Uncertainty assessment and data worth in groundwater flow and mass transport modeling using a blocking Markov chain Monte Carlo method [1]

Jianlin Fu [2]
J. Jaime Gómez-Hernández [3]

Department of Hydraulic and Environmental Engineering
Universitat Politècnica de València
Camino de Vera, s/n, 46022, Valencia, Spain
Tel. (34) 963879615
Fax. (34) 963877618

Appear in
Journal of Hydrology, 364, 328-341, 2009.
DOI: 10.1016/j.jhydrol.2008.11.014

[1]The final version for personal research use is available from fu_jianlin_ac@yahoo.com.

[2]Corresponding author. Currently with Department of Energy Resources Engineering, Stanford University, 367 Panama Street, Stanford, CA 94305, USA. E-mail: fu_jianlin_ac@yahoo.com or jianfu@dihma.upv.es

[3]E-mail: jaime@dihma.upv.es

Contents

Abstract

Groundwater flow and mass transport predictions are always subject to uncertainty due to the scarcity of data with which models are built. Only a few measurements of aquifer parameters, such as hydraulic conductivity or porosity, are used to construct a model, and a few measurements on the aquifer state, such as piezometric heads or solute concentrations, are employed to verify/calibrate the goodness of the model. Yet, at unsampled locations, neither the parameter values nor the aquifer state can be predicted (in space and/or time) without uncertainty. We demonstrate the applicability of a new blocking Markov chain Monte Carlo (BMcMC) algorithm for uncertainty assessment using, as a reference, a synthetic aquifer in which all parameter values and state variables are known. We also analyze the worth of different types of data for the characterization of the aquifer and for reduction of uncertainty in parameters and variables. The BMcMC method allows the generation of multiple plausible representations of the aquifer parameters, and their corresponding aquifer state, honoring all available information on both parameters and state variables. The realizations are also coherent with an *a priori* statistical model for the spatial variability of the aquifer parameters. BMcMC is capable of direct-conditioning (on model parameter data) and inverse-conditioning (on state variable data). We demonstrate the flexibility of BMcMC to inverse-condition on piezometric head data as well as on travel time data, what permits identification of the impact that each data type has on the uncertainty about hydraulic conductivity, piezometric head and travel time.

Keywords: Inverse modeling, Bayesian modeling, heterogeneity, stochastic hydrogeology, conditional simulation, temporal moments

1 Introduction

There is an increasing interest in evaluating the uncertainty associated with groundwater flow and mass transport predictions. Scarcity of data leads to building models that, at most, can reproduce the aquifer state at a few sampling locations. The parameters controlling the aquifer response vary in space in a non-deterministic way, generally following some kind of structural pattern overlaid by an erratic component. Scarcity of data and spatial variability makes for many alternative combinations of those parameters that could yield a model that reproduces the state measurements, i.e., piezometric heads or concentrations. Within this context, we propose to use stochastic inverse conditioning Monte Carlo methods to evaluate aquifer prediction uncertainty. Stochastic, because we will build a random function model to represent all equally likely representations of the aquifer parameter distribution; inverse, because in building such a random function model, we will use both parameter measurements (i.e., conductivity data) and aquifer state measurements (i.e., piezometric data); and conditioning, because we will force that, within measurement error, each member of the random function model will reproduce the parameter measurements, and the solution of the state equation in each member will reproduce the state measurement data.

This type of analysis has been carried out before. The originality of this work is the use of a new algorithm, based on Monte Carlo Markov chains, for the generation of the aquifer parameter realizations. As a result we will demonstrate that the new algorithm is capable of performing stochastic inverse conditioning Monte Carlo as stated above, and we will analyze the worth of different types of data for the characterization of aquifer heterogeneity and the evaluation of the uncertainty in model predictions.

The challenge we faced is the generation of realizations of conductivity following a given statistical structure, specifically, a multiGaussian random function model, conditioned to conductivity data, and "inverse conditioned" to piezometric head and travel time data. The inverse conditioning aspect of the new algorithm is our main contribution.

Probably, the most difficult part in hydrogeological inverse modeling is conditioning to mass transport information. Concentration data have been directly used for parameter inference, e.g., characterization on the spatial correlation structure and the point estimation on local parameters, by several authors (i.e., *Graham and McLaughlin*, 1989a, 1989b; *Sun and Yeh*, 1990a, 1990b; *Woodbury and Sudicky*, 1992; *Deng et al.*, 1993; *Anderman and Hill*, 1999; *Nowak and Cirpka*, 2006). *Ezzedine and Rubin* (1996) derived, using a geostatistical approach, the cross covariances between the tracer concentration data and both conductivity and head, which allows using tracer data to estimate the spatial distribution of conductivity. *Franssen et al.* (2003) used the sequential self-calibration method (*Gómez-Hernández et al.*, 1997) to generate realizations conditional to the spatially distributed concentration data with the aid of the adjoint-state method to calculate the sensitivity matrix.

Travel time data are a cheaper alternative to concentration data due to the low cost in the data acquisition. Indeed, collecting travel time of tracers only invokes couples of wells for forced-gradient flow or a series of wells distributed along a plane perpendicular to

the mean flow direction for natural-gradient flow (*Rubin and Ezzedine*, 1997). *Fernàndez-Garcia et al.* (2005) found that, even under a uniform, natural-gradient flow condition, only several full-penetrating wells are required to accurately estimate the first two moments of the breakthrough time curves (BTCs) obtained from total mass fluxes passing through the control planes. Additionally, there are also some merits in the computational aspects (*Harvey and Gorelick*, 1995). For example, the travel times are scale-independent and thus avoid the disparity problem between the model resolution and the measurement scale since the travel times are typically modeled in a Lagrangian framework rather than the grid-based Eulerian method when solving the forward transport problem.

Several authors have approached the problem of conditioning aquifer parameters on the travel time. *Vasco and Datta-Gupta* (1999) developed an asymptotic solution to the solute transport equation to calculate the sensitivities for the inversion of the tracer data. Then an iterative linearized inversion algorithm is used to infer the parameter distribution. *Wen et al.* (2002) derived sensitivity coefficients of tracer travel time with respect to permeability by tracking streamlines between well pairs. The sequential self-calibration method is then employed to construct geostatistical realizations conditional to concentration data. Results from a synthetic aquifer show that tracer concentration data carry important information on the spatial variation of permeability in the inter-well areas while the pressure data only provide information near the well-bore.

In contrast to the entire BTCs, a variety of statistical measures computed from the BTCs, e.g., the peak concentration arrival times, the percentiles of travel times, and the temporal moments of tracer data, can also be used for inverse-conditional simulation and mitigate the computational effort. Several methods based on the temporal moments and statistics of the BTCs have been used for parameter inference. *Cirpka and Kitanidis* (2001) developed a sensitivity matrix of the temporal moments of tracer data with respect to the conductivity using the adjoint-state method. On the basis of such sensitivity matrix (of the first moment), the quasi-linear geostatistical inversion or iterative cokriging method is employed to conditioning the conductivity on the tracer data. A synthetic example demonstrates a minor improvement of the integration of tracer data (in terms of the first temporal moment) into the estimate of conductivity compared to the result of head data. *Rubin and Ezzedine* (1997), *Woodbury and Rubin* (2000), and *Bellin and Rubin* (2004) proposed to use the peak concentration arrival times to infer the geostatistical models of conductivity. Actually, public officials assessing health risks associated with contaminant exposure in a drinking water supply system may be most concerned with peak concentration or the corresponding arrival time (*Lemke et al.*, 2004). Moreover, an appealing point in data acquisition is that the peak concentration arrival time is less affected by the truncated BTC records, e.g., the missing early or late arrivals due to the infrequent sampling and the insensitivity of measurements. *Wilson and Rubin* (2002) used indicator variables of solute arrivals for the inference of parameters controlling the heterogeneous structure of conductivity and the mean flow velocity.

The meaningfulness of conditioning on the various percentiles of the BTCs is apparent

in physics. The early arrivals in the BTCs follow the fastest pathways between the release source and the control plane, which are dominated by preferential flow, i.e., flow conduits. On the other hand, the late travel times reflect a more integral behavior, or even flow barriers. Therefore, different inversion results provide distinct knowledge about the flow and transport properties. High connectivity generally results in earlier breakthrough. Failing to account for such case will have too conservative conclusion in risk analysis in that the real arrival time may be much faster than that estimated one (*Gómez-Hernández and Wen*, 1998). On the other hand, low connectivity results in later breakthrough. An aquifer remediation design without considering such feature may fail because the resident contaminants will be removed more slowly than expected (*Wagner and Gorelick*, 1989). *Harvey and Gorelick* (1995) presented a method for estimating the spatial pattern of conductivity from the quartiles of solute arrive times. In a hypothetical aquifer example, they found that adding the median quartile of the BTCs to the cokriging procedure does improve the accuracy of the estimate of conductivity. But the tails of the BTCs (10 and 90 percentiles in their case) do not convey much more information about the conductivity field than the median quartile on the basis of the first-order approximation of the flow and transport equations.

The paper presents a blocking Markov chain Monte Carlo (BMcMC) algorithm tailored to the problems found in groundwater flow and mass transport inverse modeling that is used to perform an uncertainty and worth of data analysis involving conductivity, piezometric heads and travel times. Uncertainties about the different parameters and variables, at unsampled locations, are modeled, and the worth of the different types of data is evaluated under different constraints.

No conceptual model uncertainty is accounted for. That is, we assume that geometry, boundary conditions, and all intervening processes are known. We also assume known the prior model of spatial variability of conductivity, that is, the multiGaussian random function model characterizing conductivity, thus, the worth of data from diverse sources is evaluated without conceptual or structural uncertainty involved.

The remaining of the paper is organized as follows. First, the blocking Markov chain Monte Carlo method for inverse stochastic modeling of groundwater flow and mass transport is outlined. Then, several quantitative metrics are defined to measure model uncertainty and prediction uncertainty. It follows a synthetic example to illustrate the power of BMcMC in performing data integration. Then, the worth of various data is analyzed on the basis of uncertainty assessment. And finally, summary and conclusions end the paper.

2 Methodology

2.1 Blocking Markov chain Monte Carlo

Markov chain theory proves that using an appropriate transition kernel one can build a chain of realizations that, eventually, will converge to a series of random drawings from a prespecified probability distribution function (pdf). Each member of the chain is conditional to

the previous member and its value is determined through the transition kernel as a function of the previous chain member value. The transition kernel is a probability distribution function, much simpler to draw from than the target probability distribution function. Realizations can be univariate or multivariate. The two problems faced in any McMC implementation are which transition kernel to use, and how long it will take for the chain to converge.

For the specific problem of stochastic inverse conditioning in hydrogeology we wish to generate realizations of conductivity, conditional to conductivity, piezometric head, and travel times. Under some assumptions, which will be discussed later, we can build the multivariate conditional probability distribution function (cpdf) of conductivity given the different types of data. To draw directly from this cpdf is impossible, therefore, we resort to building a Markov chain of realizations that will converge to a series drawn from this cpdf.

Consider conductivity discretized at n grid nodes and modeled as a random function (RF). Consider that there are m conductivity data, and k state data (including both piezometric head and travel time). Specifically, let $\boldsymbol{x} = (x_1, x_2, ..., x_n)^T \subset R^n$ denote the RF, $\boldsymbol{x}_1 = \boldsymbol{x}_{obs} = (x'_1, x'_2, ..., x'_m)^T \subset R^m$ denote the m conductivity data, and $\boldsymbol{y} = \boldsymbol{y}_{obs} = (y_1, y_2, ..., y_k)^T \subset R^k$ denote the dependent state data. Conductivity and the state data are related through the groundwater flow and mass transport equations, succinctly represented by $\boldsymbol{y} = g(\boldsymbol{x})$.

