IDENTIFICATION OF STEADY-STATE DISTRIBUTED SYSTEMS WITH SPACE-VARIABLE PARAMETERS

G. Bastin

Laboratoire d'Automatique et d'Analyse des Systèmes, Louvain-la-Neuve, Belgium

Abstract. Identification of steady-state distributed systems with space-variable parameters is a crucial problem in aquifer hydrology. This problem is considered in a mathematical framework familiar to control engineers. Assuming a stochastic "random-walk" model for the parameter spatial variability, an optimal estimator is derived in the 1-D case, using a fixed-interval smoothing technique. An application to a numerical example demonstrates the effectiveness of the method. Some generalisations to non linear and 2-D cases are also discussed and an application to the identification of a real underground aquifer in Belgium is briefly reported.

Keywords. Distributed parameters systems, Identification, Optimal smoothing, Water resources systems.

INTRODUCTION

In this paper, we discuss the identification problem of a class of distributed systems, described by a steady-state (independent of time) partial differential equation with space variable parameters.

This problem may seem somewhat academic because identification methods and applications are most often developed for dynamic (instead of steady-state) distributed systems. Our motivation is due to the fact that, in the last decade, the problem of identifying the space-variable parameters of steady-state groundwater flow systems has received much attention as it is a fundamental step in implementing aquifer models oriented towards water supply management and control (see e.g Emsellem and de Marsily, 1971; Neuman, 1975; Cooley, 1977; Bastin and Gevers, 1977; Carotenuto, Raiconi and Di Pillo, 1978; Bastin, 1979; Neuman and Yakowitz, 1979).

The goal of the paper is twofold :

i) Set the problem in a mathematical framework familiar to control engineers.

ii) Suggest how the method might be used in solving distributed systems identification problems in other disciplines than hydrogeology.

PROBLEM STATEMENT

We consider a one-dimensional (1-D) distributed system described by the following partial differential equation :

$$\frac{\partial}{\partial x}\left\{\frac{1}{\alpha(x)}\frac{\partial y}{\partial x}\right\} + q(x) = 0 \qquad (1)$$

$$0 \leqslant x \leqslant \xi$$

with the boundary condition :

$$\frac{1}{\alpha(\xi)}\frac{\partial y}{\partial x}(\xi) = Q \qquad (2)$$

$\alpha(x)$ is a space variable parameter, $y(x)$ and $q(x)$ are respectively the state and the input of the system.

The identification problem (generally called "inverse problem") is the problem of finding the parameter function $\alpha(x)$ when $y(x)$ and $q(x)$ are given.
If (1) is integrated between 0 and ξ subject to the condition (2) the following expression is obtained :

$$\alpha(x) = \frac{\partial y}{\partial x}\left\{Q + \int_0^{\xi} q(x)dx\right\}^{-1} \qquad (3)$$

Clearly, if Q is known, the problem has a unique well defined solution $\alpha(x)$. But if Q is unknown (which is actually the case in most practical situations), the problem is said to be "ill posed" because it has an infinity of equivalent admissible solutions, all of which satisfy the system equation (1). Futhermore, even if Q is given, one must overcome another source of difficulty if one assumes that the function $y(x)$ is not exactly known but is corrupted by a measurement noise $w(x)$:

$$z(x) = y(x) + w(x) \qquad (4)$$

If we introduce the noise-corrupted function z(x) directly into (3), we obtain an inexact expression for the solution of the inverse problem :

$$\alpha^*(x) = \frac{\partial z}{\partial x} \{Q + \int_x^\xi q(x)dx\}^{-1} \quad (5)$$

The error of estimating $\alpha(x)$ by $\alpha^*(x)$ is then :

$$\alpha^*(x) - \alpha(x) = \frac{\partial w}{\partial x} \{Q + \int_x^\xi q(x)dx\}^{-1} \quad (6)$$

This error is thus proportional to the derivative of the noise ($\partial w/\partial x$) which can be arbitrary large even if the noise w(x) itself is small. The estimate $\alpha^*(x)$ exhibits unplausible large amplitude oscillations and is said "unstable" (Neuman and Yakowitz, 1979).

