INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN FLUIDS
*Int. J. Numer. Meth. Fluids* 2006; **51**:963–980
Published online 18 April 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/fld.1213

# Inversion of heterogeneous parabolic-type equations using the pilot points method

Andrés Alcolea*,†, Jesús Carrera‡ and Agustín Medina§

*School of Civil Engineering, Technical University of Catalonia, Barcelona, Spain*

SUMMARY

The inverse problem (also referred to as parameter estimation) consists of evaluating the medium properties ruling the behaviour of a given equation from direct measurements of those properties and of the dependent state variables. The problem becomes ill-posed when the properties vary spatially in an unknown manner, which is often the case when modelling natural processes. A possibility to fight this problem consists of performing stochastic conditional simulations. That is, instead of seeking a single solution (conditional estimation), one obtains an ensemble of fields, all of which honour the small scale variability (high frequency fluctuations) and direct measurements. The high frequency component of the field is different from one simulation to another, but a fixed component for all of them. Measurements of the dependent state variables are honoured by framing simulation as an inverse problem, where both model fit and parameter plausibility are maximized with respect to the coefficients of the basis functions (pilot point values). These coefficients (model parameters) are used for parameterizing the large scale variability patterns. The pilot points method, which is often used in hydrogeology, uses the kriging weights as basis functions. The performance of the method (both its variants of conditional estimation/simulation) is tested on a synthetic example using a parabolic-type equation. Results show that including the plausibility term improves the identification of the spatial variability of the unknown field function and that the weight assigned to the plausibility term does lead to optimal results both for conditional estimation and for stochastic simulations. Copyright © 2006 John Wiley & Sons, Ltd.

KEY WORDS: parameter estimation; parabolic equation; numerical modelling; groundwater

*Correspondence to: Andrés Alcolea, School of Civil Engineering, Technical University of Catalonia, Barcelona, Spain.
†E-mail: andres.alcolea@upc.edu
‡E-mail: jesus.carrera@upc.edu
§E-mail: agustin.medina@upc.edu

Contract/grant sponsor: ENRESA (Spanish Agency for Nuclear Waste Disposal)
Contract/grant sponsor: MEC (Spanish Ministry for Education and Science)

*Received 3 June 2005*
*Revised 30 January 2006*
*Accepted 5 February 2006*

Copyright © 2006 John Wiley & Sons, Ltd.

## INTRODUCTION

Parabolic equations represent natural diffusive phenomena and are used in many branches of engineering. Examples are the equations of heat conduction (industrial), groundwater flow (hydrogeology), or molecular diffusion (chemistry and contaminant transport), among others. Spatial variability of the properties (e.g. hydraulic conductivity for groundwater flow) entering those equations can be high but unknown, especially when they represent natural media. Moreover, it rules the performance of those equations [1]. Therefore, it has to be accounted for in meaningful models. Identification of the spatial variability is carried out in the context of inverse modelling [2–8] (also referred to as parameter estimation), which consists of estimating the properties as field functions from direct measurements of the properties (e.g. point values of thermal or hydraulic conductivity) and of dependent state variables (e.g. temperature and head for the heat transfer and groundwater flow equations, respectively).

The natural formulation of the inverse problem consists of assuming the state variable to be known and assume the field functions defining medium properties to be unknown. Such formulation is often ill-posed (i.e. a solution may not exist, it may not be unique, and it is usually unstable) for parabolic equations [9–11]. Moreover, it is not useful for practical purposes because the state variable is never known throughout the model domain. Therefore, one needs to parameterize the solution (i.e. to write the field functions in terms of a, hopefully small, number of parameters). Most parameterization techniques may be viewed as functional spaces where the parameters are the interpolation coefficients and the set of interpolation functions is a basis. A number of parameterization techniques have been used. Among them, the method of pilot points [12] has gained steam during recent years in hydrogeology because it is flexible and because it is formulated in a geostatistical context, so that it allows natural extensions to stochastic solutions of the governing equations. These are required when variability is important and unknown [1, 13].

The pilot points method consists of: (1) generating an initial spatially correlated field given a geostatistical model (i.e. measurements, if any, and correlation structure of the field function), (2) defining an interpolation method to obtain the value of the field functions over the model domain on the basis of their values at measurement and pilot point locations (model parameters) and (3) obtaining the value of the model parameters in such a way that the interpolated field (step 2) minimizes an objective function measuring the misfit between calculated and measured data (often, only state variables are considered). Thus, finding the optimum value of model parameters becomes an optimization problem. Note that step 3 implies the perturbation of the field generated in step 1.

As described above, the method suffers severe limitations. On the one hand, it is unstable, so that estimated pilot point values often reach non-plausible values. The inclusion of a regularization term to overcome this problem has been the subject of debate. Certes and de Marsily [14] reject the use of such a term, questioning its performance, because it depends to a large extent on the reliability of the prior estimates. RamaRao *et al.* [15] argue that the plausibility is achieved inherently, given that the initial field to be perturbed already honours (1) the available measurements of the field function and (2) the covariance structure describing the spatial variability patterns as observed in the field. Similar arguments are used by other researchers for rejecting the plausibility term [16–20]. Regularization has been used by Doherty [21]. Yet, his objective was to penalize non-homogeneity of the interpolated field rather than to include prior information about model parameters. Kowalsky *et al.* [22] used geophysical

Copyright © 2006 John Wiley & Sons, Ltd. 

measurements (Ground Penetrating Radar) in a maximum *a posteriori* geostatistical context. Recently, Alcolea *et al.* [23] proposed adding a plausibility term to the objective function, so as to penalize departures of pilot point values from their prior estimates obtained by kriging. They showed that the use of such a regularization term improves (1) the identification of spatial variability and (2) the stability of the problem, allowing the use of larger number of pilot points, thus sharpening the resolution of the spatial variability. However, they also found that including the plausibility term may lead to worse results than simply interpolating from measurements (i.e. not inverting at all) if the plausibility term is not properly weighted. Fortunately, the use of a maximum likelihood statistical framework allows the identification of the optimum weight of the plausibility term.

A second limitation of the pilot points method is related to the definition of the initial spatially correlated field: the original method of de Marsily obtained this field through conditional estimation (variants of kriging [24, 25]). This yields a single 'best' solution that minimizes the field variance and honours the available measurements of the field function. However, the resulting field is oversmoothed and does not allow a realistic representation of spatial variability. To overcome this problem, some authors [15–17, 26] included conditional simulations in the generation of the initial field, yielding a set of equally probable realizations of the field functions conditioned to available measurements. That is, each of these simulations reproduces the expected variability and honours measurements. Therefore, they can be used to evaluate the uncertainty of predictions. Unfortunately, these methods do not allow the inclusion of a plausibility term, so that they are essentially unstable.

The objective of this paper is to overcome the above limitations. Specifically, we seek a formulation of the inverse problem capable of generating equally probable simulations of the field functions (e.g. transmissivity field) that are conditioned to measurements of the medium properties and the state variables (e.g. heads). To this end, we extend the method of Alcolea *et al.* [23] to the case of conditional simulation. That is, we explore the possibility of using a regularization term in the case of seeking stochastic simulations of the properties conditioned upon point measurements of both those properties and dependent state variables.

This paper is organized as follows. First, the methodology is outlined. Second, a synthetic example using the parabolic groundwater flow equation is presented. The paper ends with a discussion of the results and some conclusions about the use of the plausibility term in the context of the pilot points method.

## METHODOLOGY

The proposed method is a modification of the pilot points method. Modifications include the use of a plausibility term and the way the vector of model parameters (value of the field functions at the pilot point locations) is updated through the optimization process. We assume that the (geostatistical) characteristics of the field functions are known. Here, the geostatistical model is defined by an autocorrelation function, but more sophisticated models may be used to represent complex heterogeneity patterns, geophysical data or known features. The procedure can be summarized as follows.

