Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Journal of Computational and Applied Mathematics 169 (2004) 1–15
JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS
www.elsevier.com/locate/cam

On the recovery of multiple flow parameters from transient head data

Ian Knowles[a,*], Tuan Le[b], Aimin Yan[a,1]

[a] *Department of Mathematics, University of Alabama at Birmingham, 452 Campbell Hall, 1530 3rd Avenue S, Birmingham, AL 35294-1170, USA*
[b] *Department of Mathematics, University of New Orleans, New Orleans, LA 70148, USA*

Received 12 November 2002; received in revised form 5 October 2003

Abstract

The problem of estimating groundwater flow parameters from head measurements and other ancillary data is fundamental to the process of modelling a groundwater system. We consider here a new method that allows for the simultaneous computation of multiple parameters as the unique minimum of a convex functional.

Keywords: Groundwater; Confined aquifer; Parameter estimation; Steepest descent

1. Introduction

We are concerned here with a new deterministic method for identifying the flow parameters in groundwater models. It is common to assume that groundwater flow in a confined isotropic aquifer is described by the equation

$$\nabla \cdot [K(x)\nabla w(x,t)] = S(x)\frac{\partial w}{\partial t} - R(x,t), \tag{1.1}$$

in which w represents the piezometric head, K the hydraulic conductivity, R the recharge–discharge, S the specific storage, $t > 0$ represents time, and x varies over some bounded region Ω of three-dimensional space representing the physical aquifer; see, for example, [2, (3.3.17)]. When the piezometric head w does not vary appreciably in the vertical dimension, the equation can be

* Corresponding author.
E-mail address: iwk@math.uab.edu (I. Knowles).
[1] Supported in part by US National Science Foundation Grants DMS-9805629 and DMS-0107492.

0377-0427/$ - see front matter
doi:10.1016/j.cam.2003.10.013

depth-averaged to obtain a two-dimensional formulation; in this case K becomes the transmissivity, S the (dimensionless) storativity, and R represents a combination of an averaged recharge-discharge and vertical leakage terms. We are interested in the problem of the simultaneous determination of the functions K, S, and R, from measured data on $w(x,t)$ at various points of Ω and over some time interval, together with values of K measured at some boundary locations. Much has already been written on various aspects surrounding this topic. In particular, the problem of the determination of K from steady state head data has been extensively (though apparently not definitively) studied, see for example [3,6,7], as well as [4,25] and the references therein for detailed survey information; the determination of K from transient data is discussed in [5,9,10,24]. There has also been some work on the determination of S [18] and R [22]. We note that there seems to be little documented work in the literature on the simultaneous determination of K, S, and R.

In [12] a new method for parameter estimation for elliptic equations (of which the steady state equation for groundwater flow is but one example) was introduced. In this method, the parameter estimation is accomplished by the minimization of a new functional which (as is shown in [12]) has the important property of being convex;[2] this gives the approach significant advantages over other methods, notably those of output least-squares type, in that the functional has a unique global minimum, with no possibility of the associated descent algorithms "getting stuck" in spurious local minima.

In this paper, we explore in detail some of the practical aspects of the descent algorithms associated with this approach. In particular, we show that the method is effective in simultaneously estimating multiple coefficients in these equations. This is important in groundwater modelling in that one cannot reasonably expect to effectively model a groundwater system without obtaining, in an appropriately objective manner, proper estimations, or measurements, of *all* of the coefficient functions in that system. In accomplishing this task, it is clear from a mathematical standpoint that steady state data is insufficient for specifying multiple coefficients. One is thus inevitably attracted to the greater information present in time varying head data. This in turn leads us to a consideration of the parabolic equation (1.1); we show in Section 2 that time varying data can be transformed to data for certain elliptic equations of the type discussed above, to which our descent methods may then be applied.

We note in passing that these methods are particularly effective when the underlying distributed parameters are discontinuous, a situation that one must assume to be the case a priori in a practical situation. We show also that the method can be adapted to obtain approximations to a time-varying recharge–discharge function $R(x,t)$; such estimates have proven particularly difficult to compute heretofore [1, p. 152]. The method also allows one to insert certain additional a priori information about the parameters being estimated directly into the algorithms. The ability to perform such insertions is an important factor in the numerical performance of the descent algorithms because the underlying problem of parameter estimation is quite ill-posed (i.e. any error in the measured data can lead to large errors in the estimated parameters), and a common, and natural, route to circumventing the numerical instabilities caused by ill-posedness is to input appropriate additional independent information (cf. [19]).

[2] To be more precise, the functional is convex under the typical conditions encountered in practice (see [12]).

2. Reformulating the time-dependent problem

By choosing new units for the time as necessary, one can assume that $0 \leqslant t \leqslant 1$. So, given measured head data $w(x,t)$, where w is considered to be a solution of Eq. (1.1), and measured values for $K(x)$ on the boundary of our physical region Ω, we seek to compute the functions K, S, and R, where for simplicity we temporarily assume that R does not depend on time, i.e. that $R = R(x)$.

We begin by transforming the solution data $w(x,t)$ of the parabolic Eq. (1.1) to data $u(x,\lambda)$, where

$$u(x,\lambda) = \int_0^1 w(x,t)\mathrm{e}^{-\lambda t}\,\mathrm{d}t \tag{2.1}$$

and $u(x,\lambda)$ satisfies an associated elliptic equation,

$$-\nabla\cdot[K(x)\nabla u(x,\lambda)] + \lambda S(x)u(x,\lambda) = R^*(x), \tag{2.2}$$

where

$$R^*(x) = \frac{\mathrm{e}^{-\lambda}-1}{\lambda}R(x) + S(x)[w(x,0) - w(x,1)\mathrm{e}^{-\lambda}]. \tag{2.3}$$

For any fixed $\lambda > 0$ it is a relatively simple matter to compute values $u(x,\lambda)$ from the known values for $w(x,t)$. We arrive then at a new problem: given $u(x,\lambda)$ for x in Ω and all $\lambda > 0$ (and K on the boundary of Ω), determine the functions K, S, and R.