The procedure followed in this paper in order to overcome both of these difficulties (non-uniqueness when Q is unknown and instability in the presence of noise w(x)) is to incorporate, into the mathematical statement of the identification, an additional "random walk" model for the spatial variability of the parameter $\alpha(x)$. This model is similar to those which are often used in identification of time varying parameters (Norton, 1975; Bohlin, 1976).

AN EQUIVALENT DISCRETE MODEL

In practical applications (as in aquifer hydrology) the data z(x) and q(x) are not given in analytical form but are measured at a finite number of discrete points in the interval $(0,\xi)$. On the other hand, the implementation of any identification method on a digital computer necessarily requires some kind of discretization. For both these reasons, it is convenient to formulate the identification method directly for a discrete equivalent model which is now presented. The system is discretized at (n+2) nodes in $(0,\xi)$:

$$x_k = \frac{k\xi}{n+1} \qquad k=0,\dots,n+1 \quad (7)$$

Then equations (1),(2) and (4) are approximated by finite differences as follows :

$$\frac{1}{\alpha_k}(y_{k+1}-y_k) - \frac{1}{\alpha_{k-1}}(y_k-y_{k-1}) + q_k = 0 \quad (8a)$$

$$\frac{1}{\alpha_n}(y_{n+1}-y_n) = Q \quad (8b)$$

$$z_k = y_k + w_k \quad (8c)$$

Eq. (8a) holds for $k=1,\dots,n-1$ and eq.(8c) for $k=0,\dots,n+1$. Define the auxiliary discrete function p_k :

$$p_k = \sum_{i=k+1}^{n} q_i \qquad p_n=0 \quad (9)$$

Then (8a) and (8b) can be combined as :

$$y_{k+1} = y_k + (Q + p_k)\,\alpha_k \qquad k=0,\dots,n \quad (10)$$

IDENTIFICATION OF A STOCHASTIC PARAMETER MODEL

The spatial variability of the parameter $\alpha(x)$ is expressed as :

$$\alpha_{k+1} = \alpha_k + \varepsilon_{k+1} \qquad k=0,\dots,n-1 \quad (11)$$

in which ε_{k+1} denotes the change of $\alpha(x)$ from x_k to x_{k+1}. The system (10),(11),(8c) can then be viewed as a second order discrete linear dynamic system with state (y_k,α_k), output z_k stochastic input ε_k and stochastic observation noise w_k.

We assume that :

(i) Q is known (well-posed inverse problem)
(ii) $W=(w_0,\dots,w_{n+1})'$ and $\mathcal{E}=(\varepsilon_1,\dots,\varepsilon_n)'$ are zero-mean gaussian random sequences with covariances :

$$R = E(WW') \qquad C = E(\mathcal{E}\mathcal{E}') \quad (12)$$

Consider the parameter vector:

$$\theta = (\alpha_0,\dots,\alpha_n,y_0) \quad (13)$$

From the system equations, $\mathcal{E}$ and W can clearly be taken as linear functions of θ :

$$\mathcal{E} = P\theta \quad \text{and} \quad W = Z-Y = Z-A\theta \quad (14)$$

where Z is the measurement vector and Y the state vector :

$$Z = (z_0,\dots,z_{n+1})' \qquad Y = (y_0,\dots,y_{n+1})'$$

It is a well known result that the optimal maximum a posteriori (MAP) estimator $\hat{\theta}$ is the value of θ that maximizes the unconditional density $f(\mathcal{E},W)$ subject to the constraints (17) or equivalently that minimizes the performance index :

$$J(\theta) = \theta'P'C^{-1}P\theta + (Z-A\theta)'R^{-1}(Z-A\theta) \quad (15)$$

This estimator is :

$$\hat{\theta} = (P'C^{-1}P + A'R^{-1}A)^{-1}A'R^{-1}Z \quad (16)$$

It has the following properties :

(i) $\hat{\theta}$ is unbiased : $E(\hat{\theta}) = E(\theta)$ (17)
(ii) The error covariance matrix is :

$$\Sigma_\theta = E\{(\theta-\hat{\theta})(\theta-\hat{\theta})'\} = (P'C^{-1}P+A'R^{-1}A)^{-1} \quad (18)$$

(iii) In the case of perfect observations the estimate α coincide with the unique exact solution α^* of the inverse problem.

