Groundwater Quality Management (Proceedings of the GQM 93 Conference held at Tallinn, September 1993). IAHS Publ. no. 220, 1994.

Automatic calibration of groundwater models by the head gradient method

DOMINIQUE THIERY

Bureau de Recherches Géologiques et Minières (BRGM), Dept. DR/HYT, BP 6009, 45060 Orléans Cedex 2, France

Abstract A groundwater model for flow and transport is reliable only if it is properly calibrated on observations. The best way to ensure a high predictive capability is to limit the number of parameters by defining homogeneous zones. These zones may be defined by an inverse model based on the gradient method. This paper describes a new scheme of inverse model for irregular finite difference 3D models in steady-state. The method determines, in each grid cell, the modification of hydraulic conductivity necessary to adjust the computed head gradients to the observed ones and yields the hydraulic conductivity in the centre of each grid cell. The stability of the results is forced through an optimization scheme simultaneously minimizing the differences of hydraulic conductivity between adjacent grid cells. After determination of the hydraulic conductivity, a procedure derives automatically optimal clusters of values defining geographical zones where homogeneous values of hydraulic conductivity may be imposed. The inverse modelling scheme is operational and applications to four real aquifer systems are presented.

INTRODUCTION

The calibration of a groundwater model consists in determining the hydrodynamic parameters (transmissivity, storage coefficient) and also recharge. Most commonly in real applications, the parameters are estimated by using a "trial-and-error" approach. This is very time consuming and depends to a great extent on the skill of the modeller. In theory, the number of parameters to be determined is not limited. For instance, in a finite difference model, considering only the calibration of the transmissivity, there are as many parameters to determine as grid cells in the model, i.e., one to several thousands of parameters.

The possibility of automatic parameter identification in distributed parameter systems, such as groundwater systems has been studied extensively during the last three decades. This problem, known as inverse modelling, is reviewed in detail by Yeh (1986), Carrera (1988) and Keidser & Rosbjerg (1991). Due to the structure of the equations, the solutions of the inverse problem are strongly affected by instability and also non-uniqueness; this is an "ill posed problem" which is why inverse modelling has up to now very seldom been used in real applications. Ensellem & De Marsily (1971) were probably the first to demonstrate that the solution must be regularized and that there is a trade-off between the accuracy and regularity. There are two main methods to achieve stability. The first uses prior information on the parameters (e.g., the transmissivity). This is the approach used by Carrera & Neuman (1986) but prior information is generally very limited in real situations and is affected by uncertainties. The other approach is to reduce drastically the number of parameters or to impose smoothness on the variation of the parameters.

Smoothness may be achieved through a flatness criterion which favours solutions where transmissivity displays no spatial oscillations, as used by Ensellem & De Marsily (1971) or by assuming a functional structure of the transmissivity distribution as used by Carrera & Neuman (1986). The reduction of the number of parameters is done by zonation, i.e. the aquifer system is divided into several zones with each zone characterized by a constant parameter. For any one type of parameter (e.g., the transmissivity) the number of unknown parameters is reduced to the number of zones. This is the approach of Khan (1986), Kessler (1987), Liu & Li (1991), Doherty (1991) and Thiéry (1993). The determination of the extension of geographical zones with homogeneous hydraulic conductivity was first approached by Ensellem & De Marsily (1971) through successive halving of the aquifer domain. The purpose of this paper is to present a new method of determining geographical zones which is operational with actual field data.

GENERAL DESCRIPTION OF THE METHOD

Among the methods which are very close to what is done by the modeller using trial-and-error are the flow net analysis described by Scott (1992) and the gradient method used by Döll (1990). Both methods are based on the principle that the flow is constant (except for recharge) in a streamtube. The method of Döll (1990) adjusts directly, in each grid cell, the value of the transmissivity in order to equal the computed head gradient to the measured one.

Our method is based on that of Döll (1990) which we have adapted in order to obtain a stable and physical spatial distribution of hydraulic conductivities. Homogeneous zones are then derived from the calculated hydraulic conductivities in each grid cell.

