



# Groundwater flow inverse modeling in non-MultiGaussian media: performance assessment of the normal-score Ensemble Kalman Filter

L. Li<sup>1</sup>, H. Zhou<sup>1</sup>, H. J. Hendricks Franssen<sup>2</sup>, and J. J. Gómez-Hernández<sup>1</sup>

<sup>1</sup>Group of Hydrogeology, Universitat Politècnica de València, 46022 Valencia, Spain

<sup>2</sup>Agrosphere, IBG-3, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

Correspondence to: L. Li (liali@upvnet.upv.es, liliangping2005@gmail.com)

Received: 1 June 2011 – Published in Hydrol. Earth Syst. Sci. Discuss.: 12 July 2011

Revised: 15 February 2012 – Accepted: 20 February 2012 – Published: 27 February 2012

**Abstract.** The normal-score ensemble Kalman filter (NS-EnKF) is tested on a synthetic aquifer characterized by the presence of channels with a bimodal distribution of its hydraulic conductivities. This is a clear example of an aquifer that cannot be characterized by a multiGaussian distribution. Fourteen scenarios are analyzed which differ among them in one or various of the following aspects: the prior random function model, the boundary conditions of the flow problem, the number of piezometers used in the assimilation process, or the use of covariance localization in the implementation of the Kalman filter. The performance of the NS-EnKF is evaluated through the ensemble mean and variance maps, the connectivity patterns of the individual conductivity realizations and the degree of reproduction of the piezometric heads. The results show that (i) the localized NS-EnKF can characterize the non-multiGaussian underlying hydraulic distribution even when an erroneous prior random function model is used, (ii) localization plays an important role to prevent filter inbreeding and results in a better logconductivity characterization, and (iii) the NS-EnKF works equally well under very different flow configurations.

## 1 Introduction

Accurate characterization of the spatial variability of hydrogeologic properties and its corresponding uncertainty is a key issue for environmental risk assessment, site remediation and restoration engineering, and the design of underground repositories for radioactive material.

In a Monte Carlo framework, heterogeneity of hydrogeologic properties is commonly characterized by the following two steps: (i) on the basis of a limited amount of direct measurements (i.e., hard data), multiple representations of aquifer properties are generated by means of the geostatistical techniques such as sequential Gaussian simulation (Deutsch and Journel, 1998), sequential indicator simulation (Gómez-Hernández and Srivastava, 1990), multiple-point geostatistical approach (Strebelle, 2002; Mariethoz et al., 2010b) or other related methods; and then (ii) on the basis of indirect measurements such as piezometric head and concentration data, inverse modeling is utilized to reduce the uncertainty by integrating these data to better characterize the spatial variability of hydrogeologic properties (e.g. for an overview see Yeh, 1986; McLaughlin and Townley, 1996; Zimmerman et al., 1998; Carrera et al., 2005; Hendricks Franssen et al., 2009; Zhou et al., 2012b).

Commonly used Monte Carlo type inverse algorithms (i.e. which generate many equally likely solutions to the inverse problems) include the self-calibration method (Sahuquillo et al., 1992; Gómez-Hernández et al., 1997; Capilla et al., 1999; Wen et al., 2002; Hendricks Franssen et al., 2003), the pilot point method (Ramarao et al., 1995; LaVenue et al., 1995; Alcolea et al., 2006), the Markov chain Monte Carlo method (Oliver et al., 1997; Fu and Gómez-Hernández, 2009; Alcolea and Renard, 2010), and the gradual deformation method (Hu, 2000; Capilla and Llopis-Albert, 2009), among others. These methods have in common that a multi-part objective function is minimized. The objective function is generally composed of the sum of squared differences (SSD) between simulated and observed

state values plus the SSD of prior and calibrated parameters values. In order to minimize this objective function the hydrogeologic parameters are modified using derivative-based methods or by sampling the posterior distribution. The main difference between the various methods is how to solve the optimization problem.

The Ensemble Kalman Filter (EnKF) (Burgers et al., 1998; Anderson, 2001; Reichle et al., 2002; Evensen, 2003) is a further alternative to generate multiple equally likely solutions to the inverse problem. The EnKF is increasingly studied in hydrogeology as well as in petroleum engineering (e.g. Wen and Chen, 2005; Chen and Zhang, 2006; Hendricks Franssen and Kinzelbach, 2008; Sun et al., 2009; Nowak, 2009; Nan and Wu, 2010; Li et al., 2012b). The attractive characteristics of the EnKF are: (i) the efficiency in computation (e.g. Hendricks Franssen and Kinzelbach, 2009 conducted a synthetic example to demonstrate that the needed CPU-time was reduced by a factor of 80 compared with the self-calibration method, to achieve comparable results); (ii) the flexibility of handling multiple sources of uncertainty, for instance, Hendricks Franssen and Kinzelbach (2008) successfully used EnKF to account for the uncertainty of both recharge and hydraulic conductivity together; Li et al. (2012a) used the EnKF to jointly calibrate porosity and hydraulic conductivity by assimilating dynamic head and multiple concentration data; (iii) real-time data assimilation without the need to store all previous states, for instance, Hendricks Franssen et al. (2011) operationally implemented the EnKF to calibrate the conductivity and leakage coefficient together in real-time in a real-world case study. On the contrary, the traditional inverse approaches mentioned above (self-calibration, pilot point, etc.) are CPU-intensive, need recalibration when new data are available and handling multiple sources of uncertainty is less straightforward with these methods.

The EnKF provides an optimal solution when the state vector follows a multiGaussian distribution and the state function is linear (Arulampalam et al., 2002). In the literature of hydrogeology, most of studies applying the EnKF technique assume that the hydraulic logconductivities follow a multiGaussian distribution (e.g. Chen and Zhang, 2006; Hendricks Franssen and Kinzelbach, 2008; Huang et al., 2008; Li et al., 2012a). The significance of accounting for non-multiGaussian distributions of hydraulic conductivity for flow and transport predictions has been stressed in many studies (e.g. Gómez-Hernández and Wen, 1998; Zinn and Harvey, 2003; Knudby and Carrera, 2005). Hence, extending the EnKF to deal with non-Gaussian state vectors would facilitate more extensive applications. Sun et al. (2009) and Zhou et al. (2011, 2012c), developed variants of the EnKF which are better accommodated to handle non-Gaussianity of parameter distributions. Sun et al. (2009) resort to couple the EnKF with a Gaussian mixture model to update the parameters of a multi-modal distribution by assimilating head data. Zhou et al. (2011, 2012c) transformed the augmented state vector (containing both the hydraulic conductivities and

piezometric heads) with marginal multi-modal distributions into a new vector with marginal Gaussian distributions, perform the EnKF on the transformed state vector and back-transform the updated vector to the original space. Both studies show that these variants of EnKF result in significant improvements in the characterization of the aquifer properties and in the predictions of flow and transport for non-multiGaussian hydraulic conductivity fields. It is worth noting that, in both studies, it was assumed that the prior random function is known, i.e. the reference field and the initial set of non-multiGaussian conductivity fields are both generated with the same geostatistical approach and the same random function model, which, in this particular case, amounted to using the same training image in the multiple-point geostatistical generator. However, in practice, it is not trivial to decide about a reasonable training image given the limited amount of information available. Partly for this reason, traditional two-point variogram-based geostatistical methods are still employed in practice. It is worth recall that the normal-score transformations used in the NS-EnKF only ensure marginal Gaussianity, with the higher-order moments not necessarily closer to the multiGaussian distribution; it is also worth recall, that the EnKF is applied on the normal-score transformed variables what amounts to having a more non-linear forecast function that before the transformations. One of the motivations of this work is to explore the capacity of the normal score EnKF proposed by Zhou et al. (2011, 2012c) to characterize non-multiGaussian media such as a highly channelized aquifer when there are no (hard) conductivity data available and the prior random function model is not the one used to generate the reference field.

Several studies (e.g. Gómez-Hernández and Wen, 1998; Western et al., 2001; Zinn and Harvey, 2003; Knudby and Carrera, 2005; Feyen and Caers, 2005; Lee et al., 2007) showed that flow and transport predictions strongly differ between multiGaussian and non-multiGaussian logconductivity fields, even when both types of fields share the same histogram and the same covariance function. These results demonstrated that the reproduction of the connectivity is of significance in practice. For the inverse conditioning, Kerrou et al. (2008) applied the self-calibration (Gómez-Hernández et al., 1997) method on a fluvial-sediment aquifer with a wrong prior model (i.e. multiGaussian instead of non-multiGaussian) and concluded that the channel structures cannot be retrieved, even when a large number of direct and indirect data are used for conditioning.

In this paper, we apply the normal-score EnKF (NS-EnKF) to a channelized aquifer characterized by a bimodal conductivity distribution, and analyze the performance of the method for fourteen scenarios which differ among them in one or several of the following aspects: the prior random function model, the boundary conditions of the flow problem, the number of piezometers used in the assimilation process, or the use of covariance localization in the implementation of the Kalman filter. None of the scenarios uses any

conductivity conditioning data. The analysis focus on the ensemble mean and variance maps, the connectivity patterns of the individual conductivity realizations and the degree of reproduction of the piezometric heads.

This paper revisits the mathematical framework of the NS-EnKF as applied in the synthetic experiment and briefly presents its numerical implementation (Sects. 2 and 3); the impact of the choice of the prior random function model is discussed in Sect. 4.1, the effect of using localization functions for the covariance while computing the Kalman gain is discussed in Sect. 4.2, the performance under different boundary conditions inducing very different flow patterns in the aquifer is discussed in Sect. 4.3, and, finally the impact of reducing the number of conditioning piezometers is discussed in Sect. 4.4. The paper ends with some conclusions (Sect. 5).

## 2 Mathematic framework

### 2.1 Flow equation

The flow equation of an incompressible fluid in saturated porous media in a Cartesian coordinate system can be obtained by combining the continuity equation and Darcy's law (Bear, 1972):

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

where  $h[L]$  is the piezometric head;  $\mathbf{K}[LT^{-1}]$  is a symmetric positive-definite rank-two tensor;  $Q[T^{-1}]$  is the volumetric source flow per unit volume;  $S_s[L^{-1}]$  is the specific storage coefficient;  $t[T]$  is the time;  $\nabla \cdot = (\partial/\partial x + \partial/\partial y + \partial/\partial z)$  is the divergence operator of a vector field, and  $\nabla = (\partial/\partial x, \partial/\partial y, \partial/\partial z)^T$  is the gradient operator of a scalar field.

### 2.2 The normal-score ensemble kalman filter with localization

The NS-EnKF method aims at generating equally-likely realizations of parameters and state variables, conditioned to real-time measurements such as piezometric head data following non-Gaussian marginal distributions. The basic algorithm is described in Zhou et al. (2011) and is summarized next for the case of assimilating observed piezometric head data at time  $t$  in a bimodal aquifer.

