text{res}}^{K} \in P_{\text{res}}$ for $\Theta_{\text{res}}$, $\Theta_{\text{res}}, K_{\text{res}}, K_{\text{res}}$. |
| $r := r_{\text{start}}$ |
| DO |
| Compute the projections $\Pi_r \Theta_{\text{res}}, \Pi_r K_{\text{res}}$ and the corresponding parameter vectors $p_{\text{res}}^{\Theta}, p_{\text{res}}^{K}$. |
| Solve the constrained minimization problem with initial values $p_{\text{res}}^{\Theta}$ and $p_{\text{res}}^{K}$. |
| Store the results in $p_{\text{res}}^{\Theta}$ and $p_{\text{res}}^{K}$. |
| Compute the sensitivity matrix $S_r$. |
| Compute the singular value decomposition of $S_r$ and $\mu_r$. |
| $r := r + \Delta r$ |
| WHILE $\epsilon \cdot \mu_r < \mu_{\text{tol}}$ and $r < r_{\text{max}}$ |

S Bitterlich and P Knabner
Multi-level identification of material laws 1019

0.2
0.25
0.3
0.35
0.4
0.45
0.5
-120
-100
-80
-60
-40
-20
0
Theta
Psi

(a)

(b)

(c)

(d)

Figure 1. Hydraulic properties and observations with 10 degrees of freedom. (a) Retention curve. (b) Hydraulic conductivity. (c) Pressure head for \( x_d = 0.0 \) cm. (d) Cumulative outflow.

We compare the identified functions and the reconstructed observations with the original coefficients and the ‘measured’ observations. Figures 1, 2 and 4, 5 and show the results with \( r = 10 \) and 34 degrees of freedom, respectively, for two different simulated experiments. We can observe that the identified coefficient functions, especially the retention curve \( \Theta \), are greatly disturbed by oscillations if the condition number of the sensitivity matrix is too high. These oscillations are influenced by the measurement errors, as a parametrization with a high number of degrees of freedom allows a better reconstruction of local distributions of the observations. This leads to the following behaviour (compare figures 3 and 6). First, the identification error decreases. Then the identification error increases if the parametrization exceeds a critical number of degrees of freedom. This critical number is to be found by the stability analysis.

**Example 1.** See figures 1–3.

**Example 2.** See figures 4–7.

If we plot analogously to the \( L \)-curve method the value of the error functional at the optimal solution point versus the condition number of the corresponding sensitivity matrix (see figure 7 for the example presented by Bitterlich and Knabner (2002), where the weighting factors are disregarded in the sensitivity matrix and the van Genuchten–Mualem parameters from example 2, but with \( m = 2.03 \), are used), we observe that the curve obtained shows behaviour similar to a \( L \) curve. Here the ‘corner’ corresponds to \( r = 6 \) degrees of freedom. But if we check the identification error of the retention curve, we find the ‘optimal’ regularization parameter \( r = 10 \). Only the identification error of the conductivity is insignificantly smaller, proportional to its size, for \( r = 6 \) degrees of freedom. The examples shown in this paper yield
the same result. Consequently, the L-curve method can, in general, not be applied to find the ‘optimal’ number of degrees of freedom.

4. Optimal design of experiments

4.1. Problem formulation and examples

On the one hand, we conclude from numerical results (shown in figure 8) that a fine enough space discretization has no significant influence on the stability of the identification procedure. On the other hand, we may suppose that the ill-posedness of the inverse problem critically depends on the experimental scenario. In case the experiment scenario influences the ill-posedness, we want to determine the hydraulic properties by that experiment which achieves the best stability of the identification procedure.

We denote the set of possible experimental scenarios by $\Xi$. For a fixed experimental scenario $\xi \in \Xi$, the multi-level algorithm computes an optimum number of degrees of freedom $r_{opt}$ by the stopping criterion. To indicate the dependence of this optimum number on the experimental scenario $\xi$ we use the notation

$$r_{opt} = r_{opt}(\xi).$$

We now formulate the following design problem.
Multi-level identification of material laws

- **Optimal experimental design problem.** Find an experimental scenario $\xi^* \in \Xi$ which minimizes the ill-posedness of the relevant inverse problem, i.e.

  $$\mu(S_{r_{opt}}(\xi^*)) = \inf_{\xi \in \Xi} \mu(S_{r_{opt}}(\xi)).$$

  (20)

  where $\mu$ stands for the condition number of the sensitivity matrix $\mu_{\text{cond}}$ or the maximum error amplification $\mu_{\text{max}}$.

  The type of shape functions for defining the coefficient spaces must not be changed in (20). In the following examples we restrict to quadratic B splines as shape functions. If we use the maximum error amplification, the condition (20) agrees with the E-optimality of the statistical design of experiments. In the same way, other common optimality criterions, e.g. described by Bandemer and Bellmann (1979), are conceivable instead of (20).

