Combining geostatistics and flow simulators to identify transmissivity

Chris Roth\textsuperscript{a}, Jean-Paul Chilès\textsuperscript{a} & Chantal de Fouquet\textsuperscript{b}

\textsuperscript{a}BRGM, Avenue Claude Guillemin, BP 6009, 45060 Orléans Cedex 2, France
\textsuperscript{b}Centre de Géostatistique, 35 rue St. Honoré, 77305 Fontainebleau, France

(Received 15 November 1996; revised 24 April 1997; accepted 4 June 1997)

The reconstruction of the transmissivity field from the more numerous experimental hydraulic head data, an inverse problem, remains the focus of continuing stochastic-based research. The difficulty of this problem arises not only from the complexity of the diffusion equation that links the two variables, but also from taking into account the physical aspects of the site under study: e.g. the boundary conditions, the effective recharge, and the geology. In practical applications, the validity of purely analytical techniques proposed to date is limited by certain simplifying assumptions, like the linearization of the flow equation, made in order to obtain a solution. For this reason, a hybrid methodology combining geostatistical techniques with deterministic numerical flow simulators is proposed. This combination allows the numerical calculation of the direct and cross covariances needed to cokrige the transmissivity from both the transmissivity and hydraulic head data. The flexibility of numerical flow simulators takes away the need for the simplifying assumptions of analytical techniques to apply the proposed methodology. © 1998 Elsevier Science Limited. All rights reserved

Key words: geostatistical simulation, inverse modelling, transmissivity, hydraulic head, flow simulators, cokriging, numerical covariances.

1 INTRODUCTION

The problem treated in this paper is one of mapping transmissivity. That is to obtain the most likely transmissivity value in the field that allows us, for example, to localise zones of high or low transmissivity. This is an estimation (by cokriging) problem and not one of simulation, which is more appropriate when one wishes to reproduce the observed head field, for example, for later velocity or transport calculations. A certain number of techniques\textsuperscript{1–3} have recently been proposed for this simulation problem.

When estimating transmissivity by cokriging, a suitable covariance model must be developed. For about 15 years, researchers have been trying to stochastically incorporate the diffusion equation, in its analytical form, in the transmissivity–hydraulic head covariance model. While a complete review of the techniques proposed to date has recently appeared\textsuperscript{4}, a summary of the main types of approach is provided here. Spectral analysis\textsuperscript{5} was the first approach used to link the transmissivity and hydraulic head covariances. The method of small perturbations was used to provide a first-order solution of the diffusion equation, the head perturbation being expressed as a function of the log-transmissivity perturbation. The framework of this work was that of two-dimensional steady-state groundwater flow with zero effective recharge for which boundary conditions are ignored, or equivalently assumed to have no influence on the hydraulic head field.

Other techniques, such as maximum likelihood procedures\textsuperscript{6,7} or intrinsic random function theory\textsuperscript{8}, were subsequently applied to the same problem. However, while the methodology changed, the basic framework of the problem and the approximations made have remained more or less the same. Some generalisations, such as including a non-constant mean transmissivity\textsuperscript{9} or a constant effective recharge term\textsuperscript{10}, have been made. A recent work\textsuperscript{11} introduces both a variable effective recharge and a fixed configuration of boundary conditions that is applicable in practice when the natural aquifer form is approximately...
rectangular. So, while progress is being made towards a general working framework, the use of a first-order solution of the diffusion equation remains a very limiting constraint in practical applications. While it has been suggested that a first-order solution is valid for a log-transmissivity variance of up to 1, in practice this variance commonly attains values between 2 and 3.

To remove the limiting assumptions associated with analytical approaches, models based on the mixture of geostatistics and numerical flow simulator results have been proposed. Two different techniques based on a similar idea have been proposed. The line of thinking is to krige an initial transmissivity field that is then iteratively modified until the numerical flow simulator results correspond with the observed head values. The modification of the initial field depends on the number, the location and the value of the fictitious transmissivity data added to the actual experimental data set. Another approach is to develop a covariance model that links the log-hydraulic conductivity to the head-squared because, as stated by the authors, “the equations of flow are simplified by the use of the head-squared instead of the head.” Monte Carlo simulations are used to express the head-squared covariance and the cross covariance between the head-squared and the log-hydraulic conductivity as a piece-wise linear function of the parameters defining the log-hydraulic head covariance. The parameters are then obtained by a maximum likelihood procedure under a multi-Gaussian hypothesis that is justified for weakly variable log-hydraulic conductivity. The resulting analytical covariances are used to cokrige the hydraulic conductivity and a constant effective recharge is finally calibrated to the cokriged field as the value that leads to the best experimental fit for the head-squared variable.