Assuming a multiGaussian RF model for $\boldsymbol{x}$, with a spatially variable mean $\boldsymbol{\mu}$ and a stationary covariance function $\boldsymbol{C_x}$, the joint probability distribution function of $\boldsymbol{x}$ given $\boldsymbol{x}_1$ is,

$$\pi(\boldsymbol{x}|\boldsymbol{x}_1, \boldsymbol{\theta}) = (2\pi)^{-\frac{n}{2}} \|\boldsymbol{C}_{\boldsymbol{x}|\boldsymbol{x}_1}\|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu}_{\boldsymbol{x}|\boldsymbol{x}_1})^T \boldsymbol{C}^{-1}_{\boldsymbol{x}|\boldsymbol{x}_1} (\boldsymbol{x} - \boldsymbol{\mu}_{\boldsymbol{x}|\boldsymbol{x}_1})\right\}, \tag{1}$$

where parameter $\boldsymbol{\theta} = \{\boldsymbol{\mu}, \boldsymbol{C_x}\}$ represents the prior information about the random function, $\boldsymbol{\mu}_{\boldsymbol{x}|\boldsymbol{x}_1}$ is the conditional mean obtained by simple kriging, and $\boldsymbol{C}_{\boldsymbol{x}|\boldsymbol{x}_1}$ is the conditional covariance obtained by simple kriging, too.

Assuming a multiGaussian distribution for the discrepancy between observed state values $\boldsymbol{y}$ and their corresponding model predictions $\boldsymbol{y}_{sim} = g(\boldsymbol{x})$, the joint pdf of $\boldsymbol{y}$ given the model parameters $\boldsymbol{x}$ is,

$$\pi(\boldsymbol{y}|\boldsymbol{x}) = (2\pi)^{-\frac{k}{2}} \|\boldsymbol{C_y}\|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\boldsymbol{y} - g(\boldsymbol{x}))^T \boldsymbol{C}^{-1}_{\boldsymbol{y}} (\boldsymbol{y} - g(\boldsymbol{x}))\right\}, \tag{2}$$

where $\boldsymbol{C_y}$ is the error covariance, generally assumed diagonal (see for instance *Carrera and Neuman* (1986) for a good justification of this choice).

Using Bayes' theorem, we can derive the posterior cpdf of $\boldsymbol{x}$ given the observations $\boldsymbol{x}_1$ and $\boldsymbol{y}$, and the prior model $\boldsymbol{\theta}$,

$$\pi(\boldsymbol{x}|\boldsymbol{x}_1, \boldsymbol{y}, \boldsymbol{\theta}) = \frac{1}{c} \times \pi(\boldsymbol{y}|\boldsymbol{x}) \times \pi(\boldsymbol{x}|\boldsymbol{x}_1, \boldsymbol{\theta}), \tag{3}$$

with $c = \int \pi(\boldsymbol{y}|\boldsymbol{x})\pi(\boldsymbol{x}|\boldsymbol{x}_1, \boldsymbol{\theta}) d\boldsymbol{x}$ being a normalization constant.

Dropping the different constants we arrive to an expression for this posterior cpdf, up to an unknown proportionality constant, given by,

$$\pi(\boldsymbol{x}|\boldsymbol{x}_1,\boldsymbol{y},\boldsymbol{\theta}) \propto \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_{\boldsymbol{x}|\boldsymbol{x}_1})^T\boldsymbol{C}^{-1}_{\boldsymbol{x}|\boldsymbol{x}_1}(\boldsymbol{x}-\boldsymbol{\mu}_{\boldsymbol{x}|\boldsymbol{x}_1}) - \frac{1}{2}(\boldsymbol{y}-g(\boldsymbol{x}))^T\boldsymbol{C}^{-1}_{\boldsymbol{y}}(\boldsymbol{y}-g(\boldsymbol{x}))\right\}. \tag{4}$$

The objective of the stochastic inverse-conditional simulation is then to draw independent, identically distributed (*i.i.d.*) samples for $\boldsymbol{x}$ from this posterior cpdf $\pi(\boldsymbol{x}|\boldsymbol{x}_1,\boldsymbol{y},\boldsymbol{\theta})$. For the sake of simplicity, $\boldsymbol{\theta}$ and $\boldsymbol{x}_1$ are dropped out in the sequel such that $\pi(\boldsymbol{x}) \equiv \pi(\boldsymbol{x}|\boldsymbol{x}_1,\boldsymbol{\theta})$ and $\pi(\boldsymbol{x}|\boldsymbol{y}) \equiv \pi(\boldsymbol{x}|\boldsymbol{x}_1,\boldsymbol{y},\boldsymbol{\theta})$.

We adopt a Markov chain Monte Carlo (McMC) algorithm employing the Metropolis-Hastings rule to explore the posterior cpdf $\pi(\boldsymbol{x}|\boldsymbol{y})$. It can be proven (*Metropolis et al.*, 1953; *Hastings*, 1970; *Geman and Geman*, 1984) that, if a series of realizations is drawn according to the following rules, it will eventually converge to a series of *i.i.d.* realizations drawn from $\pi(\boldsymbol{x}|\boldsymbol{y})$:

(1) Initialize the first realization $\boldsymbol{x}$;

(2) Generate a new realization according to the Metropolis-Hastings rule:

- Draw a candidate realization $\boldsymbol{x}^*$, conditioned on the previous realization $\boldsymbol{x}$ in the series using a transition kernel q, $\boldsymbol{x}^* \sim q(\boldsymbol{x}^*|\boldsymbol{x})$;
- Accept the candidate realization $\boldsymbol{x}^*$ with probability $\min\{1,\alpha\}$, where

$$\alpha = \frac{\pi(\boldsymbol{x}^*|\boldsymbol{y})q(\boldsymbol{x}|\boldsymbol{x}^*)}{\pi(\boldsymbol{x}|\boldsymbol{y})q(\boldsymbol{x}^*|\boldsymbol{x})}; \tag{5}$$

(3) Return to (2).

There are two critical points in the previous algorithm, how to select the transition kernel q, and how to evaluate the acceptance probability α. These two critical points have to be solved with two aims, accelerate the convergence of the chain to become a chain of realizations drawn from the sought cpdf, and favor the mixing of the chain, in the sense that the realizations explore all the space of the cdpf so that they are truly random, *i.i.d.* realizations.

There are not too many example applications of McMC in the hydrogeological literature. *Oliver et al.* (1997) were probably the first one, and they used a single-component updating kernel. With such a kernel, the candidate realization $\boldsymbol{x}^*$ is exactly equal to the previous realization except for one of the discretization cells, the value of which is drawn from a Gaussian univariate pdf with mean and variance computed by simple kriging from the conditioning $\boldsymbol{x}_1$ data. As one can quickly foresee, this method, although effective, is quite inefficient in

its convergence, particularly if the number of cells n is large, and required much computer time to generate the realizations.

Our proposal for the transition kernel follows the dissertation by *Fu* (2008) who employs a blocking scheme. The proposal candidate $\boldsymbol{x}^*$ differs from the previous realization $\boldsymbol{x}$ in an entire block, in the synthetic cases performed by *Fu* (2008), half a correlation range was the optimal block size for maximum convergence speed, although its exact size could be subject to test for the specific case at hand. The proposal kernel is a multiGaussian pdf conditioned on the conditioning $\boldsymbol{x}_1$ data, and on the values of the cells in the previous realization immediately adjacent to the block being updated. Generating the new values within the block is very simple when there are no $\boldsymbol{x}_1$ data, since, in such a case the conditioning pattern for any block is always the same and some precalculations can be carried out to speed up the generation process. When there are $\boldsymbol{x}_1$ data, the generation is a little bit less efficient, but it follows standard geostatistical techniques. The blocking scheme has proven to make the McMC simulations more efficient in other fields (*Liu*, 1996; *Roberts and Sahu*, 1997).

For the computation of the acceptance probability in (5), we should first analyze how its different components can be evaluated. It can be rewritten as,

$$\alpha = \frac{\pi(\boldsymbol{x}^*)}{\pi(\boldsymbol{x})} \frac{\pi(\boldsymbol{y}|\boldsymbol{x}^*)}{\pi(\boldsymbol{y}|\boldsymbol{x})} \frac{q(\boldsymbol{x}|\boldsymbol{x}^*)}{q(\boldsymbol{x}^*|\boldsymbol{x})}, \tag{6}$$

taking the logarithm results in,

$$\ln \alpha = \ln \pi(\boldsymbol{x}^*) - \ln \pi(\boldsymbol{x}) + \ln q(\boldsymbol{x}|\boldsymbol{x}^*) - \ln q(\boldsymbol{x}^*|\boldsymbol{x}) + \ln \pi(\boldsymbol{y}|\boldsymbol{x}^*) - \ln \pi(\boldsymbol{y}|\boldsymbol{x}). \tag{7}$$

These terms can be determined as follows:

$$\ln \pi(\boldsymbol{x}^*) = -\frac{1}{2}(\boldsymbol{x}^* - \boldsymbol{\mu})^T \boldsymbol{C}_{\boldsymbol{x}}^{-1}(\boldsymbol{x}^* - \boldsymbol{\mu}) + c_1, \tag{8}$$

$$\ln \pi(\boldsymbol{x}) = -\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^T \boldsymbol{C}_{\boldsymbol{x}}^{-1}(\boldsymbol{x} - \boldsymbol{\mu}) + c_1, \tag{9}$$

$$\ln q(\boldsymbol{x}^*|\boldsymbol{x}) = -\frac{1}{2}(\boldsymbol{x}^* - \boldsymbol{\mu}')^T \boldsymbol{C}_{\boldsymbol{x}}'^{-1}(\boldsymbol{x}^* - \boldsymbol{\mu}') + c_2, \tag{10}$$

$$\ln q(\boldsymbol{x}|\boldsymbol{x}^*) = -\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu}')^T \boldsymbol{C}_{\boldsymbol{x}}'^{-1}(\boldsymbol{x} - \boldsymbol{\mu}') + c_2, \tag{11}$$

$$\ln \pi(\boldsymbol{y}|\boldsymbol{x}^*) = -\frac{1}{2}(\boldsymbol{y} - g(\boldsymbol{x}^*))^T \boldsymbol{C}_{\boldsymbol{y}}^{-1}(\boldsymbol{y} - g(\boldsymbol{x}^*)) + c_3, \tag{12}$$

$$\ln \pi(\boldsymbol{y}|\boldsymbol{x}) = -\frac{1}{2}(\boldsymbol{y} - g(\boldsymbol{x}))^T \boldsymbol{C}_{\boldsymbol{y}}^{-1}(\boldsymbol{y} - g(\boldsymbol{x})) + c_3, \tag{13}$$

where $\pi(\boldsymbol{x}^*)$ and $\pi(\boldsymbol{x})$ are functions of the values of $\boldsymbol{x}$ over the entire domain ($\boldsymbol{x}$ and $\boldsymbol{x}^*$ are vectors of size n, and $\boldsymbol{C}_{\boldsymbol{x}}$ is a matrix of size $n \times n$), $\pi(\boldsymbol{y}|\boldsymbol{x}^*)$ and $\pi(\boldsymbol{y}|\boldsymbol{x})$ are functions of the values of $\boldsymbol{y}$ at the observation locations ($\boldsymbol{y}$ and $\boldsymbol{y}^*$ are vectors of size m, and $\boldsymbol{C}_{\boldsymbol{y}}$ is a matrix

of size $m \times m$), and $q(\boldsymbol{x}^*|\boldsymbol{x})$ and $q(\boldsymbol{x}|\boldsymbol{x}^*)$ are functions of the values of $\boldsymbol{x}$ within the n' cells of the updating block ($\boldsymbol{x}$ and $\boldsymbol{x}^*$ are, in this case, vectors of size n', and $\boldsymbol{C}'_{\boldsymbol{x}}$ is a matrix of size $n' \times n'$), $\boldsymbol{\mu}'$ and $\boldsymbol{C}'_{\boldsymbol{x}}$ are the mean and the covariance of $\boldsymbol{x}$ within the n' cells of the updating block conditioned to the immediately adjacent values of $\boldsymbol{x}$ in the previous realizations and to the conditioning data $\boldsymbol{x}_1$, and c_1, c_2 and c_3 are constants that cancel out in (7).