The domain of application of this method, which is described below, is subject to the following constraints:

(a) the aquifer must be in steady-state and only the hydraulic conductivity may be computed (assuming that the recharge is known);
(b) a smooth map of the hydraulic head must be available over the modelled domain or at least over a large part of it.

THE HEAD GRADIENT METHOD AS USED BY DÖLL (1990)

The principle of the method is very physical and very simple: consider a streamtube and neglect the recharge on a small area. The actual hydraulic head gradient $\mathbf{G_a}$ in the streamtube is determined from the observed hydraulic head maps and corresponds to an actual transmissivity $\mathbf{T_a}$; assuming initial values of transmissivities $\mathbf{T_{old}}$, the groundwater model calculates hydraulic heads, from which simulated gradients $\mathbf{G_s}$ are derived. In the streamtube of width $\mathbf{l}$, the flow Q is constant:

$$Q = T_a G_a l = T_{old} G_s l \qquad (1)$$

Assuming that the initial transmissivity was not too far from the actual one, and that the direction of the streamtube was acceptable, the following relation is immediately derived:

$$T_{new} = T_a = T_{old}\, G_s / G_a \qquad (2)$$

In a 2D or 3D model with rectangular grid cells, the same relation is applied in every direction, in order to adjust the internode transmissivity in the x, y and z directions and the process is iterative. Convergence is fast (within 5 to 10 simulations with the model) and the computed hydraulic heads are very similar to the actual ones. However, what is calculated is not the transmissivity (or the hydraulic conductivity) in the centre of each grid cell but the internode transmissivity. If **N** is the number of cells of the 3D model, there are 3 **N** such transmissivities but only **N** observations of hydraulic head. This has severe limitations:

(a) due to the large number of computed values, it is our experience that the internode transmissivities display a high spatial variability and their use for model prediction is questionable;

(b) internode transmissivity cannot be visualized on a single map but on three maps which correspond to a different anisotropy factor in each grid cell.

Our contribution was to adapt this very appealing method to get stable and smooth hydraulic conductivity in the centre of each grid cell.

METHODOLOGY OF PRESENT CONTRIBUTION

Notation

Q = flow ($L^3 T^{-1}$);
G = head gradient ($L L^{-1}$);
T = transmissivity ($L^2 T^{-1}$);
K = hydraulic conductivity ($L T^{-1}$);
R = resistance : 1/transmissivity ($T L^{-2}$);
b = width of the grid cell (L);
L = length of the grid cell (L);
h = wetted thickness (L);
i = one of the six directions (North, South, East, West, Up, Down);
w = regulation weight;
S = sum of squares to be minimized.

Direct calculation

Consider, in a finite difference scheme, one central grid cell (subscript c) and a neighbouring grid cell in direction **i**. The respective transmissivities and lengths of the cells are $\mathbf{T_c}$, $\mathbf{L_c}$ and $\mathbf{T_i}$, $\mathbf{L_i}$ and their common width is **b**. The equivalent internode transmissivity $\mathbf{T_{ei}}$ in direction **i** is calculated by harmonic mean:

$$(L_i + L_c) / T_{ei} = L_i / T_i + L_c / T_c \qquad (3)$$

If the directions of the gradients do not differ too much, equation (2) may be applied and yields:

$$T_{ei\ new} = T_{i_s}\, G_{i_s} / G_{i_a} \qquad (4)$$

where:

a = subscript for actual (measured);
s = subscript for simulation model;
T_{i_s} = internode transmissivity in the model in direction **i**.

Considering that $\mathbf{T_i}$ is known from a previous iteration, the new value $\mathbf{T_{c\ new}}$ could be derived from relation (3). However, modifying $\mathbf{T_c}$ in the central grid cell will also modify the five other internode transmissivities and the five other gradients connected to the central grid cell. Improving simulated gradient $\mathbf{G_{i_s}}$ in direction **i** may worsen other gradients. Optimization by the least squares method will be used to determine the new central transmissivity, which minimizes the sum $\mathbf{S_1}$ of the square of the deviations between observed $\mathbf{G_{i_a}}$ and simulated gradients in the six directions.

$$S_1 = \sum_{i=1}^{6} (G_{i\ new} - G_{i_a})^2 \tag{5}$$

where $G_{i_{new}}$ is the gradient in direction **i** corresponding to the optimal transmissivity.
Equation (2) yields:

$$G_{i\ new} = G_{i_s}\ T_{i_s} / T_{ei\ new} \tag{6}$$

Substituting $\mathbf{T_{ei\ new}}$ (relation (4)) in relation (6) and $\mathbf{G_{i\ new}}$ in relation (5) yields a non linear equation for $\mathbf{T_c}$.

