In order to build a mathematical groundwater flow model by the usual parameter estimation methods, a large number of measurements are necessary, distributed over the whole spatial domain of interest. In general, however, the available data are very much insufficient. Often the proposed methods assume a complete knowledge of certain parameters throughout the domain (such as piezometric head, or recharge rate).

In this paper we propose a systematic method which avoids these drawbacks. This finite difference method consists in a two-step procedure: a preliminary estimation is made of the piezometric heads and the heads of the bottom of the aquifer in a subdomain of the grid where enough measurements are available. These estimates are computed using a stochastic interpolation method. The resulting values are then used as data in a second step, in which a quadratic cost functional is minimized under the constraint that the flow equation be satisfied. This cost functional is defined from physically plausible hypotheses on the aquifer structure. The different steps of the method are described and an application to the Oylle basin is presented to illustrate the method.

I. INTRODUCTION.

Consider the two dimensional equation for steady state groundwater flow in an isotropic unconfined inhomogeneous aquifer:

$$\frac{\partial}{\partial x} \left[ K(h-s) \frac{\partial h}{\partial x} \right] + \frac{\partial}{\partial y} \left[ K(h-s) \frac{\partial h}{\partial y} \right] + w = 0$$  \hspace{1cm} (1)

where:

- \((x,y)\) are cartesian coordinates
- \(h(x,y)\) is the piezometric head
- \(s(x,y)\) is the head of the bottom of the aquifer
- \(K(x,y)\) is the hydraulic permeability
- \(w(x,y)\) is the aquifer recharge rate (through the unsaturated zone) or discharge rate.
In the last few years numerous papers dealing with the identification of such a system have been published. See, for example [1] - [6]. In such cases ([1], [2], [4], [5]) equation (1) is replaced by a finite difference equation. A square grid is then superimposed on the studied domain, each elementary square having a sidelength \( Ax = Ay \). The grid has \( M \) nodes, with \( N \) interior nodes and \( M-N \) border nodes. Certain border nodes can be fictitious [5]. The discretization of (1) on this grid gives a set of \( N \) algebraic equations at the \( N \) interior nodes:

\[
\frac{1}{2} \sum_{j \in J} \left[ K_{ij} (h_j - s_j) + K_{ij} (h_i - s_j) \right] \left[ h_j - h_i \right] + W_i (Ax)^2 = 0, \quad i=1, \ldots, N
\]  

The notation \( J \not\in i \) means that the summation is taken on \( j \) over the 4 nodes that are adjacent to node \( i \).

II. DEFINITION OF THE MODELLING PROBLEM.

We call \( \Omega \) the domain over which equation (1) is defined. The domain \( \Omega \) contains the \( M \) nodes of the discretization grid, as well as a certain number of measure points of \( h, s \) and \( K \), which usually do not coincide with the grid nodes. Let us introduce the following notations:

\( I \) is the set of indices of the grid nodes.

The first \( N \) indices will be assigned to the interior nodes.

\( I_h \) is the set of indices of the piezometric head measure points.

\( I_s \) is the set of measure point indices of the heads of the bottom of the aquifer.

\( I_K \) is the set of the permeability measure points indices. (These measurements are obtained through pumping tests).

The aquifer modelling problem can then be stated as follows:

Find numerical values for all components of the vectors:

\[
\begin{align*}
&h = (h_1, \ldots, h_M) \\
s = (s_1, \ldots, s_M) \\
&K = (K_1, \ldots, K_M) \\
&W = (W_1, \ldots, W_M)
\end{align*}
\]

such that

i) these values are compatible with the available data

ii) the flow equation (2) is satisfied.

As the problem is posed now, a few comments are necessary.

2.1. The measure points of \( h, K \) and \( s \) will most often not coincide with the grid nodes. It is therefore necessary to specify what is meant by "the model must be compatible with the available data". Most authors sidestep this problem by assuming that some (if not all) components of \( h \) are known a priori without specifying how they have been computed. It turns out that these computations are not always obvious, particularly when the number of data points is small, as will be illustrated in the example given at the end of this paper.

