EXTENSIONS OF THE DIFFERENTIAL SYSTEM METHOD FOR PARAMETER IDENTIFICATION IN AQUIFERS

Miguel Angel Moreles, Rogelio Vázquez y Fernando Avila

Comunicación Técnica No I-02-04/10-04-2002
(MB/CIMAT)

Extensions of the Differential System Method for Parameter Identification in Aquifers

Miguel Angel Moreles
CIMAT
Email: moreles@cimat.mx

Rogelio Vázquez
CICESE
Email: rvazquez@cicese.mx

Fernando Avila
CIMAT
Email: avila@cimat.mx

Abstract

We present the applicability of differential system (DS) method for identification of hydraulic conductivity and effective porosity in a phreatic aquifer. In the classical setting, the first step of the DS system is to solve an overdetermined algebraic system using least squares. A natural extension of the method is to pose a least squares problem in an appropiate functional space, we consider the space of square integrable functions in the time variable for a finite interval. Also we present different methods of approximation and discuss how to deal with noisy data.

1 Introduction

Numerical simulators of groundwater flow have been developed to study the effects of changes in geohydrologic conditions and management policy on the groundwater storage, flow directions, and water quality in aquifers. There are many difficulties to apply such models to real field situations, among others, the estimation of the model parameters is one of the most important. These simulators require the hydraulic properties, such as conductivity and storativity all over the flow domain. In order to understand the groundwater flow phenomenon is very important to know the physical properties of the media at the scale in which we want to study the problem. If model parameters are not accurately estimated at the scale of interest, the model prediction may be useless for management of groundwater resource. The inverse problem solution in geohydrology is an important part of a more general issue that concerns aquifer model validation.

Hydrologist and petroleum engineers have used numerical models for the simulation of flow and solute transport for more than three decades, to study problems related to flow in porous media. Most if not all, practical solutions to

groundwater simulation problems are obtained via one or more of the forward operators approximating the physically based mass balance relation.

There are three types of requirements for the model to be close to the real system, at least with regards to the simulated variables (the output of the model):

1. The model should include all physical processes relevant to the simulation.
2. The structure of the model should resemble that of the real system.
3. The values assigned to the variables controlling the process must be similar to their real counterparts.

These concepts are defined in the literature as: process identification, model structure identification and parameter estimation, Carrera [2]. Parameter estimation or deterministic calibration is also defined in Ginn et. al. [4].Throughout this work we follow the latter.

For the confined aquifer case, Parravicini et. al.[6], proposed a direct method, based on the solution of a Cauchy problem that allows for the determination of both transmissivity T and storativity S, when the potentials and source terms are given for three different flow conditions, at least one of them transient. Vázquez et. al. [7] developed further the method and show the advantages of using more than three flow situations data. The method essentially consists in writing the differential equation for each one of the flow conditions. Forming the set of equations as a first order partial differential system in the unknown T and an algebraic system in the unknown S. Therefore the name DS Method (for Differential System Method).

In this work we are concerned primarily with an isotropic phreatic aquifer that satisfies Dupuit assumption and for which Darcy's law and the two-dimensional approximation hold, so the flow is regulated by the Boussinesq equation, Bear [1]:

$$\frac{\partial}{\partial x}\left(K(h-\eta)\frac{\partial h}{\partial x}\right)+\frac{\partial}{\partial y}\left(K(h-\eta)\frac{\partial h}{\partial y}\right)=n_e\frac{\partial h}{\partial t}-f \tag{1}$$

where $h(x,y,t)$ is the elevation of the aquifer free surface (piezometric head, or hydraulic potential), $\eta(x,y)$ is the elevation of the bottom of the aquifer,$f(x,y,t)$ is the source term, representing a vertical flux, positive if downward.

Our purpose is to develop the DS method for the identification of the model parameters, hydraulic conductivity, $K(x,y)$, and effective porosity, $n_e(x,y)$. We will show that for the inversion, no a priori knowledge of effective porosity is needed and, moreover, the identification of hydraulic conductivity does not depend upon it.

The first step in the method requires to solve an algebraic system in the unknowns $\partial K/\partial x$, $\partial K/\partial y$, n_e when the point (x,y) is fixed. The system is overdetermined, hence, the solution is found in the least squares sense. We

shall see that a natural extension is to solve this least squares problems for more general hilbertian norms, and we show some advantages.

The aquifer under study is generated synthetically, thus the data for the DS method is obtained by solving a direct problem. We discuss briefly the effect of different discretization methods, namely, balance cell models and the finite element method (FEM). Also, we show how to deal with noisy data by means of Tikhonov regularization. Here we consider both, confined and phreatic aquifers.

We remark that the full analysis of the DS method when the generation of synthetic data is by means of the Galerkin-FEM, and the study of noisy data, is carried out in Fregoso [3] for confined aquifers, and Kú [5] for the phreatic case.

2 The continuous inverse problem

In order to show how the DS method applies to a phreatic aquifer; let us describe the method with explicit reference to an isotropic phreatic aquifer described by equation (1). For simplicity assume that the aquifer bottom coincide whit the datum ($h(x,y)=0$) then:

$$\frac{\partial}{\partial x}\left(Kh\frac{\partial h}{\partial x}\right)+\frac{\partial}{\partial y}\left(Kh\frac{\partial h}{\partial y}\right)=n_e\frac{\partial h}{\partial t}-f \tag{2}$$

Consider equation (2) and suppose that the source term $f(x,y,t_i)$, the potential $h(x,y,t_i)$ and its time derivative $\partial h(x,y,t_i)/\partial t$ are known as a function of space, at p different times t_i, $i=1,2,\ldots,p$. A set of data is given by these functions, so p sets of data are assumed to be known.