In order to obtain a linear relation, we shall use resistances **R** defined as the inverse of the transmissivities. With the same notations, equation (3) yields:

$$(L_i + L_c)\ R_{ei} = L_i\ R_c + L_c\ R_c \tag{3'}$$

$$G_{i\ new} = G_{i_s}\ R_{ei\ new} / R_{ei\ s} \tag{6'}$$

Using equation (3'), one obtains $\mathbf{R_{ei\ new}}$ and $\mathbf{R_{ei\ s}}$ and then $\mathbf{G_{i\ new}}$ as a function of the central resistance $\mathbf{R_{c\ new}}$ and $\mathbf{R_{c\ s}}$. The part of equation (5) relative to direction **i** writes:

$$(G_{i\ new} - G_{i\ a})^2 = [G_{i\ s}\ (L_i R_{i\ new} + L_c R_{c\ new}) / (L_i R_{i\ s} + L_c R_{c\ s}) - G_{i\ a}]^2 \tag{7}$$

This relation is quadratic in $\mathbf{R_{c\ new}}$; the same applies for the five other directions. The derivative of $\mathbf{S_1}$ with respect to the unknown $\mathbf{R_{c\ new}}$ is linear and yields immediately the optimal value of $\mathbf{R_{c\ new}}$ (the inverse of the optimalnodal transmissivity in the central grid cell). As the neighbouring optimal values $\mathbf{R_{i\ new}}$ are not known the process is iterative and it is our experience that about five internal iterations are enough for stabilization.

The model is then run iteratively to compute the new hydraulic heads and the new gradient $\mathbf{G_{i_s}}$.

Applications of this method to real aquifers show that a very accurate simulation of the observed hydraulic heads is obtained very quickly (after about 10 to 20 successive simulations). However, the corresponding hydraulic conductivities are erratic. In order to improve the method we incorporate a regularization term as described below.

Regularization

To obtain smoother results the optimization process has been modified and in each grid cell the optimal transmissivity is the one which minimizes the deviations of gradients and also the differences of hydraulic conductivity with the six neighbouring cells. In order to keep a linear relation, the differences of the inverse of the hydraulic conductivities (not the resistances) are minimized. A second term $\mathbf{S_2}$ is then added to $\mathbf{S_1}$:

$$S_2 = \sum_{i=1}^{6} (1/K_c - 1/K_i)^2 \tag{8}$$

Introducing the wetted thickness **h** to each grid cell gives:

$$S_2 = \sum_{i=1}^{6} (h_c R_{c_{new}} - h_i R_{i_{new}}) \tag{9}$$

In each grid cell, S (S is defined as a weighted sum of $\mathbf{S_1}$ and $\mathbf{S_2}$) is minimized. In order to have a meaningful weighting factor **w**, the terms $\mathbf{S_1}$ and $\mathbf{S_2}$ are normalized:

$$S = \underbrace{\left[\Sigma(G_{i_{new}} - G_{i_a})^2\right]}_{S_1} \Big/ \Sigma G_{i_a}^2 + w \underbrace{\left[\Sigma(h_c R_{c_{new}} - h_i R_{i_{new}})^2\right]}_{S_2} \Big/ \Sigma (h_c R_{c_{new}})^2 \tag{10}$$

minimizing **S** is equivalent to minimizing $S' = S_1 + w' S_2$, where:

$$w' = w (\Sigma G_{i_a}^2) / \Sigma (h_c R_{c_{new}})^2 \tag{11}$$

The derivative of **S'** with respect to $\mathbf{R_{c\ new}}$ is linear and yields $\mathbf{R_{c\ new}}$ and then $\mathbf{K_c}$.

With this improvement, the method works well as it will be shown in the next section.