*1. Parameterization*: A field function $f$ is expressed as the superposition of two fields: a known drift $f_D(\mathbf{x}, t)$ and an uncertain residual $f_p(\mathbf{x})$, which is a linear combination of the

Copyright © 2006 John Wiley & Sons, Ltd. 

model parameters $p_j$ (Figure 1)

$$f(\mathbf{x},t) = f_D(\mathbf{x},t) + f_p(\mathbf{x}) \tag{1}$$

*1.1. Calculation of* $f_D(\mathbf{x},t)$: The drift can be obtained through conditional estimation (kriging/cokriging) or conditional simulation, depending on whether the modeller is seeking the characterization of large scale patterns or small scale variability, respectively. Therefore, it honours hard data (i.e. direct measurements of the field function $\mathbf{f}^*$) and possibly soft data $\mathbf{g}^*$ (correlated with $\mathbf{f}$), that can be included as external drifts. In the case of conditional simulation, $f_D(\mathbf{x},t)$ reproduces the spatial variability patterns as observed in the field (honours the correlation structure as well), if the geostatistical model defined previously is informative enough. For the simple case of linear interpolation, it can be expressed as

$$f_D(\mathbf{x},t) = \sum_{i=1}^{\dim Z} \lambda_i^Z(\mathbf{x}) Z(\mathbf{x}_i,t) \tag{2}$$

where $\mathbf{x}$ is the location where $f_D$ is calculated, $t$ and $\mathbf{x}_i$ are the measurement times and locations, respectively, and $\lambda_j^Z$ are the (co-) kriging weights for the measurements [25], organized in the vector $\mathbf{Z} = (\mathbf{f}^*, \mathbf{g}^*)$. Our implementation of the methodology allows a large set of conditional estimation/simulation methods: simple kriging, residual kriging, kriging with locally varying mean, kriging with external drift, simple cokriging, ordinary cokriging, ordinary cokriging (standardized to the mean value of the primary variable $f$), among the methods for conditional estimation [24], plus a sequential simulation algorithm for conditional simulations [27].

*1.2. Parameterization of the uncertain residual* $f_p(\mathbf{x})$: It can be viewed as the perturbation that the drift requires to honour measurements of dependent state variables. It is expressed as a linear combination of the model parameters (value of the field function at the pilot point locations)

$$f_p(\mathbf{x}) = \sum_{j=1}^{N_p} \lambda_j^{pp}(\mathbf{x}) p_j \tag{3}$$

where $N_p$ is the number of pilot points used to parameterize $f_p(\mathbf{x})$ (this number does not need to be the same for all field functions representing properties) and $\lambda_j^{pp}(\mathbf{x})$ are the (co-) kriging weights for the model parameters $p_j$, calculated in the same way as $\lambda_i^Z$ for measurements. In our implementation, the location of the pilot points can be fixed or vary randomly as the optimization process proceeds.

*2. Prior estimation*: Prior estimates of the pilot point values $\mathbf{p}^*$ and the corresponding *a priori* error covariance matrix $\mathbf{V}_p$ are obtained by conditional estimation/simulation to measurements in vector $\mathbf{Z}$. Note that correlation is included during the estimation process. In fact, pilot point values should be close (i.e. highly correlated) when pilot point locations are close. Therefore $\mathbf{V}_p$ is a full matrix.

*3. Objective function*: Using maximum likelihood estimation [28], the optimum set of model parameters minimize the objective function

$$F = \sum_{i=1}^{\text{nstat}} \beta_i (\mathbf{u}_i - \mathbf{u}_i^*)^t \mathbf{V}_{u_i}^{-1} (\mathbf{u}_i - \mathbf{u}_i^*) + \sum_{j=1}^{\text{ntypar}} \mu_j (\mathbf{p}_j - \mathbf{p}_j^*)^t \mathbf{V}_{p_j}^{-1} (\mathbf{p}_j - \mathbf{p}_j^*) \tag{4}$$

Copyright © 2006 John Wiley & Sons, Ltd. 

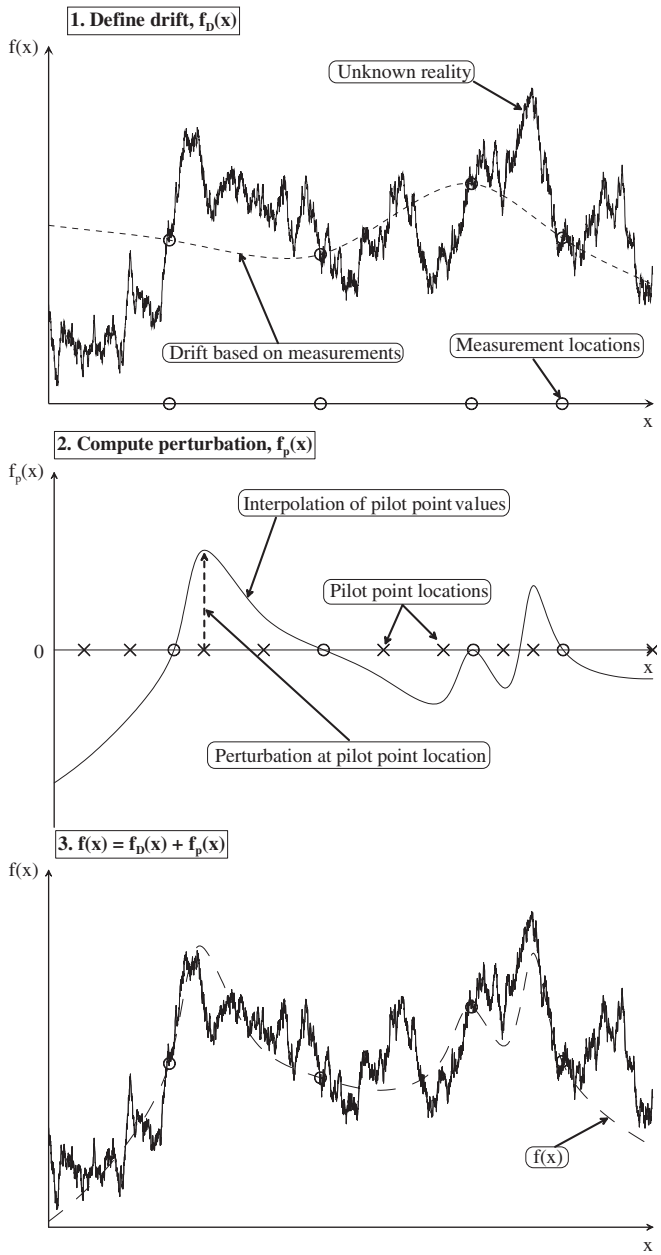


Figure 1. Schematic description of the pilot points method for defining a spatial random function $f(\mathbf{x})$, as the sum of a drift $f_D(\mathbf{x})$ and a perturbation $f_p(\mathbf{x})$. The drift is defined by conditional estimation (the smooth drift in the figure) or simulation (a sharper drift; not depicted) to available measurements. The perturbation is obtained from interpolation of the unknown pilot point values (model parameters), which are optimized so as to obtain a good fit with available indirect observations (e.g. measurements of the state variable).

