Modeling Transient Groundwater Flow by Coupling Ensemble Kalman Filtering and Upscaling

Liangping Li,[1,2] Haiyan Zhou,[1,2] Harrie-Jan Hendricks Franssen[3] and J. Jaime Gómez-Hernández,[2]

Liangping Li and Haiyan Zhou, School of Water Resources and Environment, China University of Geosciences (Beijing), Beijing 100083, China.

Liangping Li, Haiyan Zhou and Jaime Gómez-Hernández, Group of Hydrogeology, Universitat Politècnica de València, Camino de Vera, s/n, 46022 Valencia, Spain. (liali@upvnet.upv.es)

Harrie-Jan Hendricks Franssen,Agrosphere, ICG-4, Forschungszentrum Jülich GmbH, Germany

[1]School of Water Resources and Environment, China University of Geosciences (Beijing), China

[2]Group of Hydrogeology, Universitat Politècnica de València, Spain

[3]Agrosphere, ICG-4, Forschungszentrum Jülich GmbH, Germany

Abstract. The ensemble Kalman filter (EnKF) is coupled with upscaling to build an aquifer model at a coarser scale than the scale at which the conditioning data (conductivity and piezometric head) had been taken for the purpose of inverse modeling. Building an aquifer model at the support scale of observations is most often impractical, since this would imply numerical models with millions of cells. If, in addition, an uncertainty analysis is required involving some kind of Monte-Carlo approach, the task becomes impossible. For this reason, a methodology has been developed that will use the conductivity data, at the scale at which they were collected, to build a model at a (much) coarser scale suitable for the inverse modeling of groundwater flow and mass transport. It proceeds as follows: (i) generate an ensemble of realizations of conductivities conditioned to the conductivity data at the same scale at which conductivities were collected, (ii) upscale each realization onto a coarse discretization; on these coarse realizations, conductivities will become tensorial in nature with arbitrary orientations of their principal directions, (iii) apply the EnKF to the ensemble of coarse conductivity upscaled realizations in order to condition the realizations to the measured piezometric head data. The proposed approach addresses the problem of how to deal with tensorial parameters, at a coarse scale, in ensemble Kalman filtering, while maintaining the conditioning to the fine scale hydraulic conductivity measurements. We demonstrate our approach in the framework of a synthetic worth-of-data exercise, in which the relevance of conditioning to conductivities, piezometric heads or both is analyzed.

1. Introduction

In this paper we address two problems, each of which has been the subject of many works, but which have not received as much attention when considered together: upscaling and inverse modeling. There are many reviews on the importance and the methods of upscaling [e.g., *Wen and Gómez-Hernández*, 1996; *Renard and de Marsily*, 1997; *Sánchez-Vila et al.*, 2006], and there are also many reviews on inverse modeling and its relevance for aquifer characterization [e.g., *Yeh*, 1986; *McLaughlin and Townley*, 1996; *Zimmerman et al.*, 1998; *Carrera et al.*, 2005; *Hendricks Franssen et al.*, 2009; *Oliver and Chen*, 2011; *Zhou et al.*, 2011a]. Our interest lies in coupling upscaling and inverse modeling to perform an uncertainty analysis of flow and transport in an aquifer for which measurements have been collected at a scale so small that it is prohibitive, if not impossible, to perform directly the inverse modeling.

The issue of how to reconcile the scale at which conductivity data are collected and the scale at which numerical models are calibrated was termed "the missing scale" by *Tran* [1996], referring to the fact that the discrepancy between scales was simply disregarded; data were collected at a fine scale, the numerical model was built at a much larger scale, each datum was assigned to a given block, and the whole block was assigned the datum value, even though the block may be several orders of magnitude larger than the volume support of the sample. This procedure induced a variability, at the numerical block scale, much larger than it should be, while at the same time some unresolved issues have prevailed like what to do when several samples fell in the same block.

To the best of our knowledge, the first work to attempt the coupling of upscaling and inverse modeling is the upscaling-calibration-downscaling-upscaling approach by *Tran et al.* [1999]. In their approach, a simple averaging over a uniformly coarsened model is used to upscale the hydraulic conductivities. Then, the state information (e.g., dynamic piezometric head data) is incorporated in the upscaled model by the self-calibration technique [*Gómez-Hernández et al.*, 1997]. The calibrated parameters are downscaled back to the fine scale by block kriging [*Behrens et al.*, 1998] resulting in a fine scale realization conditional to the measured parameters (e.g., hydraulic conductivities). Finally, the downscaled conductivities are upscaled using a more precise scheme [*Durlofsky et al.*, 1997; *Li et al.*, 2011a] for prediction purposes. The main shortcoming of this approach is that the inverse modeling is performed on a crude upscaled model, resulting in a downscaled model that will not honor the state data accurately. *Tureyen and Caers* [2005] proposed the calibration of the fine scale conductivity field by gradual deformation [*Hu*, 2000; *Capilla and Llopis-Albert*, 2009], but instead of solving the flow equation at the fine scale they used an approximate solution after upscaling the hydraulic conductivity field to a coarse scale. This process requires an upscaling for each iteration of the gradual deformation algorithm, which is also time-consuming, although they avoid the fine scale flow solution. More recently, an alternative multiscale inverse method [*Fu et al.*, 2010] was proposed. It uses a multiscale adjoint method to compute sensitivity coefficients and reduce the computational cost. However, like traditional inverse methods, the proposed approach requires a large amount of CPU time in order to get an ensemble of conditional realizations. In our understanding, nobody has attempted to couple upscaling and the ensemble Kalman filtering (EnKF) for generating hydraulic conductivity fields conditioned to both hydraulic conductivity and piezometric head measurements. Only the work by *Peters et al.* [2010] gets close to our work as. For the Brugge Benchmark Study, they generated a fine scale permeability field, which was upscaled using a diagonal tensor upscaling. The resulting coarse scale model was provided to the different teams participating in the benchmark exercise, some of which used the EnKF for history matching. We have chosen the EnKF algorithm for the inverse modeling because it has been shown that it is faster than other alternative Monte Carlo-based inverse modeling methods (see for instance the work by *Hendricks Franssen and Kinzelbach* [2009] who show that the EnKF was 80 times fastar than the sequential self-calibration in a benchmark exercise and nearly as good).

Our aim is to propose an approach for the stochastic inverse modeling of an aquifer that has been characterized at a scale at which it is impossible to solve the inverse problem, due to the large number of cells needed to discretize the domain. We start with a collection of hydraulic conductivity and piezometric head measurements, taken at a very small scale, to end with an ensemble of hydraulic conductivity realizations, at a scale much larger than the one at which data were originally sampled, all of which are conditioned to the measurements. This ensemble of realizations will serve to perform uncertainty analyses of both the parameters (hydraulic conductivities) and the system state variables (piezometric heads, fluxes, concentrations, or others).

The rest of the paper is organized as follows. Section 2 outlines the coupling of upscaling and the EnKF, with emphasis in the use of arbitrary hydraulic conductivity tensors in the numerical model. Next, in section 3, a synthetic example serves to validate the proposed method. Then, in section 4, the results are discussed. The paper ends with a summary and conclusions.

2. Methodology

Hereafter, we will refer to a fine scale for the scale at which data are collected, and a coarse scale, for the scale at which the numerical models are built. The methodology proposed can be outlined as follows:

1. At the fine scale, generate an ensemble of realizations of hydraulic conductivity conditioned to the hydraulic conductivity measurements.

2. Upscale each one of the fine scale realizations generated in the previous step. In the most general case, the upscaled conductivities will be full tensors in the reference axes.

3. Use the ensemble of coarse realizations with the EnKF to condition (assimilate) on the measured piezometric heads.

2.1. Generation of the Ensemble of Fine Scale Conductivities

The first step of the proposed methodology makes use of geostatistical tools already available in the literature [e.g., *Gómez-Hernández and Srivastava*, 1990; *Deutsch and Journel*, 1998; *Strebelle*, 2002; *Mariethoz et al.*, 2010]. The technique to choose will depend on the underlying random function model selected for the hydraulic conductivity: multi-Gaussian, indicator-based, pattern-based, or others. In all cases, the scale at which these fields can be generated is not an obstacle, and the resulting fields will be conditioned to the measured hydraulic conductivity measurements (but only to hydraulic conductivity measurements). These fields could have millions of cells and are not suitable for inverse modeling of groundwater flow and solute transport.

2.2. Upscaling

Each one of the realizations generated in the previous step is upscaled onto a coarse grid with a number of blocks sufficiently small for numerical modeling. We use the flow upscaling approach by *Rubin and Gómez-Hernández* [1990] who, after spatially integrating Darcy's law over a block V,

$$\frac{1}{V}\int_V \bar{\mathbf{q}}dV = -\mathbf{K}^b\Big(\frac{1}{V}\int_V \overline{\nabla\mathbf{h}}\,dV\Big), \qquad (1)$$

define the block conductivity tensor ($\mathbf{K}^b$) as the tensor that best relates the block average head gradient ($\overline{\nabla\mathbf{h}}$) to the block average specific discharge vector ($\bar{\mathbf{q}}$) within the block. Notice that to perform the two integrals in the previous expressions we need to know the specific discharge vectors and the piezometric head gradients at the fine scale within the block. These values could be obtained after a solution of the flow problem at the fine scale [i.e., *White and Horne*, 1987], but this approach beats the whole purpose of upscaling, which is to avoid such fine scale numerical simulations. The alternative is to model a smaller domain of the entire aquifer enclosing the block being upscaled. In such a case, the boundary conditions used in this reduced model will be different from the boundary conditions that the block has in the global model, and this will have some impact on the fine scale values of $\overline{\nabla\mathbf{h}}$ and $\bar{\mathbf{q}}$. The dependency of the heads and flows within the block on the boundary conditions is the reason why the block upscaled tensor is referred to as non-local [e.g., *Indelman and Abramovich*, 1994; *Guadagnini and Neuman*, 1999].

For the flow upscaling we adopt the so-called Laplacian-with-skin method on block interfaces as described by *Gómez-Hernández* [1991] and recently extended to three dimensions by *Zhou et al.* [2010]. The two main advantages of this approach are that it can handle arbitrary full conductivity tensors, without any restriction on their principal directions; and that it upscales directly the volume straddling between adjacent block centers, which, at the end, is the parameter used in the standard finite-difference approximation of the groundwater flow equation (avoiding the derivation of this value by some kind of averaging of the adjacent block values). Once the interblock conductivities have been computed, a specialized code capable of handling interblock tensors is necessary. For this purpose, the public domain code FLOWXYZ3D [*Li et al.*, 2010], has been developed. The details of the upscaling approach, the numerical modeling using interblock conductivity tensors, and several demonstration cases can be found in *Zhou et al.* [2010]; *Li et al.* [2010, 2011a, b]. The resulting upscaled interblock tensors produced by this approach are always of rank two, symmetric and positive definite.

The Laplacian-with-skin method on block interfaces for a given realization can be briefly summarized as follows:

- Overlay a coarse grid on the fine scale hydraulic conductivity realization.
- Define the interblock volumes that straddle any two adjacent blocks.
- For each interblock:

  – Isolate the fine scale conductivities within a volume made up by the interblock plus an additional "border ring" or "skin" and simulate flow, at the fine scale, within this volume.

