

# **A synthetic study of the worth of transmissivity and head data to delineate groundwater protection zones in braided alluvial aquifers.**

**J. KERROU & P. RENARD**

*Centre d'Hydrogéologie, Université de Neuchâtel,  
Rue Emile Argand 11, CH-2007 Neuchâtel, Switzerland  
e-mail: [jaouher.kerrou@unine.ch](mailto:jaouher.kerrou@unine.ch)*

**I. LUNATI**

*Institute of Fluid Dynamics, Swiss Federal Institute of Technology (ETH),  
ETH-Zentrum, Sonneggstr 3, CH-8092 Zurich, Switzerland*

**H.-J. HENDRICKS-FRANSSEN**

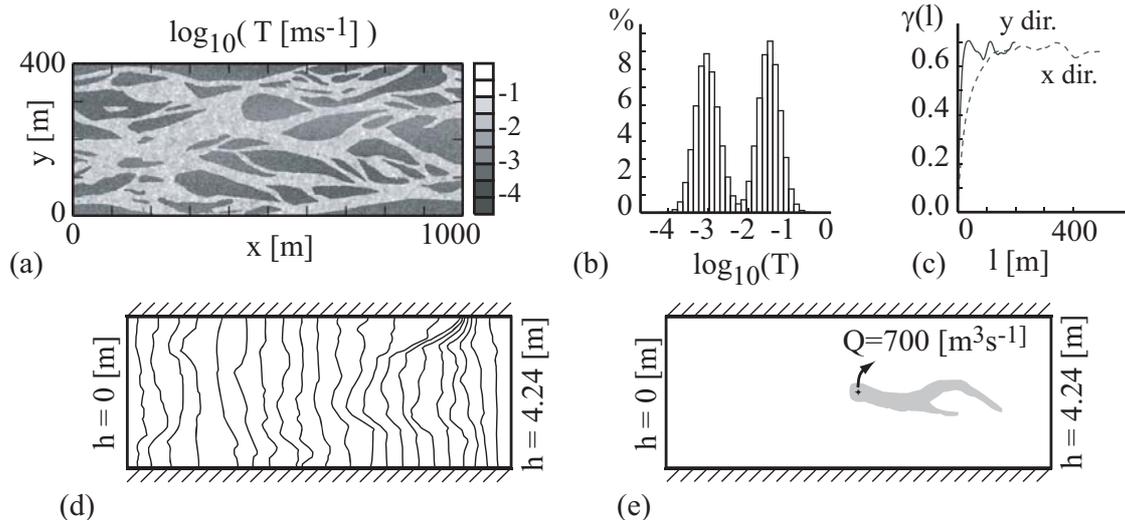
*Institute of Hydromechanics and Water Resources Management, Swiss Federal Institute of  
Technology (ETH), ETH-Hönggerberg, CH-8093 Zurich, Switzerland*

**Abstract** Kriging, stochastic conditional simulations, and stochastic inversion are compared to characterize a synthetic hypothetical braided alluvial aquifer. The performance of the three techniques is compared in terms of reproduction of a 10-days capture zone of a pumping well. The results obtained compare fairly well with previous investigations dealing with the stochastic delineation of capture zones however the presence of a large number of deterministic channels in the reference transmissivity field reduces significantly the accuracy of the three techniques as compared to previous published results.

**Key words** kriging, stochastic simulations, uncertainty, inverse, capture zones

## **INTRODUCTION**

In Switzerland, alluvial aquifers constitute the most exploited freshwater resource. At the same time, these aquifers are extremely vulnerable because they are shallow, and located in zones that are either highly urbanized or intensively used for agriculture. Furthermore, the complexity of the sedimentary processes in such regions, located very close to the Alpine arc, creates a high degree of heterogeneity leading to a large uncertainty in solute transport predictions. Finally, because of the scarcity of arable land and the subsequent economical importance of alluvial plains, the law imposes to delineate protection zones of smaller size compared to other European countries. Within this context, the present study aims at evaluating numerical techniques that can be used to characterize the aquifer in order to delineate capture zones and estimate the uncertainty of the prediction. Three approaches are considered: two *direct approaches* based only on transmissivity data (kriging and conditional simulations), and one *inverse approach* including head data. The advantage of including head data in the inverse approach has already been shown (Wen *et al*, 1996; Bakr & Butler, 2004; Stauffer *et al*, 2004). In this study, we investigate the efficiency of these techniques by considering a synthetic braided alluvial aquifer that includes some realistic deterministic structures (channels) and bimodal random fluctuations of the transmissivity field. In order to simulate the typical approach followed in practice, the study is divided in two major steps: first a characterization step to reconstruct the



**Fig. 1** (a) The synthetic transmissivity field. (b) Histogram of the log decimal of the transmissivities. (c)  $x$  and  $y$  directional variograms. (d) Head fields and boundary conditions for the uniform flow situation. (e) 10 days capture zone under the radial flow conditions.

