Calibration and Reliability in Groundwater Modelling (Proceedings of the ModelCARE 99 Conference held at Zürich, Switzerland, September 1999). IAHS Publ. no. 265, 2000.

Tracer travel time analysis for inference of aquifer geostatistical parameters

ALBERTO BELLIN

Dipartimento di Ingegneria Civile ed Ambientale, Università di Trento, via Mesiano 77, I-38050 Trento, Italy

e-mail: alberto.bellin@ing.unitn.it

YORAM RUBIN

Department of Civil and Environmental Engineering 435 Davis Hall, University of California, Berkeley, California 94720, USA

Abstract We present a new methodology for inference of the geostatistical models of an aquifer's hydrogeological parameters from the breakthrough curves measured at samplers placed over control planes. The parameters inferred are the mean velocity, the log conductivity's variance and integral scale, and the longitudinal pore-scale dispersion coefficient. The overall conclusion of the study is that the hydrogeological parameters can be inferred with reasonable accuracy by using suitable tracer tests, but that at the same time the inference of pore-scale dispersion is error prone when conducted using the same set of data.

INTRODUCTION

Stochastic modelling of flow and transport in heterogeneous geological formations calls for a geostatistical model of the hydraulic property variations which relies on suitable hydrogeological parameters. The inference of these parameters from field data represents a major challenge that limits the application of stochastic theories. Current inference methods result in large expenditure and often the applications are somewhat limited by difficulties encountered in the interpretation of the data.

To study the transport of solutes in heterogeneous geological formations, tracer tests were designed at several field sites to collect breakthrough curves (BTC) at samplers installed over control planes normal to the mean flow direction. In this paper we investigate the potential for using the breakthrough curves at the samplers for parameter inference. Specifically, we focus on the peak concentration arrival times t_p at the multilevel samplers. This paper presents the most relevant conclusions of a recent work by Bellin & Rubin (2000). The idea of using t_p was inspired by an early suggestion of Ptak & Teusch (1994) and a recent study by Rubin & Ezzedine (1997). Ptak & Teusch (1994) observed experimentally that t_p is affected by the spatial variability of the hydraulic conductivity. Rubin & Ezzedine (1997) then analysed the Cape Cod tracer test data showing that the variance of the travel time compares favourably with that of t_p.

This work revisits some of the ideas presented in Rubin & Ezzedine (1997) and investigates them rigorously both theoretically and through a numerical case study. The latter consists of simulating numerically a tracer test in a two-dimensional isotropic formation with the exponential covariance function. We note that the proposed inference methodology is not limited to the particular choice of the case study (Bellin & Rubin, 2000).

NUMERICAL EXAMPLE AND INFERENCE METHODOLOGY

The experimental setup employed in natural gradient tracer tests includes arrays of multilevel samplers installed over planes normal to the mean flow direction and equipped with several sampling ports. Solute breakthrough curves are collected at all the sampling locations. The differences between the breakthrough curves reflect both the aquifer heterogeneity and the limitation of the sampling.

Figure 1 shows three BTC collected at three samplers and the concentration flux at the control plane, which for simplicity is called the global breakthrough curve. The three samplers are a small part of the total number of multilevel samplers installed over the control plane.

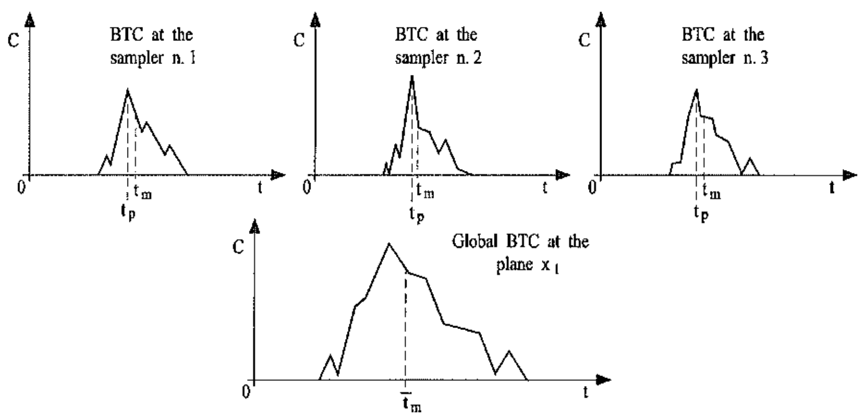

Fig. 1 Example of breakthrough curve at different sampling locations and the concentration flux across the control plane.