  – As explained in many studies [e.g, *Gómez-Hernández*, 1991; *Sánchez-Vila et al.*, 1995; *Sánchez-Vila et al.*, 2006; *Zhou et al.*, 2010; *Li et al.*, 2011a], there is a need to solve more than one flow problem in order to being able of identifying all components of the interblock conductivity tensor.

  – From the solution of the flow problems, use Equation (1) to derive the interblock conductivity tensor.

- Assemble all interblock tensors to build a realization of upscaled hydraulic conductivity tensors at the coarse scale.

The above procedure has to be repeated for all realizations, ending up with an ensemble of realizations of interblock conductivity tensors.

2.3. The EnKF with Hydraulic Conductivity Tensors

Extensive descriptions of the EnKF and how to implement it have been given, for instance, by *Burgers et al.* [1998]; *Evensen* [2003]; *Naevdal et al.* [2005]; *Chen and Zhang* [2006]; *Aanonsen et al.* [2009]. Our contribution, regarding the EnKF, is how to deal with an ensemble of parameters that, rather than being scalars, are tensors. After testing different alternatives, we finally decided not to use the tensor components corresponding to the Cartesian reference system as parameters within the EnKF, but to use some of the tensor invariants, more precisely, the magnitude of the principal components and the angles that define their orientation.

For the example discussed later we will assume a two-dimensional domain, with hydraulic conductivity tensors varying in space $\mathbf{K} = \mathbf{K}(\mathbf{x})$ of the form

$$\mathbf{K} = \begin{bmatrix} K_{xx} & K_{xy} \\ K_{xy} & K_{yy} \end{bmatrix}. \qquad (2)$$

Each conductivity tensor is converted onto a triplet $\{K_{max}, K_{min}, \theta\}$, with K_{max} being the largest principal component, K_{min}, the smallest one, and θ, the orientation, of the maximum principal component with respect to the x-axis according to the following expressions

[*Bear*, 1972]:

$$\begin{aligned}
K_{max} &= \frac{K_{xx}+K_{yy}}{2}+\left[\left(\frac{K_{xx}-K_{yy}}{2}\right)^2+\left(K_{xy}\right)^2\right]^{1/2}, \\
K_{min} &= \frac{K_{xx}+K_{yy}}{2}-\left[\left(\frac{K_{xx}-K_{yy}}{2}\right)^2+\left(K_{xy}\right)^2\right]^{1/2}, \\
\theta &= \frac{1}{2}\arctan\left(\frac{2K_{xy}}{K_{xx}-K_{yy}}\right).
\end{aligned} \tag{3}$$

After transforming all conductivity tensors obtained in the upscaling step onto their corresponding triplets, we are ready to apply the EnKF. We will use the EnKF implementation with an augmented state vector as discussed below; this is the standard implementation used in petroleum engineering and hydrogeology, although alternative implementations and refinements of the algorithm could have been used [see *Aanonsen et al.*, 2009, for a review].

Using the EnKF nomenclature, the state of the system is given by the spatial distribution of the hydraulic heads, the state transition equation is the standard flow equation describing the movement of an incompressible fluid in a fully saturated porous medium [*Bear*, 1972; *Freeze and Cherry*, 1979] (in two dimensions for the example considered later), and the parameters of the system are the spatially varying hydraulic conductivities (the storage coefficient is assumed to be homogeneous and known, and therefore, it is a parameter not subject to filtering), i.e.,

$$\mathbf{Y}_k = f(\mathbf{X}_{k-1}, \mathbf{Y}_{k-1}), \tag{4}$$

where $\mathbf{Y}_k$ is the state of the system at time step t_k, f represents the groundwater flow model (including boundary conditions, external stresses, and known parameters), and $\mathbf{X}_{k-1}$ represents the model parameters after the latest update at time t_{k-1}.

The EnKF algorithm will proceed as follows:

1. Forecast. Equation (4) is used to forecast the system states for the next time step given the latest state and the latest parameter update. This forecast has to be performed in all realizations of the ensemble.

2. Analysis. At the forecasted time step, new state observations are available at measurement locations. The discrepancy between these state observations and the forecasted values will serve to update both the parameter values and the system state at all locations in the aquifer model as follows:

(i) Build the joint vector $\mathbf{\Psi}_k$, including parameters and state values. This vector can be split into as many members as there are realizations in the ensemble, with

$$\mathbf{\Psi}_{k,j} = \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix}_{k,j} \tag{5}$$

being the j^{th} ensemble member at time t_k. Specifically, $\mathbf{X}$ (for a realization) is expressed as:

$$\mathbf{X} = [(\ln K_{max}, \ln K_{min}, \theta)_1, \ldots, (\ln K_{max}, \ln K_{min}, \theta)_{N_b}]^T \tag{6}$$

where N_b is the number of interfaces in the coarse numerical model. Notice that the logarithm of the conductivity principal components is used, since their distribution is, generally, closer to Gaussian than that of the conductivities themselves, which results in the optimality in the performance of the EnKF [*Evensen*, 2003; *Zhou et al.*, 2011b; *Schöniger et al.*, 2011].

(ii) The joint vector $\mathbf{\Psi}_k$ is updated, realization by realization, by assimilating the observations ($\mathbf{Y}_k^{obs}$):

$$\mathbf{\Psi}_{k,j}^a = \mathbf{\Psi}_{k,j}^f + \mathbf{G}_k \Big(\mathbf{Y}_k^{obs} + \boldsymbol{\epsilon} - \mathbf{H}\mathbf{\Psi}_{k,j}^f \Big), \tag{7}$$

where the superscripts a and f denote analysis and forecast, respectively; $\boldsymbol{\epsilon}$ is a random observation error vector; $\mathbf{H}$ is a linear operator that interpolates the forecasted heads to the measurement locations, and, in our case, is composed of $0's$ and $1's$ since we assume that measurements are taken at block centers. Therefore, equation (7) can be rewritten as:

$$\boldsymbol{\Psi}^a_{k,j} = \boldsymbol{\Psi}^f_{k,j} + \mathbf{G}_k\Big(\mathbf{Y}^{obs}_k + \boldsymbol{\epsilon} - \mathbf{Y}^f_{k,j}\Big), \tag{8}$$

where the Kalman gain $\mathbf{G}_k$ is given by:

$$\mathbf{G}_k = \mathbf{P}^f_k\mathbf{H}^T\Big(\mathbf{H}\mathbf{P}^f_k\mathbf{H}^T + \mathbf{R}_k\Big)^{-1}, \tag{9}$$

where $\mathbf{R}_k$ is the measurement error covariance matrix, and $\mathbf{P}^f_k$ contains the covariances between the different components of the state vector. $\mathbf{P}^f_k$ is estimated from the ensemble of forecasted states as:

$$\begin{aligned}\mathbf{P}^f_k &\approx E\left[\Big(\boldsymbol{\Psi}^f_{k,j} - \overline{\boldsymbol{\Psi}}^f_{k,j}\Big)\Big(\boldsymbol{\Psi}^f_{k,j} - \overline{\boldsymbol{\Psi}}^f_{k,j}\Big)^T\right] \\ &\approx \sum_{j=1}^{N_e} \frac{\Big(\boldsymbol{\Psi}^f_{k,j} - \overline{\boldsymbol{\Psi}}^f_{k,j}\Big)\Big(\boldsymbol{\Psi}^f_{k,j} - \overline{\boldsymbol{\Psi}}^f_{k,j}\Big)^T}{N_e},\end{aligned} \tag{10}$$

where N_e is the number of realizations in the ensemble, and the overbar denotes average through the ensemble.

In the implementation of the algorithm, it is not necessary to calculate explicitly the full covariance matrix $\mathbf{P}^f_k$, since the matrix $\mathbf{H}$ is very sparse, and, consequently, the matrices $\mathbf{P}^f_k\mathbf{H}^T$ and $\mathbf{H}\mathbf{P}^f_k\mathbf{H}^T$ can be computed directly at a strongly reduced CPU cost.

3. The updated state becomes the current state, and the forecast-analysis loop is started again.

The question remains whether the updated conductivity-tensor realizations preserve the conditioning to the fine scale conductivity measurements. In standard EnKF, when no upscaling is involved and conductivity values are the same in all realizations at conditioning locations, the forecasted covariances and cross-covariances involving conditioning points are zero, and so is the Kalman gain at those locations; therefore, conductivities remain unchanged through the entire Kalman filtering. In our case, after upscaling the fine-scale conditional realizations, the resulting ensemble of hydraulic conductivity tensor realizations will display smaller variances (through the ensemble) for the tensors associated with interfaces close to the fine scale measurements than for those far from the measurements. These smaller variances will result in a smaller Kalman gain in the updating process at these locations, and therefore will induce a soft conditioning of the interblock tensors on the fine scale measurements.

The proposed method is implemented in the C software Upscaling-EnKF3D, which is used in conjunction with the finite-difference program FLOWXYZ3D [*Li et al.*, 2010] in the forecasting step. From an operational point of view, the proposed approach is suitable for parallel computation both in terms of upscaling and EnKF, since each ensemble member is treated independently, except for the computation of the Kalman gain.

2.4. CPU time analysis

Without a CPU analysis, we can argue that the coupling of upscaling with the EnKF is of interest because it allows to analyze problems that otherwise could not be handled simply because the size of the numerical model is not amenable to the available computer resources. In our case, with our resources, we could not run any flow model with more than 10^8 nodes. However, even for those models for which we could run the fine scale flow simulation, the CPU time savings associated to the upscaling approach are considerable and worth considering for fine scale models with more than a few tens of thousands of nodes.

We performed a conservative analysis of CPU time savings in which only the CPU time spent in the flow simulations is considered, the savings will be larger when the time needed to estimate the ensemble covariance and the Kalman gain are considered. We run several flow simulations for model sizes ranging from 10^4 to 10^7 nodes, for different realizations of the hydraulic conductivities with the same statistical characteristics as the examples that will be shown later. The regression of the CPU times with respect to the number of nodes (Figure 1) gives the following expression:

$$CPUt = 10^{-5} N_{cells} \tag{11}$$

A conservative CPU time analysis has been performed in order to

3. Application Example

In this section, a synthetic experiment illustrates the effectiveness of the proposed coupling of EnKF and upscaling.

3.1. Reference Field

We generate a realization of hydraulic conductivity over a domain discretized into 350 by 350 cells of 1 m by 1 m using the code GCOSIM3D [*Gómez-Hernández and Journel*, 1993].

We assume that, at this scale, conductivity is scalar and its natural logarithm, $\ln K$, can be characterized by a multiGaussian distribution of mean -5 (ln cm/s) and unit variance, with a strong anisotropic spatial correlation at the 45° orientation. The correlation range in the largest continuity direction (x') is $\lambda_{x'} = 90$ m and in the smallest continuity direction (y') is $\lambda_{y'} = 18$ m. The Gaussian covariance function is given by:

$$\gamma(\mathbf{r}) = 1.0 \cdot \left\{ 1 - \exp\left[- \sqrt{\left(\frac{3r_{x'}}{90}\right)^2 + \left(\frac{3r_{y'}}{18}\right)^2} \right] \right\}, \tag{12}$$

with

$$\begin{bmatrix} r_{x'} \\ r_{y'} \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} r_x \\ r_y \end{bmatrix}, \tag{13}$$

and $r = (r_x, r_y)$ being the separation vector in Cartesian coordinates. The reference realization is shown in Figure 4A. From this reference realization 100 conductivity data are sampled at the locations shown in Figure 4B. These data will be used for conditioning.