3. The optimization method

We now apply the variational method proposed in [12] to (2.2). As mentioned above, we assume that $u(x,\lambda)$ is known as a solution of (2.2) for all x in the region, and all $\lambda > 0$. For each $\lambda > 0$ and functions k, s, and r let $v = u_{k,s,r,\lambda}$ be the unique solution of the boundary value problem

$$\begin{aligned} &-\nabla\cdot(k(x)\nabla v(x,\lambda)) + \lambda s(x)v(x,\lambda) = r^*(x),\\ &v|_{\partial\Omega} = u|_{\partial\Omega}. \end{aligned} \tag{3.1}$$

where

$$r^*(x) = \frac{\mathrm{e}^{-\lambda}-1}{\lambda}r(x) + s(x)[w(x,0) - w(x,1)\mathrm{e}^{-\lambda}]. \tag{3.2}$$

Notice that, in this notation $u = u_{K,S,R,\lambda}$, where K, S, and R are the functions that we seek to recover.

Consider now the functional $G(k,s,r,\lambda)$ given by

$$\begin{aligned} G(k,s,r,\lambda) = \int_\Omega & k(x)(|\nabla u|^2 - |\nabla u_{k,s,r,\lambda}|^2)\\ &+\lambda s(x)(u^2 - u_{k,s,r,\lambda}^2) - 2r^*(x)(u - u_{k,s,r,\lambda})\,\mathrm{d}x. \end{aligned} \tag{3.3}$$

This functional is a generalization of the functional used in [14] to effect numerical differentiation of a function of one variable; as is explained in the remark following [13, Theorem 2.1], the precise form arises from converting a constrained energy functional minimization to an unconstrained one

using Lagrange multipliers. It is also worth observing that the nonnegativity of this functional is equivalent to the validity of the Dirichlet principle for the associated positive self-adjoint elliptic differential operator; so the recovery of these coefficient functions via such functionals provides, roughly speaking, a kind of inverse Dirichlet principle for this situation.

The functional, H, that we actually minimize to recover the desired flow coefficients, is formed by choosing $n_{\max}$ unequal positive values of the λ parameter, $\lambda_1, \lambda_2, \ldots, \lambda_{n_{\max}}$, and then setting

$$H(k,s,r) = \sum_{i=1}^{n_{\max}} G(k,s,r,\lambda_i). \tag{3.4}$$

As we seek to determine three functions K, S, and R, it is natural to expect that one would need to use at least three of the functions $u(x,\lambda_i)$ in this process. That this is indeed the case follows from the uniqueness theorem in [12], where it is noted that one needs in addition that a certain vector field generated by the three solution functions generates no trapped orbits, a condition that is easily checked in practice via computer graphics generated directly from the computed data functions $u(x,\lambda_i)$, $1 \leqslant i \leqslant n_{\max}$ (see [15]). This condition is linked to the natural restriction on this inverse problem arising from the fact that in regions of no flow, one cannot expect to recover flow parameters by using only flow data. So, in the above we must always take $n_{\max} \geqslant 3$. In fact, it is advantageous to use $n_{\max} \gg 3$; we discuss this aspect in more detail later. We also note for later use that the same uniqueness theorem requires that K be known on the boundary of the groundwater region; further discussion on the use of prior information may be found in [8,9, Section 6].

For convenience, we list some of the properties of the functional G established in [12]. First, from [12, Theorem 2.1(i)]

$$G(k,s,r,\lambda) = \int_\Omega k(x)|\nabla(u - u_{k,s,r,\lambda})|^2 + \lambda s(x)(u - u_{k,s,r,\lambda})^2 \,\mathrm{d}x. \tag{3.5}$$

For k positive definite, $s \geqslant 0$, and $\lambda > 0$, one can see that we have $G(k,s,r,\lambda) \geqslant 0$ and we also have that $G(k,s,r,\lambda) = 0$ if and only if $u = u_{K,S,R,\lambda} = u_{k,s,r,\lambda}$. By a similar calculation to that of [12] we also have that the first variation (Gâteaux differential) of G is given by

$$\begin{aligned} G'(k,s,r,\lambda)[h_1,h_2,h_3] = \int_\Omega & (|\nabla u|^2 - |\nabla u_{k,s,r,\lambda}|^2)h_1(x) \\ & + [\lambda(u^2 - u_{k,s,r,\lambda}^2) + 2(\mathrm{e}^{-\lambda}w(x,1) - w(x,0))(u - u_{k,s,r,\lambda})]h_2(x) \\ & - 2\frac{\mathrm{e}^{-\lambda} - 1}{\lambda}(u - u_{k,s,r,\lambda})h_3(x)\,\mathrm{d}x. \end{aligned} \tag{3.6}$$

In this notation, the values of G' represent various directional derivatives for the functional G, with the functions h_i serving as the "directions" in which one might choose to vary k, s, or r; for example, if we set $h_2 = h_3 = 0$ then from Taylor's theorem for functionals, for all α small enough

$$G(k + \alpha h_1, s, r, \lambda) \approx G(k,s,r,\lambda) + \alpha G'(k,s,r,\lambda)[h_1,0,0] \tag{3.7}$$

and so a knowledge of $G'(k,s,r,\lambda)[h_1,0,0]$ allows us to estimate the difference $G(k + \alpha h_1, s, r, \lambda) - G(k,s,r,\lambda)$ when $\alpha > 0$ is not too large. In particular, in direct analogy with the gradient of a

function of several variables, we may take the function adjacent to h_1 in (3.6) to be the gradient of G with respect to k, $\nabla_k G$, i.e.

$$\nabla_k G(k,s,r,\lambda) = |\nabla u|^2 - |\nabla u_{k,s,r,\lambda}|^2. \tag{3.8}$$

Similarly

$$\nabla_s G(k,s,r,\lambda) = \lambda(u^2 - u_{k,s,r,\lambda}^2) + 2(\mathrm{e}^{-\lambda} w(x,1) - w(x,0))(u - u_{k,s,r,\lambda}), \tag{3.9}$$

$$\nabla_r G(k,s,r,\lambda) = -2\frac{\mathrm{e}^{-\lambda} - 1}{\lambda}(u - u_{k,s,r,\lambda}). \tag{3.10}$$

Exactly as in the multivariate case, these gradients allow us to use descent methods for our minimization; in particular, if we choose to set $h_2 = h_3 = 0$ and

$$h_1 = -\nabla_k G(k,s,r,\lambda),$$

we have that

$$G(k + \alpha h_1, s, r, \lambda) < G(k,s,r,\lambda)$$

for $\alpha > 0$ and not too large, and so we can (locally) minimize G in the direction of $h_1 = -\nabla_k G(k, s, r, \lambda)$ with one-dimensional search techniques. Later descent steps can minimize G in s and r as well. While the actual gradients that we use presently are somewhat different, the general idea is the same.