Copyright © 2006 John Wiley & Sons, Ltd. 

where 'nstat' denotes number of state variables $\mathbf{u}_i$ with available measurements $\mathbf{u}_i^*$ (i.e. in groundwater, $i=1$ for heads/drawdowns, $i=2$ for concentrations, etc.); 'ntypar' is the number of types of model parameters being optimized, with prior information $\mathbf{p}_j^*$ (i.e. $j=1$ for pilot points linked to transmissivities, $j=2$ for storativities, etc.). Matrices $\mathbf{V}_{\mathbf{u}_i}$ and $\mathbf{V}_{\mathbf{p}_j}$ represent our best guess of the error covariance matrices of state variables and models parameters, respectively, and $\beta_i$, $\mu_j$ are weighting scalars correcting errors in the specification of $\mathbf{V}_{\mathbf{u}_i}$ and $\mathbf{V}_{\mathbf{p}_j}$. In our synthetic example, we solve the groundwater flow equation using only drawdown data as state variable for identifying the spatial variability of the transmissivity field. Thus, we will term hereinafter $F_d$ the term of state variables and $F_p$ the term of model parameters. Assuming that error covariance matrix of drawdown is correct ($\beta_1=1$), the simplified objective function can be written as

$$F = F_d + \mu F_p = (\mathbf{s}-\mathbf{s}^*)^t \mathbf{V}_s^{-1}(\mathbf{s}-\mathbf{s}^*) + \mu(\mathbf{p}-\mathbf{p}^*)^t \mathbf{V}_{\mathbf{p}}^{-1}(\mathbf{p}-\mathbf{p}^*) \tag{5}$$

While the objective function stated in Equation (4) (or its particularization in (5)) was originally based on favouring the best match of state variables ($F_d$), while ensuring stability and plausibility of the model parameters ($F_p$), it can also be derived in a statistical framework. Gavalas *et al.* [29] derived it by maximizing the posterior pdf (probability density function) of the model parameters, MAP, while Carrera and Neuman [9] arrived to it by maximizing the likelihood of the parameters given the data, MLE. Instead, Medina and Carrera [28] prefer working with the expected value of the likelihood function, as it allows the most stable estimation of statistical parameters, i.e. $\beta_i$ and $\mu_j$. Here, we use the same formulation.

*4. Minimization*: The minimization of Equation (4) is performed by means of Levenberg–Marquardt's method. This method belongs to the Gauss–Newton family and it consists of linearizing the dependence of state variables on model parameters, while imposing that the parameter change $\Delta\mathbf{p}^k$ at the $k$th iteration is constrained.

This leads to a linear system of equations [30–32]

$$(\mathbf{H}^k + \delta^k \mathbf{I})\Delta\mathbf{p}^k = -\mathbf{g}^k \tag{6}$$

where $\mathbf{H}^k$ is an approximation of the Hessian matrix of $F$ (Equation (4)) and $\mathbf{g}^k$ its gradient at $\mathbf{p}^k$ (vector of model parameters at iteration $k$), $\mathbf{I}$ is the identity matrix and $\delta^k$ is a positive scalar (the so-called Marquardt's parameter).

When the objective function takes a form similar to Equation (5), second-order derivatives of the state variable with respect to the parameters are often neglected, and the approximation of the Hessian matrix can be written as

$$\mathbf{H}^k = 2\mathbf{J}_s^t \mathbf{V}_s^{-1}\mathbf{J}_s + 2\mu\mathbf{V}_p^{-1} \tag{7}$$

where $\mathbf{J}_s$ is the jacobian matrix (i.e. derivatives of drawdowns with respect to model parameters at iteration $k$). The latter can be calculated by direct derivation of the PDE or by the adjoint state method [33]. The gradient of the objective function can be written as

$$\mathbf{g}^k = 2\mathbf{J}_s^t \mathbf{V}_s^{-1}(\mathbf{s}-\mathbf{s}^*) + 2\mu\mathbf{V}_p^{-1}(\mathbf{p}-\mathbf{p}^*) \tag{8}$$

Copyright © 2006 John Wiley & Sons, Ltd. 

*5. Updating the vector of model parameters*: After each iteration, the vector of model parameters is updated as

$$\mathbf{p}^{k+1} = \mathbf{p}^k + \mathbf{\Delta p}^k \tag{9}$$

Prior to updating, the components of vector $\mathbf{\Delta p}^k$ are examined. If any of them is larger than a given threshold, all of them are reduced accordingly. Thus, an upper bound (per iteration) limits the maximum step size.

Steps 4 and 5 are repeated until convergence, which is checked using the criteria of Medina *et al.* [34]: (a) the maximum increment of parameters (per iteration) is very small, (b) the change in the objective function between two consecutive iterations is negligible, (c) the gradient norm is very small or (d) the ratio between the gradient norm and its value at the first iteration is small enough. The algorithm also stops if the numbers of iterations or failed iterations (those increasing the objective function) reach threshold values. In our experience, (d) is possibly the best check of convergence and, in this work, a reduction factor of $10^6$ of the norm of the gradient was adopted as indicator of convergence.

To verify uniqueness, it is advisable to repeat the estimation starting from different initial values for model parameters. Starting from the drift (zero values to model parameters) is a good strategy. Starting from large values for pilot point perturbations usually leads to a good convergence. On the contrary, starting from values that are too low often leads to poor convergence.

*6. A posteriori statistical analysis*: The optimization process is repeated using different values of the weighting scalars $\beta_i$ and $\mu_j$, whose optimum values are the ones leading to the maximum of the expected likelihood, equivalent to the minimum of the support function [28]

$$S_2 = N + \ln|\mathbf{H}| + N\ln\left(\frac{F}{N}\right) - \sum_{j=1}^{\text{ntypar}} k_j \ln \mu_j \tag{10}$$

Here, $N$ is the total number of data, $k_j$ is the number of prior information data of the $j$th parameter type and $\mathbf{H}$ is the approximation of the Hessian matrix at the end of the optimization process.

Note that the methodology for variants of conditional estimation and conditional simulation only differs in the generation of the initial drift (step 1.1). This drift is unique in the case of conditional estimation and there are many realizations for conditional simulation. In the latter case, steps 2–6 must be repeated for each realization of the initial drift.

## APPLICATION

The objective of this example is to extend the results of the previous work of Alcolea *et al.* [23] to the case of conditional simulation, exploring the possibility of using a plausibility term.

Results are explored on the basis of a synthetic example, consisting of the simultaneous interpretation of three pumping tests ruled by the parabolic groundwater flow equation

$$\nabla(\mathbf{T}\nabla h) = S\frac{\partial h}{\partial t} \quad \text{on } \Omega \tag{11}$$

Copyright © 2006 John Wiley & Sons, Ltd. 

where $\Omega$ is the flow domain, $\mathbf{T}$ is the transmissivity tensor, $S$ is storativity and $h$ is head (the state variable). Initial and boundary conditions can be written as

$$\begin{aligned} h(t=0) &= h_0(\mathbf{x}) \quad \text{on } \Omega \\ \mathbf{T}\nabla h\mathbf{n} &= \alpha(H-h)+Q \quad \text{on } \Gamma \end{aligned} \tag{12}$$

where $\Gamma$ denotes the boundary of the flow domain, $\mathbf{n}$ is a unit vector normal to $\Gamma$ and pointing outwards, $H$ and $Q$ are prescribed heads and flows, respectively, and $\alpha$ is a coefficient controlling the type of boundary condition ($\alpha=0$ for prescribed flow, $\alpha\to\infty$ for prescribed head and a mixed condition otherwise). When pumping tests are interpreted, it is useful to work with drawdowns (difference between head in presence of pumping and head in absence of pumping), denoted as 's' hereinafter. This leads to homogeneous (zero) initial and boundary drawdowns, as well as boundary flow rates. Equation (11) is solved applying the Galerkin method in space and forward finite differences in time. Elements are quadrangular bilinear.

The flow domain is a square of $400\times400\,\mathrm{m}^2$, despite it is enlarged to avoid spurious boundary effects to a global domain of $3600\times3600\,\mathrm{m}^2$ (Figure 2(a)). The finite element grid is more refined in the central part (zone of interest, Figure 2). There, the finite element mesh is structured. Outside, the element size increases as the mesh progresses towards the boundary domain (Figure 2(a)).

