

Parametric identification of a karst aquifer

G. Chiaia & G. Ranieri

*Water Engineering and Chemistry Department, 1st Engineering Faculty,
Polytechnic of Bari, Italy*

Abstract

The parametric identification of non-homogeneous aquifers is a highly complex task because it requires the performance of a large number of infield trials, which often yield results that are difficult to interpret or have doubtful reliability. Moreover, the physically continuous nature of the hydrogeological parameters to be identified (typically conductivity) means that the experimental data need to be applied to the territory, a task which is often difficult in view of their great variability.

An alternative approach that has been studied in recent years is parametric identification by means of the solution of the Inverse Problem (I.P.). This method requires a generic hydrogeological knowledge of the aquifer, and the availability of a sufficient number of piezometric survey points on the area where the identification is to be made. The method is potentially able to provide detailed information on the aquifer characteristics, at decidedly competitive costs in terms of money and time as compared with those for the application of the direct method. However, its use is largely limited to the scientific field at present, and few technical applications have been described for the management of territorial scale aquifers.

This is because by its very nature, the inverse problem has the characteristic of an intrinsically ill-posed problem, in the sense that it frequently has no unique, stable solution.

Many researchers have investigated the causes of this poor formulation of the problem, recognizing the fundamental role played by the type of boundary conditions, the accuracy of reconstruction of the piezometric surface and the type and structure of the parametric method adopted.

In the present work, after making a preliminary overview of the theoretical reasons why the I.P. is ill-posed, the above aspects are analysed in the context of a real aquifer with an extension of approximately 100 km². After describing the method used to reconstruct the piezometric surface and individuating the boundary conditions, we illustrate the effects on the solution of the I.P. of different types of parametrization of conductivity.

Keywords: inverse solution, territorial scale aquifer, piezometric surface, mathematical model.



1 Introduction

The solution of the Inverse Problem (I.P.) aims to derive, through direct observation of the values of an easily measured dependent variable (potentiometry), the spatial distribution of the parameter it depends on, i.e. conductivity. It is taken for granted, therefore, that the operator can rely on a sufficient number of piezometric survey points inside the domain under investigation.

The parametric identification procedure is normally performed using indirect methods, applying history matching type criteria to estimate the parameter $T(x,y)$. A problem that arises already at this stage of the procedure is how to model the spatial variability of the parameter being identified. In fact, conductivity is a continuous function of the position and, for the purposes of mathematical modelling of the filtration process, a finite, discrete approximation has to be made.

The process of assigning a discrete value, which in practice allows the unknown quantities of the I.P. to be reduced to a finite number, is known as parameterisation. There are two classic ways of doing this: by zoning or by using geostatistical techniques.

With the former (e.g. [4,7,8]), starting from knowledge of its geology and hydrogeology, the area under study is subdivided into a given number Z of zones, in each of which the parameter T is hypothesized to be constant, although it may vary from zone to zone.

Instead, the geostatistical technique (e.g. [12,20]) considers the parameter to be identified as a random stationary field which undergoes continuous variations in space. A series of nodes is individuated in the area of interest, each of which is associated with a basic local function. A third possible approach to parameterisation is the “hybrid” method proposed by McLaughlin [14] in which, by using a functional analysis method, the unknown value of conductivity is treated as a scale function of the position rather than as a vector.

The vector of the piezometric levels calculated is a function \mathcal{I} of the distribution T' of transmissivity.

After individuating the parameterisation criterion best suited to the case under study, this parameterisation dimension is set at Z (where Z is equal to the number of zones with the zoning technique, or to the number of nodes with the geostatistical approach), the distribution of trial T' is constituted by a vector with Z components. If \mathbf{h}_r is the vector of the identified piezometry, an “indirect” expression of the discrepancy between the real distribution of transmissivity T and the estimated value T' , is:

$$\text{Ob}(T') = [\mathbf{h}_r - F(T')]^T * \mathbf{W}_p * [\mathbf{h}_r - F(T')] \quad (1)$$

which represents the mean square difference between the calculated levels and the measured levels (the latter being weighted, if necessary, by the matrix \mathbf{W}_p). The solution to the Inverse Problem is obtained by minimizing the value (1), that takes on the significance of the objective function (O.F.).

The greatest difficulty in applying this procedure is that of identifying a suitable algorithm that can minimize the O.F. by introducing suitable variations in the components of the vector T' .

This problem can be dealt with by various approaches that differ as regards the optical focus of the mathematical tools required. An extensive list of references on this topic is available in [2,3,18].