$\hat{\theta}$ is the solution of a typical "fixed-interval smoothing problem". Expression (16), which requires the inversion of a large matrix, is not the best from a numerical viewpoint: recursive fixed-interval smoothing as presented by Anderson and Moore (1979) is certainly more efficient. However this approach is adopted here because the generalisation is straightforward when Q is unknown and also for 2-D distributed systems as we shall see later in the paper.

WHITE INPUT AND NOISE SEQUENCES

Assume that W and $\mathcal{E}$ are stationary white sequences and denotes :

$$E(w_k w_\ell) = \sigma_w^2 \delta(k,\ell) \qquad E(\varepsilon_k \varepsilon_\ell) = \sigma_\varepsilon^2(k,\ell)$$

From (14) and (16), the performance index can also be written :

$$J(\theta,Y) = \frac{1}{\sigma_\varepsilon^2} J_\varepsilon(\theta) + \frac{1}{\sigma_w^2} J_w(\theta,Y) \tag{18}$$

with $$J_\varepsilon(\theta) = \sum_{i=1}^{n} (\alpha_i - \alpha_{i-1})^2 \tag{19}$$

$$J_w(\theta,Y) = \sum_{i=0}^{n+1} (z_i - y_i)^2 \tag{20}$$

The identification may then be restated as : find optimal estimates $\hat{\theta}$ and $\hat{Y}$ that minimize $J(\theta,Y)$ under the constraint $Y=A\theta$.
The criterion J_ε is clearly a "smoothing" criterion which will, hopefully, reduce the undesirable oscillations of the solution of the inverse problem while J_w is a "calibration" criterion on the state of the system.

A NUMERICAL EXAMPLE

We assume : $\xi=1.$ $n=49$ $Q=0.05$
$q_k=0.001$ for all k

$\mathcal{E}$ and W are white gaussian sequences produced by pseudo-random number generators with standard deviations :

$$\sigma_\varepsilon^2 = 1.0 \qquad \sigma_w^2 = 1.0$$

The system equations (8c),(10),(11) are solved with initial conditions $y_0=0.0$ and $\alpha_0=4.0$
Fig.1 shows the state y_k and the perturbed output z_k. The true parameter α_k being estimated is represented on fig.2 and fig.3 (note the difference in the vertical scales). Fig.2 also shows the solution α_k^* of the inverse problem. The optimal estimator $\hat{\alpha}_k$, computed with formula (16), is shown on fig.3, together with confidence intervals ($\pm 2\sigma$) computed with formula (18). Comparison of fig.2 and fig.3 clearly demonstrates the superiority of the estimator $\hat{\alpha}_k$ (with respect to α_k^*) : the optimal solution $\hat{\alpha}_k$ is plausible and smooth, and almost all true values α_k lie inside the confidence intervals, while α_k^* exhibits very large and completely unrealistic oscillations.

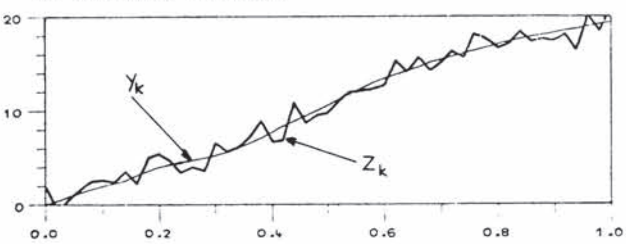
Similar simulations have been performed for a set of different values of the observation noise intensisty σ_w, and the nomalised mean-square estimation error :

$$e_\alpha = \left[\frac{1}{n+1} \sum_{i=0}^{n} \left(\frac{\hat{\alpha}_i - \alpha_i}{\alpha_i}\right)^2\right]^{1/2} \tag{21}$$