Let us introduce the notation:

$$f^i \equiv f(x,y,t_i)$$

$$h^i \equiv h(x,y,t_i)$$

$$\partial h^i/\partial t \equiv h(x,y,t_i)/\partial t$$

then equation (2) gives:

$$\frac{\partial}{\partial x}\left(Kh^i\frac{\partial h^i}{\partial x}\right)+\frac{\partial}{\partial y}\left(Kh^i\frac{\partial h^i}{\partial y}\right)=n_e\frac{\partial h^i}{\partial t}-f^i$$

Applying derivatives we obtain after some simplification

$$h^i\frac{\partial h^i}{\partial x}\frac{\partial K}{\partial x}+h^i\frac{\partial h^i}{\partial y}\frac{\partial K}{\partial y}-\frac{\partial h^i}{\partial t}n_e=-\overline{\Delta}h^iK-f^i$$

where

$$\overline{\Delta}h^i = h^i \Delta h^i + \left(\frac{\partial h^i}{\partial x}\right)^2 + \left(\frac{\partial h^i}{\partial y}\right)^2$$

Let us define

$$\mathbf{u} = (u_1, u_2, u_3) = \left(\frac{\partial K}{\partial x}, \frac{\partial K}{\partial y}, n_e\right) \tag{3}$$

And the vectors $\mathbf{z}$, $\mathbf{f}$ with components $\overline{\Delta h^i}$ and f^i respectively.

Assuming that conductivity K is known, we have for a fixed point (x, y) the following linear system.

$$\mathbf{A}\mathbf{u} = -K\mathbf{z} + \mathbf{f} \tag{4}$$

If $Rank(\mathbf{A}) = 3$ then system (4) has a unique solution in the least squares sense given by

$$\mathbf{u} = -K\mathbf{a} + \mathbf{b} \tag{5}$$

Where the three component vector function $\mathbf{a}$ and $\mathbf{b}$ are the solution of the systems:

$$\begin{aligned} \mathbf{A}\mathbf{a} &= \mathbf{z} \\ \mathbf{A}\mathbf{b} &= \mathbf{f} \end{aligned} \tag{6}$$

Now, let us rewrite equation (5) recalling definition (3) and writing down the dependence upon x explicitly. We have the following system for the first two components of $\mathbf{u}$

$$\begin{aligned} u_1 &= \frac{\partial K}{\partial x} = -Ka_1 + b_1 \\ u_2 &= \frac{\partial K}{\partial y} = -Ka_2 + b_2 \end{aligned} \tag{7}$$

And the following equation for the third component of $\mathbf{u}$

$$u_3 = n_e = -Ka_3 + b_3 \tag{8}$$

The second step for identifying the parameters consists in considering the equations in (7) as a first order differential system for K.

To solve this differential system we need Cauchy data, that is, the assignment of conductivity at a point $\mathbf{x}^0 = (x_0, y_0)$ of the domain. The Cauchy problem to solve is the following

$$\begin{aligned} u_1 &= \frac{\partial K}{\partial x} = -Ka_1 + b_1 \\ u_2 &= \frac{\partial K}{\partial y} = -Ka_2 + b_2 \\ K(x_0, y_0) &= K_0 \end{aligned} \tag{9}$$

The solution of system (9) is unique, provided it exists.

The solution at a point $\mathbf{x} = (x, y)$ is found by choosing an appropriate path joining $\mathbf{x}$ with the initial point, and integrating (9) along it.

Indeed, let $\gamma(s)$, be a path joining $\mathbf{x}^0$ with $\mathbf{x}$.

$$\gamma : [0, 1] \to \mathbb{R}^2, \quad \gamma(0) = \mathbf{x}^0, \quad \gamma(1) = \mathbf{x}$$

Let $k(s) = K(\gamma(s))$, then

$$\begin{aligned} \frac{dk}{ds} &= \nabla K(\gamma(s)) \cdot \frac{d\gamma}{ds} \\ &= (-K(\gamma(s))\mathbf{a} + \mathbf{b}) \cdot \tfrac{d\gamma}{ds} \\ &= -\mathbf{a} \cdot \tfrac{d\gamma}{ds} K(\gamma(s)) + \mathbf{b} \cdot \tfrac{d\gamma}{ds} \end{aligned}$$

Here, $\mathbf{a} \equiv \mathbf{a}(\gamma(s))$, and $\mathbf{b} \equiv \mathbf{b}(\gamma(s))$. We are led to the initial value problem

$$\frac{dk}{ds} = -\left(\mathbf{a} \cdot \frac{d\gamma}{ds}\right) k(s) + \mathbf{b} \cdot \frac{d\gamma}{ds}, \qquad k(0) = K_0 \tag{10}$$

which is easily solved.

When conductivity has been evaluated, then equation (8) is used to obtain the effective porosity at every point of the domain.

3 A Synthetic Phreatic Aquifer

The DS method, as described in the previous section, is applied to a synthetic but realistic example

The Numerical Evaluation of the Model Parameters

Consider a regularly spaced lattice of nodes, each located at the center of a square cell, with sides parallel to the orthogonal Cartesian coordinate axes, and with spacing Δx along the x and y directions.

Nodes are labeled with the ordered pair of integer numbers (m, n), $m = 1, 2, ..., M$, $n = 1, 2, ...N$, so that the pair (m, n) represents the node $x(m, n) = m\Delta x\mathbf{i} + n\Delta x\mathbf{j}$, where the vectors $\mathbf{i}$ and $\mathbf{j}$ are the unit vectors.

Recalling the discrete conservative scheme with finite differences for the unconfined aquifer Bear [1] then the model parameters that need to be identified are easily pointed out.