APPLICATION TO FOUR ACTUAL AQUIFER SYSTEMS WITH FIELD DATA

The method described above is very easy to incorporate in an existing finite differences model with little modification, and has been incorporated in the MARTHE 3D model described by Thiéry (1990). The method has been applied to the field data of four aquifer systems studied by BRGM and described in Table 1.

The results are displayed in Table 2 with the following notations:

col 2: standard deviation of the hydraulic heads in the discretized observed (target) map;

col 3: regularization weight in percent;

col 4: relaxation coefficient (this is an under-relaxation coefficient designed to slow down the modification of hydraulic conductivities and avoid oscillations);

col 5: number of direct simulations with the model;

Table 1 Description of the four aquifer systems.

Name	Situation	Surface area (km^2)	Number of grid cells	Reason of study
Thau	Hérault (F)	701	701	Water supply
Cailly	Normandy (F)	94.5	378	Water supply
Breil	Aude (F)	0.523	536	Influence of gravel pit
Island of Malta	Mediterranean	241	964	Water supply and salted wedge

Table 2 Inverse modelling applied to four aquifers.

1 Name	2 Stand. dev.obs. (m)	3 Regul. weight (%)	4 Relaxation coeff.	5 Number simul.	6 RMSE (m)	7 Normalized residual	8 Adjacent hydraulic conductivity ratio
Thau a	26.8	0	1.0	20	2.2	0.7	73
Thau b	26.8	20	0.5	12	5.0	3.5	2.2
Cailly	22.4	10	0.7	12	3.0	1.8	2.1
Breil	0.82	15	0.5	12	0.083	1.0	2.5
Malta	0.78	20	0.3	12	0.26	10.8	1.8

col 6: RMSE = Root Mean Square of Error between observed and simulated hydraulic heads;
col 7: normalized residual = square of RMSE / observed variance;
col 8: standard average ratio of hydraulic conductivities between 2 adjacent cells.

Thau aquifer

Application of the method without regularization, starting with a uniform value of hydraulic conductivity equal to $2x10^{-5}$ m s^{-1} yields the results displayed in first line of Table 2. It can be seen that the simulated heads are very close to the observed ones; the standard deviation of error is 2.2 m, which is even overcalibrated because the standard deviation of error due to interpolation between the observation wells was estimated by kriging in the range 3-5 m. Column 8 of Table 2 shows that the map of hydraulic conductivity is not acceptable because the average ratio of hydraulic conductivities between adjacent cells is as high as 73. Figure 1 displays this map which would not be used by any hydrogeologist.

A second inverse modelling was performed with a regularization weight equal to 20% (line 2 in Table 2). The normalized residual was slightly increased to 3.5% but the hydraulic conductivity map was very smooth and physical (see Fig. 1a) with a neighbouring ratio equal to 2.2.