Notice that $G'(k,s,r,\lambda) = 0$ (i.e. $G'(k,s,r,\lambda)[h_1,h_2,h_3] = 0$ for all functions h_1, h_2, h_3) if and only if

$$|\nabla u|^2 - |\nabla u_{k,s,r,\lambda}|^2 = 0,$$

$$\lambda(u^2 - u_{k,s,r,\lambda}^2) + 2(\mathrm{e}^{-\lambda} w(x,1) - w(x,0))(u - u_{k,s,r,\lambda}) = 0,$$

$$2\frac{\mathrm{e}^{-\lambda} - 1}{\lambda}(u - u_{k,s,r,\lambda}) = 0,$$

which, from the form of (3.3), is true if and only if $G(k,s,r,\lambda) = 0$; we know already that this is true if and only if $u = u_{K,S,R,\lambda} = u_{k,s,r,\lambda}$ again.

We next observe that the functional H in (3.4) has very similar properties. In particular, essentially the same argument shows that $H \geqslant 0$, and that $H(k,s,r) = 0$ if and only if $u = u_{K,S,R,\lambda_i} = u_{k,s,r,\lambda_i}$ for all $1 \leqslant i \leqslant n$, and the derivative $H'(k,s,r) = 0$ if and only if $H(k,s,r) = 0$. By choosing $n_{\max} \geqslant 3$ and assuming that the vector field condition mentioned earlier holds, it now follows from the uniqueness result [12, Theorem 3.5] that (K,S,R) is not only the unique global minimum for H, but also the unique stationary point (one can also show from the second variation for H that under the same conditions H is actually a convex functional, but we omit the details). This is the ideal context for numerical minimization and suggests a natural path to the goal of simultaneously computing the functions K, S, and R.

4. Time dependent recharge–discharge

It is not clear (and possibly not true) that the measured data in this problem uniquely determines a fully time-dependent source term $R(x,t)$. However, if we assume that R is piecewise constant in

time, then we can adapt the above procedure to recover such an R. In this case, if we assume that $0 = t_0 < t_1 < \cdots < t_m = 1$ are fixed times in our given time period, our R then takes the form

$$R(x,t) = \sum_{i=1}^{n} R_i(x)\chi_{[t_{i-1},t_i]}(t), \tag{4.1}$$

where for each i,

$$\chi_{[t_{i-1},t_i]}(t) = \begin{cases} 1 & \text{if } t_{i-1} \leqslant t \leqslant t_i, \\ 0 & \text{otherwise.} \end{cases}$$

This assumption on R in effect assumes that, over the times $0 \leqslant t \leqslant t_1$, R is "frozen" as the function $R_1(x)$ of the space variables, and over $t_1 \leqslant t \leqslant t_2$ R is $R_2(x)$, etc; each function R_i is thus a snapshot of $R(x,t)$ over a part of the time measurement period. If the time sub-intervals are chosen sufficiently small, this allows us (in theory at least) to approximate the fully time dependent R as closely as we like.

Our inverse problem may then be stated as follows: given measured head data $w(x,t)$, where w is considered to be a solution of Eq. (1.1), and measured values for $K(x)$ on the boundary of our physical region Ω, we seek to compute the functions K, S, and R_i, $1 \leqslant i \leqslant n$.

As in Section 2, we can reformulate to an elliptic equation. At this juncture it is advantageous to observe that the procedure outlined above can be applied to each interval $[t_{i-1}, t_i]$, $1 \leqslant i \leqslant n$, as a separate calculation. So we set

$$u_i(x,\lambda) = \int_{t_{i-1}}^{t_i} w(x,t)\mathrm{e}^{-\lambda t}\,\mathrm{d}t \tag{4.2}$$

and, analogous to (2.2), we obtain for each i, $1 \leqslant i \leqslant n$,

$$-\nabla\cdot[K(x)\nabla u_i(x,\lambda)] + \lambda S(x)u_i(x,\lambda) = R_i^*(x), \tag{4.3}$$

where

$$R_i^*(x) = -\frac{1}{\lambda}R_i(x)[\mathrm{e}^{-\lambda t_{i-1}} - \mathrm{e}^{-\lambda t_i}] + S(x)[w(x,t_{i-1})\mathrm{e}^{-\lambda t_{i-1}} - w(x,t_i)\mathrm{e}^{-\lambda t_i}]. \tag{4.4}$$

So now, for each i, $1 \leqslant i \leqslant n$, we are given $u_i(x,\lambda)$ and we seek K, S, and R_i.

The functional G in this case has the form given by (3.3) where now $G = G(k,s,r_i,\lambda)$ and the term r_i^* (formerly defined by Eq. (3.2)) is given by

$$r_i^*(x) = -\frac{1}{\lambda}r_i(x)[\mathrm{e}^{-\lambda t_{i-1}} - \mathrm{e}^{-\lambda t_i}] + s(x)[w(x,t_{i-1})\mathrm{e}^{-\lambda t_{i-1}} - w(x,t_i)\mathrm{e}^{-\lambda t_i}] \tag{4.5}$$

and the solutions $u_{k,s,r,\lambda}$ are written $u_{k,s,r_i,\lambda}$. The gradients $\nabla_k G$ and $\nabla_s G$ are given, as before, by (3.8) and (3.9); in place of $\nabla_r G$, we have $\nabla_{r_i} G$, $1 \leqslant i \leqslant n$, where

$$\nabla_{r_i} G(k,s,r_i,\lambda) = -2\frac{\mathrm{e}^{-\lambda t_{i-1}} - \mathrm{e}^{-\lambda t_i}}{\lambda}(u - u_{k,s,r_i,\lambda}). \tag{4.6}$$

In this case, the functional H is again given by Eq. (3.4) with $n_{\max} \geqslant 3$, and the relevant uniqueness properties giving conditions on the appropriate vector field under which this H has a unique minimum (and a unique stationary point) at (K,S,R_i) are the same as above.

In the next section, we discuss the descent process in greater detail.