The 'true' log transmissivity field ($\log_{10}T$ hereinafter) was generated with the code TRANSIN [34] by sequential simulation (Figure 2(a)) conditional to a set of measurements defining two channels of high transmissivity. The 'true' variogram (field variance minus autocorrelation function) is spherical, with a range of 200 m and a variance of 2, without nugget effect. Values of the 'true' $\log_{10}T$ field range from $-9.1$ to 0.5, with a mean value of $-4[\log_{10}(\mathrm{m}^2/\mathrm{s})]$. In this work, only the heterogeneity of the $\log_{10}T$ field was explored. Storativity was assumed to be constant and known over the whole domain, with a value of $10^{-4}$.

Thirteen measurements of $\log_{10}T$ were selected from the 'true' field as conditioning data. Measurement locations were purposefully selected in such a way that the initial drift of Equation (2) (calculated by ordinary kriging or by sequential simulation, for the cases of conditional estimation/simulation, respectively) was radically different from the 'true' field (Figure 2(b)). Note that, indeed, the high $\log_{10}T$ channels crossing the zone of interest are missed by the drift. Thus, the performance of the model is heavily dependent on the calibration of the perturbation field $f_p$. This setup was chosen to ensure that the plausibility term, which biases the solution towards the drift, would hinder finding a good solution.

Drawdown data comes from three independent pumping tests (but analysed simultaneously) in the most productive wells of the central domain (pumping rates of $10^{-2}\,\mathrm{m}^3/\mathrm{s}$ at wells B1, B2 and B3 in Figure 2). Transient drawdowns were simulated at grid nodes (Figure 3), assuming a zero drawdown as initial condition and prescribed at the boundaries of the global domain. Drawdown measurements were calculated at the thirteen points where $\log_{10}T$ measurements are available (total of 936 drawdown data). A Gaussian white noise was added to those measurements, simulating acquisition errors, with a standard deviation of 0.3 m for pumping at wells B1 and B2 and 0.15 m at well B3 (1% of the maximum drawdown at each one of the tests).

In the previous work of Alcolea *et al.* [23], the optimum weight of the plausibility term was found for the case of conditional estimation. For the purpose of this paper, we take

Copyright © 2006 John Wiley & Sons, Ltd. 

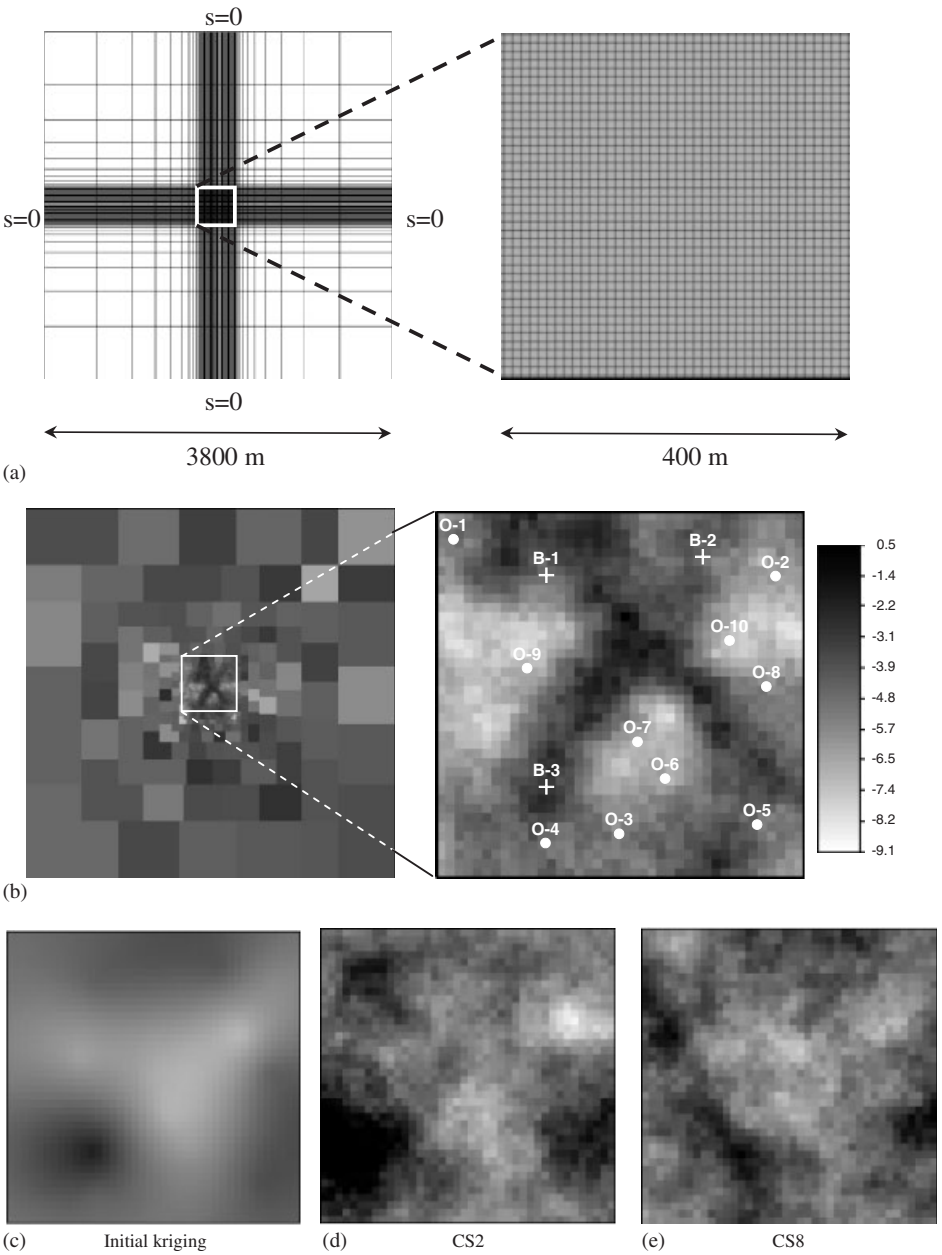


Figure 2. Test problem description: (a) finite element mesh and flow domain; (b) true $\log_{10} T$ field and location of conditioning measurements. All boundaries have a prescribed drawdown condition (zero). White square limits the zone of interest, where three pumping tests are performed independently at points B-1, B-2 and B-3. Below, initial drifts, obtained by kriging of the thirteen $\log_{10} T$ measurements (c) and by conditional simulation (cases listed in Table I). Note that they are radically different from the 'true' field depicted above.

Copyright © 2006 John Wiley & Sons, Ltd. 

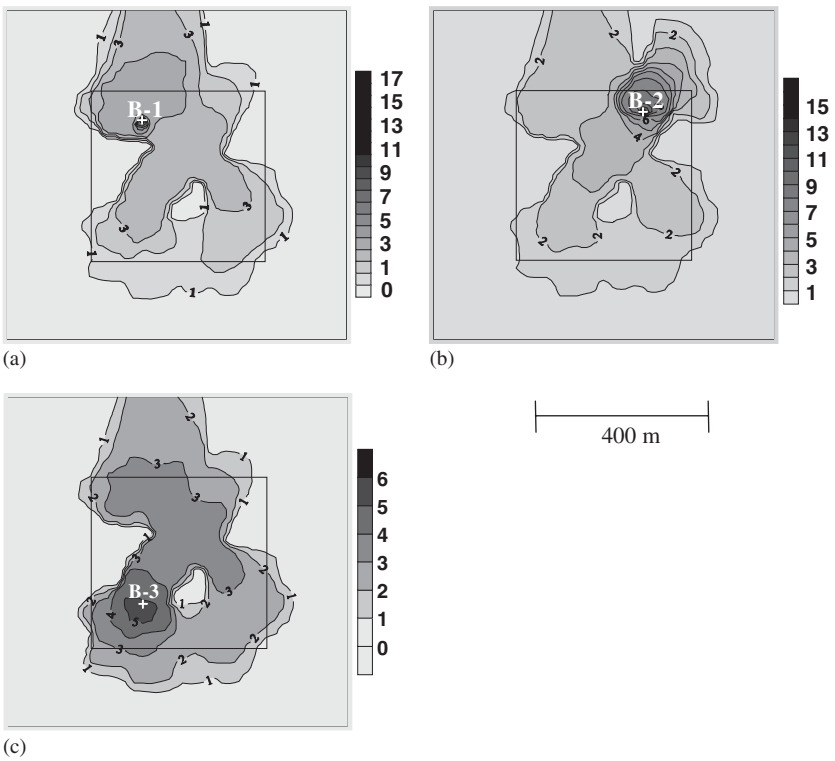


Figure 3. 'True' drawdown after pumping ($t = 7200\,\mathrm{s}$) at wells: (a) B-1; (b) B-2; and (c) B-3. The zone of interest (central square of $400 \times 400\,\mathrm{m}^2$) has been enlarged two hundred meters each side.