The forward transient groundwater flow model is run in the reference realization with the boundary conditions shown in Figure 5 and initial heads equal to zero everywhere. The total simulation time is 500 days, discretized into 100 time steps following a geometric sequence of ratio 1.05. The aquifer is confined. Specific storage is assumed constant and equal to 0.003 m^{-1}. The simulated piezometric heads at the end of time step 60 (67.7 days) are displayed in Figure 6. Piezometric heads at locations W1 to W9 in Figure 5 are sampled for the first 60 time steps to be used as conditioning data. The simulated heads at locations W10 to W13 will be used as validation data.

3.2. Hydraulic Conductivity Upscaling

For the reasons explained by *Zhou et al.* [2010]; *Li et al.* [2010], the fine scale realizations must be slightly larger than the aquifer domain in order to apply the Laplacian-with-skin upscaling approach. We assume that the aquifer of interest is comprised by the inner 320 by 320 cell domain for all realizations. Each one of these realizations is upscaled onto a 32 by 32 square-block model implying an order-of-two magnitude reduction in the discretization of the aquifer after upscaling. After several tests, the skin selected for the upscaling procedure has a width of 10 m, since it is the one that gives best results in the reproduction of the interblock specific discharges when compared to those computed on the fine scale underlying realizations.

Since the upscaling is applied to the interblock volume straddling between adjacent block centers, there are 32 by 31 column-to-column interblock tensors ($\mathbf{K}^{b,c}$) plus 31 by 32 row-to-row interblock tensors ($\mathbf{K}^{b,r}$). All interblock tensors are transformed into their corresponding triplet of invariants prior to starting the EnKF algorithm.

For illustration purposes, Figure 7 shows the resulting triplets for the reference field. This figure will be used later as the reference upscaled field to analyze the performance of the proposed method. On the right side of Figure 6, the simulated piezometric heads at the end of the 60th time step are displayed side by side with the simulated piezometric heads at the fine scale. The reproduction of the fine scale spatial distribution by the coarse scale simulation is, as can be seen, very good; the average absolute discrepancy between the heads at the coarse scale and heads at the fine scale (on the block centers) is only 0.087 m.

3.3. Case Studies

Four cases, considering different types of conditioning information, are analyzed to study the performance of the proposed approach (see Table 1). They will show that the coupling of the EnKF with upscaling can be used to construct aquifer models that are conditional to conductivity and piezometric head data, when there is an important discrepancy between the scale at which the data are collected and the scale at which the flow model is built. The cases will serve also to carry out a standard worth-of-data exercise in which we analyze the trade-off between conductivity data and piezometric head data regarding aquifer characterization.

Case A is unconditional, 200 realizations are generated according to the spatial correlation model given by Equation (11) at the fine scale. Upscaling is performed in each realization and the flow model is run. No Kalman filtering is performed.

Case B is conditional to logconductivity measurements, 200 realizations of logconductivity conditional to the 100 logconductivity measurements of Figure 3B are generated at the fine scale. Upscaling is performed in each realization and the flow model is run. No Kalman filtering is performed.

Cases A and B act as base cases to be used for comparison when the piezometric head data are assimilated through the EnKF.

Case C is conditional to piezometric heads. The same 200 coarse realizations from Case A serve as the initial ensemble of realizations to be used by the EnKF to assimilate the piezometric head measurements from locations W1 to W9 for the first 60 time steps (66.7 days).

Case D is conditional to both logconductivity and piezometric heads. The same 200 coarse realizations from Case B serve as the initial ensemble of realizations to be used by the EnKF to assimilate the piezometric head measurements from locations W1 to W9 for the first 60 time steps (66.7 days).

In Cases C and D we use the measured heads obtained at the fine scale in the reference realization as if they were measurements obtained at the coarse scale. There is an error in this assimilation that we incorporate into the measurement error covariance matrix. Specifically we here assumed a diagonal error covariance matrix, with all the diagonal terms equal to 0.0025 m^2; this value is approximately equal to the average dispersion variance of the fine scale piezometric heads within the coarse scale blocks.

3.4. Performance Measurements

Since this is a synthetic experiment, the "true" aquifer response, evaluated at the fine scale, is known. We also know the upscaled conductivity tensors for the reference aquifer, which we will use to evaluate the performance of the updated conductivity tensors produced by the EnKF.

The following criteria, some of which are commonly applied for optimal design evaluation [*Nowak*, 2010], will be used to analyze the performance of the proposed method and the worth of data:

1. Ensemble mean map. (It should capture the main patterns of variability of the reference map.)

2. Ensemble variance map. (It gives an estimate of the precision of the maps.)

3. Ensemble average absolute bias map, ϵ_X, made up by:

$$\epsilon_{X_i} = \frac{1}{N_e} \sum_{r=1}^{N_e} |X_{i,r} - X_{i,ref}|, \tag{14}$$

where X_i is the parameter being analyzed, at location i, $X_{i,r}$ represents its value for realization r, $X_{i,ref}$ is the reference value at location i, and N_e is the number of realizations of the ensemble (200, in this case). (It gives an estimate of the accuracy of the maps.)

4. Average absolute bias

$$AAB(X) = \frac{1}{N_b} \sum_{i=1}^{N_b} \epsilon_{X_i}, \tag{15}$$

where N_b is the number of interblocks when X is coarse logconductivity tensor component, or the number of blocks when X is piezometric head. (It gives a global measure of accuracy.)

5. Square root of the average ensemble spread

$$AESP(X) = \left[\frac{1}{N_b} \sum_{i=1}^{N_b} \sigma_{X_i}^2 \right]^{1/2}, \tag{16}$$

where $\sigma_{X_i}^2$ is the ensemble variance at location i. (It gives a global measure of precision.)

6. Comparison of the time evolution of the piezometric heads at the conditioning piezometers W1 to W9, and at the control piezometers W10 to W13. (It evaluates the capability of the EnKF to update the forecasted piezometric heads using the measured values.)

4. Discussion

Ensembles of coarse realizations for the four cases have been generated according to the conditions described earlier. Figure 8 shows the evolution of the piezometric heads in piezometers W1 and W9 for the 500 days of simulation; the first 60 steps (66.7 days) were used for conditioning in cases C and D. Similarly, Figure 9 shows piezometers W10 and W13; these piezometers were not used for conditioning. Figure 10 shows the ensemble mean and variance of the piezometric heads at the 60th time step, while Figure 11 shows the ensemble average absolute bias. Figure 12 shows the ensemble mean and variance of $\ln(K_{max})$ for interblocks between rows, and Figure 13 shows the ensemble average absolute bias. Finally, Table 2 shows the metric performance measurements for $\ln(K_{max})$ between rows and for piezometric heads at the 30th, 60th and 90th time steps.

4.1. The EnKF Coupled with Upscaling

The EnKF has the objective of updating conductivity realizations so that the solution of the flow equation on the updated fields will match the measured piezometric heads. Analyzing cases C and D in Figure 8, we can observe how the updated fields, when piezometric head is assimilated by the EnKF, produce piezometric head predictions that reproduce the measured values very well, particularly when compared with case A, which corresponds to the case in which no conditioning data are considered. Notice also that piezometric head data are assimilated only for the first 66.7 days (the period in which the heads are almost perfectly reproduced in the EnKF updated fields) while the rest of the simulation period serves as validation. Additional validation of the EnKF generated realizations is given in Figure 9 that shows two of the piezometers not used for conditioning; we can also observe the improvement in piezometric head reproduction for cases C and D as compared to case A. Furthermore, the analysis of Figure 10 shows how, for cases C and D, the average spatial distribution, at the end of time step 60, follows closely the reference piezometric head distribution, while the ensemble variance is reduced to very small values everywhere. The ensemble average head bias is also noticeably reduced when conditioning to heads, not just at the conditioning locations (as expected) but also elsewhere. A final analysis to show how conditioning to the heads improves the overall reproduction of the head spatial distribution is by looking at the metrics displayed in Table 2. Comparing cases B and C, it is interesting to notice the increasing impact of the conditioning to piezometric heads as time passes; at time step 30, the initial effect of just conditioning to hydraulic conductivity measurements (which occurs from time step 0) is still larger than just conditioning to the heads measured during the first 30 time steps, but at time step 60, this effect is clearly reversed, and it is maintained to time step 90 even though the heads between steps 60 and 90 are not used for conditioning. As expected, conditioning to both piezometric heads and hydraulic conductivities gives the best results in terms of smallest bias and smallest spread.

From this analysis we conclude that the EnKF coupled with upscaling is able to generate an aquifer model at a scale two orders of magnitude coarser than the reference aquifer scale that is conditional to the piezometric heads.

Besides achieving the original goal of the EnKF algorithm, it is also important to contrast the final conductivity model given by the EnKF, with the reference aquifer model. For this purpose we will compare the final ensemble of realizations obtained for cases C and D with the upscaled realization obtained from the reference, fine scale aquifer model. Conditioning to piezometric head data should improve the characterization of the logconductivities. Indeed, this is what happens as it can be seen when analyzing Figures 12 and 13 and Table 2. In these figures only the maximum component of the logconductivity tensors for the interblocks between rows is displayed, but the members of the triplet for the tensor between rows, as well as the members of the triplet for the tensors between columns, show a similar behavior. The ensemble mean maps are closer to the reference map in case that conditioning data are used; the variance maps display smaller values as compared to case A; and the bias map shows values closer to zero than in case A. All in all, we can conclude that the EnKF updates the block conductivity tensors to produce realizations which get closer to the aquifer model obtained after upscaling the reference aquifer.

There remains the issue of conditioning to the fine scale conductivity measurements. Since the fine scale conductivity measurements were used to condition the fine scale realizations, the conditioning should be noticed in the upscaled model only if the correlation scale of the conductivity measurements is larger than the upscaled block size. In such a case (as is the case for the example), the ensemble variance of the upscaled block conductivity values should be smaller for blocks close to conditioning datum locations than for those away from the conditioning points. Otherwise, if the correlation length is much smaller than the block size, then all blocks have a variance reduction of the same magnitude and the impact of the conditioning data goes unnoticed. Case B is conditioned only on the fine scale logconductivity measurements. Comparing cases A and B in Figure 12 and in Table 2 we notice that for the unconditional case, the ensemble mean of $\ln(K_{max})$ between rows is spatially homogeneous and so is the variance; however, as soon as the fine scale conductivity data are used for the generation of the fine scale realizations, the ensemble of upscaled realizations displays the effects of such conditioning, the ensemble mean starts to show patterns closer to the patterns in the upscaled reference field (Figure 7), and the ensemble variance becomes smaller for the interblocks closer to the conditioning measurements. Analyzing case D in Figure 12, which takes the ensemble of realizations from case B and updates it by assimilating the piezometric head measurements at piezometers W1 to W9, we conclude that the initial conditioning effect (to hydraulic conductivity data) is reinforced by the new conditioning data, the patterns observed in the ensemble mean maps are even closer to the patterns in the reference realization, and the ensemble variance remains small close to logconductivity conditioning locations and, overall, is smaller than for case B.