1. Initialization step. Generate the initial ensemble of log-conductivity fields (in case that there are logconductivity measurements, these fields must be conditional to them); this is achieved by any of many geostatistical simulation algorithms, such as sequential indicator simulation (e.g. Deutsch and Journel, 1998) or the multiple-point method (e.g. Strebelle, 2002). The choice of the algorithm will depend on the random function model

adopted to describe the spatial variability of logconductivity. Here, it is assumed that the scale of hydraulic conductivity measurements have a support coinciding with that of the numerical model discretization grid-blocks. If there were a discrepancy between the measurement scale and the gridblock scale, an upscaling technique would have been needed to reconcile these two scales (e.g. Zhou et al., 2010; Li et al., 2011a,b).

2. Normal-score function step. Build the logconductivity local conditional distribution functions at each grid cell from the ensemble of logconductivity realizations, and the corresponding local normal-score transfer functions ( $\phi$ ) (for the details see: Goovaerts, 1997).
3. Forecast step. For each realization, the transient groundwater flow Eq. (1) is solved from time  $t-1$  to  $t$  using standard block-centered finite differences (e.g. Harbaugh et al., 2000; Li et al., 2010). The solution can be schematically represented by:

$$\mathbf{Y}_k = f(\mathbf{X}_{k-1}, \mathbf{Y}_{k-1}) \quad (2)$$

where  $\mathbf{X}_{k-1}$  and  $\mathbf{Y}_{k-1}$  denote the hydraulic conductivity and piezometric head estimates at time  $t_{k-1}$  respectively, and  $\mathbf{Y}_k$  are the forecasted piezometric heads at time step  $t_k$ ;  $f$  represents the groundwater flow model including the boundary conditions, stresses and other known parameters.

4. Normal-score function step. Build the piezometric head local conditional distribution functions from the ensemble of forecasted realizations, and the corresponding local normal-score transfer functions ( $\varphi$ ).
5. Filter step. Update the state variables by the NS-EnKF based on the observed head data at time  $t$ .

- a. Normal-score transform each conductivity, each forecasted head, and also the observed heads

$$\hat{\mathbf{X}} = \phi(\mathbf{X}) \quad (3)$$

$$\hat{\mathbf{Y}} = \varphi(\mathbf{Y}) \quad (4)$$

$$\hat{\mathbf{Y}}^{\text{obs}} = \varphi(\mathbf{Y}^{\text{obs}}). \quad (5)$$

- b. Build the augmented normal-score transformed vector  $\hat{\Psi}$ , which includes both the transformed conductivities ( $\hat{\mathbf{X}}$ ) and the transformed forecasted heads ( $\hat{\mathbf{Y}}$ ).

$$\hat{\Psi}_{k,j} = \begin{bmatrix} \hat{\mathbf{X}} \\ \hat{\mathbf{Y}} \end{bmatrix}_{k,j}, \quad (6)$$

where  $\hat{\Psi}_{k,j}$  is the  $j$ -th ensemble member of the augmented state vector at time  $t_k$ .

**Table 1.** Parameters of the random functions describing the spatial continuity of the sand and shale logconductivities.

	Mean [ln m d <sup>-1</sup> ]	Standard deviation [ln m d <sup>-1</sup> ]	Variogram type	$\lambda_x$ [m]	$\lambda_y$ [m]	sill
Sand	2.1	0.7	exponential	144	72	1
Shale	-1.4	0.7	exponential	72	72	0.35

- c. Calculate the localized Kalman gain ( $\mathbf{G}_k$ ) (e.g. Gaspari and Cohn, 1999; Hamill et al., 2001; Chen and Oliver, 2010).

$$\mathbf{G}_k = \rho_{\hat{\mathbf{x}}\hat{\mathbf{y}}} \circ \mathbf{C}_{\hat{\mathbf{x}}\hat{\mathbf{y}}} (\rho_{\hat{\mathbf{y}}\hat{\mathbf{y}}} \circ \mathbf{C}_{\hat{\mathbf{y}}\hat{\mathbf{y}}} + \mathbf{C}_D)^{-1} \quad (7)$$

where  $\mathbf{C}_{\hat{\mathbf{x}}\hat{\mathbf{y}}}$  is the cross-covariance between transformed logconductivity and head data;  $\mathbf{C}_{\hat{\mathbf{y}}\hat{\mathbf{y}}}$  is the covariance of transformed head data,  $\mathbf{C}_D$  is the diagonal matrix of expected measurement error variances;  $\circ$  indicates the Schur product; and  $\rho_{\hat{\mathbf{x}}\hat{\mathbf{y}}}$  and  $\rho_{\hat{\mathbf{y}}\hat{\mathbf{y}}}$  are localization functions used for the  $\mathbf{C}_{\hat{\mathbf{x}}\hat{\mathbf{y}}}$  and  $\mathbf{C}_{\hat{\mathbf{y}}\hat{\mathbf{y}}}$ , respectively; the specific localization functions used in this work will be given later.

- d. Update each ensemble member of the state vector (e.g. Burgers et al., 1998; Evensen, 2003).

$$\hat{\Psi}_{k,j}^u = \hat{\Psi}_{k,j}^f + \mathbf{G}_k (\hat{\mathbf{Y}}^{\text{obs}} + \boldsymbol{\epsilon} - \hat{\mathbf{Y}}^f) \quad (8)$$

where superscripts  $u$  and  $f$  denote updated and forecasted, respectively;  $\boldsymbol{\epsilon}$  is a random observation error vector, and  $\hat{\mathbf{Y}}^f$  is a vector containing the transformed forecasted heads at sampling locations (a subset of  $\hat{\mathbf{Y}}$ ).

- e. Back transform the state vector  $\hat{\Psi}$ .

$$\mathbf{X} = \phi^{-1}(\hat{\mathbf{X}}) \quad (9)$$

$$\mathbf{Y} = \varphi^{-1}(\hat{\mathbf{Y}}) \quad (10)$$

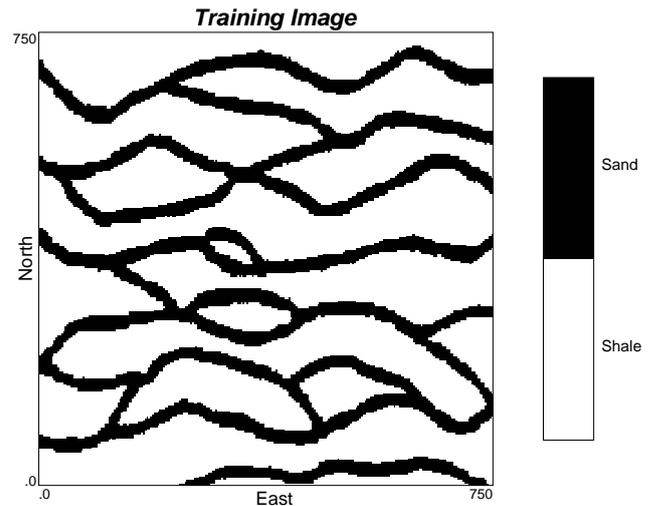
6. Go to step 3 for the next time step.

The NS-EnKF assures that the non-Gaussianity of the marginal distributions of states and parameters is preserved, through the use of the normal-score transform. The distance-dependent localization functions are used to reduce the influence of spurious correlations for large separation distances that may appear in the covariance computed through the ensemble of realizations.

### 3 Synthetic example

#### 3.1 Reference

The domain area is 300 m by 240 m by 10 m, and is discretized into  $100 \times 80 \times 1$  blocks of size  $3 \text{ m} \times 3 \text{ m} \times 10 \text{ m}$ . The reference logconductivity field is generated in two steps:

**Fig. 1.** Training image used to generate the facies distribution.

(i) the SNESIM code, a pattern-based multiple-point geostatistical algorithm, is utilized to generate a facies realization (Strebelle, 2002). (In contrast with traditional two-point variogram-based models, the multiple-point method is able to reproduce complex and realistic curvilinear structures characteristic of fluvial deposits.) We used the training image in Strebelle (2002), which is commonly used for benchmarking to compare algorithms (e.g. Wu et al., 2008; Mariethoz et al., 2010a; Zhou et al., 2012.a). This training image serves as a conceptual model for the channelized non-multiGaussian aquifer, where the channels have high conductivities representing preferential paths and are embedded in a floodplain fine-grid deposits with low conductivities (see Fig. 1); (ii) after the facies are generated, the facies are populated with realizations of conductivity fields generated using sequential gaussian simulation algorithm (Gómez-Hernández and Journel, 1993) with the parameters listed in Table 1. This procedure results in the spatial distribution of logconductivity values shown in Fig. 2a, which serves as the reference  $\ln K$  realization. This realization displays well-connected sand channels (approximately 30 % of the system) on a low-conductivity matrix. The histogram of  $\ln K$ , shown in Fig. 2b, clearly shows a bimodal logconductivity distribution typical of fluvial deposits.

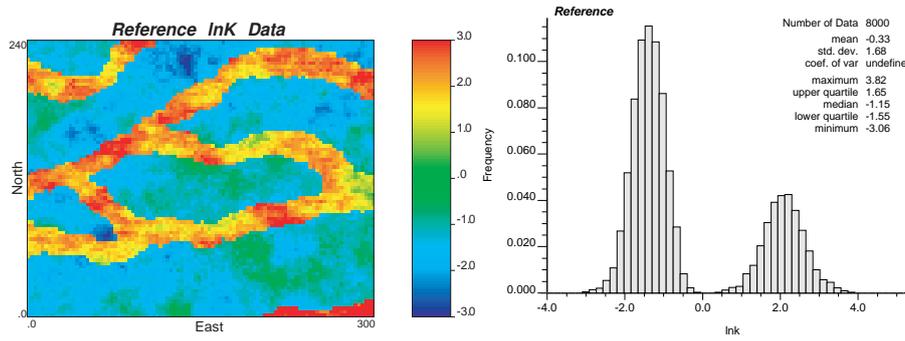


Fig. 2. Reference  $\ln K$  data.

To explore the role of boundary conditions on pattern identification, two sets of boundary conditions are considered: (i) Boundary conditions inducing parallel flow (Fig. 3a), consisting of a combination of prescribed head boundary (West), prescribed flux boundary (East) and impermeable boundaries (North and South); (ii) boundary conditions inducing radial flow (Fig. 3b), with impermeable boundaries and a set of injection/extraction wells in a typical configuration for contaminant remediation or for reservoir exploitation. For both boundary conditions, groundwater flow is simulated for a period of 500 days starting with an initial head equal to zero everywhere. The simulation time is discretized into 100 time steps, the sizes of which follow a geometric series with a ratio of 1.05. The specific storage is assumed constant and equal to  $0.003 \text{ m}^{-1}$ . The MODFLOW 2000 code (Harbaugh et al., 2000) is used to solve the transient flow Eq. (1). The head data are sampled at the first 60 time steps (approximately 67.7 days) and will be used as the conditioning data in the NS-EnKF. The spatial distribution of the sampling wells is shown in Fig. 3. Although the sampled head data are error-free since they are taken from the “true” field, in the stochastic model the sampled head data should have noise added to it. In the following data assimilation procedures it was assumed that they have a normalized measurement error with a mean of zero and a standard deviation of 0.05 m.