Let us finally mention for completeness that in [6] a doubly cubic spline function is used to interpolate piezometric heads between data points. The method we use is to first compute estimates \( \tilde{h}_i \) \( i \in I_1 \) and \( \tilde{s}_i \) \( i \in I_2 \) from the available data \( (h_i, i \in I_1) \) and \( (s_i, i \in I_2) \) on two subsets \( I_1 \) and \( I_2 \) of \( I \) using an interpolation method called "Kriging" [7]. This is an optimal
linear stochastic spatial interpolation method; the main ideas are presented in the appendix. The subsets $I_1$ and $I_2$ are chosen in such a way as to include the grid points that are either close to measure points or sufficiently surrounded by measure points so that their estimated values can be considered to be interpolated rather than extrapolated. To simplify the presentation of the method in this paper we shall consider that $I_2 = I$. In the application that we present at the end of the paper, the measure points of $s$ are actually sufficiently scattered throughout the domain so that all grid points can be considered to be well surrounded by data points. Let us finally comment that Kriging is an "exact" interpolation method, i.e. the optimal estimate in a measure point is the measure itself. It is in this sense that we can give a rigorous meaning to the condition that "the model must be compatible with the data".

Once the values $(\mathbf{h}_i \mid i \in I_1)$ and $(\mathbf{S}_k \mid k \in I_1)$ have been computed, we consider these as data for the modelling problem, i.e. we impose that the solutions $\mathbf{h}_i$ and $\mathbf{S}_k$ of the flow equation (2) coincide with $\mathbf{h}_i$ and $\mathbf{S}_k$ for $i \in I_1$ and $k \in I_2$ respectively.

2.2. In most methods that have been proposed so far ([1], [3]-[6]) the vector $W$ is assumed to be completely known. Such hypothesis is not acceptable when $W_i$ represents the recharge rate, in a given node, through an unsaturated layer of several decades. We shall therefore consider the $W_i$ values to be unknown. To simplify the presentation we shall assume that there are no pumping stations in the aquifer. The introduction of pumping does not at all modify the proposed method, provided the pumping rate is known.

2.3. As the modelling problem is posed now, it has an infinite number of solutions whatever the dimension of the sets $I_1$ and $I_2$. This difficulty in the modelling of aquifers is well known and has led many authors to formulate additional hypotheses. Chang and Yeh [5], for example, restrict all transmissivity values which are outside an a priori chosen interval. Garay, Holmes and Das [4] use a second-order polynomial representation of the aquifer transmissivity. Emsellem and De Marsily [2] search for a spatial distribution of transmissivities that is as uniform as possible. We shall develop a similar idea in this paper. One could indeed consider that the most attractive and at the same time physically plausible model is that of a homogeneous aquifer with a uniform recharge rate. Such a model, however, is usually not compatible with the data. We shall try to approach such a model as best as possible by minimizing the following cost function

$$J = \alpha C_1 J_1 + (1-\alpha) C_2 J_2$$

(4)

$C_1$ and $C_2$ are normalization constants that will be specified later. $J_1$ and $J_2$ are defined as follows

$$J_1 = \iint_A \left[ (\frac{\partial \mathbf{S}}{\partial x})^2 + (\frac{\partial \mathbf{S}}{\partial y})^2 \right] \, dxdy$$

(5)

$$J_2 = \iint_A \left[ (\frac{\partial \mathbf{W}}{\partial x})^2 + (\frac{\partial \mathbf{W}}{\partial y})^2 \right] \, dxdy$$

Actually we shall use a discretized version of these two functions.
\[ J_1 = \sum_{i \in I} \sum_{j \neq i} (K_j - K_i)^2 \]
\[ J_2 = \sum_{i = 1}^{N} \sum_{j \neq i} (W_j - W_i)^2 \]
\[ j \in N \]

The differences in the summation between \( J_1 \) and \( J_2 \) are due to the fact that the \( W_i \) need only be defined on the interior points: see equation (2).