5. A descent algorithm

We now consider some of the details of our minimization procedure. The gradients defined by (3.8)–(3.10) are commonly termed $\mathscr{L}^2$-gradients because one can write (for example)

$$G'(k,s,r,\lambda)[h_1,0,0] = (\nabla_k G, h_1)_{\mathscr{L}^2},$$

where $(\cdot,\cdot)_{\mathscr{L}^2}$ denotes the standard inner product in the Hilbert space of square integrable functions, $\mathscr{L}^2(\Omega)$. In order keep the value of K on the boundary fixed throughout the descent process, we have found it advantageous to use a class of gradients introduced in [17]. These Neuberger gradients are a type of preconditioned (i.e. smoothed) gradient that generally give superior convergence in steepest descent algorithms. We shall use the notation $\nabla_k^{\mathrm{N}} G$ to denote the Neuberger smoothing of $\nabla_k G$, defined by

$$G'(k,s,r,\lambda)[h_1,0,0] = (\nabla_k^{\mathrm{N}} G, h_1)_{\mathscr{H}^1}, \tag{5.1}$$

where the above identity is to hold for all choices of h_1 belonging to the Sobolev space $\mathscr{H}^1(\Omega)$ consisting of all functions in $\mathscr{L}^2(\Omega)$ whose derivatives also lie in $\mathscr{L}^2(\Omega)$, and $(\cdot,\cdot)_{\mathscr{H}^1}$ denotes the inner product of functions in this Sobolev space. The Neuberger gradients $\nabla_s^{\mathrm{N}} G$ and $\nabla_r^{\mathrm{N}} G$ (or $\nabla_{r_i}^{\mathrm{N}} G$, $1 \leqslant i \leqslant n$) are defined analogously. In order to compute the Neuberger gradient $\nabla_k^{\mathrm{N}} G$ (for example) we merely have to solve the boundary value problem

$$-\triangle g + g = \nabla_k G, \tag{5.2}$$

$$g|_{\partial\Omega} = 0 \tag{5.3}$$

and note from [12, Eq (3.1)] that $g = \nabla_k^{\mathrm{N}} G$; notice here that, as $g|_{\partial\Omega} = 0$, the boundary data for K is preserved during the descent process. The Neuberger gradients $\nabla_s^{\mathrm{N}} G$ and $\nabla_r^{\mathrm{N}} G$ are computed in an analogous manner.

In implementing the descent procedure when $R = R(x)$, for example, one could choose to descend by varying all of k,s,r at each descent step. However, we have found that this strategy is not particularly efficient because the rate at which H decreases with respect to k is substantially smaller than that for s and r. So our general strategy is to proceed in cycles of three, with a greater number of descent steps allocated to descent with respect to k compared to descent with respect to s or r. For a given choice of the initial functions, k_0, s_0, r_0, one could use steepest descent, beginning with the direction $-\nabla_k^{\mathrm{N}} H(k_0,s_0,r_0)$, together with a one-dimensional search routine, to line minimize H at some point (k_1,s_0,r_0), where k_1 is the latest approximation to the function K (this step would normally be repeated a predetermined number of times); this would be followed with a line minimization in the direction $-\nabla_s^{\mathrm{N}} H(k_1,s_0,r_0)$ to obtain functions (k_1,s_1,r_0), and then by another line minimization in the direction $-\nabla_r^{\mathrm{N}} H(k_1,s_1,r_0)$ to obtain functions (k_1,s_1,r_1); this three step cycle would be repeated until convergence.

In practice one gets faster (by, roughly, a factor of two) convergence with the following adaption of the standard Polak–Ribière conjugate gradient scheme [20, p. 304]. The initial search direction is $h_0 = g_0 = -\nabla_k^{\mathrm{N}} H(k_0,s_0,r_0)$. At (k_i,s_i,r_i) one uses the approximate line search routine to minimize $H(k,s,r)$ in the direction of h_i, resulting in (k_{i+1},s_i,r_i). Then $g_{i+1} = -\nabla_k^{\mathrm{N}} H(k_{i+1},s_i,r_i)$, and

$h_{i+1} = g_{i+1} + \gamma_i h_i$, where

$$\gamma_i = \frac{(g_{i+1} - g_i, g_{i+1})_{\mathscr{H}^1}}{(g_i, g_i)_{\mathscr{H}^1}} = \frac{(g_{i+1} - g_i, \nabla_k H(k_i, s_i, r_i))_{\mathscr{L}^2}}{(g_i, \nabla_k H(k_i, s_i, r_i))_{\mathscr{L}^2}}.$$

At (k_{i+1}, s_i, r_i), one uses $\nabla_s H(k_{i+1}, s_i, r_i)$ in the same way to determine (k_{i+1}, s_{i+1}, r_i), and then $\nabla_r H(k_{i+1}, s_{i+1}, r_i)$ to obtain $(k_{i+1}, s_{i+1}, r_{i+1})$, whereupon the three-step cycle repeats.

When the recharge–discharge term is time dependent (according to the discussion in Section 4) we use the same process. We discuss some of the practical issues (like how large may one choose n) in the next section.

6. Implementation and results

We describe here some of our tests involving various choices of synthetically produced data, and later we consider some partial results from well data obtained over a period of about eight months at seven monitoring wells situated in the vicinity of the campus of the University of Alabama at Birmingham.

First, some general comments. It can be seen from the form of the gradient function (3.8) that one must be able to effectively take numerical partial derivatives of the data function u in order to implement the method. In the case of synthetic data, wherein the "data" u is actually found by initially solving the appropriate parabolic equation (and is therefore a smooth function) it is appropriate to use (quadratic) interpolation procedures to obtain the desired numerical derivatives. In the case of real well data, the measurements are inevitably contaminated with noise and one has to use a more sophisticated approach. Our procedure is as follows. First at each of the measurement times the head dataset is piecewise linearly interpolated and then smoothed with the aid of the Friedrichs mollifier function

$$\rho(x) = \begin{cases} \beta \exp\left(\dfrac{-1}{\|x\|^2 - 1}\right) & \text{if } \|x\| < 1, \\ 0 & \text{otherwise,} \end{cases}$$

where β is chosen so that $\int_{R^n} \rho = 1$, to regularize the data function u by

$$u_h(x) = h^{-n} \int_\Omega \rho\left(\frac{x - y}{h}\right) u(y)\, \mathrm{d}y \tag{6.1}$$

for some small, but not too small, $h > 0$; we used $h = 0.32$ here. One can then compute the numerical derivatives of u_h using central differences and use these as approximations to the derivatives of u.

