Acknowledgement is made to C. Srinivas, B. Ramaswamy and M.F. Wheeler for the use of Figure 3 on page 246, which appears on the front cover of this book.

Computational Methods in Water Resources IX

Vol. 1: Numerical Methods in Water Resources

Editors: T.F. Russell, University of Colorado at Denver, USA
R.E. Ewing, University of Wyoming, USA
C.A. Brebbia, Wessex Institute of Technology, University of Portsmouth, UK
W.G. Gray, University of Notre Dame, USA
G.F. Pinder, University of Vermont, USA

Computational Mechanics Publications
Southampton Boston

Co-published with

Elsevier Applied Science
London New York

Kriging and Cokriging Applied to Data: Influence on the Results of a Regional Groundwater F.E.M. Model

A. Dassargues

Laboratoires de Géologie de l'Ingénieur, d'Hydrogéologie et de Prospection géophysique (L.G.I.H.), University of Liege, Bat. B19, 4000 Liège, Belgium

ABSTRACT

Before regional modelling by Finite Element Method, all possible information available from field measurements and tests has to be taken into account. As these data can be attributed to zones of limited extension, one has always to choose extrapolated or interpolated values in the zones where no data are available. At this stage, a good knowledge of the geological regional conditions is used for choosing interpretative values, but geostatistical techniques such as kriging and cokriging can be helpful in many cases. Four methods will be used and compared in order to get a maximum benefit from the data coming from the regional alluvial aquifer : (a) application of a parabolic regression between the permeability coefficient and the logarithm of the apparent electrical resistivity of the alluvial sediments and then drawing a permeability map deduced from numerous (more than 200) resistivity measurements; (b) kriging the permeability values deduced from pumping tests only ; (c) application of (a) and then kriging the permeability on basis of all the direct and indirect data ; (d) cokriging of the permeability versus the electrical resistivity in order to take advantage of this second data set. The comparison will be made before and after the running stage in the F.E.M. code. The results will help us to deduce the optimum treatment, for this case, in regard to the model calibration.

INTRODUCTION

Additionally to measured piezometric maps providing initial conditions and calibration reference targets, regional groundwater modelling needs a good knowledge of the hydrogeologic properties for the parameterization. The data about properties such as hydraulic conductivity and storage coefficient consist of point measurements. The values vary erratically in space due to the geologic variability of the domain and to measurement errors, approximations and interpretation assumptions. An easy way to get an idea of the spatial interrelations (the "spatial structure") of the measured values is to draw contour map by hand. The limited number of measured values forces the hydrogeologist to introduce a big amount of subjective interpretation in the contouring job. This interpretation is mainly influenced by the knowledge of the regional or local geological context that the hydrogeologist has in his mind on basis of all available information and on basis of his own experience. The degree of interpretation in the inter- and extrapolations is highly variable so that the accuracy or the uncertainty of such a map cannot be specified.

Geostatistical techniques such as kriging are known as providing a valuable help for the transmissivity or the hydraulic conductivity estimation in the studied aquifer (Delhomme [1], Ahmed and de Marsily [2], Neuman and Yakowitz [3], Neuman et al [4],...). The kriging technique is applied most often on data obtained from pumping tests. In practice, such data are rare but other data strongly correlated with the hydraulic conductivity can be obtained more easily and should be introduced somehow in the estimation process.

In the case of the alluvial aquifer studied here, a first set of data containing hydraulic conductivity values is completed with a second set of data containing numerous electrical resistivity values, measured by electrical soundings. In order to prepare the parameterization of the regional groundwater F.E.M. model, four techniques are used and discussed, combining parabolic regression, kriging and cokriging. This last technique allows to estimate by kriging the value of the main parameter (hydraulic conductivity) using not only the pumping well measurements (first data set) but also the electrical resistivity values (second data set).

A measured piezometric map is used as a calibration reference. The piezometric heads could also be treated by geostatistical non-stationary kriging techniques, but it does not come within the scope of this paper.

