

An Approach to Handling Non-Gaussianity of Parameters and State Variables in Ensemble Kalman Filtering

Haiyan Zhou^{a,b,*}, J. Jaime Gómez-Hernández^b, Harrie-Jan Hendricks Franssen^c, Liangping Li^b

^a*School of Water Resources and Environment, China University of Geosciences, 29 Xueyuan Lu, 100083, Beijing, China*

^b*Group of Hydrogeology, Department of Hydraulics and Environmental Engineering, Universitat Politècnica de València, 46022, Valencia, Spain*

^c*Agrosphere, IBG-3, Forschungszentrum Jülich GmbH, 52425, Jülich, Germany*

Abstract

The ensemble Kalman filter (EnKF) is a commonly used real-time data assimilation algorithm in various disciplines. Here, the EnKF is applied, in a hydrogeological context, to condition log-conductivity realizations on log-conductivity and transient piezometric head data. In this case, the state vector is made up of log-conductivities and piezometric heads over a discretized aquifer domain, the forecast model is a groundwater flow numerical model, and the transient piezometric head data are sequentially assimilated to update the state vector. It is well known that all Kalman filters perform optimally for linear forecast models and a multiGaussian-distributed state vector. Of the different Kalman filters, the EnKF provides a robust solution to address earities; however, it does not handle well non-Gaussian state-vector distributions. In the standard EnKF, as time passes and more state observations are assimilated, the distributions become closer to Gaussian, even if the initial ones are clearly non-Gaussian. A new method is proposed that transforms the original state vector into a new vector that is univariate Gaussian at all times. Back transforming the vector after the filtering ensures that the initial non-Gaussian univariate distributions of the state-vector components are preserved throughout. The proposed method is based in normal-score transforming each variable for all locations and all time steps. This new method, termed the normal-score ensemble Kalman filter (NS-EnKF), is demonstrated in a synthetic bimodal aquifer resembling a fluvial deposit, and it is compared to the standard EnKF. The proposed method performs better than the standard EnKF in all aspects analyzed (log-conductivity characterization and flow and transport predictions).

Keywords: non-Gaussian; ensemble Kalman filter; parameter identification; data assimilation; uncertainty; groundwater modeling.

*Corresponding author

Email addresses: haizh@upvnet.upv.es (Haiyan Zhou), jaime@dihma.upv.es (J. Jaime Gómez-Hernández), h.hendricks-franssen@fz-juelich.de (Harrie-Jan Hendricks Franssen), liali@upvnet.upv.es (Liangping Li)

1. Introduction

Groundwater modeling and prediction plays a critical role in decision making for groundwater management and environmental protection. In order to make reliable groundwater flow model predictions, it is necessary to account for all measured data. Although important information can be obtained from field work, it is in practice impossible to characterize an aquifer exhaustively. To best account for the state information—such as flows, hydraulic heads or concentrations—in the characterization of aquifer parameters, numerous methods of parameter identification have been proposed [e.g., for an overview see 1, 2, 3, 4, 5, 6].

In the early days of parameter identification, the aim was to obtain a single “best” estimate of the aquifer parameters. However, it has been proven that such a “best” estimate is always much smoother than the real aquifer and that flow and transport predictions performed in such estimate are very poor [7]. The alternative is to resort to Monte Carlo analysis in which multiple realizations of the aquifer parameters are built accounting for all the measured data. Each realization represents a possible case of the unknown reality. Flow and transport are modeled in each realization and all model predictions are collected to characterize uncertainty and to define an optimal prediction (e.g., pilot point method [8, 9], self-calibration [10] or gradual deformation [11]).

One such Monte Carlo-based method is the ensemble Kalman filter (EnKF) proposed by Evensen [12] and subsequently clarified by Burgers et al. [13]. The EnKF is an extension of the Kalman filter to deal with nonlinear state equations. The Kalman filter [14] proved to be a very powerful data assimilation algorithm for systems in which the relation between parameters and state is linear. This linearity allowed an exact propagation of the state covariance in time. The first attempt to deal with nonlinear transfer functions, such is the case in groundwater modeling, was the extended Kalman filter [e.g., 15, 16]. In the extended Kalman filter, the nonlinear transfer function is linearized after a Taylor expansion and this linearization is used for the covariance propagation. For highly nonlinear transfer functions, extended Kalman filter tends to deteriorate as time progresses, since the errors in the covariance propagation accumulate; at the same time, extended Kalman filter is time consuming when the aquifer is finely discretized [17].

The EnKF circumvents the problem of covariance propagation in time by working with an ensemble of realizations. In each realization the state equation is solved, and the ensemble of states is used to compute the covariance explicitly and efficiently. The EnKF has gained popularity in diverse disciplines such as oceanography, meteorology and hydrology [e.g., 18, 19, 20, 21, 22, 23, 24]. The advantages of the EnKF can be summarized as follows: first, CPU consumption is reduced mostly because of the way the state

covariance is computed (Hendricks Franssen and Kinzelbach [25] documented a reduction of the needed CPU time by a factor of 80 as compared with other Monte-Carlo type inverse modeling); second, the EnKF can easily be combined with different forecast models; third, the EnKF is capable of incorporating the observations sequentially in time without the need to store all previous states. In addition, although Evensen [12] developed the EnKF to obtain a single optimal estimate of the system state, the EnKF provides, as a by-product, the entire ensemble of states, which can be used to assess uncertainty.

Our objective with this paper is to use the EnKF for the generation of an ensemble of hydraulic conductivity realizations which are conditional to hydraulic conductivity measurements and to piezometric head data. This approach has already been described in the literature, for instance by Hendricks Franssen and Kinzelbach [26], who generated realizations of transmissivity and recharge, and by Liu et al. [27], who focused on hydraulic conductivity, dispersivity, mass transfer rate and mobile porosity ratio for a dual-domain mass transfer model at the MADE site. Example applications in petroleum engineering can be found in the works by Wen and Chen [28] and Gu and Oliver [29]. Our contribution is to develop a new approach applicable to non-Gaussian distributions of hydraulic conductivities, with an application to a bimodal distribution typical, for instance, of fluvial deposits.

It can be shown that the EnKF provides an optimal solution when the parameter vector follows a multiGaussian distribution and the state transfer function is linear [30]. In most practical applications in groundwater and petroleum engineering neither the parameters can be modeled as multiGaussian nor the transfer function is linear. The importance of accounting for the nonGaussianity of hydraulic conductivity has already been demonstrated in the literature [31, 32, 33]. To circumvent the problems of the EnKF, some authors have concentrated in the problem of non-Gaussianity in the parameters and others have focused on reparameterizing the state equation so that the relationship between model parameters and state variables is closer to linear. Sun et al. [34] worked on the non-Gaussianity aspect and took advantage of localization techniques with a Gaussian mixture model to update the parameters of a multimodal distribution. Chen et al. [35] addressed the nonlinearity problem by reparameterization.

In this work we take the route of transforming parameters and state variables so that they both follow marginal Gaussian distributions. These transformations, which are themselves highly nonlinear, will make the state transfer function even more nonlinear than for the untransformed variables, but, in return, the Kalman filtering equations will be applied on Gaussian variates. We demonstrate this approach on a synthetic aquifer resembling a highly channelized fluvial deposit. It will be shown how this approach improves the results obtained using a standard implementation of the EnKF even though the univariate transformation

applied ensures marginal Gaussian distributions but does not ensure multiGaussianity of the joint state vector.

We are aware that in another study [36] similar transformation techniques as proposed in this paper were applied but only to the state variables, not to the parameters, in the context of hydraulic tomography. The authors focus on the improvements that can be achieved with different transformation techniques, analyze under which conditions improvements can be obtained, and demonstrate a pseudo-linearizing effect of their transformations. For reference, they compared their improved method to an exhaustive particle filter. Similar applications can be found in other disciplines; for instance, in reservoir modeling, transformation from non-Gaussian distribution to Gaussian distribution is applied to the state variables, such as saturation, by Gu and Oliver [37] and, in ocean ecosystem modeling, a similar transformation is applied to chlorophyll-*a* concentration by Simon and Bertino [38]. Other transformation algorithms can be found in the literature such as in Béal et al. [39] and Bocquet et al. [40]. In contrast with all of these applications, the method proposed in this paper focuses on transforming not only the non-Gaussian distributed state variables but, most importantly, the non-Gaussian distributed model parameters, i.e., the hydraulic conductivities, which are commonly assumed to follow a log-normal distribution. These transformation of parameters and state variables will make the relationship between the transformed variables even more nonlinear (instead of more linear as pursued in the previously cited works), but the Kalman filtering equations will be applied to Gaussian variables. To our knowledge, no such transformation algorithm has been applied to handle the non-Gaussian distribution of hydraulic conductivities in the scope of the EnKF in hydrogeology.

Throughout the paper we use interchangeably the terminology from geostatistics, hydrogeology and Kalman filtering, most noticeably, a) piezometric head data assimilation is equivalent to (inverse) data conditioning, in the sense that the solution of the flow equation in each of the final realizations of hydraulic conductivity will match the measured piezometric heads, b) by forecast model, transfer function or state transition model we refer to the transient groundwater flow equation and its corresponding numerical model, and c) when referring to the flow equation we distinguish between parameters (i.e., hydraulic conductivities) and state variables (i.e., piezometric heads), whereas in the EnKF we will talk about a (joint) state vector that includes realizations of both parameters and state variables at time t . Also, since our goal is the characterization of the hydraulic conductivity spatial variability, we will be using the term hard data to refer to local measurements of hydraulic conductivity, as opposed to measurements of piezometric heads which are soft data since they do not measure directly hydraulic conductivity but serve to characterize its spatial variability.

The rest of the paper is structured as follows. After the standard EnKF is introduced, the new algorithm is explained in detail. Then a synthetic bimodal aquifer is used to evaluate the performance of the proposed method. The paper ends with a discussion and some conclusions.

2. Methodology

We first present the groundwater flow and transport equations that will be used in the synthetic example, then we follow with the description of the standard EnKF and propose the new algorithm with the transformed variables, which will be referred to as the normal-score EnKF (NS-EnKF). The flow model will be used in conjunction with ensemble Kalman filtering to obtain realizations of hydraulic conductivity conditioned to both conductivity and piezometric head data. The transport model will be used only for verification purposes to evaluate how well transport is predicted in the final conductivity fields.

2.1. Flow and transport equations

By combining mass conservation and Darcy's law, the groundwater flow equation in saturated porous media can be expressed as [41]:

$$\nabla \cdot (K \nabla h) = S_s \frac{\partial h}{\partial t} + Q \quad (1)$$

subject to initial and boundary conditions. In the differential equation, $\nabla \cdot$ is the divergence operator, ∇ is the gradient operator, K is the hydraulic conductivity [LT^{-1}], h represents hydraulic head [L], S_s is specific storage [L^{-1}], Q is the source-sink term [T^{-1}] and t is time [T]. Equation (1) is numerically solved by applying a discretization in space and time.

The governing equation for non-reactive transport in the subsurface is:

$$\phi \frac{\partial C}{\partial t} = -\nabla \cdot (qC) + \nabla \cdot (\phi \mathbf{D} \nabla C) \quad (2)$$

subject to initial and boundary conditions, where C is the concentration of solute in the liquid phase [ML^{-3}]; ϕ is porosity [-]; \mathbf{D} is the local hydrodynamic dispersion coefficient tensor [L^2T^{-1}] with principal axes parallel and perpendicular to the direction of flow and eigenvalues proportional to the components of the fluid velocity and the longitudinal and transverse dispersivities, and q is the Darcy velocity [LT^{-1}] calculated by Darcy's law: $q = -K \nabla h$.

2.2. Standard EnKF

The theory and numerical implementation of the EnKF is described extensively in Evensen [17, 42]. Here we will only recall that the EnKF deals with dynamic systems, for which state data are observed as a function of time and used to sequentially update the parameters and the state of the system. For this purpose an ensemble of realizations is generated and then each realization is updated as observation data are available.

In the EnKF, the state vector \mathbf{x} is represented by the collection of all state vectors for all N_r realizations, $\mathbf{x} = (x_1, x_2, \dots, x_{N_r})$. For realization i , x_i includes both the parameters controlling the transfer function and the state variables:

$$x_i = \begin{pmatrix} A \\ B \end{pmatrix}_i = \begin{pmatrix} (a_1, a_2, \dots, a_{N \cdot N_p})^T \\ (b_1, b_2, \dots, b_{N \cdot N_s})^T \end{pmatrix}_i \quad (3)$$

where A is the vector of model parameters, such as hydraulic conductivities and porosities, and B is the vector with the state variables, such as hydraulic heads and concentrations. The size of the state vector for each realization is determined by the number of grid cells in which the domain has been discretized (N), the number of model parameters (N_p), and the number of state variables (N_s). In our case, only one type of parameter (log-hydraulic conductivity $\ln K$) and one type of state variable (piezometric head h) are considered, therefore

$$x_i = \begin{pmatrix} (\ln K_1, \ln K_2, \dots, \ln K_N)^T \\ (h_1, h_2, \dots, h_N)^T \end{pmatrix}_i. \quad (4)$$

The final size of the ensemble state vector is equal to the size of each realization state vector multiplied by the number of realizations. In our case, it is $2 \times N \times N_r$.

The EnKF consists of two main steps: a forecast step and an update step. Both steps are to be performed in each realization. The forecast step involves the transition of the state variables from time $t - 1$ to time t ,

$$\mathbf{x}_t = F(\mathbf{x}_{t-1}) \quad (5)$$

where $F(\cdot)$ is the transfer function. In our case this transfer function leaves the log-conductivities unchanged and forecasts the piezometric heads to the next time step using the transient groundwater flow model (Equation (1)).

