How many transmissivity data are required to safely delineate a protection zone of a braided alluvial aquifer when using kriging and upscaling?

CENTRE D'HYDROGÉOLOGIE
UNIVERSITÉ DE NEUCHÂTEL

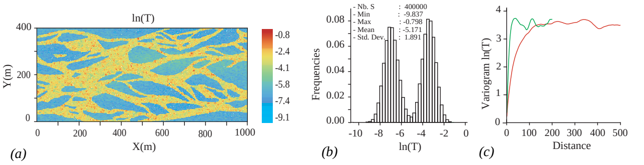
J. Kerrou [1], P. Renard [1], I. Lunati [2], S. Madier [3], and H.J. Hendricks-Franssen [4]
[1] Center of Hydrogeology, University of Neuchâtel, Neuchâtel, Switzerland
[2] Institute of Fluiddynamics, ETH Zurich, Zurich, Switzerland
[3] Strasbourg National School for Water and Environment Engineering, Strasbourg, France
[4] Institute of Hydromechanics and Water Resources Management, ETH Zurich, Zurich, Switzerland

1. Introduction

Protection zones play a key role in the protection of water-supply wells against contamination by pollutants from industry and agriculture. In Switzerland, the protection zones (SII) correspond to the 10-days isochrone area around the well. To be accurate, the forecast of these zones requires a good knowledge of the geometry, the hydrodynamic parameters, and the boundary conditions. However, usually in practice, these parameters are poorly known and the model results suffer from uncertainty. Therefore, the knowledge of subsurface heterogeneity is of prime importance to evaluate transit times of pollutants from infiltrating surfaces or riverbanks to the water-supply wells. Many studies have shown the critical impact of spatial heterogeneity on mass transport predictions (Dagan 1986). In this study, we test the efficiency of kriging estimation by taking into account upscaling on a synthetic alluvial aquifer. Then we investigate the effect of adding transmissivity data points on the accuracy of protection zone forecasts.

3. The reference

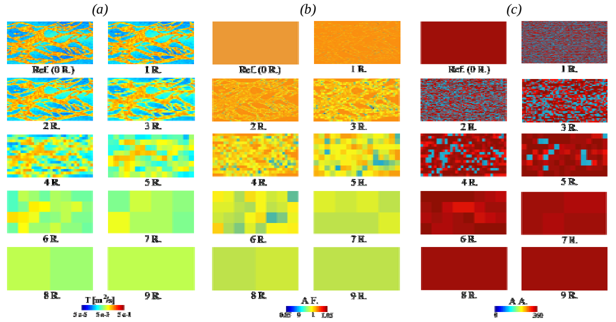
The transmissivity reference field is a 1000 m x 400 m 2D domain with a resolution of one meter. The field has been constructed from an aerial photograph of a braided river system containing lenses and channels. These two structures are then populated by MultiGaussian simulations. The results is a bimodal distribution for *ln(T)* that displays an anisotropic experimental variogram. The integral scale *i* (Gelhar 1993) is calculated in each direction: iX = 25.77 m and iY = 7.36 m.



(a) ln(T) distribution for the reference field; *(b)* ln(T) histogram; and *(c)* experimental reference field variogram

5. Upscaling

According to Renard (1997), the tensorial renormalisation (Gautier and Noetinger 1997) is one of the most accurate fast methods to calculate the equivalent permeability tensor. We therefore selected this technique and applied it on a series of successive upscaled grids. A dimensionless factor *Li* represents the size of the homogenized blocks normalized by the integral scale. As a point of comparison, we also calculate the geometric mean of the sampled transmissivities and use it in a homogeneous field.



(a) Upscaled maximum principal value of the transmissivity; *(b)* upscaled anisotropy factor; and *(c)* upscaled anisotropy angle. Calculated in this example for the reference field.

9. Results and Discussion

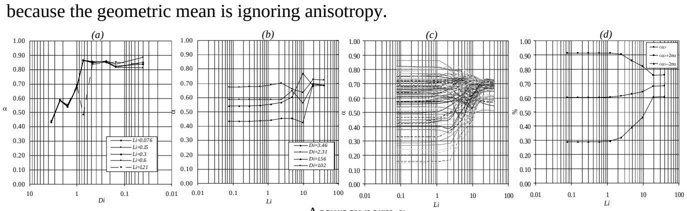
The graphs of the accuracy and error norms α and β (below) and the maps (besides) have shown that:

➢ When the average distance between samples is clearly greater than the integral scale (*Di*>1) the forecasts are improving approximately linearly with the logarithm of *Di* (graph-*a*-). When *Di* is less than 1, adding data points does not improve the forecasts.

➢ When *Di*>1, upscaling with the tensorial renormalization before calculating the protection zones usually improves the forecasts if the size of the homogenized blocks is greater than the integral scale (*Li*>1). The best accuracy is obtained with a completely homogenized domain (graph-*b*,*c*&*d*-).

➢ A sensitivity study analyzing the effect of sample location showed that the previous effect is confirmed on average (graph-*d*-) even if there are some particular cases for which the error increases with upscaling (graph-*c*-).

