Contaminant Spill in a Sandbox with Non-Gaussian Conductivities: Simultaneous Identification by the Restart Normal-Score Ensemble Kalman Filter

Zi Chen[1] · Teng Xu[2] ·
J. Jaime Gómez-Hernández[1] · Andrea Zanini[3]

Abstract The joint identification of the parameters defining a contaminant source and the heterogeneous distribution of the hydraulic conductivities of the aquifer where the contamination took place is a difficult task. Previous studies have demonstrated the applicability of the restart normal-score ensemble Kalman filter (rNS-EnKF) in synthetic cases making use of sufficient hydraulic head and concentration data. This study shows an application of the same technique to a non-synthetic case under laboratory conditions and discusses the difficulties found on its application and the avenues taken to solve them. The method is first tested using a synthetic case that mimics the sandbox experiment to establish the minimum number of ensemble members and the best technique to prevent the filter collapsing. The synthetic case shows that among different techniques based on update damping and covariance inflation, the Bauser's covariance inflation method works best in preventing filter collapse. Its application to the sandbox data shows that the rNS-EnKF can benefit from Bauser's inflation to reduce

✉ Teng Xu
   teng.xu@hhu.edu.cn

   Zi Chen
   dtpenguincz@gmail.com

   J. Jaime Gómez-Hernández
   jgomez@upv.es

   Andrea Zanini
   andrea.zanini@unipr.it

[1] Institute of Water and Environmental Engineering, Universitat Politècnica de València, Valencia, Spain

[2] State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing, China

[3] Department of Engineering and Architecture, University of Parma, Parma, Italy

the number of ensemble realizations substantially in comparison with a filter without inflation, arriving at a good joint identification of both the contaminant source and the spatial heterogeneity of the conductivities.

Keywords Inverse modeling · Forensic hydrogeology · Data assimilation · Sandbox

1 Introduction

The motivation of this paper is to advance our ability to solve the problem of the joint identification of a contaminant source in an aquifer together with the spatial distribution of hydraulic conductivities. The restart normal-score Ensemble Kalman filter (rNS-EnKF) has been tested in synthetic aquifers for the joint identification of source parameters and conductivities and in a sandbox experiment for the identification of just the source parameters (Chen et al. 2018; Xu and Gómez-Hernández 2018). In both cases, the rNS-EnKF performed well; however, it could be argued that the synthetic case was far from reality, and that the sandbox experiment used a known homogeneous conductivity. For these reasons, a new sandbox experiment was designed with a binary heterogeneous distribution of conductivity, and with the aim of testing the rNS-EnKF for the joint identification of the source and a spatially heterogeneous conductivity field.

In addition, previous experience in the application of the rNS-EnKF (Xu et al. 2013) showed the effect of filter collapse, a problem that can be tackled by the proper choice of number of ensemble realizations, covariance inflation, covariance localization or update damping. For this reason, the paper starts with the analysis of a synthetic field, resembling the new sandbox experiment, to determine the choice of number of realizations, the technique that prevents the filter from collapsing and yields an acceptable identification of both source and conductivities within reasonable computer times. Once these choices are made, the sandbox experiment is directly addressed.

The importance of contaminant source identification, for instance in relation with the protection of wellhead capture zones (Feyen et al. 2003a, b), does not need to be stressed, as it has been the subject of research for many years. The reader is referred to any of the review papers that can be found in the literature (e.g., Atmadja and Bagtzoglou 2001; Bagtzoglou and Atmadja 2005; Michalak and Kitanidis 2004; Sun et al. 2006). A very brief review, including some works that appeared after the mentioned review papers, follows.

Most contaminant source identification approaches can be classified into two main categories: optimization approaches and probabilistic approaches. In the optimization approaches, an objective function is built and the algorithm tries to minimize the discrepancies between simulated and measured concentrations using an optimization approach such as least-squares regression or maximum likelihood (e.g., Amirabdollahian and Datta 2014; Aral et al. 2001; Ayvaz 2016; Gorelick et al. 1983; Mirghani et al. 2009; Wagner 1992; Yeh et al. 2007). In the probabilistic approaches, the problem is cast in a stochastic framework and the algorithm tries to maximize the posterior probabilities of the simulated concentrations conditioned on the observed values using techniques such as those based on minimum relative entropy or the use of adjoint states

(e.g., Bagtzoglou et al. 1992; Butera et al. 2013; Koch and Nowak 2016; Neupauer and Wilson 1999; Woodbury and Ulrych 1996).

The main criticism of both approaches that can be found in the literature, and the reason it is difficult to find applications of any of those techniques in practice, is that they have worked on synthetic cases, focusing on the identification of the contaminant source parameters and assuming that aquifer hydraulic conductivities are perfectly known. But the truth is that geological properties are quite heterogeneous, only sparsely known in reality, and very influential in how the aquifer behaves (e.g., Gómez-Hernández and Wen 1994; Gómez-Hernández and Wen 1998; Knudby and Carrera 2005; Li et al. 2011; Wen et al. 1999; Zinn and Harvey 2003). Only a few papers discuss the simultaneous identification of conductivity and the contaminant source, but almost all of them are limited to either homogeneous aquifers or offer only a simplistically-described heterogeneity (Datta et al. 2009; Mahar and Datta 2000; Wagner 1992). Only the works by Koch and Nowak (2016) and Xu and Gómez-Hernández (2018) address the problem of identifying heterogeneous conductivities; the former using a Bayesian methodology, and the later using the rNS-EnKF.

This paper builds on the previous work by Chen et al. (2018) and (Xu and Gómez-Hernández 2016, 2018) in which the capabilities of the rNS-EnKF, for the purpose of the identification of the parameters defining a point contaminant source and the aquifer hydraulic conductivities, had been shown in both a synthetic case and in a laboratory experiment, and on the experience of the research team on addressing the problem of characterization of non-Gaussian conductivities (Capilla et al. 1999; Franssen and Gómez-Hernández 2002; Gómez-Hernández et al. 2003; Journel et al. 1993; Zhou et al. 2012a, b). The goal of this paper is to advance towards a practical application of the rNS-EnKF for contaminant source identification in an aquifer with sparse information about hydraulic conductivity heterogeneity. In comparison with previous papers, this paper works with data collected in a sandbox experiment, instead of with generated synthetic data, and the sandbox has a binary heterogeneous distribution (unknown to the algorithm), instead of a known homogeneous distribution. There is an additional important difference with respect to the work by Xu and Gómez-Hernández (2018), which is that no piezometric head data are available, and, therefore, the parameter identification will have to be solely based on concentration observations. This adds an additional complication to the performance of the rNS-EnKF, since an important source of information for conductivity heterogeneity identification will be missing.

In an initial attempt to apply the rNS-EnKF directly to the sandbox data, numerous problems were found related with compute time, filter collapse and filter divergence. For this reason, a decision was taken to first analyze a more controlled synthetic experiment mimicking the heterogenous sandbox, then to decide on the number of realizations and the best technique for preventing filter collapse without compromising the results (in a reasonable time, with a reasonable uncertainty). As a result, the paper contains two case studies: (i) the synthetic case, in which a sensitivity analysis is performed combining two numbers of realizations, two update damping schemes and two covariance inflation approaches, out of which the number of ensemble realizations and a filter collapse prevention technique are chosen; and (ii) the laboratory case, in which the rNS-EnKF is demonstrated using the findings from the synthetic case.

Filter collapse is dealt with by the use of covariance inflation. Several such techniques can be found in the literature (e.g., Anderson 2007; Li et al. 2009; Liang et al. 2012; Bauser et al. 2018; Hendricks Franssen and Kinzelbach 2008; Wang and Bishop 2003; Zheng 2009), of which the damping method, Wang's method and Bauser's method will be tested. These methods will be discussed in detail further on in the corresponding section.

The paper shows the power of concentration data for the joint identification of conductivities and contaminant source information in a sandbox experiment by the rNS-EnKF. After this introductory review, the paper continues with a review of the methodology and a description of the sandbox experiment and its numerical modeling, followed by the synthetic data analysis and the sandbox data analysis. The paper ends with the discussion of the results and some conclusions.

2 Methodology

2.1 Groundwater Flow and Solute Transport Equations

Water flow and contaminant transport in the sandbox are modeled using the corresponding governing equations for groundwater flow (Bear 1972) and contaminant transport (Zheng and Wang 1999)

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

$$\frac{\partial (\theta C)}{\partial t} = \nabla \cdot (\theta D \cdot \nabla C) - \nabla \cdot (\theta v C) - q_s C_s, \tag{2}$$

where S_s is specific storage $[L^{-1}]$, h is hydraulic head $[L]$, t is time $[T]$, $\nabla \cdot$ is the divergence operator, ∇ is the gradient operator, K is hydraulic conductivity $[LT^{-1}]$ and w represents distributed sources or sinks $[T^{-1}]$; θ is porosity; C is dissolved concentration $[ML^{-3}]$; D is the hydrodynamic dispersion tensor $[L^2T^{-1}]$; v is the flow velocity vector $[LT^{-1}]$ derived from the solution of the flow equation, q_s represents volumetric flow rate per unit volume of the aquifer associated with a fluid source or sink $[T^{-1}]$ and C_s is the concentration of the source or sink $[ML^{-3}]$.

The groundwater flow equation is numerically solved with MODFLOW (McDonald and Harbaugh 1988) and the contaminant transport equation with MT3DS (Zheng and Wang 1999).

2.2 The Ensemble Kalman Filter

The ensemble Kalman filter (EnKF) was developed by Evensen (1994) as an extension to the Kalman filter (KF). The main difference between the EnKF and the KF is that, in the KF, the state covariance matrix is propagated in time using an explicit expression based on a linear transition equation, while, in the EnKF, this covariance matrix is derived from the statistical analysis of an ensemble of state realizations obtained after the solution of the state equations in each realization of the ensemble. The advantage of

the EnKF over the KF is for systems in which the state transition equation is not linear; in such a case, the linear transition equation used by the KF is only an approximation and the resulting covariance deteriorates in time. By contrast, in the EnKF, since the covariance is directly calculated from actual state spatial distributions, its value is more accurate; the only limitation being that the covariance is computed from a finite ensemble of realizations (if the number of realizations is small, the resulting estimate may be also inaccurate).

Although the EnKF was initially developed to update only the state of the system as observations are gathered, it has been shown that it also can be used to update the parameters using what is called an augmented state that includes both state variables and the parameters that control them (e.g., Chen and Zhang 2006; Houtekamer and Mitchell 2001; Li et al. 2012a, c). In summary, the EnKF has been proven to be an efficient algorithm for parameter identification, for strongly non-linear state-transfer equations, (Hendricks Franssen and Kinzelbach 2009; Wen and Chen 2005a, b), and has received much attention in the last decades. Next, the algorithm is described for the case study at hand, that is, the identification of the parameters defining a contaminant source together with the identification of the conductivities in a sandbox experiment for which only concentration data are available.

First, build an augmented state vector S including the model parameters and the state variables

$$S = \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} (X_s, Z_s, I_c, I_r, T_e)^T \\ (ln K_1, ln K_1, \ldots, ln K_N)^T \\ (C_1, C_1, \ldots, C_N)^T \end{pmatrix}, \tag{3}$$

where A stands for model parameters, B for state variables, and N is the number of grid cells. In our case, the model parameters are those describing the contaminant source, X_s, Z_s, which are the contaminant source coordinates in the horizontal and vertical directions, I_c, the injection concentration, I_r, the injection rate, and T_e, the end release time, plus the hydraulic log-conductivities, $ln K$, and the state variables are the contaminant concentrations, C. The augmented state vector evolves in time, starting with an initial value at time 0, S_0.