The breakthrough curves show two distinct features that can be conveniently used to infer the hydrogeological parameters. The second moment of the BTC is much smaller than the variance of the global BTC. In addition, a large disparity is observed between the arrival times of the breakthrough curves (Rubin & Ezzedine, 1997). Specifically, we observe that the peak concentration arrival time, t_p, of the BTC varies between the samplers according to the hydraulic property variations that the solute experiences travelling towards the sampler. The same type of spatial variability is observed for the mean arrival time t_m. This suggests that the solute spreading observed at the samplers is controlled by pore-scale dispersion and small scale variations of the hydraulic conductivity, while the disparity between the arrival times at the samplers is the resulting effect of variations of the hydraulic conductivity at scales larger than the size of the sampler. However, if the size of the sampler is much smaller than the scale of heterogeneity one can assume that the second order moment of the BTC is affected by pore-scale dispersion only. Furthermore, in the Lagrangian framework, t_p can be assumed to represent the travel time of the particle injected at the initial position of the centre of mass of the solute captured by the sampler. According to this hypothesis the spread of t_p, and that of t_m too, can be used in combination with analytical expressions of the travel time moments for the inference of the hydrogeological parameters.

The interpretation of data in real life is complicated by incomplete sampling of the breakthrough curves and concentrations below the detection limit of the instruments. Furthermore, small-scale variations of the hydraulic conductivity introduce additional solute spreading in the BTC, which cannot be distinguished from that caused by pore-scale dispersion. How this affects the inference of the hydrogeological parameters is a question we want to answer with this study.

Numerical example

To answer the above fundamental questions, we simulated numerically the tracer test in a two-dimensional heterogeneous statistically isotropic formation. Pore-scale dispersion was simulated by a Brownian motion with constant longitudinal Dd_{11} and lateral Dd_{22} dispersion coefficients. The variance of the log conductivity field $Y = \ln T$ is fixed at $\sigma_Y^2 = 0.2$.

A total mass $M_0 = nC_0V_0$ of a conservative solute with constant concentration C_0, was instantaneously released within the rectangular volume V_0 of unitary thickness and sides l_1 and l_2 in the longitudinal and transverse horizontal directions, respectively. The porosity n is assumed constant throughout the formation. Furthermore, to investigate the influence of the source size on the inferred parameters, l_2 was varied between $5I_{Y,a}$ and $40I_{Y,a}$ while keeping l_1 constant and equal to $I_{Y,a}$. Additional information on the numerical setup is provided in Bellin & Rubin (2000). The breakthrough curves were collected at the sampling locations over the control plane normal to the mean flow direction. The spacing between the samplers was varied between $0.1I_{Y,a}$ and $I_{Y,a}$. To investigate how the distance from the source affects the inferred parameters, we applied the inference procedure to data collected over several control planes.