Finally, when no conductivity data are used to condition the initial ensemble of realizations, conditioning to piezometric heads through EnKF also serves to improve the characterization of the logconductivities as can be seen analyzing case C in Figure 12 and Table 2. Some patterns of the spatial variability of $\ln(K_{max})$ are captured by the ensemble mean and the ensemble variance is reduced with respect to the unconditional case, although in a smaller magnitude than when logconductivity data are used for conditioning.

From this analysis we conclude that conditioning to piezometric head data by the EnKF coupled with upscaling improves the characterization of aquifer logconductivities whether conductivity data are used for conditioning or not.

It should be emphasized that, since the EnKF algorithm starts after the upscaling of the ensemble of fine scale realizations ends, the EnKF-coupled-with-upscaling performance will be much restricted by the quality of the upscaling algorithm. It is important to use as accurate an upscaling procedure as possible in the first step of the process, otherwise the EnKF algorithm may fail. An interesting discussion on the importance of the choice of upscaling can be read in the study of the MADE site by *Li et al.* [2011a].

4.2. Worth of Data

We can use the results obtained to make a quick analysis of the worth of data in aquifer characterization, which confirms earlier findings [e.g., *Capilla et al.*, 1999; *Wen et al.*, 2002; *Hendricks Franssen*, 2001; *Hendricks Franssen et al.*, 2003; *Fu and Gómez-Hernández*, 2009; *Li et al.*, 2011c] and serves to show that the proposed approach works as expected. By analyzing Figures 8, 9, 10, 11, 12, and 13, and Table 2, we can conclude that conditioning to any type of data improves the characterization of the aquifer conductivities, and improves the characterization of the state of the aquifer (i.e., the piezometric heads). The largest improvement occurs when both, hydraulic conductivity and piezometric head measurements are used. These improvements can be seen qualitatively on the ensemble mean maps, which are able to display patterns closer to those in the reference maps; on the ensemble variance maps, which display smaller values than for the unconditional case; and on the ensemble average bias maps, which also show reduced bias when compared with the unconditional case. Quantitatively, the same conclusions can be made by looking at the metrics in the Table. The reproduction of the piezometric heads also improves when conditioning to any type of data.

It is also interesting to analyze the trade-off between conductivity data and piezometric head data by comparing cases B and C. As expected, the characterization of the spatial variability of hydraulic conductivity is better when conductivity data are used for conditioning than when piezometric head are; also, as expected, the opposite occurs for the characterization of the piezometric heads.

4.3. Other Issues

We have chosen a relatively small-sized fine scale model to demonstrate the methodology, since we needed the solution at the fine scale to create the sets of conditioning data and to verify that the coarse scale models generated by the proposed approach give good approximations of the "true" response of the fine scale aquifer. We envision that the proposed approach should be used only when the implementation of the numerical model and the EnKF are impractical at the fine scale.

To our understanding, it is the first time that the EnKF is applied on an aquifer with conductivities characterized by full tensors. The approach of representing the tensors by their invariants seems to work in this context. More sophisticated EnKF implementations, such as double ensemble Kalman filter [*Houtekamer and Mitchell*, 1998], ensemble square root filter [*Whitaker and Hamill*, 2002], Kalman filter based on the Karhunen-Loeve decomposition [*Zhang et al.*, 2007], normal-score ensemble Kalman filter [*Zhou et al.*, 2011b], could have been used, which would have worked equally well or better than the standard EnKF.

The example has been demonstrated using a reference conductivity field that was generated following a multiGaussian stationary random function. Could the method be applied to other types of random functions, i.e., non-multiGaussian or non-stationary? It could, as long as each step of the approach (see Section 2) could. More precisely, for the first step, the generation of the fine scale hydraulic conductivity measurements, there are already many algorithms that can generate realizations from a wide variety of random functions, including non-multiGaussian and non-stationary; the second step is basically deterministic, we replace an assembly of heterogeneous values by an equivalent block tensor, the underlying random function used to generate the fine scale realizations has no interference on the upscaling; however, for the third step, the application of EnKF to non-multiGaussian parameter fields is more difficult, some researchers propose moving on to particle filtering [*Arulampalam et al.*, 2002], some others have worked on variants of the EnKF to handle the non-multiGaussianity [e.g., *Sun et al.*, 2009; *Zhou et al.*, 2011b; *Schöniger et al.*, 2011; *Li et al.*, 2011d]; the non-stationarity is not an issue, since the EnKF deals, by construction, with non-stationary states.

5. Conclusion

The "missing scale" issue brought out by *Tran* [1996] is still, today, much overlooked. Data, particularly conductivity data, are collected at smaller support volumes and in larger quantities than years ago, yet, when constructing a numerical model based on these data, the discrepancy between the scale at which data are collected and the scale of the numerical model is most often disregarded.

We have presented an approach to rigorously account for fine-scale conductivity measurements on coarse-scale conditional inverse modeling. The resulting model is composed of an ensemble of realizations of conductivity tensors at a scale (much) coarser than the scale at which conductivities were measured. The ensemble of final realizations is conditioned to both conductivity and piezometric head measurements. The latter conditioning is achieved by using the ensemble Kalman filter on realizations of conductivity tensors. To handle the tensor parameters, we propose to work with the invariants of the tensors, instead of their representations on a specific reference system, this approach allows the ensemble Kalman filter to perform a tensor updating which produces realizations that are conditioned to the transient piezometric head measurements.

Acknowledgments. The authors acknowledge three anonymous reviewers and Wolfgang Nowak for their comments on the previous versions of the manuscript, which helped substantially in improving the final version. The authors gratefully acknowledge the financial support by ENRESA (project 0079000029). Extra travel Grants awarded to the first and second authors by the Ministry of Education (Spain) are also acknowledged. The second author also acknowledges the financial support from China Scholarship Council.

References

Aanonsen, S. L., G. Naevdal, D. S. Oliver, A. C. Reynolds, and B. Valles (2009), Ensemble kalman filter in reservoir engineering - a review, *SPE Journal*, *14*(3), 393–412.

Arulampalam, M. S., S. Maskell, N. Gordon, and T. Clapp (2002), A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking, *IEEE Transactions on signal processing*, *50*(2), 174–188.

Bear, J. (1972), *Dynamics of fluids in porous media.*, American Elsevier Pub. Co., New York.

Behrens, R., M. MacLeod, T. Tran, and A. Alimi (1998), Incorporating seismic attribute maps in 3D reservoir models, *SPE Reservoir Evaluation & Engineering*, *1*(2), 122–126.

Burgers, G., P. van Leeuwen, and G. Evensen (1998), Analysis scheme in the ensemble Kalman filter, *Monthly Weather Review*, *126*, 1719–1724.

Capilla, J., and C. Llopis-Albert (2009), Gradual conditioning of non-Gaussian transmissivity fields to flow and mass transport data: 1. Theory, *Journal of Hydrology*, *371*(1-4), 66–74.

Capilla, J. E., J. Rodrigo, and J. J. Gómez-Hernández (1999), Simulation of non-gaussian transmissivity fields honoring piezometric data and integrating soft and secondary information, *Math. Geology*, *31*(7), 907–927.

Carrera, J., A. Alcolea, A. Medina, J. Hidalgo, and L. Slooten (2005), Inverse problem in hydrogeology, *Hydrogeology Journal*, *13*(1), 206–222.

Chen, Y., and D. Zhang (2006), Data assimilation for transient flow in geologic formations via ensemble Kalman filter, *Advances in Water Resources*, *29*(8), 1107–1122.

Deutsch, C. V., and A. G. Journel (1998), *GSLIB, Geostatistical Software Library and User's Guide*, second ed., Oxford University Press, New York.

Durlofsky, L. J., R. C. Jones, and W. J. Milliken (1997), A nonuniform coarsening approach for the scale-up of displacement processes in heterogeneous porous media, *Adv.*

in Water Resour., *20*(5-6), 335–347.

Evensen, G. (2003), The ensemble Kalman filter: Theoretical formulation and practical implementation, *Ocean dynamics*, *53*(4), 343–367.

Freeze, R. A., and J. A. Cherry (1979), *Groundwater*, Prentice-Hall.

Fu, J., and J. Gómez-Hernández (2009), Uncertainty assessment and data worth in groundwater flow and mass transport modeling using a blocking Markov chain Monte Carlo method, *Journal of Hydrology*, *364*(3-4), 328–341.

Fu, J., H. A. Tchelepi, and J. Caers (2010), A multiscale adjoint method to compute sensitivity coefficients for flow in heterogeneous porous media, *Advances in Water Resources*, *33*(6), 698–709, doi:10.1016/j.advwatres.2010.04.005.

Gómez-Hernández, J. J. (1991), A stochastic approach to the simulation of block conductivity values conditioned upon data measured at a smaller scale, Ph.D. thesis, Stanford University.

Gómez-Hernández, J. J., and A. G. Journel (1993), Joint sequential simulation of multi-Gaussian fields, *Geostatistics Troia*, *92*(1), 85–94.

Gómez-Hernández, J. J., and R. M. Srivastava (1990), ISIM3D: an ANSI-C three dimensional multiple indicator conditional simulation program, *Computers & Geosciences*, *16*(4), 395–440.

Gómez-Hernández, J. J., A. Sahuquillo, and J. E. Capilla (1997), Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data, 1, Theory, *Journal of Hydrology*, *203*(1–4), 162–174.

Guadagnini, A., and S. P. Neuman (1999), Nonlocal and localized analyses of conditional mean steady state flow in bounded, randomly nonuniform domains: 1. theory and computational approach, *Water Resour. Res.*, *35*(10), 2999–3018.

Hendricks Franssen, H. (2001), Inverse stochastic modelling of groundwater flow and mass transport, Ph.D. thesis, Technical University of Valencia.

Hendricks Franssen, H., and W. Kinzelbach (2009), Ensemble Kalman filtering versus sequential self-calibration for inverse modelling of dynamic groundwater flow systems, *Journal of Hydrology*, *365*(3-4), 261–274.

Hendricks Franssen, H., A. Alcolea, M. Riva, M. Bakr, N. van der Wiel, F. Stauffer, and A. Guadagnini (2009), A comparison of seven methods for the inverse modelling of groundwater flow. application to the characterisation of well catchments, *Advances in Water Resources*, *32*(6), 851–872, doi:10.1016/j.advwatres.2009.02.011.

Hendricks Franssen, H. J., J. J. Gómez-Hernández, and A. Sahuquillo (2003), Coupled inverse modelling of groundwater flow and mass transport and the worth of concentration data, *Journal of Hydrology*, *281*(4), 281–295.

Houtekamer, P., and H. Mitchell (1998), Data assimilation using an ensemble kalman filter technique, *Monthly Weather Review*, *126*(3), 796–811.