Second, forecast, using the groundwater flow and transport equations, the state vector S_t at time t based on the state variable B_{t-1} and the model parameters A_{t-1} obtained at time $t-1$

$$S_t^f = \psi(A_{t-1}^a, B_{t-1}^a), \tag{4}$$

where the superscript f stands for forecasted values and a stands for updated values after assimilating the state observations; ψ represents the state-transfer function. (In the forecast step, the parameters A remain unchanged—the transfer function is the identity function—and state B evolves according to the flow and transport equations.)

Next, assimilate the state observations. The discrepancy between forecasted states and observed ones is used to update the forecasted augmented state vector according to the following expression

$$S_t^a = S_t^f + \mathbf{K}_t \left[y_t^{obs} + \varepsilon_i - \mathbf{H} S_t^f \right], \tag{5}$$

where y_t^{obs} are the observed concentrations at time step t, ε_i stands for an observation error with zero mean and covariance $\mathbf{R}_t$, $\mathbf{H}$ is the observation matrix that extracts out of the whole augmented state vector the elements at which observations were taken. $\mathbf{K}_t$ is the Kalman gain matrix

$$\mathbf{K}_t = \mathbf{P}_t^f \mathbf{H}^T [\mathbf{H} \mathbf{P}_t^f \mathbf{H}^T + \mathbf{R}_t]^{-1} \tag{6}$$

$$\mathbf{P}_t^f = \frac{1}{N_e - 1} \{ [S_t^f - \overline{S_t^f}][S_t^f - \overline{S_t^f}]^T \}, \tag{7}$$

where $\mathbf{P}_t^f$ is the experimental covariance computed from the ensemble of augmented forecasted states, and $\overline{S_t^f}$ is the experimental ensemble mean. (Notice that because observations are sparse, the observation matrix is mostly made out of zeroes, and it is not necessary to compute all the elements in $\mathbf{P}_t^f$, but only those that are multiplied by the non-zero elements of $\mathbf{H}$ in $\mathbf{P}_t^f \mathbf{H}^T$.)

2.2.1 The normal-Score EnKF

The EnKF was further extended to deal with non-Gaussian variables. The EnKF was found to be very effective for dealing with non-linear transfer functions, but it failed when the augmented state followed a non-Gaussian distribution (Zhou et al. 2014). Several approaches have been developed to address this issue: Gaussian mixture models, reparameterizations, iterative approaches, and Gaussian anamorphosis, also known as normal-score transform (e.g., Chang et al. 2010; Hendricks Franssen and Kinzelbach 2008; Kumar and Srinivasan 2019; Sun et al. 2009; Zhou et al. 2011). In this paper, the normal-score approach is used, and more precisely, the normal-score EnKF (NS-EnKF) as described by Zhou et al. (2011) or Li et al. (2012b).

The NS-EnKF is based on transforming all parameters and variables into Gaussian variates, performing EnKF in the Gaussian space, and then, backtransforming the results into the original space. The normal-score transform is a univariate transform that ensures that the transformed variates follow a Gaussian distribution, but it does not ensure that higher-order moments will follow a multi-Gaussian distribution (Jafarpour and Khodabakhshi 2011; Kumar and Srinivasan 2020); yet, the results obtained with the NS-EnKF outperform those of EnKF for clearly non-Gaussian parameters.

2.2.2 The Restart NS-EnKF

The EnKF was designed to update both parameters and state variables at each assimilation step. That is, the discrepancy between forecasted and observed variables is used to update the whole augmented state (see Eq. (5)). However, in general in the case of subsurface flow and transport, and in particular in the case at hand of contaminant source identification, the updated states could be inconsistent with the updated parameters, either because the mass conservation laws are no longer obeyed, or because

the updated state is not coherent with the updated contaminant source location. For this reason, the forecast of the augmented state to the next observation time is not done based on the updated augmented state at the previous time state; it is preferable to perform a forecast from time zero with the latest updated parameters (Camporese et al. 2011; Crestani et al. 2012; Wen and Chen 2005a). This approach is called, for this reason, the restart ensemble Kalman filter, or, in our case, the restart normal-score ensemble Kalman filter (rNS-EnKF).

The forecast function in Eq. (4) changes into

$$S^f(t) = \psi[A^a_{t-1}, B_0] = \begin{pmatrix} A^a_{t-1} \\ B_t \end{pmatrix}, \tag{8}$$

where B_0 stands for the initial contaminant concentration in the domain. The restart EnKF has been applied before, for instance, by Camporese et al. (2011) and Crestani et al. (2012).

2.2.3 Damping

One way to deal with filter collapse is to use a damping factor α, between 0 and 1, at the update step (Hendricks Franssen and Kinzelbach 2008)

$$S^a_t = S^f_t + \alpha \mathbf{K}_t \left[y^{obs}_t + \varepsilon_i - \mathbf{H} S^f_t \right]. \tag{9}$$

2.2.4 Inflation Methods

Another way to reduce filter collapse is by covariance inflation. There are several covariance inflation approaches in the literature (Anderson 2007; Bauser et al. 2018; Liang et al. 2012; Wang and Bishop 2003). In this work, two different time-dependent multiplicative covariance inflation methods are used, the one proposed by Wang and Bishop (2003) and the one by Bauser et al. (2018). In both methods, the augmented state vector should be inflated after the forecast, as follows

$$S^{inf,f}_{i,t} = \sqrt{\lambda_t}(S^f_{i,t} - \overline{S^f_t}) + \overline{S^f_t}, \tag{10}$$

where $S^{inf,f}_{i,t}$ is the inflated augmented state vector of realization i after forecast to t^{th} time step, and λ_t is the inflation factor, the computation of which depends on the approach used.

In the work by Wang and Bishop (2003), λ_t is given by

$$\lambda_t = \frac{\left(\mathbf{R}_t^{-\frac{1}{2}} d_t\right)^T \mathbf{R}_t^{-\frac{1}{2}} d_t - n_b}{trace\left\{ \mathbf{R}_t^{-\frac{1}{2}} H P_t^f \left(\mathbf{R}_t^{-\frac{1}{2}} \mathbf{H}\right)^H \right\}} \tag{11}$$

where n_b is the number of observations, and d_t is a vector with the residuals between the observation data and the mean of the forecast data at observation locations

$$d_t = y_t^{obs} - \mathbf{H} * \overline{S_t^f}. \tag{12}$$

Then, the updated augmented state vector is calculated as

$$S_{i,t}^a = S_{i,t}^{inf,f} + \lambda_t \mathbf{P}_t^f \mathbf{H}^T [\mathbf{H}\lambda_t \mathbf{P}_t^f \mathbf{H}^T + \mathbf{R}_t]^{-1} \left[y_t^{obs} + \varepsilon - \mathbf{H}S_t^{inf,f} \right]. \tag{13}$$

Wang and Bishop (2003) have already recognized that parameter λ_t could vary significantly in time, particularly at the early stages when concentrations are small everywhere. For this reason, following their recommendations, its value is restricted to be between 0.7 and 1.2.

In the work by Bauser et al. (2018), λ_t is treated as a state variable that is used to inflate the model parameters. Because it is a state variable, it is forecasted and updated using the Kalman filter formulation as follows

$$\lambda_t^f = \lambda_{t-1}^a \tag{14}$$

$$\lambda_t^a = \lambda_t^f + \mathbf{K}_{\lambda_t} [d_{\lambda_t} - h_\lambda(\lambda_t^f)], \tag{15}$$

where the superscripts f and a stand for forecasted and updated values, $\mathbf{K}_{\lambda_t}$ is the Kalman gain, d_{λ_t} is the absolute value of d_t, and $h_\lambda(\lambda_t^f)$ represents the mean residual between observation data and forecasted mean at the observation location. These values are obtained by

$$\mathbf{K}_{\lambda_t} = \mathbf{P}_{\lambda_t}^f \mathbf{H}_{\lambda_t}^T [\mathbf{H}_{\lambda_t} \mathbf{P}_{\lambda_t}^f \mathbf{H}_{\lambda_t}^T + \mathbf{R}_{\lambda_t}]^{-1} \tag{16}$$

$$(h_{\lambda_t}(\lambda_t^f))_i = [(\mathbf{R}_{\lambda_t})_{ii}]^{\frac{1}{2}}. \tag{17}$$

The covariance of the inflation parameter, $\mathbf{P}_{\lambda_t}^f$, the observation matrix $\mathbf{H}_{\lambda_t}$ and the inflation parameter observation error $\mathbf{R}_{\lambda_t}$ can be obtained from the state covariance matrix $\mathbf{P}_t^f$, the observation matrix $\mathbf{H}$ and the observation error covariance matrix $\mathbf{R}$ of the augmented state vector $\mathbf{S}$ by
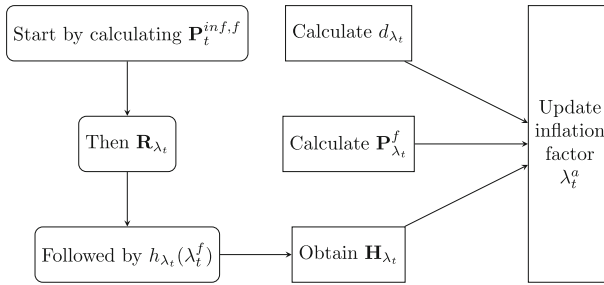
$$(\mathbf{P}_{\lambda_t}^f)_{ij} = \sigma_\lambda^2 |(\mathbf{P}_t^f)_{ij}| [(\mathbf{P}_t^f)_{ii} (\mathbf{P}_t^f)_{jj}]^{-\frac{1}{2}} \tag{18}$$

$$(\mathbf{H}_{\lambda_t})_{ij} = [2[(\lambda_t^f)_j]^{\frac{1}{2}} (h_{\lambda_t}(\lambda_t^f))_i]^{-1} \sum_m (\mathbf{H})_{ij} (\mathbf{H})_{im} (\mathbf{P}_t^f)_{jm} [(\lambda_t^f)_m]^{\frac{1}{2}} \tag{19}$$

$$(\mathbf{R}_{\lambda_t})_{ij} = |(\mathbf{R})_{ij} + (\mathbf{H}\mathbf{P}_t^{inf,f}\mathbf{H}^T)_{ij}|, \tag{20}$$

where σ_λ stands for the uncertainty about the inflation factor, which, in this case, is set to one, the same value used by Bauser et al. (2018). $\mathbf{P}_t^{inf,f}$ stands for the inflated forecast error covariance matrix, which is given by

Fig. 1 A flowchart of Bauser's method for updating the inflation factors, λ_t^a

$$\mathbf{P}_t^{inf,f} = (\sqrt{\lambda_t^f}\sqrt{\lambda_t^f}^T) \cdot \mathbf{P}_t^f. \tag{21}$$

A workflow summarizing how to apply Bauser's inflation method is shown in Fig. 1.

Finally, the updated augmented state vector is computed as

$$S_{i,t}^a = S_{i,t}^{inf,f} + \mathbf{P}_t^{inf,f}\mathbf{H}^T[\mathbf{H}\mathbf{P}_t^{inf,f}\mathbf{H}^T + \mathbf{R}_t]^{-1}\left[y_t^{obs} + \varepsilon - \mathbf{H}S_t^{inf,f}\right]. \tag{22}$$

2.2.5 Localization Methods

Covariance localization (Greybush et al. 2011) is yet another technique for tackling ensemble collapse. In this case, the localization removes spurious correlations observed in the experimental covariances, that is, corrects the experimental covariance to ensure that points that should bear no correlation have zero correlation. Experimental covariances, particularly when computed from a small ensemble size, may display unwanted spurious correlations. Covariance localization is not analyzed in this manuscript because the standard techniques for removing those spurious correlations also reduce, on occasions significantly, the correlations between locations for which the attributes are correlated. Correlations between attributes are significant in the sandbox and, for this reason, it seemed more appropriate to focus only on covariance inflation techniques rather than on covariance localization techniques, without disregarding or discarding the use of localization to improve EnKF performance in another setting.

