WATER RESOURCES RESEARCH, VOL. ???, XXXX, DOI:10.1002/,

Inverse sequential simulation: A new approach for the characterization of hydraulic conductivities demonstrated on a non-Gaussian field

Teng Xu,[1] and J. Jaime Gómez-Hernández[1]

Corresponding author: Teng Xu, Group of Hydrogeology, Universitat Politècnica de València, Camino de Vera, s/n, 46022 Valencia, Spain. (tenxu@posgrado.upv.es)

[1]Group of Hydrogeology, Research Institute for Water and Environmental Engineering, Universitat Politècnica de València, Valencia, Spain.

Abstract.

Inverse sequential simulation (iSS) is a new inverse modeling approach for the characterization of hydraulic conductivity fields based on sequential simulation. It is described and demonstrated in a synthetic aquifer with non-Gaussian spatial features, and compared against the normal-score ensemble Kalman filter (NS-EnKF). The new approach uses the sequential simulation paradigm to generate realizations borrowing from the ensemble Kalman filter the idea of using the experimental non-stationary cross-covariance between conductivities and piezometric heads computed on an ensemble of realizations. The resulting approach is fully capable of retrieving the non-Gaussian patterns of the reference field after conditioning on the piezometric heads with results comparable of those obtained by the NS-EnKF.

1. Introduction

The quality of a groundwater model, particularly when studying the fate and transport of contaminants, relies very much on the quality of the characterization of hydraulic conductivities. Many studies have shown that, unless the heterogeneity of hydraulic conductivity is well captured in the groundwater model, the resulting transport predictions could be totally wrong [*Sudicky*, 1986; *Gómez-Hernández and Wen*, 1994; *Eggleston and Rojstaczer*, 1998; *Li et al.*, 2012]; for instance, *Gómez-Hernández and Wen* [1994] show the high impact that not properly accounting for heterogeneity has in transport predictions. But not only it is important to account for heterogeneity, as important is using the most adequate heterogeneity model. For many years, the only model considered for the spatial variability of hydraulic conductivity was the mutiGaussian model of log-conductivity, until it was recognized that the spatial patterns often observed in the subsurface (i.e., channels, permeability barriers, high conductivity streaks) were better modeled using alternatives to the multiGaussian model [*Wen and Gómez-Hernández*, 1998; *Fu and Gómez-Hernández*, 2009; *Gómez-Hernández and Wen*, 1998], and, in addition, many natural heterogeneity patterns are simply unsuitable for a multiGaussian modeling. Since then, there have been many attempts to define non-multiGaussian random functions capable of capturing the spatial features difficult to capture by the multiGaussian ones, and then, to build algorithms for the spatial representation of hydraulic conductivity according to these new random function models [*Carle and Fogg*, 1996; *Strebelle*, 2002; *Mariethoz et al.*, 2010; *Haslauer et al.*, 2012].

Probably the most successful approach for the generation of realistic hydraulic conductivities is the one based on training images and multiple-point statistics [*Guardiano and Srivastava*, 1993; *Strébelle*, 2000; *Strebelle*, 2002]. The next challenge was how to use these random functions in inverse modeling, that is, how to generate realizations of hydraulic conductivity that not only are consistent with the training image and conditional to the local measurements, but also that are inverse conditioned onto observed measurements of the state variables, such as piezometric head or solute concentration. Inverse modeling in hydrogeology and petroleum engineering has a long tradition (see *Zhou et al.* [2014] for a review) but, again, most inverse models rely on the assumption that hydraulic conductivity follows a multiGaussian model. Recent attempts to couple inverse approaches and non multiGaussian random functions have been attempted by *Sun et al.* [2009]; *Sarma and Chen* [2009]; *Li et al.* [2009]; *Alcolea and Renard* [2010]; *Jafarpour and Khodabakhshi* [2011]; *Hu et al.* [2013]; *Zhou et al.* [2011, 2012a]; *Attia and Sandu* [2014], among others, with different degrees of success. Most of these approaches are extensions of algorithms that work for multiGaussian fields.

In this paper we propose a completely new algorithm that is the result of blending some of the ideas underlying multivariate Gaussian sequential simulation [*Gómez-Hernández and Journel*, 1993] and ensemble Kalman filtering [*Evensen*, 2003]. We have called this new algorithm inverse sequential simulation (iSS). The iSS algorithm aims at the characterization of hydraulic conductivity from observations of hydraulic conductivity and piezometric heads, and has been built to work for non-multiGaussian fields. As a benchmark, the algorithm will be compared with the normal-score ensemble Kalman filter (NS-EnKF) [*Zhou et al.*, 2011], which is one of the algorithms that best performs for non-multiGaussian inverse modeling. The paper continues with a description of the algorithm, followed by the comparison between NS-EnKF and iSS on a synthetic example, and ends with some discussion and conclusions.

2. Methodology

The new algorithm is a breed of sequential simulation and the normal-score ensemble Kalman filter. First, we borrow, from the ensemble Kalman filter (EnKF) [*van Loon et al.*, 2000; *Evensen*, 2003; *Blöschl et al.*, 2008; *Karri et al.*, 2014], the idea of using an ensemble of realizations to compute an experimental, non-stationary conditional cross-covariance between conductivities and piezometric heads, and also experimental non-stationary conditional auto-covariances of both conductivity and piezometric head. Second, we borrow, from the normal-score ensemble Kalman filter [*Zhou et al.*, 2011], the idea of performing a normal-score transformation and thus work with a marginally Gaussian multivariate random function. We are fully aware that a normal-score transformation only produces marginally-distributed Gaussian variables, never multiGaussian ones; however, it has been shown in the NS-EnKF that this transformation is quite effective in capturing non-Gaussian patterns [*Zhou et al.*, 2011, 2012b; *Li et al.*, 2011; *Xu et al.*, 2013]. We do not claim that higher-order moments, after the normal-score transform, will correspond to those of a multiGaussian distribution; what we claim, based on our experience working with the NS-EnKF, is that, when applying a method that is optimal for multiGaussian variables to non-multiGussian ones, the results are better if a normal-score transform is applied than if not. And third, we use standard multivariate sequential Gaussian conditional simulation [*Gómez-Hernández and Journel*, 1993; *Friedel and Iwashita*, 2013] to generate realizations of the normal scores of conductivity conditioned to the normal scores of conductivity and to the piezometric head measurements. The state equation relating conductivity and piezometric heads, with its initial conditions, boundary conditions and forcing terms, is indirectly included in the sequential simulation through the experimental conditional auto- and cross-covariances that are computed on the ensemble of realizations. When and if new head measurements are taken, the generated ensemble of conductivity realizations are used to forecast an ensemble of head realizations, new experimental covariances are computed, and a new ensemble of conductivity realizations is generated conditioned to the new head measurements.

The iSS method has been developed for its application under transient conditions, with a regeneration of the ensemble of conductivity fields each time new piezometric heads are measured. Consider that piezometric heads are collected sequentially in time. The method starts with an ensemble of conductivity fields generated according to a given random function model —for the generation of this initial set no information about piezometric heads is used, this initial set should be conditional to conductivity measurements and other soft information such as geophysical data, when available. Then, for each time step for which piezometric heads are observed, the algorithm carries out the following: (i) an ensemble of piezometric head realizations are predicted on the basis of the last ensemble of conductivity fields by means of a numerical flow model, (ii) the conductivity and head auto- and cross-covariances are computed from the ensemble of realizations — these covariances will be non-stationary, (iii) using a sequential multivariate simulation algorithm, a new ensemble of conductivity fields conditioned to the conductivity data, if any, and to the measured piezometric heads is generated. A flowchart of the iSS is included in Figure 1 and its implementation details are explained next:

Consider that there are N_e realizations in the ensemble, and that each realization is discretized into N nodes.