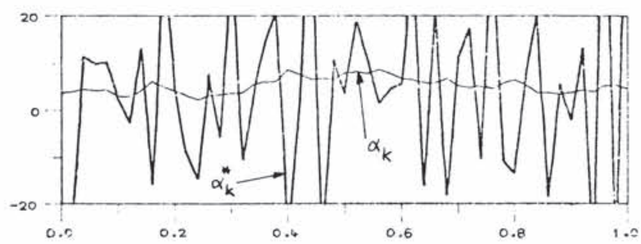
is shown on fig.3bis.

IDENTIFICATION OF UNKNOWN NOISE STATISTICS

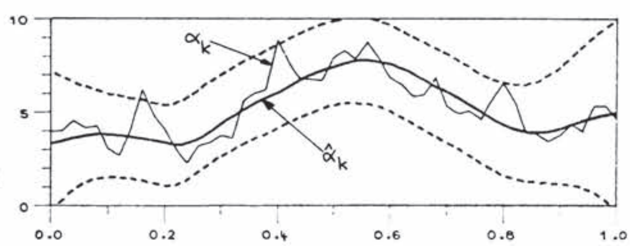
The standard deviations σ_w and σ_ε must be known in order to solve the above identification problem. In cases where the choice of σ_w cannot be based on a priori information, it may be estimated together with θ. On the other hand the selection of a value σ_ε is discussed in the next section.



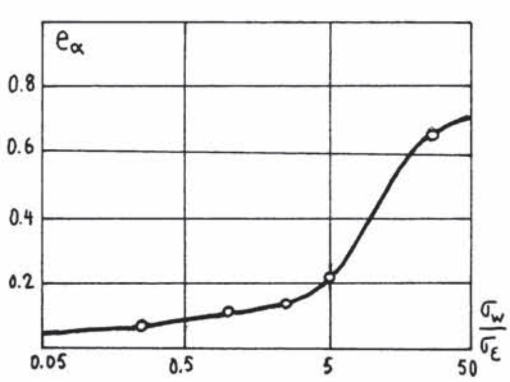
- Fig.1. -



- Fig.2. -



- Fig.3. -



- Fig.3bis. -

If σ_ε is known, one can consider the density $f(\varepsilon,W)$ as a likelihood function that must be maximized with respect to θ and σ_W.
The logarithm of $f(\varepsilon,W)$ is written :

$$-\log f(\varepsilon,W) = (n+2)\log\sigma_W + \frac{1}{2}\left(\frac{\varepsilon'\varepsilon}{\sigma_\varepsilon^2} + \frac{W'W}{\sigma_W^2}\right) + K$$

With definitions (19) and (20) this loglikelihood function can also be written :

$$-\log f(\varepsilon,W) = L(\theta,\sigma_w) = (n+2)\log(\sigma_w) + \frac{1}{2}\left(\frac{J_\varepsilon(\theta)}{\sigma_\varepsilon^2} + \frac{J_w(\theta)}{\sigma_w^2}\right) + K$$

where K is independent of θ and σ_w.
Minimisation of $L(\theta,\sigma_w)$ yields :

$$\frac{\partial L}{\partial \sigma_w} = 0 \implies \hat\sigma_w^2 = \frac{1}{n+2} J_w(\hat\theta) \qquad (21)$$

$$\frac{\partial L}{\partial \theta} = 0 \implies \hat\theta = [(\hat\sigma_w^2/\sigma_\varepsilon^2)P'P + A'A]^{-1}A'Z \qquad (22)$$