Consider an interior cell $B(m, n)$ of the discrete domain of the aquifer. The integral balance equation for $B(m, n)$ is the following.

$$\begin{aligned} & n_e(m,n)\left(h^i(m,n) - h^{i-1}(m,n)\right)\frac{(\Delta x)^2}{\Delta t_i} = -N^i(m,n) + \\ & + (KD^*)\left((m,n),(m+1,n)\right)\left(h^i(m+1,n) - h^i(m,n)\right) + \\ & + (KD^*)\left((m,n),(m-1,n)\right)\left(h^i(m-1,n) - h^i(m,n)\right) + \\ & + (KD^*)\left((m,n),(m,n+1)\right)\left(h^i(m,n+1) - h^i(m,n)\right) + \\ & + (KD^*)\left((m,n),(m,n-1)\right)\left(h^i(m,n) - h^i(m,n-1)\right) \end{aligned} \tag{11}$$

where

$$(KD^*)\left((m,n),(m+1,n)\right) = \frac{2K(m,n)K(m+1,n)}{K(m,n)+K(m+1,n)}\frac{\left(h^{i-1}(m+1,n) - h^{i-1}(m,n)\right)}{2}$$

$$(KD^*)\left((m,n),(m-1,n)\right) = \frac{2K(m,n)K(m-1,n)}{K(m,n)+K(m-1,n)}\frac{\left(h^{i-1}(m-1,n) - h^{i-1}(m,n)\right)}{2}$$

$$(KD^*)\left((m,n),(m,n+1)\right) = \frac{2K(m,n)K(m,n+1)}{K(m,n)+K(m,n+1)}\frac{\left(h^{i-1}(m,n+1) - h^{i-1}(m,n)\right)}{2}$$

$$(KD^*)\left((m,n),(m,n-1)\right) = \frac{2K(m,n)K(m,n-1)}{K(m,n)+K(m,n-1)}\frac{\left(h^{i-1}(m,n-1) - h^{i-1}(m,n)\right)}{2}$$

which is a formulation of the multiple cells models Bear [1].

The choice of space and time intervals Δx and Δt_i is based on geometrical factors and frequency of measurements, which depend upon the goals of the forecasting model.

The discrete parameters that appear in equation (11) are the internode conductivity $K\left((m,n),(m\prime,n\prime)\right)$ and the cell effective porosity $n_e(m,n)$. These are the discrete model parameters relevant for the description of groundwater flow at the given, fixed, space and time scales.

Therefore, the goal of the discrete inverse problem for the case of unconfined aquifers is the determination of the internode conductivity and the effective porosity of the cell.

Generation of Synthetic Data

The synthetic phreatic aquifer that we consider is divided into a square regular lattice of cells, with $M = N = 9$, $\Delta x = 25m$. Conductivity values in m^2/s are constant on each cell. Its values are summarized in the following table

m \ n	1	2	3	4	...
1	0.000525	0.000510	0.000495	0.000480	
2	0.000475	0.000460	0.000445	0.000430	
3	0.000425	0.000410	0.000395	0.000380	
4	0.000375	0.000360	0.000345	0.000330	
5	0.000325	0.000310	0.000295	0.000280	
6	0.000275	0.000260	0.000245	0.000230	
7	0.000225	0.000210	0.000195	0.000180	
8	0.000175	0.000160	0.000145	0.000130	
9	0.000125	0.000110	0.000095	0.000080	

5	6	7	8	9
0.000465	0.000450	0.000435	0.000420	0.000405
0.000415	0.000400	0.000385	0.000370	0.000355
0.000365	0.000350	0.000335	0.000320	0.000305
0.000315	0.000300	0.000285	0.000270	0.000255
0.000265	0.000250	0.000235	0.000220	0.000205
0.000215	0.000200	0.000185	0.000170	0.000155
0.000165	0.000150	0.000135	0.000120	0.000105
0.000115	0.000100	0.000085	0.000070	0.000055
0.000065	0.000050	0.000035	0.000020	0.000005

...

Since the domain is divided into zones of constant conductivity and the regular spacing of the cells, the internode conductivities, denoted with $K((m,n),(m\prime,n\prime))$, are the harmonic mean of the cell values. Porosity is constant on each cell as well and its values are given in the following table

$_m{}^n$	1	2	3	4
1	0.033333	0.044444	0.050000	0.053333
2	0.044444	0.066667	0.080000	0.088889
3	0.050000	0.080000	0.100000	0.114286
4	0.053333	0.088889	0.114286	0.133333
5	0.055556	0.095238	0.125000	0.148148
6	0.057143	0.100000	0.133333	0.160000
7	0.058333	0.103704	0.140000	0.169697
8	0.059259	0.106667	0.145455	0.177778
9	0.060000	0.109091	0.150000	0.184615

...

...

5	6	7	8	9
0.055556	0.057143	0.058333	0.059259	0.060000
0.095238	0.100000	0.103704	0.106667	0.109091
0.125000	0.133333	0.140000	0.145455	0.150000
0.148148	0.160000	0.169697	0.177778	0.184615
0.166667	0.181818	0.194444	0.205128	0.214286
0.181818	0.200000	0.215385	0.228571	0.240000
0.194444	0.215385	0.233333	0.248889	0.262500
0.205128	0.228571	0.248889	0.266667	0.282353
0.214286	0.240000	0.262500	0.282353	0.300000

It will become apparent that the fact that the domain is divided into zones of constant conductivity and porosity, and the correlation between these two parameters, does not play a role in the identification process.