After data are collected, the state vector is updated by the EnKF. The update process is summarized by

the following equations:

$$\begin{aligned}
\mathbf{x}_t^u &= \mathbf{x}_t^f + \mathbf{G}_t(\mathbf{z}_t + \varepsilon - \mathbf{H}\mathbf{x}_t^f) \\
\mathbf{G}_t &= \mathbf{P}_t^f \mathbf{H}^T (\mathbf{H} \mathbf{P}_t^f \mathbf{H}^T + \mathbf{R})^{-1} \\
\bar{\mathbf{x}}_t^f &\approx \frac{1}{N_r} \sum_{i=1}^{N_r} \mathbf{x}_{t,i}^f \\
\mathbf{P}_t^f &\approx \frac{1}{N_r} (\mathbf{x}_t^f - \bar{\mathbf{x}}_t^f)(\mathbf{x}_t^f - \bar{\mathbf{x}}_t^f)^T
\end{aligned} \tag{6}$$

where \mathbf{x}_t^u is the vector with the updated model parameters and state variables at time t and \mathbf{x}_t^f is the vector with the forecasted state; \mathbf{G}_t is the Kalman gain—derived after the minimization of a posterior error covariance—it multiplies the residuals between observed and forecasted values to provide an update to the latter; \mathbf{z}_t are the observations at time t ; ε is an observation error, generally characterized by a normal distribution with zero mean and a diagonal covariance \mathbf{R} (we assume that errors at different measurement locations are independent); \mathbf{H} is the observation matrix which has as many rows as number of measurements (M) and in our case consists of 0's and 1's since we assume that measurement locations coincide with the grid nodes; $\mathbf{x}_{t,i}^f$ is a realization of the ensemble of forecasted state vectors; $\bar{\mathbf{x}}_t^f$ is the ensemble mean; and \mathbf{P}_t^f is the ensemble covariance matrix of the state \mathbf{x}_t^f . It is worth noting that we do not have to compute the whole covariance matrix explicitly because we can compute directly $\mathbf{P}_t^f \mathbf{H}^T$ and $\mathbf{H} \mathbf{P}_t^f \mathbf{H}^T$ taking advantage of the fact that most of the entries of \mathbf{H} are zeroes.

The EnKF has an intrinsic feature in honoring parameter data (i.e., hydraulic conductivity measurements) as long as they remain constant during the forecast step and parameter measurement errors are neglected. In such a case the components of the covariance matrix \mathbf{P}_t^f involving parameter measurement locations are equal to zero, the Kalman gain at those locations is also zero, and therefore they remain constant also through the update step.

2.3. Normal-Score EnKF

As we have mentioned before, we are going to modify the formulation of the standard EnKF so that we ensure that the joint state vector follows marginal Gaussian distributions at all locations and all times. This is achieved by applying a normal-score transform, independently, to each element of the state vector as discussed in the appendix. This normal-score transform function (TF) has to be recomputed after each forecast step. Let $\Phi_t(\cdot)$ represent the normal-score transformation at time t so that

$$\mathbf{y}_t = \Phi_t(\mathbf{x}_t) \tag{7}$$

is a new state vector in which all variables have a marginal Gaussian distribution with zero mean and unit variance. Similarly we will have a normally distributed state vector at time $t - 1$, $\mathbf{y}_{t-1} = \Phi_{t-1}(\mathbf{x}_{t-1})$. Substituting \mathbf{y} for \mathbf{x} in the transfer function (5) we obtain

$$\mathbf{y}_t = \Phi_t(F(\Phi_{t-1}^{-1}(\mathbf{y}_{t-1}))), \quad (8)$$

which we can rewrite as

$$\mathbf{y}_t = \varphi_t(\mathbf{y}_{t-1}) \quad (9)$$

with $\varphi_t = \Phi_t \cdot F \cdot \Phi_{t-1}^{-1}$. We have replaced the original nonlinear transfer function $F(\cdot)$ with a new ear transfer function $\varphi_t(\cdot)$ that takes as input a vector of Gaussian variables and propagates it in time into a new vector which is also Gaussian. (Notice that the normal score transform only renders the variables Gaussian one by one, the multivariate properties of the state vector are also changed but not necessarily to become multiGaussian.)

The inverse of the normal-score transform function permits, at any time step and for any location, to retrieve the state vector \mathbf{x} from the vector \mathbf{y} ensuring that the marginal non-Gaussian distribution of \mathbf{x} is kept. This is particularly interesting in cases such as aquifer modeling in fluvial deposits, which are characterized by multimodal distributions of log-conductivities.

Once the new transfer function in (9) is established, the NS-EnKF just follows the same steps as the standard EnKF. However, since the transfer function depends on the normal-score transform function that, in turn, depends on the state values, there is not an explicit expression. A flow chart of the NS-EnKF is displayed in Figure 1 which consists of the following steps:

1. Generate the initial ensemble. A large number of equally likely stochastic realizations of the state vector are generated. In our case the state vector consists of log-conductivities and piezometric heads; the log-conductivities are generated using geostatistical simulation techniques [43], which are conditioned to hard data; the piezometric heads are set equal to the initial heads, which in our case are zero everywhere (alternatively the initial heads could be set equal to the steady-state solution of the flow equation for given boundary conditions and external sources, or they could be obtained from a warm-up run).
2. Forecast. For each realization, the log-conductivities stay unchanged, and the piezometric heads are the solutions of the groundwater flow equation in each realization from time $t - 1$ to time t .
3. Normal-score transform. Establish the local cumulative distribution functions (CDFs) for all the

components of the state vector from the ensemble of realizations. In our case there will be one such local CDF at each location for the log-conductivity and another one for the piezometric head. Use these local CDFs to build the normal-score transformation function and transform the state vector into a new vector, with all its components following marginal Gaussian distributions with zero mean and unit variance. It is worthwhile to mention that the log-conductivities transform functions remain unchanged during the data assimilation (TF1 in Figure 1); this helps recovering the prior model structure. On the contrary, the transformation functions for piezometric heads have to be recomputed at each time step (TF2 in Figure 1).

4. Update. State data are collected at time t . These data are normal-score transformed using the same local transformation functions computed in the previous step. In our example we only collect piezometric head data although the method would be applicable if log-conductivity data were collected, too. Next, we apply equation (6) to update the state vector.
5. Back transform. The updated state vector is back transformed using the previously constructed transformation functions. Time advances one step, the updated state vector becomes the current vector and we loop back to the forecast step.

To sum up, the proposed method applies the EnKF always to a state vector all of which components follow a marginal Gaussian distribution. Furthermore, using the normal-score transformation we ensure that the prior non-Gaussian marginals of the model parameters are kept throughout, in our case, the prior bimodal pattern of the log-conductivities. We recognize that, in the proposed method, the model parameters and state variables are transformed into Gaussian space independently of each other, which yields a state vector with univariate Gaussian marginal distribution but not necessarily jointly multiGaussian.

3. Synthetic experiment

Aquifers of fluvial deposits are a typical example of geologic formations in which conductivities have a non-Gaussian distribution. Moreover, the hydraulic conductivities in such aquifers typically show a variation over several orders of magnitude. In this section, a synthetic bimodal aquifer composed of sand and clay is built to demonstrate the effectiveness of the proposed method. The NS-EnKF is compared with the standard EnKF in order to investigate whether and to what extent the normal-score transformation improves aquifer characterization for non-Gaussian cases. The influence of several parameters (i.e., number of realizations, number of hard data and magnitude of hydraulic conductivity variance) is investigated (see Table 1 for the description of the different scenarios).

Direct sampling, a multiple-point geostatistical simulation algorithm, is used to generate facies-distribution realizations [44]. Direct sampling is a pattern-based geostatistic simulation approach that generates realizations by borrowing structures from a training image. A training image generated by FLUVSIM [45] serves as a conceptual model of the bimodal aquifer composed of sand with high permeability and floodplain fine-grid deposits (e.g., clay) with low permeability (see Figure 2). Compared with traditional variogram-based two-point geostatistics, multiple-point geostatistics is well suited for handling curvilinear structures, which cannot be characterized well with a variogram.

The numerical experiment is carried out for a synthetic aquifer of 1 m thickness extending over a domain of $500 \text{ m} \times 400 \text{ m}$ discretized in 2D into 100 columns by 80 rows (i.e., square grid cells of 5 m). The reference facies field (Figure 3A) is generated by direct sampling based on the training image of Figure 2. Then, each facies is populated with log-conductivity values. These log-conductivity values are generated by GCOSIM3D [46], independently for each facies, according to stationary multiGaussian random functions with parameters given in Table 2. The reference $\ln K$ field is plotted in Figure 3B. The histogram of $\ln K$ (Figure 3E) shows a pronounced bimodality with a global mean of $-4.43 \ln(\text{cm/s})$ and variance of $1.72 (\ln(\text{cm/s}))^2$. Hydraulic conductivity measurements are taken from this reference field at the locations shown in Figure 3C. These data will be used as conditioning hard data.

The groundwater flow equation under confined conditions is solved for the reference hydraulic conductivity field assuming impermeable boundary conditions on the northern and southern boundaries, constant prescribed head on the western boundary equal to zero meters and prescribed specific discharge through the eastern boundary of -2.2 m/d (Figure 3D). The initial head is 0 m over the area of interest. The groundwater flow equation (Equation (1)) is solved by finite differences in both space and time using the simulator developed by Fu [47]. The simulation period of 300 days is discretized into 100 time steps, the duration of which follows a geometric series with ratio of 1.05. There are no internal sinks or sources in this example. Specific storage is assumed constant and equal to 0.003 m^{-1} . Piezometers cover the aquifer regularly (see Figure 3D), and the transient evolution of the piezometric heads in the reference field at these locations is sampled. Head observations from piezometers #1 to #32 will be used as observation data (conditioning data) during the update step of the NS-EnKF. The remaining piezometers will be used for validation purposes.

Similarly as we did to generate the reference realization, one thousand facies realizations are generated by direct sampling using the training image depicted in Figure 2. Then, in each realization, both facies are populated with $\ln K$ values generated by sequential Gaussian simulation. It is important to stress that both the facies realizations and the log-conductivity realizations are conditional to the hard data sampled from

the reference field and shown in Figure 3C. The geostatistical parameters used in the sequential simulation are the same as for the reference and given in Table 2. By generating the realizations in this way, neither statistical model uncertainty nor log-conductivity measurement uncertainty are considered in this example. For illustration purposes, a randomly selected realization is shown in Figure 4 together with the ensemble mean of the 1000 realizations.

Next, hydraulic head data assimilation by the NS-EnKF is performed as illustrated in section 2.3. The transient hydraulic heads observed in the reference field during the first 50 time steps (22.8 days) at the 32 monitoring piezometers are used to update the state vector after each time step. The observed heads during the remaining 50 time steps are used for model verification. Head observation error is characterized by a normal distribution with zero mean and a standard deviation of 0.01 m.

Recall that the term assimilation, commonly used in the filtering literature, is equivalent to the term conditioning used in the geostatistical literature.

4. Performance assessment measures

Ensemble means and ensemble standard deviations will be computed. The ensemble mean map should reproduce the main spatial patterns of hydraulic conductivity, as close as permitted by the conditioning data set, and it will be visually compared with the reference map to get a qualitative assessment. The standard deviation map will indicate, locally, the degree of uncertainty and, as will be seen, will serve to delineate the edges of the channels. For a quantitative assessment we compute a precision measure given by the Average Absolute Deviation ($AAD(x)_t$), and, taking advantage of knowing the reference, we also compute a bias measure, the Average Absolute Error ($AAE(x)_t$).

$$\begin{aligned} AAD(x)_t &= \frac{1}{N_{xy}} \frac{1}{N_r} \sum_{i=1}^{N_{xy}} \sum_{j=1}^{N_r} |x_{t,i,j} - \bar{x}_{t,i}| \\ AAE(x)_t &= \frac{1}{N_{xy}} \sum_{i=1}^{N_{xy}} |\bar{x}_{t,i} - x_{t,i}^{ref}| \end{aligned} \tag{10}$$

where $x_{t,i,j}$ is the state at time step t , node i and realization j , N_{xy} and N_r are the number of grid cells and realizations, respectively, $\bar{x}_{t,i}$ represents ensemble mean at time t and node i , and $x_{t,i}^{ref}$ is the reference state at time t and node i .

The histograms of hydraulic conductivity before and after data assimilation will be compared.

Connectivity of conductivity plays a critical role for transport: early arrival times of contaminants are

very sensitive to the existence of channels characterized by connected high-conductivity values, and late arrival times are influenced by the presence of barriers of connected low-conductivity values. A series of measures have been proposed to evaluate connectivity [48] and we adopt the connectivity function as defined by Stauffer and Aharony [49] because it is straightforward and easy to compute. This connectivity function gives the probability that two points are connected within the same facies by a continuous path. Thus, connectivity is a non-increasing function of Euclidean distance. In our case, in order to calculate the connectivity function, the continuous log-conductivity fields are converted to indicator fields with values 0 and 1:

$$I(x) = \begin{cases} 1, & \text{if } \ln K \geq -4 \\ 0, & \text{otherwise} \end{cases}$$

where the threshold of -4 splits, approximately, the log-conductivities into sand and clay (Figure 3E). We use the program CONNEC3D [50] to calculate the probability that two points with log-conductivities larger than -4 are connected following a continuous path with log-conductivities larger than -4 .

Hydraulic conductivity and piezometric head conditioning is not enough to ensure good transport predictions [51]. We analyze the transport prediction ability of the updated conductivity fields by performing a transport experiment—once piezometric head is at steady state—in which particles are uniformly released over a line source at $x = 17$ m. The integrated breakthrough curves (BTC) at a control plane positioned at $x = 450$ m are computed in the reference field and compared to the BTCs in the updated fields (see Figure 5 for the experiment setup). Average Absolute Error $AAE(T_\alpha)$ and Average Absolute Deviation $AAD(T_\alpha)$ are computed for different percentiles of the cumulative mass distribution.

$$\begin{aligned} AAE(T_\alpha) &= \frac{1}{N_r} \sum_{r=1}^{N_r} |T_{r,\alpha} - T_{ref,\alpha}| \\ AAD(T_\alpha) &= \frac{1}{N_r} \sum_{r=1}^{N_r} |T_{r,\alpha} - \bar{T}_\alpha| \end{aligned} \tag{11}$$

where $T_{r,\alpha}$ is the travel time for the r^{th} realization and the α percentile of the cumulative mass distribution (α values of 5%, 50% or 95% are analyzed), $T_{ref,\alpha}$ refers to the travel times corresponding to the reference field, and \bar{T}_α represents mean travel time over the ensemble.

5. Results and discussion

5.1. Standard EnKF versus NS-EnKF

In this subsection we will show that conditioning to piezometric head data by either the EnKF or the NS-EnKF is always beneficial, since either method produces a final ensemble of realizations that better characterizes the spatial variability of hydraulic conductivity than the initial ensemble of realizations. We will also show that the NS-EnKF outperforms the standard EnKF for the fluvial aquifer used here both with regard to the characterization of hydraulic conductivity, and in the capabilities of the updated hydraulic conductivity model to predict piezometric heads at unsampled locations, and mass transport through the aquifer.