The I.P. is often ill posed; that is it hasn't a unique solution or/and the solution is not stable [16].

As shown by several authors [13–15], there is also an intrinsic instability linked to the structure of the equations governing the “direct” problem.

A big role in make the problem well posed is played by:

- The type of boundary conditions imposed
- The level (qualitative and quantitative) of knowledge of the piezometric surface
- The type of parametrization adopted.

2 Experimental processing

The case study was a karst aquifer, on which, the following aspects were analysed at the territorial scale:

- Surveying and processing the piezometric data;
- Individuating the boundary conditions;
- Parameterisation of the unknown conductivity value;

2.1 Description and diagram of the experimental site

The case study area, having an extension of 110 km², is the area shown inside the grid in figure 1.

Drinking water needs are partly satisfied by mains water lines supplied by sources outside the Apulian region and partly by pumping of the groundwater at a virtually constant rate of 315 l/s throughout the year, from the artesian wells shown in figure 1 marked as AQP.

The domain is shown in the figure as a diagram with the elements represented on a grid, deliberately built with the vertexes coinciding with points with a known piezometric value, and lying in an orthogonal direction with respect to the coastline.

In view of the large size of the area under study, modelling of the water flow was done according to the porous equivalent approach.

The piezometric survey points were obtained via ten measurement cycles on 44 sentinel artesian wells over a period of two solar years.

Subsequent processing [6] enabled us to individuate 34 points with a high degree of reliability, which were then used to calculate the piezometric surface of the aquifer in the different periods of the year.

2.2 Analysis and processing of the piezometric survey points

As already mentioned above, apart from the drinking water wells, whose position and reserves are known, there are a high number of points for irrigation which pump unknown quantities in a discontinuous manner.



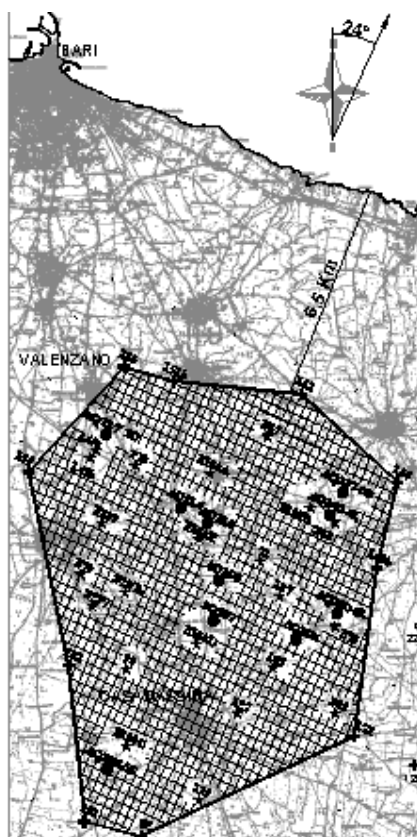


Figure 1.

This constitutes a potential cause of “ill posing” of the inverse problem because, in general, the piezometric points identified are not a unique function of the hydrogeological characteristics of the aquifer, which also depend on other factors which are difficult to measure.

To overcome this shortcoming, the parametric identification procedure was carried out on the basis of the piezometric data obtained in the month of March (figure 2), when the effects on the shape of the piezometric surface of withdrawal for irrigation can be presumed to be over.

As can be seen, the availability of only discrete, discontinuous piezometric survey points often causes ill posing of the inverse problem, due to the incorrect definition of the denominator in (4). It is therefore essential to carry out a series of processes, described in detail in a previous work [6], that yield a unique answer to the calculation of the piezometric surface in different periods of the year.

The above processing yielded the piezometric values of all the nodes in the grid in figure 2.

that shows a nugget type trend which is confirmed even in cases using different lags, demonstrates the poor spatial correlation of the parameter analysed, imputable to the high degree of anisotropy of the aquifer.

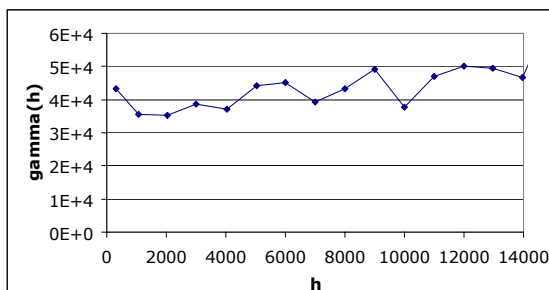


Figure 3: Semivariogram of the specific charges.