Dirichlet boundary conditions for the piezometric head have been assigned at the border of the domain. They do no vary with time. The values are as follows (in meters)

$_m{}^n$	1	2	3	4	5	6	7	8	9
1	40.00	40.05	40.10	40.15	40.20	40.25	40.30	40.35	40.40
2	39.80	x	x	x	x	x	x	x	40.20
3	39.60	x	x	x	x	x	x	x	40.00
4	39.40	x	x	x	x	x	x	x	39.85
5	39.20	x	x	x	x	x	x	x	39.70
6	39.00	x	x	x	x	x	x	x	39.55
7	39.80	x	x	x	x	x	x	x	39.40
8	38.60	x	x	x	x	x	x	x	39.30
9	38.40	38.50	38.60	38.70	38.80	38.90	39.00	39.10	39.20

The initial conditions for the piezometric head are given by the solution of a steady-state forward problem corresponding to the source term given in the following table (m^3/s)

$_m{}^n$	1	2	3	4
1	0.00003675	0.00003075	0.00002675	0.00002475
2	0.00002475	0.00001875	0.00001475	0.00001275
3	0.00001675	0.00001075	0.00000675	0.00000475
4	0.00001275	0.00000675	0.00000275	0.00000075
5	0.00001275	0.00000675	0.00000275	0.00000075
6	0.00001675	0.00001075	0.00000675	0.00000475
7	0.00002475	0.00001875	0.00001475	0.00001275
8	0.00003675	0.00003075	0.00002675	0.00002475
9	0.00005275	0.00004675	0.00004275	0.00004075

. . .

. . .

5	6	7	8	9
0.00002475	0.00002675	0.00003075	0.00003675	0.00004475
0.00001275	0.00001475	0.00001875	0.00002475	0.00003275
0.00000475	0.00000675	0.00001075	0.00001675	0.00002475
0.00000075	0.00000275	0.00000675	0.00001275	0.00002075
0.00000075	0.00000275	0.00000675	0.00001275	0.00002075
0.00000475	0.00000675	0.00001075	0.00001675	0.00002475
0.00001275	0.00001475	0.00001875	0.00002475	0.00003275
0.00002475	0.00002675	0.00003075	0.00003675	0.00004475
0.00004075	0.00004275	0.00004675	0.00005275	0.00006075

Which represents leakage in the aquifer and is present in all the situations devised.

A numerical comparison of the predicted parameters, conductivity and porosity, will be made with the reference parameters. More important for the management of the aquifer, is to verify that with the predicted parameters it is possible to recover the evolution of the aquifer. A numerical comparison is also possible, we content ourselves with a graphical comparison. We plot the contour flows of the piezometric head for synthetic aquifer as well as the contour flows of the piezometric head for the aquifer with the identified parameters. The contour flows of the initial conditions for the piezometric head are shown in Figure 1.

The transient regime is set up by a sudden start of some array of wells at $t = 0$. The array of wells are placed in the cells corresponding to the nodes: $(2, 2)$, $(3, 8)$, $(4, 8)$, $(6, 8)$, $(7, 8)$. At these nodes the source term during the transient regime is equal to 0.01 m^3/seg

Piezometric heads at different times are obtained by solving the forward problem with the reference parameters, the boundary and initial conditions, and the source term described above. The piezometric head is computed at the times in years

$$t_1 = 0.001, \quad t_2 = 0.0022, \quad t_3 = 0.0046, \quad t_4 = 0.0100.$$

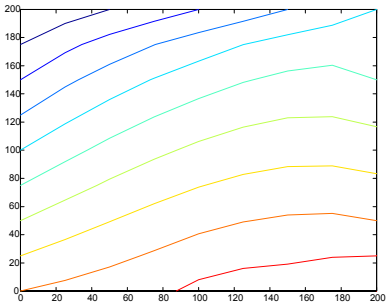

Figure 1:

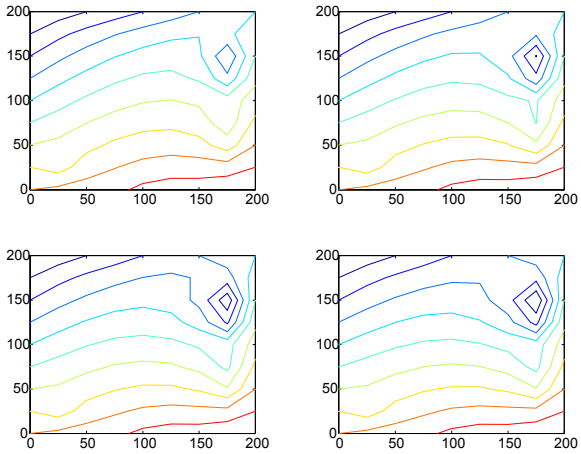

Figure 2:

The contour maps of the four transient flow situations are represented in Figure 2

4 Numerical Implementation of the DS Method

The sets of data used in the identification are given by $h(m,n,i)$, $(\partial h/\partial t)(m,n,i)$, $F(m,n,i)$, with $i = 0,1,2,3,4$. The time derivatives of the piezometric heads for the sets of data from 1 to 4 are evaluated with the backward differences.

There are two numerical problems to consider when implementing the DS method. First, the solution of the linear systems (6) , and second, the integration of equation (10) .

The systems (6) are overdetermined and ill-conditioned, the preferred algorithm is QR decomposition with householder transformations.

To integrate equation (10) we use polygonal paths through the nodes of the lattice. Moreover, **a**, **b** are approximated by constants along internode segments and denoted a_A, b_A. These quantities are scalars since between internode segments the path of integration is parallel to a coordinate axis.