The sensitivity analysis performed by *Fu* (2008) show that as soon as the domain is relatively large, inverting the covariance matrix in (8) and (9) is prohibitive, thus, these two values have to be approximated. For this purpose, we approximate both pdf values by the values of a multiGaussian distribution defined over a smaller domain centered at the block being updated. This smaller domain was made equal to twice the size of the updating block. *Fu* (2008) also found that for the initial convergence of the chain the most efficient approach was to use an independent kernel, in which case, $q(\boldsymbol{x}^*|\boldsymbol{x}) = \pi(\boldsymbol{x}^*)$ and the acceptance rate simplifies to,

$$\alpha = \frac{\pi(\boldsymbol{y}|\boldsymbol{x}^*)}{\pi(\boldsymbol{y}|\boldsymbol{x})}. \tag{14}$$

Thus, the procedure to generate the realizations is as follows: at the beginning of the chain, the proposal candidates are drawn using an independent kernel and the acceptance rate (14), then, when convergence has been achieved (as measured by the discrepancy between observed and simulated state variables), the conditional kernel is used and the acceptance rate is computed by (7) but approximating both (8) and (9) over a domain only twice as large as the updating block. Although not used in this paper, several further improvements on the methodology have been described by *Fu* (2008) and *Fu and Gómez-Hernández* (2009).

2.2 Uncertainty assessment

We will use the proposed blocking Markov chain Monte Carlo method (BMcMC) in a synthetic example to build a model of uncertainty for conductivity, piezometric heads, travel times and macrodispersion, from a limited amount of conductivity, head and travel time data. In our example we will use a synthetic aquifer to generate a synthetic reality from which samples will be taken, our synthetic reality will serve as a reference when analyzing the conditional realizations. The procedure to construct a model of uncertainty in practice would be the same, although we could not evaluate any bias with respect to reality, because reality would be unknown. In addition, in our synthetic example we assume all other parameters and model constraints known, whereas in reality there will be additional parameter and model uncertainties. Such an uncertainty model should always accompany our model predictions, so that the decision maker can assess the degree of confidence that can be deposit in the model.

The model of uncertainty is also a tool that permits the evaluation of the worth of data, and can be used for optimal network design.

In this example, in order to analyze the uncertainty due to scarcity of data, we will assume that there is no conceptual uncertainty (we know the processes involved, boundary conditions, geometry, etc.) so we can better analyze the specific worth of each data without interferences due to unaccounted uncertainties.

To measure parameter uncertainty we will compute three metrics. The first metric $I(\boldsymbol{x})_2^2$ measures the precision of the realizations since it evaluates the ensemble variance over all the cells as given by:

$$I(\boldsymbol{x})_2^2 = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{n_r}\sum_{r=1}^{n_r}(x_{i,r} - \bar{x}_i)^2, \tag{15a}$$

where n is the number of cells, n_r is the number of realizations, $x_{i,r}$ is the simulated value at cell i and realization r, and $\bar{x}_i$ is the ensemble average over all realizations at location i.

The other two metrics take advantage of our knowledge of the reference field to measure the average bias $I(\boldsymbol{x})_3$, and a combination of bias and precision $I(\boldsymbol{x})_4^2$:

$$I(\boldsymbol{x})_3 = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{n_r}\sum_{r=1}^{n_r}(x_{i,r} - x_{i,ref}), \tag{16a}$$

$$I(\boldsymbol{x})_4^2 = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{n_r}\sum_{r=1}^{n_r}(x_{i,r} - x_{i,ref})^2, \tag{16b}$$

where $x_{i,ref}$ is the reference value at location i.

Similarly, we define $I(\boldsymbol{y})_2^2$, $I(\boldsymbol{y})_3$ and $I(\boldsymbol{y})_4^2$ to measure the precision and bias of the model predictions.

2.2.1 Macrodispersion

In addition, the macrodispersion can be viewed as a parameter reflecting the spatial variability of hydraulic conductivity. Indeed, experimental and theoretical results have suggested that macrodispersion of solutes is essentially produced by the spatial variation of the fluid velocity resulting from the heterogeneity of hydraulic conductivity. We expect that as more conditioning data are used, specially data on travel times, the macrodispersion coefficient computed in the inverse conditioned realizations would be closer to the macrodispersion coefficient of the reference.

The scale- or time-dependent macrodispersion is defined as the change rate of the second-order moment of a solute plume. Extensive studies on the effects of hydraulic conductivity on macrodispersion of solutes have shown that, under steady-state flow conditions with a uniform mean hydraulic gradient in a statistically stationary media of finite correlation length of hydraulic conductivity, macrodispersion increases with time from the point at which the solute body first enters the flow domain, until after the solute cloud has traveled a few tens

of hydraulic conductivity correlation lengths when it reaches a constant asymptotic value (*Dagan*, 1984; *Khaleel*, 1994).

We compute the macrodispersion coefficients at several control planes away from the release plane using the coefficient of variation of the breakthrough curves (BTCs) as in the works by *Kreft and Zuber* (1978), *Desbarats and Srivastava* (1991), and *Wen and Gómez-Hernández* (1998),

$$A_L(\mathrm{x}) = \frac{\mathrm{x}}{2}\frac{\sigma_t^2(\mathrm{x})}{m_t^2(\mathrm{x})}, \tag{17}$$

where $A_L(\mathrm{x})$ is the apparent longitudinal macrodispersion at control plane x, $m_t(\mathrm{x})$ and $\sigma_t^2(\mathrm{x})$ are the mean and variance of travel times at x, respectively. To overcome the sensitivity of $\sigma_t^2(\mathrm{x})$ to the presence of outlier travel times, the distribution of log travel time was suggested by *Khaleel* (1994) and *Wen and Gómez-Hernández* (1998) to calculate the temporal moments,

$$m_t(\mathrm{x}) = \exp\left\{m_{\ln t}(\mathrm{x}) + \frac{1}{2}\sigma_{\ln t}^2(\mathrm{x})\right\},$$
$$\sigma_t^2(\mathrm{x}) = m_t^2(\mathrm{x})\left(\exp\left(\sigma_{\ln t}^2(\mathrm{x})\right) - 1\right),$$

$m_{\ln t}(\mathrm{x})$ and $\sigma_{\ln t}^2(\mathrm{x})$ are the mean and variance of log travel times at the displacement distance x.

3 A Synthetic Example

3.1 Reference models and conditioning data sets

A synthetic isotropic lognormal 2D confined aquifer under a uniform, natural-gradient flow condition, serves as the reference field to illustrate the effectiveness of the proposed method for inverse-conditional simulation. In this paper, we assume that $\boldsymbol{x}$ is $\ln \boldsymbol{K}$ that we aim to identify and $\boldsymbol{y}$ represents the conditioning data including head data ($\boldsymbol{h}$) and measured travel time ($\boldsymbol{t}$). Any set of coherent units will lead to the same results and conclusions, therefore, in the remainder of the paper no units are quoted, just the dimensions of the parameters.

Figure 1A shows the reference conductivity field. It was generated using the code GCOSIM3D (*Gómez-Hernández and Journel*, 1993) that generates realizations from a stationary multiGaussian distribution. It has 100 by 100 cells, each cell is of size 1 $[L]$ by 1 $[L]$ by 1 $[L]$. The mean of $\ln K$ is 0 $[\ln(LT^{-1})]$, the standard deviation of $\ln K$ is 1 $[\ln(LT^{-1})]$, and the two-point covariance is exponential with an integral scale of 50/3 $[L]$, that is, a practical range—the distance at which 95% of the correlation is lost—of 50 $[L]$ in all directions.

The use of a synthetic aquifer with a high spatial correlation is justified for two reasons: (1) it allows a better visual comparison of the simulations and the reference and (2) it is easier

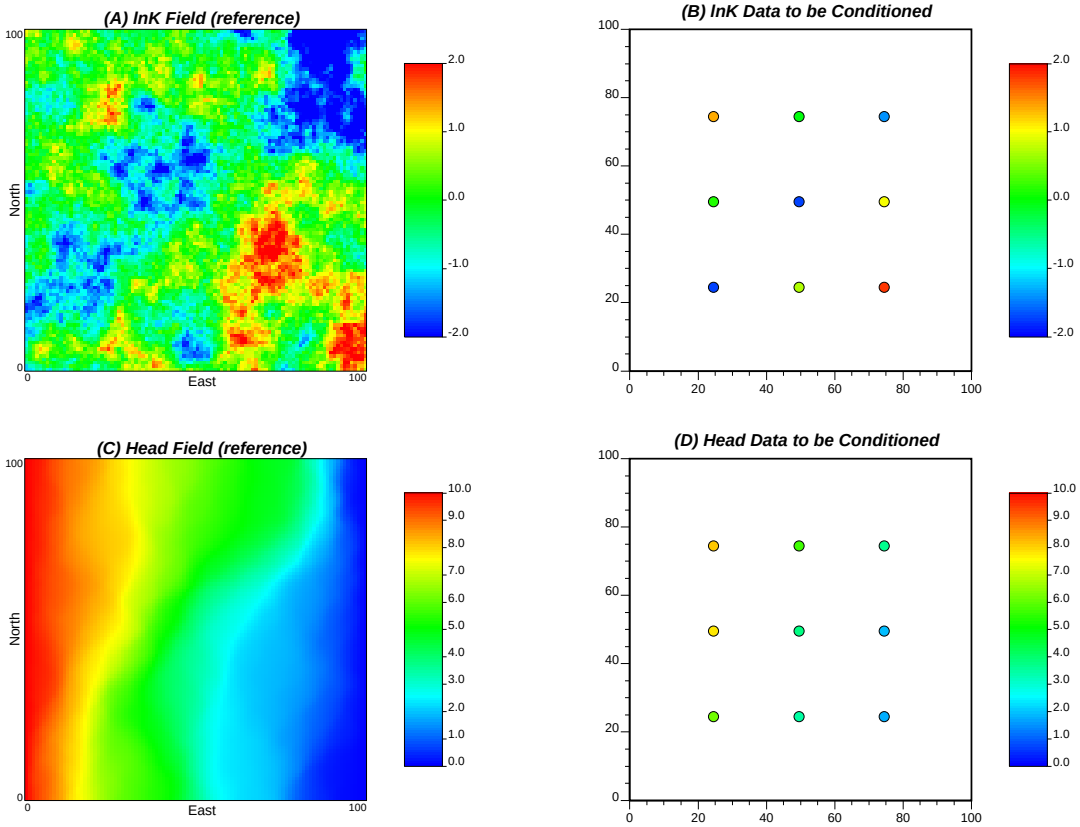


Figure 1: (A) Reference $\ln K$ field, (B) conditioning $\ln K$ data, (C) reference piezometric head (solution of the flow equation in the reference conductivity field), (D) conditioning piezometric head data

to measure the worth of data since only a few measurements are necessary to capture the essential features of the aquifer. On the other hand, the high correlation tightens the set of conditional realizations making more difficult for the BMcMC algorithm to draw acceptable candidate realizations.