2.4. As we mentioned at the beginning of this section, a set \( \{K_i| i \in I_K\} \) of permeability measurements are available. If the number of measurements is large enough, interpolated values of \( K \) can be computed at all grid points. But if only a few measure points are available, as will be the case in the example treated at the end of this paper, each measurement is transferred to the closest grid point. The set of nodes for which an estimate \( K_i \) is known will be denoted \( I_3 \).

The aquifer modelling problem is now reformulated as a constrained optimization problem with a cost function

\[ J(h, K) = (1 - \alpha) C_1 J_1 + \alpha C_2 J_2 \] (7)

The optimization is performed with respect to the parameters \( \{K_i| i \in I\} \) and \( \{W_i| i \in I\} \). The constraints are

(I) \( h_i = h_1 \), \( i \in I_1 \)
(II) \( a_i = a_1 \), \( i \in I(=I_2) \)
(III) \( K_i = K_1 \), \( i \in I_3 \)
(IV) the groundwater flow equation (2) which we shall rewrite as follows

\[ g_1(h, K, s) = -W_1 \], \( i \in I \) (8)

The normalization constants \( C_1 \) and \( C_2 \) are computed as follows

\[ C_1 = \frac{1}{n_K K^2} \], \[ C_2 = \frac{1}{n_W W^2} \] (9)

where \( n_K \) and \( n_W \) are the numbers of terms in the sums \( J_1 \) and \( J_2 \) respectively.

The terms \( \frac{J_1}{n_K} \) and \( \frac{J_2}{n_W} \) therefore represent the mean square deviation between neighbouring permeabilities and neighbouring recharge rates respectively. \( K \) and \( W \) are an approximate mean, over \( I \), of the permeability and the recharge rate, respectively. These approximate values can be adjusted in the course of the optimization procedure if so desired. The variable \( \alpha \), finally, is chosen in the interval \([0, 1]\); it is a weighting factor between the two terms of the cost function.
map no. 1: Lambert coordinates in km.
Bottom of aquifer data points are indicated: for example, at the point number 94, the head is 20 meters.
map n°2 : The studied domain $\Omega$ with:
+ grid points
* piezometry data points
$\Delta$ permeability data points.

The set $I_1$ is contoured.
3. THE PROPOSED ALGORITHM

The algorithm can be subdivided into two major parts.

3.1. Computation of the constraints (I), (II), (III), and of the starting values of the optimization.

Estimates $\hat{\alpha}$ and $\hat{\delta}$, over the whole spatial domain ($i \in I$) are computed first by "Krigeing" (see Section 2.1).

- the estimates $\hat{\delta}$ are considered as constraints (II) and are therefore directly injected in eq. (8)
- the estimates $\hat{\alpha}$ ($i \in I$) are used as starting values of the iterative optimization procedure. The subset of estimates $\hat{\alpha}_i$ ($i \in I_1$) are constraints (I) directly injected in eq. (8).

The values $\hat{\alpha}_i$ ($i \in I_1$) are directly deducted from the data (see Section 2.4) and are constraints (III) injected in (8) and in $J_1$. Estimates $\hat{\alpha}_i$ ($i \in I, \ i \notin I_1$) are subsequently computed by minimizing $J_1$ without constraints. These estimates are easily obtained as the solutions of the following system of equations:

$$\frac{\delta J}{\delta \alpha_i} = \sum_{j \notin I_1} (K_{ij} - K_{ij}) = 0, \ \text{with} \ i \in I, \ i \notin I_3 \quad (10)$$

and $K_{ij} = K_j$ for $j \in I_3$.

The set of all these estimates $\hat{\alpha}_i$ ($i \in I$) are now used as starting values for the iterative procedure.