It is possible to establish an a priori estimate for the difference between the true conductivity, $k(s)$, and the identified conductivity, $k_A(s)$. Following the proof of Gonwall's lemma we obtain

$$\begin{aligned} |k(s) - k_A(s)| \leq \quad & \exp\left(-\int_0^s a_A(r)dr\right) \cdot \\ & \cdot \left(|k(0) - k_A(0)| + \right. \\ & \left. \int_0^s \exp\left(\int_0^r a_A(p)dp\right) |k(r)| \, |a(r) - a_A(r)| + |b(r) - b_A(r)| \, dr\right) \end{aligned}$$

This suggests that among the possible integration paths that join the node (m, n), where we want to compute conductivity, to the initial node (m_0, n_0) we choose that one for which the sum $\sum_{(m_0,n_0)}^{(m,n)} |a_A|$, performed over all internode segments connecting the vertices of the path, is the smallest.

In the next table we show the identified conductivity in the inner nodes

$_m{}^n$	2	3	4
2	0.0004728	0.0003964	0.0004300
3	0.0003703	0.0003503	0.0003377
4	0.0003191	0.0003041	0.0002921
5	0.0002736	0.0002584	0.0002461
6	0.0002281	0.0002126	0.0001998
7	0.0001826	0.0001663	0.0001529
8	0.0001366	0.0001194	0.0001049

. . .

. . .

5	6	7	8
0.0003692	0.0003552	0.0003410	0.0003315
0.0003242	0.0003103	0.0002968	0.0002839
0.0002788	0.0002653	0.0002519	0.0002204
0.0002330	0.0002198	0.0002065	0.0001591
0.0001867	0.0001739	0.0001609	0.0001074
0.0001399	0.0001276	0.0001151	0.0000080
0.0000916	0.0000788	-0.0000587	-0.0000483

At first sight the prediction is correct except near the nodes where wehave wells. This is to be expected since near a well the flow is no longer laminar or aquifer-like.

Next we show a table with the absolute errors.

m \ n	2	3	4
2	0.0000128	0.0000486	0
3	0.0000397	0.0000447	0.0000423
4	0.0000409	0.0000409	0.0000379
5	0.0000364	0.0000366	0.0000339
6	0.0000319	0.0000324	0.0000302
7	0.0000274	0.0000287	0.0000271
8	0.0000234	0.0000256	0.0000251

. . .

5	6	7	8
0.0000458	0.0000448	0.0000440	0.0000385
0.0000408	0.0000397	0.0000382	0.0000361
0.0000362	0.0000347	0.0000331	0.0000496
0.0000320	0.0000302	0.0000285	0.0000609
0.0000283	0.0000261	0.0000241	0.0000626
0.0000251	0.0000224	0.0000199	0.0001120
0.0000234	0.0000212	0.0001437	0.0001183

From this table is not clear if the predictions are correct. The next table presents relative errors, which shows that the errors away from the wells are acceptable.

m \ n	2	3	4
2	0.0278	0.1091	0
3	0.0968	0.1132	0.1114
4	0.1136	0.1186	0.1148
5	0.1176	0.1239	0.1209
6	0.1226	0.1324	0.1315
7	0.1307	0.1470	0.1506
8	0.1460	0.1768	0.1932

. . .

5	6	7	8
0.1103	0.1120	0.1142	0.1041
0.1119	0.1133	0.1142	0.1128
0.1150	0.1157	0.1161	0.1839
0.1209	0.1210	0.1212	0.2766
0.1317	0.1306	0.1301	0.3684
0.1520	0.1494	0.1473	0.9333
0.2032	0.2118	1.6911	1.6893

Now we present the corresponding tables for effective porosity. The predictions are

m \ n	2	3	4
2	0.0494	0.0716	0.0867
3	0.0793	0.0918	0.0999
4	0.0805	0.1001	0.1145
5	0.0829	0.1057	0.1238
6	0.0831	0.1083	0.1290
7	0.0829	0.1090	0.1308
8	0.0827	0.1083	0.1288

. . .

5	6	7	8
0.0830	0.0873	0.0910	0.0978
0.1082	0.1160	0.1231	0.1909
0.1266	0.1379	0.1483	0.2580
0.1395	0.1542	0.1667	0.1520
0.1474	0.1656	0.1833	0.3278
0.1508	0.1722	0.1949	0.5521
0.1479	0.1684	-0.1641	-0.2139

Again we see that there are meaningless predictions near the wells. The absolute errors are

m \ n	2	3	4
2	0.0000173	0.0000084	0.0000022
3	0.0000007	0.0000082	0.0000144
4	0.0000083	0.0000142	0.0000189
5	0.0000123	0.0000193	0.0000243
6	0.0000169	0.0000251	0.0000310
7	0.0000209	0.0000310	0.0000389
8	0.0000239	0.0000372	0.0000490

. . .

5	6	7	8
0.0000122	0.0000127	0.0000127	0.0000089
0.0000168	0.0000173	0.0000169	0.0000454
0.0000216	0.0000221	0.0000214	0.0000802
0.0000272	0.0000276	0.0000277	0.0000532
0.0000344	0.0000344	0.0000320	0.0000992
0.0000437	0.0000432	0.0000385	0.0003032
0.0000573	0.0000602	0.0004130	0.0004806

and the relative errors

m \ n	2	3	4
2	0.2590	0.1052	0.0250
3	0.0082	0.0819	0.1261
4	0.0939	0.1239	0.1415
5	0.1291	0.1545	0.1641
6	0.1685	0.1881	0.1935
7	0.2011	0.2217	0.2293
8	0.2244	0.2556	0.2754

. . .

. . .

5	6	7	8
0.1282	0.1268	0.1224	0.0831
0.1344	0.1300	0.1206	0.3122
0.1457	0.1383	0.1261	0.4510
0.1630	0.1519	0.1425	0.2592
0.1891	0.1718	0.1488	0.4341
0.2246	0.2007	0.1649	1.2183
0.2791	0.2633	1.6595	1.8021

The main motivation of parameter identification for aquifers is the ability to predict the evolution of the aquifer. We do so graphically, we plot the contour lines of the different flow situations (solid line) for the piezometric head, and the contour lines using the identified parameters of the corresponding piezometric head (dashed line). See the initial condition in Figure 3. The predictions for the transient flow situations are in Figure 4. We observe that the prediction of aquifer evolution is satisfactory. This in spite of the poor identification of parameters near the wells.