The different data sets of treated equivalent hydraulic conductivity values are introduced in the finite element meshing network of the regional model. The simulation results show some differences. The advantages and disadvantages of each treatment can be deduced comparing the computed piezometric map with the reference map of the calibration. We conclude with a discussion on the contribution of such geostatistical techniques as a helpful preprocessing of the data for the calibration procedure of the F.E.M. model.

STUDY AREA AND AVAILABLE DATA (Dassargues and Lox [6])

The study area is a (17 x 2) km^2 zone in the alluvial plain of the Meuse River between the city of Liège in Belgium and the city of Maastricht in the Netherlands. An alluvial water table aquifer is lying above Paleozoic and Mesozoic semipervious formations and below thin (0 - 5 m) loam deposits. This aquifer is mainly constituted of clean gravels and silty to sandy gravels. The hydraulic conductivity is strongly influenced by the granulometric content of the gravel deposits. Silty gravels will be characterized by relatively low hydraulic conductivity values from 1 10^{-4} to 5 10^{-3} m/s when clean gravels will be affected by K values from 5 10^{-3} to 1 10^{-1} m/s.

Hydraulic conductivity measurements obtained from pumping tests are available at 22 locations and electrical apparent resistivity (measured by electrical soundings using Schlumberger profiles) are available at 220 locations, of which 22 are common to both measured properties. Piezometric data have been collected, totalling 127 points (piezometers and existing wells). Additionally, other geological data have provided about 200 points with data coming from cone penetration tests and seismic soundings. On basis of all these data, 53 interpretative transverse cross-sections have been drawn giving the spatial variations of the alluvial deposits.

TREATMENT OF HYDRAULIC CONDUCTIVITY (K)

a) Parabolic regression

As mentioned above, 22 measurements of K by pumping tests can be correlated directly with 22 apparent electrical resistivity (ρ_a) measured exactly at the same location.

Reporting the values in a (log K, ρ_a) diagram, a parabolic regression is found applying least square method :

$$\log K = a + b \cdot (\rho_a) + c \cdot (\rho_a)^2 \quad (1)$$

For the studied data, we obtain a = -4.797, b = 0.021 and c =-3.37 10^{-5} with a parabolic correlation coefficient of 0.9.

A first data set of K values is calculated applying now this regression at all the measured values of the electrical resistivity.

b) Kriging

This well-known method is applied to ln K as the logarithm of the hydraulic conductivity generally passes normality tests (Hoeksema and Kitanidis [6]). It means that the spatial structure of ln K is much better than the spatial structure of K and, consequently, the variogram shows a stronger correlation.

The definition of the variogram (γ) of a variable Z is written : $\gamma(d) = (1/2) \operatorname{var} [Z(x+d) - Z(x)]$ (2)

where d is the distance between measurement points taken by pairs. To estimate the variogram, we assume ergodicity on the increments (i.e. space averages can be used to estimate the averages in the whole set of realizations) and intrinsic hypothesis (i.e. the variance of the first order increments of Z is finite and these increments are themselves second order stationary) (de Marsily [7]). Kriging with Z = ln K will produce an unbiased estimation for the Z values and biased values for K. However, as we are preparing equivalent values of K for their introduction in a finite element mesh, assuming that the dimensions of the Representative Elementary Volume (R.E.V. after Bear and Verruijt [8]) are of the same order as the dimensions of the mesh, geometric mean of K are more appropriate than arithmetic mean to calculate equivalent values. Arithmetic mean of ln K calculated by the block kriging will automatically give geometric mean of K.

The experimental variogram has been calculated considering only K from pumping tests, using GEOEAS package (Geostatistical Environmental Assessment Software) (Englund and Sparks, [9]). A spherical model has been fitted to obtain a sill of 0.113 corresponding to a range of 3500 m, with a nugget effect of 0.04 . The cross-validation procedure has lead to obtain finally a sill of 0.093 with a nugget effect of 0.06 ,expressing that the estimator is unbiased and and that the estimation error is as small as possible .