3.2 Scenarios

Fourteen scenarios are analyzed with the characteristics indicated in the Table 2. The impact of the prior model, the use of a localization function and the different boundary conditions is tested with help of fourteen scenarios (Table 2).

The scenarios are organized to show how the characterization of the hydraulic conductivities evolves starting from an unconditional ensemble of realizations as piezometric heads are incorporated, and then when localization is used. This evolution is analyzed for each of the two flow patterns induced by the two sets of boundary conditions, and also for the cases in which the correct and an erroneous prior random functions are used.

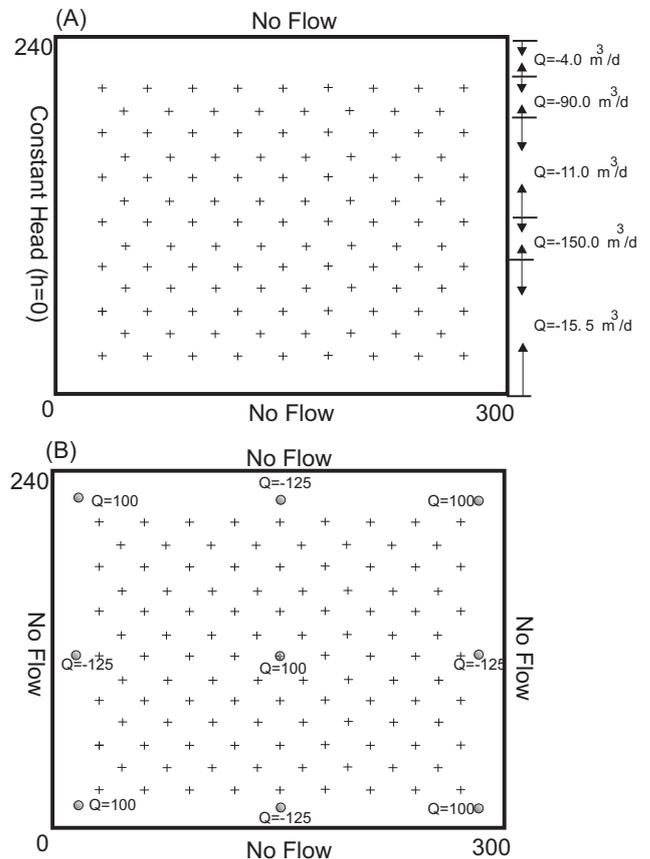
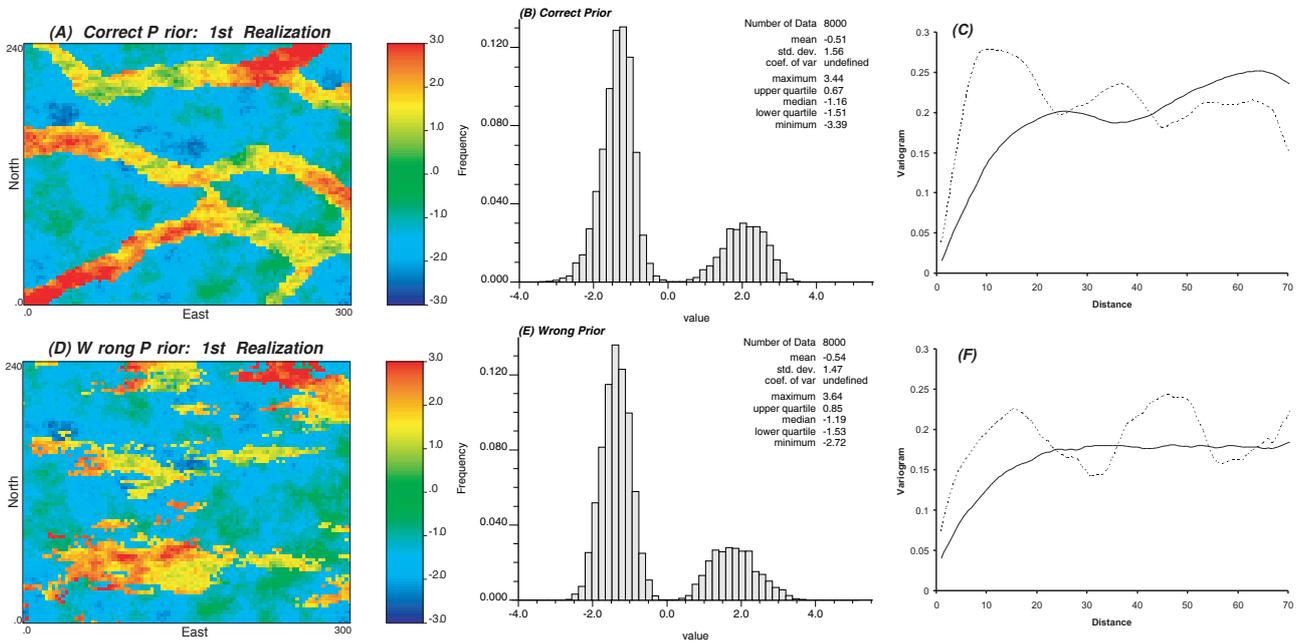


Fig. 3. Spatial distribution of sampled head data and flow configurations: (A) Parallel flow (B) Radial flow (the unit of flow in wells:  $\text{m}^3 \text{ d}^{-1}$ ).

Scenarios S1, S2, S7 and S8 are unconditional, they are the base cases containing the initial ensemble realizations to be used by the NS-EnKF in the other scenarios. They will produce the worst results in terms of conductivity characterization and piezometric head predictions and will serve as a starting point to assess the benefits of using the correct prior model, the assimilation of dynamic piezometric head information, and the use of localization.

**Table 2.** Definition of scenarios.

Scenario	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Parallel flow	✓	✓	✓	✓	✓	✓							✓	✓
Radial flow							✓	✓	✓	✓	✓	✓		
Piezometric head (111)			✓	✓	✓	✓			✓	✓	✓	✓		
Piezometric head (32)													✓	✓
Correct Prior	✓		✓	✓			✓		✓	✓			✓	✓
Wrong Prior		✓			✓	✓		✓			✓	✓		✓
NS-EnKF			✓		✓				✓		✓		✓	✓
Localized NS-EnKF				✓		✓				✓		✓		✓



**Fig. 4.** (A–C) show the spatial distribution of  $\ln K$ , histogram and variogram of the 1st realization from the correct prior model; (D–F) show the same characteristics of the 1st realization from wrong prior model. In (C) and (F), solid line and dotted line correspond to the experimental facies variogram in  $X$  and  $Y$  direction, respectively.

Scenarios S1 to S6, S13 and S14 use boundary conditions inducing parallel flow through the aquifer, and scenarios S7 to S12 boundary conditions inducing radial flow.

The prior random function models for all realizations share the same histogram and variogram but they differ in their higher order moments. Scenarios S3, S4, S9, S10, and S13, use the same random function model used to generate the reference, i.e. the facies follow the random function model implicit in the training image of Fig. 1 on which two independent multiGaussian random functions are overlaid to describe the variability of conductivity within each facies. Scenarios S5, S6, S11, S12 and S14, use a random function model for the facies based on the indicator variogram computed on the training image on which the same independent multiGaussian random functions of the reference field are overlaid. For the first set of scenarios the multiple point

code SNESIM (Strebelle, 2002) is used to generate the facies realizations, whereas the indicator simulation code ISIM3D (Gómez-Hernández and Srivastava, 1990) is used for the second set. The parameters defining the two multiGaussian random functions are given in Table 1. Figure 4 shows the first realization for the correct and the incorrect model, their histograms and their variograms. While both realizations follow the same histogram and variograms, the variogram-based sequential indicator simulation cannot generate the curvilinear patterns for the sand facies observed in the training image, which are well reproduced in the multipoint sequential simulation.

For the scenario S3, S5, S9, S11, S13 and S14 head data are assimilated without using localization whereas for the scenarios S4, S6, S10 and S12 localization is employed.

The EnKF tends to filter inbreeding if the ensemble size is small, and the heterogeneity is large (Hendricks Franssen and Kinzelbach, 2009). It is therefore expected that the NS-EnKF applied to a complex geological environment will suffer this problem, too. To analyze this aspect, a fifth-order distance-dependent localization function (Hamill et al., 2001) is used to reduce the sampling errors in the computation of the ensemble covariances. Considering previous works (Chen and Oliver, 2010), the size of the domain, the separation between piezometers, and the distance between wells for the boundary conditions inducing radial flow, the localization function is set to zero at a distance of 80 m, implying that for distances larger than that, the sampled non-stationary covariances used to compute the Kalman gain are zero. More specifically, the distance function that is used for both  $\rho_{\hat{X}\hat{Y}}$  and  $\rho_{\hat{Y}\hat{Y}}$  in Eq. (7) is:

$$\Omega(d) = \begin{cases} \frac{1}{4} \left(\frac{d}{40}\right)^5 + \frac{1}{2} \left(\frac{d}{40}\right)^4 + \frac{frac58}{8} \left(\frac{d}{40}\right)^3 - \frac{5}{3} \left(\frac{d}{40}\right)^2 + 1, & 0 \leq d < 40; \\ \frac{1}{12} \left(\frac{d}{40}\right)^5 - \frac{1}{2} \left(\frac{d}{40}\right)^4 + \frac{5}{8} \left(\frac{d}{40}\right)^3 + \frac{5}{3} \left(\frac{d}{40}\right)^2 - 5 \left(\frac{d}{40}\right) + 4 - \frac{2}{3} \left(\frac{d}{40}\right)^{-1}, & 40 \leq d \leq 80; \\ 0 & d > 80. \end{cases} \quad (11)$$

where  $d$  is the distance between observed head and  $\ln K$  data (when applied to  $\rho_{\hat{X}\hat{Y}}$ ) or between observed heads (when applied to  $\rho_{\hat{Y}\hat{Y}}[L]$ ). Here, the distance function is constant in time, isotropic, and used both for the cross-covariances between observed heads and parameters and the covariances of the head data.

Scenarios 13 and 14 stand apart from the rest of the conditional scenarios in that the number of piezometers used for assimilation is reduced to one third. Recall that no conductivity conditioning data are used for any scenario, therefore, the ability to identify the conductivity patterns is solely thanks to the dynamic head data that are assimilated at each time step, and it is expected that the quality of the characterization of the conductivities will be affected by the number of conditioning piezometers.

For all the scenarios, 1000 unconditional realizations of logconductivity are generated.