as starting point the conditional estimations using 97 pilot points located in a regular grid (Figure 5; row 1, column 3) and explore the optimum weight of the plausibility term when the initial correlated fields are drawn by conditional simulations (a total of 10 realizations). We use values of the weighting factor ranging from $10^{-1}$ to $10^{2}$. This range of values was selected by taking into account that the optimum value of $\mu$ should be one if the variogram is error-free ($\log_{10} T$ variogram used in the calibrations was the 'true' one). High values of $\mu$ give too much weight to the plausibility term. This should result in a poor identification of the spatial variability, as the field would be biased towards the initial drift (Figure 2(b)). On the contrary, small values of $\mu$ tend to disregard the plausibility term, thus risking instability.

## RESULTS

Results are explored in the same way as in the previous work [23]: qualitatively ($\log_{10} T$ maps and drawdown fits) and quantitatively. For the latter, an error vector **e** is defined as the difference between calculated and 'true' values of $\log_{10} T$ at the central part of the domain

Copyright © 2006 John Wiley & Sons, Ltd. 

(1600 blocks of $10 \times 10\,\mathrm{m}^2$). We analyse the following statistics:

(1) Total objective function and its drawdowns and parameters components ($F$, $F_d$ and $F_p$ in Equation (5), respectively). These are not good comparison criteria as they grow ($F$ and $F_d$) or decrease monotonically ($F_p$) with $\mu$.
(2) Support function of the expected likelihood (Equation (10)), whose minimization should lead to the optimum value of $\mu$.
(3) Mean error: measures the match between calculated and 'true' values of $\log_{10} T$

$$\bar{\mathrm{e}}_{\log_{10} T} = \frac{1}{1600} \sum_{i=1}^{1600} |e_i| = \frac{1}{1600} \sum_{i=1}^{1600} |\log_{10} T^{\mathrm{calc}} - \log_{10} T^{\mathrm{true}}| \tag{13}$$

We used this criterion rather than the raw one measuring the estimation bias (identical but without absolute value), given that the latter, also evaluated, was close to zero in most cases, as expected. Therefore, it did not shed new light on this research.
(4) Root mean square error of $\log_{10} T$: this is the basic raw criterion to evaluate the goodness of the identification. Theoretically, it should be smaller than the *a priori* deviation (square root of the variogram sill, $\sqrt{2}$ in this case), if conditioning is good

$$\mathrm{RMSE}_{\log_{10} T} = \left( \frac{1}{1600} \mathbf{e}^t \mathbf{e} \right)^{1/2} \tag{14}$$

As will be discussed later, mean error and root mean square error are very sensitive to the location and extreme values of the zones of high/low transmissivity.

Table I displays a comparison between the evolution of the statistics with the weight of the plausibility term for the conditional estimation and two out of ten conditional simulations. Table II summarizes the quantitative comparisons and contains the value of $\mu$ for which the estimation statistics reach their optimum value. For instance, the minimum value of $\mathrm{RMSE}_{\log_{10} T}$ is attained at $\mu = 0.1$ in simulation 5. Figure 4 displays the quantitative comparison in terms of the support function of the expected likelihood $S_2$ and the estimation errors, $\mathrm{e}_{\log_{10} T}$ and $\mathrm{RMSE}_{\log_{10} T}$. Identifications of $\log_{10} T$ are presented in Figure 5. Figure 6 displays the best matching of drawdown.

The first observation that becomes apparent from Table I is the strong effect of the dependence of the plausibility term, as occurred in the previous work. The relative importance given to this term is measured by the value of the weighting factor $\mu$. Using small values for this factor (little importance of the plausibility term, disregarding prior estimates in the optimization process and therefore, prior information) consistently leads to the best fit of drawdowns (minimum value of $F_d$) and to the worst fit of model parameters (maximum value of $F_p$), as expected. Identifications of the spatial variability (Figure 5; last row, column 2) using a weighting factor of $10^{-1}$ offer a somewhat 'lumpy' appearance, with large jumps in the calibrated transmissivity over small distances. In fact, when the drift is generated by conditional estimation, Alcolea *et al.* [23] found the optimum identification when $\mu$ equals 0.1 (the minimum value tested in this example), that yields the worst qualitative identification of the $\log_{10} T$ field in this work. In fact, values of $\mu$ lower than $10^{-1}$ were excluded from this study due to instability problems.

Similarly, large values of the weighting factor also yield poor results. The final solution tends to be close to the initial drift (Figure 5, first row), which contains little information

Copyright © 2006 John Wiley & Sons, Ltd. 

Table I. Summary of results of the sensitivity analysis to the weighting factor $\mu$, for the conditional estimation CE (from Reference [23]) and two out of ten conditional simulations CS. Minimum values for each set are written in italic characters.

| Test problem | | Objective function (Equation (5)) | | | | Estimation errors | |
|---|---|---|---|---|---|---|---|
| | Weighting factor $\mu$ | Total obj. func. ($F$) | Drawdown obj. func. ($F_d$) | Parameters obj. func. ($F_p$) | $S_2$ (Equation (8)) | $\bar{e}_{\log 10T}$ (Equation (11)) | $\mathrm{RMSE}_{\log_{10} T}$ (Equation (12)) |
| CE | $\mu \to \infty$ | $1.156 \times 10^6$ | $1.156 \times 10^6$ | — | — | 1.390 | 1.831 |
| | $10^2$ | 17426 | 5566 | *119* | 4018 | 1.525 | 2.081 |
| | $10^1$ | 3070 | 1318 | 175 | 2267 | 1.408 | 2.001 |
| | $10^0$ | 1033 | 829 | 203 | 1205 | 1.025 | 1.456 |
| | $3 \times 10^{-1}$ | 875 | 784 | 302 | 1074 | *0.950* | 1.386 |
| | $10^{-1}$ | 787 | 753 | 348 | *1007* | 0.961 | *1.331* |
| | $10^{-2}$ | 759 | 744 | 1501 | 1075 | 1.431 | 2.080 |
| | $10^{-3}$ | *741* | *737* | 3690 | 1214 | 2.016 | 2.938 |
| CS 2 | $\mu \to \infty$ | $7.677 \times 10^5$ | $7.677 \times 10^5$ | — | — | 1.71 | 2.20 |
| | $10^2$ | 8467 | 2856 | *56* | 3193 | 1.37 | 1.82 |
| | $10^1$ | 1621 | 917 | 70 | 1536 | 1.73 | 2.29 |
| | $10^0$ | 901 | 784 | 117 | 998 | *1.03* | *1.42* |
| | $10^{-1}$ | *773* | *747* | 253 | *958* | 1.38 | 1.89 |
| CS 8 | $\mu \to \infty$ | $1.233 \times 10^6$ | $1.233 \times 10^6$ | — | — | 1.64 | 2.12 |
| | $10^2$ | 9294 | 2670 | *66* | 3286 | 1.36 | 1.89 |
| | $10^1$ | 2117 | 1095 | 102 | 1804 | 1.40 | 1.94 |
| | $10^0$ | 912 | 783 | 129 | 1013 | *1.15* | *1.57* |
| | $10^{-1}$ | *774* | *750* | 244 | *933* | 1.22 | 1.67 |