This circumstance meant that to assess the aquifer we had to give up the idea of using geostatic type parameterization and to opt for the zoning technique.

It is essential to individuate the number and shape of the zones, or “structure” of the zoning, because this has a strong influence on the correct posing of the inverse problem and the physical plausibility of the solution [14].

In the case of aquifers extending over a wide area, a “direct” approach to individuating the structure of the zoning requires the execution of too high a number of infield trials to be technically and economically feasible. Nevertheless, in absence of the effects of withdrawal for irrigation and local replenishment, and at a suitable distance from drinking water pumping points, the spatial variations in the piezometric surface are largely linked to variations in conductivity.

An “indirect” approach to individuating the zoning type to be adopted can therefore be based on an analysis of the piezometric gradient, following the procedure:

- Calculate the value of the experimental gradient in all the nodes of the integration grid in figure 2
- Individuate the gradient range and subdivisions, according to the criteria illustrated below, at a suitable number of intervals
- Establish the zones by grouping together the elements of the calculation grid with gradient values falling within the same interval.

Figure 4 depicts the qualitative trend of the piezometric gradient in the integration domain.

The zone including the boundary wells 162, 166 and 168b has been excluded because their gradient values are notably lower than the mean, and indeed tend towards zero.

This zone, that probably has special hydrogeological characteristics that for economic reasons are beyond the scope of investigation of this study, has thus been excluded from all the following parametric identification processes.

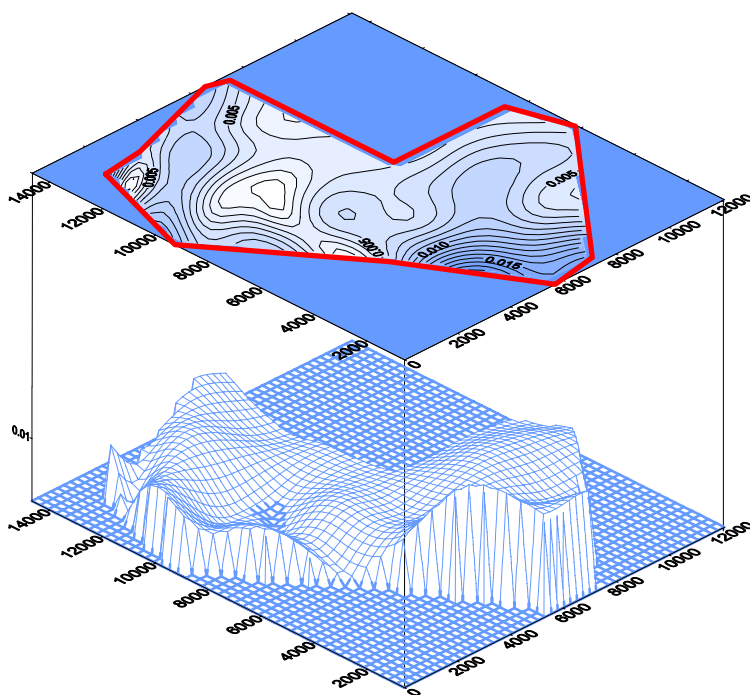


Figure 4: Piezometric gradient of the area under study.

The above procedure yields different zoning parts according to which steps are taken to subdivide the experimental gradient and to the relative criteria adopted to decide this.

With reference to the latter aspect, in the present work three different criteria were used to form the zones:

Zoning type n. 1: ordering the elements by decreasing values of the piezometric gradient and subsequent regrouping to make up zones all having the same number of elements.

Zoning type n. 2: subdividing the range of the logarithm of the piezometric gradient into an equal number of parts and grouping together all elements with gradient values within the same interval (variable elements zoning)

Zoning type n. 3: like 2 but referring to the gradient values not their logarithms.

The parameterization size, i.e. the resulting number of zones, depends on the steps adopted to subdivide the gradient range. In accordance with the indications reported by Yhe et al. [20], increasing order zoning (from 1 to 9), was created, individuating the optimal configuration by means of the sensitivity analysis reported below.

Figure 5 shows, as an example, a comparison of the different zoning types for the same parameterisation size (5 zones).