Inference methodology

The integral scale of Y is inferred by best fitting of the first order analytical expression of the transverse horizontal travel time autocovariance function (Rubin & Ezzedine, 1997):

$$\rho_\tau(x_1;r_2) = \frac{2}{X'_{11}(x_1)}\int_0^{x_1}(x_1-\xi)u'_{11}(x_1-\xi,r_2)\,\mathrm{d}\xi \tag{1}$$

with the corresponding raw autocovariance function of t_p, which is given by the following expression:

$$\rho_{t_p}(x_1;r_2) = \frac{1}{S_{t_p}^2(x_1)N(r_2)}\sum_{i=1}^{N(r_2)}\left[t_p(x_1,x_{2,i})-\bar{t}_p(x_1)\right]\left[t_p(x_1,x_{2,j})-\bar{t}_p(x_1)\right] \tag{2}$$

where $N(r_2)$ is the number of pairs with reciprocal distance equal to $r_2 = x_{2,j} - x_{1,j}$, $t_p(x_1,x_{2,i})$ is the peak concentration arrival time at the sampling location $(x_1,x_{2,i})$, and $\bar{t}_p(x_1)$ and $S_{t_p}^2(x_1)$ are the mean and variance of t_p at the control plane x_1, respectively.

In equation (1), $X'_{11} = X_{11}/\sigma_Y^2 I_Y^2$, where X_{11} is the longitudinal component of the one-particle displacement variance, and u'_{11} is the dimensionless covariance function of the longitudinal velocity. The best fitting is performed by using the Levenberg-Marquardt algorithm applied to the special case of a least squares function (Dennis & Schnabel, 1983) in the implementation of the MINPACK package (Morè *et al.*, 1980).

To infer Dd_{11} we postulate that the sampler size is smaller than the scale of the heterogeneity and that the lateral pore-scale dispersion does not influence the moments of the breakthrough curve. With these simplifying assumptions in mind we devel oped the theoretical model for the BTC at the sampling location. Dd_{11} is then inferred by matching the second temporal moment of the BTC with the model. The theoretical model and the inference methodology are discussed in Bellin & Rubin (2000).

In most practical cases Dd_{22} is one order of magnitude smaller than Dd_{11} (Dagan & Fiori, 1998), such that its influence on the travel time moments is small to negligible (Fiori, 1996). For this reason we do not attempt here the inference of Dd_{22}.

The mean velocity U and the log-transmissivity variance σ_Y^2 are obtained by solving the following nonlinear system of equations:

$$\begin{cases} \langle \tau(x_1) \rangle = \bar{t}_p(x_1) \\ \sigma_\tau^2(x_1) = S_\tau^2(x_1) \end{cases} \tag{3}$$

In equation (3), $S_\tau^2(x_1) = S_{t_p}^2(x_1) + \bar{S}_{\tau,l}^2(x_1)$ is the variance of the global BTC, where $\bar{S}_{\tau,l}^2(x_1)$ is the average of the variance of the breakthrough curve at the samplers over the control plane. In addition, $<\tau>$ and σ_τ^2 are the mean and the variance of the particle's travel time at the control plane x_1 (Rubin & Ezzedine, 1997):

$$\langle \tau(x_1) \rangle = \int_0^\infty [1 - G(t; x_1)]\,\mathrm{d}t; \quad \sigma_\tau^2(x_1) = 2\int_0^\infty t[1 - G(t; x_1)]\,\mathrm{d}t - \langle \tau(x_1) \rangle^2 \tag{4}$$

where G is the cumulative probability distribution function (CDF) of the travel time $\tau(a,x_1)$ of the particle injected at the position $x = a$. In equation (3) $\bar{t}_p$ and S_τ^2 are computed by using the actual tracer test data at the control plane x_1, and $<\tau>$ and σ_τ^2 through equation (4) depend nonlinearly on U and σ_Y^2. The solution of the system, equation (3), is obtained by using the package MINPACK (Morè *et al.*, 1980) which offers an efficient implementation of the modified Powell's algorithm (Morè *et al.*, 1980).

To what extent the inferred parameters are affected by the actual distribution of the hydraulic transmissivity and the errors introduced by the inversion procedure is an important point we want to clarify with the numerical simulations. This is accomplished through a Monte Carlo study consisting of repeated simulations of the tracer test. The uncertainty associated with the inference of the parameter θ performed by using the actual tracer test data is proportional to the coefficient of variation $CV[\theta] = \sigma[\theta]/<\theta>$, where $\sigma[\theta]$ and $<\theta>$ are the standard deviation and the expected value of θ, respectively.