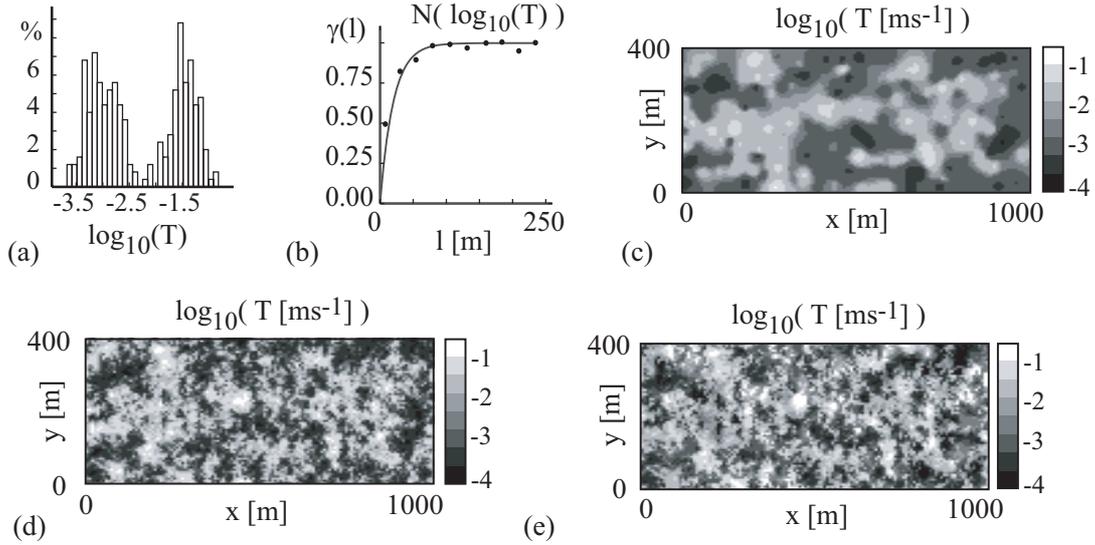
transmissivity fields from punctual transmissivity and/or head values sampled on the synthetic reality, then a forecasting step to predict the 10-days protection zone around a future pumping well. Like in real cases, the flow conditions of the forecasting step are different (new pumping well) from the flow conditions of the characterization step.

## THE VIRTUAL REALITY

The research is conducted on a synthetic and hypothetical reality (Fig. 1a-c). The medium consists of channels and lenses displaying internal heterogeneity. The reference transmissivity ( $T$ ) field is built from an air photograph displaying braided channels in the Ohau River, New Zealand (Mosley, 1982). This image was digitized and used at its real scale so that the size of the channels and lenses is realistic. The image size is 1km long and 400m wide with a resolution of 1m. Two multi-Gaussian unconditional simulations were generated to populate separately the channels and the lenses with  $T$  values. The first simulation describes the  $T$  distribution in the channels; it has an exponential variogram with a short correlation range (3 m). The second simulation describes the lenses. It has a nested variogram including one isotropic exponential model with a 3 m range, plus a cubic anisotropic model with a long range in the  $x$  direction (600 m) and a smaller one in the  $y$  direction (300 m). The overall integral scale has been estimated in the  $x$  and  $y$  direction ( $i_x = 25.7$  m,  $i_y = 7.4$  m) by numerically integrating the correlation function calculated directly on the reference  $T$  field.

On the reference transmissivity field, a uniform steady-state flow (Fig. 1d) is computed with the *fflow* code by imposing two constant head values on the east and west boundaries, whereas the north and south limits are no-flow boundaries. This flow field is used to sample head data for the inverse approach in the characterization step.