For a confidence interval chosen equal to 95 % , we obtain $\ln K = (\ln K)^* \pm 2.\sigma$ where $(\ln K)^*$ is the estimation on ln K and σ is the standard deviation of the kriged values and $K^*/e^{2\sigma} \le K \le K^* . e^{2.\sigma}$ where $e^{2.\sigma}$ is called the uncertainty for the chosen confidence interval.

The limited number of available data and their uniform spatial distribution have lead to kriged values assorted with very high uncertainties so that it was impossible to consider this K data set in the forthcoming finite element simulation.

c) Parabolic regression and kriging

Applying parabolic regression in a first step, a mean standard deviation could be calculated between the values estimated by regression and the real values. This mean variance corresponds to the error committed when determining ln K by the parabolic regression. Then drawing the experimental variogram using all the values of ln K not withstanding their respective level of uncertainty, we obtain a strong nugget effect in the variogram. It corresponds for an important part to the variance of the error committed by application of the regression. Taking into account this nugget effect in the kriging is equivalent to kriging with a constant uncertainty on the data.

The experimental variogram is fitted with a spherical model and adjusted by the cross-validation procedure to obtain a sill of 0.45 corresponding to a range of 1000 m with a strong nugget effect of 0.7 . The map of kriged hydraulic conductivity (K) joined to the uncertainty map calculated for a confidence interval of 95 % (figure 1), shows values comprised between 2 10^{-3} and 2.8 10^{-2} m/s for K and uncertainties comprised between 2.2 and 4.4 (in the boundaries area). Similar trends to the K map obtained by application of the parabolic regression are observed. However we have to remark that the calculated uncertainty is underestimated as the computation does not take into account the variable uncertainty on the electrical resistivity measurements.

d) Cokriging

Theoretically, cokriging is the best technique to improve the estimation of a variable using the information on other spatially correlated variables. In fact, no assumptions are made on the nature of the correlation between the two data sets. The basic theory of cokriging has been described by Journel and Huijbregts [10], Myers [11], De Smedt [12]. This technique improves the estimation and reduces the variance of the estimation error, but at the same time the calculation of the cross variogram and the fitting of a theoretical model can sometimes become difficult, particularly when the two variables are not strongly correlated (Ahmed and de Marsily [2]). A particular method introduced by Myers [11] is here applied on transforms of the variables : ln ρ_a and (-ln K). The experimental variograms have been drawn and fitted to spherical functions .

For our case, it has appeared that the cokriged estimates of K and their uncertainties are still largely influenced by the limited number of K data from pumping tests and their spatial distribution concentrated on 3 or 4 zones (figure 2).

Another cokriging has been computed considering that the error committed by applying the parabolic regression is constant. The variogram relative to the apparent electrical resistivities is written :

$$\gamma_{\ln \rho_a}(d) = \gamma_{\ln K} + \sigma^2_{error}\left(1 - \delta(d)\right) \qquad (3)$$

where σ^2_{error} is the error variance. The second term is considered as a pure nugget effect. The estimated values of K and their uncertainties are similar and of the same order as in the first cokriging.