➢ In all cases, the forecasts made with the geometric mean of the sample data were bad because the geometric mean is ignoring anisotropy.



Accuracy norm α

2. Methodology

To estimate the worth of adding transmissivity data and upscaling effect on improving the accuracy of the calculated protection zone, it is necessary to have a detailed solution (reference). This study includes the following steps:

- ➢ A realistic bimodal transmissivity field is generated by combining deterministic and stochastic elements. It represents a braided river alluvial aquifer environment;
- ➢ Steady-state reverse flow and advective-dispersive transport are solved in the reference T field under a uniform gradient and a single pumping well to calculate the reference protection zone;
- ➢ Successive data sets of punctual transmissivities are randomly sampled from the T reference;
- ➢ Experimental variograms are calculated, modeled and used to krige a number of T fields;
- ➢ 9 successive stages of upscaling are applied on every estimated T field;
- ➢ Protection zones are calculated by solving again the steady-state reverse flow and advective-dispersive transport in each estimated T field (including the upscaled T fields);
- ➢ Calculated protection zones are compared with the reference and error indicators are calculated.

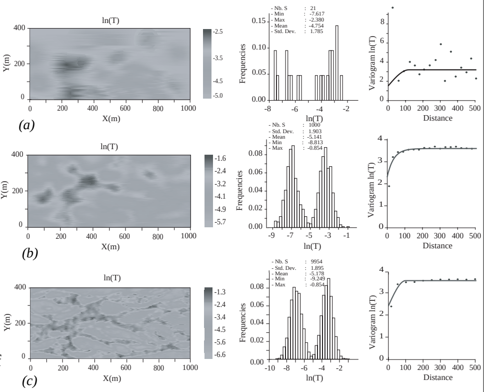
4. Sampling the reference

Successive data sets are taken randomly from the transmissivity reference field. Average distances between samples are calculated (table besides). We define *Di* the dimensionless average distance between the samples normalized by the integral scale. If it is greater than 1, sample points cannot resolve the heterogeneity that has a length scale on the order of the integral scale.

Number of samples	Average distance Dx [m]	Di1=Dx/ιX [ ]
21	89.33	3.467
44	59.54	2.310
96	40.36	1.566
250	26.37	1.023
416	19.16	0.743
1000	12.51	0.486
4000	6.33	0.246
9954	4.02	0.156
400000	1.00	0.039

6. T field estimation

An experimental variogram for each ln(T) set of samples has been calculated. Directional variograms (30° and 45°…) were tested but rejected because they didn't show anisotropy. A spherical model with nugget effect has been kept for all sets of data. Specific model parameters were inferred for every case and cross-validated. The interpolation was then conducted by lognormal kriging. The figure shows the data histograms, variograms models and the estimated T field. Note that the channels of the reference are reproduced only when a high number of samples is used.



Histogram, variogram model and estimated (ln)T field for *(a)* 21, *(b)* 1000, and *(c)* 9954 samples

7. Protection zone calculation

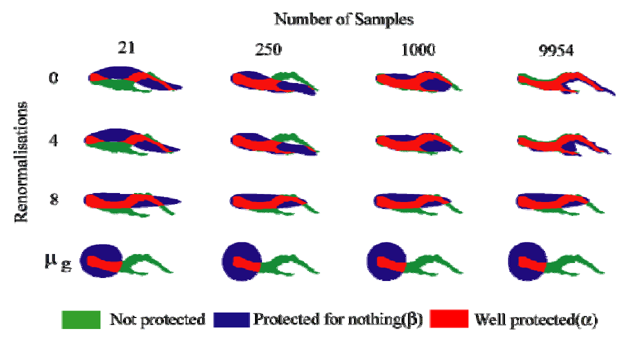
The protection zone is calculated by solving the 2D steady-state reverse groundwater flow and advective-dispersive transport with the finite element code Feflow. The result is the distribution of average life expectancies. We then select all the grid nodes having average life expectancies lower than 10 days. The same method is applied for the reference and for all the estimated and upscaled T fields.

8. Error and accuracy norms

To quantify and compare the error we decided to use two error norms:

α = the percentage of the reference protection zone that has been correctly forecasted. (red area normalized by red and green)

β = the percentage of the calculated zone that is protected for nothing normalized by the calculated zone (blue area normalized by blue plus red)



References

Dagan G. (1986) Statistical theory of groundwater flow and transport; pore to laboratory, laboratory to formation and formation to regional scale. Water Resources Research. 22, 1205–1355

Gautier Y., **Noetinger B.** (1997) Preferential flow-paths detection for heterogeneous reservoirs using a new renormalisation technique. Porous Media, 26, 1-23

Gelhar L.W. (1993) Stochastic Subsurface Hydrology. Prentice Hall New Jersey

Renard P., **Le Loc'h G.**, **Ledoux E.**, **De Marsily G.** (2000) A fast algorithm for the estimation of the equivalent hydraulic conductivity of heterogeneous media. Water Resources Research 36(12), 3567-3580