In the forecasting step, the flow field has the same boundary conditions plus an additional constant-rate pumping well located in the middle of the domain. The 10-



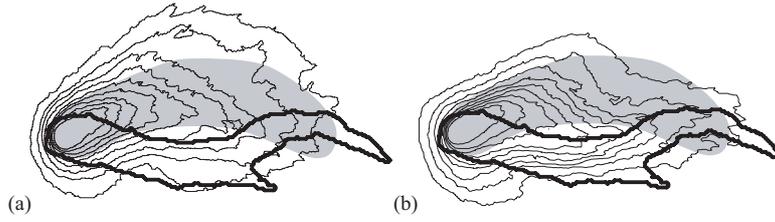
**Fig. 2** The characterization step for 250  $T$  data. (a) Histogram of the 250  $T$  data, (b) Experimental and model variograms of the Gaussian transform of the decimal log of the  $T$  data. (c) Kriging. (d) One simulation conditioned only to the 250  $T$  values. (e) The same simulation conditioned to the 250  $T$  and 1000  $h$  measurements.

days capture zone of this well is calculated by solving the Kolmogorov backward equation (Uffink, 1989). The mean life expectancy is calculated with the *gwad* finite element code (Cornaton, 2004). The 10 days capture zone is defined as the region where these values are lower than 10-days (Fig. 1e). The longitudinal and transversal dispersivities are very small (2 and 0.2 m), leading to a narrow dispersion of the life expectancy values around the mean.

## SIMULATING THE CHARACTERIZATION AND FORECAST STEPS

To simulate the aquifer characterization, the reference transmissivity field is sampled at random locations within the cells of a regular grid. We obtain three data sets with 21, 250, and 1000  $T$  measurements. The mean distance between the samples is respectively 89, 26, and 12 meters. If we normalize these distances by the integral scale in the  $x$  direction, we obtain a mean dimensionless distance between the samples  $d=3.5$ , 1, and 0.5 respectively. In other words, the three data sets represent measurements whose mean spacing is 3 times the integral scale in the  $x$  direction, on the order of the integral scale, and half of the integral scale, respectively. Note that the reference image is anisotropic and the integral scale in the  $y$  direction is much smaller, even smaller than the mean distance between the samples in the 1000  $T$  measurements case. The same sampling method is applied for the piezometric heads. Sampling locations are the same location as for the transmissivities.

Each transmissivity data set is intended to mimic a set of experimental data and it is analyzed separately. The  $T$  values are transformed into a normal variable  $N$  via a Gaussian anamorphosis of the decimal log of  $T$ . An experimental variogram is calculated and a model fitted to this data. The variogram model is used for the three characterization techniques. As a first direct approach, kriging of the  $N$  values is used. This map is transformed back to get an estimation of the  $T$  field (Fig. 2c). As a second direct approach, 100 conditional stochastic simulations of  $N$  are generated and



**Fig. 3** Examples of the probability maps calculated in the case of 21 T measurements. The thick line represents the reference. The grey area, the zone forecasted by kriging. The thin lines represent the isoprobability contours for every value between 0.1 to 0.9 with a step of 0.1 when using (a) 100 conditional simulations only with the 21 T, (c) 100 conditional simulations with the 21 T and the 1000 head measurements.

transformed back into  $T$  (Fig. 2d). As a third approach, the previous stochastic simulations are conditioned in addition by the head data (Fig. 2e) with the *inverto* code (Hendricks Franssen, 2001). In this case, for each ensemble of stochastic simulations corresponding to a given  $T$  data set, the conditioning on heads ( $h$ ) is made successively for all the  $h$  data sets that contain the same number of head data than the  $T$  data or more. At the end of the process, we have 3 kriged transmissivity fields, 300 simulations conditioned only on  $T$ , and 600 simulations conditioned on  $T$  and heads. For each of these transmissivity fields, the 10-days capture zone is calculated with the same approach that was used for the reference field.