3 Sandbox Experiment

A contaminant experiment was carried out in a sandbox with sodium fluorescein as the tracer. The size of the sandbox was 120 cm by 14 cm by 70 cm. Two reservoirs with constant water levels at 62.5 cm and 60.6 cm with respect to the bottom of the sandbox were set at the upstream and downstream boundaries, respectively. (Notice that the experiment was performed with the upstream boundary on the right side of the sandbox, and all figures are represented in this way.) These two tanks define prescribed
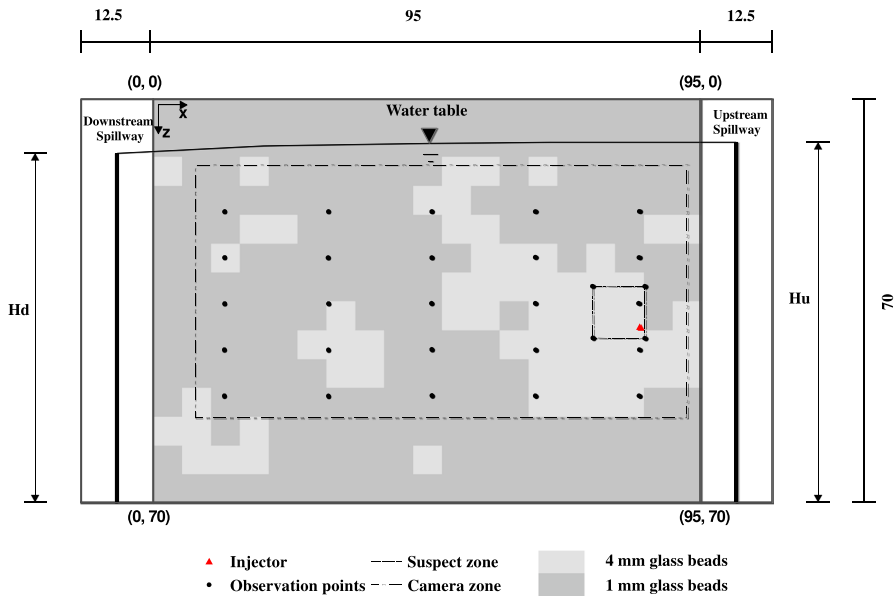
Fig. 2 Sketch of the experimental device (view from the camera side inside the darkroom). H_u and H_d stand for the constant head boundaries, the dashed rectangle corresponds to the area captured by the camera in which concentrations will be monitored, the red triangle is the release location, and the small square around the red dot indicates the release suspect location during the identification process. Units are in cm. Pairs of numbers in parenthesis refer to row and column pairs in the numerical model

head boundaries; the bottom of the sandbox was impermeable and the top boundary was the phreatic surface. Between the upstream and downstream tanks, the area filled with sand was 95 cm by 10 cm by 70 cm, which, for the purpose of modeling, is discretized into 95 columns, 1 row, and 70 layers of equal-sized cells of 1 cm by 10 cm by 1 cm. The sandbox was filled with glass beads of two different diameters, 1 mm and 4 mm, according to a spatial arrangement generated using a truncated Gaussian simulation (Journel and Isaaks 1984) with the first quartile as the truncation threshold, resulting in a large-bead proportion of 0.25. The spatial distribution of the glass beads in the sandbox can also be seen in Fig. 2. An injector was located at column 86, layer 40, at the position identified with a red dot in the figure. The whole sandbox was placed in a darkroom with a blue light source that was used to excite the injected fluorescein. Pictures of the plume, as it evolved in time, were taken and luminosity values were converted into concentration after a calibration procedure following Citarella et al. (2015).

The hydraulic properties of the beads (Table 1) had been characterized before with the same sandbox equipment (e.g., Cupola et al. 2015; Citarella et al. 2015). The hydraulic conductivity of the large beads was estimated as 10.4 cm·s^{-1}, and that of the small beads as 0.65 cm·s^{-1}. The porosity is constant, independent of the bead size, and equal to 0.37. The longitudinal dispersivity within the large beads was estimated as 0.25 cm, and within the small beads as 0.106 cm. The ratio of transverse to longitudinal dispersivity is constant and equal to 0.45.

Table 1 Parameters used in the groundwater flow and transport models

Hydr. conduct., K 4 mm beads	10.4 cm·s^{-1}
Hydr. conduct., K 1 mm beads	0.65 cm·s^{-1}
Porosity, ϕ	0.37
Long. disp., α_L 4 mm beads	0.2 cm
Long. disp., α_L 1 mm beads	0.106 cm
TRVT, α_T/α_L	0.45

Although after processing the pictures the spatial distribution of concentration is fully known within the entire central area of the sandbox (dashed rectangle in Fig. 2), in order to mimic a potential sampling campaign in the field, only the concentrations observed at the twenty nine dots identified as observation points in the figure will be used for the purpose of identifying both the hydraulic conductivity and the contaminant source parameters. The release lasted 1200 s, the fluorescein concentration was 20 mg/l and the injection rate 2.60 cm$^3 \cdot$ s^{-1}. Observations were taken every 30 s until after 3000 s from the beginning of the injection, for a total of 100 observations at each observation point.

The number of observation locations was large enough to allow us to arrive at acceptable results. Previous studies (Xu et al. 2013) have shown that there is a threshold number of observation locations below which the identification becomes impossible, due to lack of information by which to perform any identification. The number of observations and their regular pattern used here may not be realistic in a practical case, but it should always be borne in mind that without enough information, identification by the EnKF or any other approach will not be possible.

4 Definition of Scenarios and Ensemble Initialization

On a first attempt to apply the rNS-EnKF directly with the observed sandbox concentrations, some difficulties were found, mostly related to filter collapse. These difficulties led us to perform a synthetic experiment, prior to applying the filter to the real data, to analyze the impact of the number of ensemble realizations and the use of different approaches to prevent filter collapse. For this purpose, a reference set of synthetic concentrations was generated by solving, numerically, the flow and transport equations in a field with the same spatial distribution of conductivities as the sandbox, the same boundary conditions, and the same solute injection pulse. Then, six scenarios ($S1 - S6$) were analyzed with different ensemble sizes and different damping and inflation methods. More precisely, two ensemble sizes were tested (500 and 1000), two values for the damping coefficient (damping with a factor of 0.1 and with a factor of 0.5) and two covariance inflation methods (Wang's method and Bauser's method). After the analysis of the results using the synthetic reference, the conclusion was reached, as discussed below, that Bauer's inflation method was the best method to prevent filter collapse, thus two additional scenarios ($R1 - R2$) were run using the experimental data to test Bauer's inflation approach. The combination of ensemble sizes and inflation methods for the different scenarios is shown in Table 2.

Table 2 Definition of scenarios

Scenario Synthetic	Inflation method	Ensemble size
$S1$	No inflation	500
$S2$	No inflation	1000
$S3$	Damping factor = 0.1	500
$S4$	Damping factor = 0.5	500
$S5$	Wang's method	500
$S6$	Bauser's method	500
Experimental		
$R1$	No inflation	1000
$R2$	Bauser's method	500

The initial ensembles of log-conductivity realizations are the same for all scenarios (for the scenarios of 500 realizations only the first 500 of a total of 1000 realizations are retained. The choice of the first 500 is arbitrary; any subset of 500 could have been used without loss of generality). They are generated using a Gaussian random function with a mean equal to the weighted mean of the bead log-conductivities, 1.07 ln cm $\cdot$ s^{-1}, and a variance equal to the variance of a binary Gaussian mixture of two facies with the means and proportions of the sandbox and an internal variance of one within each facies, i.e., 1.55 (ln cm $\cdot$ s^{-1})^2. The correlation range of the log-conductivities is isotropic and equal to 15 cm. Previous studies (Xu et al. 2013), in which no conditioning conductivity values had been used—as in this case—have shown that the initial ensemble of log-conductivities is not as important as a sufficient number of observations of the state of the aquifer.

Similarly, the initial ensembles of source locations and pulses are the same for all scenarios. They are generated within suspect ranges that are defined using uniform distributions. The suspect source location (X_s, Z_s), in cm, ranges in U[78, 86] $\times$ U[38, 47] (see Fig. 2), the suspect injection rate ranges in U[2, 3] cm$^3 \cdot$ s^{-1}, the suspect injection concentration ranges in U[5, 25] mg/l and the suspect final release time ranges in U[1050, 1250] s (see Table 3). These parameters are generated independently of one another and of the log-conductivities. These ranges are used exclusively for the generation of the initial ensembles; afterwards, the updated parameter values are not restricted by any bounds. These ranges are chosen considering that, in a real case, there is always some information about when and where the contamination entered the system. It could be argued that the ranges should have been larger. From previous works on the impact that the choice of the initial ensemble of realizations has in the performance of the EnKF (Xu et al. 2013), it could be anticipated that the use of wider or shorter ranges would have little impact on the final results.

5 Performance Evaluation

The rNS-EnKF was applied to each scenario assimilating the observed concentrations at the points indicated in Fig. 2 at each time step. No log-conductivity or piezometric

Table 3 Suspect ranges of source parameters for the generation of the initial ensemble of realizations and their true values

Parameter	Actual value	Suspect range
Xs - x-coordinate of source (cm)	86	$78 - 87$
Zs - z-coordinate of source (cm)	40	$38 - 47$
Ir (cm$^3 \cdot$ s^{-1})—injection rate	2.60	$2 - 3$
Ic (mg/l)—injection concentration	20	$5 - 25$
Te (s)—final release time	1200	$1050 - 1250$

head data were observed at any time. After assimilating the concentration data at the end of each time step, the filter provided an ensemble of updated parameters, which were analyzed in different ways:

1. Computing the ensemble mean and variance of the contaminant source parameters at the end of each time step. The ensemble mean can be interpreted as a parameter estimate and the variance as a measure of the estimation uncertainty.
2. Visually analyzing the spatial variability of the cell ensemble mean and ensemble variance of log-conductivities with respect to the reference log-conductivity spatial distribution.
3. Computing the root mean-squared error (RMSE), the ensemble spread (ES), and the ratio RMSE/SE of log-conductivities as given by

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\ln K_i^{ref} - \overline{\ln K_i})^2}, \tag{23}$$

$$\text{ES} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \sigma^2_{\ln K_i}}, \tag{24}$$

with n being the number of cells over which the averages are computed, $\ln K_i^{ref}$ is the reference log-conductivity value at cell i, $\overline{\ln K_i}$ is the average of the ensemble of log-conductivity realizations at cell i, and $\sigma^2_{\ln K_i}$ is the variance. The RMSE measures the accuracy of the ensemble average as an estimate of the reference field, and the ES measures the uncertainty associated with such an estimate. The ratio RMSE/ES is a measure of filter inbreeding, which may cause the filter to collapse, and should, ideally, be close to one (e.g., Liang et al. 2011; Xu et al. 2013).

6 Results

As mentioned above, two analyses have been performed, a preliminary one using synthetic data to decide on the number of realizations and on a method to prevent

filter collapse, followed by a specific analysis of the data collected from the sandbox experiment.

6.1 Analysis of the Synthetic Data

The synthetic analysis is performed on six scenarios, with combinations between two numbers of realizations and five alternatives to prevent filter collapse, as given in Table 2. Recall that the reference for the synthetic case comes from a numerical simulation of flow and transport with the same characteristics as the sandbox experiment.

