Joint identification of contaminant source and aquifer geometry in a sandbox experiment with the restart Ensemble Kalman filter

Zi Chen[a], J. Jaime Gómez-Hernández[a], Teng Xu[a,*], Andrea Zanini[b]

[a]Institute of Water and Environmental Engineering, Universitat Politècnica de València, Valencia, Spain
[b]Dipartimento di Ingegneria e Architettura, Università degli Studi di Parma, Parma, Italy

Abstract

Contaminant source identification is a key problem in handling groundwater pollution events. The ensemble Kalman filter (EnKF) is used for the spatiotemporal identification of a point contaminant source in a sandbox experiment, together with the identification of the position and length of a vertical plate inserted in the sandbox that modifies the geometry of the system. For the identification of the different parameters, observations in time of solute concentration are used, but not of piezometric head data since they were not available. A restart version of the EnKF is utilized because it is necessary to restart the forecast from time zero after each parameter update. The results show that the restart EnKF is capable of identifying both contaminant source information and aquifer-geometry-related parameters together with an uncertainty estimate of such identification.

Keywords: Inverse modeling; Observation error; Groundwater laboratory experiment; Stochastic hydrogeology

1. Introduction

The problem of identifying a contaminant source in an aquifer using solute concentration data has been the subject of attention for many years (e.g., Atmadja and Bagtzoglou,

*Corresponding author
Email address: tenxu@posgrado.upv.es (Teng Xu)

2001; Michalak and Kitanidis, 2004; Bagtzoglou and Atmadja, 2005; Sun et al., 2006, and references therein). Briefly, the proposed methods could be grouped into two categories: optimization approaches and probabilistic approaches. The main difference between the two approaches is that the optimization approaches cast the problem as a deterministic one in which parameters are found that minimize a given objective function, whereas the probabilistic approaches cast the problem in a stochastic framework and the parameters to estimate become random variables. In the first category, Gorelick et al. (1983) identified the groundwater pollution source information through an optimization model using linear programming and multiple regression; Wagner (1992) employed a non-liner maximum likelihood method to estimate source location and flux; Mahar and Datta (2000) used a nonlinear optimization model for estimating the magnitude, location and duration of groundwater pollution sources with binding equality constraints; Yeh et al. (2007) developed a hybrid approach, which combines simulated annealing, tabu search and a three-dimensional groundwater flow and solute transport model to solve the source identification problem; and Ayvaz (2010) utilized a harmony search-based simulation-optimization model to determine the source location and release histories by using an implicit solution procedure. In the second category, Bagtzoglou et al. (1992) applied a particle method to estimate, probabilistically, source location and spill-time history; Woodbury and Ulrych (1996) used a minimum relative entropy approach to recover the release and evolution histories of a groundwater contaminant plume in a one-dimensional system; Neupauer and Wilson (1999) employed a backward location model based on adjoint state method (BPM-ASM) to identify a contaminant source; Butera et al. (2013) utilized a simultaneous release function and source location identification (SRSI) method to identify the release history and source location of an injection in a groundwater aquifer; and Koch and Nowak (2016) derived and applied a Bayesian reverse-inverse methodology to infer source zone architectures and aquifer parameters.

The ensemble Kalman filter (EnKF), which could be included in the group of probabilistic approaches mentioned above, has recently addressed the problem of contaminant source

identification. The EnKF introduced by Evensen (2003) has gained much popularity in recent years for its efficiency in solving inverse problems in different fields such as oceanography, meteorology and hydrology (Houtekamer and Mitchell, 2001; Li et al., 2012a; Xu et al., 2013b). The advantages of the EnKF can be summarized as follows (Chen and Zhang, 2006; Zhou et al., 2011): computational efficiency when compared with other inverse approaches, easy integration with different forecast models, ability to account for model and observation errors, and easy uncertainty characterization since the final outcome is always an ensemble of realizations. In hydrogeology, the EnKF has been mainly applied for the identification of aquifer parameters such as hydraulic conductivity or porosity (Li et al., 2012b; Xu et al., 2013a; Zhou et al., 2014; Xu and Gómez-Hernández, 2015, 2016a). Recently, Xu and Gómez-Hernández (2016b) demonstrated the possibility to apply the EnKF for the identification of a contaminant source in a deterministic synthetic aquifer, and later Xu and Gómez-Hernández (2018) showed that the method can be also applied for the simultaneous identification of hydraulic conductivities and the parameters defining a contaminant source also in a synthetic aquifer.

