

## **Inverse modelling of groundwater flow and groundwater mass transport**

**JOHAN VALSTAR**

*Netherlands Institute of Applied Geoscience-TNO, National Geological Survey, PO Box 80015, 3508 TA Utrecht, The Netherlands*

e-mail: [j.valstar@nitg.tno.nl](mailto:j.valstar@nitg.tno.nl)

**Abstract** An inverse algorithm for groundwater flow and groundwater mass transport modelling has been tested for a rather heterogeneous aquifer. Results show that a limited number of measurements (four head and 45 concentration measurements) are sufficient to provide detailed information about the head and concentration distribution. The estimate of the conductivity values showed a significant improvement as the estimate captures most of the large-scale structure of the real conductivity field. Variation on a scale that is smaller than the density of the measurements cannot be estimated.

**Key words** data analysis; groundwater modelling; heterogeneity; inverse modelling

### **INTRODUCTION**

Groundwater flow and groundwater mass transport models try to predict the behaviour of groundwater and groundwater contaminants as successfully as possible. Unfortunately, model predictions and especially predictions of contaminant transport, are not reliable in many cases. One of the reasons for erroneous predictions is the heterogeneity of the subsurface, which in most cases is not well known. Direct measurements of soil properties at such a small scale that they describe the heterogeneity well enough demands an enormous amount of measurements and consequently, the associated costs are high.

In this research, an alternative method of obtaining information about the heterogeneity of the subsurface (inverse modelling) has been applied. Inverse models use measurements of the state variables that are related to the properties that are to be estimated. In this case, groundwater head and concentration measurements are used to obtain better estimates of soil properties, such as hydraulic conductivity and sorption coefficients. As it is impossible to obtain all the information needed to describe the heterogeneity of the conductivity completely, part of this heterogeneity is upscaled to effective parameters. In this study, the influence of the heterogeneity of the conductivity at a scale smaller than the grid size is simulated using the concept of dual porosity which includes a kinetic exchange between the mobile and immobile domains (Harvey & Gorelick, 2000).

### **INVERSE MODEL**

A Bayesian approach has been used, in which the following objective function will be minimized (Tarantola, 1987):

$$J = (z - M(h, c))^T P_z^{-1} (z - M(h, c)) + (\alpha - \bar{\alpha})^T P_\alpha^{-1} (\alpha - \bar{\alpha})$$

where  $J$  is the objective function;  $z$  are the measurements;  $M(h, c)$  are the model predictions of the measurements;  $P_z$  is the measurement error covariance matrix;  $\alpha$  are the model parameters to be estimated;  $\bar{\alpha}$  is the prior mean of the model parameters and  $P_\alpha$  is the prior covariance of the model parameters; and superscript  $T$  denotes the transpose of the vector.

This objective function incorporates both residuals between measurements and model predictions and deviations from the *a priori* mean of the model parameters. The measurement residuals are weighted by the inverse of the covariance of the measurement errors, whereas the parameter deviations are weighted by the inverse of the *a priori* covariance function (Tarantola, 1987).

The minimization of the objective function is computer-time consuming for large-scale real-world problems because the number of unknown parameters becomes very large. However, it can be proved that the number of unknown parameters can be reduced to the number of measurements that are available using an optimal parameterization technique and consequently the computer time is reduced considerably. Details about this inverse technique can be found in Valstar (1997) for inverse groundwater flow problems, and Valstar (2001) for inverse groundwater flow and transport problems. This technique is an extension of the representer technique used in oceanography (Bennett, 1992) and techniques applied in groundwater transport problems (Reid, 1996; Sun, 1998; McLaughlin & Townley, 1996).

## MODEL SET-UP

A two-dimensional heterogeneous conductivity field was first generated and then the transport of a contaminant was simulated. Groundwater flow was driven by infiltration to the top of the aquifer and a constant flux across the model boundary on the right-hand side. The left boundary is a constant head boundary and the lower boundary is a no flux boundary. The source is modelled as an initial plume that is solely present in the mobile domain (width 15 m, depth 2.0 m) at the top of the aquifer in the right-hand part of the model domain. The model domain is discretized by a rectangular grid of 100 cells in the horizontal direction, and 25 cells in the vertical direction. The dimensions of the cells are 1.0 m by 0.25 m. The total simulation time is 400 days, which has been divided into 120 time steps.

### Hydraulic conductivity

A synthetic hydraulic conductivity field has been generated as a log-normally distributed random field with an exponential anisotropic covariance structure. The mean of the natural logarithm of the conductivity is  $\ln(1.0 \text{ m day}^{-1})$ , while its variance is 2.0 and the horizontal and vertical correlation lengths are 50 and 2.0 m respectively. The real synthetic natural logarithms of the conductivity values are shown in Fig. 1.