The high correlation produces a relatively large correlation range to sampling space ratio, which is always an important factor for model identification. Generally speaking, compared to the sampling spacing, a smaller prior correlation length yields an inverse estimate closer to the prior mean except in the immediate vicinity of measurements (*McLaughlin and Townley*, 1996). In principle, parameter variability with a correlation length smaller than the sampling spacing cannot be determined. Conversely, a larger prior correlation length yields a smooth estimate which varies gradually over the region of interest. Besides, it was recognized that the uncertainties of models and their responses remain significant even with a large number of hard conditioning data (*Harter and Yeh*, 1996; *van Leeuwen et al.*, 2000; *Scheibe and Chien*, 2003). *Eggleston et al.* (1996) found that for the Cape Code aquifer the estimation error is relatively insensitive to the number of hard data above a threshold of three measurements per integral volume. In this synthetic example we will keep two samples per correlation range.

A single-phase flow experiment is set up as follows. The upper and lower boundaries are no flow boundaries. The left and right sides have prescribed constant heads equal to 10 $[L]$ and 0 $[L]$, respectively. The confined steady-state flow problem is solved by a block-center finite-difference simulator. The reference head field obtained by solving the flow equation using the reference $\ln K$ field is shown in Figure 1C. From both the reference logconductivity and piezometric head fields, nine samples are taken that will be used for conditioning purposes at the locations and with the values shown in Figure 1B and Figure 1D.

With the $\ln K$ field, and the corresponding head field, the velocity field at cell's interface is obtained using Darcy's law and a constant porosity of 0.3. The conservative transport problem is then solved by the constant-displacement random-walk particle-tracking algorithm (*Wen and Gómez-Hernández*, 1996). The effects of pore-scale dispersion and molecular diffusion are neglected. In this case, the solute particles move along the streamlines of the steady-state velocity field. Hence the solute plume is confined transversally by the two no-flow boundaries. Two thousand particles, uniformly distributed along the left boundary, are tracked until they arrive at the control plane located at the right boundary. Travel times are recorded at different control planes for the purpose of computing the apparent macrodispersion at various distances from the source. The first and second moments of the breakthrough curve in the reference field at the right boundary, and five of its percentiles will serve as conditioning data; their values are shown in Table 1.

BMcMC is now used to generate realizations of logconductivity conditioned to the nine logconductivity data and the nine piezometric head data depicted in Figure 1, and the statistics of the travel time BTC at the right boundary shown in Table 1, all of them from an underlying multilogGaussian random function with zero mean, unit variance and an exponential isotropic covariance with a range of one half the simulation domain.

Table 1: *Statistics of the reference travel time BTC at the right boundary*

Statistic	Travel time $[T]$
05 percentile of the BTC	912
25 percentile of the BTC	1014
50 percentile of the BTC	1081
75 percentile of the BTC	1232
95 percentile of the BTC	1823
1st moment	1192
$\sqrt{\text{2nd moment}}$	313

We are not seeking "the best" logconductivity spatial distribution, but any realization that could be drawn from the random function and honor the conditioning data up to within measurement errors. At the end we will have an ensemble of realizations, each of which with the proper patterns of spatial variability, and each of which consistent with the conditioning data. These ensemble of conditional realizations serves to perform an uncertainty analysis on logconductivity, piezometric head, travel time and macrodispersion.

3.2 Uncertainty assessment

Six scenarios are analyzed. Several combinations of conditioning data sets are considered, from unconditional realizations to realizations conditioned to all data types. The conditioning information included in each scenario is indicated in Table 2. This table also shows the relative errors used to build the diagonal $\boldsymbol{C_y}$ matrix in (12) and (13). As mentioned previously, the large correlation of the reference field makes difficult for the BMcMC algorithm to sample the conditional realizations; for this purpose, the relative errors used are large so that realizations will be accepted even though the reproduction of head and travel times is not exact.

Table 2: *Scenarios and conditioning data types*

Scenario	Conditioning	Error measurements for building $\boldsymbol{C_y}$
1	$\boldsymbol{x}\|-$	—
2	$\boldsymbol{x}\|\boldsymbol{x}_1$	—
3	$\boldsymbol{x}\|\boldsymbol{h}$	$\sigma_h^2 = 0.2$
4	$\boldsymbol{x}\|\boldsymbol{t}$	$\sigma_t^2 = 1.0$
5	$\boldsymbol{x}\|\boldsymbol{h}, \boldsymbol{t}$	$\sigma_h^2 = 0.2$, $\sigma_t^2 = 1.0$
6	$\boldsymbol{x}\|\boldsymbol{x}_1, \boldsymbol{h}, \boldsymbol{t}$	$\sigma_h^2 = 0.2$, $\sigma_t^2 = 1.0$

With these restrictions, generating 100 realizations for each of the six scenarios can be carried out in an ordinary desktop PC in little time. For each set of 100 realizations the three metrics defined earlier are computed for logconductivity and piezometric head. Also, for each set of 100 realizations the ensemble mean and ensemble variance of logconductivity and of piezometric head are calculated. Also, for each realization, the apparent macrodispersion for planes at different distances from the source were computed using (17), and their ensemble means are recorded.

Table 3: *Precision, bias and combined measure of uncertainty for* $\ln K$

Scenario	Model	$I(\boldsymbol{x})_2^2$	$I(\boldsymbol{x})_3$	$I(\boldsymbol{x})_4^2$
1	$\boldsymbol{x}\|-$	1.02	0.22	1.45
2	$\boldsymbol{x}\|\boldsymbol{x}_1$	0.91	0.22	1.24
3	$\boldsymbol{x}\|\boldsymbol{h}$	0.89	0.25	1.22
4	$\boldsymbol{x}\|\boldsymbol{t}$	0.88	0.54	1.42
5	$\boldsymbol{x}\|\boldsymbol{h},\boldsymbol{t}$	0.85	0.30	1.19
6	$\boldsymbol{x}\|\boldsymbol{x}_1,\boldsymbol{h},\boldsymbol{t}$	0.81	0.32	1.13

Table 3 shows the value of the three metrics computed on the log of hydraulic conductivity. (Notice that, since neither $\bar{x}_i$ nor $x_{i,ref}$ are spatially uniform, $I(\boldsymbol{x})_4^2 \neq I(\boldsymbol{x})_3^2 + I(\boldsymbol{x})_2^2$). Figure 2 shows the ensemble average of the 100 logconductivity realizations for all six scenarios. Figure 3 shows the ensemble variance for the same 100 logconductivity realizations and scenarios. Figure 4 shows the histogram of all the generated conductivity values for all 100 realizations, and Figure 5 show the variograms for the reference field in the x- and y-directions, and the average variograms of all 100 realizations for each scenario.

Regarding the statistical structure of the realizations and considering that, as we assumed, no uncertainty about the prior RF model will be considered (that is, the prior RF model used for all BMcMC runs is the same as the one used to generate the reference realization), it is not surprising that all histograms for all scenarios are roughly Gaussian with zero mean and unit variance, with possibly the exception of the set of realizations which are only conditioned to travel time statistics which show a slightly large mean (also noticeable in the $I(\boldsymbol{x})_3$ metric) and a smaller variance. The mean variograms for all scenarios are very similar and close to the variogram of the reference realization, and actually they are almost identical for distances below λ (= 16.66). This indicates that the BMcMC algorithm does generate realizations with the desired statistical properties: indeed, in BMcMC we draw candidates from a distribution and then the drawing is accepted or rejected, but the realization is never modified (or perturbed) to make it conditional to the piezometric and travel time data, as is done in other stochastic inverse-conditional approaches.

Analyzing Table 3 we notice that the intrinsic variability of logconductivity about its ensemble mean $I(\boldsymbol{x})_2^2$ decreases as more information is added: with no conditional data