  In general, the function $\mu(S_{r_{opt}}(\xi))$ is neither differentiable nor continuous in $\xi \in \Xi$. A special feature is that $\mu(S_{r_{opt}}(\xi))$ is continuous only if $r_{opt}$ is independent of $\xi$, i.e. $r_{opt}(\xi)$ is constant. Consequently, customary numerical minimization schemes do not work. In practice, it is often sufficient to consider only a finite number of varying experimental scenarios and compare one with another. To demonstrate this we give the following example. (The data for the examples of this section are, in each case, generated by using the van Genuchten–Mualem parameters from example 2 but with $m = 2.03$.)

  **Example 3.** So far we have used measurements of the pressure head $\omega_{\psi}^* = \omega_{\psi}^*(0)$ at the top of the column for identification of the hydraulic properties. But the experimental measurement method allows pressure head measurements for an arbitrary space coordinate $x_d$. Therefore, we can specify a new experimental scenario by changing the measurement scheme to another

  ![Figure 3. Stability and error analysis. (a) Condition number of the sensitivity matrix. (b) Maximum error amplification. (c) Identification error of the retention curve. (d) Identification error of the hydraulic conductivity.](image-url)
Figure 4. Hydraulic properties and observations with 10 degrees of freedom. (a) Retention curve. (b) Hydraulic conductivity. (c) Pressure head for \( x_d = 0 \) cm. (d) Cumulative outflow.

coordinate \( x_d \) for the pressure head measurement, i.e. a set of possible experimental scenarios can be described by

\[
\Xi_d = \{ (\omega \psi(x_d), \omega q) : 0 \leq x_d < L \}. \tag{21}
\]

The numerical identification procedure involves a discretization of the direct problem (3). For one spatial dimension, the hybrid mixed finite element discretization approximates the pressure head constant on each element and in each node, such that, for a given finite element grid with \( N \) elements, it is suggested to let \( \Xi_d \) contain \( 2N \) different experimental scenarios (note that \( x_d < L \)) by setting

\[
\Xi_d = \left\{ (\omega \psi(x_d), \omega q) : x_d = \frac{i L}{2N}, i = 0, \ldots, 2N - 1 \right\} \tag{22}
\]

if equidistant nodes are used.

By a step-by-step shift of the coordinate \( x_d \) from the upper bound to the lower bound of the soil column, we can examine the stability of the identification procedure for each experimental scenario from \( \Xi_d \). Figure 9 compares the trend of the condition number and the maximum error amplification for six different locations of the pressure head measurement during the multi-level algorithm. In figure 10, \( \mu_{\text{cond}}(S_{\text{opt}}(\xi)) \) and \( \mu_{\text{max}}(S_{\text{opt}}(\xi)) \) for \( \xi \in \Xi_d \) are depicted for a further example of a soil column of length \( L = 15 \) cm with a space discretization by 100 elements. We have used \( r_{\text{max}} = 10 \) and \( \frac{1}{2} \cdot \mu_{\text{tol}} = 10000 \) for \( \mu_{\text{cond}} \) and \( \frac{1}{2} \cdot \mu_{\text{tol}} = 100 \) for \( \mu_{\text{max}} \), respectively, in the multi-level algorithm.

The above results show that the upper part of the column (0 \( x_d < \frac{L}{2} \)) is more qualified for the pressure head measurement than the lower part (\( \frac{L}{2} \leq x_d < L \)). If 0 \( x_d < \frac{L}{2} \), then the location \( x_d \) of the pressure head measurement does not essentially influence the stability of the
identification procedure. Figure 10 also shows the effect of the piecewise constant function $r_{opt}(\xi)$. For that example the multi-level algorithm yields

$$ r_{opt} = \begin{cases} 10 & x_d \in [0, 14.25], \\ 4 & x_d \in (14.25, 14.55), \\ 3 & x_d \in [14.55, 15], \end{cases} \quad \text{for } \mu_{\text{cond}} $$

and

$$ r_{opt} = \begin{cases} 10 & x_d \in [0, 14.25], \\ 3 & x_d \in (14.25, 15), \end{cases} \quad \text{for } \mu_{\text{max}}. $$

A decrease of $r_{opt}(\xi)$ corresponds to an instant decrease of $\mu_{\text{cond}}(\xi)$ and $\mu_{\text{max}}(\xi)$, respectively. But the values of $\mu_{\text{cond}}(\xi)$ and $\mu_{\text{max}}(\xi)$ cannot fall below the tolerance values.