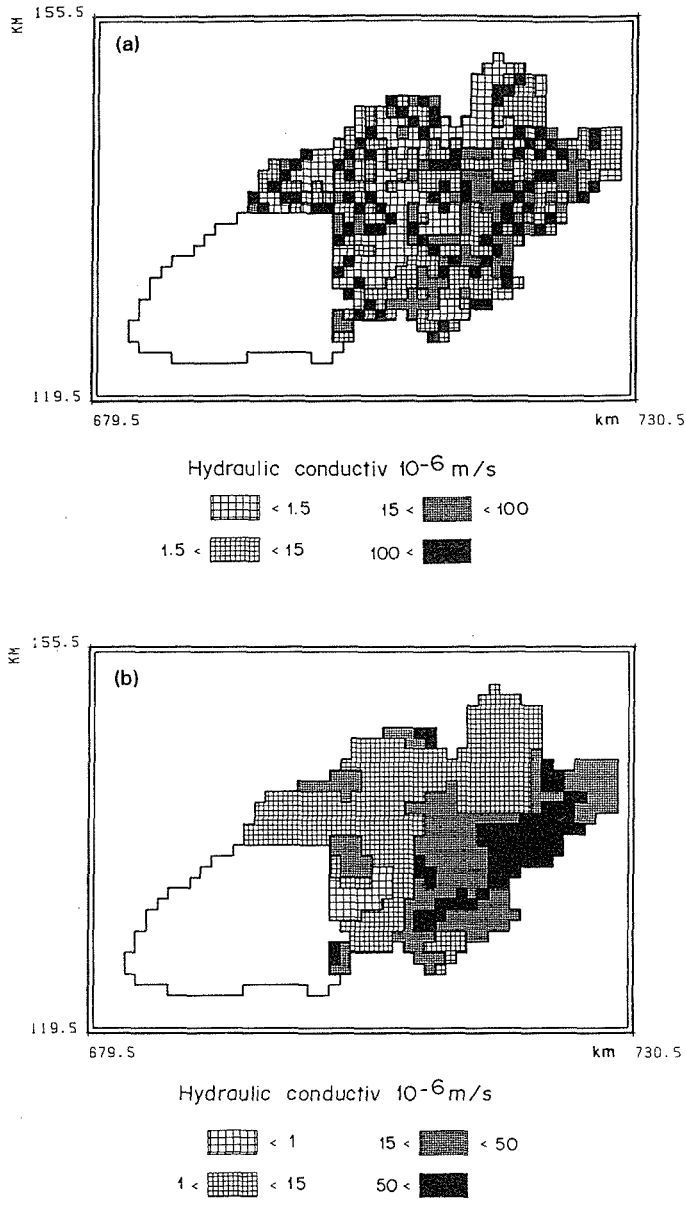


Fig. 1 Thau aquifer: geographical zones of hydraulic conductivity identified by inverse modelling: a) without regulation; b) with a regulating weight of 20%.

Other aquifers

The inverse modelling method has also been applied to the three other aquifers and Table 2 shows that good results were obtained with regularization weights in the range 10-20% (with relaxation coefficients in the range 0.3-0.7 and after less than 15 direct simulations). Figure 2 displays the maps of hydraulic conductivities determined by inverse modelling.