Figures 3 and 4 focus on the source parameters; they provide the ensemble mean and the ensemble variance, respectively, of all five source parameters, after the update at each time step for all six scenarios. The ranges of the ensemble variances were very different for each parameter; for this reason the results are displayed after standardization by the ensemble variances of the initial ensembles. It is hard to argue which scenario performs best. Scenario $S3$, the one with a damping factor of 0.1, can be discarded since it is the one that ends with the highest variances for most of the parameters. Scenario $S5$, the one with Wang's inflation method, should also be discarded because it collapses the ensemble after a few time steps, as shown by the rapid decrease of the ensemble variance to zero for almost all parameters. Scenario $S2$, with no inflation, but 1000 realizations—double the rest of the scenarios—performs well in that it provides an estimate close to the true values, and the variance decreases in time consistently and similarly to the rest of the scenarios. Scenario $S1$, with no inflation and 500 realizations shows some filter collapse, which does not happen as quickly as for $S5$ but ends with similar magnitudes for the ensemble variances. Scenario $S4$, with a damping factor of 0.5, does a good job in the estimation of the source parameters, except for Ic, but the final uncertainties are the largest after $S3$ for most of the parameters. Finally, scenario $S6$, with Bauser's inflation method, could be considered as the one with the best performance, since it provides very good estimates for all parameters, except for Ir, and it has low final uncertainties without filter collapse. All methods estimate the vertical position Zs of the release point lower in the sandbox than its real position, this behavior can be produced by local velocity variations induced by the proximity of the injection to the boundary between two cells with different glass bead diameters, which are not resolved by the observations.

Figure 5 shows the ensemble mean and Fig. 6 the ensemble variance of the initial $\ln K$ realizations and of the updated ones computed at the 90th time step for all synthetic scenarios. The ensemble mean and ensemble variance of the initial $\ln K$ are almost homogeneous and equal to their prior values, since no conditional data of $\ln K$ is employed. After assimilating all concentration data during 90 time steps, the ensemble mean of the updated $\ln K$ conductivities can capture the main patterns of variability of the glass bead distribution with a substantial reduction of the ensemble variance in most of the sandbox. A comparison among the different scenarios shows that, again, $S3$ performs worst, with the worst estimation of $\ln K$ and the largest estimation variances, and $S5$ shows filter collapse at most locations. Of the remaining scenarios, $S2$ and $S6$ give the best results, with $S2$ being slightly better in $\ln K$ pattern estimation thanks to the larger number of ensemble members. For a more quantitative evaluation of the
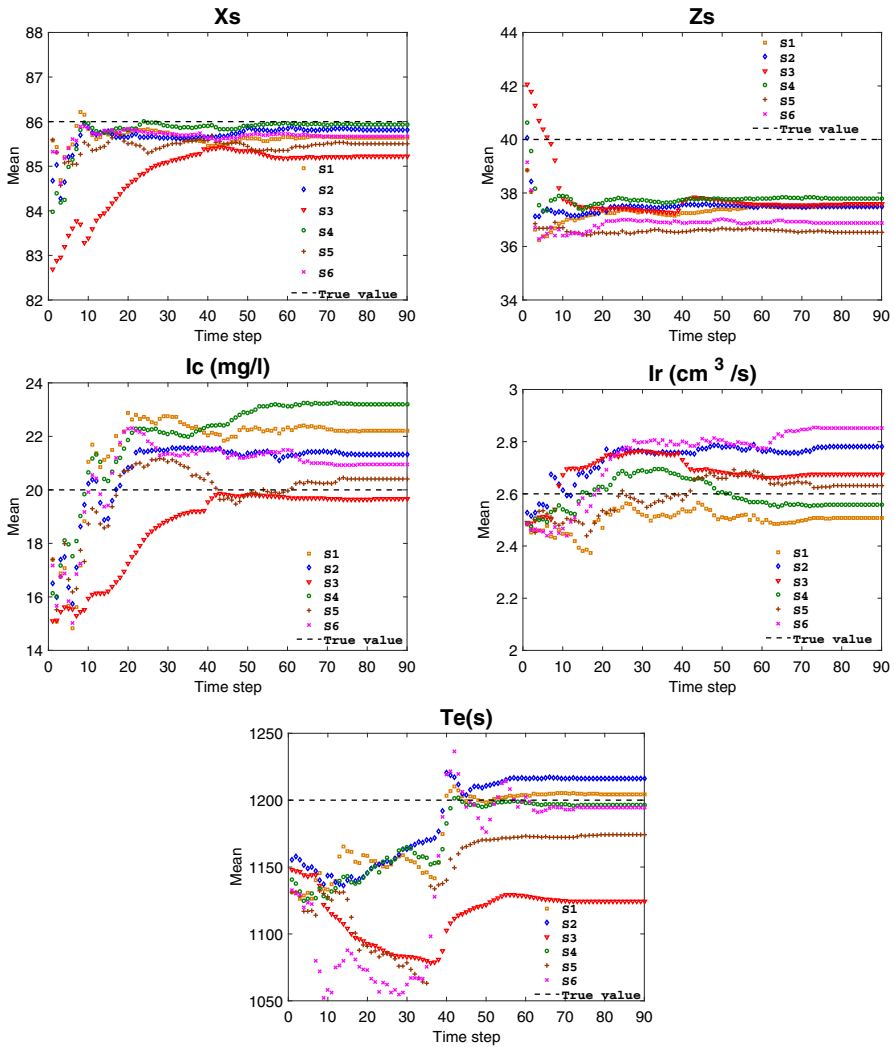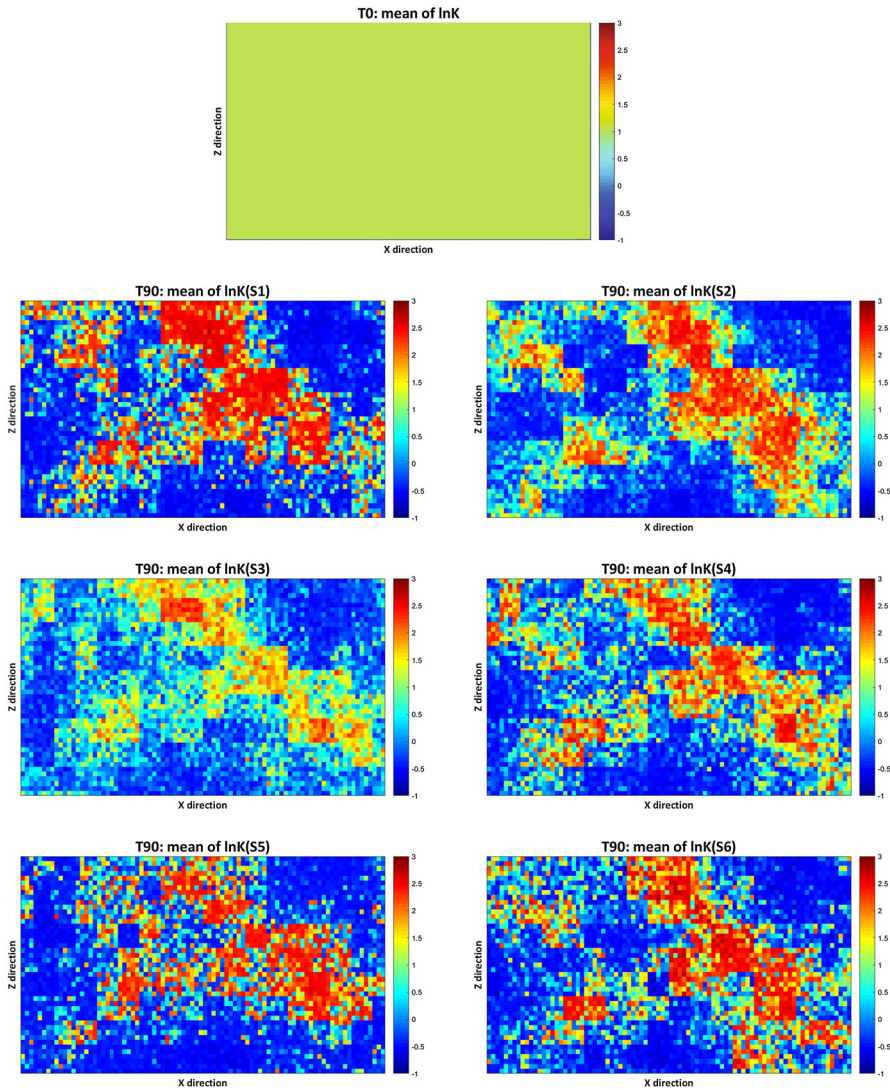
Fig. 3 Time evolution of the ensemble means of the updated contaminant source parameters for all synthetic scenarios ($S1 - S6$)

identification of $\ln K$, Figure 7 shows how the three statistics RMSE, ES and RMSE/ES evolve in time as the data assimilation proceeds. The best performance would be for the lowest values of RMSE and ES and the closest-to-one RMSE/ES ratio. The two best scenarios are $S2$ and $S6$, with $S6$ having the RMSE/ES ratio closest to one.

Taking into consideration the performance of the rNS-EnKF for the different synthetic scenarios, the two scenarios that will be analyzed with the experimental data are the non-inflation method with 1000 realizations, referred to as $R1$, and the Bauser's inflation method with 500 realizations, referred to as $R2$.
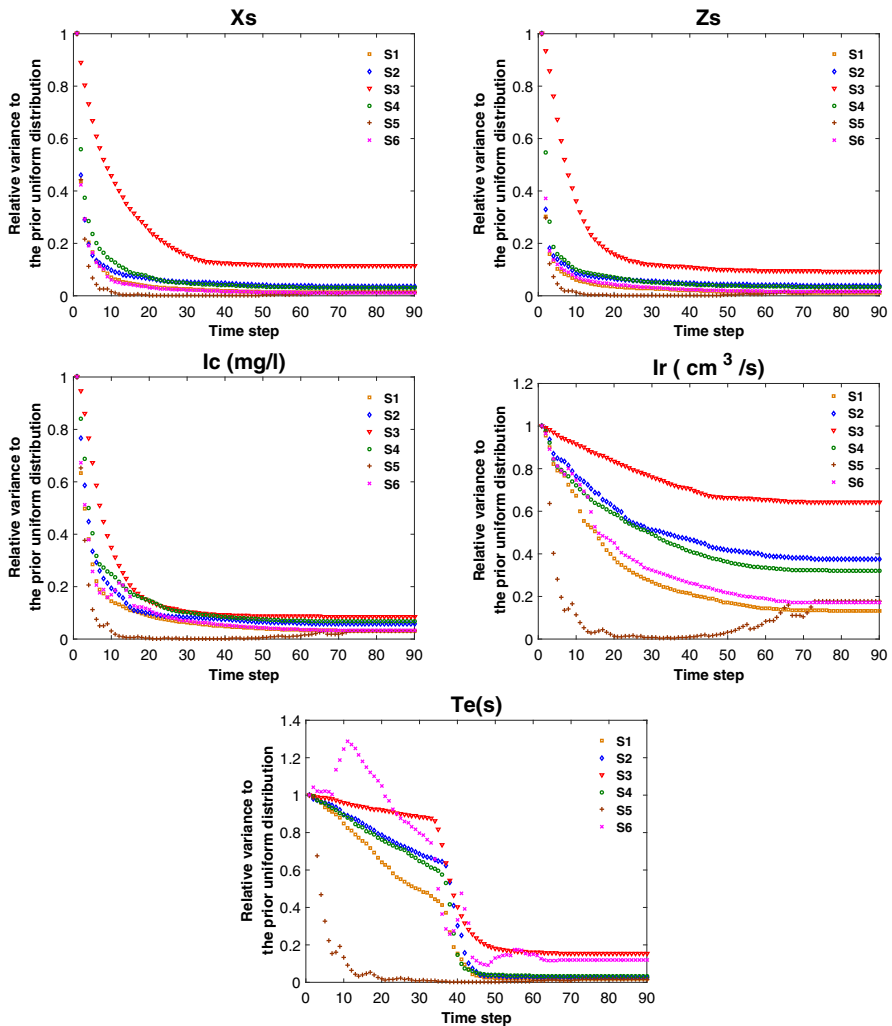
Fig. 4 Time evolution of ensemble variances of the updated contaminant source parameters for all synthetic scenarios($S1 - S6$). Each variance plot has been standardized by the variance of the initial ensemble

6.2 Analysis of the Sandbox Data

The difficulties found on the first attempt in applying the rNS-EnKF to the sandbox data must be due to observation errors in the concentrations. According to earlier work (Chen et al. 2018), an underestimation of the observation error will force the filter to fit the concentrations too closely, producing biased estimates of the parameters, and an overestimation of the observation error will allow too loose a fit, producing estimates with large uncertainty. Since the same sandbox equipment as Cupola et al. (2015) and Chen et al. (2018) was used, the same observation error distribution with a mean of 0 mg/l and a standard deviation of 1 mg/l were retained for this analysis.