## RESULTS

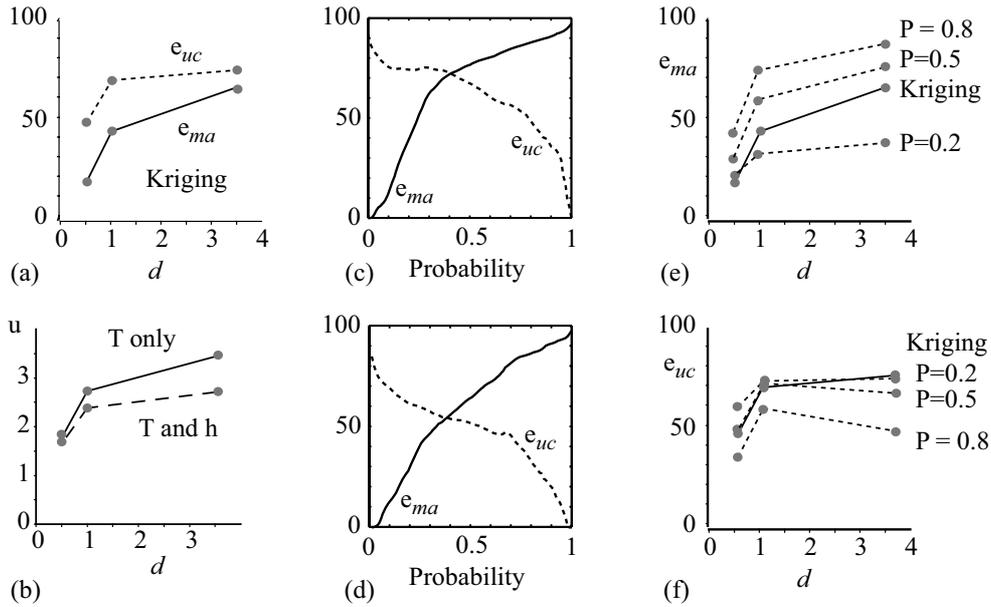
The first striking result when looking at the transmissivity maps and capture zones is that the forecasts differ significantly from the reference (Fig. 1, 2 and 3). To quantify this discrepancy, we introduce two error norms that allow to compare one forecasted capture zone with the reference:

$$e_{ma} = \frac{N}{B+N} = \text{missed area} \quad e_{uc} = \frac{P}{B+P} = \text{useless cost} \quad (1)$$

where  $N$  [ m<sup>2</sup> ] is the area of the protection zone that is not correctly identified by the forecast,  $B$  [ m<sup>2</sup> ] is the area of the reference protection zone that is correctly forecasted,  $P$  [ m<sup>2</sup> ] is the area wrongly forecasted as belonging to the protection zone. In other words,  $e_{ma}$  is the percentage of the reference that has not been identified.  $e_{uc}$  is the percentage of the forecast that is unnecessarily protected. In the direct conditional-simulation approach and in the inverse approach, several simulations have been constructed for a given data set. In these two cases, the uncertainty is estimated by constructing the maps of the probability for a point to belong to the capture zone. Examples of these maps are shown in Fig. 3. In addition to the two error norms presented above, we propose to quantify the uncertainty with the following dimensionless number:

$$u = \frac{I}{N+B} = \text{uncertainty} \quad (2)$$

where  $N+B$  [ m<sup>2</sup> ] is the area of the reference protection zone, and  $I$  [ m<sup>2</sup> ] is the area located between the 0.9 and 0.1 isoprobability contours. The missed area  $e_{ma}$  and



**Fig. 4** Synthesis of the results. (a) Error norms versus mean distance between the samples. (b) Uncertainty versus mean distance between the samples. (c) Error norms versus probability threshold for the simulations with 21 T values. (d) Same graph but 1000 additional head measurements used to condition the T fields. (e,f) Error norms versus  $d$  for different probability thresholds. (See comments in the text)

useless cost  $e_{uc}$  errors are calculated as well for the different levels of probability. Figure 4 summarizes the results of our calculations for the three approaches considered.

When the direct approach with Kriging is employed, we find that the error decreases as expected when we add  $T$  data (Fig. 4a). The reduction becomes stronger when the average distance between the data point is lower than the integral scale. The behaviour is similar for the missed area ( $e_{ma}$ ) and the useless cost ( $e_{cu}$ ). Note that even in the best case (1000  $T$  measurements), there is still about 50 % of the forecasted protection zone that is useless ( $e_{cu}=0.5$ ) even if a large part of the reference protection zone is correctly identified ( $e_{ma}=0.3$ ). When stochastic simulations are employed, the uncertainty reduces, as expected, when the distance between the samples reduces (Fig. 4b). The reduction is stronger when the mean sample distance is lower than the integral scale ( $d < 1$ ). Like previous authors (Wen *et al.*, 1996; van Leewen *et al.*, 2000; Bakr & Butler, 2004), we observe that the inverse approach allows reducing the uncertainty. This reduction is more important when the  $T$  measurements are too distant to allow capturing the main transmissivity field structure. What is specific to our example is that the level of uncertainty remains very significant even when a large number of samples is used. The uncertain area has still a size of 1.8 times the area of the reference protection zone for 1000 samples.