5.1.1. Characterization of hydraulic conductivity

Regarding the characterization of hydraulic conductivities, we compare the initial (unconditional) ensemble of realizations and the final ensembles once the updating (conditioning) process is completed, focusing on: (i) the two performance measures defined earlier, $AAE(\ln K)$ and $AAD(\ln K)$, (ii) individual realizations from the ensembles, (iii) ensemble averages, (iv) log-conductivity histograms, (v) probability maps of the two facies, and (vi) connectivity in the horizontal direction.

Table 3 lists the values of $AAE(\ln K)$ and $AAD(\ln K)$; these values show, quantitatively, that conditioning to piezometric heads by either the standard EnKF or the NS-EnKF improves the characterization of the spatial distribution of hydraulic conductivities. A reduction on the AAE indicates a higher precision, while a reduction on the AAD implies a smaller bias. These values also show that the NS-EnKF performs better than the standard EnKF.

Figure 6 shows a randomly selected realization and the ensemble mean after conditioning to piezometric heads for the two filtering approaches; the reference field is also shown. We notice that the continuity of the sand channels in the NS-EnKF case is better reproduced than for the standard EnKF both for the single realization and for the ensemble mean. In any case, it is quite remarkable how well the main patterns of continuity on the reference field are captured by the ensemble Kalman filters.

A more pronounced difference between the final results obtained by the two filters is observed when analyzing the log-conductivity histograms. Figure 7 displays the histogram of the initial ensemble of log-conductivity (recall that both filters start with the same ensemble of initial realizations conditioned to the hard log-conductivity data) and the histograms of the final ensembles. The initial ensemble pattern is maintained by the NS-EnKF as expected, since, this is a built-in feature of the new algorithm. However, the

standard EnKF produces a histogram that gets close to Gaussian: the two initial modes are smeared out and the final range is increased.

Another way to look at how the evolution of the histograms impacts the characterization of the spatial distribution of the hydraulic conductivities is to retrieve the facies distribution by using the value of $-4 \ln(\text{cm/s})$ as a cut-off to identify the sand and clay facies. By transforming all realizations into 0's (if $\ln K$ is less than -4) and 1's (if $\ln K$ is greater than -4), and then computing the ensemble average, the resulting average maps would represent the probability that at each location the log-conductivity is larger than -4 , which, in this case, could be interpreted as that the facies is sand. Figure 8 shows the probability maps obtained in the initial ensemble, and on the final ensembles. In the initial ensemble, the hard log-conductivity data are enough to allow distinguish the location of the sand channels, but with still high uncertainty on their width. In the final ensembles, all channels are better delimited, with a very small area with probabilities away from 0 or 1 in the results of the NS-EnKF as compared with the probabilities obtained by the standard EnKF.

In a fluvial aquifer the connectivity induced by the sand channels is quite important in terms of transport predictions. Figure 9 shows connectivity as a function of horizontal separation, for each one of the realizations in the initial ensemble and in the two final ensembles. Also the connectivity of the ensemble averages and of the reference field are shown. Connectivity is computed on the indicator maps obtained above and measures the probability that any two points, both of them with log-conductivities above $-4 \ln(\text{cm/s})$, and which are a certain horizontal distance apart, are connected by a continuous path of log-conductivities above $-4 \ln(\text{cm/s})$. Looking only at the ensemble averages, we can conclude that the standard EnKF tends to produce realizations which are, on average, less connected, for all distances, than the reference field, with the NS-EnKF producing an average connectivity which is closer to the reference one than the standard EnKF; at the same time, the high connectivity values observed in the initial ensemble average for short distances is lost after conditioning to piezometric heads by any of the filters. Looking at the individual realizations, we can notice that the spread of the curves is progressively reduced from the initial ensemble to the ensemble obtained by the standard EnKF with the minimum spread for the ensemble obtained by the NS-EnKF. It remains to analyze the apparent splitting of the NS-EnKF ensemble of connectivity curves into two subsets, one set with curves about that of the reference field, and another one with curves higher than the reference.

In all the aspects analyzed, the NS-EnKF outperforms the standard EnKF with respect to the characterization of the spatial variability of log-conductivities.

5.1.2. Prediction ability of the updated model

Regarding the prediction capabilities of the updated models, we will analyze piezometric head predictions at both conditioning piezometers and control piezometers, we will also analyze the solute predictions for the mass transport experiment described earlier.

Figure 10 shows the hydraulic head evolution with time at two of the observation wells. Results are shown for the initial ensemble of realizations (conditioned only on log-conductivity) and for Scenarios 1 and 2 (standard EnKF and NS-EnKF) after conditioning to 32 piezometers up to the 50th time step (22.8 days). Head uncertainty is significantly reduced after data assimilation. Conditioned piezometric heads, up to the 50th time step, are almost perfectly reproduced in both scenarios. However, piezometric head prediction after the 50th time step (when observation data are not used for conditioning anymore) exhibits lower uncertainty for the NS-EnKF than for the standard EnKF. We have also checked the head evolution at 8 control piezometers, two of which are shown in Figure 11. Like at the observation locations, piezometric head uncertainty is reduced, after data assimilation, in both scenarios, and the NS-EnKF gives more precise predictions than the standard EnKF.

The updated $\ln K$ fields are further evaluated by using them as input for the solute transport prediction exercise described at the end of Section 4. The random walk particle tracking program RW3D [52, 53, 54] is used to solve Equation (2). For the sake of simplicity, only advection is considered in the experiment. Table 3 shows the values of $AAE(T_\alpha)$ and $AAD(T_\alpha)$ for the 5th, 50th and 95th percentiles of the breakthrough curves. We can see that, when using the NS-EnKF, both the bias and uncertainty are substantially reduced with respect to the corresponding values computed on the initial ensemble. The standard EnKF also improves bias and uncertainty compared with the initial ensemble, although this improvement is less than that of the NS-EnKF. Figure 12 summarizes the breakthrough curves for the initial ensemble of realizations (conditioned only on log-conductivity) and for Scenarios 1 and 2 (standard EnKF and NS-EnKF) after data assimilation. The NS-EnKF gives better transport prediction than the standard EnKF, the breakthrough curve uncertainty, as measured by the width of the confidence interval given by the 5th and 95th percentiles, is reduced significantly, and the median BTC is almost identical to the reference one.

In summary, with respect to prediction of flow and transport, the standard EnKF is able to predict head evolution properly, and makes a reasonable prediction of solute transport, but cannot match the NS-EnKF, which performs better in both cases.

5.2. Parameter sensitivity on NS-EnKF

From the previous section we can conclude that the proposed approach, the NS-EnKF, is better than the standard EnKF. Next, we will analyze the sensitivity of the NS-EnKF to different parameters, i.e., time span of data assimilation, number of realizations, number of conditioning hard data and log-conductivity variance.

5.2.1. Effect of time span of head data assimilation

In this section we will discuss the impact of the duration of the piezometric head data assimilation period on the spatial characterization of log-conductivity. Figures 13 and 14 show the ensemble mean and standard deviation of log-conductivity for the initial ensemble and after assimilating head data up to time steps 1, 10 and 50 (that is, 0.11, 1.45 and 22.8 days).

From the analysis of this figure, we note that just a single time step is needed to obtain an ensemble mean that displays channels similar to those in the reference field. Then, as more observation heads are assimilated, the ensemble mean tends to resemble better the reference; however, the ensemble mean does not change as much between the 10th and the 50th time steps as it did between the 1st and the 10th.

The length of the assimilation period has a larger impact in the ensemble standard deviation maps. These maps show how the uncertainty on log-conductivity becomes smaller as the time span of head data assimilation increases. As mentioned above, the ensemble mean maps change little between steps 10 and 50, but the standard deviation maps do change substantially, with an overall reduction of its values everywhere and with the highlight of the most likely locations for the channel borders given by the highest standard deviation values. In these maps, we can also observe that the standard deviation is zero at log-conductivity conditioning data locations, implying that $\ln K$ are honored in all realizations.

We conclude that, for the specific flow setup discussed in this paper, a short span of piezometric head data is enough to obtain an ensemble mean capturing the main patterns of variability of log-conductivity, both in terms of facies delineation and of within-facies log-conductivity spatial variability; however, the ensemble of final realizations gets more precise when a longer time span is used for conditioning, resulting in an ensemble standard deviation map that could be used to draw the borders of the sand/clay interface.

5.2.2. Effect of size of the ensemble

All results shown until now were for an ensemble of 1000 realizations. To evaluate the impact of ensemble size, we performed an analysis with an ensemble of 200 realizations, referred to as Scenario 3 in Table 1. The ensemble log-conductivity mean, together with a few realizations before and after assimilating the

piezometric head data for the first 50 time steps are displayed in Figure 15. For this number of realizations, and for our example, which has a relatively high log-conductivity variance, the problem of filter inbreeding [26] is apparent. The final updated realizations are almost identical to each other, and virtually equal to the ensemble mean after data assimilation. The average absolute deviation $AAD(\ln K)$ is only 0.02 (see Table 4, Scenario 3), which represents an unreasonable reduction of uncertainty. A series of algorithms have been proposed to cope with the problem of filter degeneration caused by a small ensemble size [55, 56, 26, 57]. We have not investigated how these algorithms would perform in conjunction with the NS-EnKF, and we conclude that the ensemble size has to be carefully chosen, particularly for high-variance log-conductivity fields.

5.2.3. Effect of number of hard conditioning data

Maintaining an ensemble size of 1000 realizations, we reduce the number of $\ln K$ conductivity conditioning hard data. It could be argued that the good results observed in the previous sections were due to the large number of conditioning data. This large number yields an initial ensemble average that already captures the main patterns of the reference $\ln K$ variability. For this reason we have reduced the number of conditioning data to 20 regularly-spaced samples taken from the reference field (Scenario 4, see Table 1). Then, we have generated a new set of initial $\ln K$ realizations by direct sampling using the same training image of Figure 2. The ensemble average of this new set of initial realizations can be seen in Figure 16. This mean map displays hardly any information about the channels of the reference, even though all 1000 realizations share the same 20 conditioning data.

It is remarkable to see how, assimilating the piezometric head data, the resulting ensemble mean map displays the most important features in the reference, although the channels are not as connected as in the reference. However, the relative improvement of this updated ensemble mean with respect to the initial one is very important. We attribute this behavior to the coupling of the EnKF and the normal-score transforms applied to all members of the state vector as described earlier.

Individual realizations, such as number 550th, resemble the main channel patterns of the reference; however, they are noisier than similar realizations generated with a larger number of $\ln K$ conditioning data, and, somehow, the channel connectivity is obscured. The metrics obtained for the flow and transport exercises are given in Table 4 (Scenario 4). We can see that all of them decrease their values with respect to the values computed in the initial ensemble, and that the final values, although larger, are comparable to those obtained when conditioning to 80 $\ln K$ data (Scenario 2 in Table 3).

We conclude that, overall, reducing the number of hard conditioning data is not an obstacle for log-

conductivity characterization by the NS-EnKF.

5.2.4. Effect of $\ln K$ variance

Our final analysis focuses on the impact of a very high variance of $\ln K$ on the performance of the NS-EnKF (Scenario 5, see Table 1). To keep the rest of the parameters as similar as possible to those used in Scenario 2, we have simply scaled the reference realization, as well as the 1000 initial realizations for Scenario 2, so that both reference and initial realizations have a variance of 9.92, while keeping the same mean and the same spatial patterns. Then, we have run the same flow exercise in the new reference field, and we have sampled the same piezometers for their usage as conditioning data by the NS-EnKF. Figure 17 shows the reference field with its histogram, initial realization number 550, the initial ensemble mean, the updated realization number 550 after assimilating the piezometric head data up to time step 50, and the updated ensemble $\ln K$ mean. We can see that even for this very large variance, the NS-EnKF is able to generate realizations which display the bimodal characteristic of the reference realization with, approximately, the same distribution of channels. We attribute the good behavior of the method to the normal-score transformation and to the large conductivity contrast that there is between sand and clay. The measures quantifying bias and uncertainty of log-conductivity characterization and flow and transport predictions are shown in Table 4 (Scenario 5). We can see that the biases and uncertainties are significantly reduced; however, we notice that the final bias and uncertainty values for the travel times after the NS-EnKF are still very high, which can be attributed to the extreme range of log-conductivities. This large range of variability induces zones of very low K and of very high K in all realizations; a small deviation of the proportions of the very low (or very high) log-conductivities with respect to the reference field can produce a large deviation on the travel times. The improvement of transport prediction is also well noticed when analyzing the BTCs (see figure 18).

We conclude that the NS-EnKF can handle high variance cases, particularly for bimodal cases with contrasting facies.

5.3. Discussion

We have demonstrated that the proposed NS-EnKF works on the fluvial aquifer used as a reference in this synthetic exercise. We have also shown that the method is quite robust for scenarios under conditions less favorable than the ones of the reference scenario (Scenario 2). The starting point is the less amenable for a Kalman filter: the state transition equation is a complex ear function, and both log-conductivities and piezometric heads can be statistically characterized by a multivariate distribution which is far from the

multiGaussian one. Kalman filter performance is the best for linear state transition equations and for multi-Gaussian random functions. We have chosen to transform both parameters and state variables (all collected in a joint state vector) into new ones which follow Gaussian marginal distributions. This transformation does not imply that the bivariate, trivariate and higher-order distributions of the transformed random function are multiGaussian. We have also chosen the ensemble Kalman filter to condition to piezometric head data, because, out of the different Kalman filters, it is the most robust against earities of the state transition function.

The main conclusion that we derive from this work is that, in a similar way to the EnKF is robust to earities; the marginal transformation into Gaussian deviates of conductivities and piezometric heads is robust to non-multiGaussianity. It could be argued that the results are so good because the number of hard conditioning data is large. There are two replies to this argument: (i) the number of hard conditioning data is the same for the standard EnKF as for the NS-EnKF in the examples, yet NS-EnKF always performs better, and (ii) one of the scenarios analyzed reducing substantially the number of conditioning data yields similar results. In addition, we have also performed an analysis with no hard data, which is included in a paper currently under review [58], and the results are remarkably good; not as good as we present here, but good enough to warrant that the benefits of the NS-EnKF are worth considering.