Fig. 5 Ensemble mean of the initial $\ln K$ realizations and the updated $\ln K$ realizations of all synthetic scenarios($S1 - S6$) at the $90th$ time step

Figures 8 and 9 show the evolution of the ensemble mean and the ensemble variance, respectively, of the contaminant source parameters for the two sandbox scenarios ($R1$, $R2$). Both approaches perform well with mean estimates close to the true values and estimation variances close to zero for all parameters. It seems that the injection concentration and the injection rate are more difficult to identify; they have the largest estimation error and the largest estimation variance. However, if the mass loading rate is computed, that is the product of the injection rate times the injection concentration; its mean and variance is similar to those of the other contaminant parameters. This
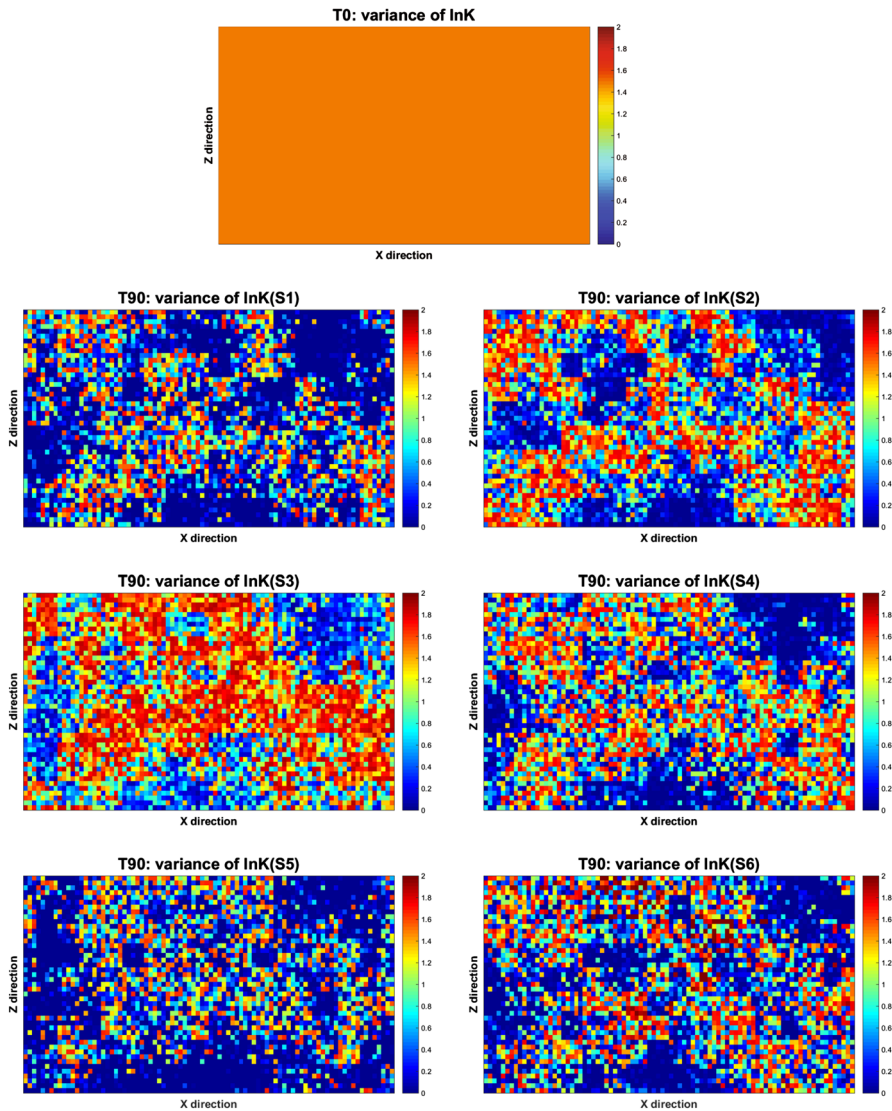
Fig. 6 Ensemble variance of the initial $\ln K$ realizations and the updated $\ln K$ realizations of all synthetic scenarios($S1 - S6$) at the $90th$ time step

result seems to indicate that there may be some indetermination in the identification of parameters Ic and Ir that disappears when the subject of identification is its product. Disregarding parameters Ic and Ir, it can be concluded that both scenarios perform equally and, therefore, that Bauser's inflation method can make up for the reduction from 1000 realizations to 500 realizations with similar performance.

Figure 10 shows the ensemble mean and variance of $\ln K$ for scenarios $R1$ and $R2$ at the 90th time step. Figure 11 shows the ensemble mean of the absolute differences
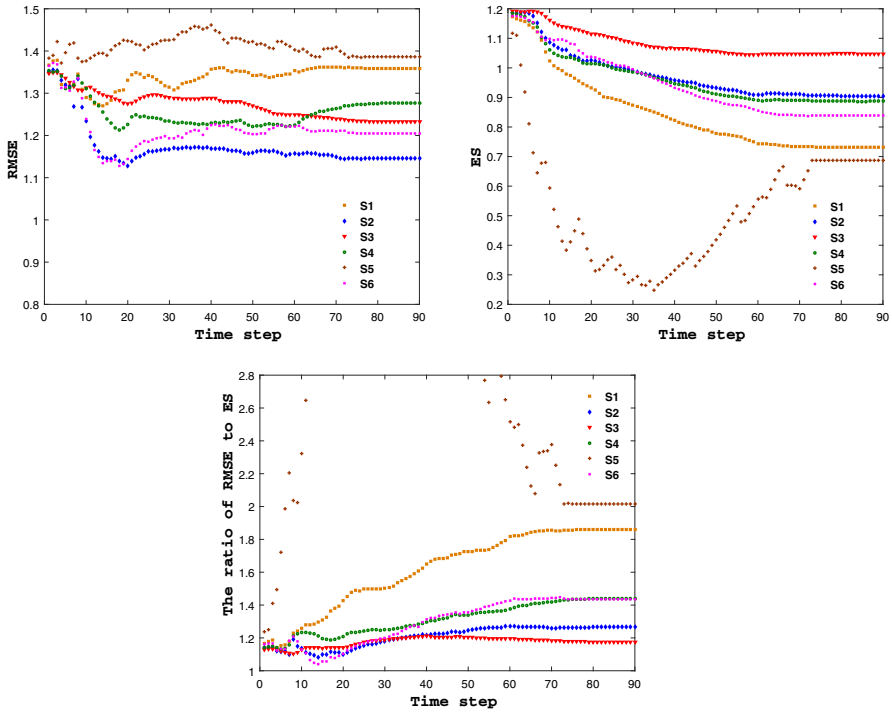
Fig. 7 Time evolution of RMSE, ES and the ratio of RMSE to ES for all synthetic scenarios (S1−S6)

between the reference and updated $\ln K$ maps at the 90th time step. Both scenarios capture the main patterns of variability of $\ln K$ and the ensemble variance is substantially reduced in the areas of low conductivity. This is mainly because there is a strong correlation between low concentrations and low conductivities; the algorithm forces conductivities to be low at the locations where the realization predict large concentrations when the observed values are low or zero. Comparing the two scenarios, variance reduction is larger for scenario $R2$ and the absolute deviations between reference and estimated conductivities are smaller for $R2$, implying again that Bauser's inflation method is a valuable approach to reducing ensemble size and achieving results similar to (or better that) those obtained when a larger ensemble is used. Figure 12 shows the evolution in time of the RMSE, ES and RMSE/ES ratio for scenarios $R1$ and $R2$. Again, scenario $R2$ performs remarkably well as compared to scenario $R1$, with a similar RMSE, smaller ES and a ratio RMSE/ES not too far from one.

Figure 13 shows the evolution of the contaminant plume in the sandbox at the 10th, 40th, 60th and 90th time steps, while Fig. 14 is the ensemble mean of the contaminant plumes computed on the initial ensemble of realizations. Figures 15 and 16 show the ensemble mean of the contaminant plumes for scenarios $R1$ and $R2$, respectively, at the same times steps as in Fig. 13 computed with all the parameters updated at the 90th time step. The comparison of the simulated plumes with the observed ones is very favorable, demonstrating that the estimated parameters are conditioned on the
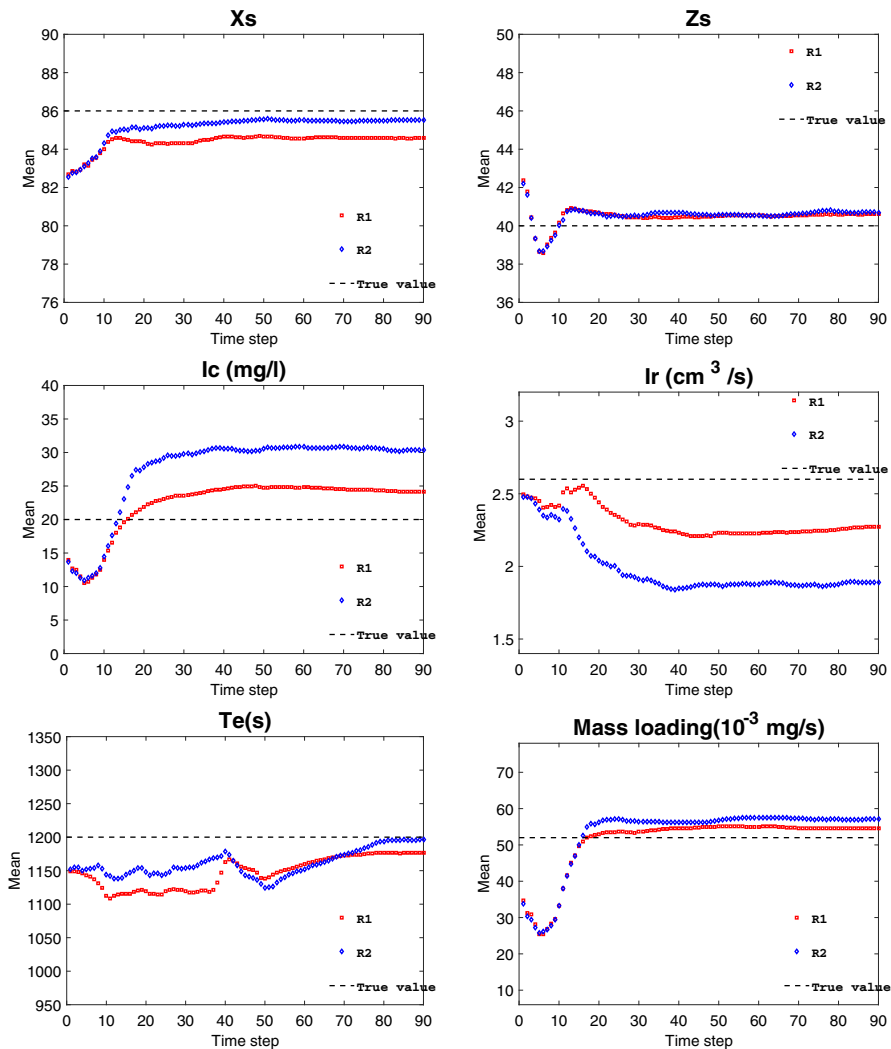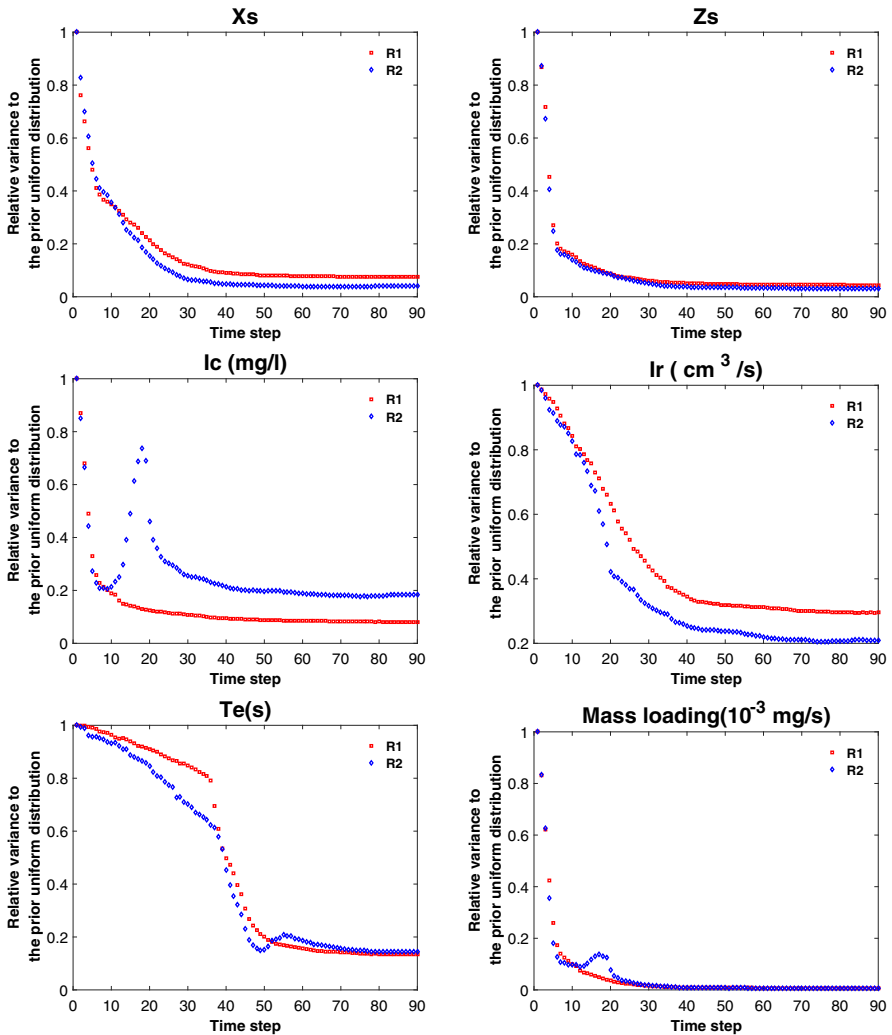
Fig. 8 Time evolution of the ensemble means of the updated contaminant source parameters for the two sandbox scenario ($R1$, $R2$). Also shown is the mass loading rate $Ic \cdot Ir$