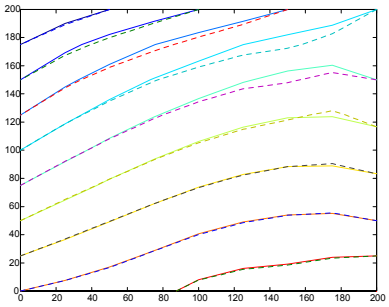

Figure 3:

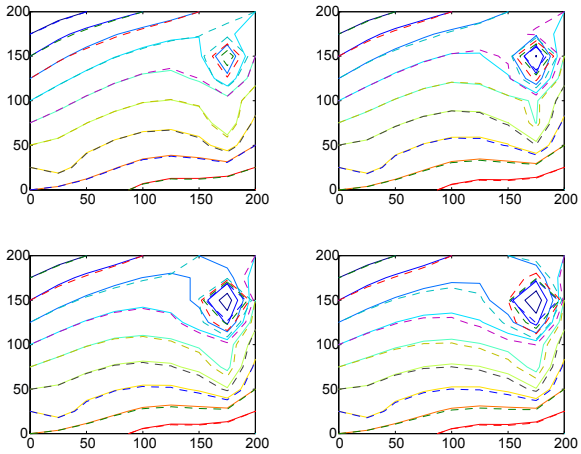

Figure 4:

5 An alternative approach to the DS Method

The continuous problem

Let us proceed as in Section 2 and consider the equation (2) in the form

$$h\frac{\partial h}{\partial x}\frac{\partial K}{\partial x} + h\frac{\partial h}{\partial y}\frac{\partial K}{\partial y} - \frac{\partial h}{\partial t}n_e = -\overline{\Delta}hK - f$$

where

$$\overline{\Delta}h = h\Delta h + \left(\frac{\partial h}{\partial x}\right)^2 + \left(\frac{\partial h}{\partial y}\right)^2$$

As before define

$$\mathbf{u} = (u_1, u_2, u_3) = \left(\frac{\partial K}{\partial x}, \frac{\partial K}{\partial y}, n_e \right)$$

Let us fix (x, y) again. In this case we assume that h, ∇h, $\overline{\Delta} h$ and f are known in (x, y, t) for $t \in I = [0, T]$, and belong to $L^2(0, T)$.

Recall that

$$L^2(0,T) = \left\{ \varphi : \int_0^T |\varphi(t)|^2 \, dt < +\infty \right\}$$

with inner product

$$\langle \varphi, \psi \rangle = \int_0^T \varphi(t)\psi(t) dt$$

and induced norm

$$\|\varphi\|^2 = \langle \varphi, \varphi \rangle = \int_0^T |\varphi(t)|^2 \, dt$$

Consider the map

$$\mathcal{R} : \mathbb{R}^3 \to L^2(0,T)$$

given by

$$\mathcal{R}(u_1, u_2, u_3) = \frac{1}{2} \left\| h \frac{\partial h}{\partial x} u_1 + h \frac{\partial h}{\partial y} u_2 - \frac{\partial h}{\partial t} u_3 + \overline{\Delta h} K + f \right\|^2$$

Assuming that K is known, the minimum satisfies $\nabla \mathcal{R}(u_1, u_2, u_3) = 0$, which leads us to a system

$$\mathbf{A}\mathbf{u} = -K\mathbf{z} + \mathbf{f} \tag{12}$$

If $Rank(\mathbf{A}) = 3$ this system has a unique solution given by

$$\mathbf{u} = -K\mathbf{a} + \mathbf{b}$$

Where the three component vector function $\mathbf{a}$ and $\mathbf{b}$ are the solutions of the systems

$$\mathbf{Aa} = \mathbf{z}$$

$$\mathbf{Ab} = \mathbf{f}$$

Once we solve these algebraic systems, we mimic the second step in the DS method as in Section 2.

Numerical Implementation

Throughout the development of the DS method with L^2-norm, we assumed that the functions involved were known for all times. In practice this is not the case, instead we have observation for a finite number of time situations, that is, we know, $h(m,n,i)$, $(\partial h/\partial t)(m,n,i)$, $F(m,n,i)$, with $i = 0, 1, \ldots, p$. The time derivatives of the piezometric heads for the sets of data from 1 to p are evaluated with the backward differences.

First we have to solve the linear system (12). When using the L^2-norm, the system reads

$$\begin{bmatrix} \langle hh_x, hh_x\rangle & \langle hh_x, hh_y\rangle & \langle hh_x, -h_t\rangle \\ \langle hh_x, hh_y\rangle & \langle hh_y, hh_y\rangle & \langle hh_y, -h_t\rangle \\ \langle hh_x, -h_t\rangle & \langle hh_y, -h_t\rangle & \langle h_t, h_t\rangle \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = -K \begin{bmatrix} \langle hh_x, \overline{\Delta}h\rangle \\ \langle hh_y, \overline{\Delta}h\rangle \\ \langle -h_t, \overline{\Delta}h\rangle \end{bmatrix} + \begin{bmatrix} \langle hh_x, f\rangle \\ \langle hh_y, f\rangle \\ \langle -h_t, f\rangle \end{bmatrix} \tag{13}$$

The components of the system (13) are integrals in t, this are approximated numerically by interpolating with linear functions at time t_1, t_2, t_3 and t_4. The system is again ill-conditioned and solved with QR decomposition as before.