A different technique, also based on several Monte Carlo simulations, is proposed in this paper. Given the transmissivity covariance model, a series of transmissivity fields is simulated and the corresponding head fields obtained by a numerical flow simulator. However, the form of the hydraulic head itself is largely determined by the boundary conditions, the effective recharge and the mean transmissivity of the aquifer, and as such masks the influence of the smaller-scale head variations which result from the local conditions, the effective recharge and the mean transmissivity of the aquifer. So, to extract the maximum amount of information from the data, a covariance model linking the log-transmissivity to the head variation will be developed. So the parallel series of fields, or what we shall term joint flow simulations, are used to numerically calculate the direct and cross covariances that lead to the cokriging of the transmissivity field from both the transmissivity and head variation data. The steps of the method are: (1) defining the transmissivity variogram from experimental data and geostatically simulating a large number of transmissivity fields; (2) obtaining the corresponding head fields by numerical flow simulator; (3) calculating the mean head field and, in turn, the resulting series of head variation fields; (4) numerically calculating the required direct and cross covariances for all pairs of points; and (5) estimating the log-transmissivity by cokrigeing from the log-transmissivity and the head variation data, using the numerical covariances.

Given the numerical-based character of the technique proposed, the convergence rate of the numerical covariances and the positive definiteness of the resulting numerical cokriging covariance matrix will be discussed in some detail. The method will be applied to a sandy aquifer from the north of France for which special attention is paid to the variographic analysis of the log-transmissivity given that the measurements taken are neither punctual nor accurate.

2. JOINT FLOW SIMULATIONS

2.1 Geostatistical transmissivity simulations (step 1)

With a view to the case study later treated, the scalar transmissivity is modelled as being lognormally distributed and stationary over the aquifer. The mean and variance of the normally distributed log-transmissivity \( \ln T \), are denoted by \( \mu \) and \( \sigma^2 \), respectively. The transmissivity can be geostatically simulated once its covariance model has been fixed from either the experimental data or from other aquifers presenting similar hydrodynamic characteristics. However, given the skewness of the experimental data, there is a greater likelihood of having a few very large values that may obscure the inference of a covariance model from the experimental covariance. Working in terms of the Gaussian log-transmissivity minimizes the sensitivity of the experimental covariance. Moreover, the convergence rate of Gaussian simulations is orders of magnitude higher than that of the equivalent lognormal variable because of the influence of the high values that make up the tail of the lognormal distribution. Hence, choosing the first variable of interest as \( \ln T \) limits the number of simulations needed for the numerically calculated covariances to converge. Working with \( \ln T \) or with its fluctuations about the constant mean \( \mu \) leads to the same result.

The transmissivity fields are non-conditionally simulated using a method based on the discrete spectral density, which ensures the normality of the probability law of \( \ln T \) and hence the lognormality of the reconstructed transmissivity. These simulations reproduce the experimental covariance and histogram. If we had used a series of conditional transmissivity simulations, the simulated value at the data points is always equal to the experimental value itself. So the numerical covariance between experimental points, calculated as the covariance between two constant values, is equal to zero. Similarly the numerical covariance between an experimental point and a point to be estimated is the covariance between a constant and the point to be estimated, and is also strictly equal to zero. Hence only non-conditional simulations allow us to numerically reproduce the covariance model.

The transmissivity simulation provides punctual values of
In $T$ at a very fine scale over the entire field. From this almost continuous $T$ field, two sets of $T$ values are required for the cokriging equations. Firstly, the $T$ values at the points to be estimated need to be selected. These points, which normally constitute a regular grid, will be denoted by the coordinate vector $x_i = (x_i, y_i)$. Secondly, the $T$ value associated to each point $x_i$ must be obtained where $x_i = (s(x_i), y_i)$ represents the location of the experimental transmissivity data. This involves regrouping the simulated punctual $T$ values around $x_i$ to form a value that is representative of the zone of influence of the given well test. This point, the regularized nature of the transmissivity data, will be discussed in more detail in Section 4.1.