observed concentrations, and that they are capable of giving a good prediction of contaminant movement.

7 Discussion and Conclusions

Xu and Gómez-Hernández (2018) showed the capabilities of the restart normal-score ensemble Kalman filter (rNS-EnKF) for the simultaneous identification of source parameters and hydraulic conductivities in synthetic aquifers. This work presents the first attempt to apply it to a non-synthetic exercise. An aquifer is mimicked by a

Fig. 9 Time evolution of the ensemble variances of the updated contaminant source parameters for the two sandbox scenario ($R1$, $R2$). Also shown is the mass loading rate $Ic \cdot Ir$. Notice that each ensemble variance has been normalized by their values at time zero

laboratory sandbox in which geometry, initial and boundary conditions are known. The first finding was that it was not straightforward to apply the approach to the collected data; working under laboratory conditions does not preclude measurement errors and other errors, which prevented the filter from working properly on first attempts. The filter would collapse, even for large ensemble sizes. This led us to conduct a preliminary analysis of a synthetic case using solute concentrations generated by a numerical model, thus getting rid of model or measurement errors. Six scenarios were compared in this synthetic exercise, showing the importance of a good selection of an approach to prevent filter collapse. Of the four alternative approaches, Bauser's covariance inflation
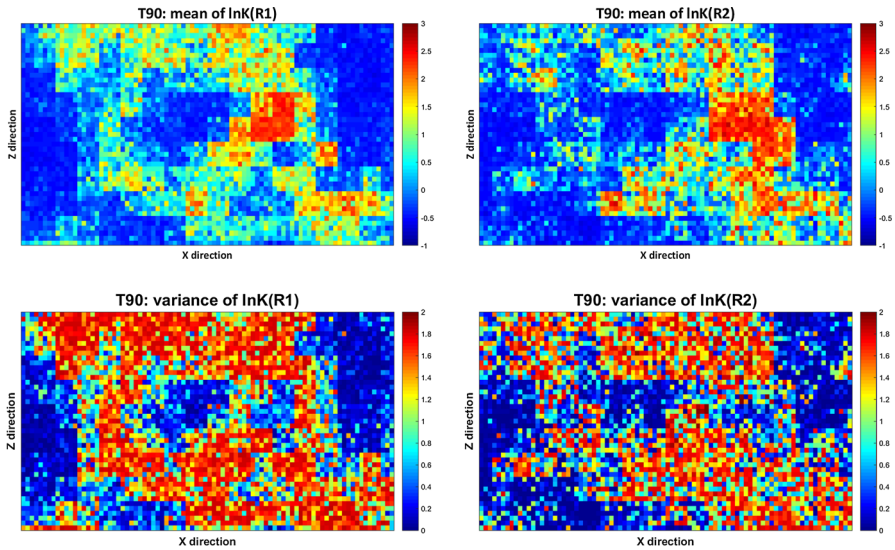
Fig. 10 Ensemble mean (top row) and ensemble variance (bottom row) of updated $\ln K$ of scenarios $R1$ and $R2$ at the $90th$ time step
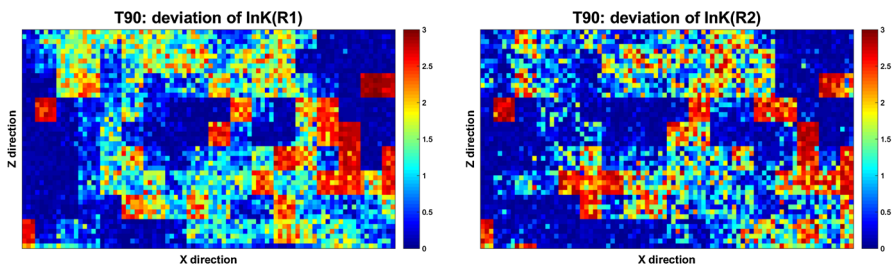
Fig. 11 Ensemble mean of the absolute deviation between reference and updated $\ln K$ in scenarios $R1$ and $R2$ at the $90th$ time step

method emerged as the most appropriate, allowing us to reduce the ensemble size from 1000 members (without inflation) to 500 (with inflation) to yield similar results. In these synthetic scenarios, it could also be observed that the horizontal coordinate of the source was well identified, but that the vertical one was estimated a little bit downwards from the original position. The explanation must be due to the closeness of the source to a boundary between the large glass beads and the small ones. The synthetic results also showed that it is difficult to identify a binary conductivity field starting from a continuous distribution of log-conductivities, yet the two main zones of high and low conductivities were well captured in the different scenarios, with the scenario having 1000 realizations performing best, followed by the scenario with 500 realizations and using Bauser's covariance inflation method.

The application of Bauser's inflation and 500 realizations to the data observed in the sandbox was compared with a non-inflated filter and 1000 realizations, with comparable results. The identification of the source parameters was good in both cases, even
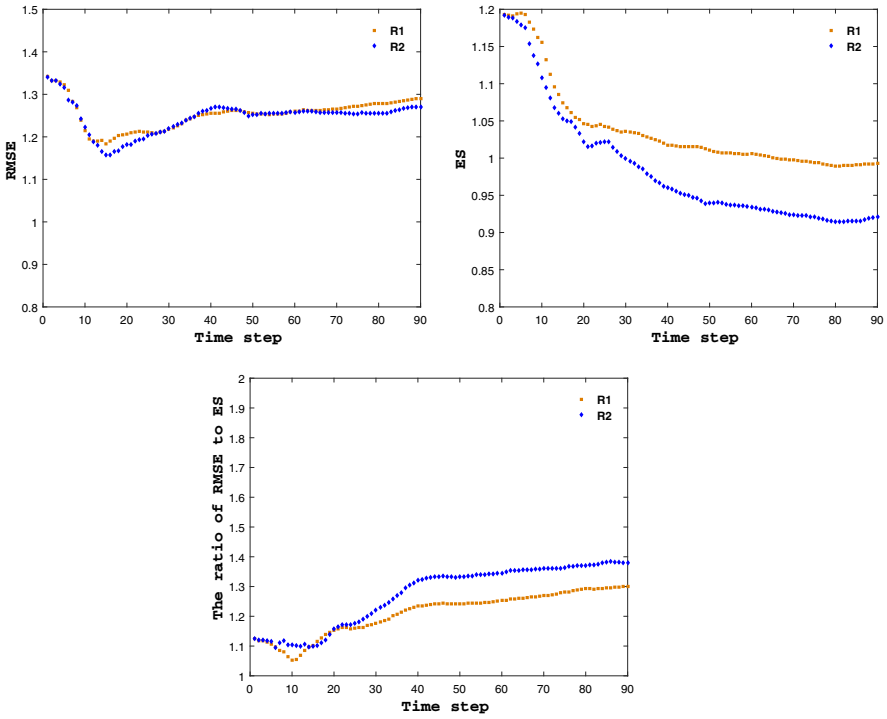
Fig. 12 Time evolution of RMSE, ES and the ratio of RMSE to ES for scenarios R1 and R2
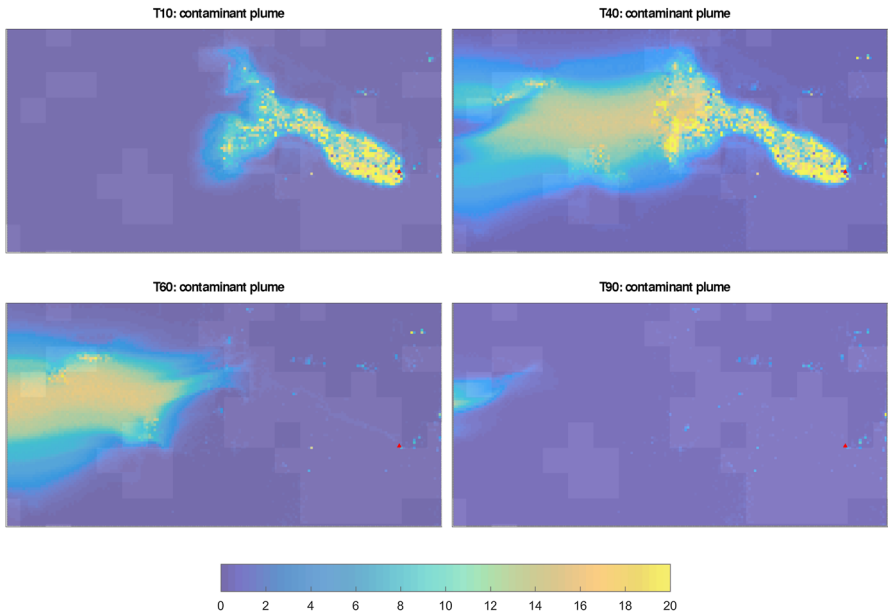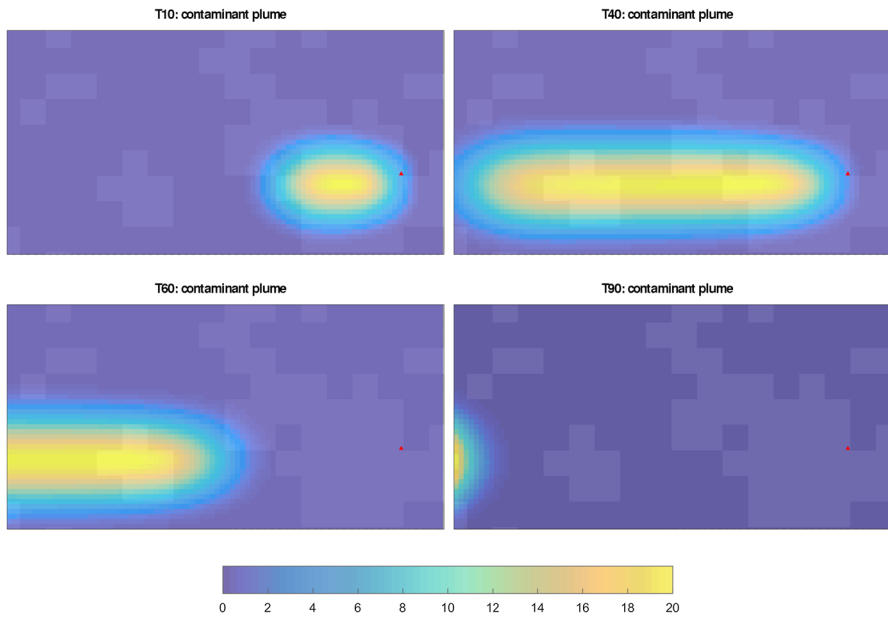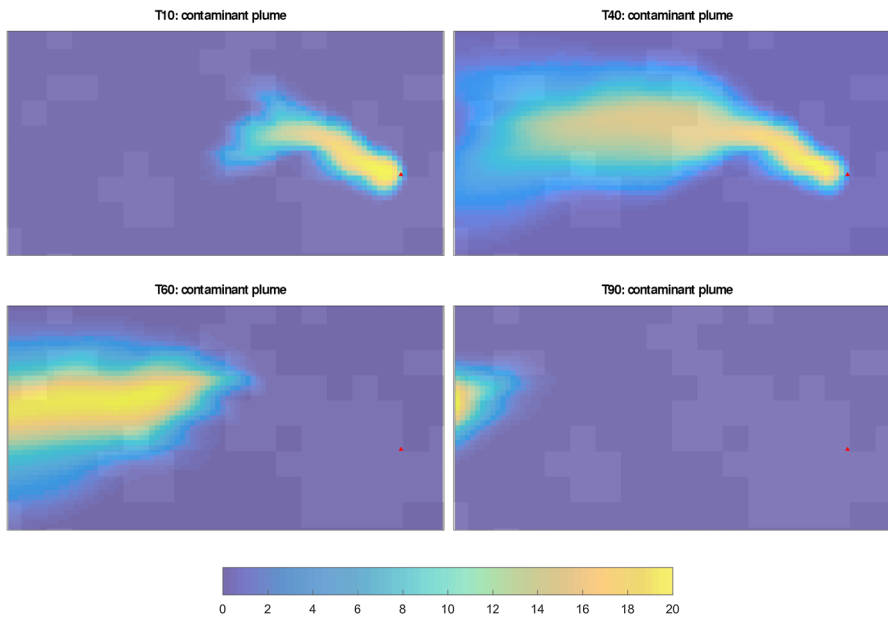