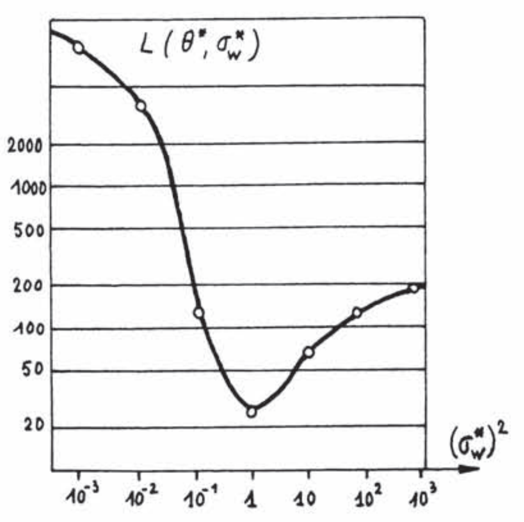
In this case, the optimal estimator becomes a <u>non-linear</u> function of the measurement Z (because $\hat\sigma_w$ is itself a function of $\hat\theta$). Equations (21) and (22) suggest the following procedure to solve the problem :

a) solve equation (22) for a sequence of tentative values σ_w^* : each trial yields a tentative estimate $\hat\theta^*$.
b) select the optimal value $\hat\sigma_w$ as the value of σ_w^* that minimizes the loglikelihood function $L(\theta,\sigma_w)$ or equivalently which verifies (21).

This procedure has been applied to our numerical example :

$L(\hat\theta^*,\sigma_w^*)$ is represented in fig.4

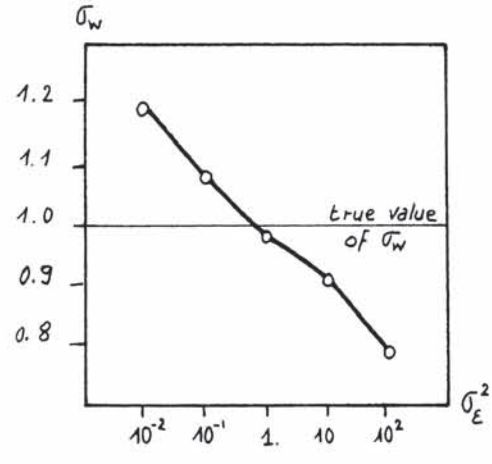
The optimal estimated standard deviation $\hat\sigma_w=0.98$ is very close to the true one ($\sigma_w=1.0$) and , clearly, the solution is unique although the estimator is non-linear with respect to Z.



- Fig.4. -

SENSITIVITY OF THE SOLUTION TO THE VALUE OF σ_ε^2

In the numerical example of the previous sections, we did assume that σ_ε is exactly known ($\sigma_\varepsilon=1.0$).
Let us now consider that the value of σ_ε^2 which is assumed known is in fact inexact (because, for example, it has been approximately estimated from a priori empirical information on the spatial variability of the parameter). In order to test the sensitivity of the solution of our numerical example, the identification procedure was performed for a sequence of inexact values of σ_ε^2 and the results are illustrated by fig.5, where the estimated noise standard-deviation $\hat\sigma_w$ is drawn as a function of σ_ε^2. Clearly, even if the variance σ_ε^2 is taken ten times too large or ten times too small, the error on σ_w remains smaller than ten percent of the true value. In the same way, the parameter estimate $\hat\theta$ and the state estimate $\hat Y$ are very insensitive to σ_ε^2 and remain acceptable even if σ_ε^2 is taken one or two orders of magnitude away from its true value.



- Fig.5. -

Futhermore, in case where the choice of σ_ε^2 cannot be based on prior information on the parameter, it is still possible to infer an estimate of σ_ε^2 from the measurements Z if σ_w is given. Therefore, we define the following auxiliary random sequences :

$$\zeta_k = \frac{z_{k+1}-z_k}{Q+p_k} + \frac{z_{k-1}-z_k}{Q+p_{k-1}} \qquad (23)$$

$$\omega_k = \frac{w_{k+1}-w_k}{Q+p_k} + \frac{w_{k-1}-w_k}{Q+p_{k-1}} \qquad (24)$$

$(k = 1,\dots,n)$

It is easy to show that :