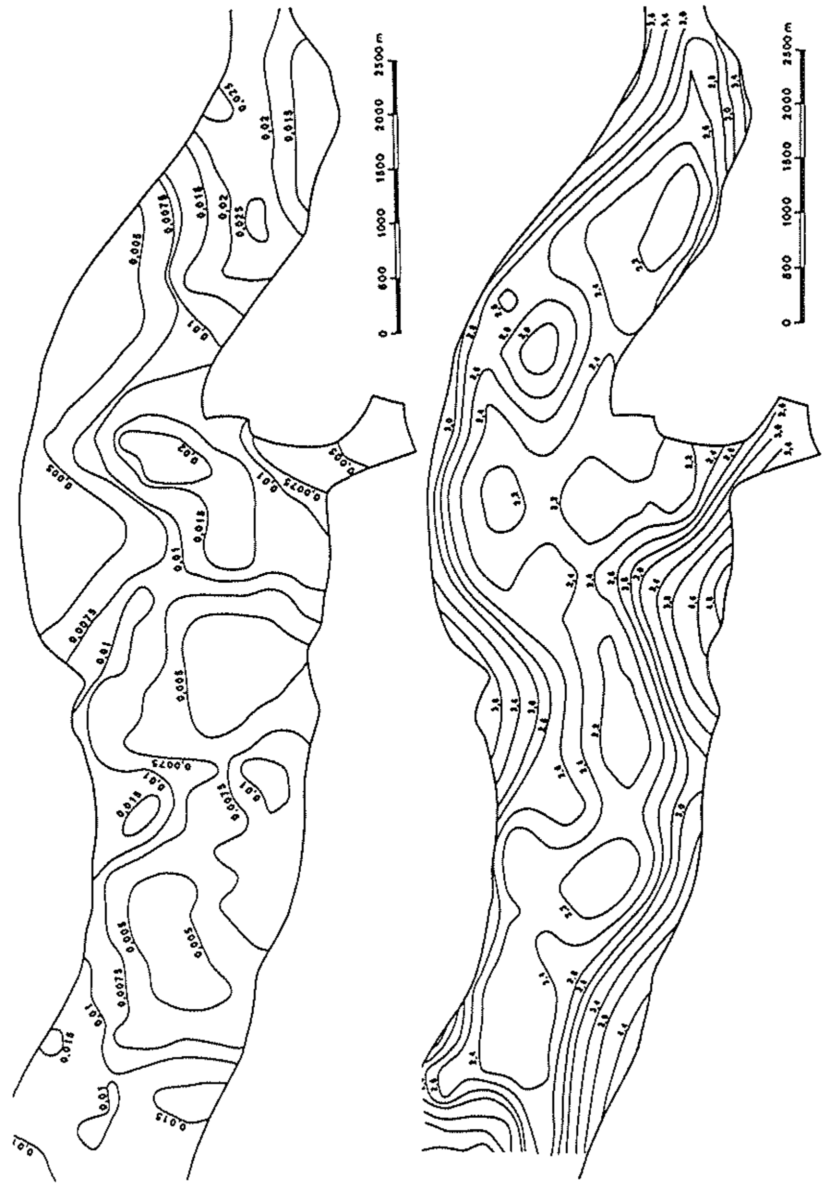


Fig.1 Map of block kriged K (above) and uncertainties with a 95 % confidence interval using method c .