Fig. 13 Reference contaminant plume evolution at the 10th, 40th, 60th and 90th time steps in the sandbox. Red triangle denotes the real injector
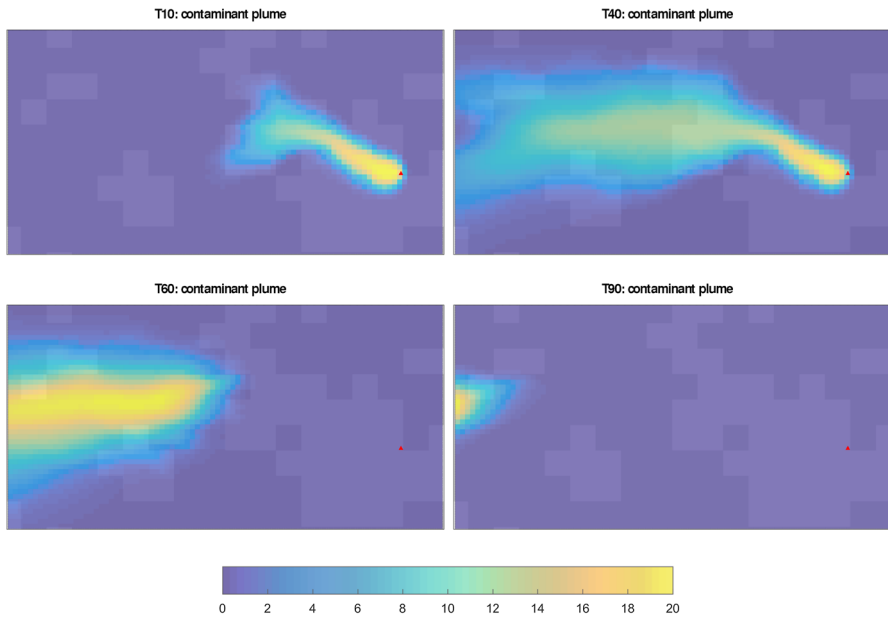
Fig. 14 Ensemble mean of contaminant plume evolution of the initial realizations at the 10th, 40th, 60th and 90th time steps. Red triangle denotes the real injector

Fig. 15 Ensemble mean of contaminant plume evolution of scenario $R1$ at the 10th, 40th, 60th and 90th time steps with all parameters updated after the 90th time step. Red triangle denotes the real injector

Fig. 16 Ensemble mean of contaminant plume evolution of scenario $R2$ at the 10th, 40th, 60th and 90th time steps with all parameters updated after the 90th time step. Red triangle denotes the real injector

for the vertical coordinate of the injection. A better identification of the source vertical position in the sandbox than in the synthetic exercises could be explained by the larger measurement error variance used in the sandbox observations than in the synthetic scenarios. A larger measurement error gives the filter more flexibility to update the parameters to fit the observations, while resulting in a larger variance on the ensemble of final parameters. It was also evident that the estimation of both injection rate and injection concentration were biased; a further analysis showed that there is a degree of indetermination in the estimation of these two parameters, since the parameter that really matters is their product, the mass loading rate. The mass loading rate is well estimated with no bias and little uncertainty. As in the synthetic case, the estimation of a binary conductivity field by a continuous one is almost impossible, but the final ensemble of log-conductivities displays enough spatial heterogeneity to distinguish two main areas of high and low conductivities. More importantly, the solution of the mass transport equation in the final conductivity fields yields a contaminant plume that moves in space and time in pattern very similar to the one observed in the sandbox, particularly in comparison with the mean plume estimate based on the initial ensemble of conductivities.

It is important to notice that, in the sandbox experiment, the only available data were concentration data; no observations of either conductivities or piezometric heads were available. In a practical case, both conductivity and piezometric head data could also be assimilated, resulting in an improved estimation of all parameters being identified as shown, for instance by Wen et al. (1996). In all cases, the number and distribution of the observations will be critical; an interesting continuation of this work would be

to perform a sensitivity analysis on the number and geometry of observation locations together with the inclusion of piezometric head and conductivity data.

In conclusion, the rNS-EnKF has been demonstrated to work for the joint identification of a contaminant source and conductivities beyond the synthetic exercises where it had been tested previously. The demonstration is still far from field conditions, where boundary and initial conditions, forcing terms or geometry are not necessarily known, but the sandbox exercise included a binary heterogeneous conductivity spatial distribution, which is always difficult to identify. Further work should focus on the application of the rNS-EnKF to a field case.

References

Amirabdollahian M, Datta B (2014) Identification of pollutant source characteristics under uncertainty in contaminated water resources systems using adaptive simulated anealing and fuzzy logic. Int J GEOMATE 6(1):757–762

Anderson JL (2007) An adaptive covariance inflation error correction algorithm for ensemble filters. Tellus, Ser A: Dynam Meteorol Oceanogr 59(2):210–224. https://doi.org/10.1111/j.1600-0870.2006.00216.x

Aral MM, Guan J, Maslia ML (2001) Identification of contaminant source location and release history in aquifers. J Hydrol Eng 6(3):225–234. https://doi.org/10.1061/(ASCE)1084-0699(2001)6:3(225)

Atmadja J, Bagtzoglou AC (2001) State of the art report on mathematical methods for groundwater pollution source identification. Environ Forens 2(3):205–214. https://doi.org/10.1006/enfo.2001.0055

Ayvaz MT (2016) A hybrid simulation-optimization approach for solving the areal groundwater pollution source identification problems. J Hydrol 538:161–176. https://doi.org/10.1016/j.jhydrol.2016.04.008

Bagtzoglou AC, Atmadja J (2005) Mathematical methods for hydrologic inversion: the case of pollution source identification. Water Pollut 5:65–96. https://doi.org/10.1007/b11442

Bagtzoglou AC, Dougherty DE, Tompson AFB (1992) Application of particle methods to reliable identification of groundwater pollution sources. Water Resour Manage 6(1):15–23. https://doi.org/10.1007/BF00872184

Bauser HH, Berg D, Klein O, Roth K (2018) Inflation method for ensemble Kalman filter in soil hydrology. Hydrol Earth Syst Sci 22(9):4921–4934. https://doi.org/10.5194/hess-22-4921-2018

Bear J (1972) Dynamics of Fluids in Porous Media. American Elsevier, Amsterdam

Butera I, Tanda MG, Zanini A (2013) Simultaneous identification of the pollutant release history and the source location in groundwater by means of a geostatistical approach. Stochast Environ Res Risk Assess 27(5):1269–1280. https://doi.org/10.1007/s00477-012-0662-1

Camporese M, Cassiani G, Deiana R, Salandin P (2011) Assessment of local hydraulic properties from electrical resistivity tomography monitoring of a three-dimensional synthetic tracer test experiment. Water Resour Res 47(12):1–15. https://doi.org/10.1029/2011WR010528

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

Chang H, Zhang D, Lu Z (2010) History matching of facies distribution with the EnKF and level set parameterization. J Comput Phys 229(20):8011–8030. https://doi.org/10.1016/j.jcp.2010.07.005

Chen Y, Zhang D (2006) Data assimilation for transient flow in geologic formations via ensemble Kalman filter. Adv Water Res 29(8):1107–1122. https://doi.org/10.1016/j.advwatres.2005.09.007

Chen Z, Gómez-Hernández JJ, Xu T, Zanini A (2018) Joint identification of contaminant source and aquifer geometry in a sandbox experiment with the restart ensemble kalman filter. J Hydrol 564:1074–1084

Citarella D, Cupola F, Tanda MG, Zanini A (2015) Evaluation of dispersivity coefficients by means of a laboratory image analysis. J Contam Hydrol 172:10–23. https://doi.org/10.1016/j.jconhyd.2014.11.001

Crestani E, Camporese M, Baú D, Salandin P (2012) Ensemble Kalman filter versus ensemble smoother for assessing hydraulic conductivity via tracer test data assimilation. Hydrol Earth Syst Sci Discuss 9(11):13083–13115. https://doi.org/10.5194/hessd-9-13083-2012

Cupola F, Tanda MG, Zanini A (2015) Laboratory sandbox validation of pollutant source location methods. Stochast Environ Res Risk Assess 29(1):169–182. https://doi.org/10.1007/s00477-014-0869-4