Table II. Values of the weighting factor $\mu$ for which the estimation errors are minima (CE and CS denote conditional estimation and conditional simulation, respectively). The tested values of $\mu$ were $10^{-1}$, $10^0$, $10^1$ and $10^2$. The optimum identification of the spatial variability, as measured by the support function of the expected likelihood $S_2$ (Equation (10)) is always attained when $\mu$ equals 0.1.

| Test problem | $\bar{e}_{\log_{10} T}$ (Equation (13)) | $\mathrm{RMSE}_{\log_{10} T}$ (Equation (14)) |
|---|---|---|
| CE | $10^{-1}$ | $10^{-1}$ |
| CS 1 | $10^0$ | $10^0$ |
| CS 2 | $10^0$ | $10^0$ |
| CS 3 | $10^{-1}$ | $10^{-1}$ |
| CS 4 | $10^0$ | $10^0$ |
| CS 5 | $10^{-1}$ | $10^{-1}$ |
| CS 6 | $10^2$ | $10^2$ |
| CS 7 | $10^{-1}$ | $10^{-1}$ |
| CS 8 | $10^0$ | $10^0$ |
| CS 9 | $10^0$ | $10^0$ |
| CS 10 | $10^2$ | $10^2$ |

Copyright © 2006 John Wiley & Sons, Ltd. 

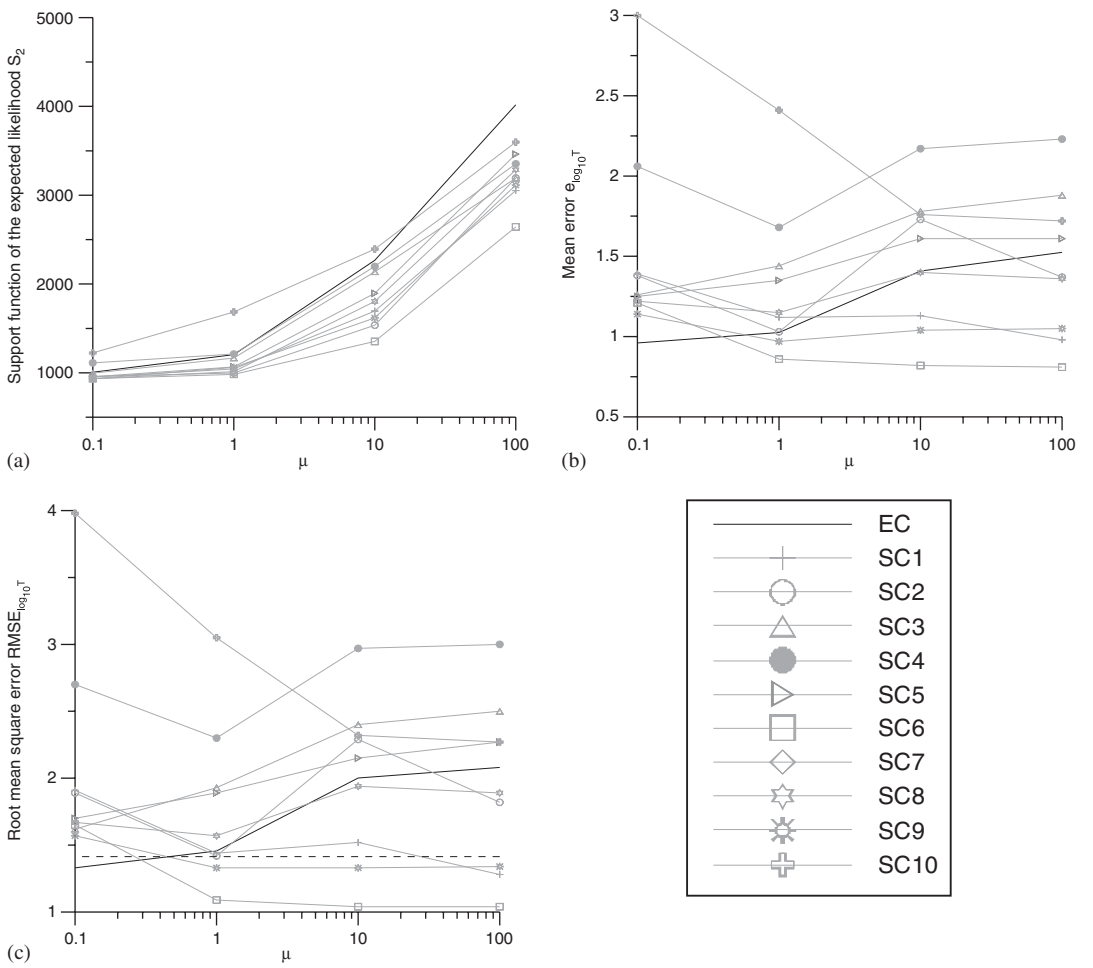


Figure 4. Support function of the expected likelihood and estimation errors versus $\mu$ and: (a) support function $S_2$; (b) mean error $\bar{e}_{\log_{10}T}$; and (c) $\mathrm{RMSE}_{\log_{10}T}$ (dashed horizontal line displays theoretical threshold value of $\sqrt{2}$).

about the actual variability of the 'true' field. However, estimation errors are sometimes smaller when $\mu$ equals $10^2$ than the ones in the case of $10^1$ (see CS2 in Table I and Figures 4(b) and (c)). We attribute this effect to the sensitivity of the estimation errors to the geometrical definition and extreme values of the high transmissivity channels (i.e. a small error in the position or the inclination of the channels may lead to large values of the estimation errors). As displayed in Figure 5 (column 2), the identification of the $\log_{10}T$ field in row 2 ($\mu = 10^2$) is worse than the one in row 3 ($\mu = 10^1$), although its estimation errors are smaller. This effect is not reproduced for the case of conditional estimation.

Optimum identifications, as measured by criteria $S_2$, are obtained when $\mu$ equals $10^{-1}$ in the ten conditional simulations, the same result attained by Alcolea *et al.* [23] for conditional

Copyright © 2006 John Wiley & Sons, Ltd. 