6. Conclusions

The EnKF has gained popularity in diverse disciplines because of its computational efficiency, flexibility and capacity for uncertainty assessment. However, although the EnKF is relatively robust for nonlinear model dynamics, it performs not optimally for non-multiGaussian parameter distributions. In this work, a new approach (NS-EnKF) is proposed that always works on state variables with marginal Gaussian distributions. Gaussianity is achieved by applying a normal-score transform to all variables, prior to performing the updating step in Kalman filtering. We recognize that the normal-score transform of the state variables renders them just univariate Gaussian, and that the multivariate distribution may remain far from the optimal multiGaussian distribution; yet, the new state variables are closer to multiGaussian than prior to the transformation, and, the histograms of the state variables (conductivity and piezometric head in our case) are preserved throughout the entire process. We have demonstrated how to implement the NS-EnKF to generate log-conductivity realizations which are conditional to log-conductivity data and also to transient piezometric head data. The method not only is able to honor the conditioning data, but it preserves the bimodal characteristics of the underlying reference field, performs well in the context of hydraulic head predictions

outside the conditioning period, and also in the context of mass transport prediction. We have also analyzed the impact in the performance of the NS-EnKF of the following parameters: the length of the piezometric head sampling period, the number of conditioning log-conductivity data, the number of realizations in the ensemble and the variance of log-conductivity.

The following conclusions can be drawn from the simulation experiments:

- The proposed method, the NS-EnKF, gives better results than the standard EnKF in terms of reproducing the bimodal $\ln K$ histogram, the connectivity of the $\ln K$ fields and in terms of predicting hydraulic heads and conservative transport. Both methods are capable of generating $\ln K$ realizations which honor the conditioning data, namely log-conductivity and transient head data. On the contrary, the standard EnKF cannot preserve the bimodal histogram of $\ln K$, underestimates connectivity and gives less precise and less accurate transport predictions than the NS-EnKF for the transport exercise discussed.
- The hydraulic head measurements during the early stage of the assimilation process by the NS-EnKF are enough to capture the main patterns of variability of $\ln K$; however, by extending the assimilation time span, the uncertainty on $\ln K$ is largely decreased, and the ensemble standard deviation map serves to delineate the interfaces between sand and clay.
- In this experiment good results were obtained with 1000 realizations, but results deteriorate for 200 realizations. Careful choice of the ensemble size is important in the application of the NS-EnKF.
- A large number of log-conductivity conditioning data helps in delineating the spatial variability patterns before the hydraulic head is assimilated by the NS-EnKF; however, we found that even when the number of hard data is smaller, piezometric head observations are enough to capture the complex patterns exhibited in this particular example.
- The NS-EnKF also performs well in strongly heterogeneous formations with $\ln K$ variance close to 10.

The univariate Gaussian distribution is considered in the present method and bivariate or even multivariate Gaussian distribution will be investigated in the future.

Acknowledgements. The authors gratefully acknowledge the financial support by ENRESA (project 0079000029). The financial aid from the China Scholarship Council (CSC) to the first author is

appreciated and extra travel grants from the Ministry of Education (Spain) awarded to the first and fourth authors are also acknowledged.

Appendix A: Normal-score transform

Normal-score transform is a tool relating any distribution function $F(x)$ to a standard Gaussian function $G(y)$ [59]. It is described in Figure 19, where the processes of normal-score transforming variable x onto a Gaussian deviate y , and its inverse, are depicted. In other words, the two random variables x and y are linked through their cumulative probability distributions:

$$\begin{aligned} F(x) &= G(y) \\ y &= G^{-1}[F(x)] \\ x &= F^{-1}[G(y)] \end{aligned} \tag{A-1}$$

where y follows a standard normal distribution with zero mean and unit variance and is the normal-score transform of x . More detailed information can be found in Goovaerts [59] and Deutsch and Journel [43].

For each one of the components of the state vector (hydraulic heads or hydraulic log-conductivities), the ensemble values are used to estimate a non-parametric cumulative distribution function $F(x)$. Estimation of $F(x)$ amounts to sort all values for a given state vector component, and assign to each value a cumulative probability equal to $i/(N_r + 1)$, where i is the ordinal position after sorting, and N_r is the number of realizations in the ensemble. These distribution functions, which have to be computed for each location and each time step, are used to normal-score transform the components of the state vector, then the NS-EnKF is performed, and then a back transform is done to retrieve the state vector in its original space to feed it to the transfer function and perform the next forecast step.

In the back transform process, since $F(x)$ is non-parametrically defined, there is a need to establish rules of interpolation to retrieve the state vector value x given a cumulative probability $F(x)$, which, most likely, will not correspond to any of the cumulative probabilities estimated from the values of the ensemble. Besides, since the Gaussian deviates could result in cumulative probabilities that are outside of the range of cumulative probabilities estimated from the ensemble, there is a need to explicitly state the absolute minimum and maximum of each state vector component, and the corresponding interpolation rules for the upper and lower tails of the cumulative distribution function. Different interpolation schemes are possible and the reader is referred to Deutsch and Journel [43] for the details.

Considering the way the cumulative distribution function $F(x)$ is estimated, the backtransform $F^{-1} \cdot G(y)$ always exists; however, the normal-score transform $G^{-1} \cdot F(x)$ is undefined at conditioning locations for which all ensemble values are identical, and thus, $F(x)$ is a step function. At those locations, we have chosen to assign a normal score by taking the average of the normal scores of the state variables in the neighboring cells.

- [1] Yeh, W.W.. Review of parameter identification procedures in groundwater hydrology: The inverse problem. *Water Resour Res* 1986;22(2):95–108.
- [2] McLaughlin, D., Townley, L.R.. A reassessment of the groundwater inverse problem. *Water Resour Res* 1996;32(5):1131–1161.
- [3] De Marsily, G., Delhomme, J.P., Coudrain-Ribstein, A., Lavenue, A.M.. Four decades of inverse problems in hydrogeology. Theory, modeling, and field investigation in hydrogeology: a special volume in honor of Shlomo P Neumans 60th birthday 2000;348:1–17.
- [4] Carrera, J., Alcolea, A., Medina, A., Hidalgo, J., Slooten, L.J.. Inverse problem in hydrogeology. *Hydrogeol J* 2005;13(1):206–222. doi:10.1007/s10040-004-0404-7.
- [5] Hendricks Franssen, H.J., Alcolea, A., Riva, M., Bakr, M., van der Wiel, N., Stauffer, F., et al. A comparison of seven methods for the inverse modelling of groundwater flow. Application to the characterisation of well catchments. *Adv Water Resour* 2009;32(6):851–872.
- [6] Oliver, D., Chen, Y.. Recent progress on reservoir history matching: a review. *Computat Geosci* 2010;doi: 10.1007/s10596-010-9194-2.
- [7] Gómez-Hernández, J.J., Wen, X.. Probabilistic assessment of travel times in groundwater modeling. *Stoch Hydrol Hydraul* 1994;8(1):19–55.
- [8] RamaRao, B., LaVenue, A., De Marsily, G., Marietta, M.. Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields, 1, theory and computational experiments. *Water Resour Res* 1995;31(3):475–493.
- [9] Alcolea, A., Carrera, J., Medina, A.. Inversion of heterogeneous parabolic-type equations using the pilot points method. *Int J Numer Meth Fl* 2006;51(9-10):963–980.
- [10] Gómez-Hernández, J.J., Sahuquillo, A., Capilla, J.E.. Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data—I. Theory. *J Hydrol* 1997;203(1-4):162–174.

- [11] Hu, L.. Gradual deformation and iterative calibration of gaussian-related stochastic models. *Math Geol* 2000;32(1):87–108.
- [12] Evensen, G.. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J Geophys Res* 1994;99(C5):10143–10162.
- [13] Burgers, G., Jan van Leeuwen, P., Evensen, G.. Analysis scheme in the ensemble Kalman filter. *Mon Weather Rev* 1998;126(6):1719–1724.
- [14] Kalman, R.E.. A new approach to linear filtering and prediction problems. *J Basic Eng-T ASME* 1960;82:35–45.
- [15] Leng, C., Yeh, H.. Aquifer parameter identification using the extended Kalman filter. *Water Resour Res* 2003;39(3). doi: 10.1029/2001WR000840.
- [16] Yeh, H., Huang, Y.. Parameter estimation for leaky aquifers using the extended Kalman filter, and considering model and data measurement uncertainties. *J Hydrol* 2005;302(1-4):28–45.
- [17] Evensen, G.. The Ensemble Kalman Filter: Theoretical formulation and practical implementation. *Ocean Dynam* 2003;53(4):343–367.
- [18] Bertino, L., Evensen, G., Wackernagel, H.. Sequential data assimilation techniques in oceanography. *Int stat Rev* 2003;71(2):223–241.
- [19] Houtekamer, P.L., Mitchell, H.L.. A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon Weather Rev* 2001;129:123–137.
- [20] Moradkhani, H., Sorooshian, S., Gupta, H.V., Houser, P.R.. Dual state-parameter estimation of hydrological models using ensemble Kalman filter. *Adv Water Resour* 2005;(28):135–147.
- [21] Wen, X., Chen, W.. Real-time reservoir model updating using ensemble Kalman filter with confirming option. *SPE J* 2006;11(4):431–442.
- [22] Chen, Y., Zhang, D.. Data assimilation for transient flow in geologic formations via ensemble Kalman filter. *Adv Water Resour* 2006;29:1107–1122.
- [23] Nowak, W.. Best unbiased ensemble linearization and the quasi-linear Kalman ensemble generator. *Water Resour Res* 2009;45. doi:10.1029/2008WR007328.

- [24] Li, L., Zhou, H., Hendricks Franssen, H.J., Gómez-Hernández, J.J.. Modeling transient groundwater flow by coupling ensemble kalman filtering and upscaling. submitted to Water Resour Res 2010;.
- [25] Hendricks Franssen, H.J., Kinzelbach, W.. Ensemble Kalman filtering versus sequential self-calibration for inverse modelling of dynamic groundwater flow systems. J Hydrol 2009;365(3-4):261–274.
- [26] Hendricks Franssen, H.J., Kinzelbach, W.. Real-time groundwater flow modeling with the Ensemble Kalman Filter: Joint estimation for states and parameters and the filter inbreeding problem. Water Resour Res 2008;44. doi: 10.1029/2007WR006505.
- [27] Liu, G., Chen, Y., Zhang, D.. Investigation of flow and transport processes at the MADE site using ensemble Kalman filter. Adv Water Resour 2008;31:975–986.
- [28] Wen, X.H., Chen, W.H.. Some practical issues on real-time reservoir model updating using ensemble Kalman filter. SPE J 2007; 12(2): 156–166.
- [29] Gu, Y., Oliver, D.. An iterative ensemble kalman filter for multiphase fluid flow data assimilation. SPE J 2007;12(4):438–446.
- [30] Arulampalam, M.S., Maskell, S., Gordon, N., Clapp, T.. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. IEEE T Signal Proces 2002;50(2):174–188.
- [31] Journel, A.G., Deutsch, C.V.. Entropy and spatial disorder. Math Geol 1993;25(3):329–355.
- [32] Gómez-Hernández, J.J., Wen, X.H.. To be or not to be multi-Gaussian? A reflection on stochastic hydrogeology. Adv Water Resour 1998;21(1):47–61.
- [33] Zinn, B., Harvey, C.. When good statistical models of aquifer heterogeneity go bad: A comparison of flow, dispersion, and mass transfer in connected and multivariate Gaussian hydraulic conductivity fields. Water Resour Res 2003;39(3). doi: 10.1029/2001WR001146.
- [34] Sun, A., Morris, A., Mohanty, S.. Sequential updating of multimodal hydrogeologic parameter fields using localization and clustering techniques. Water Resour Res 2009;45(7). doi: 10.1029/2008WR007443.
- [35] Chen, Y., Oliver, D.S., Zhang, D.. Data assimilation for nonlinear problems by ensemble Kalman filter with reparameterization. J Petrol Sci Eng 2009;66:1–14.

- [36] Schöniger, A., Nowak, W., Hendricks Franssen, H.J.. Parameter estimation by ensemble Kalman filters with transformed data: approach and application to hydraulic tomography. submitted to Water Resour Res 2010;.
- [37] Gu, Y., Oliver, D.S.. The ensemble Kalman filter for continuous updating of reservoir simulation models. *Journal of Energy Resources Technology* 2006;128(1):79–87.
- [38] Simon, E., Bertino, L.. Application of the Gaussian anamorphosis to assimilation in a 3-D coupled physical-ecosystem model of the North Atlantic with the EnKF: a twin experiment. *Ocean Sci* 2009;(5):495–510.
- [39] Béal, D., Brasseur, P., Brankart, J.M., Ourmières, Y., Verron, J.. Characterization of mixing errors in a coupled physical biogeochemical model of the North Atlantic: implications for nonlinear estimation using Gaussian anamorphosis. *Ocean Sci* 2010;6:247–262.
- [40] Bocquet, M., Pires, C.A., Wu, L.. Beyond gaussian statistical modeling in geophysical data assimilation. *Mon Weather Rev* 2010;138(8):2997–3023.
- [41] Bear, J.. *Dynamics of fluids in porous media*. New York, 764pp: American Elsevier Pub. Co.; 1972. ISBN 9780444001146.
- [42] Evensen, G.. *Data assimilation: The ensemble Kalman filter*. 279pp: Springer Verlag; 2007.
- [43] Deutsch, C.V., Journel, A.G.. *GSLIB, Geostatistical Software Library and User’s Guide*. New York, 384pp: Oxford University Press; second ed.; 1998.
- [44] Mariethoz, G., Renard, P., Straubhaar, J.. The direct sampling method to perform multiple-point geostatistical simulations. *Water Resour Res* 2010;46. doi: 10.1029/2008WR007621.
- [45] Deutsch, C., Tran, T.. FLUVSIM: a program for object-based stochastic modeling of fluvial depositional systems. *Comput Geosci* 2002;28:525–535.
- [46] Gómez-Hernández, J.J., Journel, A.G.. Joint sequential simulation of Multi-Gaussian fields. In: Soares, A., editor. *Geostatistics Tróia ’92*; vol. 1. Dordrecht: Kluwer Academic Publishers; 1993, p. 85–94.
- [47] Fu, J.. *A Markov chain Monte Carlo method for inverse stochastic modeling and uncertainty assessment*. Ph.D. thesis; Technical University of Valencia. 136pp; 2008.