### 3.3 Evaluation criteria

For every scenario, the following criteria are used to assess the results (Chen and Zhang, 2006; Zhou et al., 2012c):

1. Root mean square error (RMSE). RMSE measures the accuracy of the estimation.

$$RMSE(X) = \left[ \frac{1}{N_b} \sum_{i=1}^{N_b} (\bar{X}_i - X_{ref,i})^2 \right]^{1/2} \quad (12)$$

where  $X_i$  is either logconductivity  $\ln K$  or hydraulic head  $h$  at location  $i$ ,  $\bar{X}_i$  represents its ensemble mean value at node  $i$ ,  $X_{ref,i}$  is the reference value at location  $i$ ,  $N_b$  is the number of nodes.

2. Ensemble spread (ES). ES indicates the uncertainty (precision) of the estimation.

$$ES(X) = \left[ \frac{1}{N_b} \sum_{i=1}^{N_b} \sigma_{X_i}^2 \right]^{1/2} \quad (13)$$

where  $\sigma_{X_i}^2$  is the ensemble variance at location  $i$ .

The smaller the values for RMSE and ES, the better the prediction of variable  $X$ . As discussed in Chen and Zhang (2006), when RMSE and ES have a similar magnitude the resulting ensemble variance provides a more realistic measure of the uncertainty associated to the ensemble mean estimate.

3. Connectivity. The flow and transport behaviors are strongly impacted by the presence of connected high-conductivities and connected low-conductivities (e.g. Wen and Gómez-Hernández, 1998; Knudby and Carrera, 2005; Fernández-García et al., 2010). Here, we adopt the connectivity function defined by Pardo-Igúzquiza and Dowd (2003) to evaluate the connectivity. In this work, two steps are needed to analyze the connectivity for the continuous variable:

- i. An indicator image is obtained as:

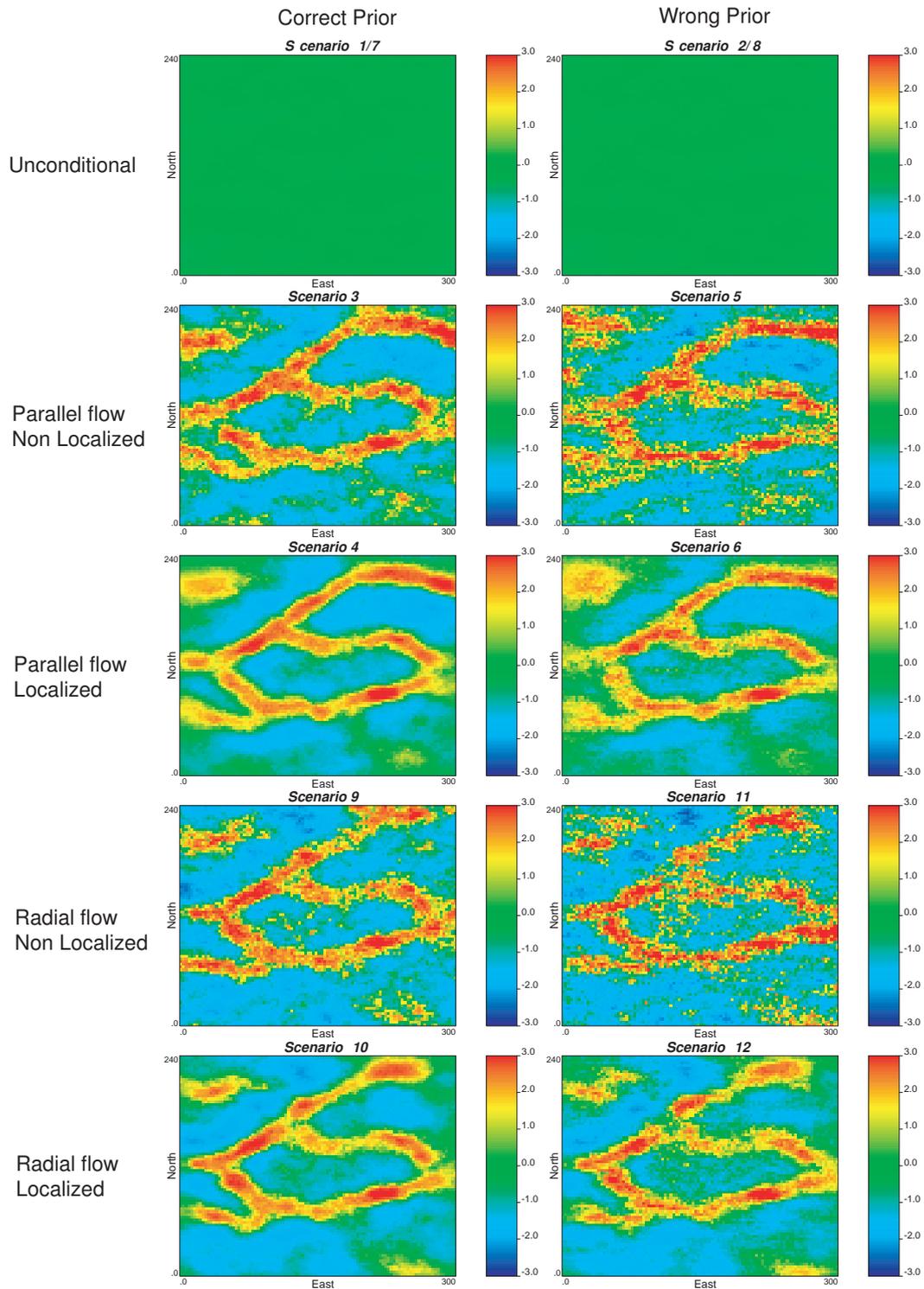
$$I(x) = \begin{cases} 1, & \text{if } \ln K \geq \alpha \\ 0, & \text{otherwise} \end{cases} \quad (14)$$

where  $\alpha$  is the threshold that classifies the logconductivity into sand and shale. In our case,  $\alpha$  is set to zero since this value approximately splits the histogram of logconductivity in two (see Fig. 2).

- ii. The code CONNEC3D (Pardo-Igúzquiza and Dowd, 2003), a computer program for connectivity analysis of 3-D random model, is used to calculate the probability that two points with logconductivity larger than  $\alpha$  are connected. The result is a function of distance that gives the probability that two points located in the sand facies are connected by a continuous path in that same facies.

## 4 Results and discussion

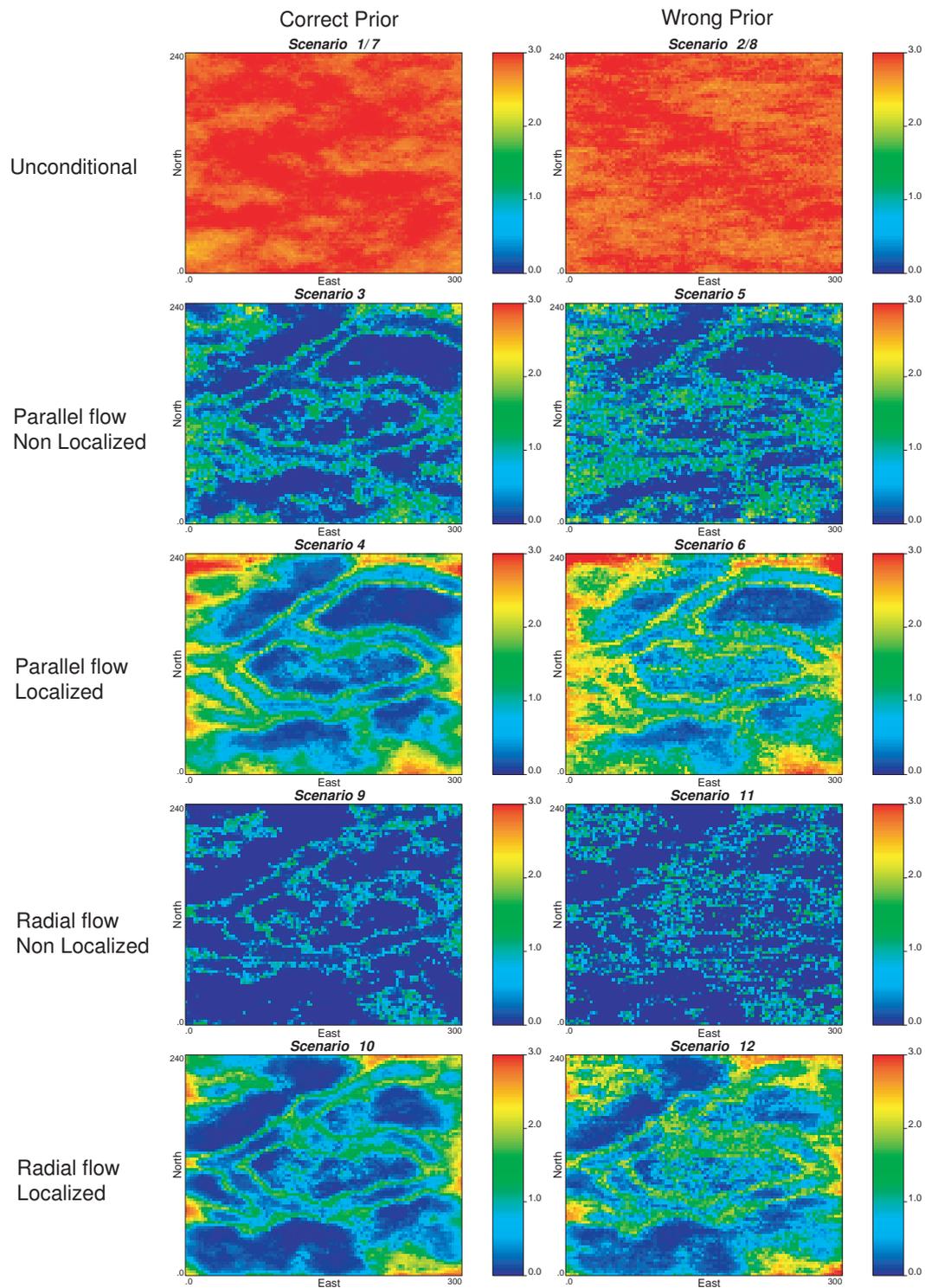
Ensembles of conductivity realizations are generated for the fourteen scenarios according to the configurations described earlier. As mentioned earlier, scenarios S13 and S14 stand apart in that they are the only ones using a smaller number of conditioning piezometric heads, for this reason most of the figures group the results for scenarios S1 to S12. Scenarios S13 and S14 are analyzed later by themselves. Figures 5 and 6 show the ensemble average and ensemble variance of logconductivity at time step 60 for the first twelve scenarios. Figure 7 shows the evolution of RMSE and ES



**Fig. 5.** Ensemble average logconductivity fields at time step  $nT = 60$  for the different scenarios.

as function of the updating time step. Figure 8 displays the connectivity function for sand as a function of the horizontal distance. Figure 9 shows the RMSE and ES of the predicted heads computed on the updated hydraulic conductivities after

60 time steps (67.7 days) by rerunning the flow model from  $t = 0$ .