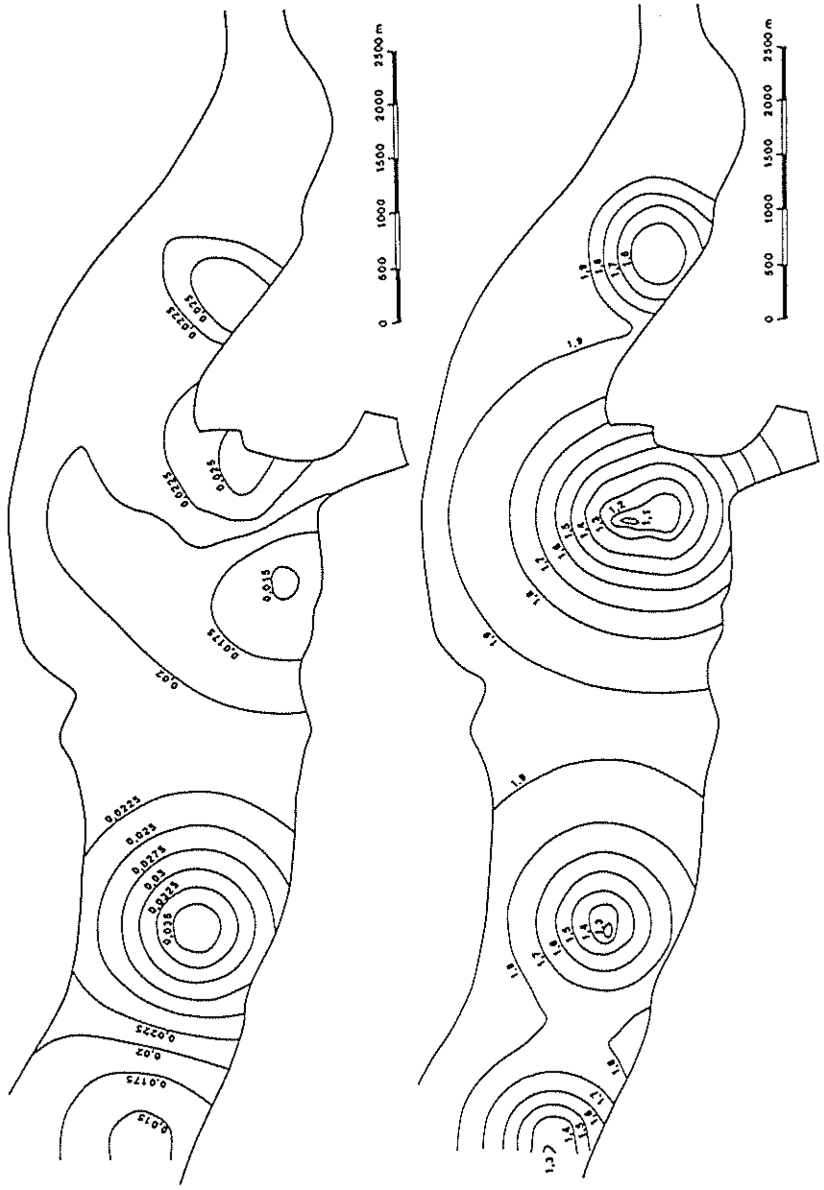


Fig.2 Map of block kriged K (above) and uncertainties with a 95 % confidence interval using method d .

SIMULATIONS USING A FINITE ELEMENT CODE

Our non linear code, called LAGAMINE (Charlier [13]), has already been used to simulate different groundwater situations (Dassargues [14]). In this case, a non linear storage law has been used to simulate the water table conditions of the alluvial aquifer (Dassargues et al.[15]).

The 3D meshing network and the numerous constraints are accurately represented in the model : complex geometrical features, uniform recharge, leakage from the canal, infiltration from the hills feeding the alluvial plain, local uplift infiltration from a confined aquifer in karstified limestones underlying the alluvial plain, impervious banks in some places with drain tubes along the River Meuse , pumping and recharge. In such a complex case study, the non linear code with implementation of the storage law for modelling the water table, is very useful (Dassargues [14]) to avoid the need for an automatic meshing procedure on a regional domain with highly heterogeneous conditions and more than 2350 finite elements disposed in four layers.

Independently to the geostatistical study, the calibration of the model has been completed by "trial and error" method using 26 "materials" of different hydraulic conductivity. The spatial distribution and the values of K have been adapted at each run until the calibration has been considered as satisfactory. An agreement between measured and computed values has been reached with a mean discrepancy value of 0.10 m. This steady state calibration provided the awaited flow results at both internal and external boundaries. One of the aims of the supported part of the study was indeed to determine the flow entering in the River Meuse along this portion of its alluvial plain: a flow of 5.28 m^3/s was found.

Then, all other parameters and constraints maintained unchanged, simulations have been run using the treated values for K. As the finite elements of the 3D mesh do not correspond exactly to the kriging blocks, an interface program assign to each element the kriged K value of the block on basis of barycenter computations.