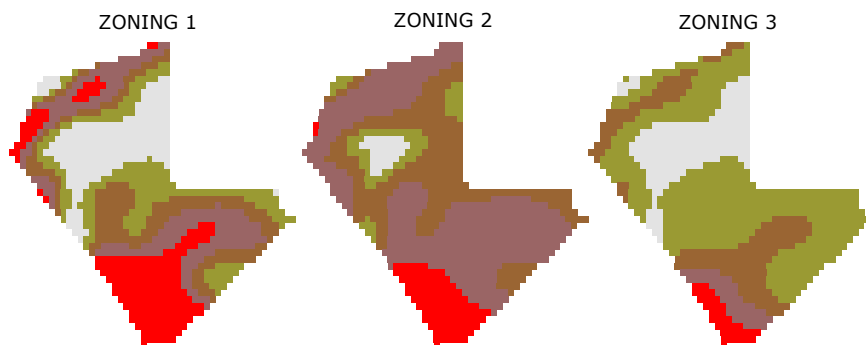


Figure 5: Comparison of zoning types according to different criteria for subdividing the piezometric gradient range. The parameterization size in the case under study is equal to 5.

2.4 Algorithm for the solution of the inverse problem

In the present work, particular attention was paid to the methodology adopted to solve the inverse problem, with the aim of individuating simple, repeatable procedures that could enable technicians responsible for managing aquifers to apply these parametric identification measures to real aquifers.

Processing was done with SGH_D2 software [13], available on the WWW. It is very simple to use and so for technical applications it is preferable to other codes, such as *sutra-1* or *modflow*, which are certainly more refined but are better employed for scientific purposes.

The algorithm used to minimize the objective function is the downhill-simplex, which is particularly efficacious when there are only a small number of parameters to be optimised, as in the present case.

The inverse problem was solved by using the finite elements scheme shown in figure 1, which has a rectangular grid where $Dx = 241\text{m}$ and $Dy = 305\text{m}$, and adopting the three different zoning techniques described in the previous section, operating with parameterization sizes ranging from 1 (homogeneous aquifer) to 9.

The software used limits the piezometric control points that can be used to build the lens function to 30. These were selected from the 1295 nodes on the grid depicted in figure 2, according to the criterion of equidistribution.

2.5 Definition of the boundary conditions

The piezometric levels observed in the wells on the boundary of the integration domain (see figure 2) were used to define the Dirichlet boundary conditions.

It was a more complex task to individuate the Neumann conditions, owing to the need to quantify the water flow in and out of the integration domain.

To calculate this it was necessary to discover the hydrological balance in the territorial layer immediately above the study area, shown by the crossing dashed lines in figure 6. This layer is laterally bounded by the subterranean watershed of

the aquifer and in a longitudinal direction, lies parallel to the mean direction of the groundwater flow. Being 20 km wide, it is representative of the mean hydrogeological conditions in the territory.

The aquifer recharging process can be roughly deduced from the following expression:

$$N = P - (E + R + A) \quad (2)$$

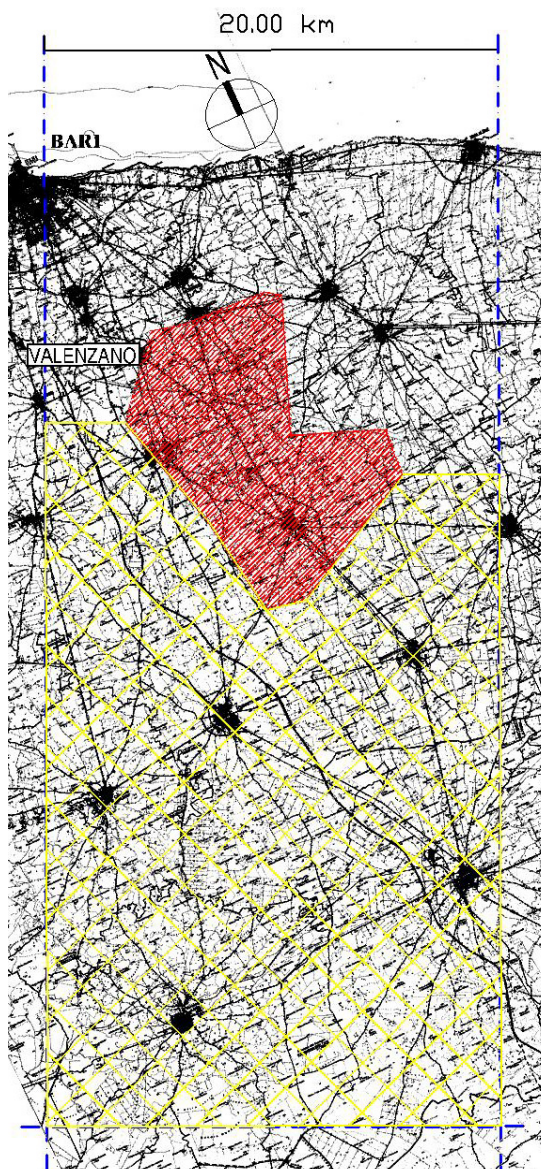


Figure 6: Recharging area of the domain under study.



where: N = recharging of the groundwater; P = rainwater height; E = true evapo-transpiration; R = streaming; A = variation of the water capacity over the terrain. Troisi and Vurro [17] obtained the value of (2) (in mm/year) for each hydrographic-thermometric station of the Murgia basin, calculated by adopting two distinct procedures for assessing the evapo-transpiration terms.