All the works mentioned above were tested in synthetic cases. Only a few works can be found in the literature for laboratory or field cases. Woodbury et al. (1998) extended the minimum relative entropy (MRE) method to recover the release history of a contaminant and applied it to reconstruct the release history of a 1,4-dioxane plume observed at the Gloucester Landfill in Ontario, Canada. Michalak (2003); Michalak and Kitanidis (2004) employed a Bayesian inverse formulation to estimate the contaminant history of trichloroethylene (TCE) and perchloroethylene (PCE) in an aquifer at the Dover Air Force Base, Delaware, a site that had already been analyzed by Liu and Ball (1999) in the same context of source identification. Cupola et al. (2015b,a) compared the source location identification (SRSI) method to the backward probability model based on the adjoint state method (BPM-ASM) with data taken from a sandbox experiment. Zanini and Woodbury (2016) also used data from a sandbox experiment to apply an empirical Bayesian method combined with Akaike's Bayesian

Information Criterion (ABIC) to deduce the release history of a groundwater contaminant.

The main objective of this paper is to assess the performance of the restart EnKF (r-EnKF) for the identification of contaminant source parameters and aquifer geometry with data from a sandbox experiment. The source parameters of interest are the release location, release starting and ending times, and contaminant load, and regarding the geometry the method should try to retrieve the position and length of a plate that is inserted about the center of the sandbox and induces a deflection of the flowlines towards the bottom of the sandbox. The state information assimilated by the r-EnKF is limited to concentration data at a few observation points, since no piezometric head data were available.

The paper is organized as follows, first, the state equations and the fundamentals of the r-EnKF will be recalled, second, the sandbox characteristics are described together with the numerical model used to reproduce its behavior, third, the r-EnKF is tested with data from a synthetic experiment that mimics the sandbox experiment with the aim to verify if the r-EnKF is capable of identifying the kind of parameters sought, and four, the r-EnKF is applied with observation values taken from the sandbox experiment, the problems encountered are analyzed, alternative approaches are discussed and the final results presented. The paper ends with a summary and conclusions on the main findings.

2. Methodology

2.1. Groundwater Flow and Solute Transport Equation

The sandbox will be modeled as a two-dimensional system in the XZ plane, where an inert contaminant spreads due to advection and dispersion under a steady-state flow. The dimension of the sandbox in the y direction is small enough to assume that the state variables are constant along any line for any given (x, z) value. The governing equations are:

$$\frac{\partial}{\partial x}\left(K_x \frac{\partial h}{\partial x}\right) + \frac{\partial}{\partial z}\left(K_z \frac{\partial h}{\partial z}\right) + w = 0 \tag{1}$$

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

where K_x and K_z are the principal components of the hydraulic conductivity tensor in the x and z spatial coordinates respectively $[LT^{-1}]$ which are assumed aligned with the coordinate system of reference in the entire domain; h is the hydraulic head $[L]$; w represents distributed sources or sinks $[T^{-1}]$; t is time $[T]$; θ represents the porosity of the medium; C is dissolved concentration $[ML^{-3}]$; $\nabla\cdot$ is the divergence operator; ∇ is the gradient operator; D represents the hydrodynamic dispersion coefficient tensor $[L^2T^{-1}]$; v is the flow velocity vector $[LT^{-1}]$ derived from the solution of the flow model; q_s represents volumetric flow rate per unit volume of aquifer associated with a fluid source or sink $[T^{-1}]$ and C_s is the concentration of the source or sink $[ML^{-3}]$.

The flow equation is solved using MODFLOW (McDonald and Harbaugh, 1988), and the transport equation is solved using MT3DS (Zheng and Wang, 1999).