There is no change when integrating the corresponding Cauchy problem.

Parameter Identification

Next we present the tables of the identified parameters, as well as the tables of errors. First the table for conductivity

m \ n	2	3	4
2	0.00029191	0.000445	0.00038077
3	0.00036835	0.00035074	0.00034553
4	0.00031122	0.0003095	0.00031062
5	0.00027155	0.00026996	0.00026711
6	0.00023007	0.00022536	0.0002185
7	0.00018559	0.000178624	0.00016871
8	0.00014005	0.000130535	0.00011794

. . .

5	6	7	8
0.00038564	0.00040672	0.00041533	0.00040893
0.00034363	0.00035522	0.00035748	0.0004394
0.00031738	0.00032548	0.0003296	8.1903
0.0000002	0.00026904	0.00028467	0.00030661
0.000212	0.00021001	0.0002134	0.00029135
0.00015859	0.00015049	0.00014255	0.00011805
0.00010412	8.8846e-005	5.86e-005	5.0739e-005

. . .

Observe that in the node $(3, 8)$, where there is a well located, we have an unrealistic identification.

Let us look at the corresponding table for porosity.

m \ n	2	3	4
2	0.045446	0.073957	0.081039
3	0.075467	0.094369	-0.003454
4	0.087773	0.04075	0.081392
5	0.022469	0.070784	0.10514
6	0.059599	0.088017	0.11483
7	0.067653	0.094547	0.11941
8	0.070582	0.098384	0.1219

. . .

5	6	7	8
0.037925	0.084428	0.107	0.10359
0.089526	0.12655	0.14796	0.18364
0.12082	0.16172	0.19782	17191
0.13502	0.17512	0.23987	0.3004
0.14272	0.18265	0.24638	0.31801
0.14458	0.17762	0.22642	0.25531
0.14227	0.15949	0.12279	0.228785

. . .

In agreement with the previous table, we have a meaningless prediction at the node $(3, 8)$. There is also an incorrect prediction at node $(3, 4)$.

As before, we compare the tables for errors. The absolute errors are

m \ n	2	3	4
2	0.00016809	0	4.9232e-005
3	4.1653e-005	4.4259e-005	3.4471e-005
4	4.8782e-005	3.5499e-005	1.9381e-005
5	3.8446e-005	2.5041e-005	1.289e-005
6	2.9928e-005	1.9639e-005	1.1504e-005
7	2.4409e-005	1.6382e-005	1.1293e-005
8	1.9954e-005	1.4466e-005	1.2061e-005

. . .

5	6	7	8
2.9358e-005	6.7217e-006	3.0334e-005	3.8934e-005
2.1368e-005	5.2161e-006	2.2481e-005	0.0001194
2.3787e-006	2.5477e-005	4.4597e-005	8.19
2.0168e-007	1.9039e-005	4.9666e-005	8.6613e-005
2.9981e-006	1.0012e-005	2.8401e-005	0.00012135
6.4055e-006	4.8503e-007	7.5455e-006	1.95e-006
1.0877e-005	1.1154e-005	2.64e-005	1.9261e-005

Better conclusions can be drawn by looking at the table of relative errors.

m \ n	2	3	4
2	0.36542	0	0.11449
3	0.10159	0.11205	0.090714
4	0.13551	0.1029	0.058731
5	0.12402	0.084886	0.046036
6	0.11511	0.080161	0.050016
7	0.11623	0.08401	0.06274
8	0.12471	0.099768	0.09278

. . .

5	6	7	8
0.070743	0.016804	0.078789	0.10523
0.058542	0.014903	0.067106	0.37311
0.0075515	0.084922	0.15648	30333
0.00076107	0.076156	0.21135	0.3937
0.013944	0.050058	0.15352	0.71382
0.038821	0.0032335	0.055892	0.01625
0.094583	0.11154	0.31059	0.27516

. . .

We conclude that identification with the L^2-norm is better than before. The same holds true for porosity. Here is the table of absolute errors.

m \ n	2	3	4
2	0.021221	0.006043	0.00785
3	0.0045332	0.0056309	0.11774
4	0.001116	0.073536	0.051941
5	0.072769	0.054216	0.043007
6	0.040401	0.045316	0.045168
7	0.036051	0.045453	0.050283
8	0.036085	0.047071	0.055883

. ..

5	6	7	8
0.057313	0.015572	0.0032961	0.0030723
0.035474	0.0067844	0.0079573	0.038182
0.027332	0.0017243	0.028118	17191
0.031643	0.0067009	0.045423	0.09527
0.039098	0.017353	0.030995	0.089443
0.049859	0.037762	0.0069176	0.006423
0.062861	0.069081	0.1261	0.037886

. . .

and the table with absolute errors.

m \ n	2	3	4
2	0.31832	0.075538	0.088312
3	0.056664	0.056309	1.0302
4	0.012555	0.64344	0.38956
5	0.76407	0.43373	0.2903
6	0.40401	0.33987	0.2823
7	0.34764	0.32466	0.29631
8	0.3383	0.32362	0.31434

. ..

5	6	7	8
0.60179	0.15572	0.031783	0.028803
0.28379	0.050883	0.056838	0.2625
0.18449	0.010777	0.1657	96699
0.18986	0.036855	0.23361	0.46444
0.21504	0.086767	0.1439	0.39131
0.25642	0.17532	0.029647	0.025806
0.30645	0.30223	0.50666	0.14207

. . .

We see that the predictions are correct except at the nodes $(3, 4)$ and $(3, 8)$. This is remarkable since we have good predictions for 46 out of 48 nodes.

Finally, figures 5 and 6, show that the aquifer evolution is predicted correctly.