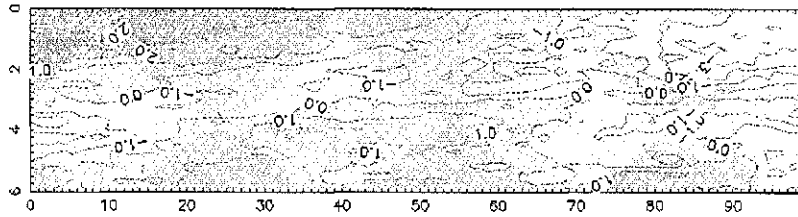


Fig. 1 Synthetic natural logarithms of the conductivity distribution.

### Exchange rate coefficient and effective immobile porosity

The dual porosity parameters, the exchange rate coefficient and the effective immobile porosity have been generated as log-normally distributed random parameters as well. The mean of the natural logarithm of the exchange rate coefficient is  $\ln(0.001 \text{ day}^{-1})$  whereas it has an exponential covariance structure with variance of 1.0 and correlation lengths of 50 and 2.0 m in the horizontal and vertical directions, respectively. The effective immobile porosity is generated as a spatially constant, but random log-normally distributed parameter. The mean of the natural logarithm of the effective immobile porosity is 0.0 (–) and its variance is 1.0. Results showed that concentrations were not very sensitive to these dual porosity parameters in this example (Valstar, 2001) Consequently, their estimation in this example is very difficult because the concentration measurements provide little information about the parameters. Therefore, most emphasis in this paper is addressed at the estimation of conductivity.

### Known parameters

The other model parameters comprise the mobile porosity and the longitudinal and transverse dispersivity which are known to be spatially constant. The mobile porosity is 0.20 and the longitudinal and transverse dispersivities are 0.25 and 0.025 m, respectively.

## RESULTS AND DISCUSSION

### Forward modelling

The synthetic model parameters are used in the forward model to obtain the real heads and the real mobile and immobile concentrations. The groundwater heads are shown in Fig. 2. The head distribution indicates small gradients in the left part of the domain, where the conductivity values are large. The head gradients are higher in the right part of the domain, as the conductivity values are lower. The upper right part shows a very large head gradient, which is due to the small conductivity values and the flux that is forced into this region due to the constant flux boundary condition. The groundwater heads have little variation in the vertical direction.

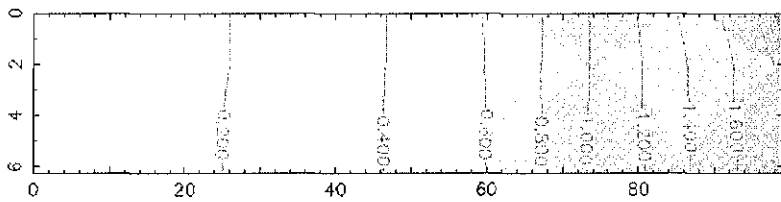


Fig. 2 Real groundwater head distribution.

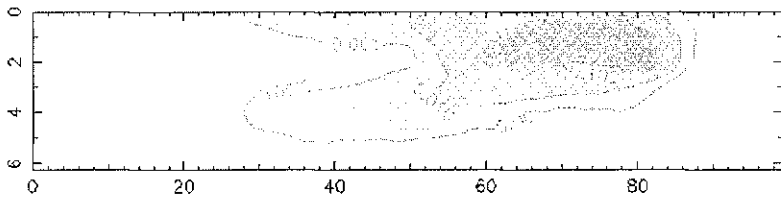


Fig. 3 Synthetic mobile concentration distribution at the end of the simulation.

The mobile concentrations after 400 days are shown in Fig. 3. The distribution of the mobile concentration clearly shows that the bulk of the plume has not travelled a large distance as initially, it is trapped in the low conductivity zone, close to the initial plume position. As soon as it reaches a high conductivity region, the contamination travels faster but the concentrations in the permeable region are much lower. It is clear that two fronts of the plume develop: one in the lower part of the domain and one just below the surface; both these areas have a relatively high conductivity, compared with the low conductivity region in between.

Series of measurements were collected from the head and concentration distributions after 400 days and from concentration distributions at earlier time steps. Synthetic measurement errors were added to these values. The measurement errors are Gaussian distributed with a standard deviation of 1 cm for the head measurements and the standard deviation was set at 10% of the real value of the concentration for the concentration measurements.

### Prior estimates

Using the prior mean estimates for the model parameters, the prior estimates of the state variables are calculated with the forward model. The prior estimates of heads are shown in Fig. 4. The prior estimates of the mobile concentrations at the end of the simulation are shown in Fig. 5.