- [48] Knudby, C., Carrera, J.. On the relationship between indicators of geostatistical, flow and transport connectivity. *Adv Water Resour* 2005;28(4):405–421.
- [49] Stauffer, D., Aharony, A.. *Introduction to percolation theory*. Taylor and Francis, London. 181pp; 1994.
- [50] Pardo-Igúzquiza, E., Dowd, P.A.. CONNEC3D: a computer program for connectivity analysis of 3D random set models. *Comput Geosci* 2003;29:775–785.
- [51] Meier, P.M., Medina, A., Carrera, J.. Geostatistical inversion of cross-hole pumping tests for identifying preferential flow channels within a shear zone. *Ground water* 2001;39(1):10–17.
- [52] Fernàndez-Garcia, D., Illangasekare, T., Rajaram, H.. 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 Resour Res* 2005;41(3). doi: 10.1029/2004WR003125.
- [53] Salamon, P., Fernàndez-Garcia, D., Gómez-Hernández, J.. A review and numerical assessment of the random walk particle tracking method. *J Contam Hydrol* 2006;87(3-4):277–305.
- [54] Li, L., Zhou, H., Gómez-Hernández, J.J.. Transport upscaling using multi-rate mass transfer in three-dimensional highly heterogeneous porous media. *Adv Water Resour* 2010;34(4):478–489.
- [55] Houtekamer, P., Mitchell, H.. Data assimilation using an ensemble Kalman filter technique. *Mon Weather Rev* 1998;126:796–811.
- [56] Hamill, T., Whitaker, J., Snyder, C.. Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon Weather Rev* 2001;129:2776–2790.
- [57] Chen, Y., Oliver, D.S.. Cross-covariances and localization for EnKF in multiphase flow data assimilation. *Computat Geosci* 2010;14(4):579–601.
- [58] Zhou, H., Li, L., Gómez-Hernández, J.J., Hendricks Franssen, H.J.. Pattern identification in bimodal aquifer with normal-score ensemble kalman filter. *Math Geosci* 2011;under review.
- [59] Goovaerts, P.. *Geostatistics for natural resources evaluation*. Oxford University Press, USA. 483pp; 1997.

Table 1: Scenario definition. Parameters that change from one scenario to another.

Scenario	Hard $\ln K$ data	$\ln K$ Variance	No. realizations	EnKF	NS-EnKF
1	80	1.72	1000	×	
2	80	1.72	1000		×
3	80	1.72	200		×
4	20	1.72	1000		×
5	80	9.92	1000		×

Table 2: Parameters defining the multiGaussian random function used to generate the log-conductivity within each of the two facies.

Facies	Variogram type	Mean ($\ln(\text{cm/s})$)	Standard deviation ($\ln(\text{cm/s})$)	λ_x^* (m)	λ_y^* (m)
clay (background)	exponential	-5.5	0.5	50	50
sand (channel)	exponential	-3.0	0.5	100	50

* ranges in the x and y directions

Table 3: Comparison of the initial ensemble (only conditioned to conductivity data), standard EnKF and NS-EnKF (Scenarios 1 and 2 in Table 1, which are conditioned to both conductivity and piezometric head data).

Metrics	Initial ensemble	Scenario 1	Scenario 2
AAE($\ln K$)	0.78	0.65	0.64
AAD($\ln K$)	0.75	0.52	0.40
AAE(h)	0.48	0.12	0.13
AAE($T_{5\%}$)	11.49	10.76	7.80
AAE($T_{50\%}$)	60.26	42.79	27.91
AAE($T_{95\%}$)	332.29	164.16	94.41
AAD($T_{5\%}$)	9.50	10.63	3.91
AAD($T_{50\%}$)	60.01	34.80	27.90
AAD($T_{95\%}$)	220.37	134.39	79.39

Table 4: Performance measures for Scenarios 3, 4 and 5.

Metrics	Initial ensemble	Scenario 3	Initial ensemble	Scenario 4	Initial ensemble	Scenario 5
AAE($\ln K$)	0.78	0.85	1.05	0.77	1.87	1.57
AAD($\ln K$)	0.76	0.02	1.00	0.53	1.81	1.00
AAE(h)	0.52	0.52	1.34	0.14	0.77	0.43
AAE($T_{5\%}$)	-	-	12.01	8.54	9.59	10.77
AAE($T_{50\%}$)	-	-	79.66	91.09	1986.06	1345.58
AAE($T_{95\%}$)	-	-	176.32	71.45	19313.82	4664.59
AAD($T_{5\%}$)	-	-	12.10	7.70	9.51	3.68
AAD($T_{50\%}$)	-	-	70.44	32.67	1668.17	831.16
AAD($T_{95\%}$)	-	-	158.70	71.20	13698.18	3134.55

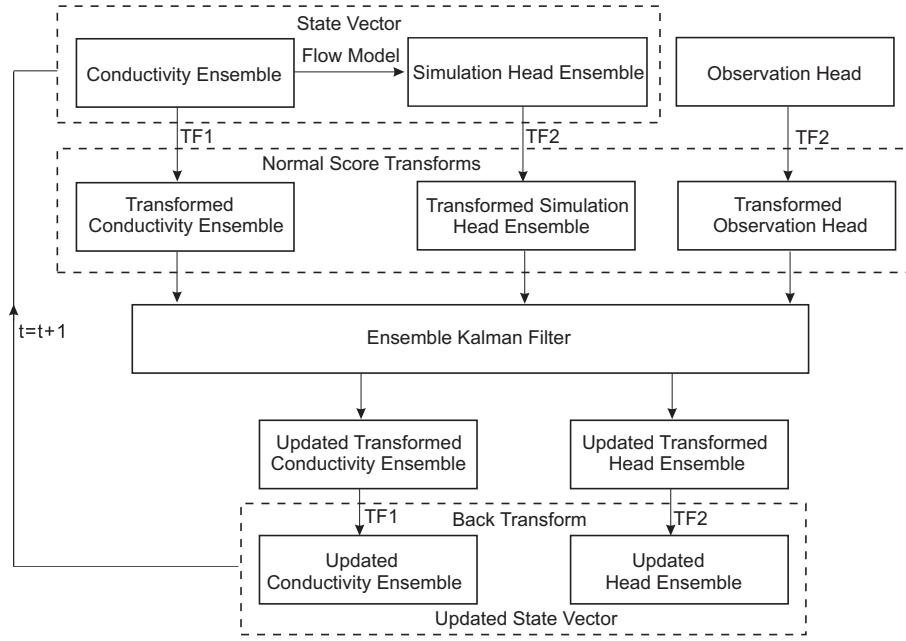


Figure 1: Flow chart of the NS-EnKF.

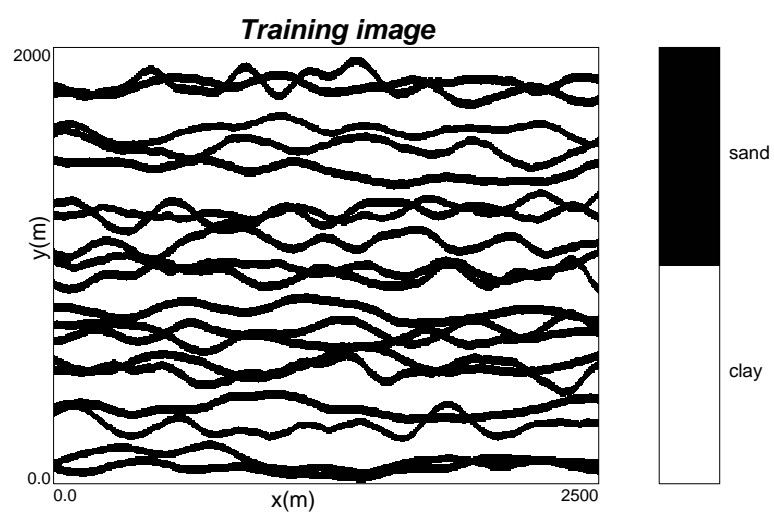


Figure 2: Training image used to generate the ensemble of binary facies realizations.

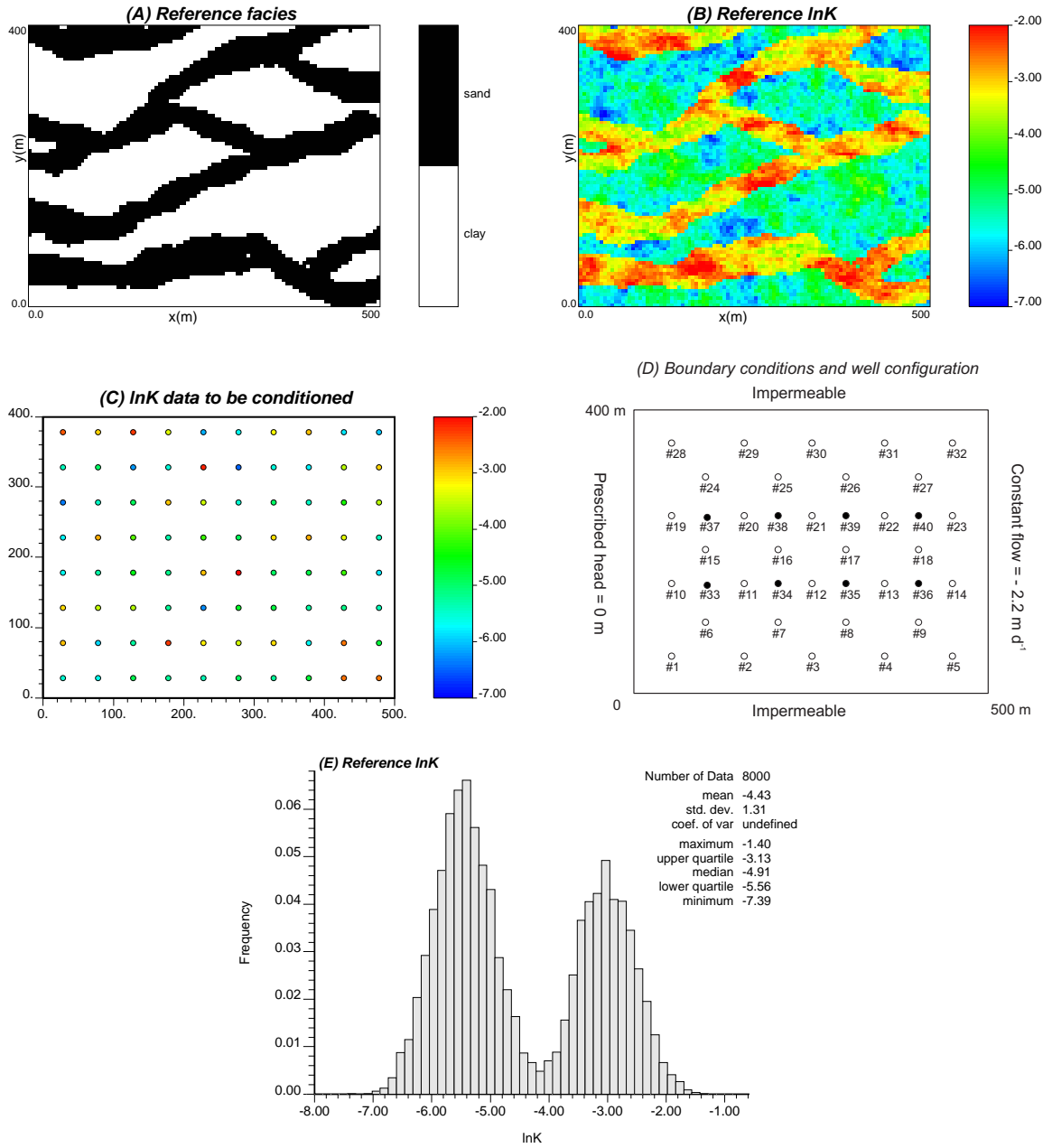


Figure 3: (A) Reference facies field; (B) reference log-conductivity field (ln(cm/s)); (C) hydraulic conductivity measurements used for conditioning; (D) flow boundary conditions and location of piezometers (open circles for observation, solid circles for prediction); (E) histogram of reference log-conductivity.

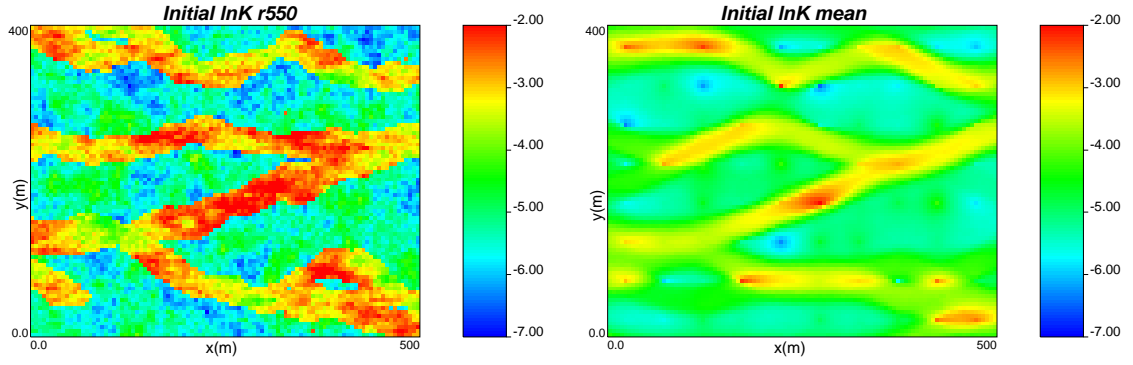


Figure 4: A randomly selected $\ln K$ realization (550th) from the initial ensemble and the ensemble mean of all 1000 initial realizations for Scenarios 1 and 2.

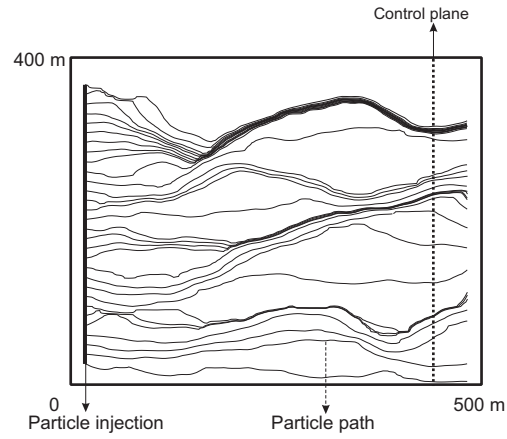


Figure 5: Configuration of the transport prediction experiment showing the line source where particles are injected and the control plane where breakthrough curves are measured (a few particle paths are shown for illustration).