We remark that the solution of the direct problem, was carried out without modifying the identified parameters, that is,without disregarding the meaningless predictions. To simulate a physically correct model, it is necessary to replace those incorrect values using, for instance, some technique of interpolation.

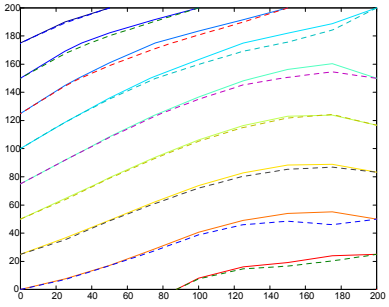

Figure 5:

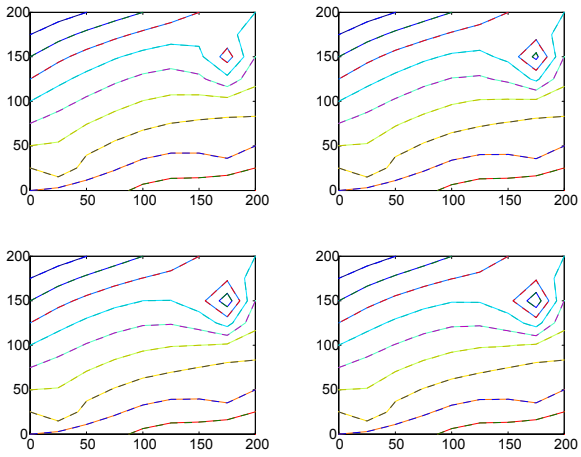

Figure 6:

6 Finite Elements and Noisy Data

The direct problem in aquifer simulation involves the numerical solution of partial differential equations (PDE) of parabolic type. The solution presented above is by the Finite Difference method. Another approach is the Galerkin approximation followed by the Finite Element Method (FEM), Galerkin-FEM for short. It is well known that finite differences is not well suited to work with irregular geometries, hence, in this case the Galerkin-FEM is preferred. As a consequence, it is natural to develop the DS method when the synthetic data is generated with Galerkin-FEM.

In real cases, besides the fact that the aquifer's domain is irregular, the data collected is not noisy free in general. In order to apply the DS method to a case as realistic as possible, we may add noise to the data generated synthetically. For instance, in Vazquez et al [7] the DS method is applied successfully to a confined aquifer, then correlated and uncorrelated noise is added to the piezometric head before inversion. It is shown that the DS method is not as effective. We obtain the same conclusion for the phreatic aquifer under study.

Our ultimate goal is to apply this theory to real aquifers, an step in that direction is to implement the DS method with data generated synthetically by the Galerkin-FEM approach, and then add noise to the piezometric head. The same conclusions hold, the DS method is only effective when the data is noise free. Since the model of aquifer evolution is a PDE, some smoothness is assumed in the functions involved. Although somewhat artificial, noisy data can be regularized by Tikhonov's method. The identification improves greatly. A full analysis of this approach for the confined aquifer in Vazquez et al [7] is carried out in Fregoso [3]. The analogue for the phreatic aquifer presented in this work is developed in Kú [5].

7 Comments and conclusions

The DS method, like any other inverse method in geophysics, needs high quality data. Thanks to recent improvements in instrumentation, it is nowadays possible. However, data have to be carefully analyzed before its use for inverse methods. Also, it is fundamental to collect independent sets of data.

The DS method can be applied to small subregions of an aquifer. In particular we can assess where the independence condition is verified by analyzing the collected data. The DS method identifies the parameters directly at the scale determined by the spacing of the observation points where data are collected.

Some other features of the method are:
(i) uses those measurable quantities that are usually collected in aquifers under control or that can be interpolated from these,
(ii) no prior information on effective porosity is required,
(iii) no initial guess of the unknown parameters is required,
(iv) no forward problem solutions are needed,

(v) the identification of hydraulic conductivity does not depend upon effective porosity, even when only transient data are used,
(vi) the internode hydraulic conductivity and cell effective porosity are ready-to-use parameters for the numerical implementation of the forward problem with conservative schemes,
(vii) takes into account several flows with different directions, all over the aquifer, using data on the whole flow field.

Acknowledgments

This work was initially supported by project CIMAT-CICESE 2000. The extensions developed in Section 5, as well as the master theses described in Section 6, were granted support by project CIMAT-CICESE 2001.

References

[1] J. Bear: *Hydraulics of Groundwater*; Mc Graw Hill; New Yor. (1979)

[2] J. Carrera; *State of the art of the inverse problem applied to the flow and solute transport equations, Groundwater Flow and Quality Modeling*; E. Custodio, A. Gurgui and J. P. Lobo-Ferreira (Eds.), D. Reidel, Publishing Co., Dordrecht, The Netherlands, 549-583. (1988)

[3] E. Fregoso: *Identificación de transmisividad y coeficiente de almacenamiento en un acuífero confinado y la aproximación de Galerkin*; Master Thesis; CIMAT. (2001)

[4] T. R. Ginn, J. H. Cushman & M. H. Houck; *A continuous-time inverse operator for groundwater and contaminant transport modeling: Deterministic case*, Water Resources Research, Vol. 26 No. 2, 241-2652. (1990)

[5] R. Kú: *Una formulación en elemento finito del sistema diferencial para la identificación de parámetros en un acuífero no confinado*; Master Thesis; CIMAT. (2001)

[6] G. Parravicini, M. Giudici, G. Morossi & G. Ponzini; *Minimal a priori asignment in a direct method for determining phenomenological coefficients uniquely*; Inverse Problems, 11, 611-629. (1995)

[7] R. Vázquez, M. Giudici, G. Ponzini, G. Parravicini; *The differential system method for the identification of transmissitivity and storativity*; Transport in Porous Media, 26, 339-371. (1997)