Hu, L. Y. (2000), Gradual deformation and iterative calibration of gaussian-related stochastic models, *Math. Geology*, *32*(1), 87–108.

Indelman, P., and B. Abramovich (1994), Nonlocal properties of nonuniform averaged flows in heterogeneous media, *Water Resour. Res.*, *30*(12), 3385–3393.

Li, L., H. Zhou, and J. J. Gómez-Hernández (2010), Steady-state groundwater flow modeling with full tensor conductivities using finite differences, *Computers & Geosciences*, *36*(10), 1211–1223.

Li, L., H. Zhou, and J. J. Gómez-Hernández (2011a), A comparative study of three-dimensional hydrualic conductivity upscaling at the macrodispersion experiment (MADE) site, columbus air force base, mississippi (USA), *Journal of Hydrology*, *404*(3-4), 135–142.

Li, L., H. Zhou, and J. J. Gómez-Hernández (2011b), Transport upscaling using multi-rate mass transfer in three-dimensional highly heterogeneous porous media, *Advances in Water Resources*, *34*(4), 478–489.

Li, L., H. Zhou, J. J. Gómez-Hernández, and H. J. Hendricks Franssen (2011c), Jointly mapping hydraulic conductivity and porosity by assimilating concentration data via ensemble kalman filter, *Journal of Hydrology*, p. submitted.

Li, L., H. Zhou, H. J. Hendricks Franssen, and J. J. Gómez-Hernández (2011d), Groundwater flow inverse modeling in non-multigaussian media: performance assessment of the normal-score ensemble kalman filter, *Hydrology and Earth System Sciences Discussions*, *8*(4), 6749–6788, doi:10.5194/hessd-8-6749-2011.

Mariethoz, G., P. Renard, and J. Straubhaar (2010), The direct sampling method to perform multiple-point geostatistical simulaitons, *Water Resources Research*, *46*, W11,536, doi:10.1029/2008WR007621.

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

Naevdal, G., L. Johnsen, S. Aanonsen, and E. Vefring (2005), Reservoir monitoring and continuous model updating using ensemble kalman filter, *SPE Journal*, *10*(1).

Nowak, W. (2010), Measures of parameter uncertainty in geostatistical estimation and geostatistical optimal design, *Math. Geology*, *49*, 199–221.

Oliver, D., and Y. Chen (2011), Recent progress on reservoir history matching: a review, *Computational Geosciences*, *15*(1), 185–221.

Peters, L., R. J. Arts, G. K. Brower, C. R. Geel, S. Cullick, R. J. Lorentzen, Y. Chen, N. Dunlop, F. Vossepoel, R. Xu, P. Sarman, A. A. H. H., and A. Reynolds (2010), Results of the brugge benchmark study for flooding optimization and history matching, *SPE Reservoir Evaluation & Engineering*, *13*(3), 391–405, doi:10.2118/119094-PA.

Renard, P., and G. de Marsily (1997), Calculating equivalent permeability: A review, *Advances in Water Resources*, *20*(5-6), 253–278.

Rubin, Y., and J. J. Gómez-Hernández (1990), A stochastic approach to the problem of upscaling of conductivity in disordered media, theory and unconditional numerical simulations, *Water Resour. Res.*, *26*(4), 691–701.

Sánchez-Vila, X., J. Girardi, and J. Carrera (1995), A synthesis of approaches to upscaling of hydraulic conductivities, *Water Resources Research*, *31*(4), 867–882.

Sánchez-Vila, X., A. Guadagnini, and J. Carrera (2006), Representative hydraulic conductivities in saturated groundwater flow, *Reviews of Geophysics*, *44*(3), doi:10.1029/2005RG000169.

Schöniger, A., W. Nowak, and H. J. Hendricks Franssen (2011), Parameter estimation by ensemble Kalman filters with transformed data: approach and application to hydraulic tomography, p. submitted to Water Resources Research.

Strebelle, S. (2002), Conditional simulation of complex geological structures using multiple-point statistics, *Mathematical Geology*, *34*(1), 1–21.

Sun, A. Y., A. P. Morris, and S. Mohanty (2009), Sequential updating of multimodal hydrogeologic parameter fields using localization and clustering techniques, *Water Re-*

sources Research, *45*, 15 PP.

Tran, T. (1996), The 'missing scale' and direct simulation of block effective properties, *Journal of Hydrology*, *183*(1-2), 37–56, doi:10.1016/S0022-1694(96)80033-3.

Tran, T., X. Wen, and R. Behrens (1999), Efficient conditioning of 3D fine-scale reservoir model to multiphase production data using streamline-based coarse-scale inversion and geostatistical downscaling, in *SPE Annual Technical Conference and Exhibition*.

Tureyen, O., and J. Caers (2005), A parallel, multiscale approach to reservoir modeling, *Computational Geosciences*, *9*(2), 75–98.

Wen, X. H., and J. J. Gómez-Hernández (1996), Upscaling hydraulic conductivities: An overview, *J. of Hydrology*, *183*(1-2), ix–xxxii.

Wen, X. H., C. Deutsch, and A. Cullick (2002), Construction of geostatistical aquifer models integrating dynamic flow and tracer data using inverse technique, *Journal of Hydrology*, *255*(1-4), 151–168.

Whitaker, J., and T. Hamill (2002), Ensemble data assimilation without perturbed observations, *Monthly Weather Review*, *130*(7), 1913–1925.

White, C. D., and R. N. Horne (1987), Computing absolute transmissibility in the presence of Fine-Scale heterogeneity, *SPE 16011*.

Yeh, W. (1986), Review of parameter identification procedures in groundwater hydrology: The inverse problem, *Water Resources Research*, *22*(2), 95–108.

Zhang, D., Z. Lu, and Y. Chen (2007), Dynamic reservoir data assimilation with an efficient, dimension-reduced kalman filter, *SPE Journal*, *12*(1), 108–117.

Zhou, H., L. Li, and J. J. Gómez-Hernández (2010), Three-dimensional hydraulic conductivity upscaling in groundwater modelling, *Computers & Geosciences*, *36*(10), 1224–

1235.

Zhou, H., J. J. Gómez-Hernández, and L. Li (2011a), Inverse methods in hydrogeology: evolution and future trends, *Journal of Hydrology*, p. Submitted.

Zhou, H., J. J. Gómez-Hernández, H.-J. Hendricks Franssen, and L. Li (2011b), An approach to handling Non-Gaussianity of parameters and state variables in ensemble Kalman filter, *Advances in Water Resources*, *34* (7), 844–864.

Zimmerman, D., G. De Marsily, C. Gotway, M. Marietta, C. Axness, R. Beauheim, R. Bras, J. Carrera, G. Dagan, P. Davies, et al. (1998), A comparison of seven geostatistically based inverse approaches to estimate transmissivities for modeling advective transport by groundwater flow, *Water Resources Research*, *34* (6), 1373–1413.

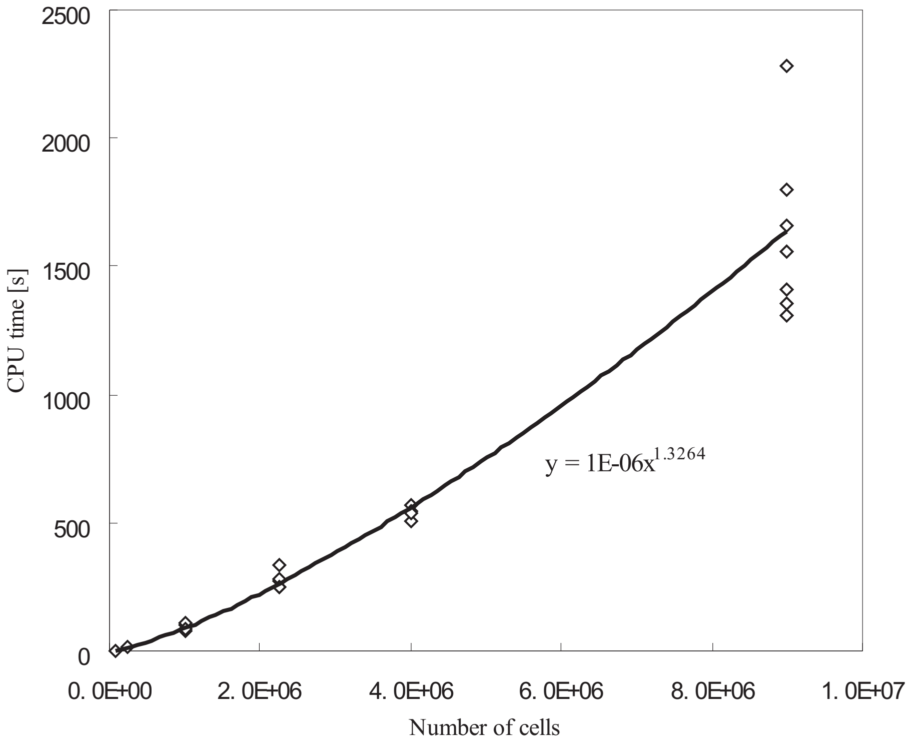

Figure 1. Regression of CPU time versus number of cells from several runs of MODFLOW on heterogeneous realizations.

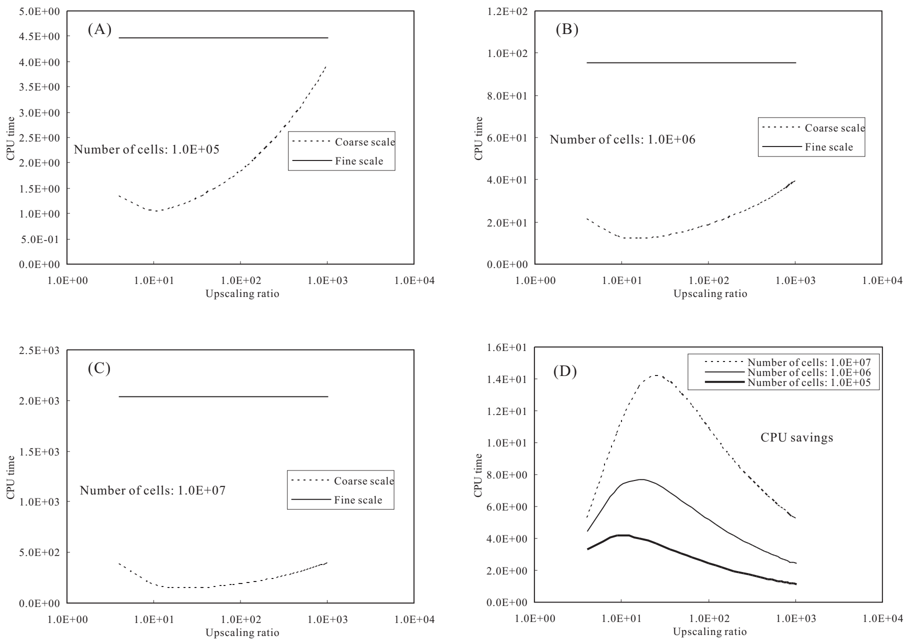

Figure 2. CPU time as a function of upscaling ratio for a single time step modeling. Plots A, B and C show, for different fine scale model sizes, the CPU time needed to run one single time step in a fine scale model and the CPU times needed for the upscaling plus running the flow model for different sizes of the upscaling block. Plot D shows the ratio of CPU time between the fine scale and the coarse scale, the larger the ratio, the larger the savings.