The area under study belongs to the territory served by the stations of Altamura, Masseria Mercadante, Cassano Murge, Casamassima, Turi and Gioia del Colle. The annual recharge load was obtained by reconstructing the Thiessen Polygons for each station, individuating the portions falling within the zone of interest, and multiplying the surface by the relative recharge value. From this sum was then subtracted the total quantity of drinking water withdrawn by the Acquedotto Pugliese from 6 wells situated in the countryside of Acquaviva and Gioia del Colle, which extract a total volume of 75.5 l/s, equal to 2.35 Mm³/year.

This yielded the net annual average recharge volume in the layer above the integration domain. Divided by the width of 20 km, this made it possible to calculate a specific daily charge of q_I [m³/day*m], entering through the uphill boundary of the domain.

The inflow charge in each of the 50 cells making up the uphill boundary of the integration domain was calculated by multiplying this value by their transverse width, equal to 241 m. The overall mean inflow charge in the domain Q_I is, of course, equal to $q_{Ic} \times 50$.

The annual mean outflow volume from the integration domain (Q_O) was calculated by adding the inflow volume ($Q_I \times 360$) to the innate recharge value of the domain (having a surface of 89.17 km² all falling within the Thiessen Polygon of Casamassima) and subtracting the drinking water extracted, equal to 180.8 l/s = 5.62 Mm³/year. The discharge q_{oc} from each of the 50 cells on the downhill boundary below the domain is equal to $Q_O / (360 \times 50)$.

A knowledge of the charge crossing the area under study and of the underlying mean gradient allowed us to estimate the equivalent conductivity value of the aquifer as equal to 1670 m²/day: this was used as the initial value in the parametric identification process.

2.6 Results of the parametric identification

Figure 7 shows the results of the parametric identification made with the three different zoning criteria, varying the parameterization size.

Firstly, it can be seen that the conductivity value obtained with the homogeneous aquifer hypothesis (parameterization size = 1), equal to 1580 m²/day, is very close to the mean value obtained above.

Referring to zoning 1, it is clear that the different parameterization sizes (at 1, 2, ..., 9 zones) yield conductivity values with a mean increase proportional to the number of zones: this confirms the hypothesis underlying the zoning criterion adopted, which postulates a direct link between conductivity and the piezometric gradient value, which increases with the number of zones.

Again referring to zoning 1, it can be seen that while the parameterization size varies, the conductivity values remain within a physically plausible range (500–6200 m²/g), compatible with the hydrogeological features of the aquifer.



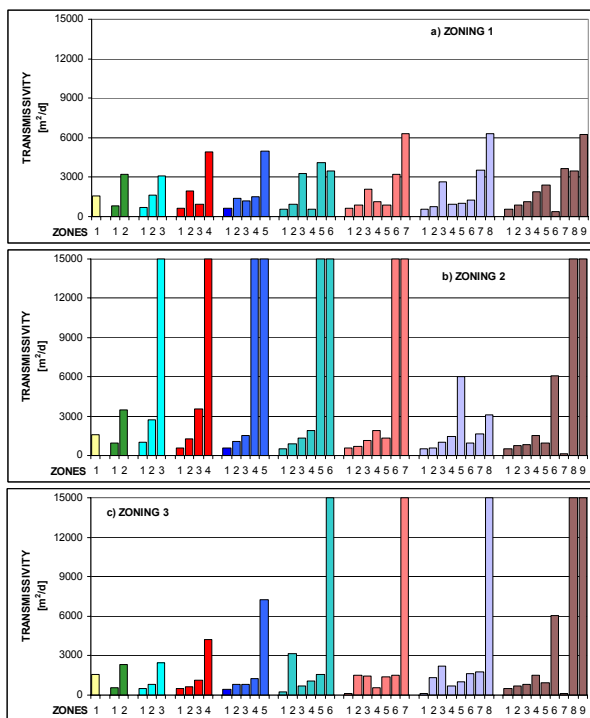


Figure 7: Results of parametric identification with the three zoning criteria at changing parameterization sizes.