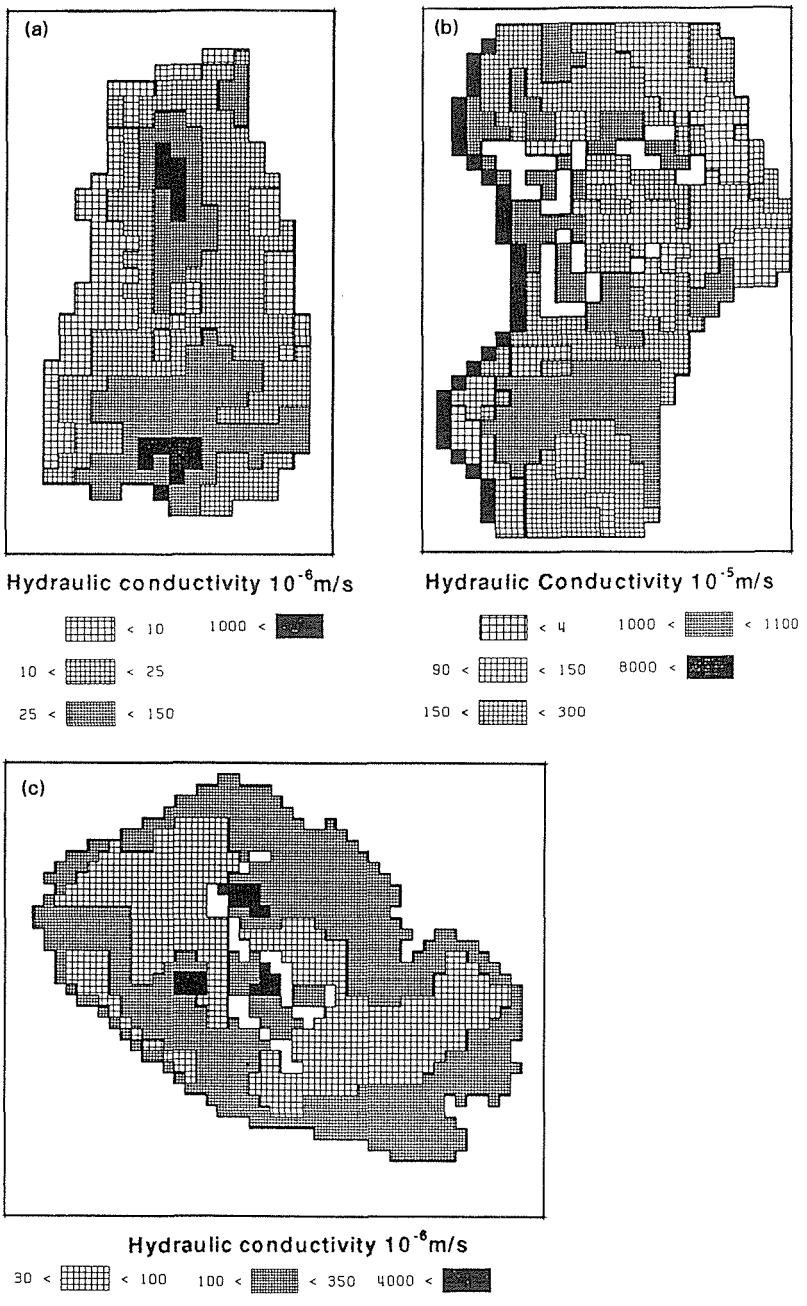


Fig. 2 Geographical zones of hydraulic conductivity identified by inverse modelling: a) Cailly aquifer; b) Breil aquifer; c) Malta.

Final calibration

Homogeneous zones of hydraulic conductivities have been determined automatically from the identified hydraulic conductivity in each grid cell. The nominal value of hydraulic conductivity of each zone has then been optimized

Table 3 Subsequent optimization by geographic zones.

Name	Number of zones	Number of simulations	Initial normalized residual (%)	Final normalized residual (%)
Thau	5	57	3.9	1.4
Cailly	4	51	4.9	1.0
Breil	5+1	53	4.5	0.9
Malta	5	62	18.1	10.8

in order to obtain the best simulation of the observed hydraulic heads. The optimization by zone is described by Thiéry (1993). Table 3 displays the results of the optimization by zones and Fig. 3 compares the simulated and observed heads.