1. Initialization step. We need to start from an ensemble of conductivity fields. This ensemble should be generated with the algorithm that is most adequate for the type of heterogeneity that describes the conductivity spatial variability. This ensemble can be made conditional to hard measurements of conductivity, and also to soft measurements such as those derived from geophysics. Measurement errors are easily accounted for by simply adding a random error (drawn from the measurement error distribution function) to each hard measurement prior to using it for the generation of each realization of the ensemble. For the purpose of illustration, we choose a formation with channel-like features that introduce a high connectivity of the facies in the direction of flow. We select this type of heterogeneity because it is well known that it is difficult to capture by multiGaussian-based approaches. In our synthetic case, the initial ensemble of realizations is generated in two steps. In the first step, an ensemble of binary facies realizations is constructed using single normal equation simulation *Strébelle* [2000] —a very efficient implementation of sequential normal-equation simulation, first developed by *Guardiano and Srivastava* [1993] and improved by *Strébelle* [2000]. Then, each facies (channel/sand and non-channel/shale) is independently populated with log-conductivity values using sequential Gaussian simulation; the conductivities of each facies have very distinct mean values, ensuring that each realization has a clearly bimodal distribution, with the highest-value mode in the channel elements and the lowest-value mode in the non-channel elements. The specific parameters used for the generation of the initial ensemble of log-conductivities are described in the next section. At the end of this step there is an ensemble of hydraulic conductivity fields

that will be denoted by $\mathbf{K}^0$, with $K^0_{i,j}$ being the conductivity for realization i at node j. The superindex is used for the time coordinate and zero indicates that these are the initial conductivity estimates. For notation purposes, we will use $K_{i,\cdot}$ to denote realization i of the ensemble, and $K_{\cdot,j}$ to denote the set of N_e conductivity values collected from all realizations at node j. During the initialization step it is also necessary to specify the initial piezometric heads $\mathbf{H}^0$, as well as boundary conditions and forcing terms necessary to solve the transient groundwater flow equation.

2. Forecasting step. In this step, the simulated piezometric heads ($\mathbf{H}^{t+1}$) are calculated for the $(t+1)^{\text{th}}$ time step based on the piezometric heads and the hydraulic conductivity estimates from the t^{th} time step using a transient flow model:

$$H^{t+1}_{i,\cdot} = \psi(H^t_{i,\cdot}, K^t_{i,\cdot}), \quad i = 1, \ldots, N_e. \tag{1}$$

The groundwater flow equation, represented by $\psi(\cdot)$ has to be solved independently for each realization of the ensemble.

3. Normal-score transformation step. A normal-score transformation will be applied to all the conductivity values of all the realizations:

$$\widetilde{K^t_{i,j}} = G^{-1}(F_j(K^t_{i,j})), \quad i = 1, \ldots, N_e; \quad j = 1, \ldots, N. \tag{2}$$

where $\widetilde{\mathbf{K}}^{\mathbf{t}}$ and $\mathbf{K}^t$ are the normal-score transformed hydraulic conductivity vector and the hydraulic conductivity vector estimates at time step t, respectively; $F(\cdot)$ is a vectorial normal-score transform function, with N components, one for each location. The normal-score transform function is, generally, a non-parametric function that is built as described in Appendix A. After the normal-score transform of all the elements in all realizations, the transformed ensemble of normal-score conductivity realizations will follow a marginal Gaussian distribution with zero mean and unit variance.

4. Covariance calculation. As it will be explained later, for the updating step the normal-scored conductivity auto-covariance and the cross-covariances between normal-scored conductivity and piezometric heads will be needed. These covariances are non-stationary and need to be computed accounting for the locations of each variable. The procedure is described next. First consider the augmented variable vector

$$\mathbf{S} = \begin{bmatrix} \widetilde{\mathbf{K}} \\ \mathbf{H} \end{bmatrix} \tag{3}$$

with N_e realizations of $2N$ variables, the covariance between any two variables $S_{.,k}$ and $S_{.,l}$ is given by

$$C_{k,l} = \frac{1}{N_e} \sum_{m=1}^{N_e} (S_{m,k} - \langle S_{.,k} \rangle)(S_{m,l} - \langle S_{.,l} \rangle) \quad k = 1, \ldots, 2N; l = 1, \ldots, 2N \tag{4}$$

with

$$\langle S_{.,*} \rangle = \frac{1}{N_e} \sum_{m=1}^{N_e} S_{m,*} \tag{5}$$

Since piezometric heads change in time, and hydraulic conductivities are also updated in time, the covariance is recalculated at each time step and, therefore, it is time dependent. The covariance matrix $\mathbf{C}$ contains $2N \times 2N$ elements; however, it will be explained later that not all elements have to be computed, and the effective number of elements that must be calculated is $2N \times 2N'$, with N' being one or two orders of magnitude smaller than N.

5. Sampling step. Piezometric heads are sampled at a few locations N_h at time step $t + 1$

6. Update step. In this step, a new ensemble of $\widetilde{\mathbf{K}}$ is generated conditioned to the hard conductivity measurements and to the just sampled piezometric heads. Sequential multivariate Gaussian simulation is used as outlined next; the reader interested in knowing all the implementation details of the algorithm is referred to [*Gómez-Hernández and Journel*, 1993; *Deutsch and Journel*, 1992; *Delbari et al.*, 2009]. The steps of the sequential simulation algorithm to generate realization $\widetilde{K_{i,\cdot}}$ (the superindex $t+1$ is omitted for clarity):

(i) Assign the normal-score transformed values of the conditioning conductivity measurements to the closest nodes in the grid.

(ii) Assign the observed piezometric heads to the closest nodes in the grid.

(iii) Generate a random path through all N grid nodes to be simulated.

(iv) Visit a node along the random path. At the node location, search, within a predefined search neighborhood, all $\widetilde{K}$ values already in the grid, and all observed piezometric heads. Then, compute the conditional distribution function given the $\widetilde{K}$ and H data found. Under the assumption of multivariate Gaussianity, this conditional distribution function is Gaussian and its mean and variance are given by the solution of a set of simple kriging equations [e.g., *Deutsch and Journel*, 1992; *Goovaerts*, 1997]. Denoting the node for which the conditional distribution is to be computed by j, the row vector of n conditioning data (normal-scored conductivities and piezometric heads) by $\mathbf{S}_{(n)}$, the covariance matrix between any two variables at the conditioning locations by $\mathbf{C}_\alpha$, and the covariance column vector between the conditioning locations and the location being estimated as $\mathbf{C}_{j,\beta}$, the conditional mean at j is given by

$$m_{\widetilde{K_{i,j}}} = \langle \widetilde{K_{i,j}} \rangle + \mathbf{C}_\alpha^{-1} \mathbf{C}_{j,\beta} (\mathbf{S}_{(n)} - \langle \mathbf{S}_{(n)} \rangle), \tag{6}$$

where $\langle \cdot \rangle$ refers to the average value computed through the ensemble of realizations at a given location as in Eq. (5), and the conditional variance is given by

$$\sigma^2_{\widetilde{K_{i,j}}} = C_{j,j} - \mathbf{C}_{j,\beta}^T \mathbf{C}_\alpha^{-1} \mathbf{C}_{j,\beta}. \tag{7}$$

where T denotes transpose. Note that the use of a search neighborhood limits the pairs of variables for which their covariance is needed while building C_α and $C_{j,\beta}$, it is for this reason that the number of covariance values that have to be precomputed at step 4 is limited to $2N \times N'$, with N' being the number of nodes within the search neighborhood (generally much smaller than N).

(v) Draw a random number λ from a standard Gaussian distribution with zero mean and unit variance, and generate $\widetilde{K_{i,j}}$ as

$$\widetilde{K_{i,j}} = m_{\widetilde{K_{i,j}}} + \lambda \sqrt{\sigma^2_{\widetilde{K_{i,j}}}} \tag{8}$$

(vi) Assign $\widetilde{K_{i,j}}$ to node j and return to step (d) to visit another node until all nodes in realization i have been visited.

The update step is repeated for all realizations in the ensemble.

7. Back transformation step. Back transform the just generated normal-score transformed conductivities into conductivities using the inverse of the previously computed normal-score transform functions:

$$K_{i,j} = F_j^{-1}(G(\widetilde{K_{i,j}})), \quad i = 1, \ldots, N_e; \quad j = 1, \ldots, N. \tag{9}$$

8. Go back to the step 2 and repeat the process for as many time steps as there are observed piezometric heads.

The main difference between the iSS algorithm and the NS-EnKF algorithm is in the updating step. The updating step in the NS-EnKF (as in any other variant of the ensemble Kalman filter) is based on the premise that if there is a departure between forecasted piezometric heads and observed ones it is because there must be a departure between the conductivity estimates and their real values, and this departure can be computed by simple cokriging of the head departures. In the NS-EnKF, at each time step, there is a refinement of the conductivity fields according to the expression

$$\widetilde{K_{i,j}^{t+1}} - \widetilde{K_{i,j}^{t}} = \mathbf{C}_\alpha^{-1}\mathbf{C}_{j,\beta}(\mathbf{S}_{(n)} - \mathbf{S}_{i,(n)}^f), \quad , i = 1, \ldots, N_e; j = 1, \ldots, N \tag{10}$$

where $\mathbf{S}_{(n)}$ is a vector with all the observed piezometric heads, $\mathbf{S}_{i,(n)}^f$ is a column vector containing the forecasted piezometric heads at observation locations for realization i, C_α is a matrix with the covariances of forecasted heads at observation locations, and $C_{j,\beta}$ is a vector with the cross-covariances between normal-scored conductivity at location j and piezometric heads at observation locations.