2.2. The Ensemble Kalman Filter

The ensemble Kalman filter was first introduced by Evensen (2003) to circumvent the difficulty of propagating covariances in time in the original and extended Kalman filter formulations. The restart EnKF (r-EnKF) has proven its capacity for contaminant source identification in synthetic cases (Xu and Gómez-Hernández, 2016b, 2018); now, we propose to test the r-EnKF in a sandbox experiment. For this specific case, there will be eight parameters to identify, six related to the contaminant source, and two related to aquifer geometry. In the first group, they are the contaminant source location (Xs, Zs), the injection concentration Ic, the injection rate Ir, plus the starting Ts and ending Te release times. In the second group, the algorithm will try to identify the position along the x direction Xb and the total depth Zb of a vertical plate inserted about the center of the sandbox to deflect the flowlines. The rest of the parameters defining the flow and transport conditions in the sandbox are not subject to identification and are equal to their observed values as explained

in the description of the experiment in the next section. The r-EnKF is shortly described next.

In the ensemble Kalman filter with extended state vector, we deal with two types of variables, the system parameters subject of identification, of which there could be observations or not, and the state of the system, of which there will be observations. The state is forecasted in time solving the corresponding state equations, with the latest parameter update, up to the specific time steps when observations are collected; these observations are assimilated by the filter and serve to update the parameters and the state of the system. In the restart filter, state variables are not updated, only system parameters are, because the system state forecast for the next observation time is restarted from time zero to make sure that the forecasted system state is fully coherent with the state equations, and, in our case, with the updated **contaminant source**. (In the original implementation of the filter, both state and parameters are updated, and the state system is forecasted from the last updated state values using the last updated parameters.) The r-EnKF is an iterative algorithm that cycles forecast and data assimilation (with the corresponding parameter update) until all observations have been accounted for. The implementation of the r-EnKF for the identification of the eight parameters described above can be summarized as follows **(Evensen, 2003; Xu and Gómez-Hernández, 2016b)**:

1. Generate an initial ensemble of parameter values. An ensemble of N_e realizations of eight-tuples of the parameters to be identified is generated. Parameter values are drawn, independently, from uniform distributions defined between first-guess minimum and maximum values—there are no restrictions on these uniform distributions, their range can be wider or narrower than the one used in this paper, and they do not have to necessarily contain the "real" value, they are simply used to initialize the algorithm. We build N_e vectors S_i with the eight parameters for each realization:

$$S_i = [Xs_i, Zs_i, Xb_i, Zb_i, Ic_i, Ir_i, Ts_i, Te_i]^T \tag{3}$$

where i is the realization index and the superscript T stands for transpose.

2. Repeat for each system state observation time. Forecast the state. For each ensemble member, forecast the system state, that is, the concentrations in the aquifer, for the t^{th} observation time using the values of the parameters from the last update (or the initial parameters for the first observation time). In the original implementation of the EnKF, the system state at the t^{th} observation time is forecasted based on the concentrations at the $(t-1)^{th}$ observation time and using the last updated parameters; however, it is virtually impossible to account for an update of the source location or the injection time unless the state equation is solved from time zero, thus the need to restart the simulation from time zero (Xu and Gómez-Hernández, 2016b). The forecast of concentrations is given by

$$C_i^f(t) = \psi\left[C_0, S_i^a(t-1)\right], \tag{4}$$

where the superscripts f and a refer to forecasted and updated values after assimilation, respectively; ψ represents the numerical model that forecast, in time, concentrations, on a grid with N_m nodes; C_i is an $N_m \times 1$ column vector containing the forecasted concentrations at all the discretization nodes of the numerical model for realization i; S_i^a is the vector with the last updated parameters; C_0 is the initial contaminant concentration of the domain, which is the same for all realizations. The forecast of the parameters is simply

$$S_i^f(t) = S_i^a(t-1). \tag{5}$$

3. Parameters update. First compute the parameter covariance through the ensemble of forecasted realizations

$$P_S^f(t) = \frac{1}{N_e}\sum_{i=1}^{N_e}\left\{\left[S_i^f(t) - \overline{S_i^f(t)}\right]\left[S_i^f(t) - \overline{S_i^f(t)}\right]^T\right\} \tag{6}$$

with

$$\overline{S_i^f(t)} = \frac{1}{N_e}\sum_{i=1}^{N_e} S_i^f(t), \tag{7}$$

where P_S^f is an 8×8 matrix of parameter covariances and $\overline{S_i^f(t)}$ is an 8×1 column vector of parameter averages. Then, compute the parameter-concentration cross-covariances but only with the forecasted concentration values that fall at concentration observation locations for time t (for the sake of simplicity, we will assume that observations are taken coinciding with some of the numerical model nodes, if not, there will be a need to provide a linear averaging procedure to estimate concentrations at observation locations from model concentration forecasts)