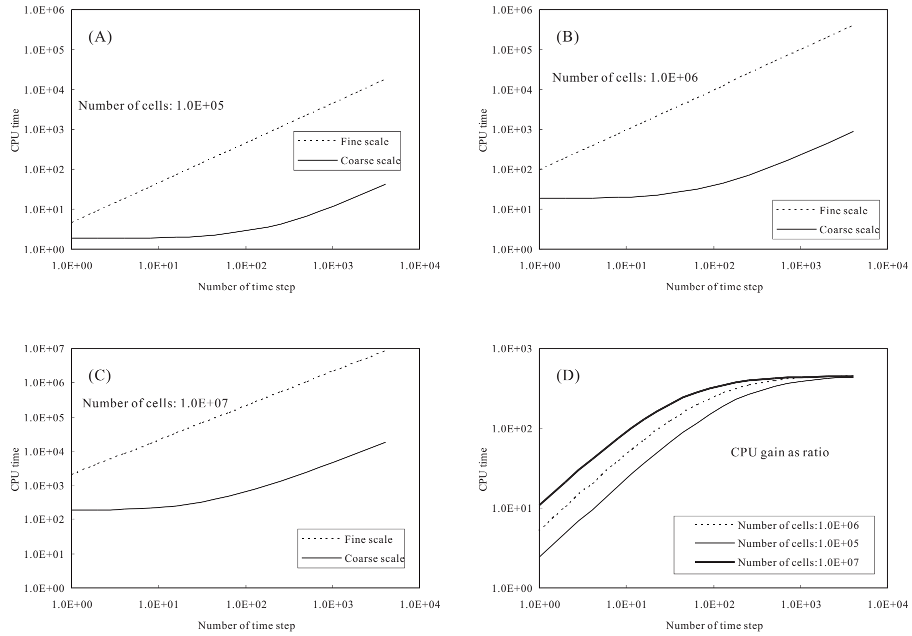

Figure 3. CPU time as a function of the number of time steps modeled. Plots A, B and C show, for a fine scale model of 10^5 nodes, and for an upscaled model of 10^3 blocks (upscaling ratio of 100), the CPU time needed to run the fine and coarse scale models as a function of the number of time steps. Plot D shows the ratio of CPU time between the fine scale and the coarse scale, the larger the ratio, the larger the savings.

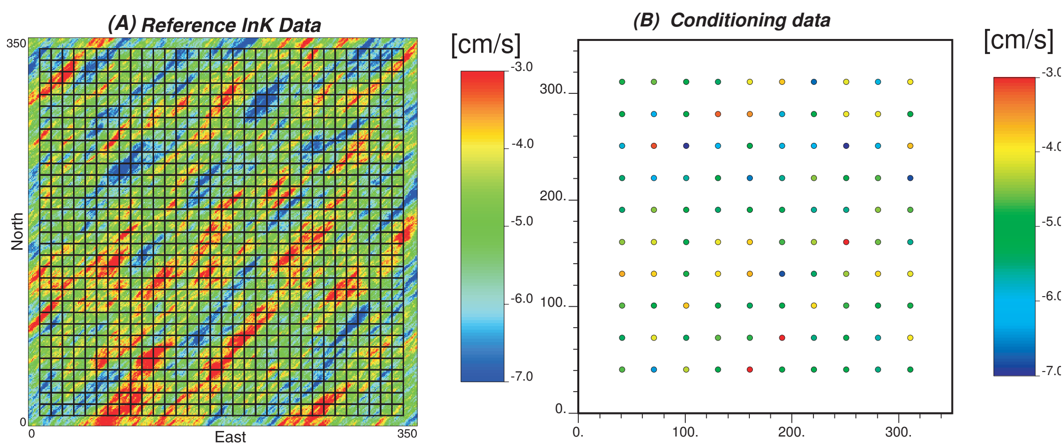

Figure 4. (A) Reference ln$K$ field overlaid with the discretization of the numerical model at the coarse scale. (B) Conditioning ln$K$ data.

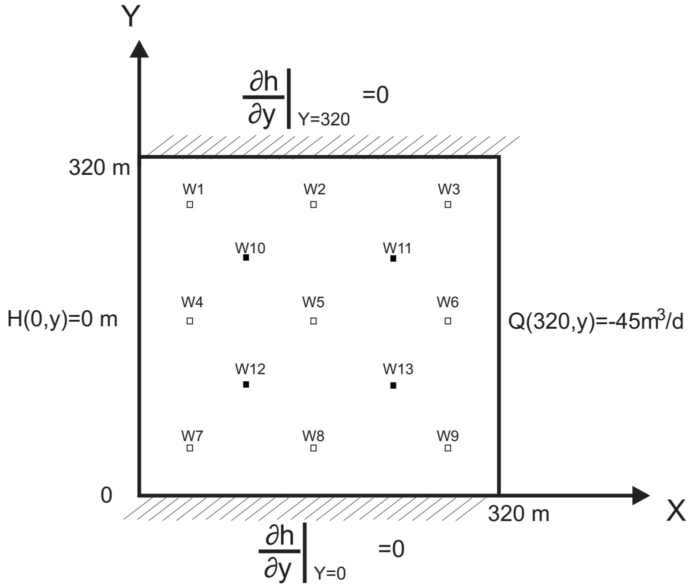

Figure 5. Sketch of the flow problem with boundary conditions, observation and prediction wells. Empty squares correspond to the piezometric head observation wells (W1-W9); filled squares correspond to the control wells (W10-W13).

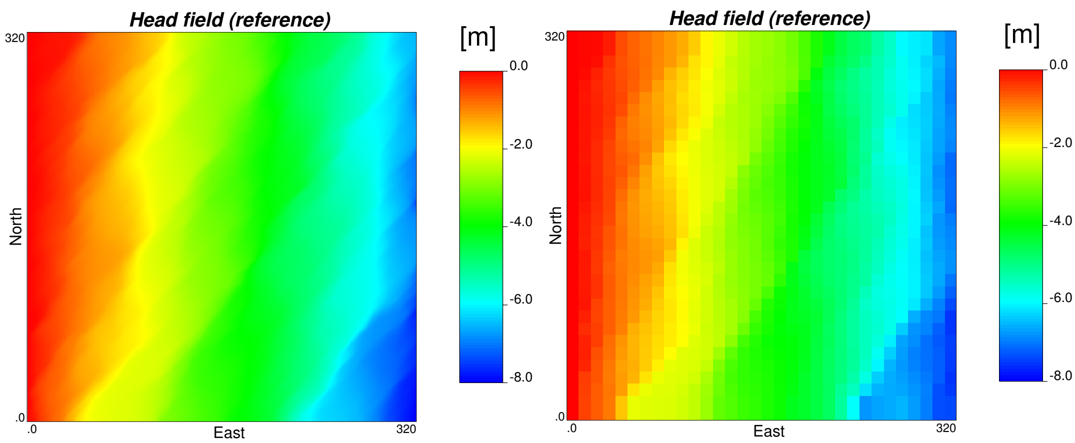

Figure 6. Reference piezometric head at the 60th time step. Left, as obtained at the fine scale; right, as obtained at the coarse scale

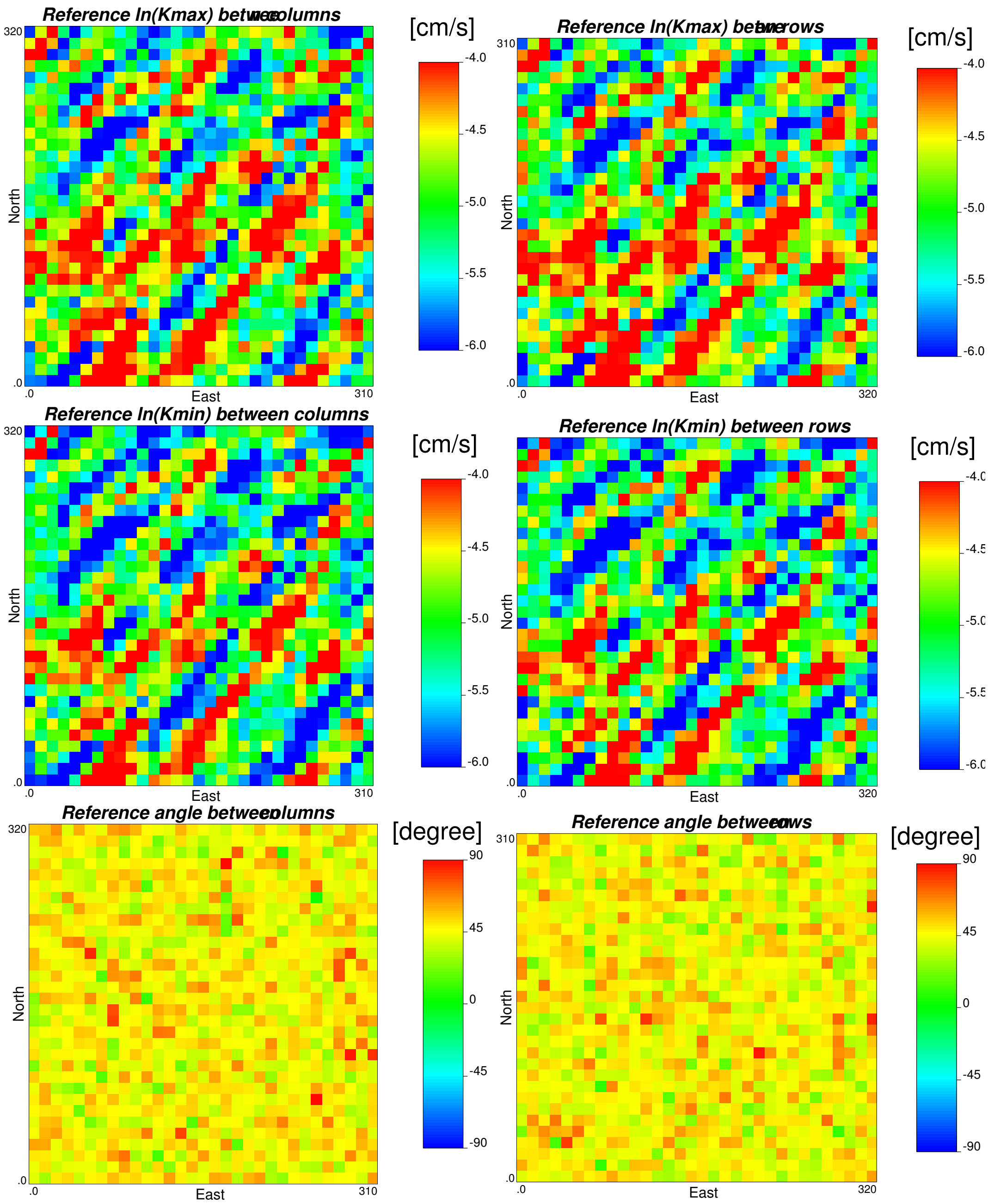

Figure 7. Upscaled values for the interblock tensor components: $\ln(Kmax)$, $\ln(Kmin)$ and rotation angle for the maximum component measured from the x-axis θ (in degrees), for both the interblocks between columns and the interblocks between rows. Upscaling method used: Laplacian with a skin of 10 m