Vice-versa, the structure of zoning types 2 and 3 is poorly suited, in the case under study, to modelling the spatial variability of the conductivity parameter: this is demonstrated by the instability of the solution of the inverse problem, which gives rise, in agreement with reports by other authors [18], to values at the upper limit of the variability range of the parameter, set at $15000 \text{ m}^2/\text{g}$.

This instability is particularly evident in the case of adoption of zoning type 2, where excessive values are present even at low parameter sizes, while with zoning type 3 excessive values appear when the parameter size is 6 or more.

The zoning geometry that gives rise earliest to instability of the solution to the inverse problem is shown in figure 8. It is evident that in zoning 3, where the instability occurs, there is a small number of elements (33 elements, as compared with 626 elements in zoning 1 and 406 elements in zoning 3) and that in this zoning there is no defined boundary condition of the flows.

Detailed analysis of the other zoning types generating instability of the I.P. solution confirmed that the phenomenon arises in the case of zones with few elements.

The transmissivity value obtained for a given cell in the domain according to the different zoning type and parameterization size is shown in figure 9. It can

be seen that before the above-described instability sets in, the transmissivity values obtained with the different zoning types are fairly similar, demonstrating a certain “robustness” of the algorithm used for the calculation.

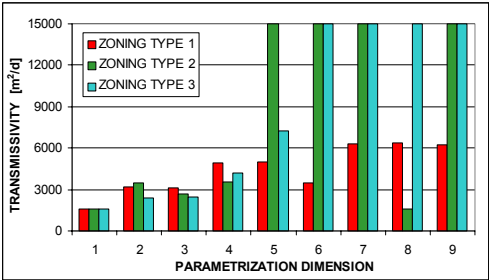


Figure 8: Transmissivity values for the some cell varying parameter dimension.

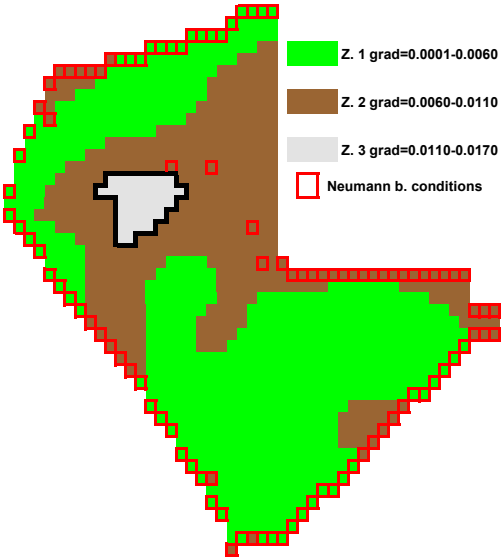


Figure 9: Zoning geometry that gives rise earliest to instability in the I.P. solution.

2.7 Choice of the solution

The results of the processing described in the above section showed that zoning 1 is best able to describe the aquifer, as it can support high parameterization sizes without developing numerical instability.

The solution to the I.P. must therefore be found among the 9 possible configurations depicted in figure 7a, characterized by different parameterization sizes.

In view of the intrinsic ill posing of the I.P. frequently referred in this work, the choice cannot be based on deterministic criteria, but must rely on the concept of optimisation, which yields not “the solution”, but “the best among the available solutions”.

A classic methodology was proposed by Yhe and Yoon [19], based on comparison of the errors associated with modelling of the aquifer and with the uncertainty of the parameterisation.

The former, as is well known, is the objective function (2); the latter, which affects the stability of the solution, is the norm of the matrix of covariance of conductivity. This can be calculated by the relation:

$$\text{Cov}(\mathbf{T}^*) = \frac{J(\mathbf{T}^*)}{M-L} \cdot [\mathbf{J}_D^T \cdot \mathbf{J}_D]^{-1} \quad (3)$$

where \mathbf{T}^* is obtained by perturbing solution \mathbf{T} of the inverse problem with the suitable sized noise [1], $J(\mathbf{T}^*)$ is the value of the objective function, \mathbf{J}_D indicates the Jacobean matrix, \mathbf{J}_D^T its transposition, and M and L are the piezometric survey points (30 in the case under study) and the parameterisation size (ranging from 1 to 9), respectively.

The size of the matrix (3) ($L \times L$), and hence the numerical value of its norm, depend on the parameterization size adopted, which hampers comparison between the different configurations.