From the technical point of view the pressure head can be measured simultaneously at more than one space coordinate. The experimental set-up used by Durner et al (1999b) records tensiometric pressure head data at two depths, one in the upper half and one in the lower half of a soil column with length $L = 15$ cm.

Thus, it is also possible to identify the hydraulic properties by using pressure head data from different spatial points. But case studies with artificial data as well as with real data show that, if the pressure head is well fitted at one space coordinate in the upper half of the column, then the pressure head is also matched in the lower half of the column. Hence, using additionally pressure head data does not, in general, yield better identification results. However, the identification without any pressure head data does not guarantee adequate results for the hydraulic properties if they are parametrized by means of $B$ splines.
Example 4. Another possibility for changing the experimental scenario is offered by varying the suction head. Geoscientists often drain the soil column by decreasing the pressure head at the bottom of the column step by step. This leads to so-called multi-step outflow experiments. Figures 11–13 present the numerical results for a multi-step outflow experiment with a suction head from 15 to $-100$ cm by five pressure head steps with $T = 20$ h and a 5% Gaussian noise. Already a parametrization with ten degrees of freedom shows an unexpectedly large identification error of the hydraulic conductivity.
We want to examine the influence of the number of steps in $h(t)$ on the stability of the identification procedure. Here the suction head is to be assumed as non-increasing in time, i.e. $\delta h(t, t') := h(t) - h(t') \geq 0$ for all $t < t'$ with $t, t' \in [0, T]$ (23)
Figure 11. Hydraulic properties and observations with 6 degrees of freedom for a multi-step outflow experiment. (a) Retention curve. (b) Hydraulic conductivity. (c) Pressure head for \( x_d = 0.0 \) cm. (d) Cumulative outflow.

and \( h(T) \) is set to be a predefined value \( \hat{h} < 0 \). Furthermore the pressure steps are chosen to be equidistant in time.

So we arrive at the following discrete but infinite set:

\[
\Xi_h = \{ \{h(t)\} : h \text{ is a step function on an equidistant partition with } \delta h(t,t') \geq 0, \ t < t', \ h(0) = \psi_0(L), \ h(T) = \hat{h} \}.
\]  

Figure 14 compares the stability of the identification for suction head functions with 1, 3, 7 and 15 steps, and furthermore for a continuous suction head which can be interpreted as the limiting function of multi-step suction head functions. From it we conclude that an outflow experiment with a continuous suction head provides the best stability, independent of the number of degrees of freedom. The more suction head steps are used the better stability of the identification procedure is achieved.

4.2. A stochastic algorithm in a search for solving the design problem

To solve the optimal design problem (20) for connected or (and) infinite sets \( \Xi \) we adapt the method of simulated annealing from combinatorial optimization. The simulated annealing algorithm is described in table 2. Here, it involves a Monte Carlo method with which the experimental scenario \( \xi' \) is generated randomly in a prescribed neighbourhood \( U_{\tau_k}(\xi) \) of \( \xi \). This neighbourhood is dependent on the actual temperature \( \tau_k \).

The algorithm accepts a neighbour scenario \( \xi' \) as the next scenario \( \xi \) if the stability of the identification problem corresponding to \( \xi' \) is better than that corresponding to the current scenario \( \xi \). Otherwise it will accepts \( \xi' \) with the probability \( e^{-\Delta s/\tau_k} \). As the temperature
Figure 12. Hydraulic properties and observations with 10 degrees of freedom for a multi-step outflow experiment with 5% Gaussian noise. (a) Retention curve. (b) Hydraulic conductivity. (c) Pressure head for \( \chi_d = 0 \). (d) Cumulative outflow.

Table 2. Simulated annealing algorithm.

1. Initializing
   \( k := 0 \)
   Select an initial temperature \( \tau_k = \tau_{\text{start}} > 0 \), a temperature step size \( \Delta \tau > 0 \) and \( l > 0 \).
   Generate an initial (random) scenario \( \xi \in \Xi \).

2. Generation of neighbour scenarios and cooling schedule
   WHILE \( \tau_k > 0 \)
   Perform the following \( l \) times.
   Generate a random scenario \( \xi' \in U_{\tau_k}(\xi) \subset \Xi \).
   \( \Delta \mu = \mu(\tilde{S}_{\text{opt}}(\xi')) - \mu(\tilde{S}_{\text{opt}}(\xi)) \)
   If \( \Delta \mu < 0 \) or \( e^{-\Delta \mu / \tau_k} \geq \text{RANDOM}[0, 1] \) then \( \xi := \xi' \).
   \( k := k + 1 \)
   \( \tau_k := \tau_{k-1} - \Delta \tau \)

decreases, so does the acceptance probability of a scenario that reduces the stability of the corresponding inverse problem. In this way we can avoid local optima. Besides the step size \( \Delta T \), which defines the cooling schedule, the number \( l \) of examined neighbour scenarios \( \xi' \) is a further control parameter of the simulated annealing algorithm which strongly influences the final experimental scenario. For more details on simulated annealing see, for example, Katayama and Narihisa (2001), Kirkpatrick et al (1983) and Kuhn (1992).