When comparing these prior estimates with the real heads and concentrations it is obvious that there are significant differences. The prior heads are much higher than the real heads, while the two concentration fields have a different shape for the plume and moreover the maximum of the prior mobile concentrations is at quite a different location than the maximum of the real mobile concentrations.

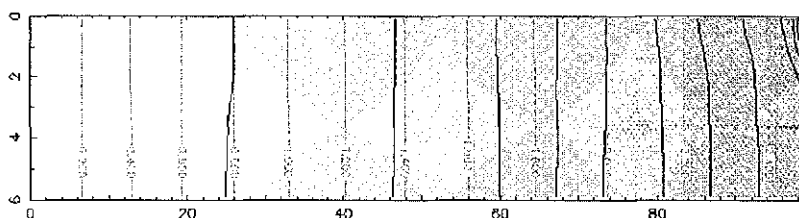


Fig. 4 Prior (grey scaled areas) and real (solid line) head distribution.

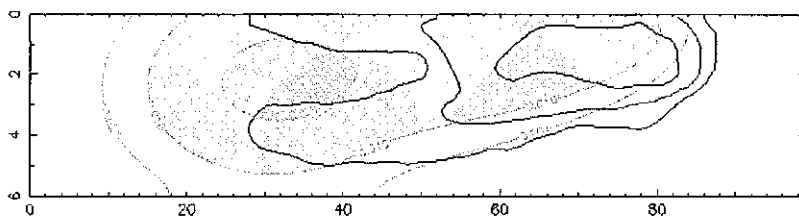


Fig. 5 Prior (grey scaled areas) and real (solid line) mobile concentration distribution at the end of the simulation.

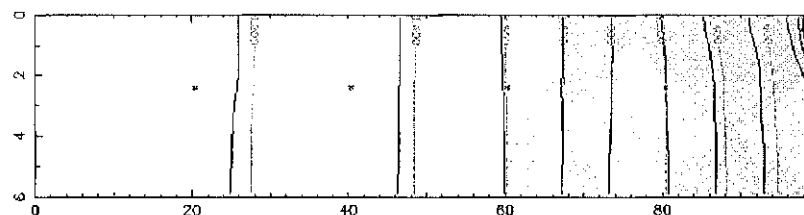


Fig. 6 Posterior (grey scaled areas) and real (solid line) head distribution; the asterisks denote the locations of the head measurements.

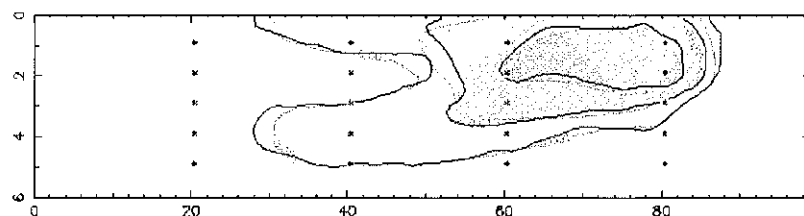


Fig. 7 Posterior (grey scaled areas) and real (solid line) mobile concentration distribution at the end of the simulation; the asterisks denote the locations of the mobile concentration measurement locations.

## Posterior estimates

In this example, four head and 45 mobile concentration measurements were used in the inverse algorithm. The locations of the head measurements are shown in Fig. 6. The mobile concentrations were taken at three different time steps. Ten, 15 and 20 measurements were sampled after 20, 50 and 120 time steps, respectively. The 20 locations at which measurements were sampled after 120 time steps are shown in Fig. 7. After 20 and 50 time steps, the mobile concentrations were sampled at the 10 and 15 most upstream locations.