### 2.2 Numerical hydraulic head calculation (steps 2 and 3)

The hydraulic head field corresponding to a given transmissivity field is obtained via a numerical flow simulator that, in our case, is based on a finite differences algorithm. A quick reminder of the cokriging (CK) procedure is presented for the example where $ln T$ obtained at each finite differences grid node as the average head value over the series of independent head simulations.

\[
\text{div}\{T(x, y) \text{ grad } H(x, y)\} + Q(x, y) = 0
\]

where $Q$ is the effective recharge assumed to be known. From this equation, we propose to use the variations, denoted by $\phi$, of the hydraulic head about its mean $h$ as the second variable of interest. These local variations best reflect the influence of local variations of the log-transmissivity. The value of the hydraulic head itself essentially provides the global form of the flow that somewhat masks the influence of the smaller-scale head variations which result from the local variations of the log-transmissivity $ln T$.

The mean hydraulic head, $h(x) = E(H(x))$, can be calculated at each finite differences grid node as the average head value over the series of independent simulations. The value of the head variation $\phi$ is then obtained as the difference between the hydraulic head and its mean: $\phi(x) = H(x) - h(x)$. Variographic studies of the head perturbations show it to vary in a much smoother fashion than $ln T$. This smoothness justifies calculating $\phi$ at the experimental hydraulic head data points $x_i$ by a simple linear interpolation of the four nearest grid node $\phi$ values.

### 3 THE NUMERICALLY OBTAINED COKRIGING SYSTEM

#### 3.1 Calculating numerical covariances (step 4)

Values of $ln T(x_i)$ at the points to be estimated, as well as the $ln T(x_j)$ and $\phi(x_j)$ values at the experimental data points have been obtained for each joint flow simulation. Over a series of such simulations, the direct and cross covariances required for estimation by cokriging can be numerically calculated for each pair of points considered. The covariance of $\phi$ between two experimental points: $\text{Cov} \{\phi(x_j), \phi(x_j')\} = c_{\phi}(\beta, \beta')$ is presented as an example:

\[
c_{\phi}(\beta, \beta') = \frac{1}{N} \sum_{k=1}^{N} \{\phi(x_j) - \bar{\phi}(x_j)}\{\phi(x_j') - \bar{\phi}(x_j')\}
\]

where $\bar{\phi}(x_j) = \frac{1}{N} \sum_{i=1}^{N} \phi_i(x_j)$ (1) where $N$ is the number of independent joint flow simulations, and $\phi(x_j)$ the numerical average of $\phi$ at the point $x_j$ calculated over the $N$ simulations. By analogy, the other direct covariances $\text{Cov} \{ln T(x_j), ln T(x_j')\} = c_{ln T}(\alpha, \alpha')$ and $\text{Cov} \{ln T(x_j), ln T(x_j')\} = c_{ln T}(\alpha, \alpha')$ and the two cross covariances $\text{Cov} \{\phi(x_j), ln T(x_j)\} = c_{\phi,ln T}(\beta, \alpha)$ and $\text{Cov} \{\phi(x_j), ln T(x_j')\} = c_{\phi,ln T}(\beta, \alpha')$ are obtained to provide all the necessary elements to estimate the transmissivity by cokriging.

#### 3.2 Cokriging the transmissivity (step 5)

A quick reminder of the cokriging (CK) procedure is presented for the example where $ln T$ is estimated at the point $x_0$. The estimator is given by:

\[
\text{ln } T(x_0) = \sum_{i} \lambda_i \text{ln } T(x_i) + \sum_{j} \nu_j \phi(x_j)
\]

where the $\lambda_i$ and $\nu_j$ the unknown cokriging weights assigned to the experimental values of ln $T$ and $\phi$, respectively, are specific to the point $x_0$ to be estimated. If the covariance matrices are denoted by $C = [c_{ij}]$, then these weights are obtained by solving the following linear cokriging system:

\[
\begin{bmatrix}
C_{ln T(\alpha, \alpha')} & C_{ln T(\beta, \alpha')} & \mathbb{I} & 0 \\
C_{ln T(\alpha, \beta')} & C_{ln T(\beta, \beta')} & 0 & 1 \\
\mathbb{I} & 0 & 0 & 0 \\
0 & 1 & 0 & \omega \\
0 & 0 & 1 & \tau \\
\end{bmatrix}
\begin{bmatrix}
\lambda_i \\
\nu_j \\
\omega \\
\tau \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
C_{\phi,ln T(\alpha, \alpha')} \\
C_{\phi,ln T(\beta, \alpha')} \\
\mathbb{I} \\
0 \\
\end{bmatrix}
= \begin{bmatrix}
C_{ln T(\alpha, \alpha')} \\
C_{ln T(\beta, \alpha')} \\
\mathbb{I} \\
0 \\
\end{bmatrix}
\]
where the Lagrange multipliers $\omega$ and $\tau$ are introduced to ensure that the estimation is unbiased. Besides the estimated value, the stochastic framework of cokriging theory also provides the associated estimation variance $\sigma(x_0)_{\text{CK}}^2 = \text{Var}\{\ln T(x_0) - \ln T(x_0)^{\text{CK}}\}$.

$$\sigma(x_0)_{\text{CK}}^2 = c_{\ln} r(x_0, x_0) - \sum_{\alpha} \lambda_{\alpha} c_{\ln} r(x_0, \alpha) - \sum_{\beta} w_{\beta} c_{\ln} r(\beta, x_0) - \omega - \tau \quad (3)$$

This estimation variance is guaranteed to be non-negative when cokriging covariance matrix is positive semi-definite. This condition, automatically verified when using standard analytical covariance models, is also satisfied for numerical covariances calculated from expressions like eqn (1), as will be shown for the submatrix $C_{\phi} = \text{Cov}\{\phi(x_2), \phi(x_2')\}$. The matrix $C_{\phi}$ is defined as positive semi-definite if, for any vector $\xi$, it can be shown that $\xi^T C_{\phi} \xi$ is non-negative. According to eqn (1) we can write:

$$\xi^T C_{\phi} \xi = \frac{1}{N} \sum_{k=1}^{N} \left\{ \sum_{\beta} \xi_\beta (\phi(x_2)) - \phi(x_2) \right\} \left\{ \sum_{\beta} \xi_\beta (\phi(x_2') - \phi(x_2')) \right\} \geq 0,$$

$\forall$ vectors $\xi$

where again $N$ is the number of independent simulations. Hence $C_{\phi}$ is positive semi-definite. This proof can be generalized to show that the whole cokriging covariance matrix is positive semi-definite, which guarantees that the numerical cokriging system is itself valid. This all-important property is ensured because the same number of independent calculation pairs (equal to the number of simulations) is used to obtain the numerical covariances.

4 THE CASE STUDY

The case study is a sandy aquifer located in the north of France. While this aquifer represents part of the regional water table, hydrogeologically speaking, it can be treated as an individual entity. The principal sources of effective recharge are from rainfall where the aquifer is unconfined and overlaid by sandy formations, and from the adjoining aquifers further upstream. The aquifer, while unconfined for the most part, is occasionally confined above by a layer of clay. Another clay layer separates the aquifer in question from other underlying aquifers. Fig. 1 shows the aquifer, its location and the boundary conditions as input into the numerical flow simulator. The shaded grey zones represent the location of impermeable boundaries, while the remainder are prescribed head boundaries. Also, the presence of a few gullies inside the domain imposes the maximum head value there. The 45° rotation is made so that the main directions of anisotropy of the transmissivity variogram, as discussed in Section 4.1, are along the horizontal and vertical.

![Aquifer case study with boundary conditions: 41 transmissivity data (●) and 153 hydraulic head measurements (+).](image-url)
The simulation grid and finite differences grid are aligned to the new axes after this rotation.

It is the flexibility of the methodology based on joint flow simulations, as opposed to analytical models, that allows these complex physical features of the aquifer to be taken into account. The irregular boundary conditions, the effective recharge, and the impermeability of overlying geological layers are all used as supplementary information by the numerical flow simulator during the calculation of each hydraulic head field. The subject of numerous geological and hydrogeological studies, this aquifer has been very well sampled. The position of the 41 transmissivity values and the 153 hydraulic head measurements is also shown in Fig. 1. The transmissivity ranges from $2 \times 10^{-6}$ to $1.3 \times 10^{-3} \text{ m}^2/\text{s}$ with an average value of $2.1 \times 10^{-4} \text{ m}^2/\text{s}$.