3.2. The iterative optimization procedure.

A gradient method is used in order to minimize the following Lagrangian:

$$L = J + \sum_{i \in I} \lambda_i (\hat{\alpha}_i - \hat{\alpha}_i) \quad (11)$$

- preliminary step: starting values $\hat{\alpha}$ and $\hat{\delta}$.
- 1st step: compute $\hat{\alpha}_i = -\frac{\delta L}{\delta \lambda_i}, \ i = 1, \ldots, N$
- 2nd step: compute $\delta J/\delta \alpha_i, \ i = 1, \ldots, N$
- 3rd step: compute the Lagrange coefficients so that $\frac{\delta L}{\delta \lambda_i} = 0, \ i = 1, \ldots, N$
- 4th step: compute $\frac{\delta L}{\delta \alpha_i}, \ i \in I$ and $i \notin I_1$
- 5th step: compute $\frac{\delta L}{\delta \alpha_i}, \ i \in I$ and $i \notin I_3$
- 6th step: compute the descent directions by the conjugate gradient method:
  - $d_i^h, \ i \in I$ and $i \notin I_1$
  - $d_i^K, \ i \in I$ and $i \notin I_3$
- 7th step: the new $\hat{\alpha}$ and $\hat{\delta}$ estimates are $\hat{\alpha}_i = \hat{\alpha}_i - \tau d_i^h$.
Figure 1: Evolution of the piezometric head in 4 wells (meters above sea level).
Table 1: Piezometric data.

<table>
<thead>
<tr>
<th>Well's number</th>
<th>Piezometric head (meters above sea level)</th>
<th>Datum</th>
</tr>
</thead>
<tbody>
<tr>
<td>69</td>
<td>94.90</td>
<td>21.10.75</td>
</tr>
<tr>
<td>70</td>
<td>118.86</td>
<td>21.10.75</td>
</tr>
<tr>
<td>71</td>
<td>83.34</td>
<td>12.11.75</td>
</tr>
<tr>
<td>72</td>
<td>87.14</td>
<td>12.11.75</td>
</tr>
<tr>
<td>73</td>
<td>89.18</td>
<td>12.11.75</td>
</tr>
<tr>
<td>74</td>
<td>89.69</td>
<td>12.11.75</td>
</tr>
<tr>
<td>75</td>
<td>107.40</td>
<td>12.11.75</td>
</tr>
<tr>
<td>76</td>
<td>109.85</td>
<td>12.11.75</td>
</tr>
<tr>
<td>77</td>
<td>97.21</td>
<td>12.11.75</td>
</tr>
<tr>
<td>78</td>
<td>113.88</td>
<td>13.10.75</td>
</tr>
</tbody>
</table>

Table 2: Permeability data.

<table>
<thead>
<tr>
<th>Well's number</th>
<th>Permeability (meters/day)</th>
</tr>
</thead>
<tbody>
<tr>
<td>69</td>
<td>13.99</td>
</tr>
<tr>
<td>70</td>
<td>4.96</td>
</tr>
<tr>
<td>75</td>
<td>3.50</td>
</tr>
<tr>
<td>79</td>
<td>13.82</td>
</tr>
<tr>
<td>80</td>
<td>3.72</td>
</tr>
</tbody>
</table>

Map n°3: Bottom of the aquifer by Kriging (isovales x 10^2 meters above sea level).

Map n°4: Piezometry by Kriging (isovales x 10^2 meters above sea level).
and \( \hat{K}_1 = \hat{K}_1 - \tau d_1 \). Then \( w_i^+ = -g - (\hat{h}^+ \cdot \hat{k}^+) = -g(\tau) \).

A nearly optimal gain \( \tau \) is computed by first expanding \( g(\tau) \) around the previous value \( w_i \) up to the first order term in \( \tau \). \( \tau \) is then approximated by a quadratic function of \( \tau \), whose minimum \( \tau^* \) is chosen as the gain.

- \( \text{8th step} \): compute new estimates \( h_i^* - \tau^* d_j \), \( i \in I \) and \( \hat{h}_i^* \)
  \( \hat{c}_i^* - \tau^* c_j^* \), \( i \in I \) and \( \hat{c}_i^* \).

- go to \( \text{1st step} \).

Remark.