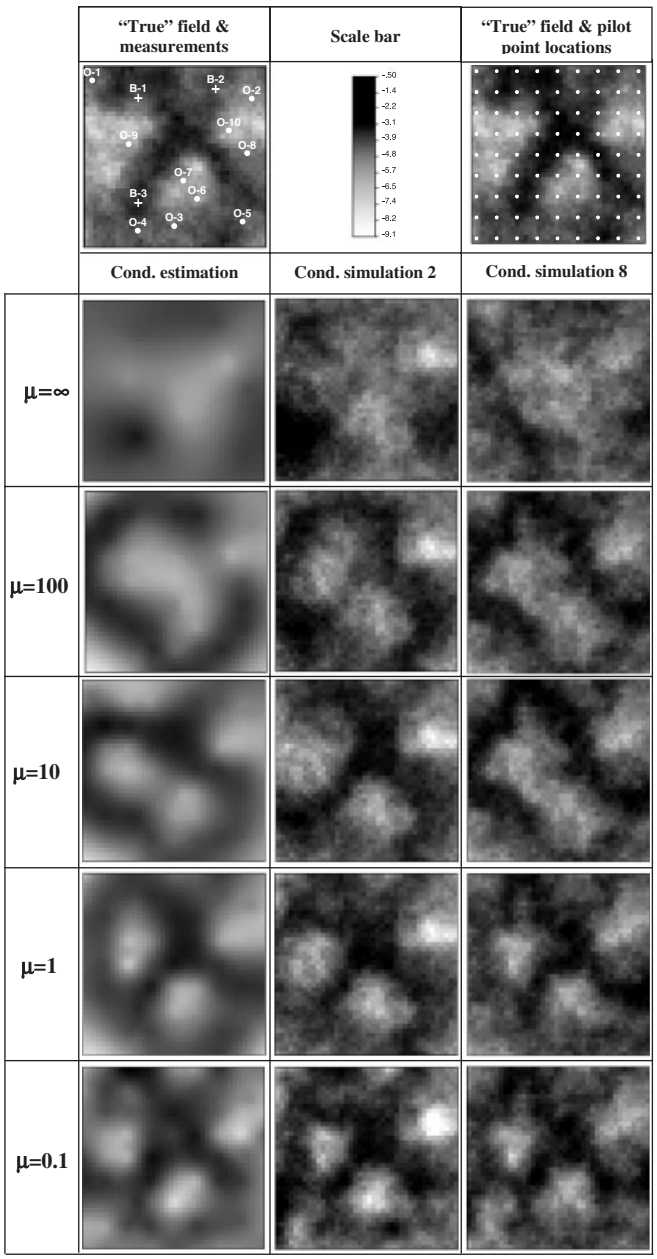


Figure 5. Qualitative comparison of the results: Row 1. 'True' field, $\log_{10} T$ conditioning measurements, common scale bar and situation of pilot points. Row 2: drifts to be perturbed (column 1 obtained by ordinary kriging of the $\log_{10} T$ measurements; columns 2 and 3 by sequential simulation). Rows 3–6 display the results after conditioning to $\log_{10} T$ and drawdown measurements with varying weight $\mu$. The 'true' field is resembled when optimum weight is assigned, as measured by $S_2$ (in all cases, when $\mu$ equals 0.1). Conditional estimation resembles the large scale patterns of the 'true' field, despite the identifications of the spatial variability are oversmoothed.

Copyright © 2006 John Wiley & Sons, Ltd. 

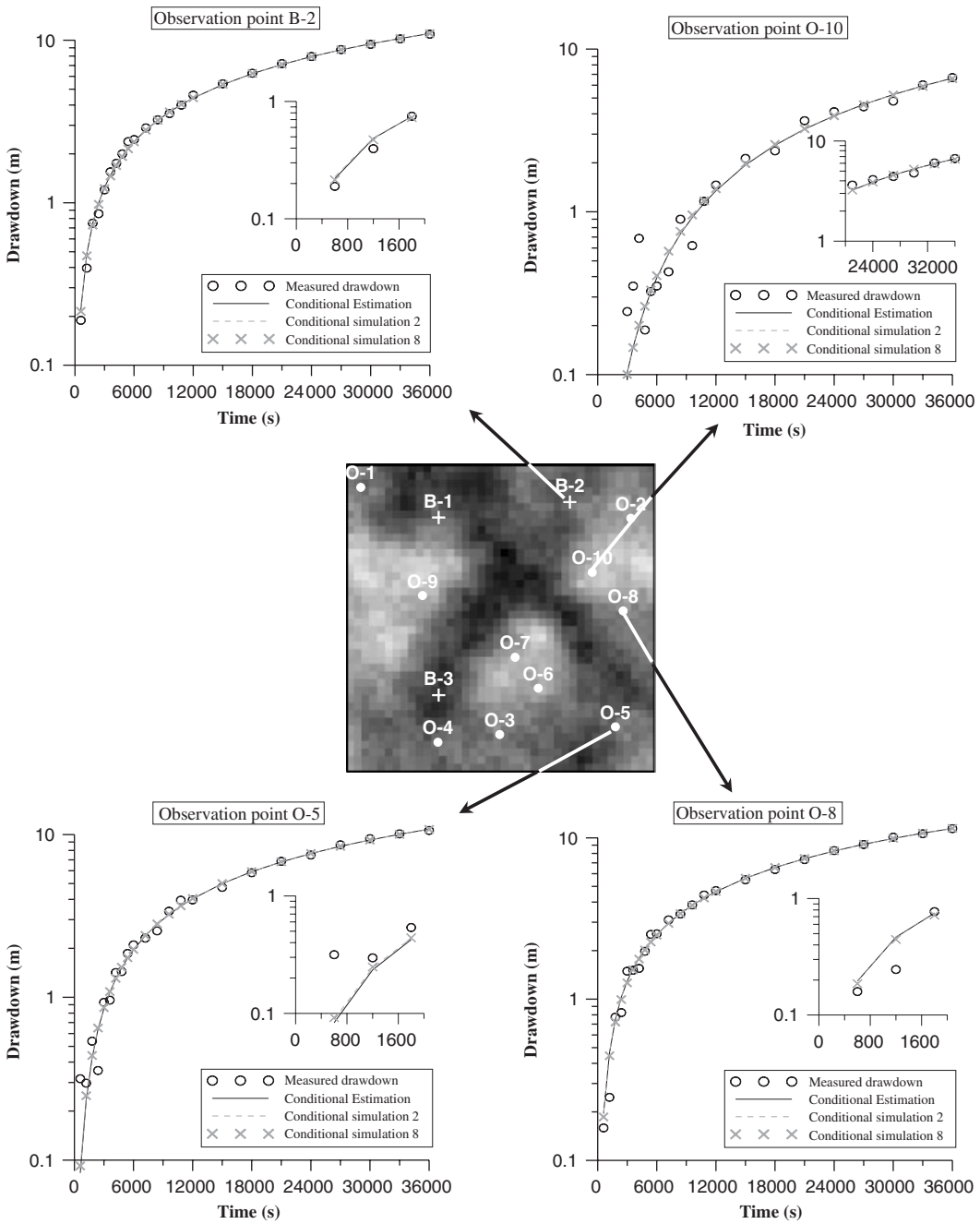


Figure 6. Time evolution of measured (circles) and computed drawdowns in response to pumping at B-3 at selected observations points. Results of conditional estimation (black line) and two of the conditional simulations are presented. Note that the fit of drawdown data is almost the same in the three cases.

Copyright © 2006 John Wiley & Sons, Ltd. 

estimation. This is important because it suggests that the modeller does not need to identify the optimum weight for each conditional simulation (usually a large number), but to obtain it just once using the method in its variant of conditional estimation.

Another result shared by conditional estimation and simulation is that, if the plausibility term is not properly weighted, the identification of the spatial variability is worse than the initial drift, as measured by estimation errors (Table I). However, the use of the methodology in a maximum likelihood framework allows the estimation of the weighting factor $\mu$. Therefore, the use of a plausibility term is advisable, independently of how the drift was calculated.

Figure 6 displays the best drawdown fit ($\mu$ equals 0.1) for the conditional simulations at Table I and the conditional estimation. Calculated drawdowns are very similar in all cases and fit the data. In fact, drawdown objective functions were very similar in all cases (Table I). Despite the best match to drawdown data is obtained when the plausibility term is neglected, the drawdown fits for the optimum identification of the $\log_{10} T$ field (optimum weighting scalar $\mu$) were nearly as good as the best ones ($\mu \to 0$) and the simulation was stable.

## CONCLUSIONS

The pilot points method provides a powerful tool for the identification of the field functions ruling the behaviour of a PDE. The suggested approach includes a plausibility term in the optimization process and two ways for calculating the initial drift, by conditional estimation or simulations conditioned to the direct measurements of the field function. Conditional estimation leads to optimal results (i.e. minimum estimation errors) but fails to reproduce small scale variability, which may be important when using the model for predictions. Instead, conditional simulation seeks a set of equally likely realizations of the field conditioned to all available information. Both variants have been tested on a synthetic example using the parabolic groundwater flow equation, examining the role of the plausibility term.