Computed piezometric maps, obtained with data treated by the methods a , b , c , and d (described here above), are compared with the reference piezometric map. As an example of results, the figure 3 shows the computed piezometric map obtained with K data kriged after application of the parabolic regression (method c), in regard to the reference piezometric map.

Looking at all the results, it has appeared clearly that in regard to the calibration procedure the treatment methods c and d were the more reliable. The simulations with minima and maxima values of K considering the confidence interval of 95 % provide variation ranges in flow results. For example, the flows entering in the River Meuse are computed from 3.79 to 6.38 m^3/s with K treated by the method c and from 4.60 to 7.04 m^3/s with K treated by the method d, in regard to 5.28 m^3/s computed by the calibration described above.

The uncertainty characterizing the K values has been translated somehow in terms of uncertainty in the flow results. The piezometric maps obtained with the K_{min} and K_{max} data sets (from methods c and d) were not better than those obtained before and, additionally, we have to remark that the results of these simulations correspond not necessarily to the extremes in the computed piezometric levels.

Comparing the results of the simulations with treated K (methods c and d) to the results of the successive runs needed by the "trial and error" calibration, we have observed that the first run with treated K was already more accurate than the eleventh run of the normal calibration (figure 4). It is of course a very subjective comparison.

CONCLUSIONS

It should be difficult to choose definitely an ideal method to prepare optimally the K values data set in order to facilitate the calibration procedure. It is clear that the application of methods such as kriging or cokriging based on a sufficient number of data points is positive to shorten the calibration. However, it could be difficult to start a manual trial and error calibration with kriged or cokriged values of K . Indeed, many slightly different K values transform the domain in a kind of mosaic difficult to modify later.

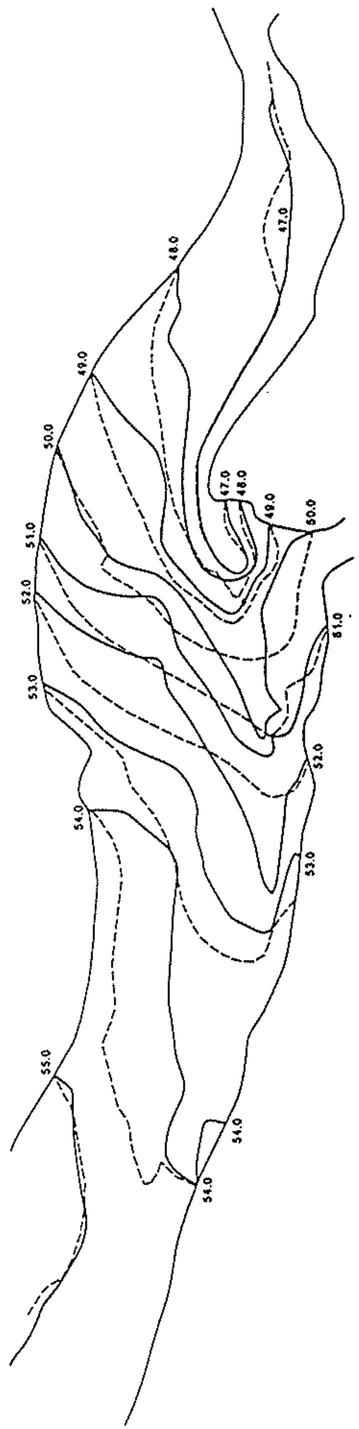
An automatic calibration procedure should be adopted (Samper et al [16]): the reference piezometric levels are recorded at different target points (Woessner [17]) and the calibration is realized automatically by successive adjustments of the K values until optimization of an "objective function" (comparing computed and measured values) is reached (Ouboter and De Rooij [18]). This kind of automatic calibration can be easily chosen for simple case studies but for a 3D regional case with complex geometrical conditions and constraints, it would be hazardous to consider as unique criterion, the minimization of a function depending only of the computed and measured values at the target points. For this reason, the only way to continue a classical trial and error calibration is to apply a kind of zonation approach (Keidser et al [19]) where the finite elements affected by K values comprised in defined ranges are grouped in the same "material" class of a determined K value.

