



## Analysis of Inverse Stochastic Moment Equations of Groundwater Flow in Heterogeneous Porous Media

Marco Panzeri, Emanuela Bianchi Janetti, Monica Riva, and Alberto Guadagnini

Dipartimento di Ingegneria Idraulica, Ambientale, Infrastrutture Viarie, Rilevamento (DIHAR), Politecnico di Milano, Piazza L. Da Vinci, 32, 20133 Milano, Italy

We present a nonlinear stochastic inverse algorithm that allows conditioning estimates of hydraulic heads, fluxes and their associated uncertainty on information about log-hydraulic conductivity ( $Y$ ) and head ( $h$ ) data collected in a randomly heterogeneous aquifer.

The methodology is framed in a Maximum Likelihood (ML) context and allows estimating jointly the spatial distribution of  $Y$ , the underlying variogram parameters as well as the variance of measurement errors associated with  $Y$  and  $h$  information. Estimates of prediction errors of hydraulic heads and fluxes are then calculated a posteriori, upon solving equations satisfied by the corresponding co-variances.

The methodology is based on recursive approximations of exact nonlocal first and second conditional stochastic moment equations of groundwater flow. Log-hydraulic conductivity is parameterized geostatistically on the basis of available measured values and a set of unknown values at discrete "pilot points". Whereas prior values of  $Y$  at pilot points are obtained by generalized kriging, posterior estimates are obtained through a ML fit of computed to measured heads. The ML-based objective function includes a regularization/plausibility term penalizing large departures of  $Y$  estimates from their prior information. The posterior estimates are then projected onto the computational grid by kriging. The variance of measurement errors and the key parameters of the  $Y$  variogram are evaluated separately to avoid bias and instability. We develop an efficient and accurate method for their estimation by embedding the geostatistical inversion of the flow process within a genetic search procedure.

The methodology is illustrated by means of a set of synthetic examples and field studies. We explore the influence of the number of pilot points and the relevance of the order of approximation of the governing mean flow equation on the identification of the system's parameters. We study the ability of different model selection criteria to properly identify the variance of measurement errors and the parameters of the  $Y$  variogram. We show theoretically and demonstrate by way of synthetic examples that minimization of the Negative Log-Likelihood (NLL) function, as well as model selection criteria such as HIC, BIC, AIC, and AICc, which are linearly dependent on NLL, are not suitable criteria to correctly identify these statistical parameters. On the other hand, we illustrate that the Kashyap's Bayesian criterion (KIC), supported by the adoption of a second-order closure of the mean flow equations, allows a proper estimation of the variance of measurement errors and of the  $Y$  variogram parameters.