An interesting aspect of the stochastic simulations is that the resulting probability map can be used as a tool to help decision makers. The decision maker can define a level of risk (a probability value) that he is ready to accept and based on this probability, identify the area on the map corresponding to this level of risk. By doing so, the error becomes a function of this probability value. A risk prone decision maker will select a high probability value (0.9 for example) and will apply the protection measures only to this area. As we can see in Fig. 4c, in this case, the decision maker will miss a large part of the zone that he should have protected ( $e_{ma}$  close to 1), but he

will reduce the surface that is unnecessarily ( $e_{uc}$  close to 0). It is interesting to notice that adding head data allows reducing considerably the errors for all probability values if the number of  $T$  data is small (comparison of Fig. 4c and 4d). Fig. 4e and 4f demonstrate that adding data points reduces the missed area for all probability levels, however the useless cost error has significantly increased when going from 21 to 250  $T$  measurements. A relatively surprising result is that the Kriging forecast performs as well as the simulations. Finally, it is important to note that all the errors remain rather high in this study and even for the case with the highest number of samples. If we analyse in detail the results of the inverse approach, the statistics show that the characterisation of the  $T$  field improves considerably when adding more data, but the head conditioning is not very efficient to improve the  $T$  field except in the case of 21  $T$  measurements. In the case of a 1000  $h$  samples for example, the head field is very well reproduced but the improvement of the  $T$  field is rather limited.

## CONCLUSION

The performances of kriging, stochastic simulations, and stochastic inversion to characterize a braided alluvial aquifer and to forecast protection zones have been compared. If most trends in our results are similar to those already published, the most important point is that even with a large number of data points, it remains difficult for the geostatistical techniques to identify the deterministic structures that control the shape of the protection zone. With a thousand  $T$  data, and even by taking a safe probability value of 0.2, the forecast still miss about 30% of the reference capture zone and in this case overestimate the area that should be protected by 60%. This is a rather disappointing result: it demonstrates that an accurate forecast in this specific example requires an amount of data that is not realistic. We think that there are two reasons for this behaviour. First, geostatistics cannot create information at a scale which is below the resolution of the sampling. Second, the fact that we are using a conceptual stochastic model that does not include explicitly the channels has probably a high impact as well. This result should not however be misunderstood. We are convinced that stochastic techniques have a major role to play in applied studies like the delineation of protection zones. However, the practitioner must be aware that the accuracy of the forecast will depend on the amount of data available and on the adequacy of the model with the real field structure.

## REFERENCES

- Bakr, M. I., & Butler, A. P. (2004) Worth of head data in well-capture zone design: deterministic and stochastic analysis. *Journal of Hydrology*. **290**(3-4), 202-216.
- Cornaton, F. (2004) *Deterministic models of groundwater age, life expectancy and transit time distributions in advective-dispersive systems*. PhD Thesis, University of Neuchâtel, Switzerland.
- Hendricks Franssen, H. J. W. M. (2001) *Inverse stochastic modelling of groundwater flow and mass transport*. PhD Thesis. Technical University of Valencia, Spain.
- Mosley, M. P. (1982) Analysis of effect of changing discharge on channel morphology and instream uses in a braided river, Ohau River, New Zealand. *Water Resour. Res.* **18**(4) 800-812.
- Stauffer, F., Hendricks Franssen, H.-J., and Kinzelbach, W. (2004) Semianalytical uncertainty estimation of well catchments: Conditioning by head and transmissivity data. *Water Resour. Res.* **40**(W08305): doi:10.1029/2004WR003320.
- Uffink, G.J.M. (1989) Application of the Kolmogorov's backward equation in random walk simulation of groundwater contaminant transport. In Kobus, H.T., & Kinzelbach, W. (Eds) *Contaminant Transport in Groundwater*, Balkema, 283-289.
- van Leeuwen, M., Butler, A. P., te Stroet, C. B. M., and Tompkins, J. A. (2000) Stochastic determination of well capture

zones conditioned on regular grids of transmissivity measurements. *Water Resour. Res.* **36**(4) 949-957.

Wen X. H., Gomez-herandez, J. J., Capilla, J. E., and Sahuquillo, A. (1996) Significance of conditioning to piezometric head data for predictions of mass transport in groundwater modelling. *Mathematical Geology* **28**(7), 951-868.