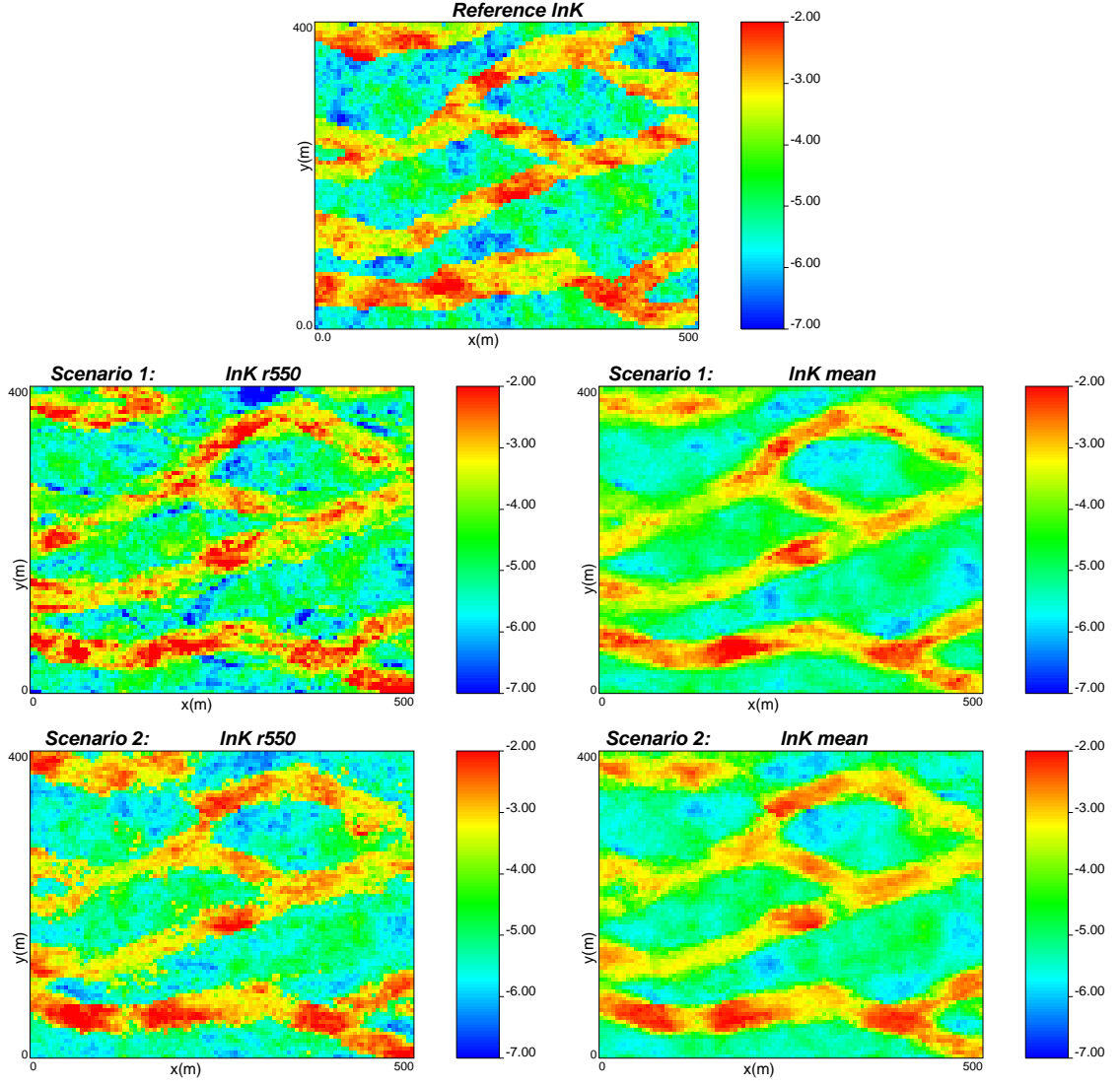


Figure 6: The reference field together with a randomly selected realization (550th) of $\ln K$ and the ensemble mean of $\ln K$ for Scenarios 1 and 2 (standard EnKF and NS-EnKF) after 50 times steps of piezometric head data assimilation.

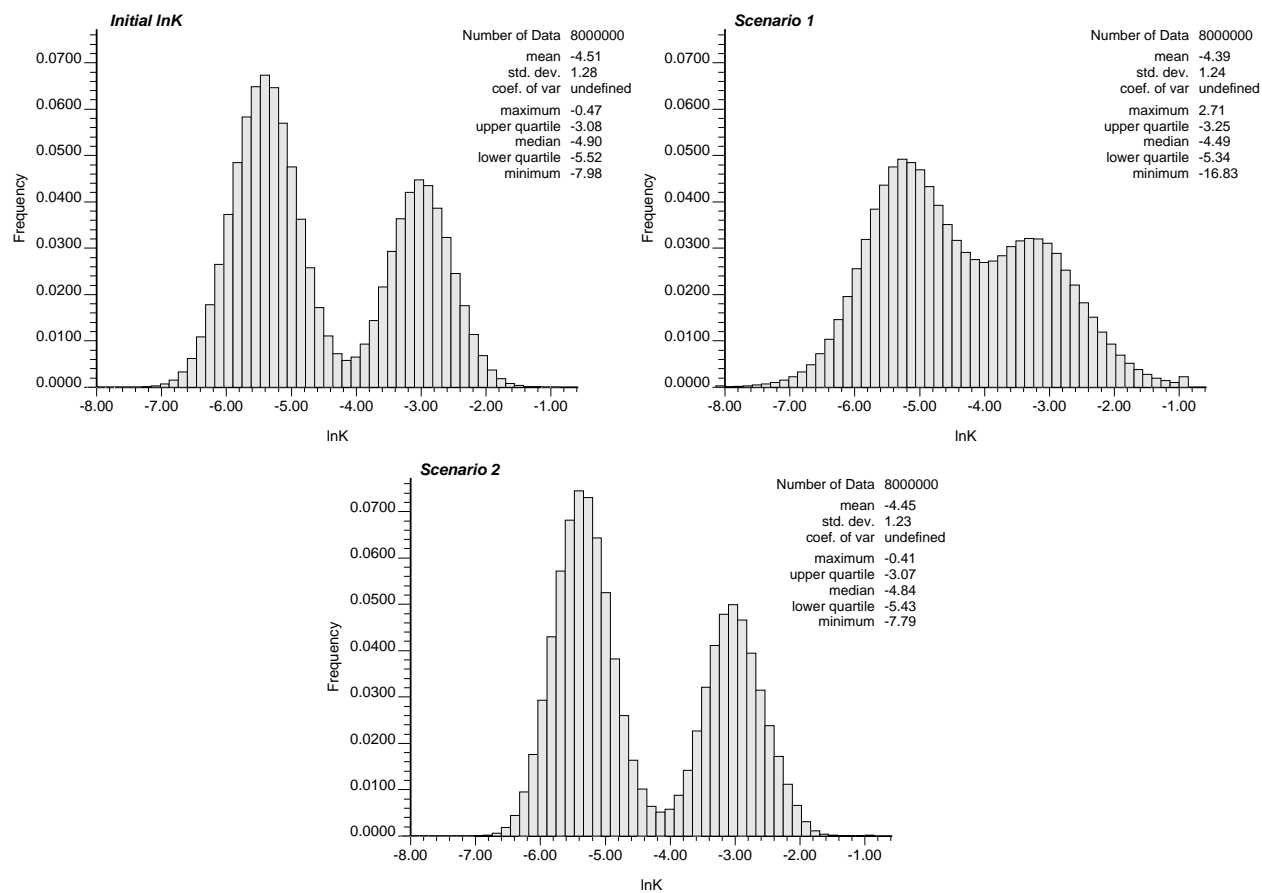


Figure 7: Log-conductivity histograms for the initial ensemble of realizations and for Scenarios 1 and 2 (standard EnKF and NS-EnKF) after data assimilation.

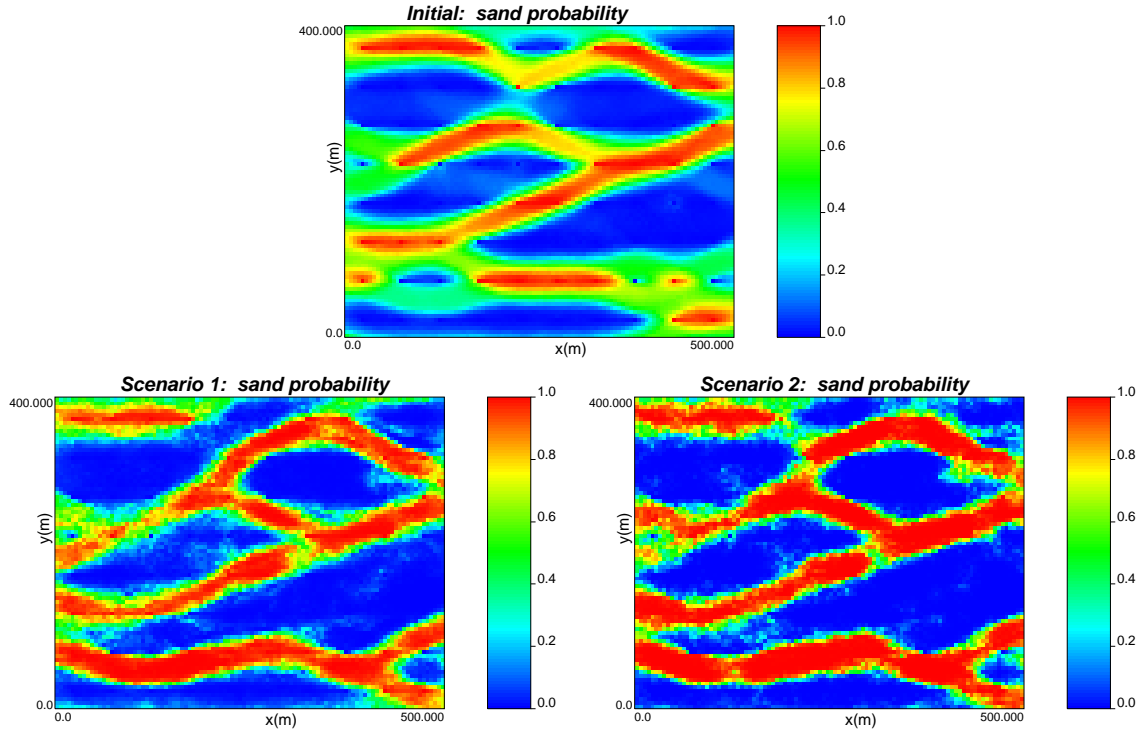


Figure 8: Probability that log-conductivity is larger than $-4 \ln(\text{cm/s})$ for the initial ensemble of realizations, and for the ensembles corresponding to Scenarios 1 and 2 (standard EnKF and NS-EnKF) after 50 times steps of piezometric head data assimilation.

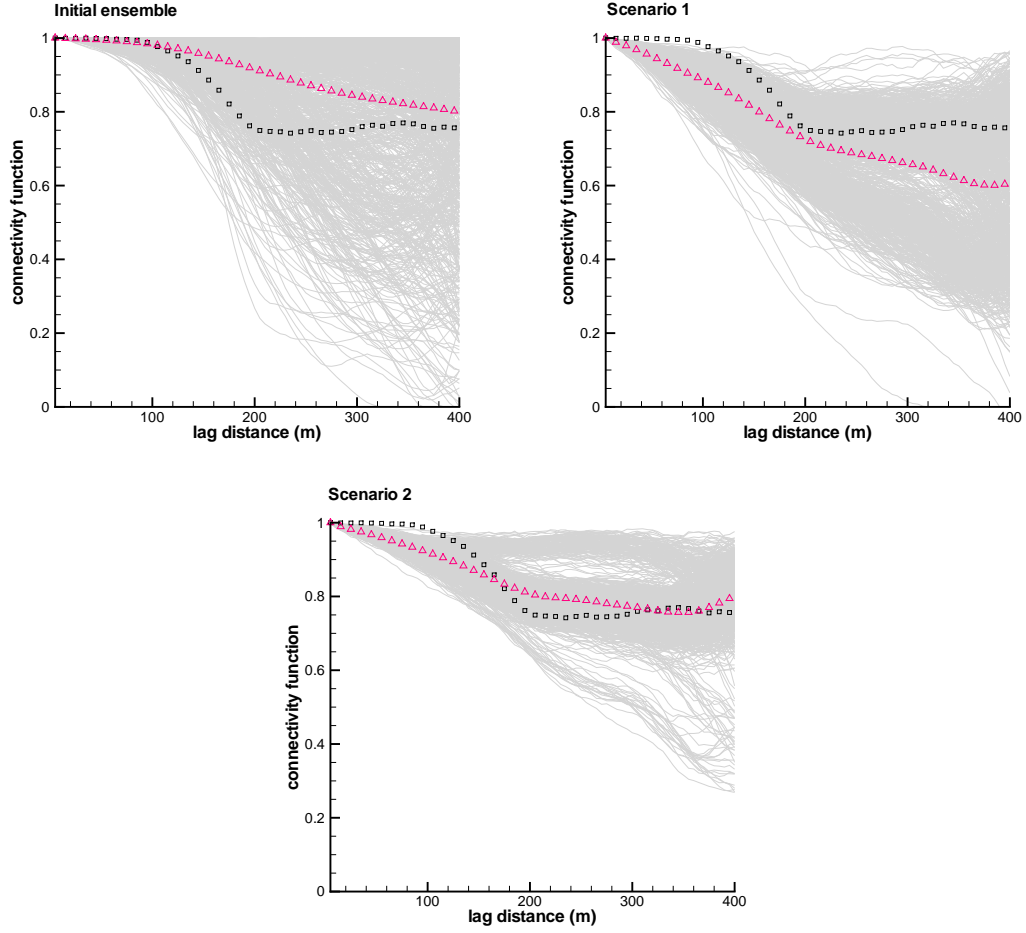


Figure 9: Connectivity, as a function of the separation distance, measuring the probability that two points aligned in the x direction are connected by a continuous path of log-conductivities larger than $-4 \ln(\text{cm/s})$. Gray curves correspond to individual realizations, their mean is given by the triangles (red) and the squares (black) correspond to the reference. Results are shown for the initial ensemble of realizations (conditioned on log-conductivity) and for Scenarios 1 and 2 (standard EnKF and NS-EnKF) after data assimilation.