DISCUSSION OF THE RESULTS AND CONCLUSIONS

In the following the performance of the inference methodology is analysed by considering the ratio between the expected value of the inferred parameters and the corresponding actual values $<\theta>/\theta_a$ and also the coefficients of variation.

The expected value of σ_Y^2, obtained by solving the nonlinear system of equations (3) in repeated simulations of the tracer test, resulted much larger than the actual value (Bellin & Rubin, 2000). The explanation is in the nonlinear dependence of σ_τ^2 on I_Y, which in combination with the fact that the inferred I_Y fluctuates around the actual value, results in inferred σ_Y^2 which on average are larger than what was expected according to $<S_\tau^2>$ and the actual value of I_Y (Bellin & Rubin, 2000). Substituting in equation (3) $S_\tau^2(x_1)$ with S_{tp}^2, the expected value of the inferred σ_Y^2 results in a better match with the actual value. Thus, the results discussed in the following Section are then obtained by using S_{tp}^2.

The density of the sampling locations is an important factor influencing the accuracy of the inference procedure. Preliminary simulations, discussed in the work by Bellin & Rubin (2000), showed that the inferred parameters are insensitive to the density of the sampling locations if the spacing between the samplers is smaller than $I_{Y,a}$. Another important factor influencing the inferred parameters is the source size. Bellin & Rubin (2000) showed that a transverse source size larger than $15I_{Y,a}$ should be employed in order to get a good match between the expected value of the inferred parameters and the corresponding actual value. Furthermore $CV[\theta]$ decreases as the source size increases and ergodic conditions are approached.

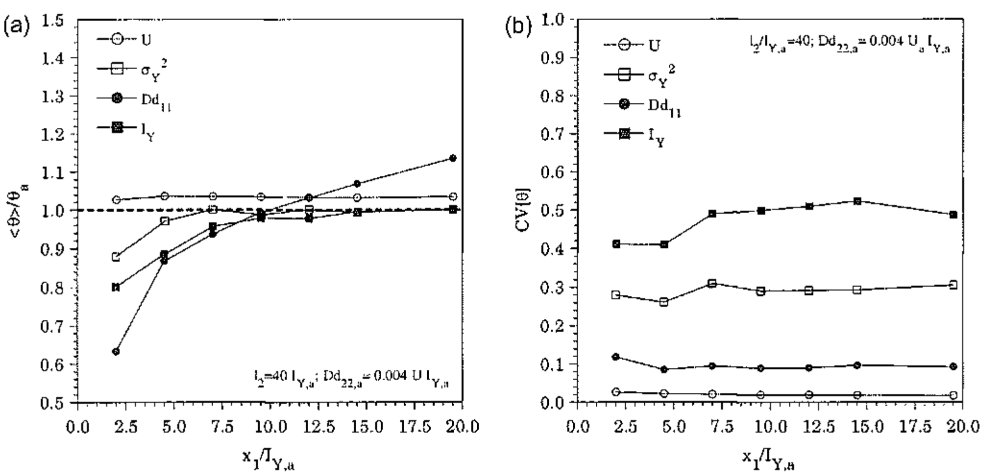

Fig. 2 Inferred parameters versus the distance from the source: (a) ratio between the expected and the actual values of the inferred parameters; (b) coefficient of variation of the inferred parameters.