In fact, more often, the kriging and cokriging techniques are applied in order to process conditional simulations (de Marsily [7]). Equiprobable realizations of the hydraulic conductivity field with the "conditions" corresponding to the measured values are simulated in the groundwater model. Each simulation relative to each realization of the K set of data provides an equiprobable answer (Henriques [20], Hoeksema and Clapp [21], Schafmeister and Pekdeger [22]).

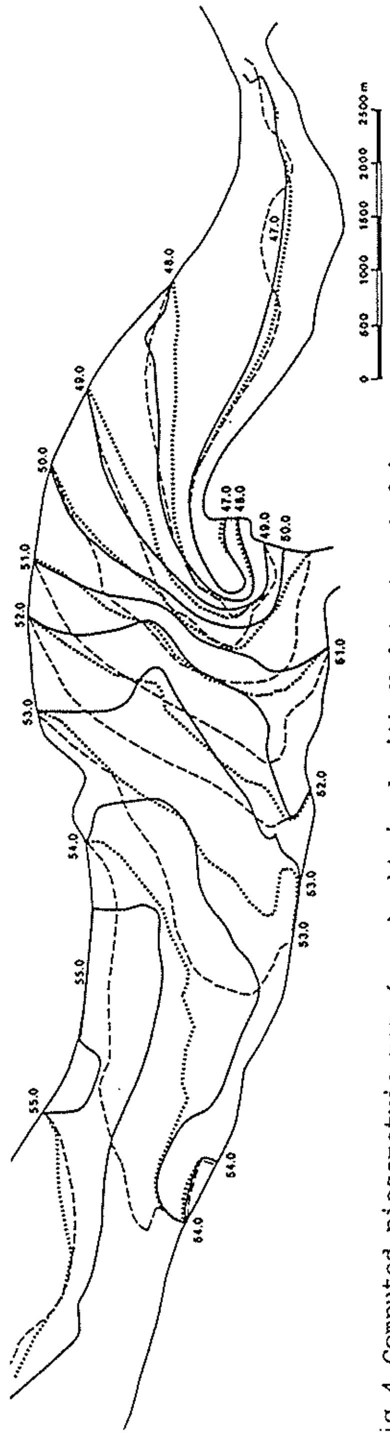
We have to observe that in the literature, these equiprobable results of simulations based on kriged data, are very rarely or never compared (until now) to a reference measured situation, chosen for calibration. In fact, kriging techniques and conditional simulations appear as to be used, in practice, in the cases where no calibration is possible by an evident lack of measured piezometric data.

ACKNOWLEDGEMENTS

A large part of this study has been supported by the Public Works Ministry of Belgium.



ig.3 Computed piezometric map (——) obtained with K data treated by method c, in regard to the reference piezometric map (----).



ig.4 Computed piezometric map (——) obtained with K data treated by method d, compared to the reference peizometric map (----) and to the computed map at the 11th run of the trial and error calibration (......).