We have found that, neglecting the plausibility term, which is the standard approach in the context of pilot points, favours the best match of drawdown data, but often leads to an unstable identification of the model parameters. Large values give too much importance to the plausibility term, which biases the solution towards the drift. If the geostatistical model contains little information of the variability patterns (as in our synthetic example), the solution yields poor identifications of the spatial variability. In fact, a disturbing finding is that, in most cases, conditioning the fields to state variable data worsens the results if the plausibility term is not weighted properly. Fortunately, the use of a statistical framework (maximum likelihood in this case) allows the estimation of the optimum weight of the plausibility term, and therefore, its use is recommended. However, to search this optimum weight for each conditional simulation can be tedious.

A key finding of this work is that the optimum value of the weighting factor (as measured by the support function of the expected likelihood) was the same as the one obtained using conditional estimation. This frees the modeller of the burden of having to seek the optimum weight at each conditional simulation (usually a large number).

Good fits to measured state variable were obtained when neglecting (assigning low weight to) prior information. Still, fits nearly as good were obtained with stable simulations when moderate weights were assigned to prior information. We stress that the prior information provides a valuable data of aquifer heterogeneity, even when it is poorly informative of the

Copyright © 2006 John Wiley & Sons, Ltd. 

actual variability. Thus, the use of a plausibility term including it (usually disregarded in the context of pilot points) needs to be considered.

ACKNOWLEDGEMENTS

This work was funded by ENRESA (Spanish Agency for Nuclear Waste Disposal) and MEC (Spanish Ministry for Education and Science).

REFERENCES

1. Freeze RA. A stochastic-conceptual analysis of one-dimensional groundwater flow in nonuniform homogeneous media. *Water Resources Research* 1975; **11**(5):725–741.
2. Carrera J, Alcolea A, Medina A, Hidalgo J, Slooten LJ. Inverse problem in hydrogeology. *Hydrogeology Journal* 2004; **13**(1):206–222.
3. Carrera J. State of the art of the inverse problem applied to the flow and solute transport problems. *Groundwater Flow and Quality Modeling*, *NATO ASI Series* 1987; 549–585.
4. Ouyang T. Analysis of parameter-estimation heat-conduction problems with phase-change using the finite-element method. *International Journal for Numerical Methods in Engineering* 1992; **33**(10):2015–2037.
5. Schnur DS, Zabaras N. An inverse method for determining elastic-material properties and a material interface. *International Journal for Numerical Methods in Engineering* 1992; **33**(10):2039–2057.
6. de Marsily G, Delhomme JP, Delay F, Buoro A. 40 years of inverse problems in hydrogeology. *Comptes Rendus de l'Academie des Sciences Series IIA Earth and Planetary Science* 1999; **329**(2):73–87.
7. McLaughlin D, Townley LLR. A reassessment of the groundwater inverse problem. *Water Resources Research* 1996; **32**(5):1131–1161.
8. Yeh WWG. Review of parameter estimation procedures in groundwater hydrology: the inverse problem. *Water Resources Research* 1986; **22**(2):95–108.
9. Carrera J, Neuman SP. Estimation of aquifer parameters under transient and steady-state conditions. 1. Maximum likelihood method incorporating prior information. *Water Resources Research* 1986; **22**(2):199–210.
10. Carrera J, Neuman SP. Estimation of aquifer parameters under transient and steady-state conditions. 2. Uniqueness, stability and solution algorithms. *Water Resources Research* 1986; **22**(2):211–227.
11. Carrera J, Neuman SP. Estimation of aquifer parameters under transient and steady-state conditions, 3. Application to synthetic and field data. *Water Resources Research* 1986; **22**(2):228–242.
12. de Marsily GH, Lavedan G, Boucher M, Fasanino G. Interpretation of interference tests in a well field using geostatistical techniques to fit the permeability distribution in a reservoir model. In *Geostatistics for Natural Resources Characterization*. Part 2, Verly D *et al*. (eds). D. Reidel Publication Co.: Dordrecht, 1984; 831–849.
13. Narayanan VAB, Zabaras N. Stochastic inverse heat conduction using a spectral approach. *International Journal for Numerical Methods in Engineering* 1992; **60**(9):1569–1593.
14. Certes C, de Marsily GH. Application of the pilot point method to the identification of aquifer transmissivities. *Advances in Water Resources* 1991; **14**(5):284–300.
15. RamaRao BS, Lavenue M, de Marsily GH, Marietta MG. Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields. 1. Theory and computational experiments. *Water Resources Research* 1995; **31**(3):475–493.
16. Capilla JE, Gómez-Hernández JJ, Sahuquillo A. Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data. 2. Demonstration on a synthetic aquifer. *Journal of Hydrology* 1997; **203**:175–188.
17. Gómez-Hernández JJ, Sahuquillo A, Capilla JE. Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data. 1. Theory. *Journal of Hydrology* 1997; **204**(1–4):162–174.
18. Lavenue M, RamaRao BS, de Marsily GH, Marietta MG. Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields. 2. Application. *Water Resources Research* 1995; **31**(3):495–516.
19. Lavenue M, de Marsily GH. Three-dimensional interference test interpretation in a fractured aquifer using the pilot-point inverse method. *Water Resources Research* 2001; **37**(11):2659–2675.
20. Wen XH, Deutsch CV, Cullick AS. Construction of geostatistical aquifer models integrating dynamic flow and tracer data using inverse technique. *Journal of Hydrology* 2002; **255**:151–168.
21. Doherty J. Groundwater model calibration using pilot points and regularization. *Ground Water* 2003; **41**(2): 170–177.
22. Kowalsky MB, Finsterle S, Rubin Y. Estimating flow parameter distributions using ground-penetrating radar and hydrological measurements during transient flow in the vadose zone. *Advances in Water Resources* 2004; **27**:583–599.

Copyright © 2006 John Wiley & Sons, Ltd. 

23. Alcolea A, Carrera J, Medina A. Pilot points method incorporating prior information for solving the groundwater flow inverse problem. *Advances in Water Resources* 2006, in press.
24. Gu L. Moving kriging interpolation and element-free Galerkin method. *International Journal for Numerical Methods in Engineering* 2003; **56**(1):1–11.
25. Krige DG. A statistical approach to some mine evaluation and allied problems on the Witwatersrand. *Ph.D. Thesis*, University of the Witwatersrand, Johannesburg, 1951.
26. Hendricks-Franssen HJ. Inverse stochastic modelling of groundwater flow and mass transport. *Ph.D. Thesis*, Technical University of Valencia, Spain, 2001.
27. Liu YH, Journel A. Improving sequential simulation with a structured path guided by information content. *Mathematical Geology* 2004; **36**(8):945–964.
28. Medina A, Carrera J. Geostatistical inversion of coupled problems: dealing with computational burden and different types of data. *Journal of Hydrology* 2003; **281**:251–264.
29. Gavalas GR, Shaw PC, Seinfield JH. Reservoir history matching by Bayesian estimation. *Society of Petroleum Engineers Journal* 1976; **261**:337–350.
30. Marquardt DW. An algorithm for least-squares estimation of non-linear parameters. *Journal of the Society for Industrial and Applied Mathematics* 1963; **11**(2):431–441.
31. Cooley RL. Regression modeling of groundwater flow. *USGS Open Report* 1985; 85–180.
32. Bereaux Y, Clermont JR. Numerical-simulation of complex flows of non-newtonian fluids using the stream tube method and memory integral constitutive-equations. *International Journal for Numerical Methods in Fluids* 1995; **21**(5):371–389.
33. Galarza GA, Carrera J, Medina A. Computational techniques for optimization of problems involving non-linear transient simulations. *International Journal for Numerical Methods in Engineering* 1999; **45**(3):319–334.
34. Medina A, Alcolea A, Carrera J, Castro LF. Modelos de flujo y transporte en la geosfera: Código Transin IV. In *IV Jornadas de Investigación y Desarrollo Tecnológico de Gestión de Residuos Radiactivo de ENRESA. Publicación técnica 9/2000*, 2000; 195–200.

Copyright © 2006 John Wiley & Sons, Ltd. 