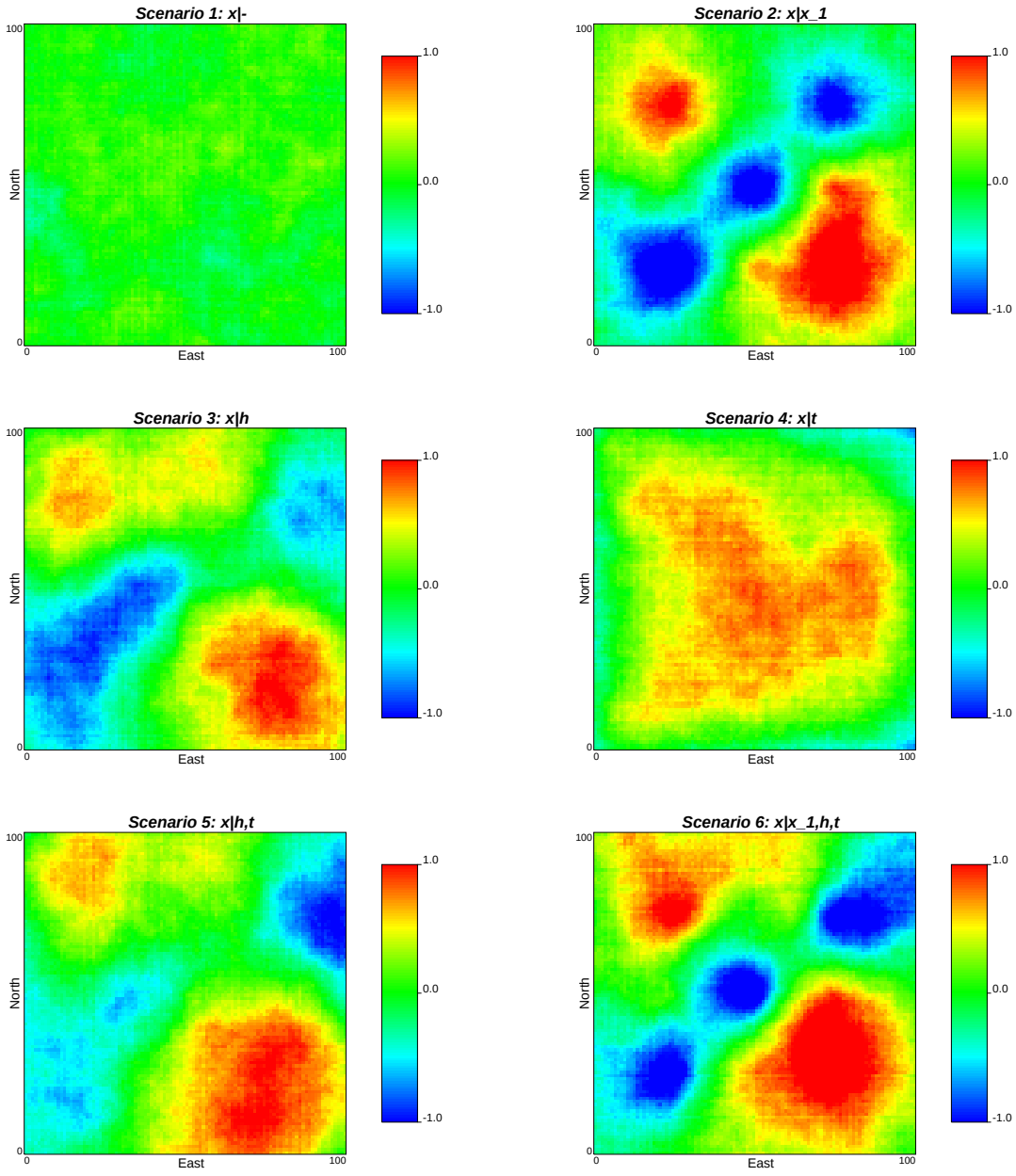


Figure 2: Ensemble average logconductivity fields for the different scenarios

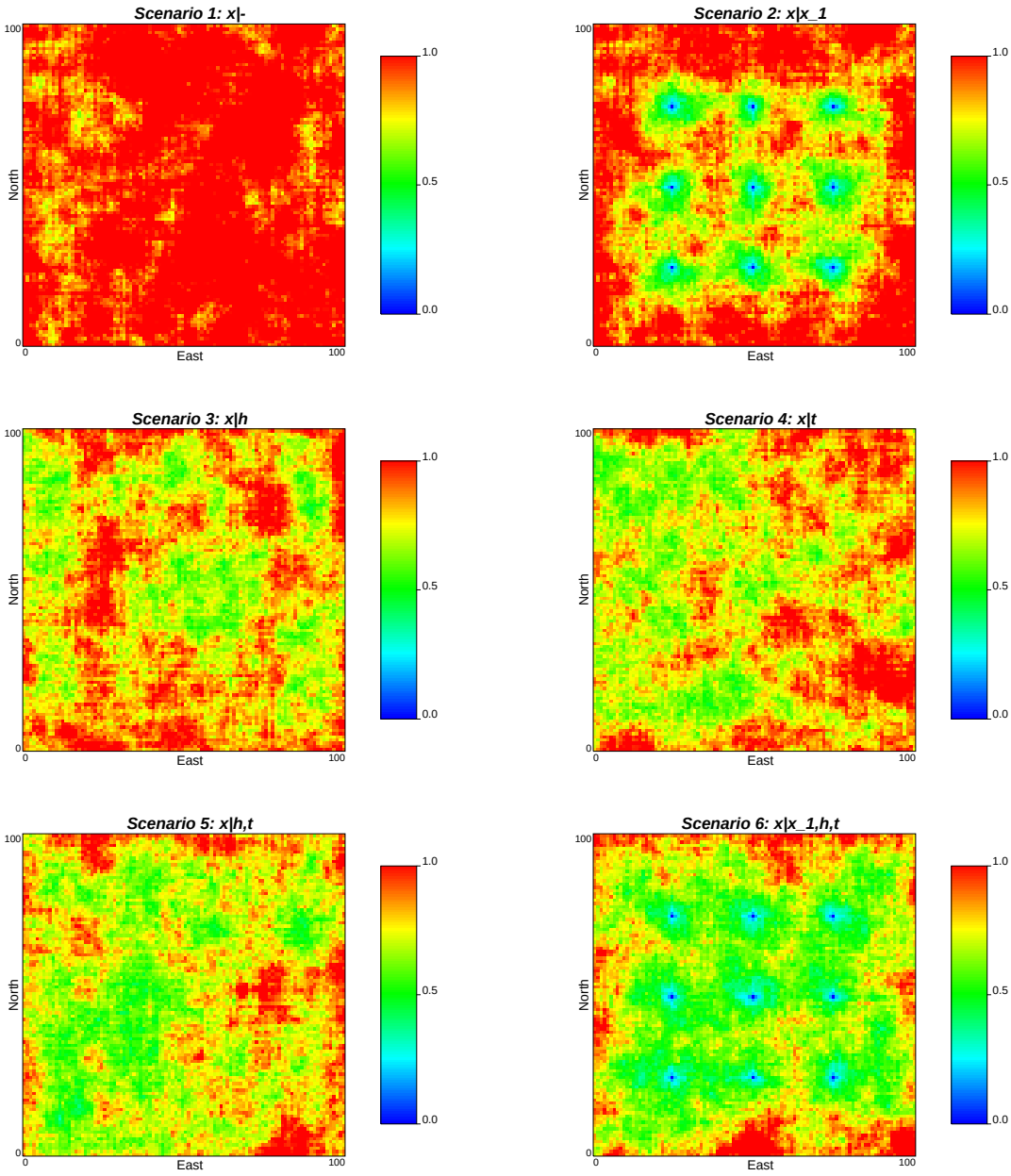


Figure 3: Ensemble variance of logconductivity fields for the different scenarios data

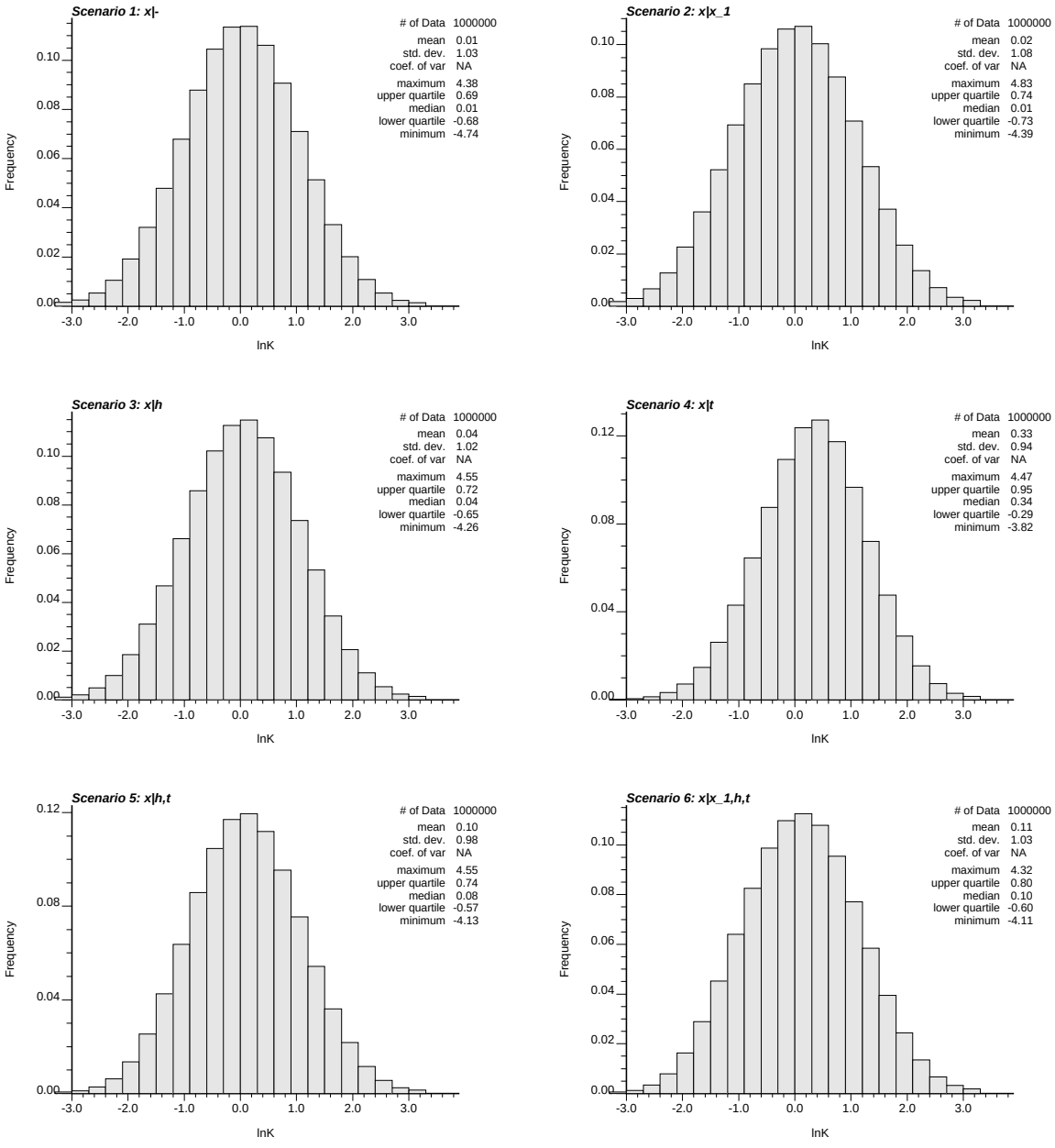
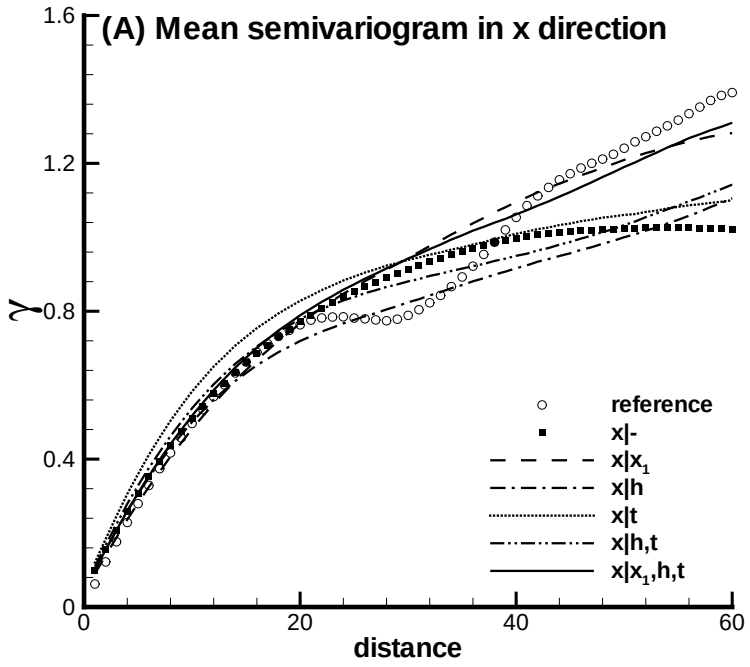


Figure 4: The histograms of lnK computed over all 100 realizations for the different scenarios



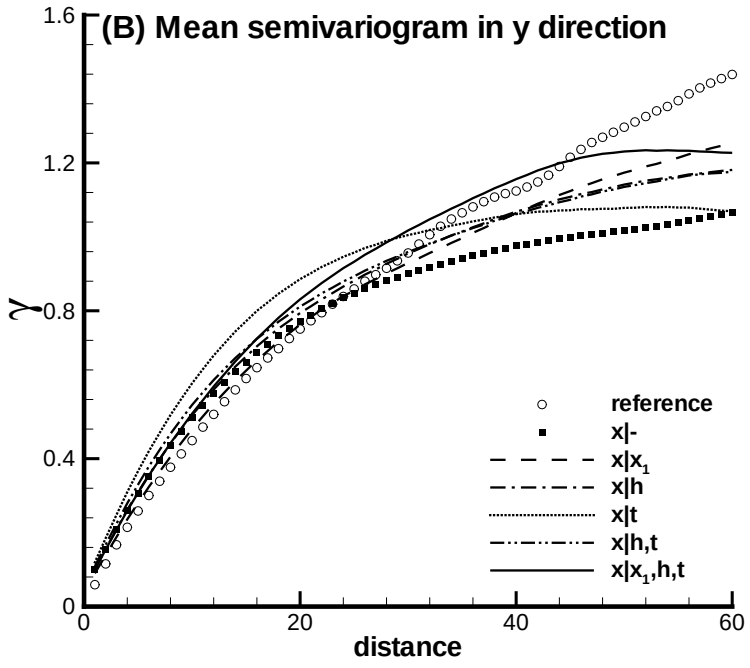


Figure 5: Ensemble average variograms of $\ln K$ for the different scenarios

(scenario 1) it is equal to the prior variance, and it is reduced to 0.91 when the 9 conductivity data are used. It seems that both head and travel time data impose a higher restriction on the alternative realizations that can reproduce such data, as indicated by a lower value of $I(\boldsymbol{x})_2^2$, and, as would be expected, when several data types are combined for conditioning, the ensemble variance reduces even further. In all cases, the ensemble of realizations are biased with respect to the reference as indicated by the values of $I(\boldsymbol{x})_3$, part of this behavior can be explained by the fact that the reference realization, although drawn from a multiGaussian RF with zero mean and unit variance, has a mean of 0.05 and a variance of 1.11 (this departure in the realizations statistics with respect to the RF model ones is expected given the large correlation range with respect to the size of the field). The fact that BMcMC draws all its realizations using an a priori RF with a smaller mean and a smaller variance than the reference should explain most of the observed bias. The metric $I(\boldsymbol{x})_4^2$ which is a combined measure of precision and bias summarizes the influence of the different types of data in getting an ensemble of realizations as close as the data permit to the reference: its value goes down from 1.45 with no data to 1.13 when all data are accounted for.