We used several public domain PDE packages to solve the equations. For the elliptic boundary value problems, we mainly used the FIVE POINT STAR finite difference solver from the ELLPACK system [21]; to obtain parabolic synthetic data, we used the PDETWO solver [16]. Both of these solvers performed impeccably on the problems we considered. All the computations were performed on the UAB Department of Mathematics Sun Unix and Beowulf systems.

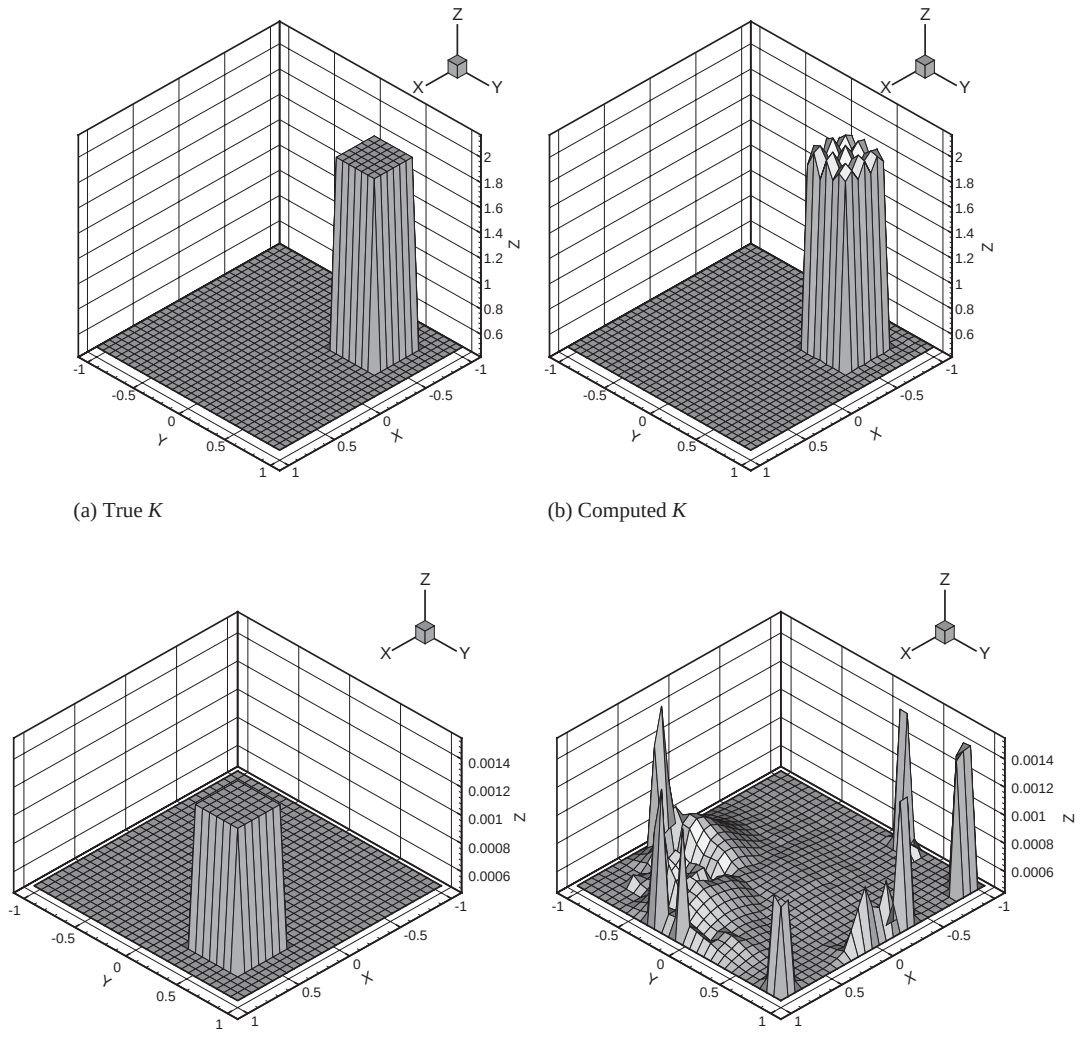
Parameter identification problems of the type considered here fall under the general heading of ill-posed inverse problems. From a practical standpoint, the fall-out from this observation is that one cannot expect to carry out these computations in a stable fashion without directly confronting

this issue. Many general methods for dealing with ill-posedness have been proposed, including Tikhonov regularization [23], limiting the number of grid points (ill-posedness tends to become more pronounced as the grids become finer), and limiting the number of iterations in iterative estimation procedures, and even casting out the direct approach in favour of a statistically based approach [11]. In the groundwater problem, there are difficulties associated with each of these choices: all Tikhonov regularization methods make use of a regularization parameter whose critical value must be known quite accurately for the method to be effective, and this can be problematical in the case of sparse, noisy aquifer data; if one limits the grid size too severely, the model error may increase unacceptably; if one limits the number of iterations, one may not be able to extract all of the information in the data.

In the case of the present algorithm, we observed in our early trials that the main symptom of ill-posedness in the computations was a tendency of the computed values for the hydraulic conductivity, K, to slowly become unbounded below. As the elliptic solvers are quite sensitive to a loss of positive definiteness for K, the program would crash quite quickly when negative values of K were encountered. Now with field data, one generally can input a reasonable estimate for a positive lower bound, $c > 0$, for the conductivity. We then modified the program so that at each descent step the values for k_i smaller than c were set equal to c (and similar cutoffs were incorporated into the computations of the other coefficients, whenever justifiable on physical grounds). The effect was quite dramatic: the algorithm became extremely stable, and we were able to let it run over hundreds of thousands of descent steps without serious degradation of the resulting images. In particular, it now became possible to simultaneously recover multiple coefficients, albeit at the cost of an increasing amount of computer time as the number of coefficients increased. It should be noted that in a typical least-squares minimization it is common to see large oscillations in the parameter values with unboundedness both from above and below. It appears that in our case, if one is to extrapolate from the computations exhibited here, the combination of an enforced lower bound and the convexity of the functional essentially eliminates the tendency for the parameter values to become unbounded above.

We also found that increasing the value of $n_{\max}$ in the defining equation for H (Eq. (3.4)) substantially improved the images; this is in line with the observation that ill-posedness is in some sense a manifestation of information loss, and so it makes sense that one should always strive to add information whenever possible. In the results below we typically used $n_{\max} = 20$, and we chose the λ_j so that $0 < \lambda_j \leqslant 1$. As λ is the transformed time parameter, it is not unreasonable to expect that using an even greater value of $n_{\max}$ would correspond to increasing the time resolution in the parabolic equation and should give even better results. In general, the method is flexible enough to allow the inclusion of multiple datasets, so that one may further decrease the natural ill-posedness associated with groundwater data. Mathematically, as we are minimizing a convex functional the only manifestation of ill-posedness is the "flatness" of H in a neighbourhood of the unique global minimum, and one would expect less flatness in the presence of additional data.