To circumvent this shortcoming the authors [19] suggest that parameterisation should be adopted, transforming the conductivity vector \mathbf{T}^* (that in (3) has size L), into a new vector of size N , and N number of elements of the calculation grid.

For this purpose a transformation matrix \mathbf{G} , is introduced, of size $N \times L$, whose lines are all null elements except one, which is equal to the unit. This matrix associates each of the N elements of the domain with one of the conductivity L deriving from the parameterisation considered. In this way the new vector \mathbf{T}_e is obtained, for each element in the domain:

$$\mathbf{T}_e = \mathbf{G} \cdot \mathbf{T}^*$$

So (3) is changed to:

$$\text{Cov}(\mathbf{T}_e) = \frac{J(\mathbf{T}^*)}{M-L} \cdot \mathbf{G} \cdot [\mathbf{J}_D^T \cdot \mathbf{J}_D]^{-1} \cdot \mathbf{G}^T \quad (3')$$

The objective function shows a marked decreasing trend for parameterisation sizes between 1 and 4, while subsequent increases up to value 8 do not substantially improve the precision of the modelling but do notably increase the uncertainty of the parameterisation.

Combined analysis of the two curves in figure 10 shows that 5 is the ideal parameterisation size. In fact, at this configuration both the objective function and the norm associated with the matrix of covariance of conductivity have optimal values.

The optimal solution of the inverse problem for the domain under study is shown in figure 11. The mean value of conductivity, that can be calculated as the simple arithmetical mean of the five conductivities thanks to the equal number of elements making up the different zones, is equal to approximately 1900 m²/day, very close to the previously obtained values. It can be seen that the distribution

of conductivity is coherent with the indications deriving from the shape of the piezometric surface. The presence of a low conductivity zone uphill justifies the high gradient values found in this area (figure 5), while the high value of conductivity T5 explains the presence, in the central zone of the domain, of a layer with very minor gradient values.

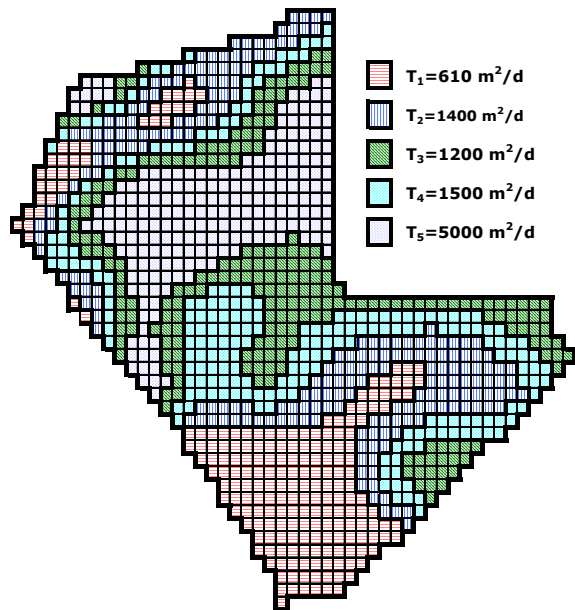


Figure 10: Optimal solution of the Inverse Problem.

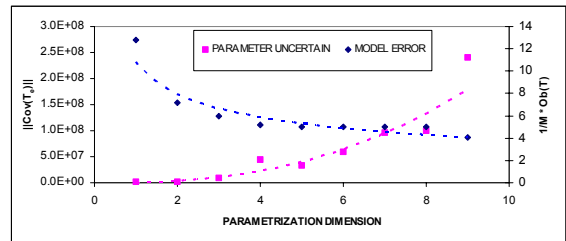


Figure 11: Error of the model and parameterization uncertainty.

3 Conclusions

To apply parametric identification methods based on the solution of the inverse problem to a real aquifer, it is firstly necessary to gain a knowledge of a series of essential elements that are needed to limit the effects of the intrinsic “ill posing” of the I.P.



In particular, theoretical analyses have demonstrated that correct posing of the problem depends very strongly on the level of knowledge of the aquifer piezometry, of the nature of the boundary conditions and of the parameterisation type adopted to assign discrete conductivity values.

In-depth study of these aspects has been made in the present work, focusing on a real aquifer with a planimetric extension in the order of 100 km².

The methods used to interpolate the piezometric survey points and to make a preventive hydrogeological assessment of the territory have been analysed in detail. The results of this analysis indicate that zoning is the parameterisation method best suited to the aquifer.