### 4.1 Variographic analysis of the log-transmissivity

The variographic analysis of the transmissivity needs to take two factors into account: the error component associated with each experimental transmissivity value and the regularized nature of this data, that is, the fact that they do not represent punctual (i.e. point-support) values. Each transmissivity value is representative of the area around the well over which the influence of the pumping test is felt by a modification to the pre-pumping aquifer level. This area, defined by the so-called radius of action of the well, i.e. the distance from the well beyond which the drawdown caused by the pumping test is negligible, is the support over which the transmissivity measurement has been taken. Under certain simplifying hypotheses the radius of action can be approximated as $R_a = 1.5 \sqrt{t/t_S}$ where $t$ is the pumping time and $t_S$ the storage coefficient. Each transmissivity measurement is therefore assumed to be representative of the average transmissivity within the circle of radius $R_a$ centred on the well.

The distribution of the experimental $R_a$ is presented in Fig. 2, where the size of each open circle is proportional to the value of $R_a$. The lack of an accompanying piezometer prevents us from calculating the radius of action for the five wells marked by a dot. Two distinct populations can be identified: the group of wells to the northwest of the aquifer having an average $R_a$ of about 120 m, and the rest for which the average $R_a$ is about 20 m. The first group is due to the presence of recent superficial alluvial deposits only found in that part of the aquifer. To account for these differing support sizes, the transmissivity data is divided into two groups, with the mean $R_a$ of each group being assigned to all wells within that group as the support size associated with the transmissivity values.

Two series of experimental variograms are calculated. The first series is calculated from only the southern population (34 data), for which the average circular support has a radius of about 20 m. The second series is then calculated from all the data. In this case the average support size more than doubles. The regularization of a point variogram model allows us to correctly model the changing behaviour of the experimental variogram according to the size of the support considered. The link between the point variogram $\gamma$ and its regularized counterpart $\gamma_R$ is

Fig. 2. Proportional representation of the experimental radius of action: five missing values (●).
given by:

\[
\gamma_R(h) = \text{measurement error variance} + \frac{1}{V_{R_a} V_{R_a^*}} \int_{V_{R_a}} \int_{V_{R_a^*}} \gamma(u - v) \, du \, dv,
\]

where \(V_{R_a}\) and \(V_{R_a^*}\) are the areas of the circular supports whose radii are \(R_a\) and \(R_a^*\), respectively. These circles are centred on two points separated by the distance vector \(h\).

The measurement error variance, due to the incertitude associated with the transmissivity data, is discussed in the following paragraph. The sill and range of the point variogram model are chosen so that the resulting \(\gamma_R(h)\) correctly fits the experimental variograms for different support sizes. As an example Fig. 3 shows the omnidirectional variogram for all data fitted by a spherical variogram model: the continuous line represents the regularized variogram model \(\gamma_R(h)\) for which the support of the data is taken into account, and the dotted line is its point support equivalent \(\gamma(h)\). Anisotropic behaviour was seen from the individual directional variograms with the range of continuity in the northeast–southwest direction being about 2.5 times that seen in the northwest–southeast direction. As expected, the punctual model has both a higher sill and a less regular behaviour near the origin than its regularized counterpart. This punctual model, used to stochastically simulate the punctual transmissivity fields, produces a more variable and more locally erratic simulation than that which would have been obtained had a punctual model been directly fitted to the regularized data.

The apparent nugget effect in Fig. 3 has been fitted to take measurement errors of the data into account. Transmissivity values are calibrated from the experimental drawdown curve for each well test, so as to obtain the most satisfactory theoretical fit of the experimental curve. They therefore necessarily contain an error component that depends inversely on how well the experimental drawdown curve represents the reality. Tests on the calibrated transmissivity values of the site\(^{24}\) have shown that an overall factor of 5 defines a realistic confidence interval around the transmissivity data, i.e. the ratio between the highest and lowest plausible transmissivity values for a given well test. This multiplicative confidence interval is converted into an additive one when dealing with the log-transmissivity. The resulting measurement error variance is calculated from standard normal confidence intervals and incorporated in the log-transmissivity variogram model as a discontinuity of 0.16 at the origin.

The log-transmissivity was simulated on a regular 5 \(\times\) 5-m grid using the punctual variogram model presented in Fig. 3. Several upscaling techniques\(^{25}\) have been proposed for well transmissivity data. Here we chose to obtain the simulated experimental data as the local equivalent transmissivity, the geometric mean, of the simulated points within the radius of action of each well. This radius varies between 11 and 180 m. So, between 13 and 970 punctual values are used to calculate the simulated values assigned to the data points.