3. Synthetic Example

A synthetic bimodal confined aquifer composed of 35% high permeability sand and 65% low permeability shale is constructed on a 50 m by 50 m square discretized into a grid of 50 by 50 by 1 cells. The thickness of the confined aquifer is assumed to be 5 m. (The actual units are irrelevant for the purpose of the study, as long as consistency among them is kept.) The construction of the reference field is done in two steps. First, a two-facies field with 9 conditional data (Figure 2) is generated via the SNESIM code by *Strebelle* [2002] using the training image in [*Strebelle*, 2002] (see Figure 3). Second, this binary field is populated independently for each facies with log-conductivity values using the sequential Gaussian simulation code GCOSIM3D *Gómez-Hernández and Journel* [1993] with the parameters shown in Table 1. The resulting reference log-conductivity field and its histogram are shown in Figures 4 and 5. The two figures show that the distribution of log-conductivity is clearly non-Gaussian, the histogram has two modes (one for each facies) and the global mean and standard deviation are -0.9 ln(m/d), and 2.9 ln(m/d), respectively.

The transient groundwater flow simulator MODFLOW [e.g., *McDonald and Harbaugh*, 1984; *Harbaugh et al.*, 2000] is used to solve the transient groundwater equation. The model boundary is impermeable (see Figure 4). Figure 6 shows the locations of wells, including 25 observation wells, 2 injection wells and 3 pumping wells. Observation wells #6 and #7 will be used as calibration wells (post audit) and will not be used for conditioning. The injection rates at the two injection wells #1 and #2 are 16 m^3/d, and 15 m^3/d, respectively. The pumping rates at the three pumping wells #3, #4, and #5 are 7.5 m^3/d, 7.5 m^3/d and 14.5 m^3/d, respectively. The initial head is set to 8 m throughout the study domain. Specific storage is set to 0.03 m^{-1}. The total simulation time is 500 days, discretized into 100 time steps of increasing size following a geometric series with ratio 1.02 (the length of the first time step results in 1.60 days). The piezometric heads simulated in the reference field are sampled at the observation wells for each time step and used as input data for both algorithms.

As already mentioned, the performance of the iSS will be compared to that of the NS-EnKF. We will analyze two scenarios, scenario S0 with the results obtained applying the NS-EnKF, and scenario S1 with the results of the iSS. The data supplied to both algorithms are the same. Both algorithms use the same initial ensemble of 600 conductivity realizations. The initial ensemble of conductivity realizations is generated in a manner similar to which the reference field was generated: first an ensemble of facies realizations, conditional to the same 9 facies values as the reference, is built using SNESIM; then, each facies is independently (and unconditionally) populated using GCOSIM3D with the same parameters in Table 1. We acknowledge that, in practice, we will never have access to the underlying statistics of the conductivity field; however, we choose to use the same random function model to generate the reference field and the initial fields of the ensemble to test the method, making sure that whatever departures there are between the reference and the simulations are due to the implementation of the algorithm and not to uncertainty in the underlying random function model.

4. Analysis

Both the NS-EnKF and the iSS have been used to incorporate the observed piezometric heads during the first 50 time steps (135.4 days) for the generation of an ensemble of conductivity realizations. These realizations are analyzed by looking at several aspects:

1. Histogram of the ensemble. Figure 7a shows the log-conductivity histogram for the initial ensemble. Figure 7b,c display the log-conductivity histograms of the updated ensemble after the 50th time step for scenario S0, and scenario S1, respectively. We can

see from Figure 7b,c that the histogram of log-conductivity with its bimodality is retained after 50 time steps in both scenarios.

2. Ensemble mean and ensemble variance. Figure 8a,b shows the ensemble mean and variance for the initial ensemble. The only conditioning data used for the generation of the initial ensemble is the facies type at 9 locations. This limited information is not enough to control the spatial heterogeneity of each realization; therefore, the ensemble mean only shows some localized high and low values and the ensemble variance is quite high everywhere, with only some small values around the facies conditioning locations. Figures 9 and 10 show the ensemble mean and ensemble variance, after 10 and 50 time steps, respectively. As time passes, and more piezometric heads are used to characterize the ensemble, the realizations of the ensemble are more alike, resulting in ensemble means that delineate the locations of the channels much better than in the initial set, and ensemble variances with zones of virtually no variance and small variances elsewhere. The areas with the highest ensemble variances are those areas with conductivity values which are not as sensitive to the piezometric heads at observation locations as the rest of the aquifer; in a practical case, one could propose the sampling of conductivities at those locations. Both methods perform equally well, with the highest variance reductions for the longer times. The only significative difference is that the ensemble means obtained with the iSS algorithm appear to have a slightly larger short scale variability than the ensemble means obtained with the NS-EnKF.

3. Root mean square error (RMSE) and ensemble spreading (ES) of log-conductivity. In synthetic examples like this one, we can calculate the deviation of the realizations from the “truth”, since we have access to the underlying conductivity distribution from

which the piezometric heads have been observed. The RMSE measures the accuracy of the algorithm in reproducing the reference field, and the ES measures the precision of the ensemble of realizations. The RMSE and ES are given by

$$RMSE = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(\ln K_j^{ref} - \langle \ln K_j \rangle)^2}, \tag{11}$$

$$ES = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\sigma^2_{\ln K_j}}, \tag{12}$$

where $\ln K_j^{ref}$ is the $\ln K$ value at node j in the reference field, $\langle \ln K_j \rangle$ is the ensemble mean, and $\sigma^2_{\ln K_j}$ is the ensemble variance. As discussed by *Chen and Zhang* [2006] when the RMSE and the ES have a similar magnitude, the resulting ensemble variance provides a realistic measure of the uncertainty associated to the ensemble mean estimate.

Figure 11 shows the evolution in time of the RMSE, ES, and the ratio of RMSE to ES for both methods. We can see that the RMSE corresponding to the NS-EnKF is smaller than that for the iSS, indicating that, on average, the realizations obtained with the NS-EnKF are closer to the reference than those obtained by iSS. The ES is similar in both cases. And the ratio RMSE/ES remains closer to 1 for the NS-EnKF than for the iSS, indicating a better characterization of the uncertainty with the former approach. Yet, these differences are small, and the absolute values of RMSE and ES sustain the conclusions derived from the visual analysis of the ensemble means and variances discussed before.

4. Reproduction of the piezometric heads at the two calibration wells. Figure 12 shows the evolution of the piezometric heads for the initial ensemble of conductivity realizations at the two calibration wells #6 and #7 —these wells were not used for conditioning. Since, no piezometric information was used to generate the initial conductivity ensembles, the

spread among the individual responses of each realization is quite large. Figure 13 shows the evolution of the piezometric heads in the updated ensembles for both approaches. The vertical dashed line indicates the end of the use of the observed piezometric heads as conditioning data. Comparing Figures 12 and 13, the effect of conditioning on piezometric heads is patent, with a significant reduction of the spread of the piezometric head curves in all realizations about the reference one for both approaches.

5. Solute breakthrough curves (BTCs). A transport prediction experiment of an inert solute is carried out to further evaluate the goodness of the characterization of conductivity by the updated conductivity realizations. For this purpose, 10 000 conservative particles are released along a vertical line at $x = 2.5$ m and the arrival times are recorded at two control planes, located at $x = 15$ m and $x = 45$ m (see Fig. 4). The random walk particle tracking program RW3D [*Fernàndez-Garcia et al.*, 2005] is used to solve the transport equation. Porosity is assumed constant and equal to 0.3. Figure 14 shows the BTCs at the two control planes for the initial ensemble. Figure 15 shows the BTCs at the two control planes using the conductivity fields updated after 50 times steps. We can see that the uncertainty about the BTC predictions is significantly reduced after conditioning to the piezometric head data. Comparing Figure 15b,d with Figure 15a,c, we can find that the iSS performs a little better than the NS-EnKF, since the spread of the BTC predictions is smaller, and the median BTCs is closer to the reference BTCs. The reproduction of the BTCs is not as good as the reproduction of the piezometric heads, since concentration data were not used for conditioning, this explains that although the spread of the BTCs is significantly reduced, there is a larger bias in the estimation of the reference BTC by the median of the BTCs computed in the upated conductivity fields than in the initial

fields, which we do not attribute to the goodness of any of the two approaches. The next step would be to include some concentration data as conditioning data and check how well either method is able to produce conductivity fields that are conditional to BTC information.