A more qualitative analysis, which leads to similar conclusions, can be done analyzing Figures 2 and 3. First of all, it is important to recall that the ensemble average is not a realization, therefore it does not display the degree of spatial variability of individual realizations (or the reference); yet, it can be used to check whether the long trend patterns of the reference are captured by the ensemble. Notice that the reference field has a very distinctive pattern for the highs and lows characterized by a zone of high values in the lower right corner, and a zone of low values in the upper right corner. Notice, too, that since the ensemble fields are smoother representations of the reference, the scale used in Figure 2 is narrower than the one in Figure 1 in order to appreciate better the overall patterns in the ensemble averages. When no conditioning data are used, no overall pattern is captured, even though each individual realizations will depict a pattern of variability similar to that of the reference. The spatial distributions of highs and lows will be randomly located in the ensemble of realizations leading to a flat ensemble mean as shown in Figure 2. When only the $\boldsymbol{x}_1$ data are used, we get the typical kriging map: the overall patterns are more consistently reproduced through the ensemble of realizations as long as the conditioning data samples them. Conditioning only to piezometric head is enough to introduce the patterns of highs and lows necessary to bend the head isolines implied by the nine head data (notice in the reference head field that the head gradient is not uniform with its highest values in the lower left and upper right corners). Piezometric head is the result of a convolution applied on all hydraulic conductivity data; for this matter, even though they are not sufficient to identify exactly local values, they are very good in capturing overall patterns. When conditioning to travel time only, and considering that the data is taken at the right boundary only, it is impossible to delineate local features in the realizations. Precisely, this difficulty in delineating local features induces the bias in the realizations, unless the high values of the lower right quadrant in the reference are captured by the conditioning algorithm. There is a need to increase the overall mean of conductivity in order to match the statistics of the reference BTC curve. It

is interesting to see how conditioning in head and travel times induces already the proper patterns of variability, in this case, local head information complement the global information provided by the travel times. And, of course, the best results are obtained when all data are accounted for.

From Figure 3 we can appreciate how the introduction of conditioning data forces the generation of conditional realizations more alike to each other. While no conditioning data is used, the overall variability of the realizations is equal to the prior variance. As soon as conductivity data are used, the variance map has the typical bull-eye look of a kriging map, with zero variance at sampling locations and growing away from them. More interesting is to notice how the variance is reduced in a more uniform manner when conditioning to piezometric head data and/or travel time data. Piezometric head samples, although local, carry information about conductivity over a much larger distance than conductivity samples, likewise travel time data. However, the minimum variance values, when conditioning on either head or time never reach zero anywhere, but, in contrast, they induce variance reductions near the edges of the aquifer, something that conductivity data alone cannot. Again, the highest reductions in conductivity variance occur for the case in which all data are used for conditioning.

Regarding the contribution of travel time to the reduction of uncertainty in conductivity, we find, from scenario 4 (even from scenario 5) that uncertainty reduction is larger near the upstream boundary than near the downstream one, what agrees with the finding presented by *Franssen et al.* (2003) who reached this conclusion in the basis of inverse-conditioning to spatial concentration data by the sequential self-calibration method (*Gómez-Hernández et al.*, 1997). The reason why it happens is still not very clear to the authors but *Franssen et al.* (2003) attributed this phenomenon to more sensitivity of concentration to flow velocity in the upstream zone.

Analysis of Table 4, containing the metrics regarding piezometric head, shows that head data is the most efficient piece of information in reducing the overall variability of predicted heads about their reference values, and also the bias. Scenarios 3, 5, and 6 have noticeably lower values for all three metrics than the other scenarios that do not use head as conditioning data. Taking as the reference the unconditional case, the reduction in all three metrics after conditioning to conductivity (scenario 3) is much smaller than after conditioning to head (scenario 4) indicating that, for the purpose of best characterization of piezometric head, piezometric data is the most appropriate.

From Figure 6 we can see that for the unconditional case, even though each realization will have undulating patterns of piezometric head similar to those in the reference, the random location of high and low gradients makes that the ensemble average appear to correspond to a uniform field. As soon as local data is introduced, the local patterns in piezometric head emerge, with an ensemble average for scenario 6 which is very close to the reference piezometric head of Figure 1. Travel time information, not being local, does not achieve, by itself only (scenario 4), a significant departure from the ensemble field for the unconditional case.

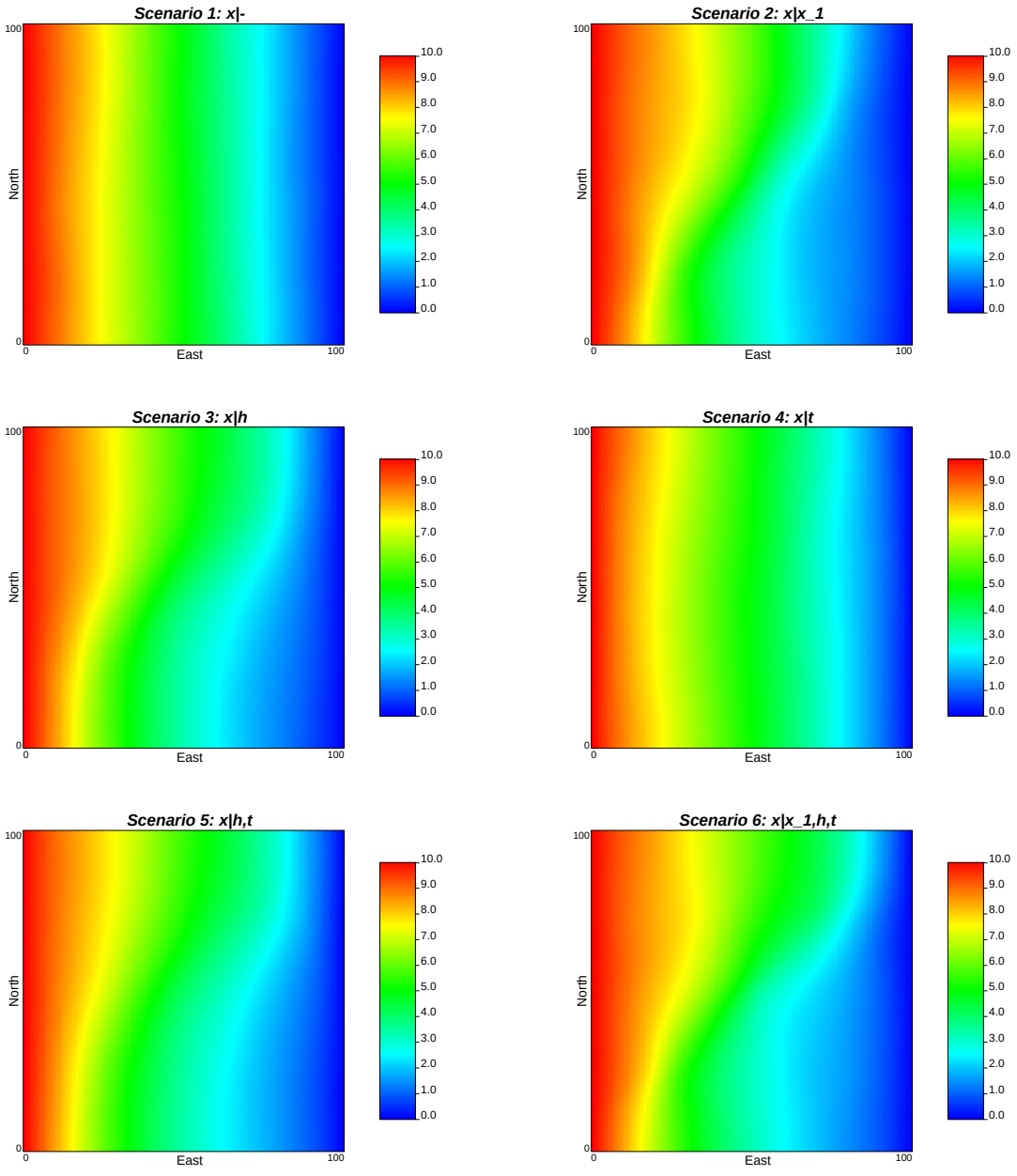


Figure 6: Ensemble average piezometric head fields for the different scenarios

Table 4: *Precision, bias and combined measure of uncertainty for predicted head*

Scenario	Model	$I(\boldsymbol{h})_2^2$	$I(\boldsymbol{h})_3$	$I(\boldsymbol{h})_4^2$
1	$\boldsymbol{x}\|-$	0.99	0.18	1.36
2	$\boldsymbol{x}\|\boldsymbol{x}_1$	0.73	-0.10	0.88
3	$\boldsymbol{x}\|\boldsymbol{h}$	0.40	-0.10	0.55
4	$\boldsymbol{x}\|\boldsymbol{t}$	0.65	0.18	1.12
5	$\boldsymbol{x}\|\boldsymbol{h},\boldsymbol{t}$	0.41	-0.10	0.57
6	$\boldsymbol{x}\|\boldsymbol{x}_1,\boldsymbol{h},\boldsymbol{t}$	0.36	-0.03	0.47

It is very interesting to see how the piezometric head ensemble variance evolves for the different scenarios (Figure 7). For the unconditional case we retrieve the classical result, no variance at boundaries with prescribed head, increasing away from the boundaries to a maximum at the center. Introducing conductivity data reduces the variance but still leaves some areas with high uncertainty about heads. The picture changes substantially as soon as head data are used. For scenarios 3, 5 and 6, the variability in head among realizations is very small as indicated by the close to zero value of the variance over most of the domain.

3.2.1 Conditioning and spreading

Rubin (1991) examined the impact of conditioning on $\ln K$ and h measurements in tracer plume migration. In his study, he also includes the temporal moments of BTCs into the conditioning procedure and investigates its impact on the prediction of solute plume spreading. While *Rubin*'s approach is based on the linearization of the flow and transport equations, our approach does not linearize them. We would like to compare our results with those by *Rubin*.

By setting various control planes in the mean flow direction, the macrodispersion coefficients for the different scenarios may be computed using (17). Travel time BTCs are observed at each of 10 control planes between 0.1λ and 2.0λ of the source, and the corresponding apparent macrodispersion is computed for each plane, for each realization, for each scenario. Also from these BTCs, the fifth and ninety-fifth percentile are retrieved. These travel times are not used for conditioning. The travel data used for conditioning are the statistics of the BTC at the right boundary, only.