**Fig. 6.** Ensemble variance logconductivity fields at time step  $nT = 60$  for the different scenarios.

#### 4.1 Effect of prior random function model

For the scenarios S1, S2, S7 and S8 no conditioning data were used. For those scenarios, the ensemble average of hydraulic logconductivity is approximately spatially uniform

and equal to the prior mean (see Fig. 5) and, similarly, the ensemble variance map is approximately equal to the prior variance (see Fig. 6). As expected, the unconditional average maps do not capture any spatial pattern, although individual realizations do display the spatial variability according to

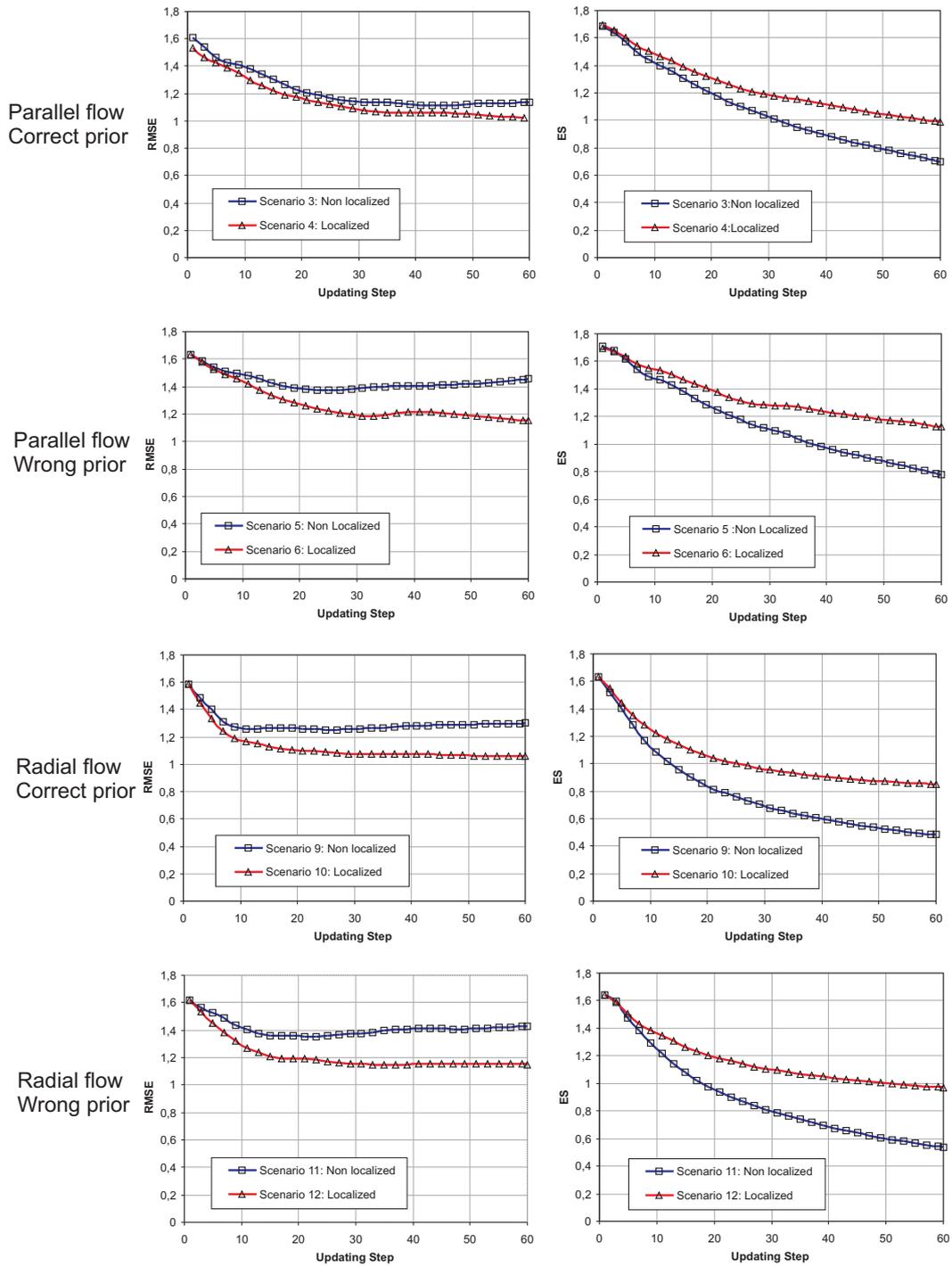
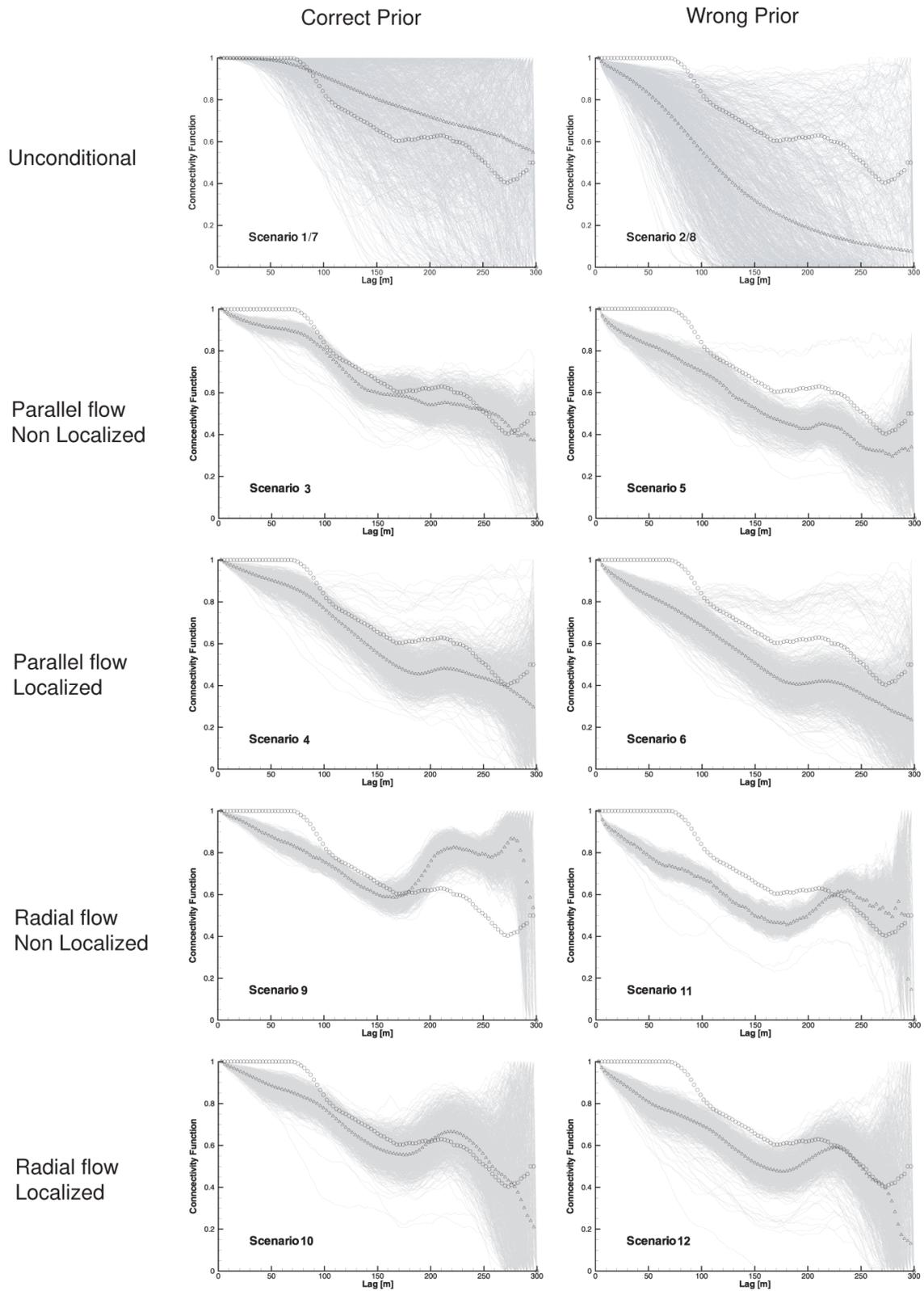


Fig. 7. RMSE and ES of  $\ln K$  for the different cases.

their prior random function models (see the first realization in Fig. 4a and d).

For the scenarios where head data are used for assimilation, the ensemble means of hydraulic conductivity (estimated after 60 time steps of assimilation), depict the facies distribution well in visual comparison with the reference.

The ensemble variances of logconductivity are clearly reduced in comparison with the prior variances and display a feature with higher variance at the boundaries of the channels and at places far away from the head measurements such as the boundaries of the domain. Comparing the two columns on Figs. 5 and 6, it is quite striking to realize that, even



**Fig. 8.** Connectivity function of logconductivity fields at time step  $nT = 60$  for the different scenarios. Gray curves correspond to individual realizations, their mean is given by the triangles and the circles correspond to the reference.