5. Discussion

From the above analysis we can conclude that the quality of the ensemble of realizations generated by iSS is comparable to that of the ensemble generated by the NS-EnKF. The main difference between the two approaches is that iSS is a stochastic simulation approach, whereas the NS-EnKF works by progressively refining an initial ensemble of realizations on the basis of the discrepancy between forecasted and predicted piezometric heads. In the iSS, at each time step, a new ensemble of realizations is generated using as conditioning data the last set of observed piezometric heads; the ensemble of realizations keeps improving as time progresses, through the updating of the experimental ensemble non-stationary covariances. We start with an initial unconditional random function model; at the initialization step, the initial ensemble of realizations already produces a random function model that is conditional on conductivity data; and then, after each time step, the random function model is modified by making it conditional to the observed piezometric heads, too. Furthermore, since iSS is based on the sequential simulation algorithm, which uses a search neighborhood to decide the information to use when computing the cumulative distribution function at a given location, it avoids the use of spurious correlations for long distances, since conditioning data which are far from the point being simulated are never used, solving, in this way, a common problem of the ensemble Kalman filter implementations. There is nothing that prevents the use of the iSS algorithm to condition not only on the observed piezometric heads at time $t+1$, but simultaneously on piezometric heads measured at time t, and at other previous times (something that can also be achieved with the use of ensemble smoothers); including these additional past data would imply only computing the cross-covariances between normal-scored conductivities and the forecasted piezometric heads at the earlier steps, and could improve the final characterization of conductivities. The power of iSS, as compared with other stochastic simulation techniques, is based on the same principle that has made the ensemble Kalman filter so successful, the use of non-stationary covariances experimentally computed from a set of ensemble realizations of conductivity and the resulting ensemble of piezometric heads obtained after running a numerical flow model on the conductivity realizations. Notice that iSS, being an inverse simulation technique, is neither an optimization algorithm nor a filter, it is a stochastic simulation technique.

The code we have developed for the purposes of demonstrating the algorithm is far from been optimal and therefore it does not make too much sense to benchmark it in terms of CPU times. However, we can anticipate that the most time consuming step of the workflow is the computation of the non-stationary covariances; it is this step that should attract the largest optimization efforts at the beginning, once this step has been optimized, we should optimize the simulation algorithm, even considering the use of alternative covariance-based stochastic simulation ones.

This paper presents a new approach for inverse modeling in a stochastic context, which is conceptually very simple and easy to implement. At this stage, the results do not indicate that it should be used for inverse modeling as a replacement of the NS-EnKF, since it performs equally well. However, we believe that the novelty of the approach will warrant extensions and refinements that will yield it superior to existing methods. The algorithm is still in its early stages, but we are confident that it could be used for inverse stochastic simulation under a wide variety of scenarios, not limited to the field of hydrogeology.

6. Summary and conclusions

A new inverse stochastic simulation method, the inverse sequential simulation (iSS), has been proposed for the purpose of generating a set of hydraulic conductivity realizations that, when used to model groundwater flow, can reproduce observed piezometric heads. The method is based on the sequential simulation paradigm making use of the non-stationary covariance experimentally inferred form an ensemble of conductivity fields and the corresponding piezometric head responses, in a manner similar as how it is done in ensemble Kalman filtering. We have benchmarked the technique against the normal-score ensemble Kalman filter (NS-EnKF), possibly the current algorithm that best performs for inverse modeling using transient piezometric heads. The iSS compares very well with the NS-EnKF in all aspects analyzed (histogram reproduction, ensemble mean and variance, reproduction of observed piezometric heads, and reproduction of breakthrough curves), and it can be considered an alternative approach for inverse stochastic simulation. The new algorithm has room for expansion and has the potential of application in other fields.

Appendix A: Normal-score transform

We have N_e realizations, each one discretized into N nodes. Let $K_{i,j}$ the conductivity for realization i at node j. For each node, determine the experimental cumulative distribution function (cdf) from the set of values $\{K_{i,j}, i = 1, \ldots, N_e\}$:

$$F_j(k) = Prob(K \leq k), j = 1, \ldots, N, \tag{A1}$$

these cdfs are generally non-parametric, and they are defined in tabular form by pairs of $(k, F_j(k))$ values. Let $G(y)$ be the cumulative distribution function of a variable y having a Gaussian distribution of zero mean and unit variance.

The normal-score transformation is given by:

$$\widetilde{K_{i,j}} = G^{-1}(F_j(K_{i,j})), \quad i = 1, \ldots, N_e; \quad j = 1, \ldots, N. \tag{A2}$$

Similarly, the normal-score back transform is given by:

$$K_{i,j} = F_j^{-1}(G(\widetilde{K_{i,j}})), i = 1, \ldots, N_e; j = 1, \ldots, N. \tag{A3}$$

Acknowledgments.

The first author acknowledges the financial support from the China Scholarship Council (CSC). Financial support to carry out this work was also received from the Spanish Ministry of Economy and Competitiveness through project CGL2011-23295. We also want to thank Associate Editor Alberto Bellin and three anonymous reviewers for their valuable and constructive comments. All data used in this analysis are available from the authors.

References

Alcolea, A., and P. Renard (2010), Blocking moving window algorithm: Conditioning multiple-point simulations to hydrogeological data, *Water Resources Research*, *46*(8).

Attia, A., and A. Sandu (2014), A sampling filter for non-gaussian data assimilation, *arXiv preprint arXiv:1403.7137*.

Blöschl, G., C. Reszler, and J. Komma (2008), A spatially distributed flash flood forecasting model, *Environmental Modelling & Software*, *23*(4), 464–478.

Carle, S. F., and G. E. Fogg (1996), Transition probability-based indicator geostatistics, *Mathematical Geology*, *28*(4), 453–476.

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.

Delbari, M., P. Afrasiab, and W. Loiskandl (2009), Using sequential gaussian simulation to assess the field-scale spatial uncertainty of soil water content, *Catena*, *79*(2), 163–169.

Deutsch, C. V., and A. G. Journel (1992), *Geostatistical software library and user&s guide*, vol. 1996, Oxford university press New York.

Eggleston, J., and S. Rojstaczer (1998), Identification of large-scale hydraulic conductivity trends and the influence of trends on contaminant transport, *Water Resources Research*, *34*(9), 2155–2168.

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

Fernàndez-Garcia, D., T. Illangasekare, and H. Rajaram (2005), Differences in the scale dependence of dispersivity and retardation factors estimated from forced-gradient and uniform flow tracer tests in three-dimensional physically and chemically heterogeneous porous media, *Water resources research*, *41*(3), W03,012.

Friedel, M. J., and F. Iwashita (2013), Hybrid modeling of spatial continuity for application to numerical inverse problems, *Environmental Modelling & Software*, *43*, 60–79.

Fu, J., and J. 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), 328–341.

Gómez-Hernández, J., and X.-H. Wen (1994), Probabilistic assessment of travel times in groundwater modeling, *Stochastic Hydrology and Hydraulics*, *8*(1), 19–55.

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 X.-H. Wen (1998), To be or not to be multiGaussian: A reflection in stochastic hydrogeology, *Advances in Water Resources*, *21*(1), 47–61.

Goovaerts, P. (1997), *Geostatistics for Natural Resources Evaluation*, 496 pp., Oxford University Press, New York.

Guardiano, F., and R. Srivastava (1993), Multivariate geostatistics: beyond bivariate moments, in *Geostatistics-Troia*, edited by A. Soares, pp. 133–144, Kluwer Academic Publ, Dordrecht.

Harbaugh, A., et al. (2000), *MODFLOW-2000, the US Geological Survey modular ground-water model: User guide to modularization concepts and the ground-water flow process*, US Geological Survey.

Haslauer, C., P. Guthke, A. Bárdossy, and E. Sudicky (2012), Effects of non-gaussian copula-based hydraulic conductivity fields on macrodispersion, *Water Resources Research*, *48*(7).

Hu, L., Y. Zhao, Y. Liu, C. Scheepens, and A. Bouchard (2013), Updating multipoint simulations using the ensemble kalman filter, *Computers & Geosciences*, *51*, 7–15.

Jafarpour, B., and M. Khodabakhshi (2011), A probability conditioning method (PCM) for nonlinear flow data integration into multipoint statistical facies simulation, *Mathematical Geosciences*, *43*(2), 133–164.

Karri, R. R., X. Wang, and H. Gerritsen (2014), Ensemble based prediction of water levels and residual currents in singapore regional waters for operational forecasting, *Environmental Modelling & Software*, *54*, 24–38.