Figure 8 shows the ensemble average apparent macrodispersion coefficients for all control planes and all scenarios and the macrodispersion coefficients obtained in the reference field. The first, and most noticeable, observation is that macrodispersion estimation is clearly enhanced when travel time data is used for conditioning. This observation is consistent with the work by *Woodbury and Sudicky* (1992). Based on the bromide and chloride tracer tests performed at the Borden aquifer in Ontario, Canada, they found that conditioning to the spatial moments of concentration data enhances the estimation on the rate of plume

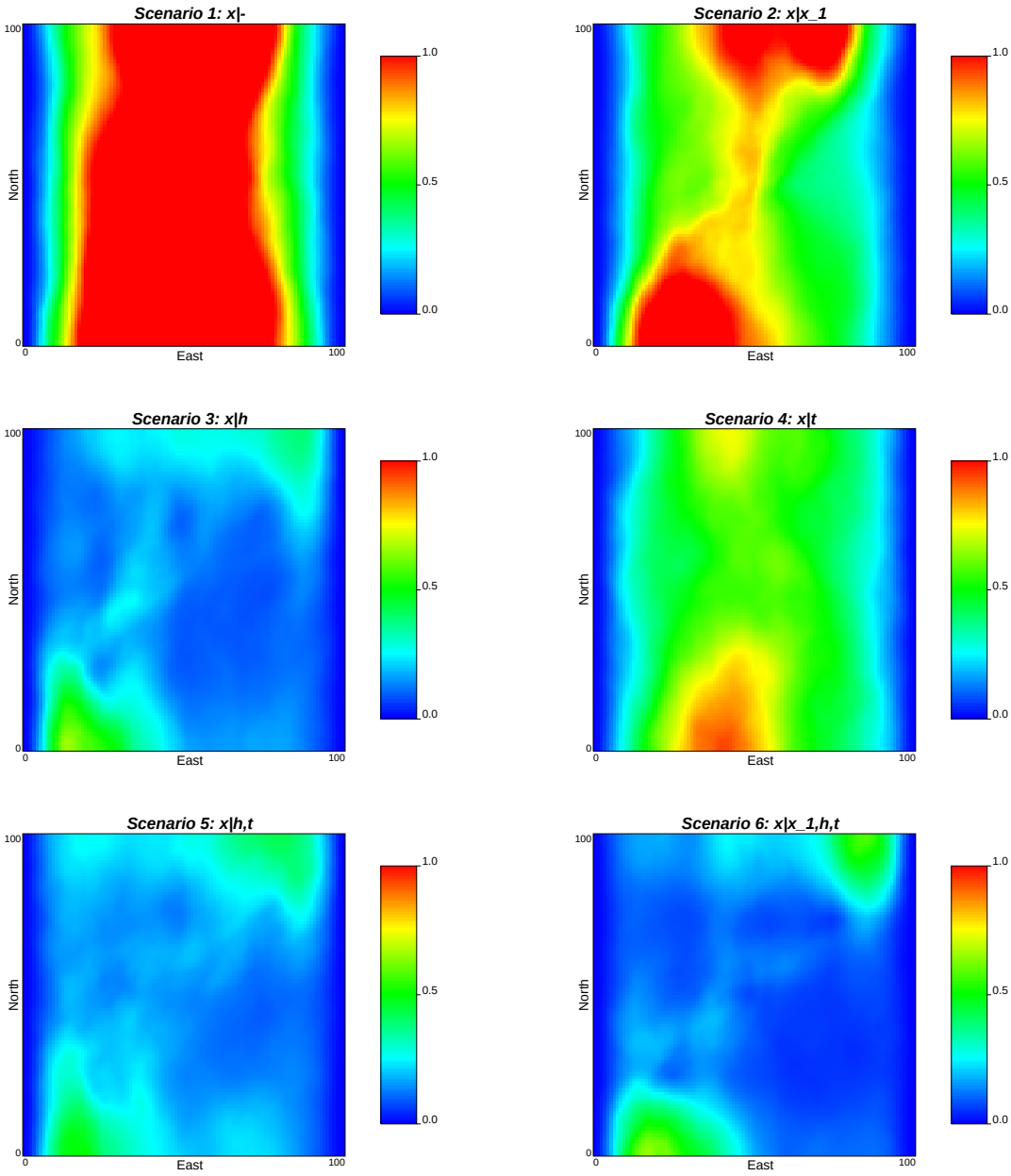


Figure 7: Ensemble variance of piezometric head fields for the different scenarios

spreading in the longitudinal direction. Therefore, although travel times alone do not help much in identifying conductivities or heads as was shown earlier, they do carry important information about solute spreading, whether used alone or with other types of data.

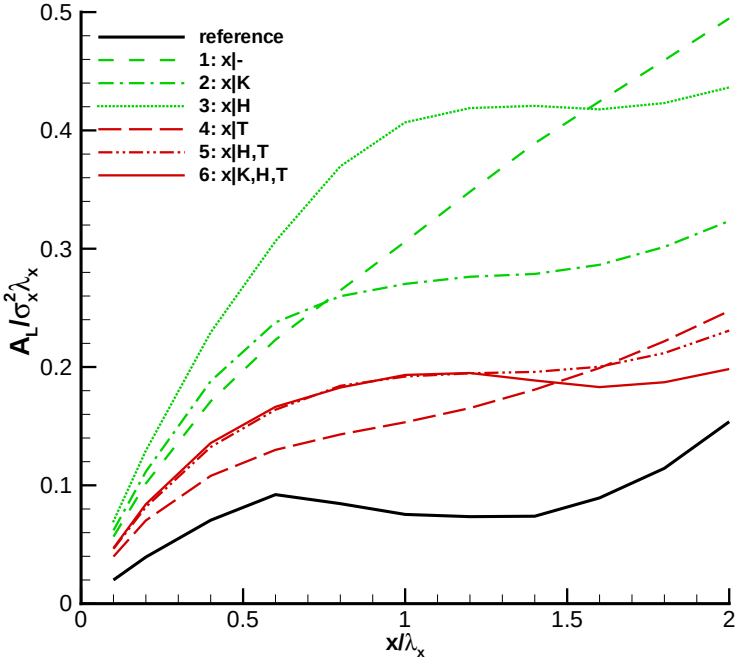


Figure 8: Ensemble average apparent macrodispersion coefficients as a function of the distance from the source for the different scenarios and the reference

When comparing conductivity and head, it appears that conductivity data alone contains more information about macrodispersion than head data alone. *Rubin* (1991) also found that introducing head data does not cause any considerable improvement for the estimation on the solute plume migration.

Figure 9 shows the evolution of the ensemble average of the logarithm of the fifth $t_{5\%}$ and ninety-fifth $t_{95\%}$ percentiles of travel time as the particles move away from the source for all scenarios. All scenarios underestimate the early arrival (Figure 9A) with small differences among them. For the early arrival of particles, within two integral scales of the source it seems that conditioning to data does not modify much the front tail of the BTCs. On the contrary, the late arrival is best estimated when all conditioning data are used (Figure 9B), although the difference between scenarios remains small.

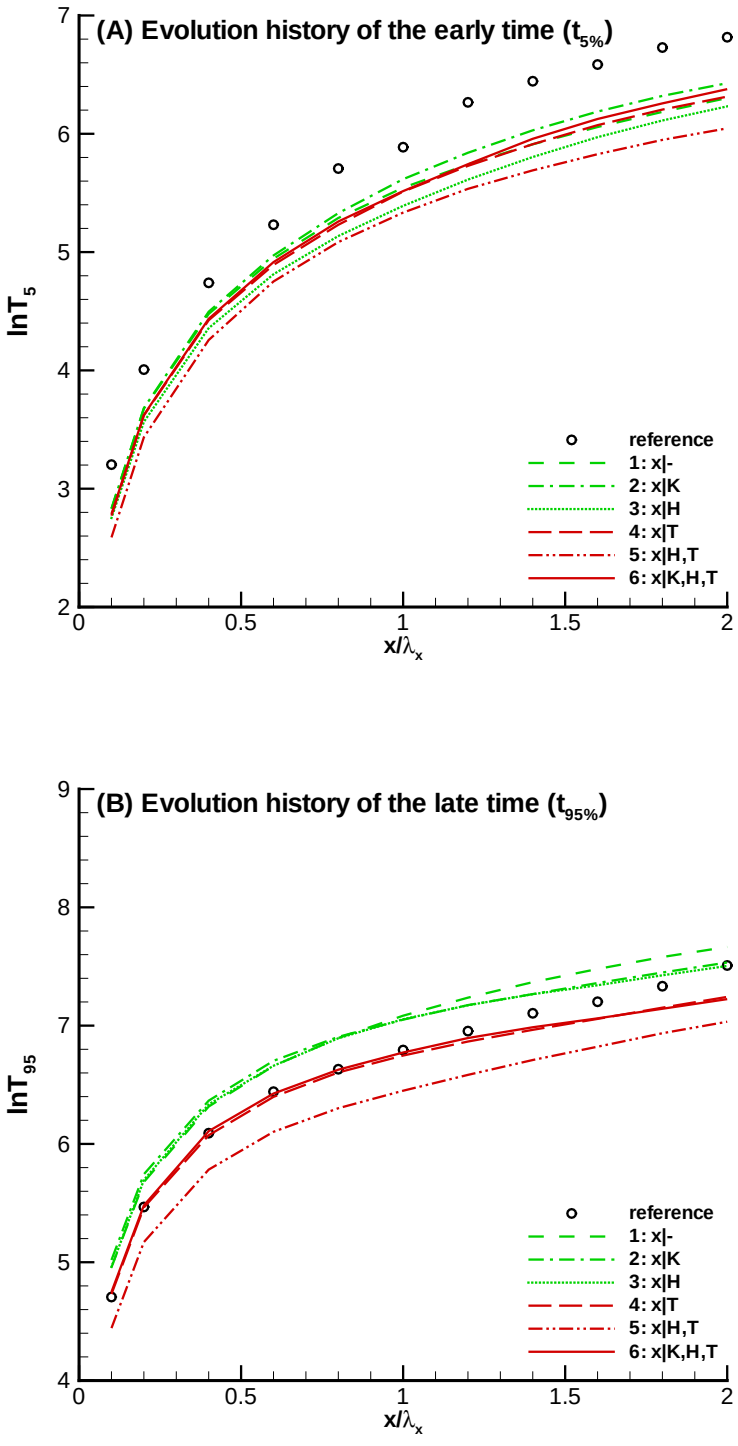


Figure 9: Ensemble average log-travel times at different control planes for all scenarios: (A) Early arrival time $t_{5\%}$, (B) late time $t_{95\%}$

4 Summary and Conclusions

We have presented and demonstrated a new algorithm for the stochastic generation of logconductivity realizations conditioned to logconductivity data, to piezometric head and travel time data based on the theory of blocking Markov chain Monte Carlo that, to the best of our knowledge, had not been used in hydrogeology earlier. The most interesting characteristic of this method is that uses a sampling algorithm for conditioning to state data (heads and travel times) instead of the common optimization approach used by other methods capable of conditioning to state data.

We have used a synthetic example to demonstrate the applicability of BMcMC and to analyze the interplay of different types of conditioning data in the characterization of the aquifer parameters and of its state. We found that individually, logconductivity data is the type of data that most reduces the uncertainty about logconductivity spatial distribution; and that piezometric head data is the type of data that most reduces the uncertainty about the piezometric head spatial distribution. However, all of them combined have always a reinforcing effect in characterizing any parameter or state variable.

It was found that logconductivity samples have mostly a local impact in uncertainty reduction, whereas piezometric head samples have a more intense impact. Travel time data, since it was collected at the edge of the aquifer, carries only global information about the spatial distribution of conductivities or heads, for the purpose of better characterizing the aquifer it should be used in combination with other data types.