Datta B, Chakrabarty D, Dhar A (2009) Simultaneous identification of unknown groundwater pollution sources and estimation of aquifer parameters. J Hydrol 376(1–2):48–57. https://doi.org/10.1016/j.jhydrol.2009.07.014

Evensen G (1994) Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J Geophys Res 99(C5):10143. https://doi.org/10.1029/94JC00572

Feyen L, Gómez-Hernández JJ, Ribeiro P Jr, Beven KJ, De Smedt F (2003a) A Bayesian approach to stochastic capture zone delineation incorporating tracer arrival times, conductivity measurements, and hydraulic head observations. Water Resour Res 39(5):1126. https://doi.org/10.1029/2002WR001544

Feyen L, Ribeiro P Jr, Gomez-Hernandez J, Beven KJ, De Smedt F (2003b) Bayesian methodology for stochastic capture zone delineation incorporating transmissivity measurements and hydraulic head observations. J Hydrol 271(1–4):156–170

Franssen HH, Gómez-Hernández J (2002) 3d inverse modelling of groundwater flow at a fractured site using a stochastic continuum model with multiple statistical populations. Stochast Environ Res Risk Assess 16(2):155–174

Gómez-Hernández J, Wen XH (1994) Probabilistic assessment of travel times in groundwater modeling. Stochast Hydrol Hydraul 8(1):19–55

Gómez-Hernández J, Franssen HJH, Sahuquillo A (2003) Stochastic conditional inverse modeling of subsurface mass transport: a brief review and the self-calibrating method. Stochast Environ Res Risk Assess 17(5):319–328

Gómez-Hernández JJ, Wen XH (1998) To be or not to be multi-Gaussian? A reflection on stochastic hydrogeology. Adv Water Resour 21(1):47–61. https://doi.org/10.1016/S0309-1708(96)00031-0

Gorelick SM, Evans B, Remson I (1983) Identifying sources of groundwater pollution: an optimization approach. Water Resour Res 19(3):779–790. https://doi.org/10.1029/WR019i003p00779

Greybush SJ, Kalnay E, Miyoshi T, Ide K, Hunt BR (2011) Balance and ensemble kalman filter localization techniques. Mon Weather Rev 139(2):511–522

Hendricks Franssen HJ, Kinzelbach W (2008) 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(9):1–21. https://doi.org/10.1029/2007WR006505

Hendricks Franssen HJ, Kinzelbach W (2009) Ensemble kalman filtering versus sequential self-calibration for inverse modelling of dynamic groundwater flow systems. J Hydrol 365(3–4):261–274

Houtekamer PL, Mitchell HL (2001) A sequential ensemble kalman filter for atmospheric data assimilation. Mon Weather Rev 129(1):123–137. https://doi.org/10.1175/1520-0493(2001)129<0123:ASEKFF>2.0.CO;2

Jafarpour B, Khodabakhshi M (2011) A Probability Conditioning Method (PCM) for nonlinear flow data integration into multipoint statistical facies simulation. Math Geosci 43(2):133–164. https://doi.org/10.1007/s11004-011-9316-y

Journel A, Isaaks E (1984) Conditional indicator simulation: application to a saskatchewan uranium deposit. J Int Assoc Math Geol 16(7):685–718

Journel AG, Gomez-Hernandez JJ et al (1993) Stochastic imaging of the wilmington clastic sequence. SPE format Evaluat 8(01):33–40

Knudby C, Carrera J (2005) On the relationship between indicators of geostatistical, flow and transport connectivity. Adv Water Resour 28(4):405–421. https://doi.org/10.1016/j.advwatres.2004.09.001

Koch J, Nowak W (2016) Identification of contaminant source architectures—A statistical inversion that emulates multiphase physics in a computationally practicable manner. Water Res Res 52(2):1009–1025. https://doi.org/10.1002/2015WR017894

Kumar D, Srinivasan S (2019) Ensemble-based assimilation of nonlinearly related dynamic data in reservoir models exhibiting non-gaussian characteristics. Math Geosci 51(1):75–107. https://doi.org/10.1007/s11004-018-9762-x

Kumar D, Srinivasan S (2020) Indicator-based data assimilation with multiple-point statistics for updating an ensemble of models with non-Gaussian parameter distributions. Adv Water Resour 141:103611. https://doi.org/10.1016/j.advwatres.2020.103611

Li H, Kalnay E, Miyoshi T (2009) Simultaneous estimation of covariance inflation and observation errors within an ensemble Kalman filter. Q J R Meteorol Soc 135(639):523–533. https://doi.org/10.1002/qj.371

Li L, Zhou H, Gómez-Hernández JJ (2011) A comparative study of three-dimensional hydraulic conductivity upscaling at the macro-dispersion experiment (made) site, Columbus air force base, Mississippi (USA). J Hydrol 404(3–4):278–293

Li L, Zhou H, Gómez-Hernández JJ, Hendricks Franssen HJ (2012a) Jointly mapping hydraulic conductivity and porosity by assimilating concentration data via ensemble Kalman filter. J Hydrol 428–429:152–169. https://doi.org/10.1016/j.jhydrol.2012.01.037

Li L, Zhou H, Hendricks Franssen HJ, Gómez-Hernández JJ (2012b) Groundwater flow inverse modeling in non-MultiGaussian media: performance assessment of the normal-score Ensemble Kalman Filter. Hydrol Earth Syst Sci 16(2):573–590. https://doi.org/10.5194/hess-16-573-2012

Li L, Zhou H, Hendricks Franssen HJ, Gómez-Hernández JJ (2012c) Modeling transient groundwater flow by coupling ensemble kalman filtering and upscaling. Water Resour Res 48(1):W01537. https://doi.org/10.1029/2010WR010214

Liang X, Zheng X, Zhang S, Wu G, Dai Y, Li Y (2011) Maximum likelihood estimation of inflation factors on error covariance matrices for ensemble kalman filter assimilation. Q J R Meteorol Soc 138(662):263–273

Liang X, Zheng X, Zhang S, Wu G, Dai Y, Li Y (2012) Maximum likelihood estimation of inflation factors on error covariance matrices for ensemble Kalman filter assimilation. Q J R Meteorol Soc 138(662):263–273. https://doi.org/10.1002/qj.912

Mahar PS, Datta B (2000) Identification of pollution sources in transient groundwater systems. Water Resour Manage 14(3):209–227. https://doi.org/10.1023/A:1026527901213

McDonald JM, Harbaugh AW (1988) A modular three-dimensional finite-difference flow model. Techniq Water Resour Investig US Geol Surv Book 6:586. https://doi.org/10.1016/0022-1694(86)90106-X

Michalak AM, Kitanidis PK (2004) Estimation of historical groundwater contaminant distribution using the adjoint state method applied to geostatistical inverse modeling. Water Resour Res. https://doi.org/10.1029/2004WR003214

Mirghani BY, Mahinthakumar KG, Tryby ME, Ranjithan RS, Zechman EM (2009) A parallel evolutionary strategy based simulation-optimization approach for solving groundwater source identification problems. Adv Water Resour 32(9):1373–1385. https://doi.org/10.1016/j.advwatres.2009.06.001

Neupauer RM, Wilson JL (1999) Adjoint method for obtaining backward-in-time location and travel time probabilities of a conservative groundwater contaminant. Water Resour Res 35(11):3389–3398. https://doi.org/10.1029/1999WR900190

Sun AY, Painter SL, Wittmeyer GW (2006) A constrained robust least squares approach for contaminant release history identification. Water Res Res 42(4):1–13. https://doi.org/10.1029/2005WR004312

Sun AY, Morris AP, Mohanty S (2009) Sequential updating of multimodal hydrogeologic parameter fields using localization and clustering techniques. Water Resour Res 45(7):1–15. https://doi.org/10.1029/2008WR007443

Wagner BJ (1992) Simultaneous parameter estimation and contaminant source characterization for coupled groundwater flow and contaminant transport modelling. J Hydrol 135(1–4):275–303. https://doi.org/10.1016/0022-1694(92)90092-A

Wang X, Bishop CH (2003) A comparison of breeding and ensemble transform kalman filter ensemble forecast schemes. J Atmos Sci 60(9):1140–1158. https://doi.org/10.1175/1520-0469(2003)060<1140:ACOBAE>2.0.CO;2

Wen XH, Chen WH (2005) Some practical issues on real-time reservoir model updating using ensemble Kalman filter. Paper presented at the International Petroleum Technology Conference, Doha, Qatar, November 2005. Paper Number: IPTC-11024-MS. https://doi.org/10.2523/IPTC-11024-MS

Wen XH, Chen WH (2006) Real-time reservoir model updating using ensemble Kalman filter with confirming option. SPE J 11(4):431–442. https://doi.org/10.2118/92991-PA

Wen XH, Jaime Gómez-Hernandez J, Capilla JE, Sahuquillo A (1996) Significance of conditioning to piezometric head data for predictions of mass transport in groundwater modeling. Math Geol 28(7):951–968. https://doi.org/10.1007/BF02066011

Wen XH, Capilla JE, Deutsch C, Gómez-Hernández J, Cullick A (1999) A program to create permeability fields that honor single-phase flow rate and pressure data. Comp Geosci 25(3):217–230

Woodbury AD, Ulrych TJ (1996) Minimum relative entropy inversion: theory and application to recovering the release history of a groundwater contaminant. Water Resour Res 32(9):2671–2681

Xu T, Gómez-Hernández JJ (2016) Joint identification of contaminant source location, initial release time, and initial solute concentration in an aquifer via ensemble Kalman filtering. Water Resour Res. https://doi.org/10.1002/2014WR016618.Received

Xu T (2017) Gómez-Hernández JJ (2018) Simultaneous identification of a contaminant source and hydraulic conductivity via the restart normal-score ensemble Kalman filter. Adv Water Resour 112:106–123. https://doi.org/10.1016/j.advwatres.2017.12.011

Xu T, Gómez-Hernández JJ, Zhou H, Li L (2013) The power of transient piezometric head data in inverse modeling: an application of the localized normal-score EnKF with covariance inflation in a heterogenous bimodal hydraulic conductivity field. Adv Water Res 54:100–118. https://doi.org/10.1016/j.advwatres.2013.01.006

Yeh HD, Chang TH, Lin YC (2007) Groundwater contaminant source identification by a hybrid heuristic approach. Water Resour Res 43(9):1–16. https://doi.org/10.1029/2005WR004731

Zheng C, Wang PP (1999) MT3DMS: A Modular Three-Dimensional Multispecies Transport Model (December):219

Zheng X (2009) An adaptive estimation of forecast error covariance parameters for Kalman filtering data assimilation. Adv Atmos Sci 26(1):154–160. https://doi.org/10.1007/s00376-009-0154-5

Zhou H, Gómez-Hernández JJ, Hendricks Franssen HJ, Li L (2011) An approach to handling non-Gaussianity of parameters and state variables in ensemble Kalman filtering. Adv Water Resour 34(7):844–864. https://doi.org/10.1016/j.advwatres.2011.04.014

Zhou H, Gómez-Hernández JJ, Li L (2012a) A pattern-search-based inverse method. Water Resour Res 48(3):W03505. https://doi.org/10.1029/2011WR011195

Zhou H, Li L, Franssen HJH, Gómez-Hernández JJ (2012b) Pattern recognition in a bimodal aquifer using the normal-score ensemble kalman filter. Math Geosci 44(2):169–185

Zhou H, Gómez-Hernández JJ, Li L (2014) Inverse methods in hydrogeology: evolution and recent trends. Adv Water Resour 63:22–37. https://doi.org/10.1016/j.advwatres.2013.10.014

Zinn B, Harvey CF (2003) 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(3):137–147. https://doi.org/10.1029/2001WR001146