### 4.2 Convergence of the log-transmissivity simulations

Before going on to the more complex step of cokriging with numerical covariances, kriging with numerical covariances should firstly be shown to be equivalent to kriging with the corresponding analytical covariance model. This equivalence depends on the convergence of the two-point statistics calculated over the series of log-transmissivity simulations, that can be tested from the evolution of the numerically obtained variance of \(\ln T\) as a function of the number of simulations \(N\). From eqn (1), for a given point \(x_0\):

\[
\text{Var}(\ln T(x_0))_N = \frac{1}{N} \sum_{k=1}^{N} (\ln T(x_0) - (\ln T(x_0))_N)^2
\]

Standard statistics theory, that provides the probability
distribution of the numerical variance calculated for a given point \( x_0 \): \( \text{Var}(\ln T(x_0)) = \sigma^2/N \), allows the construction of confidence intervals around \( \sigma^2 \), the theoretical variance of \( \ln T \). We consider 10 different points chosen at random in the field and calculate \( \text{Var}(\ln T(x)) \) for each of them. Fig. 4 presents the evolution of \( \text{Var}(\ln T(x)) \) as a function of \( N \), the underlying simulated variance \( \sigma^2 = 2.9 \) and the 95% confidence interval (the dashed curves) around it. The width of this interval, i.e. the expected fluctuations of \( \text{Var}(\ln T(x)) \), decreases slowly so that even after 400 simulations the residual fluctuations can represent up to 14% of the true variance. This slow convergence rate, that depends on \( 1/N \), implies that even by doubling the number of simulations \( (N = 800) \), fluctuations of up to 11% can still be expected.

We note here that the convergence of the numerical variance depends on the convergence of the numerical mean as seen by the form of eqn (4) in the case of the head variation. Any residual fluctuations in the numerical mean are necessarily carried by the numerical variance. So the convergence of the numerical variance implicitly implies the convergence of the corresponding mean. For this reason, behaviour of the numerical mean has not been considered here. These comments also apply to the head variation simulations discussed in the next section.

To visualise the effect of these residual fluctuations, \( \ln T \) is kriged with both the analytical covariances and the numerical covariances calculated from the 400 simulations (Fig. 5). It is reminded that, when considering the transmissivity data in units of \( 10^{-6} \text{ m}^2/\text{s} \), \( \ln T \) ranges between 0.79 and 7.18, with an average value of 4.37. While globally the two kriged maps present the same characteristics, locally the numerically kriged map displays fluctuations not seen when kriging with analytical covariances. These fluctuations, quantified as the standardised difference between the two maps for each point, were shown to be normally distributed with a zero mean and a standard deviation of about 8% over the field. So, when compared with the analytically based method, the numerical kriging result is globally unbiased but presents local fluctuations that are the same order of magnitude as those displayed in the calculation of the numerical covariances themselves.

4.3 Convergence of the hydraulic head

The hydraulic head field corresponding to a simulated transmissivity field is obtained via a slightly modified finite differences algorithm that produces a globally unbiased result. The dimensions of the finite differences grid mesh \((20 \times 50 \text{ m})\) are chosen to reflect the ratio of the longest and shortest ranges of continuity seen in the anisotropic log-transmissivity variogram. There are 40 punctual simulated transmissivity values within each grid mesh. The interface transmissivity, calculated as the geometric mean of the 20 simulated transmissivity values within a diamond-shaped support that extends from one mesh centre to the next, is directly entered into the numerical flow simulator26.

The series of hydraulic head fields are obtained for a constant set of physical parameters (boundary conditions, effective recharge, etc.). The numerical head variation covariances are calculated after having subtracted the mean head from each head field. As for \( \ln T \), the convergence of the head variation is tested by the evolution of its numerically obtained variance as a function of the number of simulations. However, the distribution of \( f(x_0) \) depends on the location of the point \( x_0 \) considered11. The head variation is highly skewed with a low variance near the prescribed head boundaries but has a symmetrical distribution and a high variance over the central part of the aquifer.

Two series of tests were made (a) for five central points and (b) for five points very near to a prescribed head boundary. Note that the vertical scale of Fig. 6(a) is 10 times that of Fig. 6(b). As the exact distribution of \( \phi \) is not known, and
because its variance is not constant over the field, it is not possible to associate confidence intervals to these results. The behaviour for the symmetrically distributed central points is similar to that of the normally distributed ln $T$. On the other hand, for those points near a prescribed head boundary, being highly sensitive to the occasional large value, the evolution of the numerical variance presents some very important jumps even towards the end of the 400 simulations. This means that more simulations would be needed to dramatically reduce the fluctuations in the numerical covariances when there are experimental data points near a prescribed head boundary.