In Figs. 1 and 2, we demonstrate the simultaneous recovery of eight coefficient functions from known data on the solution $w(x,t)$, where x takes values in a two-dimensional region. We deliberately chose discontinuous K, S, and $R = R(x,t)$ for this test, both because the recovery of discontinuous functions is more difficult than the recovery of smooth ones, and because in the field, subsurface parameters are unlikely to be smooth functions. We assume that R has the form (4.1) where the time interval $0 \leqslant t \leqslant 1$ is divided into six equal sub-intervals (so, $n = 6$). In order to investigate



(a) True K

(b) Computed K

(c) True S

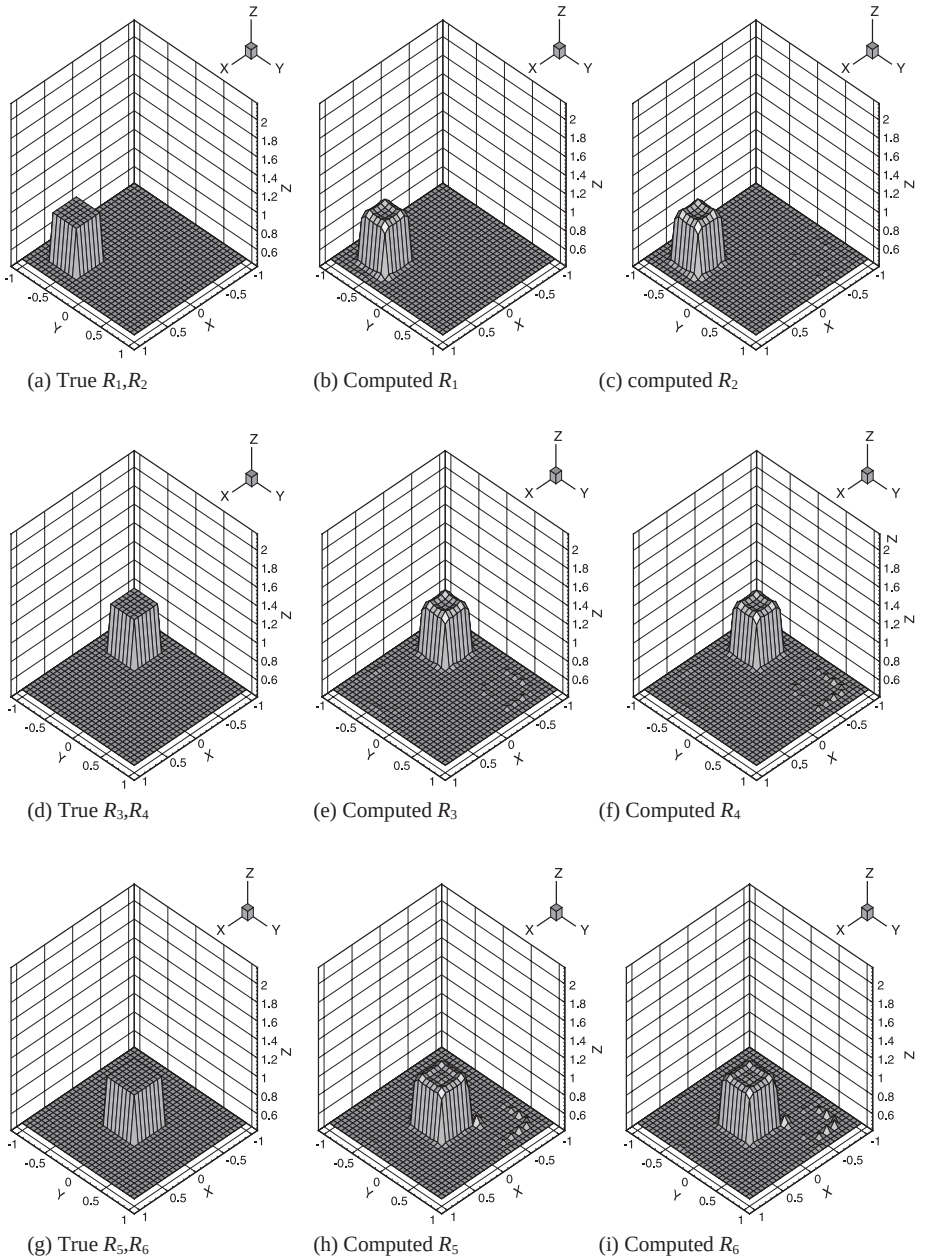
(d) Computed S

Fig. 1. Recovery of K and S given $w(x, y, t)$.

"edge" effects, we further assume that $R_1 = R_2$, $R_3 = R_4$, and $R_5 = R_6$. So, we seek to recover three different functions, R_1, R_3, and R_5.

As can be seen, the recovery of K is good, as the discontinuity is quite clear, and the height is accurate.

The true and recovered functions $R_i(x)$ are shown in Fig. 2. On our multiprocessor Beowulf system the task of computing each R_i was sent to an individual processing node. So, the massive computational task involved in the computing a large number of recharge parameters is readily scalable.



(a) True R_1,R_2 (b) Computed R_1 (c) computed R_2

(d) True R_3,R_4 (e) Computed R_3 (f) Computed R_4

(g) True R_5,R_6 (h) Computed R_5 (i) Computed R_6

Fig. 2. Recovery of $R(x, y, t)$ given $w(x, y, t)$.

The computed S is more of a problem. The main difficulty here appears to be that small errors in the computed K, and, to a lesser extent, R seem to have noticable effects on the recovery of S, because we have recovered this true S quite well when K and R are assumed known, and only S is being recovered, as may be seen in Fig. 3. On the other hand, it is worth noting that even though the

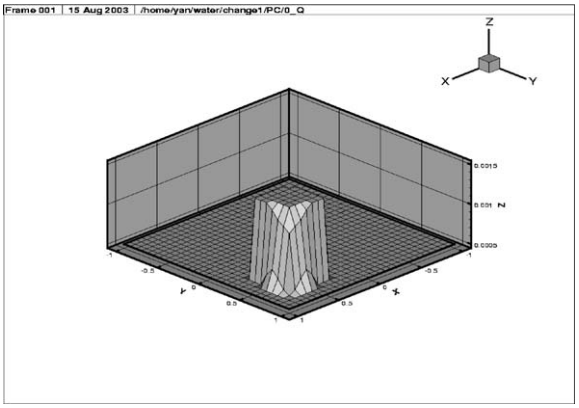


Fig. 3. Recovery of S given K.