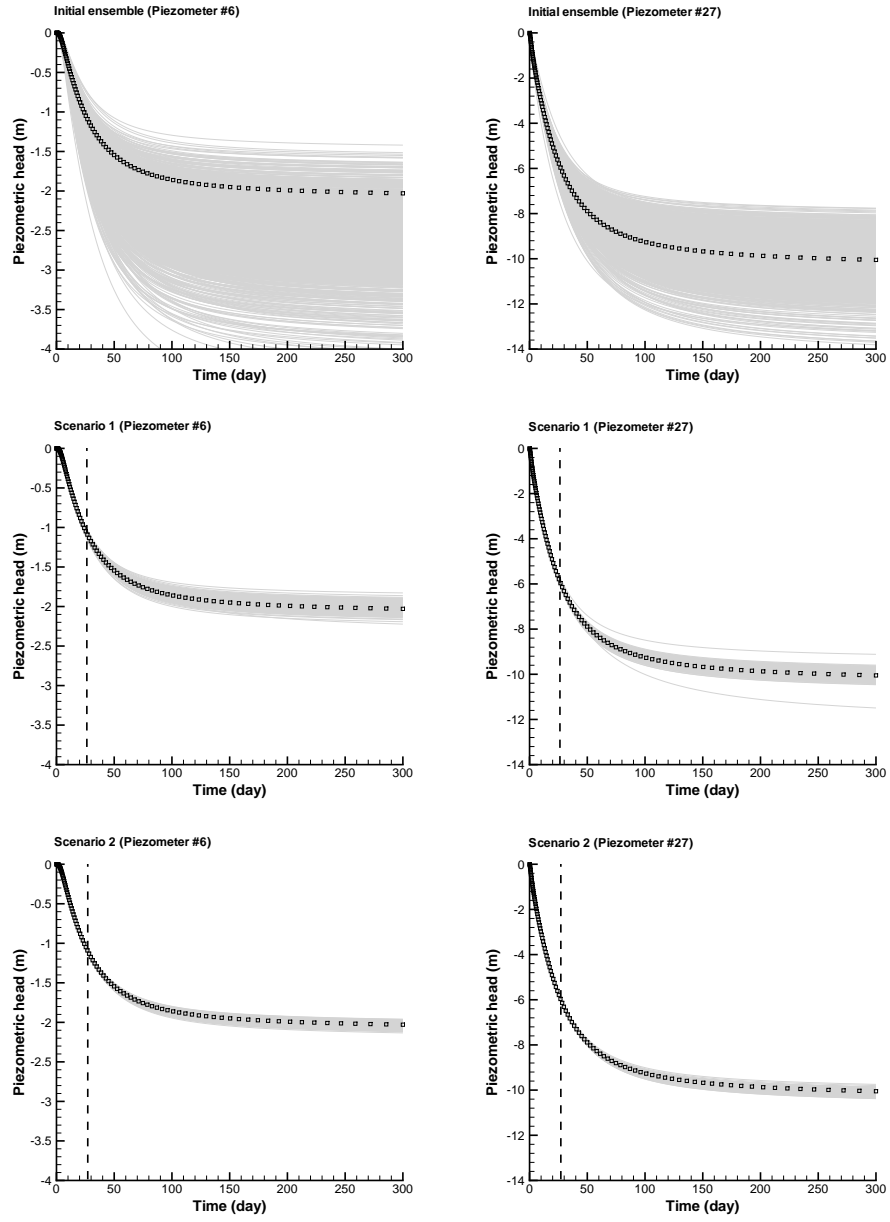


Figure 10: Piezometric head evolution at two of the thirty-two conditioning piezometers. Results are shown for the initial ensemble of realizations (conditioned only on log-conductivity) and for Scenarios 1 and 2 (standard EnKF and NS-EnKF, conditioned to transient piezometric heads up to 22.8 days). The vertical dashed lines indicate the conditioning period. Squares represent the piezometric head in the reference field.

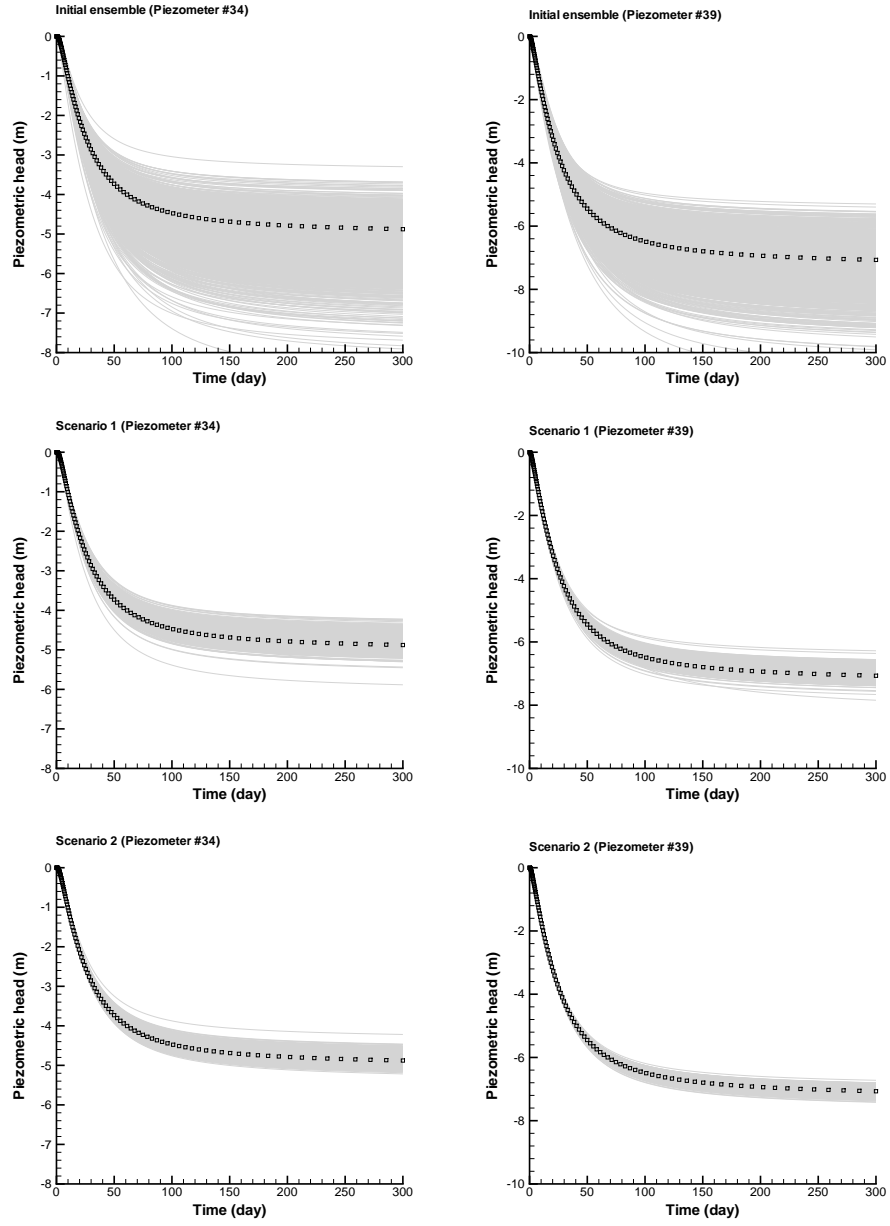


Figure 11: Piezometric head evolution at two of the eight validation piezometers. Results are shown for the initial ensemble of realizations (conditioned only on log-conductivity) and for Scenarios 1 and 2 (standard EnKF and NS-EnKF) after data assimilation. Squares represent the piezometric head in the reference field.

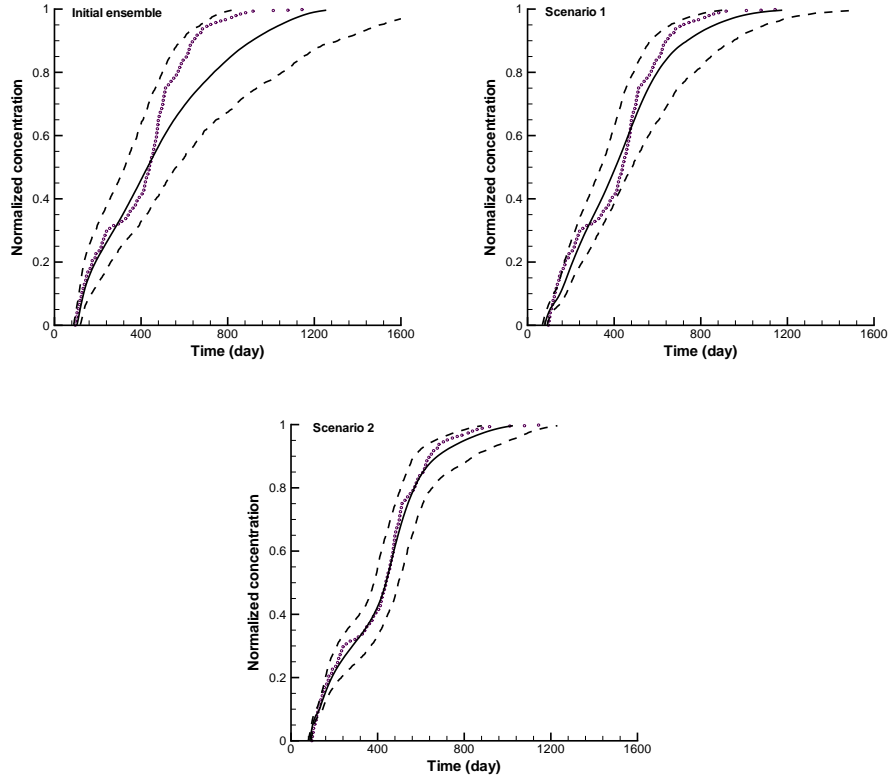


Figure 12: Summary of the breakthrough curves. The 5th percentile, the median, and the 95th percentile of the travel times are computed as a function of normalized concentration. Dashed lines correspond to the 5th and 95th percentiles, the solid line corresponds to the median, and the dotted line is the breakthrough curve in the reference. Results are shown for the initial ensemble of realizations (conditioned only on log-conductivity) and for Scenarios 1 and 2 (standard EnKF and NS-EnKF) after data assimilation.

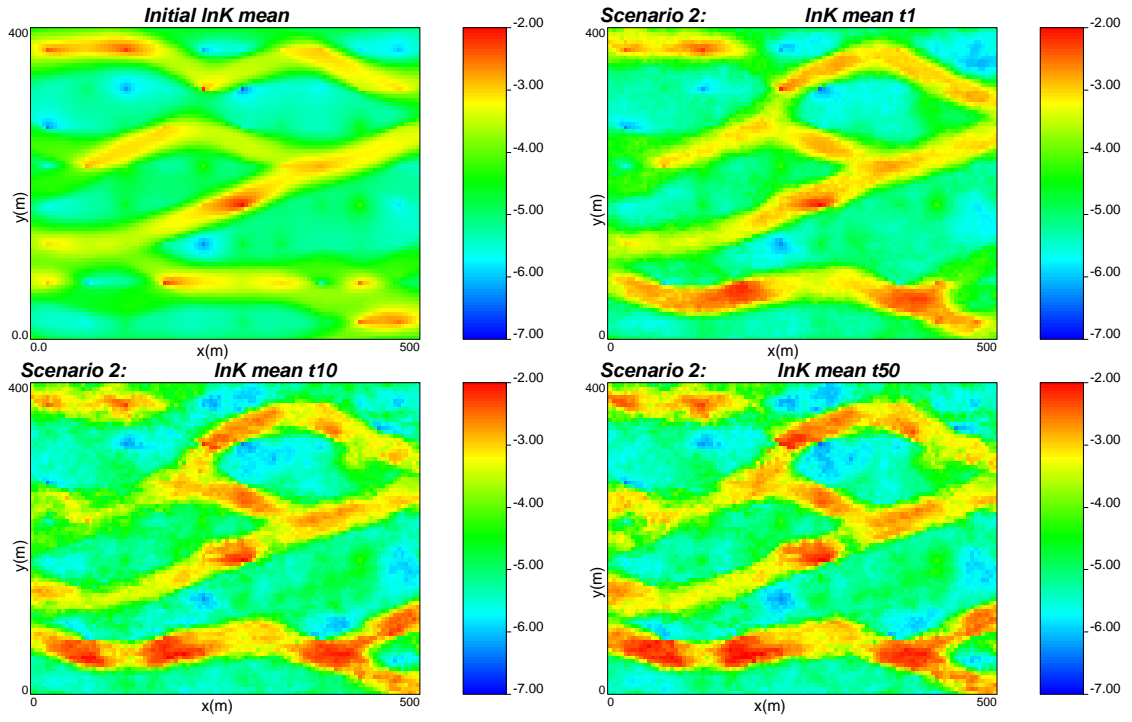


Figure 13: Evolution of log-conductivity ensemble mean as data are assimilated for time steps 1, 10 and 50 (Scenario 2).

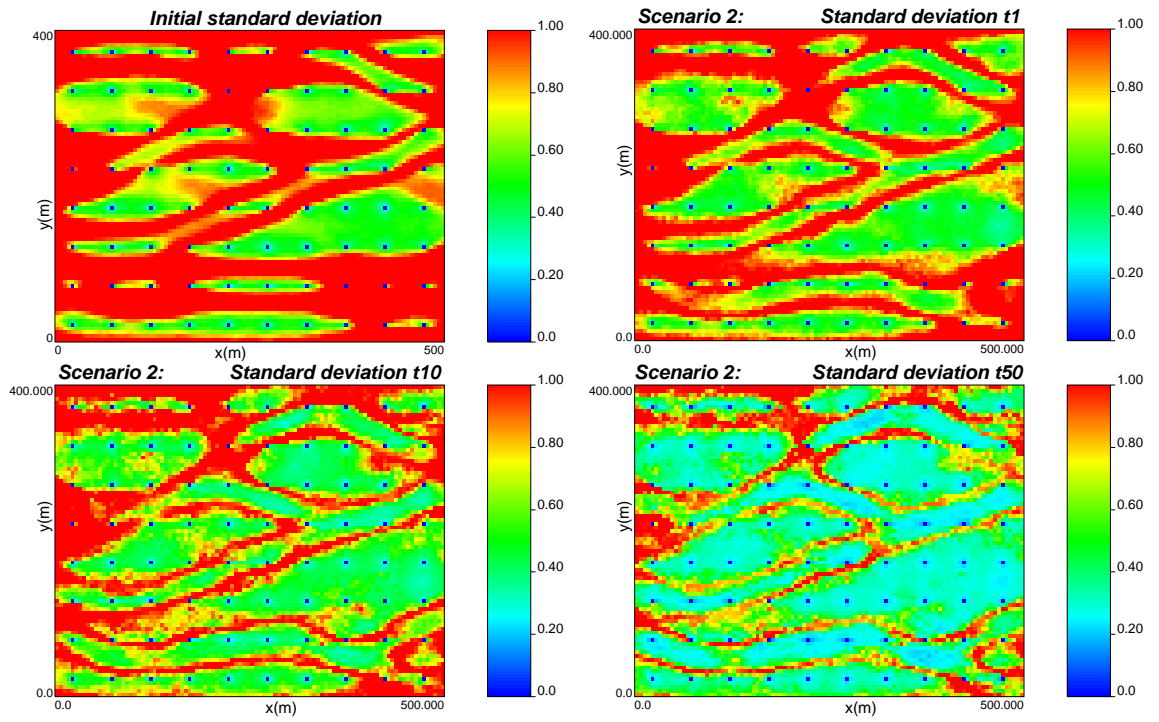


Figure 14: Evolution of log-conductivity ensemble standard deviation as data are assimilated for time steps 1, 10 and 50 (Scenario 2).