It should be noted that in our method, and in contrast with most other methods proposed so far, the boundary conditions need not be known and the flow equation (2) does not have to be solved for either \( h \) or \( K \) at each iteration of the optimization algorithm. We only need to compute the residuals \( W_i \) of the equation, all other quantities being known. This is an important computational advantage, which is due to the hypotheses we have chosen and to the use of the Lagrangian method.

4. A PRACTICAL EXAMPLE

The method has been applied to the identification of a rectangular portion of the "Bruxellion" aquifer in the Dyle river basin (Belgium). The studied domain can be located on map \( n^2 \) and is represented on a larger scale on map \( n^2 \). The discretization grid has been superimposed on the domain. It contains \( N=68 \) nodes, with \( N=42 \) interior nodes \( (\Delta x = \Delta y = 500 \text{ meters}) \).

4.1. Available data.

- head of the bottom of the aquifer: 35 measure points (see map \( n^2 \)). Most of these measures have been made and provided by the Service Géologique de Belgique.
- piezometric head: ten wells (see map \( n^2 \)) have been measured on a more or less regular basis since 1971. Figure 1 shows the evolution of the piezometric head in 4 of these wells, as a matter of illustration. In the example we present here, we have used the data obtained in November and December 1975 (see table 1).
- permeability: five pumping tests (see map \( n^2 \)) were made by Lapanaia \( L \), and horizontal permeability coefficients were computed by Dagan and Foulston methods (see table 2).

The set of indices defined in (3) are now as follows:

- \( I = \{1, 2, \ldots, 68\} \) with \( N=42 \)
- \( I_h = \{69, 70, 71, 72, 73, 74, 75, 76, 77\} \)
- \( I_k = \{69, 70, 75, 79, 80\} \)
- \( I_s = \{81, \ldots, 115\} \)

4.2. Computation of the starting values and constraints.

1) The estimates \( \hat{h}_1, \ldots, \hat{h}_{68} \) and \( \hat{c}_1, \ldots, \hat{c}_{68} \) are computed by "Kriging" with the hypothesis that both the drift and the variogram are linear. The resulting estimated values are presented on maps \( n^2 \) and 4. The subset \( I_1 \) has been chosen as \( I_1 = \{3, 4, 9 \text{ to } 12, 15 \text{ to } 30, 35, 36, 42\} \).
Fig. 2: Final value of cost functional $J$ versus $\alpha$.

Fig. 3: Mean square deviation between permeabilities at neighbouring nodes.

Fig. 4: Mean square deviation between recharge flows at neighbouring nodes.

Fig. 5: Estimated mean recharge rate.
Map n°5 : Algorithm starting values of permeability (x10 meters/day)

Map n°6 : Optimized piezometry for $\alpha = 0.5$.

Map n°7 : Optimized permeability for $\alpha = 0.5$.

Map n°8 : Optimized recharge rate (x10 cm/year).
2) The set $I_3$ contains those nodes that are closest to permeability measure points: $I_3 = \{3, 21, 23, 29, 68\}$. The estimates $\hat{K}_n (I_4 I_3)$ are then computed (see section 3.1). The results are presented on maps $n^s$.  

4.3. Results of the iterative optimization procedure.

The procedure has been experimented with different values of $a$ ranging from 0.2 to 0.8. In all cases we have used an approximate mean permeability $K$ of 0 m/day and an approximate mean recharge rate $W$ of 25 cm/year (see 2.4). The procedure stops when, for 10 successive iterations, the relative variations of the cost $J$ between 2 successive iterations is less than $5 \times 10^{-2}$. The results are presented on figures 2, 3, 4 and 5; they call for the following comments.

1) The values of $a$ which are less than 0.2 must be rejected because they lead to negative recharge rates $W_i$ in several nodes, which is not admissible.

2) In all cases we have obtained a negative permeability at node $n^s$. This indicates that the constraints are too stringent. In order to have a plausible solution, we have therefore replaced this permeability by the average taken over the four neighboring permeabilities.