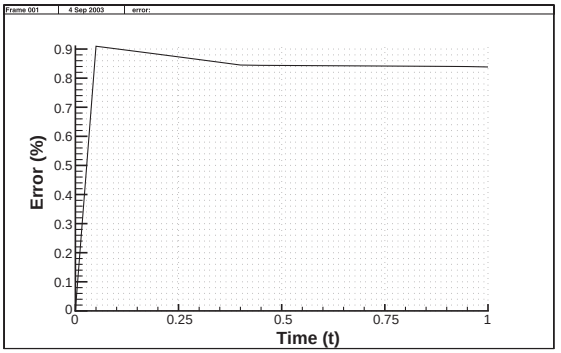


Fig. 4. Model error.

S recovery does not appear as effective as the others, the model seems to be relatively insensitive to this error, probably because the values of S are so small in the first place. The model error is shown below in Fig. 4. Here we have graphed the maximum relative error between the model values and the "true" head data, over the space grid points as a function of time. This shows that the model formed from the above K, S, and R well approximates the original problem.

The data for Figs. 1 and 2 was obtained by using the PDE package PDETWO [16] to solve the 2-d parabolic equation (1.1) over a square region

$$\{(x, y) : -1 \leqslant x \leqslant 1, -1 \leqslant y \leqslant 1\}$$

and with $0 \leqslant t \leqslant 1$; the chosen time step was $h = 10^{-7}$ and we used a 30×30 grid on the spatial domain. We used the initial condition

$$w(x, y, 0) = 2 + 0.5 \cos \pi x \cos \pi y$$

(to simulate slowly varying head data), and boundary conditions

$$w(x, -1, t) = 2 - (0.5 - t) \cos \pi x,$$

$$w(x, 1, t) = 2 - (0.5 - t) \cos \pi x,$$

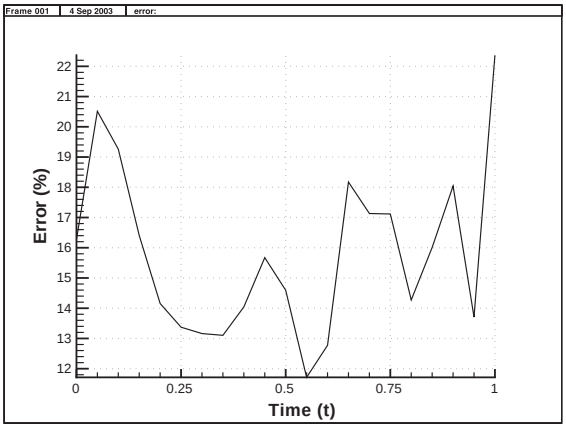


Fig. 5. Sparse data test.

$$w(-1, y, t) = 2 - (0.5 - t)\cos \pi y,$$

$$w(1, y, t) = 2 - (0.5 - t)\cos \pi y.$$

This parabolic data was then transformed to elliptic data $u(x, y, \lambda)$ via formula (2.1) and a Simpson's rule quadrature.

A valid criticism of the tests done thus far is that practical head data is both sparse and noisy, so that in particular one does not have head data at every point on a 30×30 grid as we assumed above. In the next test, we first computed $w(x, t)$ at discrete times on the same 30×30 grid for given K, S, and $R(x)$ as above, where we now consider these values as our "true" head data at each discrete time. Then we discarded 99% of the interior head values, keeping a regular grid containing nine interior values, and on the boundary we kept the corresponding boundary data points to complete the regular grid. To the surviving head values we added 20% relative error and then piecewise two-dimensional linearly interpolated these data values to obtain our synthetic "measured" head dataset. From this data the K, S, and R were recovered as above, and used to produce the model head values. The maximum relative error over the space grid points as a function of time, between these model values and the "true" head data is graphed in Fig. 5 above. As can be seen, the maximum error is comparable to the added noise. This and other similar trials indicate that the recovery process appears to be stable with respect to sparse well sites and head measurement error.

We also investigated the practical utility of the method by means of measurements of flow data gathered from seven wells located on the UAB campus over a period of about 8 months, and assuming a confined depth-averaged two-dimensional model for the aquifer. The results are shown in Fig. 6. The data was first interpolated piecewise linearly on the irregular triangular grid formed in a rectangle containing the measurement points, and then smoothed and differentiated by using the technique outlined in [13, Section 5]. Each side of the rectangle contained one measurement point, and at each of these boundary points we also measured the value of K (in units of feet/minute) using a standard bail test in which the head is measured in the pumping well [15, p. 58]; here the storativity S is dimensionless and R has units of feet/minute. We ran the computation first under the

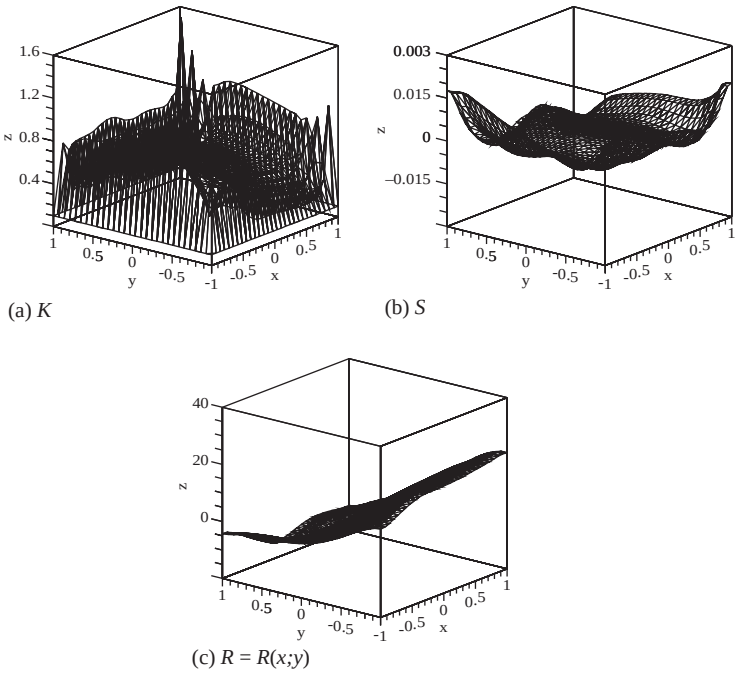


Fig. 6. UAB data.

assumption that $R = R(\boldsymbol{x})$, and later assumed that R had the form (4.1) with $n = 2, 3$. In each case the computed values for K and S were essentially the same, with the computed functions $R_i(\boldsymbol{x})$ showing some modest variation. It is found in practice that values for S typically lie in the range 0.00005–0.005. The computed values of S are approximately consistent (see, for example, [1, p. 41]) with the sand and clay mixture that is known to constitute much of the subsurface region under UAB [15, p. 56].