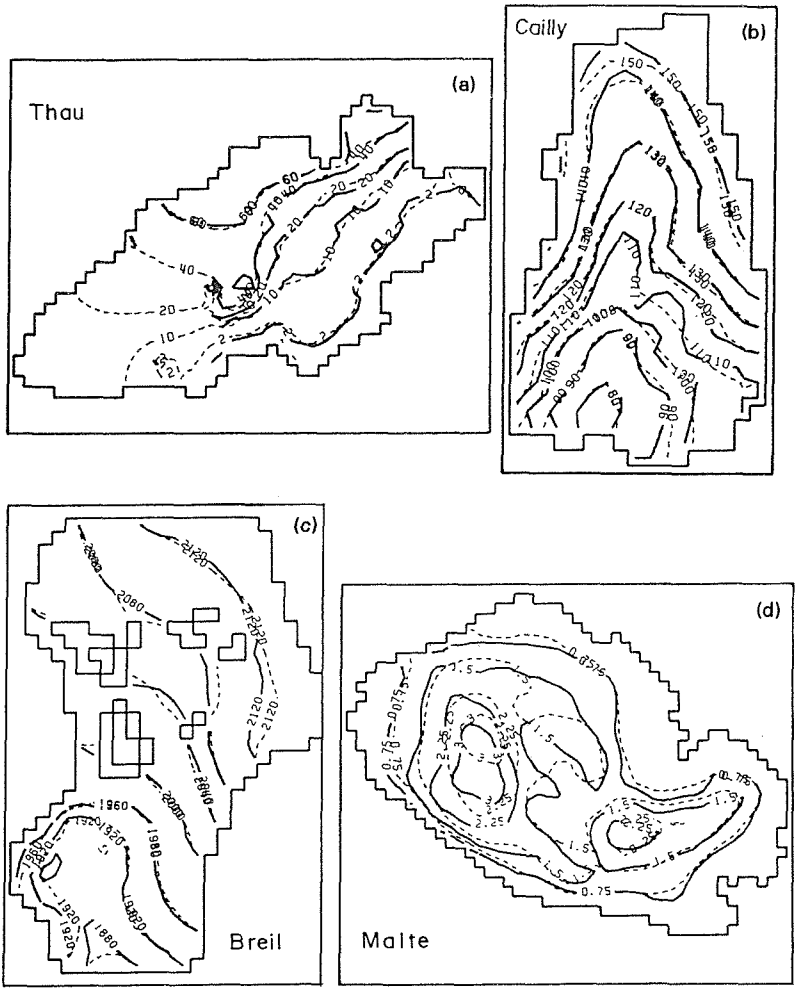


Fig. 3 Comparison of observed (solid line) and simulated (dashed) hydraulic heads, after optimizing homogeneous hydraulic conductivity in the zones identified by inverse modelling. a) Thau; b) Cailly; c) Breil; d) Malta.

CONCLUSIONS

The head gradient method is an efficient way of determining homogeneous zones of hydraulic conductivity in an aquifer system. Combined with an optimization by geographic zone, this method is helpful for automatic calibration of models of real aquifers. It is hoped that it will enable the hydrogeologist to spend more time on the understanding of the hydrogeological scheme rather than on the less gratifying task of calibrating the model.

Acknowledgements This is BRGM contribution no. 93004; this work was financially supported by a BRGM research project.

REFERENCES

Carrera, J. & Neuman, S.P. (1986) Estimation of aquifer parameters under transient and steady state conditions: 1 Likehood method incorporating pior information; 3 Application to synthetic field data. *Wat. Resour. Res.* **22**(2), 199-210 and 228-242.

Carrera, J. (1988) State of the art of the inverse problem applied to the flow and solute transport equations. In: *Groundwater Flow and Quality Modelling* (ed. by E. Custodio, Guirgui & Ferreira.), D. Reidel Publ. Comp., 549-583.

Doherty, J. (1991) MODINV-software for MODFLOW parameter optimisation. Modelling in groundwater resources. Proc. Int. Conf. on Modelling Groundwater Flow and Pollution, Nanjing Univ. Press, China, 397-402.

Döll, P. (1990) An automated trial-and-error procedure to estimate transmissivity from hydraulic head contour maps. In: *Calibration and Reliability in Groundwater Modelling,* Int. Conf. *ModelCARE 90,* The Hague, September, Vol. of Poster Papers, 109-118.

Ensellem, Y. & De Marsily, G. (1971) An automatic solution for the inverse problem. *Wat. Resour. Res.* **7**(5), 1264-1283.

Keidser, A. & Rosbjerg, D. (1991) A comparison of four inverse approaches to groundwater flow and transport parameter Identification. *Wat. Resour. Res.* **27**(9), 2219-2232.

Kessler, N. (1987) Calage semi-automatique des modèles hydrodynamiques. Description du logiciel MICA et deux exemples d'application. Unpubl. Report BRGM n° 87 SGN 871 EAU.

Khan, I.A. (1986) Inverse problem in groundwater: a) model development, b) model application. *Ground Water* **24**(1), 32-48.

Liu, Y. & Li, Q. (1991) A Newton-Khan's algorithm solution to the inverse problem in groundwater. In: *Modelling in Groundwater Resources.* Proc. Int. Conf. on Modelling Groundwater Flow and Pollution, Nanjing Univ. Press, China, 132-138.

Scott, D.M. (1992) An evaluation of flow net analysis for aquifer identification. *Ground Water* **30**(5).

Thiéry, D. (1990) Logiciel MARTHE. Modélisation d'aquifère par un maillage rectangulaire en régime transitoire pour le calcul hydrodynamique des ecoulements. Version 4.3, BRGM report No R32210.

Thiéry, D. (1993) Calibration of groundwater models by optimisation of parameters in homogeneous geological zones. Submitted to the Int. Conf. on "Stochastic and Statistical Methods in Hydrology and Environmental Engineering", June 1993, Univ. of Waterloo, Ontario.

Yeh, W.W.G. (1986) Review of parameter identification procedures in groundwater hydrology: The inverse problem. *Wat. Resour. Res.* **22**(2), 95-108.