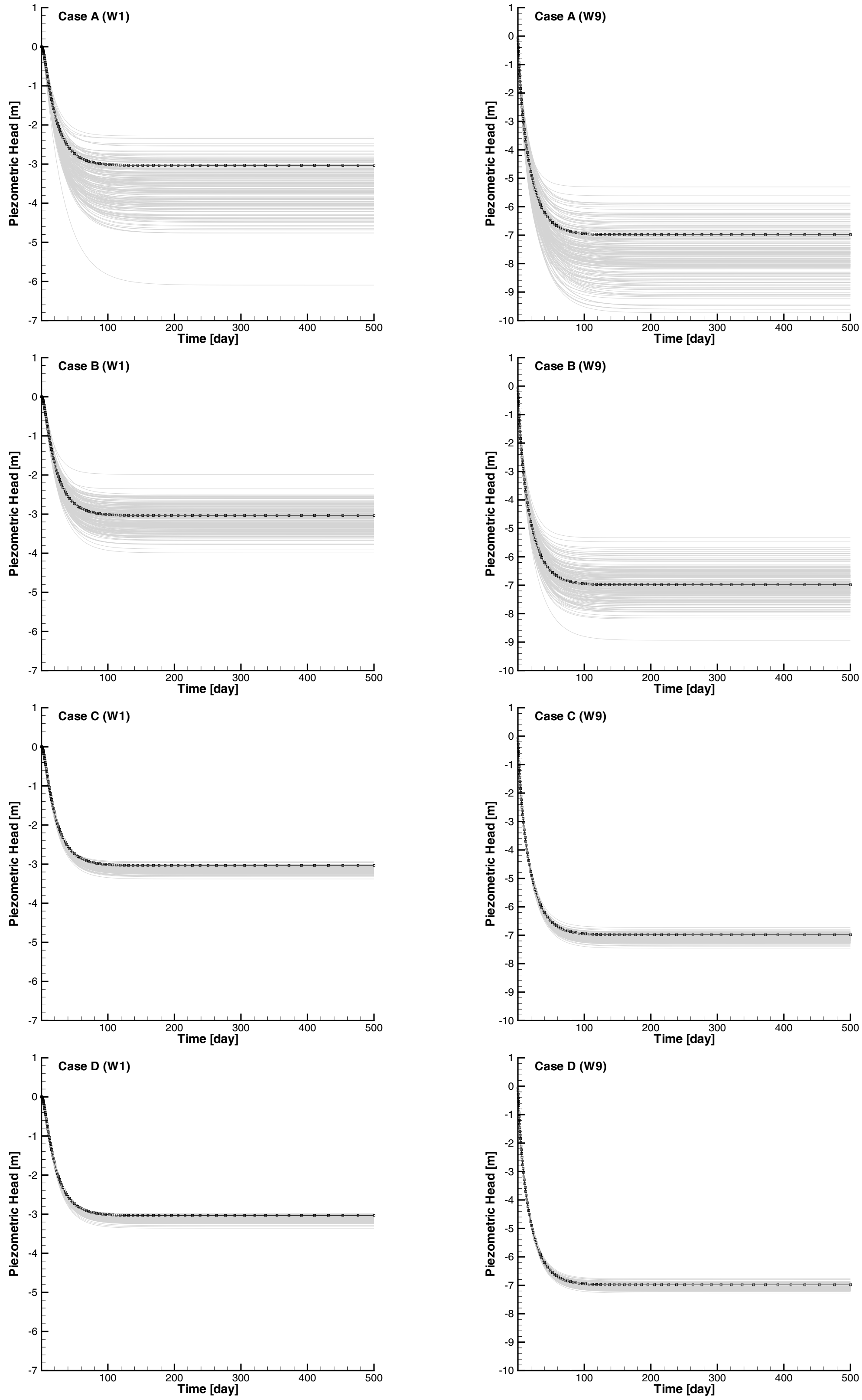

Figure 8. Piezometric head time series in the reference field and simulated ones for all cases at wells W1 (left column) and W9 (right column). The piezometric heads measured at these wells during the first 67.7 days were used as conditioning data for cases B and D.

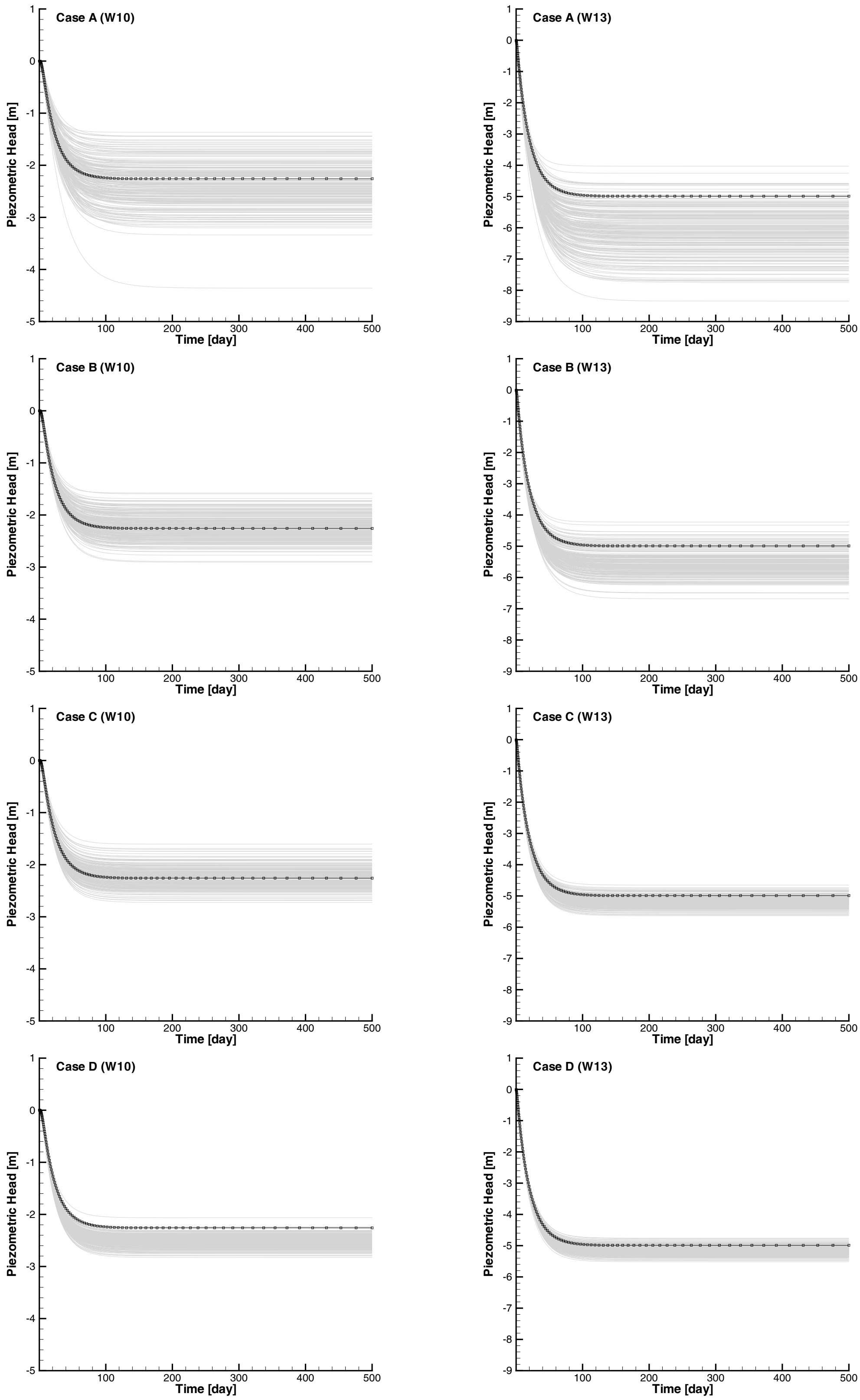

Figure 9. Piezometric head time series in the reference field and simulated ones for all cases at control wells W10 (left column) and W13 (right column). These wells were not used as conditioning data for any case.

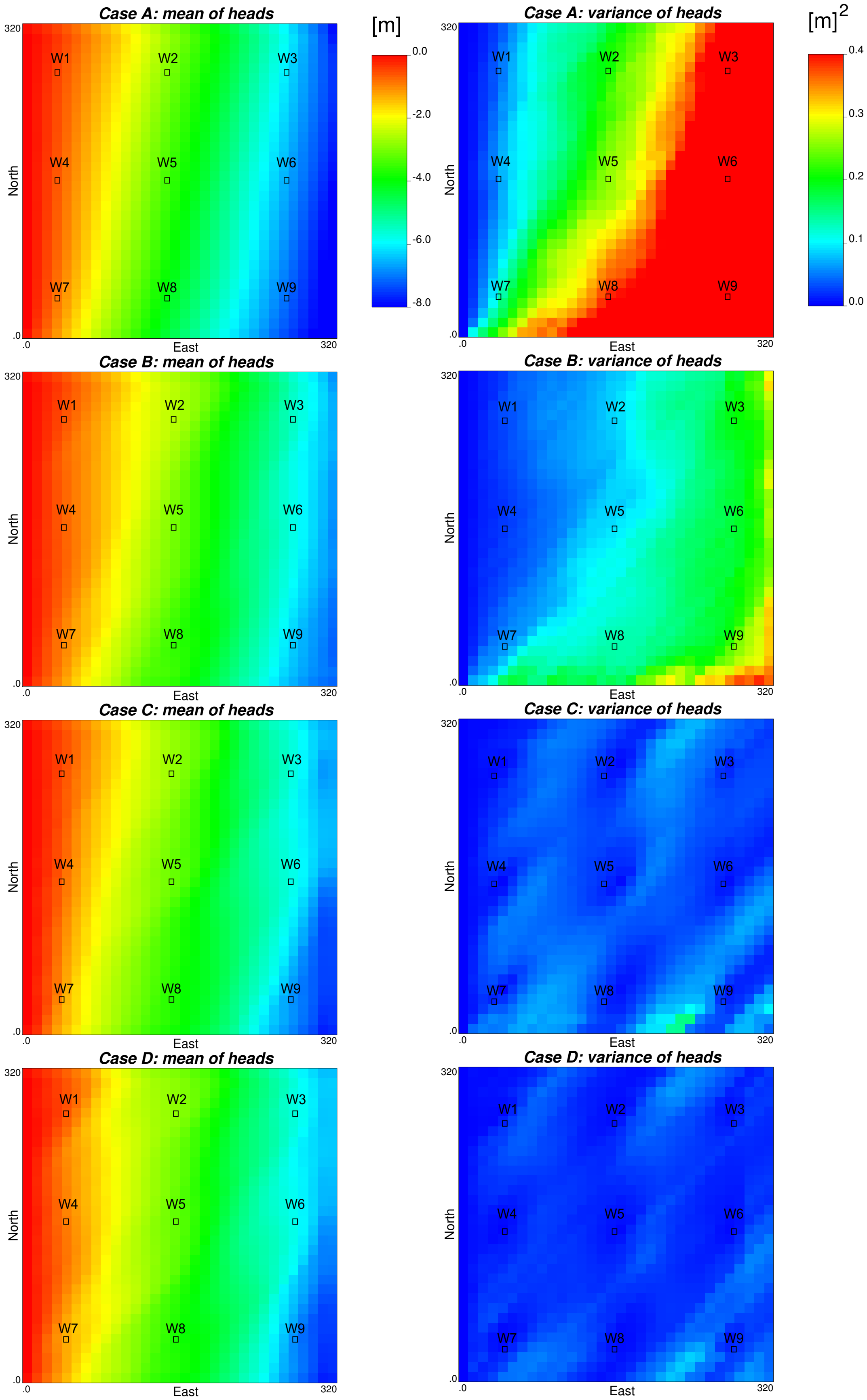

Figure 10. Ensemble average and variance of piezometric heads for the different cases.