Li, H., E. Kalnay, and T. Miyoshi (2009), Simultaneous estimation of covariance inflation and observation errors within an ensemble kalman filter, *Quarterly Journal of the Royal Meteorological Society*, *135*(639), 523–533.

Li, L., H. Zhou, H. Franssen, and J. Gómez-Hernández (2011), Groundwater flow inverse modeling in non-multigaussian media: performance assessment of the normal-score ensemble kalman filter, *Hydrol. Earth Syst. Sci. Discuss*, *8*(4), 6749–6788.

Li, L., H. Zhou, J. Gómez-Hernández, and H. Hendricks Franssen (2012), Jointly mapping hydraulic conductivity and porosity by assimilating concentration data via ensemble kalman filter, *Journal of Hydrology*.

Mariethoz, G., P. Renard, and J. Straubhaar (2010), The direct sampling method to perform multiple-point geostatistical simulations, *Water Resources Research*, *46*(11), W11,536.

McDonald, M., and A. Harbaugh (1984), *A modular three-dimensional finite-difference ground-water flow model*, Scientific Publications Company.

Sarma, P., and W. Chen (2009), Generalization of the ensemble kalman filter using kernels for nongaussian random fields, in *SPE Reservoir Simulation Symposium*.

Strébelle, S. (2000), Sequential simulation drawing structures from training images, Ph.D. thesis, Stanford University. 187pp.

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

Sudicky, E. A. (1986), A natural gradient experiment on solute transport in a sand aquifer: Spatial variability of hydraulic conductivity and its role in the dispersion process, *Water Resources Research*, *22* (13), 2069–2082.

Sun, A., A. Morris, and S. Mohanty (2009), Sequential updating of multimodal hydrogeologic parameter fields using localization and clustering techniques, *Water Resources Research*, *45* (7), W07,424.

van Loon, M., P. J. Builtjes, and A. Segers (2000), Data assimilation of ozone in the atmospheric transport chemistry model lotos, *Environmental Modelling & Software*, *15* (6), 603–609.

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* (1), 129–156.

Xu, T., J. Jaime Gómez-Hernández, H. Zhou, and L. Li (2013), The power of transient piezometric head data in inverse modeling: an application of the localized normal-score enkf with covariance inflation in a heterogenous bimodal hydraulic conductivity field, *Advances in Water Resources*, *54*, 100–118.

Zhou, H., J. Gómez-Hernández, H. Hendricks Franssen, and L. Li (2011), An approach to handling non-gaussianity of parameters and state variables in ensemble kalman filtering, *Advances in Water Resources*, *34* (7), 844–864.

Zhou, H., J. J. Gómez-Hernández, and L. Li (2012a), A pattern-search-based inverse method, *Water Resources Research*, *48*(3).

Zhou, H., L. Li, H. Hendricks Franssen, and J. Gómez-Hernández (2012b), Pattern recognition in a bimodal aquifer using the normal-score ensemble kalman filter, *Mathematical Geosciences*, pp. 1–17.

Zhou, H., J. J. Gómez-Hernández, and L. Li (2014), Inverse methods in hydrogeology: Evolution and recent trends, *Advances in Water Resources*, *63*, 22–37.

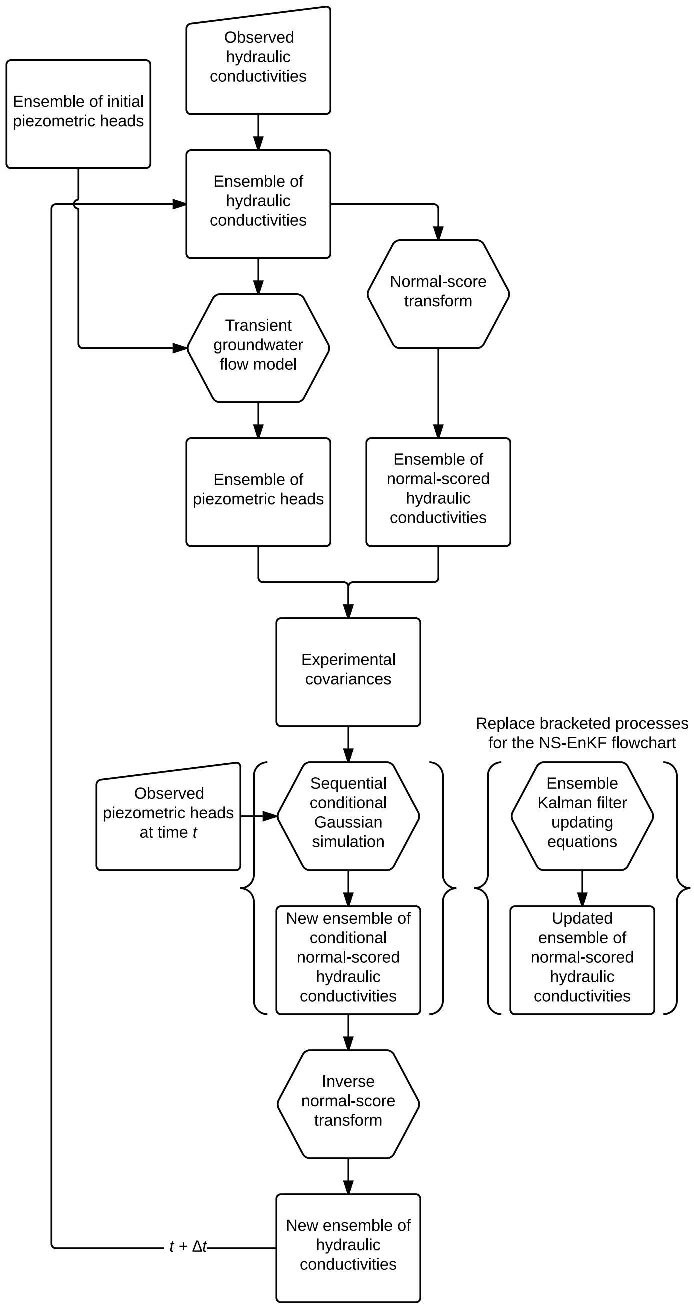


Figure 1. Flowchart of the iSS. For the flowchart of the NS-EnKF, replace the bracketed processes for the floating ones in brackets on the right.

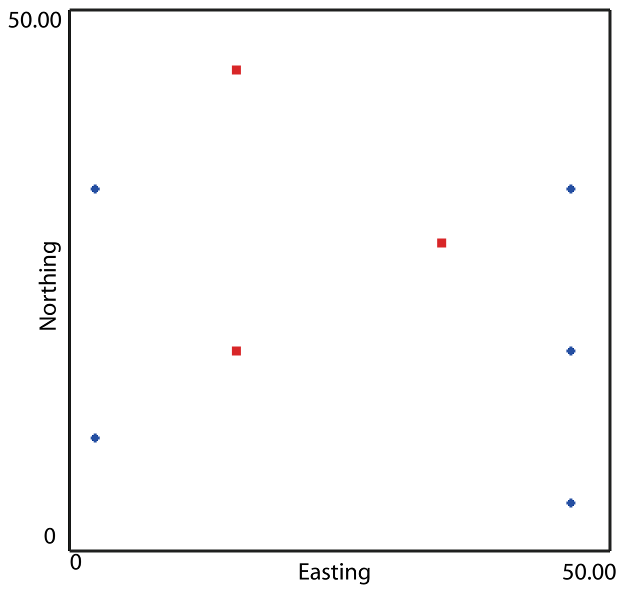


Figure 2. Location of the conditional data. The red nodes denote shale; the green nodes denote sand.

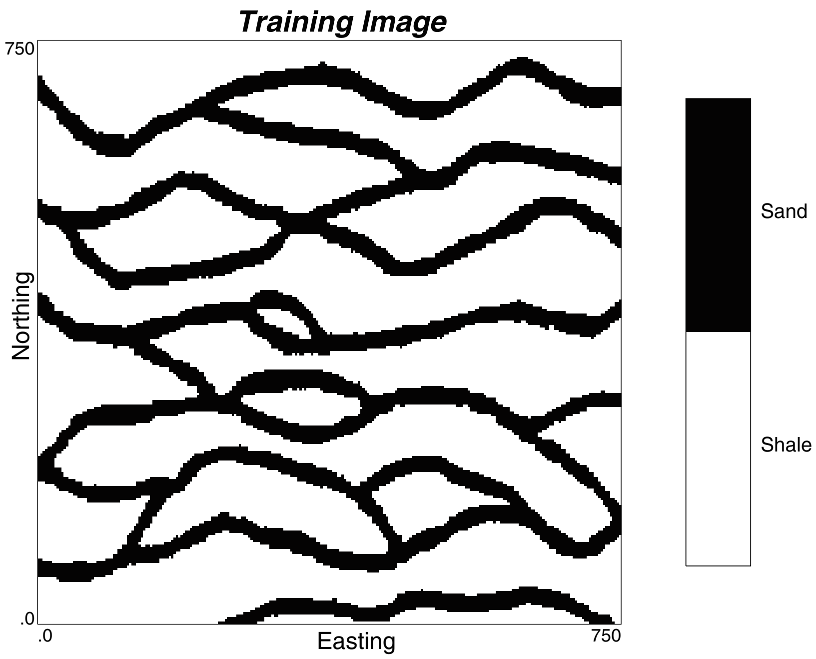


Figure 3. Training image used to generate the ensemble of binary facies realizations.