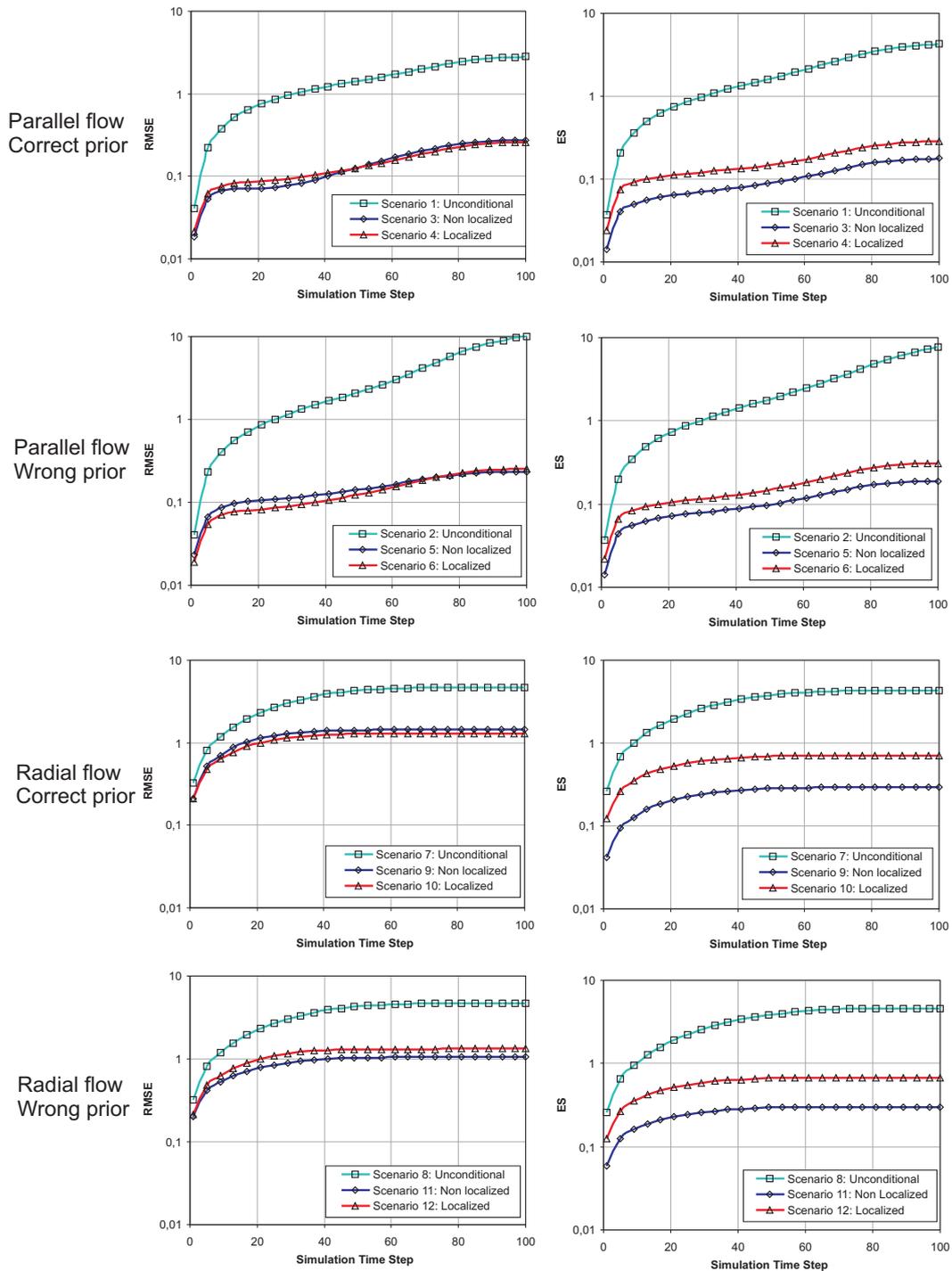


Fig. 9. RMSE and ES of simulated heads for the different scenarios.

for the wrong prior random function model, assimilating the piezometric head data with the NS-EnKF, yields ensembles of realizations which, on average, capture the location of the channels in the reference field and display the highest uncertainty at the boundary channels.

For a more quantitative analysis, Fig. 7 shows how both the RMSE and the ES evolve for the different conditional scenarios as a function of the updating step. It is not surprising that the wrong prior always gives higher values for the average departure between individual realizations and the reference and also for the average spread, being most noticeable for the case of parallel flow.

Regarding the reproduction of the connectivity function, Fig. 8 shows the individual connectivity functions for all realizations, their average and the connectivity of the reference field for the first 12 scenarios. The discrepancy between the results obtained with the correct and the wrong prior are more significant in this case, something that is quite understandable since the wrong prior only uses the indicator variogram of the facies distribution as the starting set of ensemble realizations. The average connectivity of the wrong prior is significantly smaller than the reference connectivity for the unconditional case; then, conditioning to piezometric heads helps increasing the average connectivity as well as the spread of the individual connectivity functions, however, the connectivity remains below that of the reference for all cases. The average connectivity of the correct prior starts slightly above the reference one with a very wide spread of the individual functions, and gets closer to the reference when the piezometric heads are assimilated.

Finally, regarding the reproduction of the piezometric heads, Fig. 9 shows the RMSE and ES as a function of the updating step for the first 12 scenarios. In these figures we have as a reference the metric values for the unconditional cases to show the important reduction in both metrics induced by conditioning to the piezometric heads. Like in Fig. 7, these metrics are average values computed over the entire aquifer, not just for the conditioning piezometers. The behavior of the wrong prior is very similar to that of the correct prior with the only noticeable difference that the localization tends to worsen the RMSE for the wrong model whereas for the correct prior, localization improves the RMSE.

In summary, the NS-EnKF displays a great potential to capture the patterns of variability of logconductivities on the sole basis of piezometric information for clearly non-multiGaussian fields even when the prior random function model is not fully consistent with the reference realization.

## 4.2 Effect of localization

It is well known that the standard EnKF has the problem of filter inbreeding, which leads to a final ensemble of realizations too similar among them resulting in the underestimation of uncertainty of flow and transport predictions (Hendricks Franssen and Kinzelbach, 2009; Devegowda et al., 2010; Zhou et al., 2011). The ensemble size we have used is large, 1000 realizations, yet, when analyzing the piezometric head covariances, and cross-covariances with logconductivity we observe the same type of spurious correlation values that, eventually, lead to filter inbreeding; therefore, we have decided to apply localization techniques and compare the resulting ensembles with and without localization. In general, localization should increase the ensemble spread and, at the same time, reduce the RMSE, that is, there is a gain in accuracy at the cost of precision, which also helps to avoid filter inbreeding.

Analyzing Figs. 5 and 6 we observe that the ensemble mean hydraulic conductivity maps produced with the localized NS-EnKF are smoother than the same mean maps without localization, and capture better the main channels of the reference field. All artifacts in the non-localized realizations are removed after localization; the local ensemble variances in the localized ensembles increase, particularly significant values appear in the northwestern corner of the parallel flow exercise an area without conditioning data and too close to a prescribed head boundary where the sensitivity of the piezometric head to a change in logconductivity is close to zero. These results may indicate that even in this case with a large number of realizations, the ensemble variance estimate obtained from a non-localized NS-EnKF may be too optimistic.

The effect of localization is also quite clear when analyzing the RMSE and ES of the calibrated conductivity fields (see Fig. 7). On one hand, localization induces a reduction of the RMSE in all cases, and an increase in the ES. In all cases, these changes result in values of RMSE and ES that are very similar in value at the end of the updating steps, which is indicative, according to Chen and Zhang (2006), of reduced filter inbreeding. Taking the ratio ES/RMSE as an index of filter inbreeding, this is worse for the radial flow scenarios without localization. In previous synthetic studies, Chen and Zhang (2006) and Hendricks Franssen and Kinzelbach (2008) have found that 200 realizations are enough to avoid filter inbreeding without localization for multiGaussian fields. Here we observe that filter inbreeding can be serious when EnKF is used without localization in a fluvial-type aquifer.

Regarding sand connectivity it is not clear whether localization improves its characterization. The spread of the connectivity functions is much larger for the localized ensemble than for the non localized one, and, at least for parallel flow, the average of the connectivity functions does not get closer to the real one after localization; the small spread exhibited by the non localized connectivity functions makes the reference connectivity one fall outside the cloud of individual functions for the radial flow case. After localization, the reference connectivity is within the cloud for the correct prior function model, while still falls partly outside for the wrong prior. The falling of the reference connectivity function outside the cloud of individual functions is also an indication of underestimation of the ensemble variability induced by the non localized NS-EnKF.

Finally, localization has a minimal impact in the reproduction of the piezometric heads as seen in Fig. 9, the RMSE is almost the same with and without, although localization serves to reduce the RMSE for the correct prior scenarios, but increases it for the wrong prior ones. Localization, as expected, increases the ES in all cases.

It is concluded that, for this particular case, 1000 realizations is not enough to avoid filter inbreeding in the application of the EnKF. For this reason we recommend the use

of localization together with the NS-EnKF to improve the results.

### 4.3 Effect of boundary conditions

The NS-EnKF has been tested for two sets of boundary conditions which basically induce parallel flow and radial flow, the rest of the setup being the same.

The characterization of the hydraulic conductivities is good under both flow conditions. The mean of the final ensemble of realizations gets closer to the reference field for the parallel flow case than for the radial flow one, as can be observed by looking at the channels identified in Fig. 5, their width is more uniform and their extent go through the entire aquifer for parallel flow. As already mentioned, the local ensemble variance (Fig. 6) show the highest values for the cells close to the prescribed head boundary and extending through the shale area in the parallel flow case as it would be expected given the lack of sensitivity of head to conductivity changes at those locations.

There is not a significantly different behavior on the evolution of the RMSE and the ES (Fig. 7) for the two sets of boundary conditions. Just notice that the final values of both metrics tend to converge to the same value after localization for both flow conditions, while prior to localization they are different (larger RMSE and smaller ES for radial flow than for parallel flow).

Regarding the reproduction of the connectivity, it is apparently more difficult to reproduce it for radial flow case than for parallel flow; for radial flow, there seems to be an important filter inbreeding that leaves the reference connectivity function outside the cloud of individual functions; apart from that, there is no significant difference on the performance of the NS-EnKF in characterizing the connectivity under radial or parallel flow conditions.

Finally, regarding the reproduction of the piezometric heads (Fig. 9), the main difference between parallel and radial flow is that for radial flow the overall reduction of RMSE with regard to the unconditional realizations is smaller than for parallel flow, but this is easily understandable since the piezometric head fluctuations with respect to the initial state are stronger for radial flow than for parallel due to the several pumping and recharge wells used for the radial flow scenarios.

We conclude that the NS-EnKF performs similarly well for either radial flow or parallel flow conditions. However, from the analysis of the  $RMSE(\ln K)$  and  $ES(\ln K)$ , the radial flow case apparently is more prone to filter inbreeding than the parallel flow case. Inbreeding that disappears after localizing the filter.

### 4.4 Effect of number of conditioning piezometers

Scenarios 1 to 12 have been analyzed considering that there are 111 piezometers (Fig. 3) providing dynamic hydraulic

head measurements during the first 66.7 days (first 60 simulation time steps), which are assimilated through the NS-EnKF to update an ensemble of hydraulic conductivity realizations. For this number of piezometers, the NS-EnKF seems to perform reasonably well for a variety of scenarios, including the case in which an incorrect prior random function model is used. Since there are no logconductivity data, the information provided by the piezometer is crucial for the identification of the hydraulic conductivity spatial patterns. We decided to reduce the number of piezometers to one third of the original ones and keep only 32 of them (Fig. 10a) to analyze two more scenarios, under parallel flow and without localization; scenario 13 assumes the correct prior random function model and scenario 14 the wrong one.