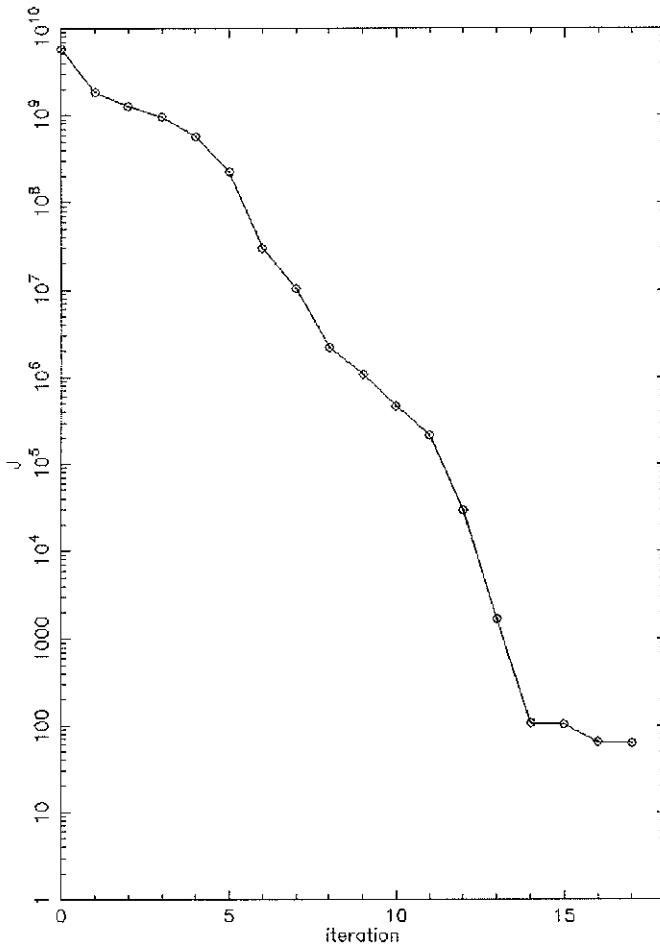


Fig. 8 Evolution of the objective function.

The model converged in 17 iterations. The evolution of the objective function is shown in Fig. 8, which shows that the value of the objective function decreases over approximately eight orders of magnitude.

The posterior head distribution is shown in Fig. 6. The posterior head estimates show a large improvement compared with the prior head estimates (Fig. 4). It fails to predict the high head gradients in the upper right corner of the domain, but that is not surprising as no head measurements were used in this part of the domain.

The posterior mobile concentration distribution after 120 time steps is shown in Fig. 7. The posterior mobile concentrations also show a large improvement when compared with the prior estimates in Fig. 5. The occurrence of the greatest concentrations close to the location of the initial plume, and the development of two fronts downstream are predicted.

The posterior estimates of the natural logarithms of the conductivity are shown in Fig. 9. The distribution of the posterior natural logarithms of the conductivity shows a

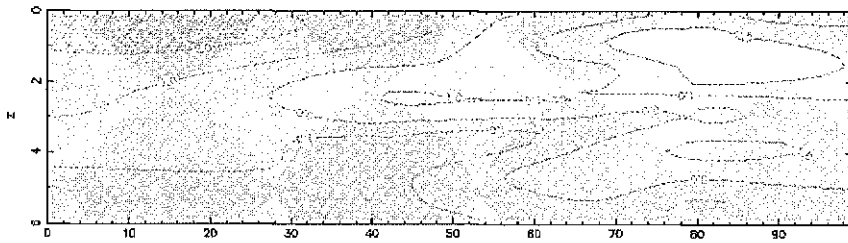


Fig. 9 Posterior natural logarithms of the conductivity distribution.

low conductivity area in the upper right part of the domain, which is close to the source. High conductivity areas are predicted in the upper left corner of the domain and the lower left part of the domain, which are the regions where the two fronts of the plume start to develop. It is also clear that the posterior conductivity distribution is a smooth distributed field, when compared to the real conductivity field. Variation of model parameters on a scale that is smaller than the density of measurement network is not recaptured as no information about the smaller scale variation is available. The inverse model, however, provides estimates that are comparable to up-scaled, effective parameters.

## REFERENCES

- Bennett, A. F. (1992) *Inverse Methods in Physical Oceanography*. Cambridge University Press, Cambridge, UK.
- Boggs, J. M., Young, S. C., Beard, L. M., Gelhar, L. W., Rehfeldt, K. R. & Adams, E. E. (1992) Field study of dispersion in a heterogeneous aquifer. I. Overview and site description. *Wat. Resour. Res.* **28**, 3281–3291.
- Harvey, C. & Gorelick, S. M. (2000) Rate limited mass transfer or macrodispersion: which dominates the plume evolution at the Macrodispersion Experiment (MADE) site? *Wat. Resour. Res.* **36**, 637–650.
- McLaughlin, D. & Townley, L. R. (1996) A reassessment of the groundwater inverse problem. *Wat. Resour. Res.* **32**, 695–704.
- Reid, L. (1996) A functional inverse approach for three-dimensional characterization of subsurface contamination. PhD Thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.
- Sun, C. C. (1998) A stochastic inverse approach for characterizing soils and groundwater contamination at heterogeneous field sites. PhD Thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.
- Tarantola, A. (1987) *Inverse Problem Theory*. Elsevier, The Netherlands.
- Valstar, J. R. (1997) Efficient inverse modeling: a representer approach. In: *Proc. IAMG 97* (Third Annual Conf. International Association for Mathematical Geology) (ed. by V. Pawlowsky-Glahn), 767–771. CIMNE, Barcelona, Spain.
- Valstar, J. R. (2001) Inverse modeling of groundwater flow and transport. PhD Thesis, Delft University of Technology, Delft, The Netherlands.