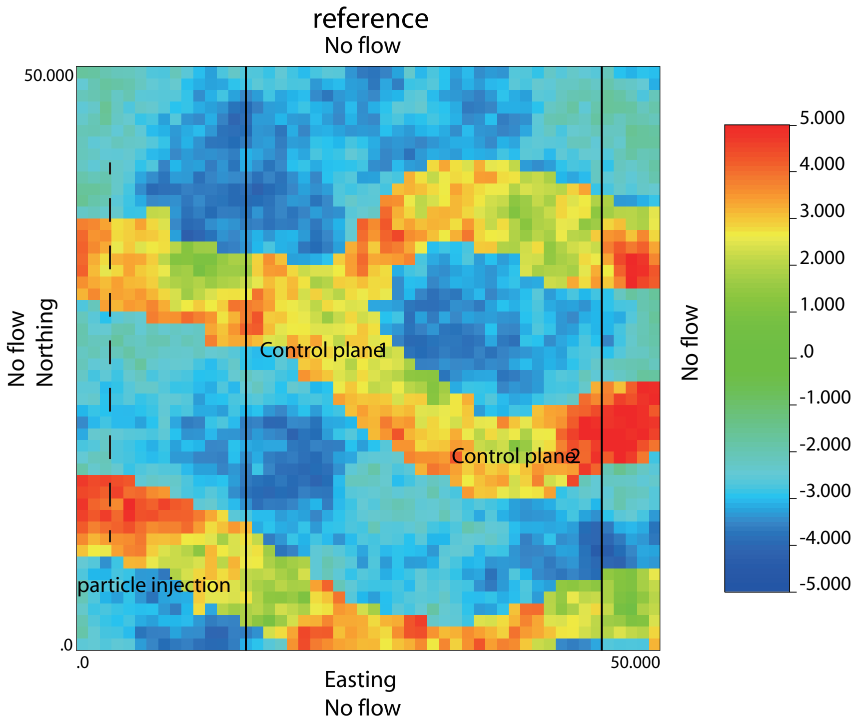


Figure 4. Reference ln$K$ field. It shows the boundary conditions; it also shows the source line (dashed) and the control planes at which breakthrough curves are computed (solid lines) in the transport experiment.

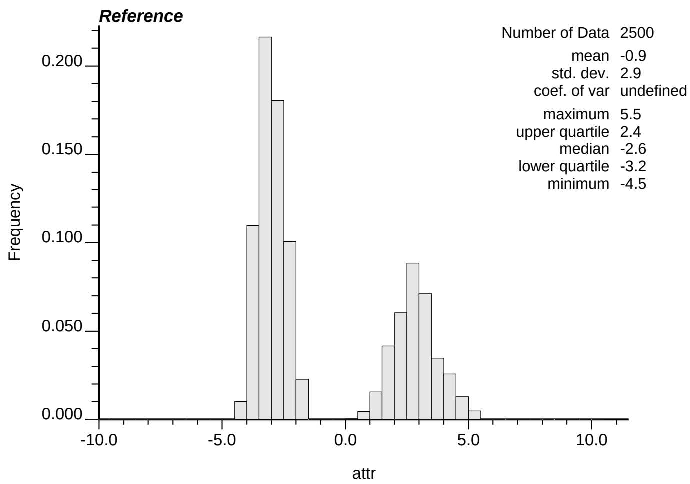


Figure 5. The histogram of the reference log-conductivity field.

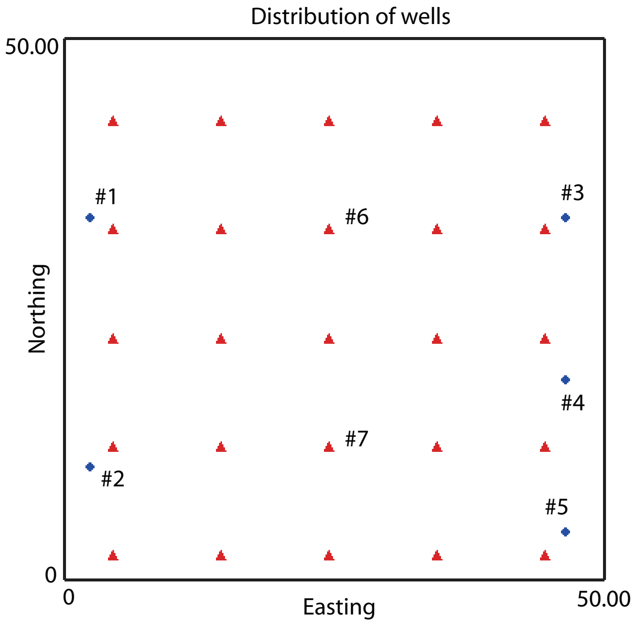


Figure 6. Location of wells. Red triangles denote observation wells; blue squares denote injection (#1, #2) and pumping wells (#3, #4 and #5). The observation wells #6, #7 are used as calibration wells.

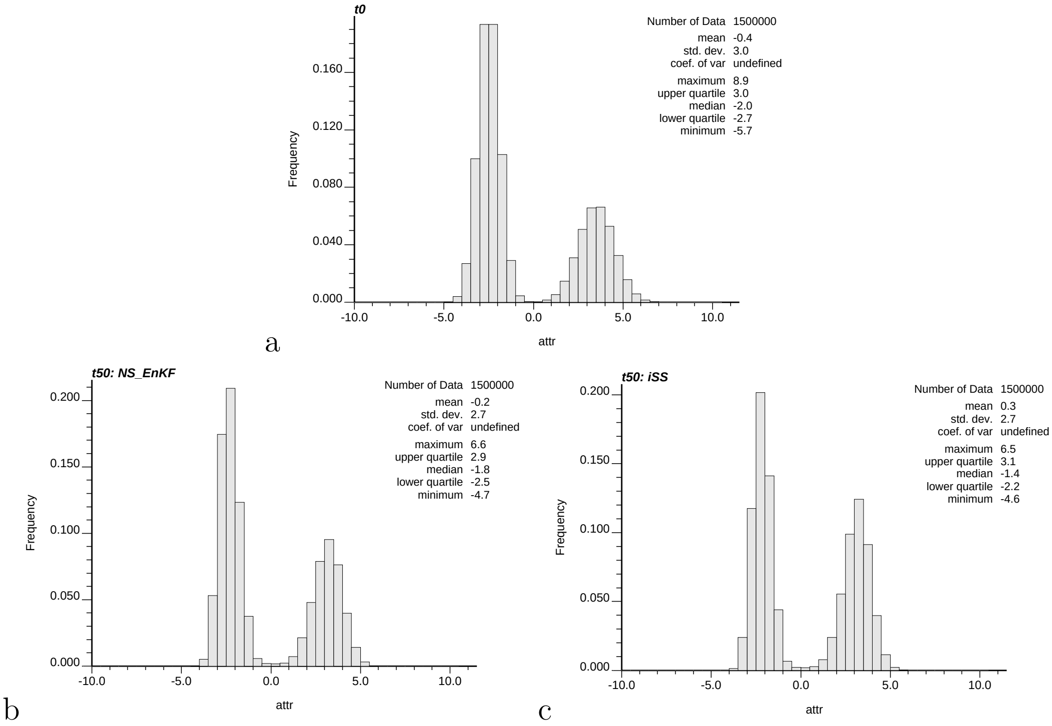


Figure 7. Log-conductivity histograms of the initial ensemble of realizations and of the updated ensemble of realizations after the 50th assimilation step of the two scenarios.