(i) $E(\zeta_k) = E(\omega_k) = 0$ for all k
(ii) $\zeta_k = \omega_k + \varepsilon_k$
(iii) $E(\omega_k\varepsilon_\ell) = 0$ for all (k,ℓ)

Then, obviously, $\Omega=(\omega_1,\ldots,\omega_n)'$ and $\mathcal{Z}=(\zeta_1,\ldots,\zeta_n)'$ are gaussian zero-mean non stationary random sequences with covariances

$$E(\Omega\Omega') = R_\omega$$

$$E(\mathcal{Z}\mathcal{Z}') = \sigma_\varepsilon^2 I_n + R_\omega$$

and the probability density of $\mathcal{Z}$ is :

$$f(\mathcal{Z}) = \{(2\pi)\det(\sigma_\varepsilon^2 I_n + R_\omega)\}^{-n/2} \times \exp\{-\frac{1}{2}\mathcal{Z}'(\sigma_\varepsilon^2 I_n + R_\omega)^{-1}\mathcal{Z}\} \quad (25)$$

The covariance R_ω can be derived from σ_w^2 while $\mathcal{Z}$ contains only linear combinations of measurements z_k. Hence a maximum-likelihood estimate σ_ε^2 may be obtained by maximizing $f(\mathcal{Z})$ with respect to σ_ε^2.

IDENTIFICATION OF STEADY-STATE GROUNDWATER FLOW SYSTEMS.

Consider that equations (1) and (2) describe the flow in an unconfined inhomogeneous 1-D groundwater system. The state y(x) is the "hydraulic head", the inverse parameter $\{\alpha(x)\}^{-1}$ is the "transmissivity", q(x) is the input flow-rate and Q is the lateral flow rate.
With the stochastic approach of previous sections, a plausible solution to the identification of this groundwater flow system is obtained provided that (i) the lateral flow-rate Q is known and (ii) the inverse transmivity is gaussian process. Unfortunately, these requirements are somewhat unrealistic. Indeed, the flow rate Q can almost never be evaluated in practical situations and, on the other hand, it is generally accepted by soil scientists that the transmissivity is a lognormal process, not an "inverse normal" process. We will therefore modify the identification method to take these facts into account.

a) The transmissivity is a lognormal proces

We denote $\tau_k = \log(\alpha_k)^{-1}$

and we replace equations (10) and (11) by :

$$y_{k+1} = y_k + (Q+p_k)\exp(-\tau_k) \quad (26)$$

$$\tau_{k+1} = \tau_k + \varepsilon_{k+1} \quad (27)$$

With this model, the MAP approach leads to the minimization of the following performance index (please compare with(18)) :

$$J(\theta,Y) = \frac{1}{\sigma_\varepsilon^2}\sum_{i=1}^{n}(\tau_i-\tau_{i-1})^2 + \frac{1}{\sigma_w^2}\sum_{i=0}^{n+1}(z_i-y_i)^2 \quad (28)$$

Here $\theta=(\tau_0,\ldots,\tau_n,y_0)$ and the minimization is constrained by equation (26)(instead of (10)) which is non linear with respect to τ_k.

b) Q is unknown.

The constraint (26) is simply a rewriting of the basic equations (8a-b). The unknown Q is eliminated from the problem if the constraint (26) is replaced by the flow equation (8a) and if the flow rate Q is evaluated using (8b), only after θ and Y have been estimated. Obviously, equation (8a) must be written in terms of τ_k.

With these modifications, it is no longer possible to derive an explicit expression of $\hat{\theta}$ (similar to (22)). The solution must be computed by an iterative algorithm.

APPLICATION TO A REAL TWO-DIMENSIONAL AQUIFER.