However, travel time conditioning has an important impact in the reproduction of the apparent macrodispersion within the aquifer, and, in consequence, it is important for the better reproduction of plume spreading.

The results would probably not have been so good if the sampling density of conductivity and head had been larger (closer to the correlation range) as observed by *Dagan* (1985) and *Rubin and Dagan* (1992).

We have not discussed the specifics of how to use these models of uncertainty for network design, but it is easy to understand that the maps in Figure 3 and Figure 7 can guide in determining the best new sampling locations. For instance, it is clear that there are two zones, in the lower left and the upper right corners, that would benefit from two piezometric samples. These types of maps, together with the algorithm to generate conditional realizations, and the proper network design methodology can be used to minimize sampling costs while estimating aquifer responses to a specified precision (*James and Gorelick*, 1994). Apart from the economical factor in network design (*James and Gorelick*, 1994; *Criminisi et al.*, 1997), most algorithms construct the objective function either based on error reduction of state variables or on the decrease of prediction uncertainty (e.g., *McKinney and Loucks*, 1992), but seldom both. Our results shows that conditioning to different combinations of data types can meet these two objectives.

Acknowledgements

A doctoral fellowship and an extra travel grant awarded to the first author by the Universitat Politècnica de València, Spain, is gratefully acknowledged. The second author had a sabbatical grant by Technical University of Valencia during the preparation of this manuscript.

References

[1] Anderman, E.R., and M.C. Hill, 1999. A new multistage groundwater transport inverse method: Presentation, evaluation, and implications, *Water Resources Research*, 35(4), 1053-1063.

[2] Bellin, A., and Y. Rubin, 2004. On the use of peak concentration arrival time for the inference of hydrogeological parameters, *Water Resources Research*, 40(7), W07401.

[3] Carrera, J., and S.P. Neuman, 1986. Estimation of aquifer parameters under transient and steady state conditions, 1. Maximum likelihood method incorporating prior information, *Water Resources Research*, 22(2), 199-210.

[4] Cirpka, O.A., and P.K. Kitanidis, 2001. Sensitivity of temporal moments calculated by the adjoint-state method and joint inversing of head and tracer data, *Advances in Water Resources*, 24(1), 89-103.

[5] Criminisi, A., T. Tucciarelli, and G.P. Karatzas, 1997. A methodology to determine optimal transmissivity measurement locations in groundwater quality management models with scarce field information, *Water Resources Research*, 33(6), 1265-1274.

[6] Dagan, G., 1984. Solute transport in heterogeneous porous formations, *Journal of Fluid Mechanics*, 145, 151-177.

[7] Dagan, G., 1985. Stochastic modeling of groundwater flow by unconditional and conditional probabilities: the inverse problem, *Water Resources Research*, 21(1), 65-72.

[8] Deng, F.W., J.H. Cushman, and J.W. Delleur, 1993. Adaptive estimation of the log fluctuating conductivity from tracer data at the Cape Cod site, *Water Resources Research*, 29(12), 4011-4018.

[9] Desbarats, A.J., and R.M. Srivastava, 1991, Geostatistical characterization of groundwater flow parameters in a simulated aquifer, *Water Resources Research*, 27(5), 687-698.

[10] Eggleston, J.R., S.A. Rojstaczer, and J.J. Peirce, 1996. Identification of hydraulic conductivity structure in sand and gravel aquifers: Cape Cod data set, *Water Resources Research*, 32(5), 1209-1222.

[11] Ezzedine, S., and Y. Rubin, 1996. A geostatistical approach to the conditional estimation of spatially distributed solute concentration and notes on the use of tracer data in the inverse problem, *Water Resources Research*, 32(4), 853-861.

[12] Fernandez-Garcia, D., T.H. Illangasekare, and H. Rajaram, 2005. Differences in the scale-dependence of dispersivity estimated from temporal and spatial moments in chemically and physically heterogeneous porous media, *Advances in Water Resources*, 28(7), 745-759.

[13] Franssen, H.J.H., J.J. Gómez-Hernández, and A. Sahuquillo, 2003. Coupled inverse modeling of groundwater flow and mass transport and the worth of concentration data, *Journal of Hydrology*, 281, 281-295.

[14] Fu, J., 2008. A Markov Chain Monte Carlo Method for Inverse Stochastic Modeling and Uncertainty Assessment. Ph. D. Thesis. Universidad Politécnica de Valencia, Valencia, Spain, 140 pp.

[15] Fu, J., and J.J. Gómez-Hernández, 2009. A blocking Markov chain Monte Carlo method for inverse stochastic hydrogeological modeling, *Mathematical Geosciences*, in press.

[16] Geman, S., and D. Geman, 1984. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images, *IEEE transactions on Pattern Analysis and Machine Intelligence*, 6(6), 721-741.

[17] Gómez-Hernández, J.J., and A.G. Journel, 1993. Joint simulation of multiGaussian random variables. In Amilcar Soares, editor, Geostatistics Troia '92, volume 1, pages 85-94. Kluwer.

[18] Gómez-Hernández, J.J., A. Sahuquillo, and J.E. Capilla, 1997. Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data: I. Theory, *Journal of Hydrology*, 203, 162-174.

[19] Gómez-Hernández, J.J., and X.-H. Wen, 1998. To be or not to be multi-Gaussian? A reflection on stochastic hydrology, *Advances in Water Resources*, 21(1), 47-61.

[20] Graham, W., and D. McLaughlin, 1989a. Stochastic analysis of nonstationary subsurface solute transport: 1. Unconditional Moments, *Water Resources Research*, 25(2), 215-232.

[21] Graham, W., and D. McLaughlin, 1989b. Stochastic analysis of nonstationary subsurface solute transport: 2. Conditional Moments, *Water Resources Research*, 25(11), 2331-2355.

[22] Harter, T., and T.-C.J. Yeh, 1996. Conditional stochastic analysis of solute transport in heterogeneous, variably saturated soils, *Water Resources Research*, 32(6), 1597-1609.

[23] Harvey, C.F., and S.M. Gorelick, 1995. Mapping hydraulic conductivity: Sequential conditioning with measurements of solute arrival time, hydraulic head, and local conductivity, *Water Resources Research*, 31(7), 1615-1626.

[24] Hastings, W.K., 1970. Monte Carlo sampling methods using Markov chains and their application, *Biometrika*, 57(1), 97-109.

[25] James, B.R., and S.M. Gorelick, 1994. When enough is enough: The worth of monitoring data in aquifer remediation design, *Water Resources Research*, 30(12), 3499-3513.

[26] Khaleel, R., 1994. Scale and directional dependence of macrodispersivities in colonnade networks, *Water Resources Research*, 30(12), 3337-3355.

[27] Kreft, A., and A. Zuber, 1978. On the physical meaning of dispersion equation and its solutions for different initial and boundary conditions, *Chemical Engineer Sciences*, 33, 1471-1480.

[28] Lemke, L.D., W.A. Barrack II, L.M. Abriola, and P. Goovaerts, 2004. Matching solute breakthrough with deterministic and stochastic aquifer models, *Ground Water*, 42(6), 920-934.

[29] Liu, J.S., 1996. Metropolized independent sampling with comparisons to rejection sampling and importance sampling, *Statistics and Computing*, 6(2), 113-119.

[30] McKinney, D.C., and D.P. Loucks, 1992. Network design for predicting groundwater contamination, *Water Resources Research*, 28(1), 133-147.

[31] McLaughlin, D., and L.R. Townley, 1996. A reassessment of the groundwater inverse problem, *Water Resources Research*, 32(5), 1131-1161.

[32] Metropolis, N., A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, 1953. Equations of state calculations by fast computing machines, *Journal of Chemical Physics*, 21(3), 1087-1092.

[33] Nowak, W., and O.A. Cirpka, 2006. Geostatistical inference of hydraulic conductivity and dispersivities from hydraulic heads and tracer data, *Water Resources Research*, 42(8), W08416.

[34] Oliver, D.S., L.B. Cunha, and A.C. Reynolds, 1997. Markov chain Monte Carlo methods for conditioning a log-permeability field to pressure data, *Mathematical Geology*, 29(1), 61-91.

[35] Roberts, G.O., and S.K. Sahu, 1997. Updating schemes, correlation structure, blocking and parameterization for the Gibbs sampler, *Journal of Royal Statistical Society B*, 59(2), 291-317.

[36] Rubin, Y., 1991. Prediction of tracer plume migration in disordered porous media by the method of conditional probabilities, *Water Resources Research*, 27(6), 1291-1308.

[37] Rubin, Y., and G. Dagan, 1992. Conditional estimation of solute travel time in heterogeneous formations: Impact of transmissivity measurements, *Water Resources Research*, 28(4), 1033-1040.

[38] Rubin, Y., and S. Ezzedine, 1997. The travel times of solutes at the Cape Cod tracer experiment: Data analysis, modeling, and structural parameters inference, *Water Resources Research*, 33(7), 1537-1547.

[39] Scheibe, T.D., and Y.-J. Chien, 2003. An evaluation of conditioning data for solute transport prediction, *Ground Water*, 41(2), 128-141.

[40] Sun, N.-Z., and W.W.-G. Yeh, 1990a. Coupled inverse problems in groundwater modeling: 1. Sensitivity analysis and parameter identification, *Water Resources Research*, 26(10), 2507-2525.

[41] Sun, N.-Z., and W.W.-G. Yeh, 1990b. Coupled inverse problems in groundwater modeling: 2. Identifiability and experimental design, *Water Resources Research*, 26(10), 2527-2540.

[42] van Leeuwen, M., A.P. Butler, C.B.M. te Stroet, and J.A. Tompkins, 2000. Stochastic determination of well capture zones conditioned on regular grids of transmissivity measurements, *Water Resources Research*, 36(4), 949-957.

[43] Vasco, D.W., and A. Datta-Gupta, 1999. Asymptotic solutions for solute transport: A formalism for tracer tomography, *Water Resources Research*, 35(1), 1-16.

[44] Wagner, B.J., and S.M. Gorelick, 1989. Reliable aquifer remediation in the presence of spatial variable hydraulic conductivity: from data to design, *Water Resources Research*, 25(10), 2211-2225.

[45] Wen, X.-H., and J.J. Gómez-Hernández, 1996. The constant displacement scheme for tracking particles in heterogeneous aquifers, *Ground Water*, 34(1), 135-142.

[46] Wen, X.-H., and J.J. Gómez-Hernández, 1998. Numerical modeling of macrodispersion in heterogeneous media: a comparison of multi-Gaussian and non-multi-Gaussian models, *Journal of Contaminant Hydrology*, 30, 129-156.

[47] Wen, X.-H., C.V. Deutsch, and A.S. Cullick, 2002. Construction of geostatistical aquifer models integrating dynamic flow and tracer data using inverse technique, *Journal of Hydrology*, 255, 151-168.

[48] Wilson, A., and Y. Rubin, 2002. Characterization of aquifer heterogeneity using indicator variables for solute concentrations, *Water Resources Research*, 38(12), 1283.

[49] Woodbury, A.D., and Y. Rubin, 2000. A full-Bayesian approach to parameter inference from tracer travel time moments and investigation of scale effects at the Cape Cod experimental site, *Water Resources Research*, 36(1), 159-171.

[50] Woodbury, A.D., and E.A. Sudicky, 1992. Inversion of the Borden Tracer Experimental Data: Investigation of stochastic moment models, *Water Resources Research*, 28(9), 2387-2398.