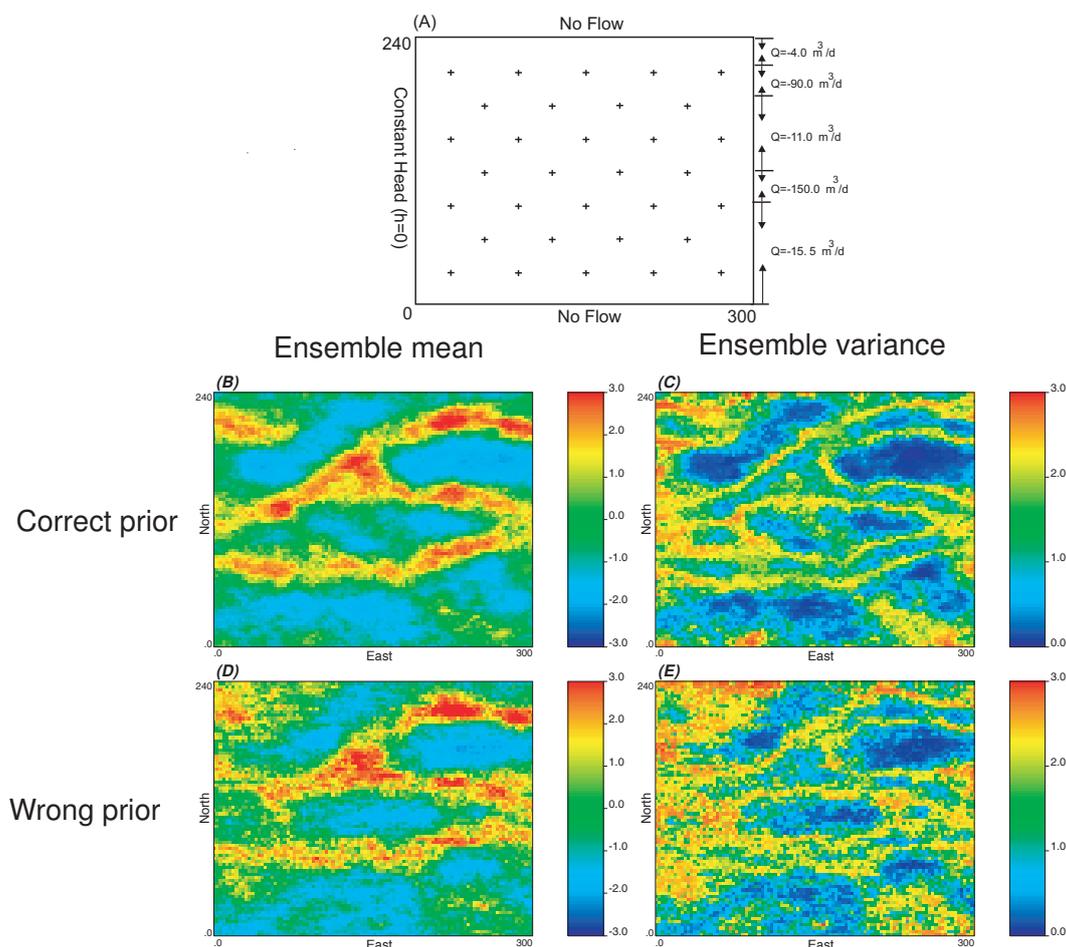
Maps B and D in Fig. 10 show the ensemble mean of logconductivity and maps C and E the ensemble variance for the correct and wrong prior, respectively. As expected, the characterization of the hydraulic conductivity field is worse and the ensemble variance is larger when fewer piezometers are used; however, the characterization is not bad. The overall patterns are well captured and the highest local variances occur at channel borders, with a much better performance for the correct prior random function than for the wrong one. Although not shown, the connectivity in scenarios 13 and 14 is clearly deteriorated when the number of piezometers goes down to 32.

We conclude that the number of piezometers is important for the characterization of the logconductivity, particularly when the prior random function is wrong.

### 4.5 Discussion

This paper presented a performance assessment of the NS-EnKF, a real-time data assimilation algorithm, in a two dimensional bimodal aquifer, and focused on four aspects: (1) the impact of random function choice; (2) the effect of introducing a localization function (3) the flow regimes and (4) the number of piezometers used during assimilation.

The results show that assimilating a sufficiently large number of transient piezometric head data with the NS-EnKF allows a good characterization of non-multiGaussian patterns such as those depicted by channelized aquifers, even if the prior geostatistical model is erroneous and based on two-point indicator variograms. This result may have important implications for groundwater modeling studies, since it implies that a sufficiently large number of piezometers may offset the negative impact of an erroneous choice of the random function model used to generate the starting ensemble of conductivity realizations. For those who are not keen on using training images to characterize the spatial variability of the hydraulic conductivity, a variogram-based simulation algorithm may serve as the starting point, leaving to the NS-EnKF the task to identify the curvilinear features that the variogram-based algorithm cannot.



**Fig. 10.** (A) flow configuration and spatial distribution of sampled head data. (B) the ensemble mean of logconductivity with correct prior model. (C) the ensemble variance of logconductivity with correct prior model. (D) the ensemble mean of logconductivity with wrong prior model. (E) the ensemble variance of logconductivity with wrong prior model.

Although it might be argued that we had access to the “correct” prior histogram in the wrong prior analysis, if the contrast in conductivities between the channel and non-channel facies is large enough, the really important a priori information that would be needed would be the proportions of each facies and the contrast between their average conductivities.

Our results are more optimistic than the ones from Kerrou et al. (2008), who concluded that an erroneous prior geostatistical model cannot be corrected by inverse modeling, even if many hydraulic head data are available. We attribute this impossibility to the inverse modeling method utilized and claim that the NS-EnKF can help to produce an acceptable characterization of the hydraulic conductivities even when the prior random function model is incorrect.

We also demonstrate the importance of coupling the NS-EnKF with a distance-dependent localization function, even for a case such as this in which the ensemble consisted of 1000 realizations. A future research topic is the improvement of the localization functions, allowing for anisotropy

and non-stationarity (i.e. time and/or space dependence), and with the option to have different formulations for piezometric head covariances and piezometric-conductivity cross-covariances.

Apparently, the NS-EnKF works equally well for both parallel and radial flow conditions. And, it is clear that for a case such as this one in which the conductivity identification is based solely on the information conveyed by the dynamic piezometric head data, the number of piezometers is very important, more so, when the prior random function model is wrong.

## 5 Conclusions

The complexity of subsurface geological structures calls for a characterization technique, which can incorporate the conductivity measurement data as well as the responses of aquifers in an efficient and effective way. The NS-EnKF, a

real-time data assimilation technique which couples EnKF with normal score transformation, presented by Zhou et al. (2011). The method avoids the classical multiGaussianity inherent to most simulation algorithms.

In this paper, we have analyzed 14 scenarios to present a detailed analysis of the impact of (i) prior model choice, (ii) use of localization functions, (iii) flow regime, and (iv) number of piezometers used, on the performance of the NS-EnKF algorithm in a synthetic 2-D bimodal aquifer. The following conclusions can be drawn from the simulation exercises:

- The NS-EnKF performs relatively well when an incorrect prior geostatistical model is used. Results are worse than when the correct prior is used, but acceptable.
- Coupling the NS-EnKF with a distance-dependent localization function improves both the characterization of conductivity and the prediction of groundwater flow.
- The performance of NS-EnKF was similar for two different flow scenarios (parallel flow and radial flow), in both cases, the patterns of conductivity are identified very well.
- The quality of the characterization is directly related to the number of piezometers used, more so when a wrong a priori model is used.

*Acknowledgements.* The authors gratefully acknowledge the financial support by the Spanish Ministry of Science and Innovation through project CGL2011-23295. The two anonymous reviewers are gratefully acknowledged for their comments which helped improving the final version of the manuscript.

Edited by: S. Attinger

## References

- Alcolea, A. and Renard, P.: Blocking Moving Window algorithm: Conditioning multiple-point simulations to hydrogeological data, *Water Resour. Res.*, 46, W08511, doi:10.1029/2009WR007943, 2010.
- Alcolea, A., Carrera, J., and Medina, A.: Pilot points method incorporating prior information for solving the groundwater flow inverse problem, *Adv. Water Resour.*, 29, 1678–1689, 2006.
- Anderson, J.: An ensemble adjustment Kalman filter for data assimilation, *Mon. Weather Rev.*, 129, 2884–2903, 2001.
- Arulampalam, M., Maskell, S., Gordon, N., and Clapp, T.: A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking, *IEEE T. Signal Process.*, 50, 174–188, 2002.
- Bear, J.: *Dynamics of fluids in porous media*, American Elsevier Pub. Co., New York, 1972.
- Burgers, G., van Leeuwen, P., and Evensen, G.: Analysis scheme in the ensemble Kalman filter, *Mon. Weather Rev.*, 126, 1719–1724, 1998.
- Capilla, J. and Llopis-Albert, C.: Gradual conditioning of non-Gaussian transmissivity fields to flow and mass transport data: 1. Theory, *J. Hydrol.*, 371, 66–74, 2009.
- Capilla, J. E., Rodrigo, J., and Gómez-Hernández, J. J.: Simulation of non-Gaussian transmissivity fields honoring piezometric data and integrating soft and secondary information, *Math. Geol.*, 31, 907–927, 1999.
- Carrera, J., Alcolea, A., Medina, A., Hidalgo, J., and Slooten, L.: Inverse problem in hydrogeology, *Hydrogeol. J.*, 13, 206–222, 2005.
- Chen, Y. and Oliver, D.: Cross-covariances and localization for EnKF in multiphase flow data assimilation, *Computational Geosciences*, 1–23, 2010.
- Chen, Y. and Zhang, D.: Data assimilation for transient flow in geologic formations via ensemble Kalman filter, *Adv. Water Resour.*, 29, 1107–1122, 2006.
- Deutsch, C. V. and Journel, A. G.: *GSLIB, Geostatistical Software Library and User's Guide*, 2nd Edn., Oxford University Press, New York, 1998.
- Devegowda, D., Arroyo-Negrete, E., and Datta-Gupta, A.: Flow relevant covariance localization during dynamic data assimilation using EnKF, *Adv. Water Resour.*, 33, 129–145, 2010.
- Evensen, G.: The ensemble Kalman filter: Theoretical formulation and practical implementation, *Ocean Dynam.*, 53, 343–367, 2003.
- Fernández-García, D., Trinchero, P., and Sanchez-Vila, X.: Conditional stochastic mapping of transport connectivity, *Water Resour. Res.*, 46, W10515, doi:10.1029/2009WR008533, 2010.
- Feyen, L. and Caers, J.: Multiple-point geostatistics: a powerful tool to improve groundwater flow and transport predictions in multi-modal formations, *Geostatistics for Environmental Applications*, Springer, Berlin, Heidelberg, New York, 197–208, 2005.
- Fu, J. and Gómez-Hernández, J. J.: A Blocking Markov Chain Monte Carlo Method for Inverse Stochastic Hydrogeological Modeling, *Math. Geosci.*, 41, 105–128, 2009.
- Gaspari, G. and Cohn, S.: Construction of correlation functions in two and three dimensions, *Q. J. Roy. Meteorol. Soc.*, 125, 723–757, 1999.
- Gómez-Hernández, J. J. and Journel, A. G.: Joint sequential simulation of multi-Gaussian fields, *Geostat. Troia*, 92, 85–94, 1993.
- Gómez-Hernández, J. J. and Srivastava, R. M.: ISIM3D: An ANSI-C Three Dimensional Multiple Indicator Conditional Simulation Program, *Comput. Geosci.*, 16, 395–440, 1990.
- Gómez-Hernández, J. and Wen, X.: To be or not to be multi-Gaussian? A reflection on stochastic hydrogeology, *Adv. Water Resour.*, 21, 47–61, 1998.
- Gómez-Hernández, J. J., Sahuquillo, A., and Capilla, J. E.: Stochastic Simulation of Transmissivity Fields Conditional to Both Transmissivity and Piezometric Data, 1, Theory, *J. Hydrol.*, 203, 162–174, 1997.
- Goovaerts, P.: *Geostatistics for natural resources evaluation*, Oxford University Press, USA, 1997.
- Hamill, T., Whitaker, J., and Snyder, C.: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter, *Mon. Weather Rev.*, 129, 2776–2790, 2001.
- Harbaugh, A. W., Banta, E. R., Hill, M. C., and McDonald, M. G.: *MODFLOW-2000, the U.S. Geological Survey modular groundwater model*, U.S. Geological Survey, Branch of Information Services, Reston, VA, Denver, CO, 2000.
- Hendricks Franssen, H. and Kinzelbach, W.: Real-time groundwater flow modeling with the Ensemble Kalman Filter: Joint estimation of states and parameters and the filter