The performance index (28) can be taken as a dicretized form of the following continuous function :

$$J(\theta,Y) = \frac{1}{\sigma_\varepsilon^2}\int_0^\xi(\frac{\partial\tau}{\partial x})^2 dx + \frac{1}{\sigma_w^2}\int_0^\xi\{z(x)-y(x)\}^2 \quad (29)$$

This continuous form provides a guide for a pragmatic generalisation of the MAP identification approach to two-dimensional (2-D) steady-state distributed systems. Indeed, assume that $\tau(x_1,x_2)$, $y(x_1,x_2)$, $z(x_1,x_2)$ are defined on a 2-D domain S. Then a 2-D analog of (29) is :

$$J(\theta,Y) = \frac{1}{\sigma_\varepsilon^2}\iint_S\{(\frac{\partial\tau}{\partial x_1})^2 + (\frac{\partial\tau}{\partial x_2})^2\}dS + \frac{1}{\sigma_w^2}\iint_S\{z(x_1,x_2)-y(x_1,x_2)\}^2 dS \quad (30)$$

On the other hand, steady-state flow in an inhomogeneous isotropic aquifer is classically described by the following 2-D partial differential equation :

$$\nabla(T\nabla y) + q = 0 \quad (31)$$

$T(x_1,x_2)=\exp\{\tau(x_1,x_2)\}$ is transmissivity, $y(x_1,x_2)$ is hydraulic head, $q(x_1,x_2)$ is input flow-rate.
By a straightforward extension of the previous 1-D approach, the identification problem may be stated as : find $\hat{\theta}$ and $\hat{Y}$ that minimize J(θ,Y) under the constraint of the flow equation (31).
Obviously, the minimization is performed numerically using finite difference approximations of both the performance index and the flow equation.
The method has been applied succesfully to the identification of a real groudwater flow system in the Dyle river basin (Belgium). Detailed descriptions of the aquifer geological characteristics, of the discretization procedure, of the available data and of the optimization algorithm are given in other publications (Bastin, 1979; Bastin and Duqué, 1981).
Here we limit ourselves to the presentation of a few typical results, for a small rectan-

gular portion (3 km x 5.5 km) of the aquifer. An observed hydraulic head contour map $z(x_1,x_2)$ is shown on fig.6. This map has been computed by interpolation between point-wise data (in 28 wells).σ_w is evaluated to about 3.5 meters. Asequence of identifications has been performed for increasing values of σ_ε^2 (from 0.2 to 2000). The results are summarized on table 1 which shows how J_ε , ey and $|z-y|$max vary with σ_ε^2. ey is the mean-square deviation between simulated hydraulic heads $\hat{y}(x_1,x_2)$ and the observed ones. For all values of σ_ε^2 that we have tested, ey is very small (maximum 1 meter) compared with the "a priori" uncertainty on the observed hydraulic head (expressed by σ_w=3.5 meters). On the other hand, spatial smoothing (expressed by J_ε) of the parameter is well illustrated on fig.7,8 and 9 where the spatial behavior of $T(x_1,x_2)$ is drawn for three values of σ_ε^2. More complete results on this application are discussed in Bastin and Duqué (1981).

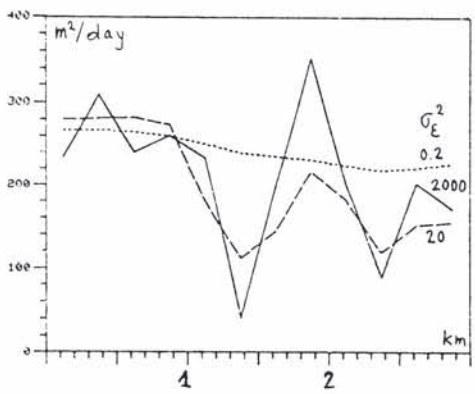


Fig.7. Estimated transmissivity for three values of σ_ε^2 along the cross-section A-A'.

REFERENCES

Anderson, B.D.O.,Moore,J.B.(1979). Optimal Filtering. Prentice Hall.

Bastin,G. and Gevers,M.(1977). Joint use of space interpolation and optimization methods for steady-state aquifer modelling. In Van Steenkiste Ed., Modelling of environmental systems. North-Holland.

Bastin, G. (1979). Approches déterministe et stochastique pour l'identification de systèmes hydrogéologiques.Thèse de doctorat.