Fig. 5. (a) Kriged ln $T$ using analytically obtained covariances. (b) Kriged ln $T$ using numerically obtained covariances.
4.4 Transmissivity estimation

Given the relatively important residual fluctuations seen in the numerically calculated variance of $f$, it is to be expected that the corresponding cokriged transmissivity estimation will also present some artificially introduced local variability. On the numerically cokriged log-transmissivity field (Fig. 7), this local variability is especially noticeable in the zone around the western impermeable boundary where the estimated $\ln T$ takes on a very spotty appearance. This may be due to the influence of some neighbouring head variation data that are close to a prescribed head boundary.

The striking feature of the cokriged result is the increased variability of the $\ln T$ estimate when compared with the kriged result (Fig. 6). The kriging weights for our case study were always seen to be between 0 and 1. In this case the resulting kriging estimator, a linear interpolator of the $\ln T$ data, is limited in value by the range of the experimental data themselves. The cokriging estimator, on the other hand, takes both the $\ln T$ and $\phi$ data into account with the weight assigned to the second variable being determined by the cross correlations which can be either positive or negative, according to the relative position of the points considered. There is therefore no obvious a priori limit to the range of values, i.e. the variability, of the cokriging estimate. This can be seen in the southern part of the aquifer where the kriged estimate is replaced by a much lower cokriging estimate. This is especially pleasing as this low cokriged transmissivity agrees with hydrogeological studies of the aquifer that confirm the existence of important head gradients in this part of the aquifer that can only come from very low transmissivity values. Similarly, the zone of high cokriged transmissivity seen through the central part and to the north corresponds well with the manual calibration of the aquifer in a zone that is known to have a smooth head profile. So, the cokriged map presents features that, coming from the influence of the head data and therefore unseen on the kriged map, are consistent with the hydrogeologist’s manual interpretation of the aquifer’s characteristics.

It must be stressed, however, that the cokriged map does...
not reproduce the reality. It can only be used as an indicator of the global behaviour in the transmissivity representing the most likely value of the transmissivity at any point, given the neighbouring transmissivity and head values. In particular the linear cokriging estimate only respects the degree of linear relationship between the two variables, and not the non-linear relationship given by the flow equation. The solution of the flow equation in the cokriged transmissivity field will not produce heads equal to the measured head values. Should one be interested in the correct reproduction of the head field from the resulting transmissivity field, simulations should be considered.

5 CONCLUSION

The combination of geostatistics and numerical flow simulators thus provides a powerful technique for reconstructing the transmissivity field from the available transmissivity and head data. Using a probabilistic framework to model the spatial behaviour of the transmissivity, the corresponding hydraulic head field is determined by a numerical flow model based on a modified finite differences algorithm. A series of such joint flow simulations leads to the numerical calculation of the direct and cross covariances needed to cokrige the transmissivity. The cokriging estimator is preferred to the kriging estimator so as to incorporate the influence of the more numerous and more precise hydraulic head data in the estimated transmissivity field. For the practical application presented here, it was shown that a large number of joint flow simulations is necessary for the numerically calculated covariances to converge. This is the sole disadvantage of the numerical method when compared with available first-order analytic methods: it is computer intensive with respect to the number of simulations required and the space required to store the necessary numerical covariances.

On the other hand, the proposed method is very flexible. The standard first-order approximations, multi-Gaussian assumptions and constraints on boundary conditions or the effective recharge of analytical approaches are no longer needed when solving the diffusion equation with a numerical flow simulator, and could not realistically be applied to the case study treated. So, for those cases where the use of quicker analytical techniques is not possible, the proposed numerical method provides a solution that is consistent with the physical relationships that govern groundwater flow.

ACKNOWLEDGEMENTS

This work was financed by the BRGM. The authors would also like to thank the French Government Agency ANDRA for the use of the data, and P. Delfiner, J.P. Delhomme, G. de Marsily and G. Matheron for their comments. Thanks also go to the reviewers for their remarks that have improved the presentation of this paper.

REFERENCES

1. RamaRao, B.S., LaVenue, A.M., de Marsily, G. and Mar- jetta, M.G. Pilot point methodology for automated