REFERENCES

1. Delhomme, J.P., Applications de la théorie des variables régionalisées dans les sciences de l'eau, Thèse de Docteur-Ingénieur, Université Pierre et Marie Curie, Paris 6., 1976 .
2. Ahmed,S., and de Marsily,G., Comparison of geostatistical methods for estimating transmissivity using data on transmissivity and specific capacity, Water Resources Research, vol.23, n°5, pp.1717-1737 , 1987.
3. Neuman and Yakowitz, A statistical approach to the inverse problem of aquifer hydrology, 1, Theory, Water Resour. Research,15(4),845-860,1979.
4. Neuman S.P., G.E. Fogg and E.A. Jacobson, A statistical approach to the inverse problem of aquifer hydrology,2,case study,W.R.R.16(1),33-38,1980.
5. Dassargues, A., et Lox, A., Modélisation mathématique de la nappe alluviale de la Meuse en aval de Liège (Belgique),"Le système hydrologique dans la région frontalière Liège-Maasbracht; résultats des recherches 1985-1990", Rapport et notes n°26 , CHO-TNO, Delft, pp. 27-54, 1991.
6. Hoeksema, R.J., and Kitadinis, P.K.,Analysis of spatial structure of properties of selected aquifers,Water Resources Research, 21, 563-572, 1985.
7. de Marsily, G.,Quantitative Hydrogeology, Academic Press, 1986.
8. Bear, J., and Verruijt, A., Modeling groundwater flow and pollution , Reidel, 414 p, 1987.
9. Englund, E., and Sparks, A., GEO EAS , Geostatistical Environmental Assessment Software,User's guide,U.S. Environmental Protection Agency,Environmental Monitoring Systems Laboratory,Las Vegas,Nevada 89193-3478,1988.
10.Journel, A.G., and Huibregts, Ch.J., Mining geostatistics,Academic Press, 600p, 1978.
11.Myers, D.E., Cokriging : Methods and alternatives, in The Role of Data in Scientific Progress, edited by Glaeser, P.S., Elsevier, pp.425-428, 1985.
12.De Smedt, F., Multivariate geostatistics, ERASMUS course textbook, Laboratory of Hydrology, Vrije Universiteit Brussel, 31 p., 1991.
13.Charlier, R., Approche unifiée de quelques problèmes non-linéaires de mécanique des milieux continus par la méthode des éléments finis,Doctorat n° 109,Faculté des Sciences Appliquées, Université de Liège,Belgium,1987.
14.Dassargues, A., Paramétrisation et simulation des réservoirs souterrains, Discrétisation du domaine, Préparation statistiques des données, couplages et non linéarités des paramètres, PhD Thesis, Applied Sciences Faculty of the University of Liège,Belgium, 1991.
15.Dassargues, A., Radu, J.P.and Charlier, R. Finite element modelling of a large water table aquifer in transient conditions,Adv.in Water Resources, Vol. 11, June, pp. 58-66, 1988.
16.Samper, J., Carrera, J. Galarza, G., and Medina, A., Application of an automatic calibration technique to modelling an alluvial aquifer, ModelCARE 90 : Calibration and Reliability in Groundwater Modelling,The Hague, pp. 87-95, 1990.
17.Woessner,W.W.,Setting calibration targets and assessing model calibration -room for improvement:an example from North-America,ModelCARE 90:Calibration and Reliability in Groundwater Modelling, The Hague,pp.279-288, 1990.
18.Oubouter, M.R.L. and De Rooij, N.M., Use of an autocalibration to evaluate the consistency of a data set in relation to geochemichal model formulations, ModelCARE 90 : Calibration and Reliability in Groundwater Modelling, The Hague, pp.195-202, 1990.
19.Keidser, A.,Rosbjerg, D.,Jensen, K.H.,and Bitsch, K., A joint kriging and zonation approach to inverse groundwater modelling,ModelCARE 90: Calibration and Reliability in Groundwater Modelling,The Hague,pp. 171-183,1990.
20.Henriquez, A., Saving computer and engineering time in reservoir simulation, note prepared for International Workshop on Computer Applications in Reservoir Modeling, I.B.M. European Petroleum Application Center (E.P.A.C.), Stavanger, 1991.
21.Hoeksema, R.J., and Clapp, R.B., Calibration of groundwater flow models using Monte Carlo simulations and geostatistics,ModelCARE 90 :Calibration and Reliability in Groundwater Modelling, The Hague, pp. 33-42, 1990.
22.Schafmeister,M.Th.,and Pekdeger,A., Regionalization of hydraulic aquifer properties-Optimization by geostatistical simulation techniques,ModelCARE 90 : Calibration and Reliability in Groundwater Modelling, The Hague, pp. 447-455, 1990.