3) The mean square deviation between permeabilities at neighbouring nodes (see fig. 3) is very insensitive to the value given to $a$. Indeed it varies between 3.47 m/day for $a = 0.3$ and 4.2 m/day for $a = 0.9$.

4) Similarly the computed mean recharge rate (see fig. 5) is almost totally insensitive to the value of $a$; it is in all cases close to 27 cm/year. It is interesting to mention that, using a global rainfall-riverflow model, Hultot, Duplessis and Laurent [8] have estimated the mean recharge rate of the "Bruxellion" aquifer at 27.12 cm/year. This value, which is remarkably close to ours, has been obtained through a totally independent method.

5) On the other hand the mean square deviation between recharge rates at neighbouring nodes (see fig. 4) is extremely sensitive to the value given to $a$. The individual values $W_i$ at the nodes must therefore be interpreted with great caution. However the model does give us an idea of the global trend of the recharge rate distribution.

Finally, the computed maps of piezometry, permeability and recharge rate for $a = 0.5$ are presented (see maps $n^s, n^6, n^7$). Let us note that all the maps have been drawn by a X-Y plotter linked to an automatic cartography program [10].

Appendix: Interpolation by Kriging.

We give here a very short and incomplete presentation of the Kriging method. For more informations, see reference [7].

Let $u= (x,y)$ be a point and $z(u)$ a real function in a twodimensional space. The value of this function is known at $n$ data points: $z_1 = z(u_1), \ldots, z_n = z(u_n)$.

The function $z$ is viewed as a realization of a stochastic process $Z(u)$ which is the sum of two terms:

$$Z(u) = m(u) + Y(u) \quad (A1)$$

where $m(u)$ is the mean, also called the drift, of the process : $m(u) = E[Z(u)]$; and $Y(u)$ is a purely random component. The half-variance of increments is called
"the variogram" : $\gamma_{ij} = E \{ [Y(u_i) - Y(u_j)]^2 \}$. 

Some additional hypotheses are made : 

1) The drift varies slowly in space and is of the form : 

$$m(u) = \sum_{l=1}^{L} a_l f^l(u)$$

where the functions $f^l(u)$ are chosen a priori.

2) The variogram $\gamma_{ij}$ is 2nd order stationary, i.e. it depends only on the euclidean distance $d_{ij}$ between the points $u_i$ and $u_j$ :

$$\gamma_{ij} = \gamma(d_{ij})$$

In the application developed in the present paper, we consider the following particular case :

- first order drift : $m(x,y) = a_0 + a_1 x + a_2 y$ (A4)

- linear variogram : $\gamma_{ij} = \alpha d_{ij}$ (A5)

The constants $a_0$, $a_1$, $a_2$, $\alpha$ are unknown.

Now let us define the linear estimator of $z_0$ :

$$\hat{z}_0 = \sum_{i=1}^{n} \lambda_i z_i$$

(A6)

The coefficients $\lambda_i$ are chosen so that this estimation is unbiased and minimum variance.

1) unbiased : $E(\hat{z}_0) = \sum_{i=1}^{n} \lambda_i E(z_i) = E(z_0)$

(a_0, a_1, a_2), being unknown, the $\lambda_i$ must obey the following equalities

$$\sum_{i=1}^{n} \lambda_i = 1, \sum_{i=1}^{n} \lambda_i x_i = x_0, \sum_{i=1}^{n} \lambda_i y_i = y_0$$

(A7)

2) minimum-variance :

$$\text{VAR} [\hat{z}_0 - z_0] = \sum_{i=0}^{n} \sum_{j=0}^{n} \lambda_i \lambda_j d_{ij}$$

(with $\lambda_0 = -1$) (A8)

Minimizing $\text{VAR} [\hat{z}_0 - z_0]$ with the constraints (A6), one computes the $\lambda_i$.

Then the optimal linear estimate of $z_0$ is :

$$\hat{z}_0 = \sum_{i=1}^{n} \lambda_i z_i$$

(A9)
References