Example 5. In the following, we neglect the condition \( \delta h(t, t') \geq 0 \) for \( t < t' \). We test the simulated annealing algorithm for the design problem (20) with
Figure 13. Stability analysis for the multi-step outflow experiment. (a) Condition numbers. (b) Maximum error amplifications. (c) Identification error of the retention curve. (d) Identification error of the hydraulic conductivity.

Figure 14. Dependence of the stability of the identification on the number of steps for decreasing the pressure head at the outflow boundary. (a) Condition numbers. (b) Maximum error amplifications.

\[ \Xi_{h, pwc} = \{ h(t) : h(t) \in C[0, T], h(t)|_{(t_i, t_{i+1})} \in P^0(t_i, t_{i+1}) \} \]

\[ i = 0, \ldots, M - 1, h(0) = \psi_0(L), h(T) = h, \psi_0(L) \geq h(t) \geq \hat{h} \}, \quad (25) \]

for a given partition \( 0 = t^0 < \cdots < t^M = T \) of the time interval \([0, T]\). Here \( P^0 \) is the set of constant functions.

We use a soil column of length \( L = 15 \) cm with \( \hat{h} = -100 \) cm and \( T = 40 \) h. The partition of \([0, T]\) is chosen to be equidistant:

\[ t^i := i \frac{T}{M}, \quad i = 0, \ldots, M \]
Multi-level identification of material laws

Figure 15. Initial values and final results of the simulated annealing algorithm with piecewise constant suction head $h(t)$.

Figure 16. Range of the accepted objective function values dependent on the temperature $\tau_k$.

with $M = 16$. The control parameter of the simulated annealing algorithm is set to be

$$
\tau_{\text{start}} = 1000, \quad \Delta \tau = 20, \quad l = 20
$$

and for the neighbourhood $U_{\tau_k}(h(t))$ we use

$$
U_{\tau_k}(h(t)) := \left\{ \hat{h}(t) \in \Xi_{h,\text{pwc}} : h(t) - \gamma \tau_k \frac{\psi_0(L) - \hat{h}}{\tau_{\text{start}}} \leq \hat{h}(t) \leq h(t) + \gamma \tau_k \frac{\psi_0(L) - \hat{h}}{\tau_{\text{start}}} \right\}
$$

(26)

with $\gamma = 0.25$. The initial suction head is shown in figure 15 and the final suction head of the simulated annealing algorithm, as well. The condition number of the sensitivity matrix was chosen as the objective function. For the stopping criterion of the multi-level algorithm we have used $\frac{1}{c} \cdot \mu_{\text{tol}} = 10000$ and $r_{\text{max}} = 10$. Figure 16 shows the range of accepted values of the objective function $\mu_{\text{cond}}(S_{\nu+}(\xi))$ for $\xi \in \Xi_{h,\text{pwc}}$ dependent on the temperature $\tau_k$. We can see that the search space is reduced for increasing cooling.

5. Conclusions

We have developed an efficient numerical procedure for the identification of coefficients of nonlinear parabolic differential equations from column experiments. A hierarchical parametrization of the unknown coefficients by spline functions allows a step-by-step
correction of the reconstruction of experimental observations. The ill-posedness of the inverse problems is examined by a linear stability analysis. By means of this analysis several experimental designs can be examined.

Numerical experiments for the identification of the hydraulic properties from soil column outflow experiments show that the identification procedure is almost uninfluenced by the discretization level and is robust to Gaussian noise. Optimal design problems produce recommendations for appropriate suction head functions and pressure head measurement positions.

References

Bandemer H and Bellmann A 1979 Statistische Versuchsanalyse (Leipzig: Teubner)
Hanke M, Neubauer A and Scherzer O 1995 A convergence analysis of the Landweber iteration for nonlinear ill-posed problems Numer. Math. 72 21
Isakov V 1998 Inverse Problems for Partial Differential Equations (Berlin: Springer)
Kirkpatrick S, Gelatt C D Jr and Vecchi M P 1983 Optimization by simulated annealing Science 220 671
Kuhn H 1992 Heuristische Suchverfahren mit simulierter Abkühlung Wirtschaftswissenschaftliches Studium 8 387
Louis A K 1989 Inverse und schlecht gestellte Probleme (Stuttgart: Teubner)
Tarantola A 1987 Inverse Problem Theory (Amsterdam: Elsevier)
Vogel C R 1996 Non-convergence of the L-curve regularization parameter selection method Inverse Problems 12 535