- inbreeding problem, *Water Resour. Res.*, 44, W09408, doi:10.1029/2007WR006505, 2008.
- Hendricks Franssen, H. and Kinzelbach, W.: Ensemble Kalman filtering versus sequential self-calibration for inverse modelling of dynamic groundwater flow systems, *J. Hydrol.*, 365, 261–274, 2009.
- Hendricks Franssen, H., Gómez-Hernández, J., and Sahuquillo, A.: Coupled inverse modelling of groundwater flow and mass transport and the worth of concentration data, *J. Hydrol.*, 281, 281–295, 2003.
- Hendricks Franssen, H., Alcolea, A., Riva, M., Bakr, M., van der Wiel, N., Stauffer, F., and Guadagnini, A.: A comparison of seven methods for the inverse modelling of groundwater flow. Application to the characterisation of well catchments, *Adv. Water Resour.*, 32, 851–872, 2009.
- Hendricks Franssen, H. H. P., Kaiser, U. K., Bauser, G., Stauffer, F., Mueller, R., and Kinzelbach, W.: Operational real-time modeling with EnKF of variably saturated subsurface flow including stream-aquifer interaction and parameter updating, *Water Resour. Res.*, 47, W02532, doi:10.1029/2010WR009480, 2011.
- Hu, L. Y.: Gradual Deformation and Iterative Calibration of Gaussian-Related Stochastic Models, *Math. Geol.*, 32, 87–108, 2000.
- Huang, C., Hu, B. X., Li, X., and Ye, M.: Using data assimilation method to calibrate a heterogeneous conductivity field and improve solute transport prediction with an unknown contamination source, *Stoch. Environ. Res. Risk Assess.*, 23, 1155–1167, 2008.
- Kerrou, J., Renard, P., Hendricks Franssen, H., and Lunati, I.: Issues in characterizing heterogeneity and connectivity in non-multiGaussian media, *Adv. Water Resour.*, 31, 147–159, 2008.
- Knudby, C. and Carrera, J.: On the relationship between indicators of geostatistical, flow and transport connectivity, *Adv. Water Resour.*, 28, 405–421, 2005.
- LaVenue, A. M., Ramarao, B. S., de Marsily, G., and Marietta, M. G.: Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields, 2, Application, *Water Resour. Res.*, 31, 495–516, 1995.
- Lee, S., Carle, S., and Fogg, G.: Geologic heterogeneity and a comparison of two geostatistical models: Sequential Gaussian and transition probability-based geostatistical simulation, *Adv. Water Resour.*, 30, 1914–1932, 2007.
- Li, L., Zhou, H., and Gómez-Hernández, J. J.: Steady-state groundwater flow modeling with full tensor conductivities using finite differences, *Comput. Geosci.*, 36, 1211–1223, doi:10.1016/j.cageo.2010.04.002, 2010.
- Li, L., Zhou, H., and Gómez-Hernández, J. J.: A comparative study of three-dimensional hydraulic conductivity upscaling at the Macrodispersion Experiment (MADE) site, Columbus Air Force Base, Mississippi (USA), *J. Hydrol.*, 404, 278–293, doi:10.1016/j.jhydrol.2011.05.001, 2011a.
- Li, L., Zhou, H., and Gómez-Hernández, J. J.: Transport Upscaling Using Multi-rate Mass Transfer in Three-dimensional Highly Heterogeneous Porous Media, *Adv. Water Resour.*, 34, 478–489, doi:10.1016/j.advwatres.2011.01.001, 2011b.
- Li, L., Zhou, H., Gómez-Hernández, J. J., and Hendricks Franssen, H.: Jointly Mapping Hydraulic Conductivity and Porosity by Assimilating Concentration Data via Ensemble Kalman Filter, *J. Hydrol.*, doi:10.1016/j.jhydrol.2012.01.037, in press, 2012a.
- Li, L., Zhou, H., Hendricks Franssen, H., and Gómez-Hernández, J. J.: Modeling transient flow by coupling Ensemble Kalman Filtering and upscaling, *Water Resour. Res.*, 48, W01537, doi:10.1029/2010WR010214, 2012b.
- Mariethoz, G., Renard, P., and Caers, J.: Bayesian inverse problem and optimization with iterative spatial resampling, *Water Resour. Res.*, 46, W11530, doi:10.1029/2010WR009274, 2010a.
- Mariethoz, G., Renard, P., and Straubhaar, J.: The Direct Sampling method to perform multiple-point geostatistical simulations, *Water Resour. Res.*, 46, W11536, doi:10.1029/2008WR007621, 2010b.
- McLaughlin, D. and Townley, L.: A reassessment of the groundwater inverse problem, *Water Resour. Res.*, 32, 1131–1161, 1996.
- Nan, T. and Wu, J.: Groundwater parameter estimation using the ensemble Kalman filter with localization, *Hydrogeol. J.*, 19, 1–15, 2010.
- Nowak, W.: Best unbiased ensemble linearization and the quasi-linear Kalman ensemble generator, *Water Resour. Res.*, 45, W04431, doi:10.1029/2008WR007328, 2009.
- Oliver, D., Cunha, L., and Reynolds, A.: Markov chain Monte Carlo methods for conditioning a permeability field to pressure data, *Math. Geol.*, 29, 61–91, 1997.
- Pardo-Igúzquiza, E. and Dowd, P.: CONNEC3D: a computer program for connectivity analysis of 3D random set models, *Comput. Geosci.*, 29, 775–785, 2003.
- Ramarao, B. S., LaVenue, A. M., de Marsily, G., and Marietta, M. G.: Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields, 1, Theory and computational experiments, *Water Resour. Res.*, 31, 475–493, 1995.
- Reichle, R., Walker, J., Koster, R., and Houser, P.: Extended versus ensemble Kalman filtering for land data assimilation, *J. Hydrometeorol.*, 3, 728–740, 2002.
- Sahuquillo, A., Capilla, J. E., Gómez-Hernández, J. J., and Andreu, J.: Conditional simulation of transmissivity fields honouring piezometric head data, in: *Hydraulic Engineering Software IV, Fluid Flow Modeling*, edited by: Blair, W. R. and Cabrera, E., vol. II, Elsevier Applied Science, London, UK, 201–214, 1992.
- Strebelle, S.: Conditional simulation of complex geological structures using multiple-point statistics, *Math. Geol.*, 34, 1–21, 2002.
- Sun, A. Y., Morris, A. P., and Mohanty, S.: Sequential updating of multimodal hydrogeologic parameter fields using localization and clustering techniques, *Water Resour. Res.*, 45, W07424, doi:10.1029/2008WR007443, 2009.
- Wen, X. and Gómez-Hernández, J.: Numerical modeling of macrodispersion in heterogeneous media: a comparison of multi-Gaussian and non-multi-Gaussian models, *J. Contam. Hydrol.*, 30, 129–156, 1998.
- Wen, X., Deutsch, C., and Cullick, A.: Construction of geostatistical aquifer models integrating dynamic flow and tracer data using inverse technique, *J. Hydrol.*, 255, 151–168, 2002.
- Wen, X. H. and Chen, W.: Real-time reservoir model updating using ensemble Kalman filter, in: *SPE reservoir simulation symposium*, 2005.
- Western, A., Blöchl, G., and Grayson, R.: Toward capturing hydrologically significant connectivity in spatial patterns, *Water Resour. Res.*, 37, 83–97, 2001.

- Wu, J., Boucher, A., and Zhang, T.: A SGeMS code for pattern simulation of continuous and categorical variables: FILTERSIM, *Comput. Geosci.*, 34, 1863–1876, 2008.
- Yeh, W.: Review of parameter identification procedures in groundwater hydrology: The inverse problem, *Water Resour. Res.*, 22, 95–108, 1986.
- Zhou, H., Li, L., and Gómez-Hernández, J. J.: Three-dimensional hydraulic conductivity upscaling in groundwater modelling, *Comput. Geosci.*, 36, 1224–1235, doi:10.1016/j.cageo.2010.03.008, 2010.
- Zhou, H., Gómez-Hernández, J. J., Hendricks Franssen, H., and Li, L.: An approach to handling non-Gaussianity of parameters and state variables in Ensemble Kalman Filtering, *Adv. Water Resour.*, 34, 844–864, doi:10.1016/j.advwatres.2011.04.014, 2011.
- Zhou, H., Gómez-Hernández, J. J., and Li, L.: A pattern search based inverse method, *Water Resour. Res.*, doi:10.1029/2011WR011195, in press, 2012a.
- Zhou, H., Gómez-Hernández, J. J., and Li, L.: Inverse methods in hydrogeology: trajectory and future trends, *J. Hydrol.*, in review, 2012b.
- Zhou, H., Li, L., Gómez-Hernández, J. J., and Hendricks Franssen, H.: Pattern identification in bimodal aquifer by normal score Ensemble Kalman Filter, *Math. Geosci.*, 44, 169–185, doi:10.1007/s11004-011-9372-3, 2012c.
- Zimmerman, D., De Marsily, G., Gotway, C., Marietta, M., Axness, C., Beauheim, R., Bras, R., Carrera, J., Dagan, G., Davies, P., Gallegos, P., Galli, A., Gomez-Hernandez, J., Grindrod, P., Gutjahr, A., Kitanidis, P., Lavenue, A., McLaughlin, D., Neuman, S., RamaRao, B., Ravenne, C., and Rubin, Y.: A comparison of seven geostatistically based inverse approaches to estimate transmissivities for modeling advective transport by groundwater flow, *Water Resour. Res.*, 34, 1373–1413, 1998.
- Zinn, B. and Harvey, C. F.: 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.*, 39, 1051, doi:10.1029/2001WR001146, 2003.