The irregular appearance of K on the boundary can be traced to the fact that we had only one measured value for K on each side of our rectangular region, because we were unable to carry out the transmissivity measurement at three of our seven wells. We have observed from our tests on synthetic data that in such cases, while the values of K near the boundary were in general not reliable, interior values of K were usually quite good (see for example [14, Fig. 3]). In all cases the minimizations were very stable, and "ran out of steam" after about 150 iterations, which indicates that the parameters shown in Fig. 6 are plausible candidates for the "best fit" for this dataset. As our data did not come with any solute information, we were not able to complete the task of forming a complete flow/solute model at this point in time. We hope to remedy this in a future study. We note in passing that one can use similar techniques to recover not only the full hydraulic conductivity tensor, but also solute parameters like the porosity and the hydrodynamic dispersion tensor as well various solute source terms.

In summary, these tests indicate that verifiably reliable recovery of multiple subsurface parameters, which of course includes the solution of the full inverse groundwater problem discussed earlier, may now be possible.

Finally, we take the opportunity here to thank the referees for their careful reading of the original manuscript; their suggestions led to substantial improvements in the final version of the paper.

References

[1] M.P. Anderson, W.W. Woessner, Applied Groundwater Modeling, Academic Press, New York, 1992.
[2] J. Bear, A. Verruijt, Modeling Groundwater Flow and Pollution, D. Reidel Publishing Co., Dordrecht, 1992.
[3] H. Ben Ameur, G. Chavent, J. Jaffré, Refinement and coarsening indicators for adaptive parametrization: application to the estimation of hydraulic transmissivities, Inverse Problems 18 (3) (2002) 775–794.
[4] J. Carrera, State of the art of the inverse problem applied to the flow and solute equations, in: E. Custodio (Ed.), Groundwater Flow and Quality Modelling, Kluwer Academic, Dordrecht, 1988, pp. 549–583.
[5] J. Carrera, S.P. Neuman, Estimation of aquifer parameters under transient and steady-state conditions: 2. Uniqueness, stability and solution algorithms, Water Resour. Res. 22 (1986) 211–227.
[6] G. Chavent, Identification of distributed parameter systems: about the output least square method, its implementation and identifiability, in: R. Iserman (Ed.), Proceedings of the Fifth IFAC Symposium on Identification and System Parameter Estimation, Vol. 1, Pergamon Press, Oxford, 1980, pp. 85–97.
[7] G. Chavent, On the theory of practice of nonlinear least-squares, Parameter identification in ground water flow, transport, and related processes, Part I, Adv. Water Res. 14 (2) (1991) 55–63.
[8] Y. Emsellem, G. de Marsily, An automatic solution for the inverse problem, Water Resour. Res. 7 (5) (1971) 1264–1283.
[9] T.R. Ginn, J.H. Cushman, M.H. Houch, A continuous-time inverse operator for groundwater and contaminant transport modeling: deterministic case, Water Resour. Res. 26 (1990) 241–252.
[10] D.L. Hughson, A. Gutjahr, Effect of conditioning randomly heterogeneous transmissivity on temporal hydraulic head measurements in transient two-dimensional aquifer flow, Stochastic Hydrol. Hydraul. 12 (1998) 155–170.
[11] A.G. Journel, J.C. Huijbregets, Mining Geostatistics, Academic Press, San Diego, CA, 1978.
[12] I. Knowles, Uniqueness for an elliptic inverse problem, SIAM J. Appl. Math. 59 (4) (1999) 1356–1370, Available online at http://www.math.uab.edu/knowles/pubs.html.
[13] I. Knowles, Parameter identification for elliptic problems, J. Comput. Appl. Math. 131 (2001) 175–194.
[14] I. Knowles, R. Wallace, A variational solution of the aquifer transmissivity problem, Inverse Problems 12 (1996) 953–963.
[15] T.A. Le, An inverse problem in groundwater modeling. Ph.D. Thesis, University of Alabama at Birmingham, 2000.
[16] D. Melgaard, R.F. Sincovec, General software for two-dimensional nonlinear partial differential equations, ACM Trans. Math. Software 7 (1) (1981) 106–125.
[17] J.W. Neuberger, Sobolev Gradients in Differential Equations, in: Lecture Notes in Mathematics, Vol. 1670, Springer, New York, 1997.
[18] S.P. Neumann, Perspective on “delayed yield”, Water Resour. Res. 15 (1979) 899–908.
[19] L.E. Payne, Improperly Posed Problems in Partial Differential Equations, SIAM, Philadelphia, 1975.
[20] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, Numerical Recipes: The Art of Scientific Computing, Cambridge University Press, Cambridge, 1989.
[21] J.R. Rice, R.F. Boisvert, Solving Elliptic Problems Using ELLPACK, Springer, Berlin, 1985.
[22] J. Simmers, Estimation of Natural Groundwater Recharge, in: NATO ASI Series C, Vol. 222, D. Reidel Publishing Co., Dordrecht, 1988.
[23] A.N. Tikhonov, V.Y. Arsenin, Solutions of Ill-Posed Problems, V.H. Winston & Sons, Washington, DC, 1977.
[24] G.R. Vázquez, M. Guidici, G. Parravicini, G. Ponzini, The differential system method for the identification of transmissivity and storativity, Trans. Porous Media 26 (1997) 339–371.
[25] W.W.-G. Yeh, Review of parameter identification procedures in groundwater hydrology: the inverse problem, Water Resour. Res. 22 (2) (1986) 95–108.