Bastin,G. and Duqué,C.(1981).Modelling of steady-state groundwater flow systems. Proc. of IASTED symposium, Davos,Switzerland.

Böhlin,T. (1976). Four cases of identification of changing systems. In Identification : Advances and Case Studies. Academic Press.

Carotenuto,L.,Raiconi,G.,DiPillo,G.(1978). A regularized solution to the identification problem for the distributed parameter model of an underground aquifer. Proc. IFAC Triennal world congress, Helsinki.

Emsellem,Y and deMarsily,G.(1971). An automatic solution to the inverse problem. Water Resources Research, 7(5).

Neuman,S.P.(1975). Role of subjective value judgement in parameter estimation. In Van Steenkiste ed. Modelling and Simulation of Water Resources Systems. North-Holland.

Neuman,S.P. and Yakowitz,S.(1979). A statistical approach to the inverse problem of aquifer hydrology. Water Rsources Research. 15(4)

Norton J.P.(1975). Optimal smoothing in the identification of linear time varying systems Proc.IEE., 122(6).

σ_ε^2	J_ε	ey (cm)	$\|z-y\|_{max}$ (cm)
0.2	0.13	98	437
2	4.30	77	320
20	30.20	33	143
200	70.49	12	36
2000	119.40	3.6	11

Table 1.

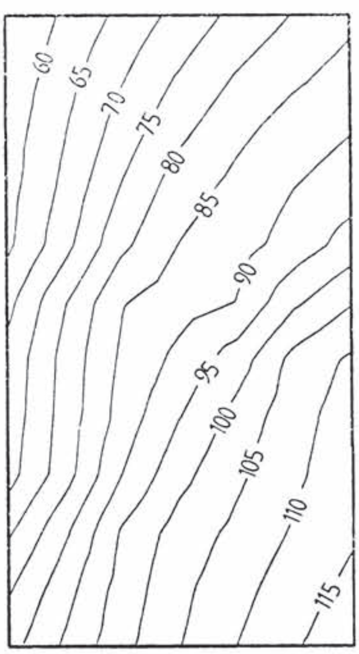


FIG.6. Observed hydraulic heads contour map (meters)

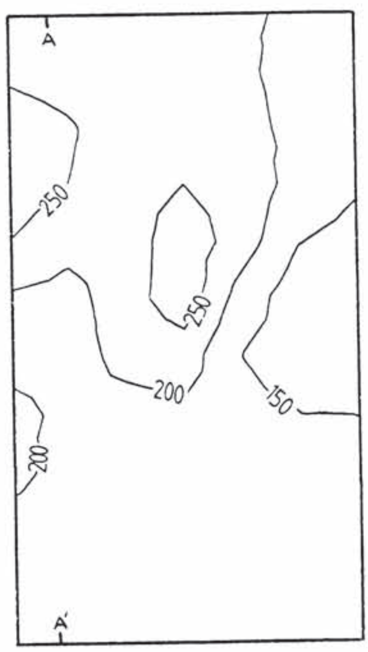


Fig.8. Estimated transmissivity map (m^2/day) σ_ε^2=2

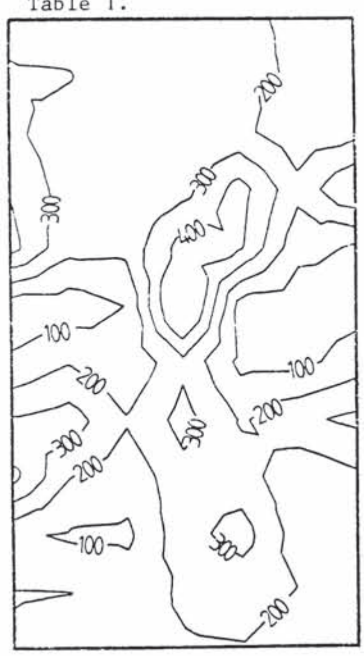


Fig.9. Estimated transmissivity contour map (m^2/day) σ_ε^2=200

For Discussion see page 200