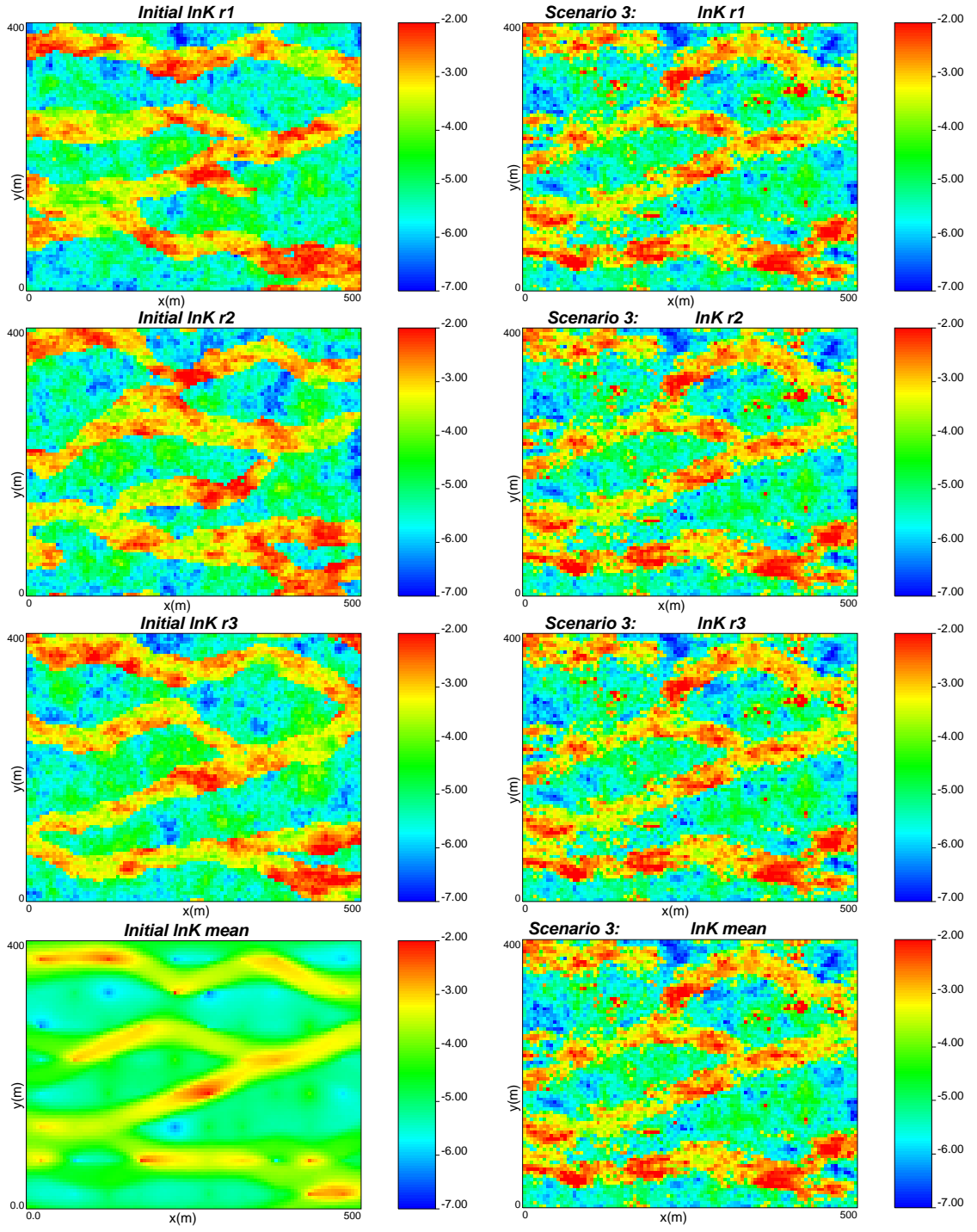


Figure 15: The first three realizations and the ensemble mean before (left column) and after (right column) 50 times steps of piezometric head data assimilation using the NS-EnKF with an ensemble size of 200 realizations (Scenario 3). Notice how filter inbreeding effects the final results making all the realizations virtually the same.

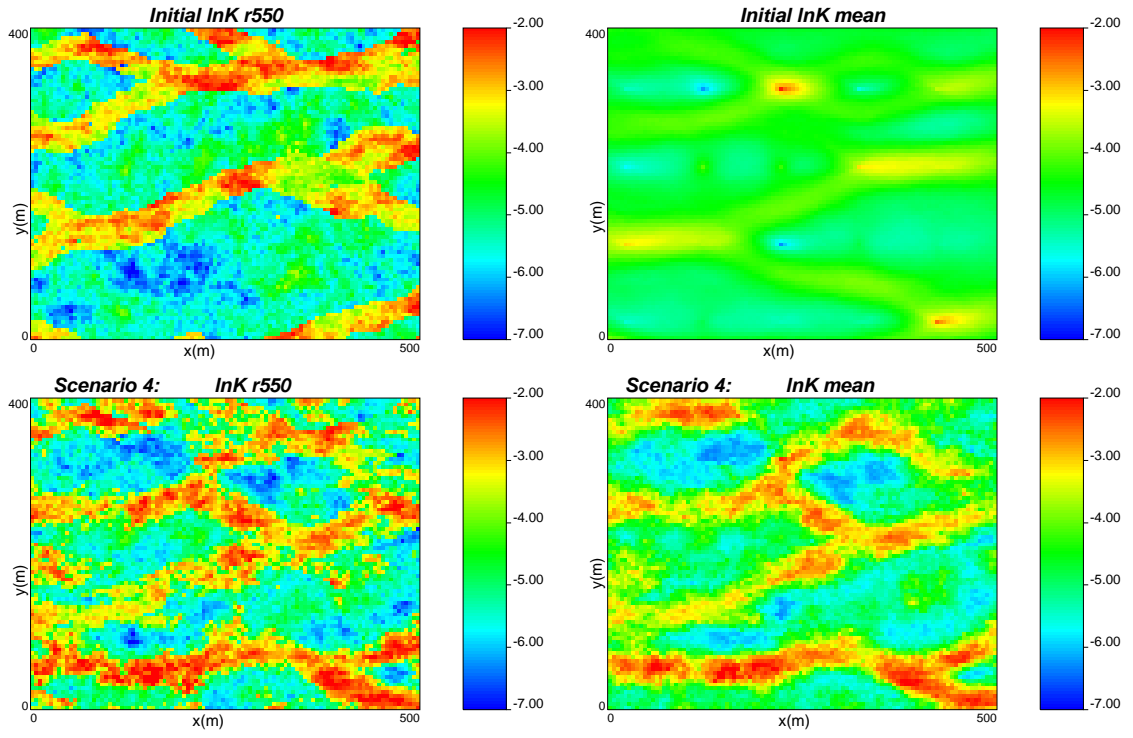


Figure 16: A randomly selected realization (550th) and the ensemble mean before and after 50 time steps of piezometric head data assimilation using the NS-EnKF for the scenario where only 20 log-conductivity conditioning data are used (Scenario 4).

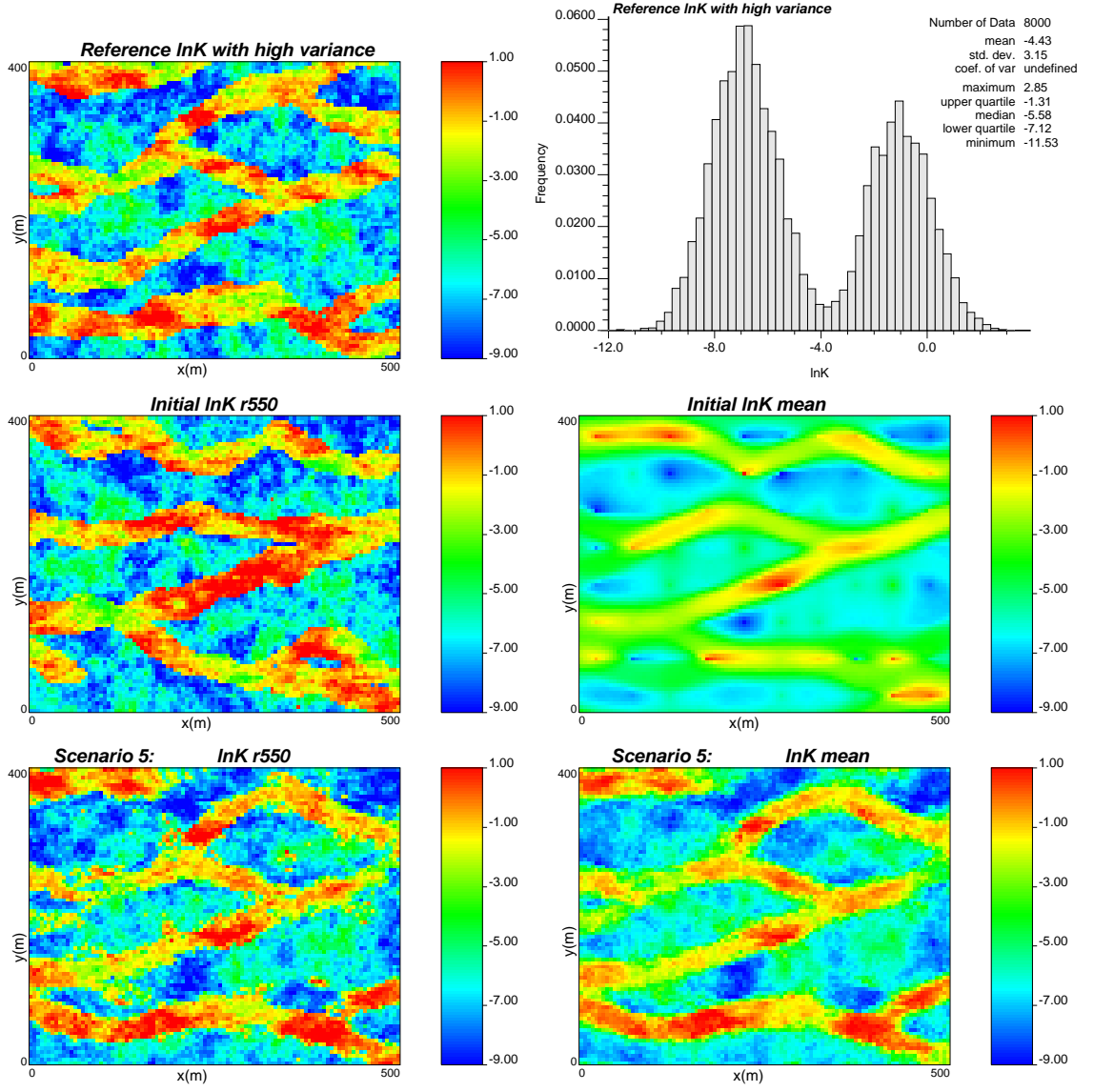


Figure 17: New reference log-conductivity field with larger variance ($9.92 \ln(\text{cm/s})^2$) and corresponding histogram, together with a realization (550^{th}) and the ensemble mean before and after 50 times steps of piezometric head data assimilation using the NS-EnKF (Scenario 5).

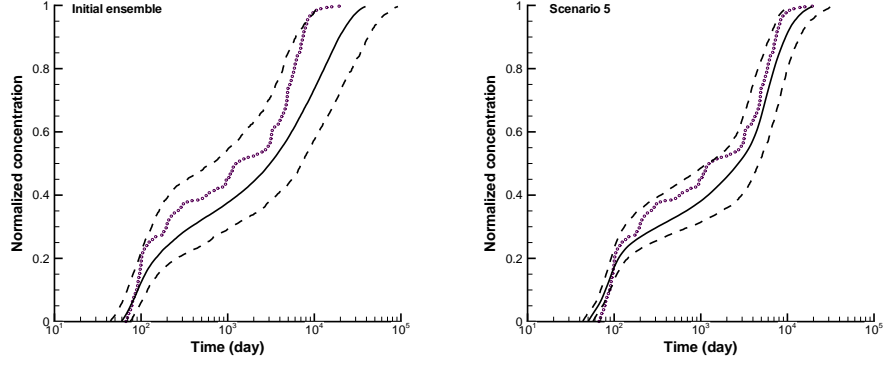


Figure 18: Summary of the breakthrough curves for Scenario 5. The 5th percentile, the median, and the 95th percentile of the travel times are computed as a function of normalized concentration. Results are shown for the initial ensemble of realizations (conditioned only on log-conductivity) and for the realizations after data assimilation.

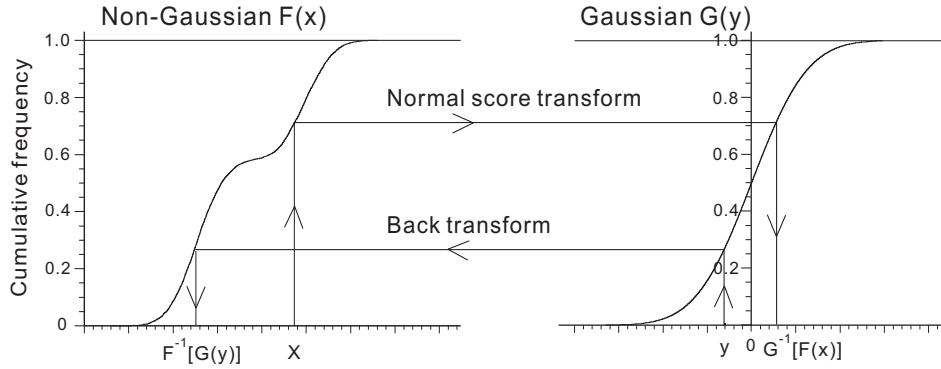


Figure 19: Sketch of normal-score transformation.