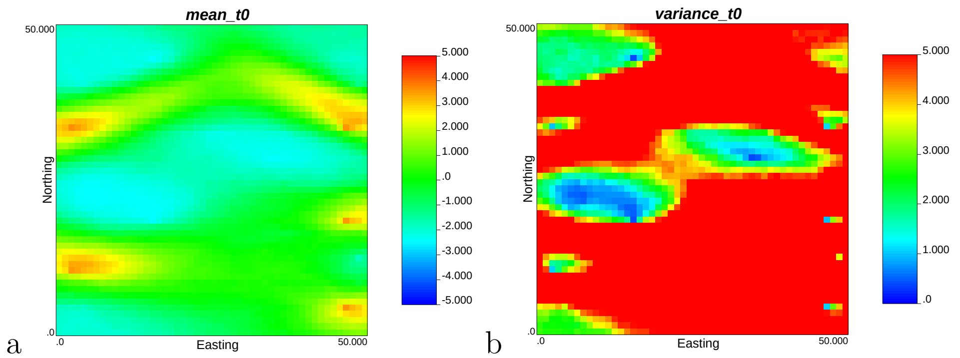


Figure 8. Ensemble mean and ensemble variance of ln$K$ for the initial realizations.

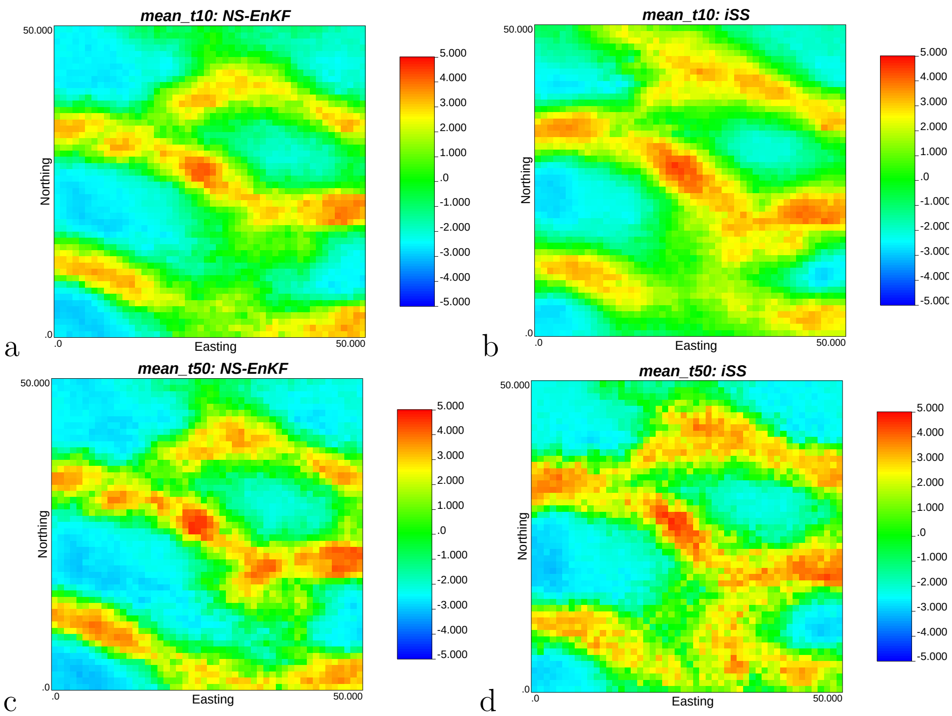


Figure 9. Ensemble mean of ln$K$ after assimilating observation heads at the 10th and 50th time steps for the two scenarios.

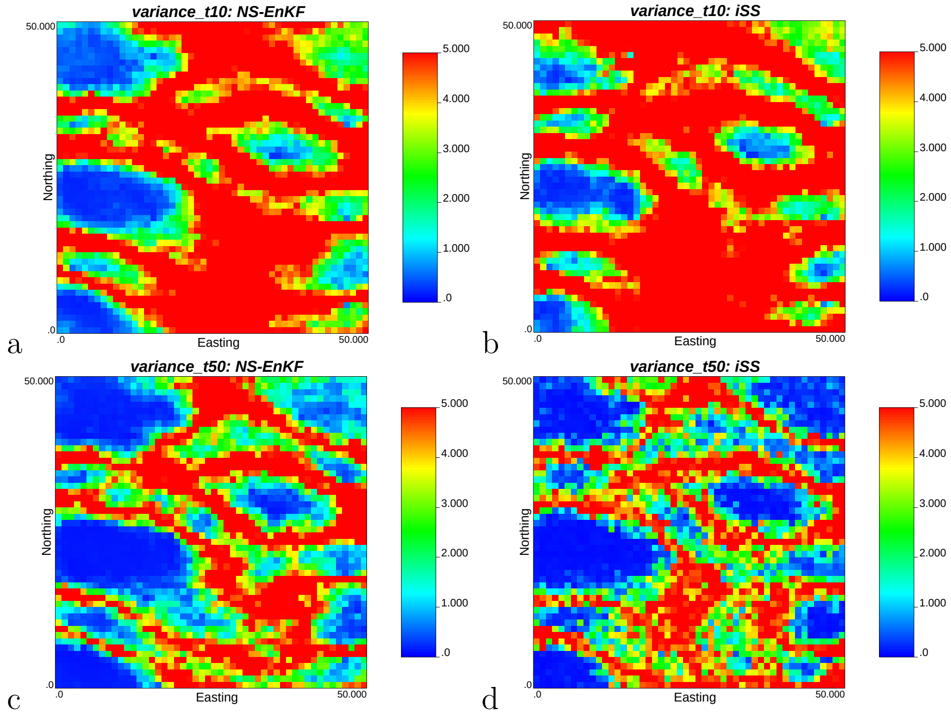


Figure 10. Ensemble variance of $\ln K$ after assimilating observation heads at the 10th and 50th time steps for the two scenarios.

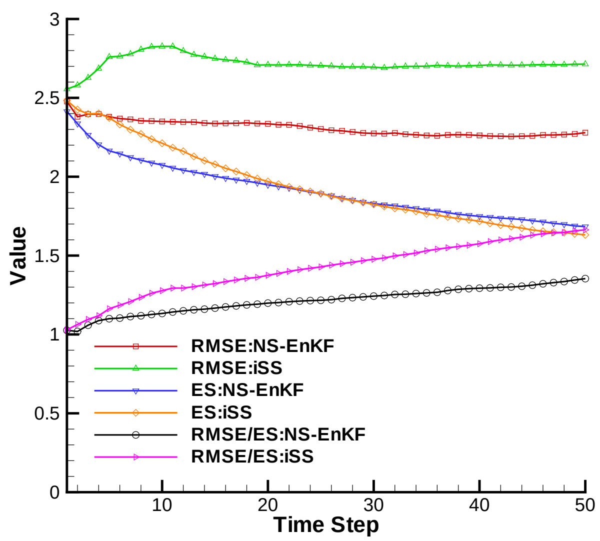
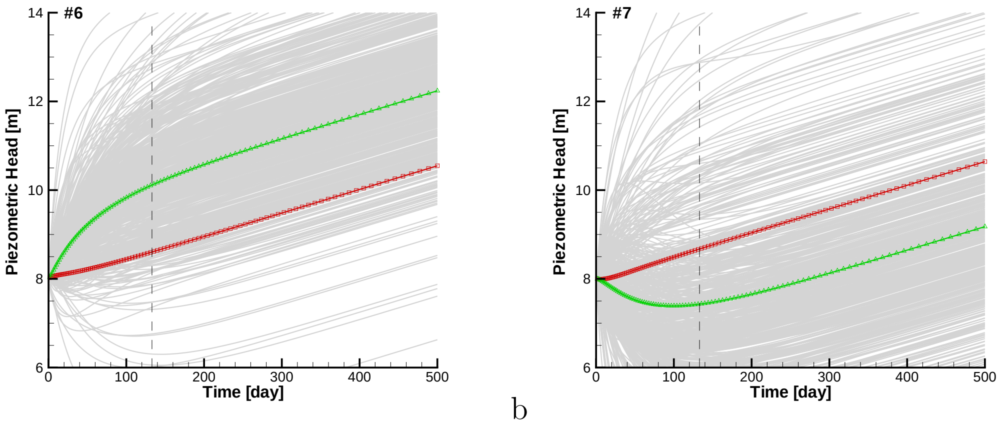


Figure 11. RMSE, ES, and RMSE/ES.



a b

Figure 12. Evolution in time of the piezometric head at the two calibration wells for the initial ensemble of log-conductivity realizations. The red square line corresponds to the piezometric head in the reference, the gray lines correspond to the realizations and the vertical dashed lines marks the end of the conditioning period, the green delta line corresponds to the average of the gray lines.