The effects of this method on the solution of the I.P. have been investigated by comparing different zoning techniques, all based on an “a priori” knowledge of the shape of the piezometric surface. Zoning techniques involving the aggregation of a variable number of elements gave rise to an unstable solution of the I.P. and were therefore rejected.

To pose the I.P. correctly, boundary flow conditions need to be imposed, and hence the hydrogeological balance of the area under study had to be identified.

The optimal solution of the I.P. was obtained by means of the sensitivity analysis proposed by Yhe et al. [20].

References

- [1] Bard Y., 1974. Nonlinear parameter estimation. Ed. John Wiley – New York.
- [2] Butera I., Grella E., Maione U., Tanda M.G., 1988. Applicazione comparativa di tre metodi di risoluzione del problema inverso in un acquifero eterogeneo bidimensionale. *Atti XXVI Convegno di Idraulica e Costruzioni Idrauliche*. Catania 9-12 settembre 1998.
- [3] Binley A., 1993. Groundwater model parameter identification. *Basic Lectures of European Training Course on Parameter estimation in groundwater flow and pollution modelling*. Ed. Bios - Cosenza (ITALY).
- [4] Carrera J., Neuman S.P., 1986. Estimation of aquifer parameters under transient and steady state conditions. (1-2-3). *Water Resource Research*, Vol 22, n. 2 pp. 199-242.
- [5] Chiaia G., Di Santo A., Ranieri M. 1999. Interpretazione di misure sperimentali della falda pugliese. *Giornata Mondiale dell'Acqua – Roma*.
- [6] Chiaia G., Ranieri M. 2004. Elaborazione di rilievi piezometrici finalizzate alla identificazione parametrica di un acquifero a scala territoriale. *L'acqua* n. 4 pp. 29-40.
- [7] Cooley R.L., 1982. Incorporation of prior information on parameters into non linear regression groundwater flow models. 1, Theory. *Water Resource Research*, Vol 18, n. 4 pp. 965-976.
- [8] Cooley R.L., 1983. Incorporation of prior information on parameters into non linear regression groundwater flow models. 2, Applications. *Water Resource Research*, Vol 19, n. 3 pp. 662-676.



- [9] Desbarats A.J., Logan C.E., Hinton M.J., Sharpe D.R., 2002 On the kriging of water table elevations using collateral information from digital elevation model. *Journal of Hydrology Vol. 255* pagg. 25-38.
- [10] Gambolati G., Volpi G., 1979 Groundwater Contour Mapping in Venice by Stochastic interpolators 1. Theory. *Water Resource Research, Vol 15, n. 2 pp. 281-290.*
- [11] Goovaerts P., 1997 Geostatistics for Natural Resource Evaluation. *Oxford University Press-New York.*
- [12] Hoeksema R.J., Kitandis P.K., 1984. An application of the geostatistical approach to the inverse problem in two-dimensional groundwater modeling. *Water Resource Research, Vol 20, n. 7 pp. 1003-1020.*
- [13] Karovent T. 2001. Subsurface and groundwater hydrology: basic theory and application of computational methods. *Pubblicazione elettronica.*
- [14] McLaughlin D., Townley L.R., 1996. A reassessment of the groundwater inverse problem. *Water Resource Research, Vol 32, n. 5 pp. 1131-1161.*
- [15] Sun N.Z., 1994. Inverse problems in groundwater modelling. *Kluwer Academic Publishers - London.*
- [16] Tikhonov A.N., Arsenin V.Y., 1977. Solutions of ill-posed problems. *Halsted Press/Wiley – New York.*
- [17] Troisi S., Vurro M., 1987. Bilancio idrologico per la valutazione della ricarica naturale: una applicazione ai bacini delle Murge e del Salento – *Idrotecnica, n.2, pp.47-64.*
- [18] Yeh W. W-H., 1986. Review of parameter identification procedures in groundwater hydrology: the Inverse Problem. *Water Resource Research, Vol 22, n. 2 pp. 95-108.*
- [19] Yeh W. W-H., Yoon Y.S., 1981. Aquifer parameter identification with optimum dimension in parameterisation. *Water Resource Research, Vol 17, n. 3 pp. 664-672.*
- [20] Yeh W. W-H., Yoon Y.S., Lee K.S., 1983. Aquifer parameter identification with Kriging and optimum parameterisation *Water Resources Research Vol. 19 n.1 pp. 225-233.*
- [21] Yakowitz, Noren, 1976. On the identification of inhomogeneous parameters in dynamic linear partial differential equations. *Journal Math. An. Appli. 53.*