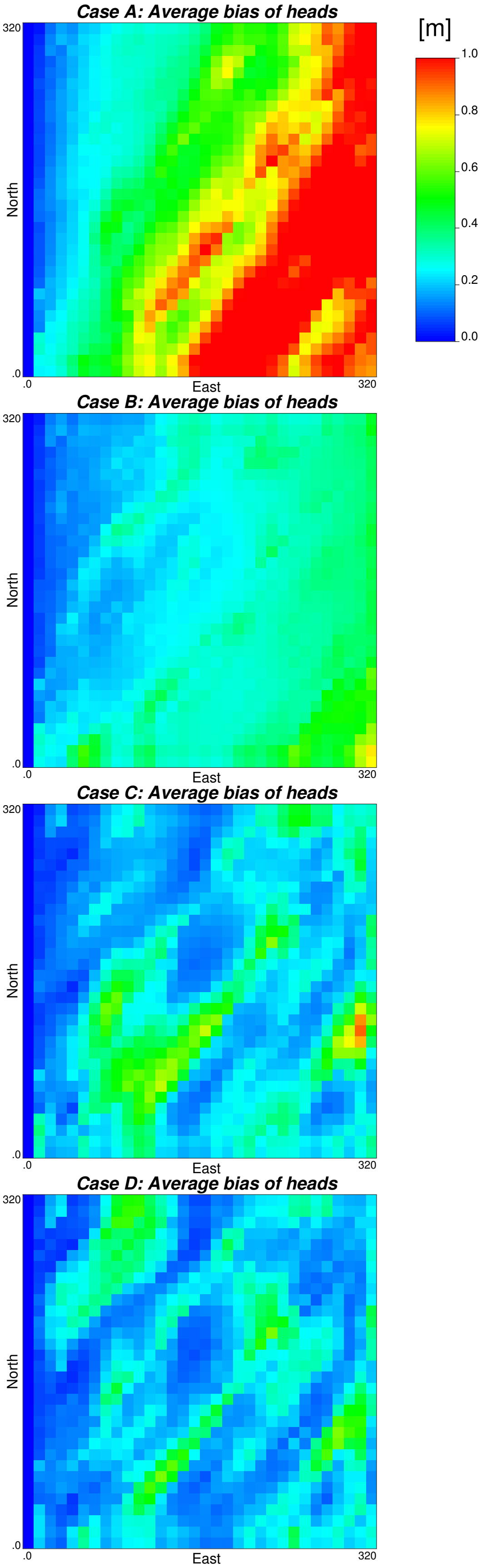

Figure 11. Ensemble average absolute bias of piezometric heads for the different cases.

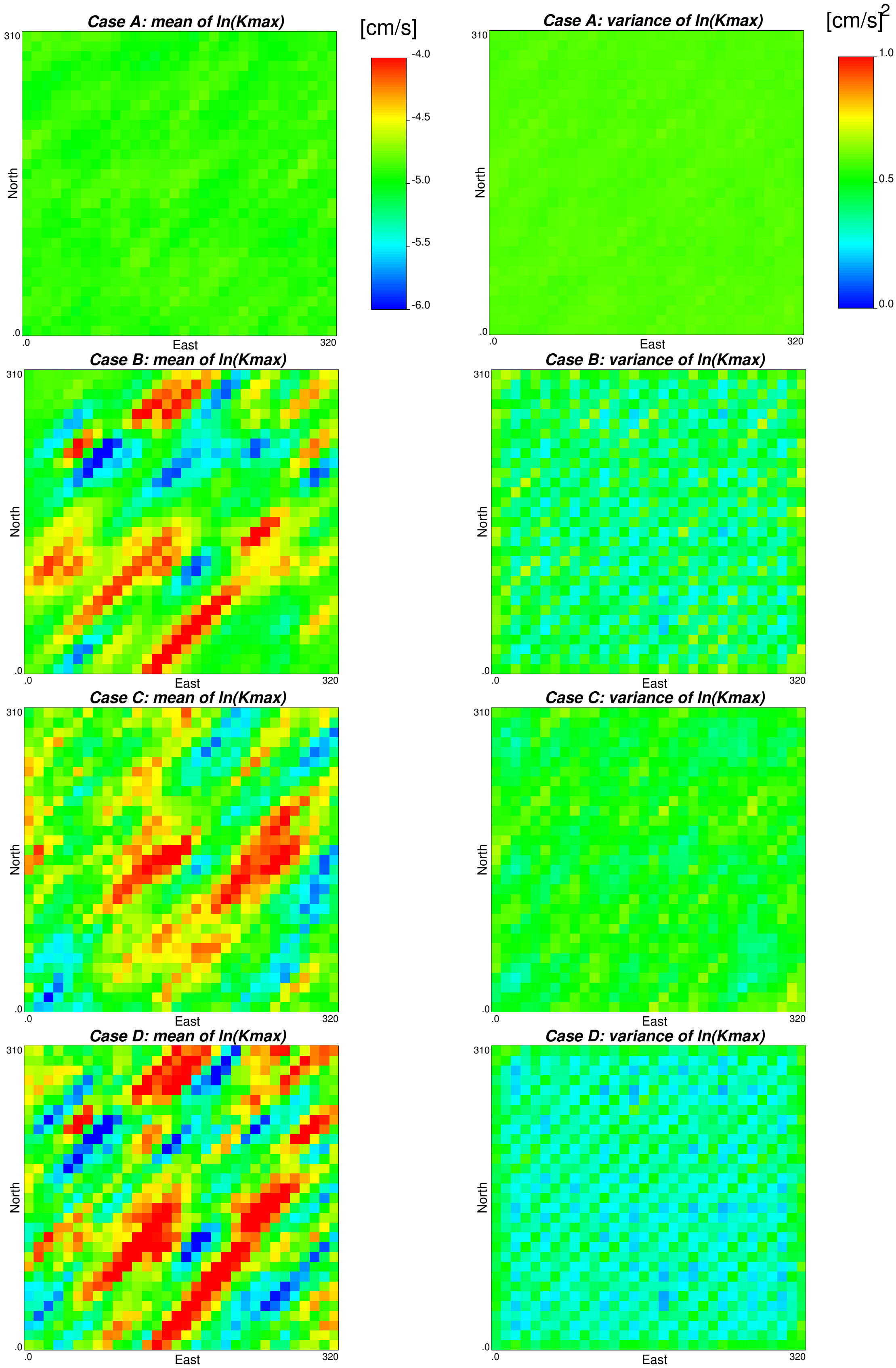

Figure 12. Ensemble average and variance of $\ln(Kmax)$ for the different cases.

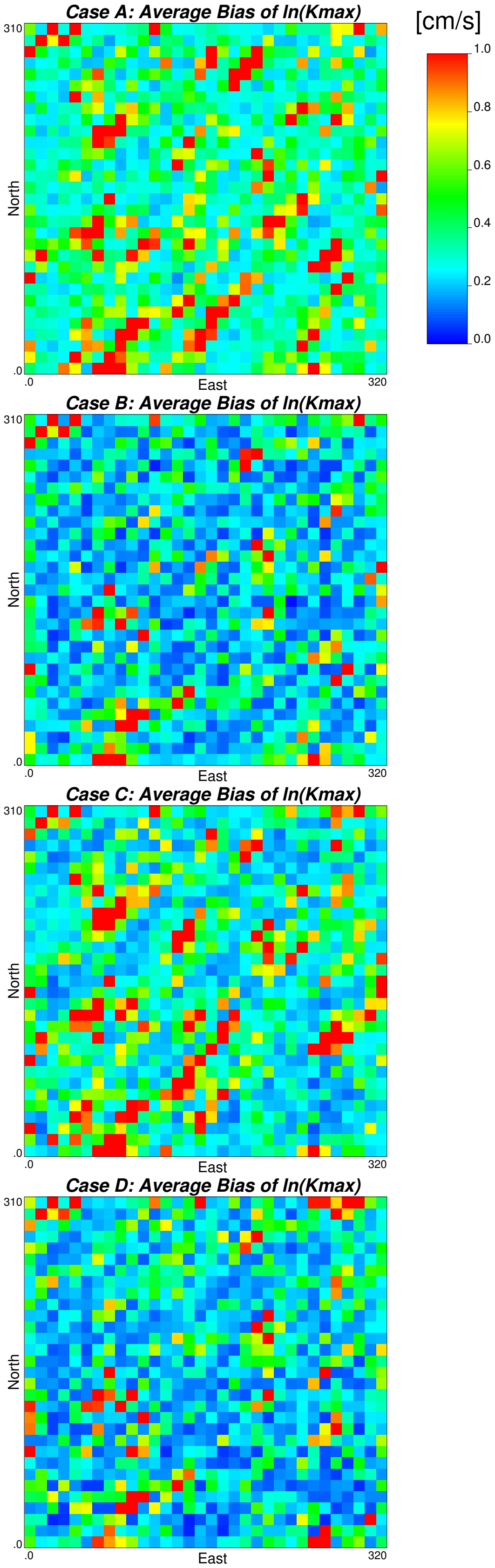

Figure 13. Ensemble average absolute bias of $\ln(Kmax)$ for the different cases.

Table 1. Definition of Cases depending on the the different sets of conditioning data.

Conditioning Data	Case A	Case B	Case C	Case D
Hydraulic conductivities(K)	No	Yes	No	Yes
Dynamic piezometric heads(h)	No	No	Yes	Yes

Table 2. Bias and spread of predicted heads at time steps 30, 60 and 90 and of updated loghydraulic conductivity $\ln K_{max}^{b,r}$ at time step 60.

	Case A	Case B	Case C	Case D
$AAB(h_{nt=30})$	0.189	0.119	0.124	0.118
$AESP(h_{nt=30})$	0.201	0.132	0.111	0.086
$AAB(h_{nt=60})$	0.580	0.256	0.224	0.195
$AESP(h_{nt=60})$	0.533	0.323	0.186	0.146
$AAB(h_{nt=90})$	0.672	0.281	0.236	0.204
$AESP(h_{nt=90})$	0.627	0.355	0.195	0.153
$AAB(\ln K_{max})$	0.452	0.306	0.417	0.296
$AESP(\ln K_{max})$	0.805	0.660	0.702	0.594