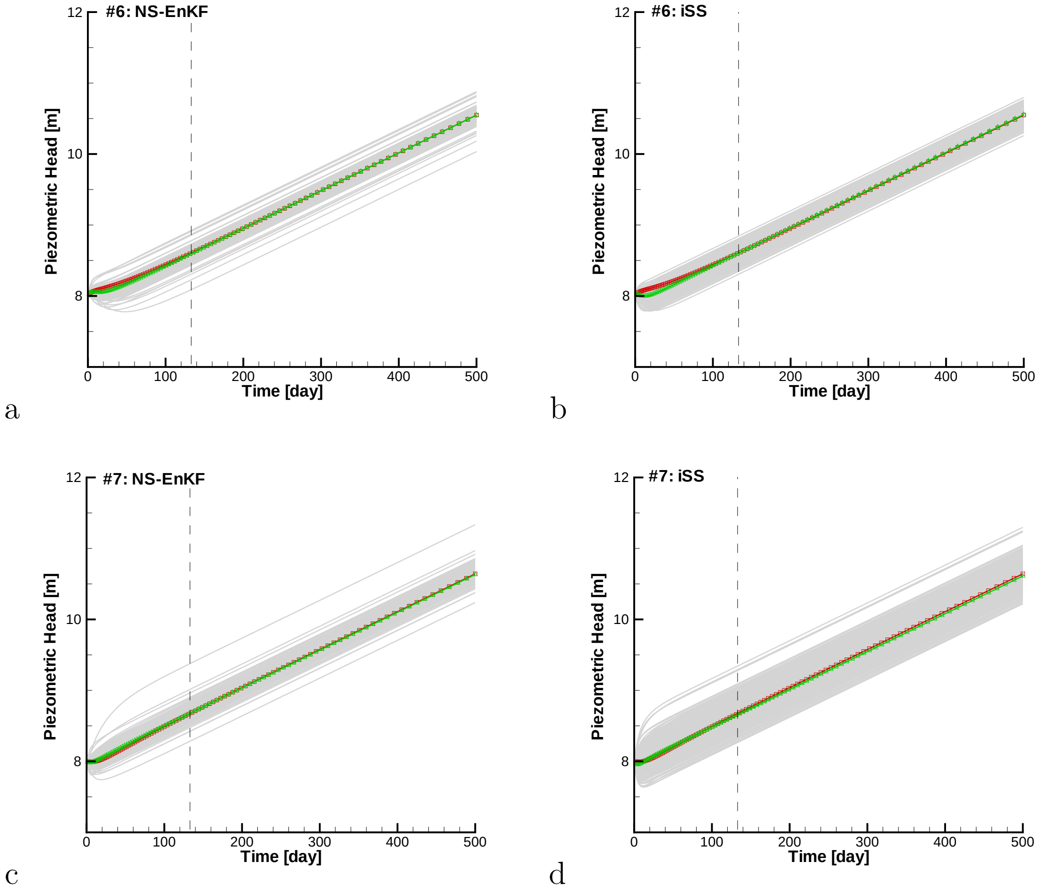
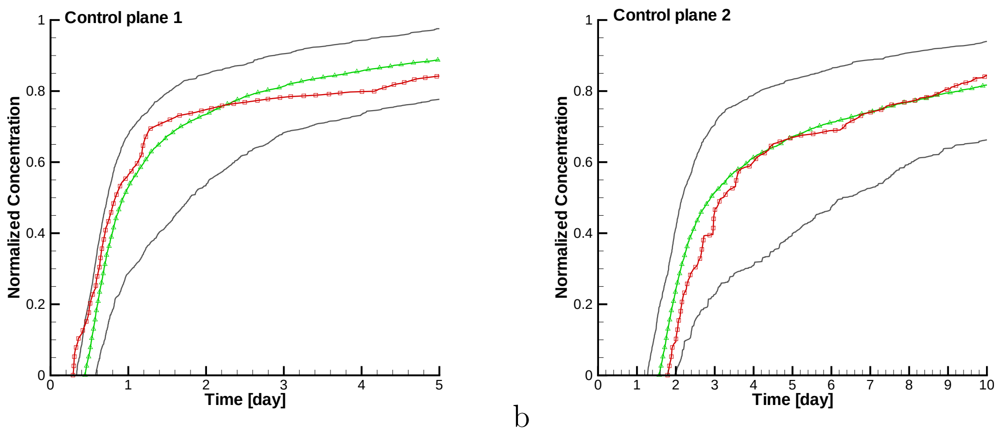
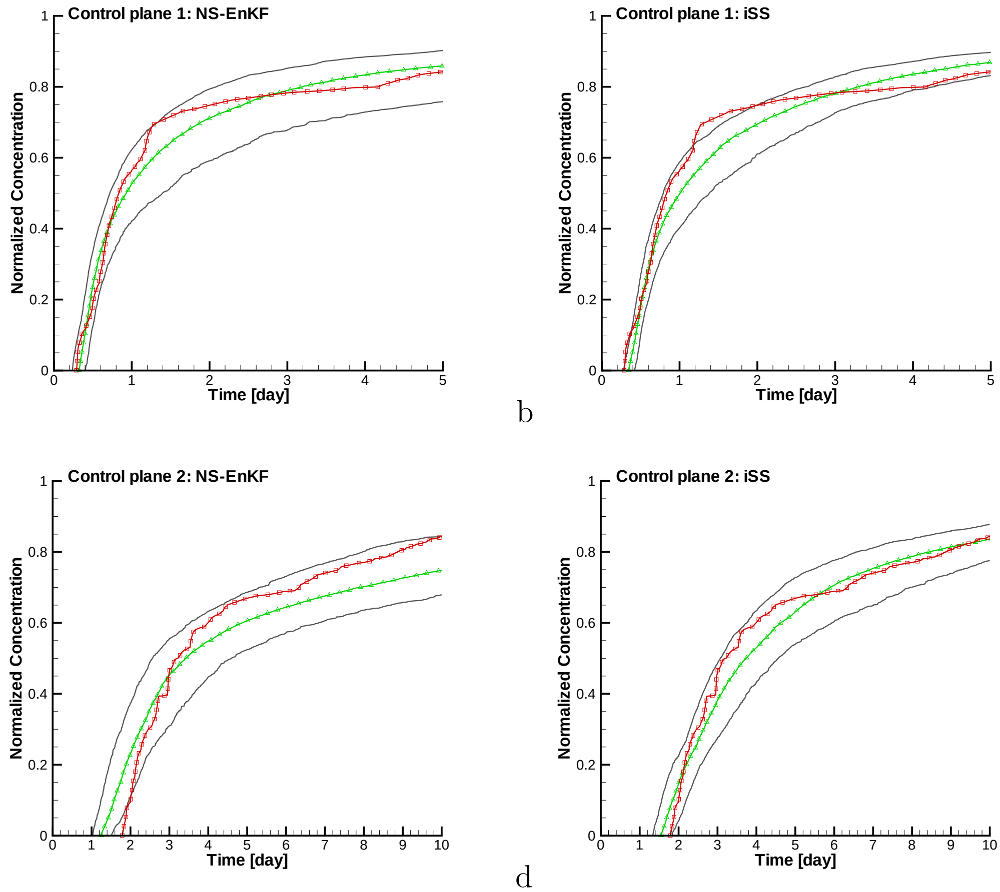


Figure 13. Evolution in time of the piezometric head at the two calibration wells after the 50th time step. The red square line corresponds to the piezometric head in the reference, the gray lines correspond to the realizations and the vertical dashed lines marks the end of the conditioning period, the green delta line corresponds to the average of the gray lines.



a b

Figure 14. Breakthrough curves (BTC) at the two control planes for the initial $\ln K$ realizations. The red square line corresponds to the BTCs in the reference. The black lines correspond to the 5 and 95 percentiles of all realization BTCs, and the green delta line corresponds to the median.



a b c d

Figure 15. Breakthrough curves (BTC) at the two control planes for the ln$K$ realizations updated after the 50th time step. The red square line corresponds to BTCs in the reference. The black lines correspond to the 5 and 95 percentiles of all realization BTCs, and the green delta line corresponds to the median.

Table 1. Parameters of the random functions describing the spatial continuity of the sand and shale log-conductivities

Facies	Proportion	Mean [ln [m/d]]	Std.dev [ln [m/d]]	Variogram type	λ_x [m]	λ_y [m]	sill
Sand	0.35	3.5	1.0	exponential	20	20	1
Shale	0.65	-2.5	0.6	exponential	20	20	0.35