Figures 2(a) and 2(b) show the expected value and the coefficient of variation of the inferred parameters versus the distance from the source. The transverse source size is fixed at $l_2 = 40I_{Y,a}$ and the spacing between the sampling locations is $\Delta s = 0.5I_{Y,a}$. Although at short distances from the source $<I_Y>$ is smaller than the actual value, it increases with the distance reaching for $x_1 > 7.5I_{Y,a}$ a stable limit which is slightly smaller than $I_{Y,a}$. Furthermore, for $x_1 > 5I_{Y,a}$ the relative difference between $<I_{Y,a}>$ and the actual value is smaller than 4%. Unfortunately the large coefficient of variation at all the control planes suggest that the inference of $I_{Y,a}$ from the actual tracer test data is uncertain. The inferred σ_Y^2 performs similarly to I_Y except for the fact that the relative difference between the expected value and the corresponding actual value reduces showing values smaller than 3% for $x_1 > 4.5I_Y$. Although $CV[\sigma_Y^2]$ is smaller than $CV[I_Y]$, it is still large leading to an uncertain σ_Y^2 when the inference is performed by using the actual tracer test data. However, Bellin & Rubin (2000) showed that when the inferred parameters are averaged over several control planes, the coefficients of variation reduce to values comparable with those obtained by Woodbury & Sudicky (1991) from extensive core analysis of the Borden site material.

Figures 2(a) and 2(b) show that U is inferred accurately at all the control planes with relative differences between $<U>$ and the actual value less than 4%, and $CV[U] < 0.02$. The inference of pore-scale dispersion is more problematic. Although $CV[Dd_{11}]$ is smaller than 0.12 and decreases slightly with the distance from the source, $<Dd_{11}>$ increases with distance. The mixing between adjacent stream tubes, which is induced by lateral pore-scale dispersion and streamline tortuosity, in combination with small scale variations of the hydraulic transmissivity, is the cause of the additional solute's spreading observed at the samplers. Figures 2(a) and 2(b) show also that at intermediate distances, where the processes producing additional spreading of the BTC are still not relevant, the inferred Dd_{11} is satisfactory in each realization.

The overall conclusions of this work are as follows. The mean velocity is inferred accurately irrespective of the source size and the distance from the source. The inference of the longitudinal pore-scale dispersion is error prone and acceptable values are obtained only at intermediate distances from the source. The other parameters should be inferred by using breakthrough curves collected at planes which distance from the source is larger than $x_1 = 4.5I_Y$ in tracer tests conducted with transverse source sizes larger than $20I_Y$.

Acknowledgements The first author acknowledges the support of the programme MURST-COFIN97 Monitoraggio e controllo del trasporto di contaminanti per la salvaguardia e la gestione delle risorse idriche sotterranee (Monitoring and control of contaminant transport for the management and the safety of groundwater). Support for the second author was made available through the Hydrologic Science Programme of NSF, grant EAR-9628306

REFERENCES

Bellin, A. &. Rubin, Y. (2000) On the use of tracer data for hydrogeological parameters inference. *Wat. Resour. Res.* (submitted)

Dagan, G. & Fiori, A. (1997) The influence of pore-scale dispersion on concentration statistical moments in transport through heterogeneous aquifers. *Wat. Resour. Res.* **33**(7), 1595–1605.

Dennis, J. E. & Schnabel, R. B. (1983) *Numerical Methods for Unconstrained Optimization and Nonlinear Equations.* Prentice-Hall, Englewoods Cliffs, New Jersey, USA.

Fiori, A. (1996) Finite-Peclet extension of Dagan's solutions to transport in anisotropic heterogeneous formations. *Wat. Resour. Res.* **32**(1), 193–198.

Morè, J. J., Garbow, B. S. & Hillstrom, K. E. (1980) User Guide for MINPACK-1. *Argonne National Laboratory Tech. Report. N. ANL-80-74.*

Ptak, T. & Teusch, G. (1994) Forced and natural gradient tracer tests in a highly heterogeneous porous aquifer: instrumentation and measurements. *J. Hydrol.* **159**, 79–104.

Rubin, Y. & Ezzedine, S. (1997) The travel times of solutes at the Cape Cod Tracer Experiment: data analysis, modeling, and structural parameters inference. *Wat. Resour. Res.* **33**(7), 1537–1547.

Woodbury, A. D. & Sudicky, E. A. (1991) The geostatistical characteristics of the Borden aquifer. *Wat. Resour. Res.* **27**(4), 533–546.