$$P_{SC}^f(t) = \frac{1}{N_e}\sum_{i=1}^{N_e}\left\{\left[S_i^f(t) - \overline{S_i^f(t)}\right]\left[C_i^f(t) - \overline{C_i^f(t)}\right]^T\right\} \tag{8}$$

with

$$\overline{C_i^f(t)} = \frac{1}{N_e}\sum_{i=1}^{N_e} C_i^f(t), \tag{9}$$

where P_{SC}^f is an $8 \times N_o$ matrix of parameter-concentration cross-covariances, with N_o being the number of nodes of the numerical model at which observations are taken at time step t, and $\overline{C_i^f(t)}$ is an $N_o \times 1$ column vector of average concentrations. Next, compute the $8 \times N_o$ Kalman gain matrix $K(t)$ as

$$K(t) = P_S^f(t)[P_{SC}^f(t) + R(t)]^{-1} \tag{10}$$

where $R(t)$ is **an** $N_o \times N_o$ **diagonal** observation error covariance matrix **(implying that there is no correlation between observation errors)** and proceed to update the parameter values, realization by realization by

$$S_i^a(t) = S_i^f(t) + K(t)\left[d_i(t) - C_{io}^f(t)\right], \tag{11}$$

where $d_i(t)$ is an $N_o \times 1$ vector of observed concentrations (including observation errors with covariance given by $R(t)$) and $C_{io}^f(t)$ is an $N_o \times 1$ vector of forecasted concentrations.

4. Go back to step 2 and repeat the whole process until all observations are assimilated.

3. Experimental Case

3.1. Description of the experiment

A single point pollution experiment was performed in a sandbox using sodium fluorescein as tracer. The sandbox is built in plexiglass and has external dimensions of 120 cm × 14 cm × 70 cm as sketched in Fig. 1. The internal volume of 96 cm × 10 cm × 70 cm is filled with constant-diameter spherical glass beads. There are two reservoirs at the edges of the box imposing constant water levels of 60.7 cm and 53.6 m upstream and downstream, respectively. An injector was set up at the upstream part of the sandbox at the location indicated by a red square in the figure, and a plastic plate was vertically inserted inside the glass beads in the middle of the sandbox, whose position and length is also shown in the figure. The experimental equipment was placed in a dark box and a digital camera was used to capture, every 5 s, the fluorescein luminosity within the rectangular zone of 85 cm by 44 cm marked with a ticked rectangle in Fig. 1. The pictures were then processed and the fluorescein luminosity transformed into concentrations after a calibration procedure, as described by Citarella et al. (2015). In this case, eight different fluorescein concentrations ($C = 0; 2.5; 5; 10; 20; 25; 30; 35$ mg/l) were used to calibrate and generate the luminosity-concentration curves in each picture pixel. The total experiment time lasted 1965 s, the injection started at time 120 s and finished at time 1000 s. During the experiment, the rate and concentration of the injection were also recorded.

It is very important to note that there are no piezometric head observations. The design of the tank did not allow for those observations. Had there been piezometric head data, they

could have been assimilated in the filter and, without doubt, would have helped in improving the identification (as shown by Xu and Gómez-Hernández (2018)).

3.2. Numerical Model

Since the thickness of the sandbox along the y axis is relatively small, we can assume that the variability of piezometric heads and concentration along this direction is negligible. Therefore, a two-dimensional groundwater flow and transport model in the XZ plane is built. The upstream and downstream vertical boundaries are set as constant prescribed piezometric head values, and the bottom boundary is impermeable while the top boundary is the phreatic surface. The model corresponds to the yellowish area in Fig. 1, where the coordinates of the four model corners are given. The tank is filled with homogeneous spherical glass beads with a conductivity of 0.58 cm/s and a porosity of 0.37. The vertical plastic plate was inserted at a distance of 52 cm from the left boundary and its length is of 42 cm. It is modeled as an impermeable barrier, which will deflect the flowlines towards the bottom of the sandbox. The sandbox is discretized into 96 columns, one row, and 70 layers; the size of each cell is $(\Delta x, \Delta y, \Delta z) = (1, 10, 1)$ cm. The total simulation time is 1800 s and is discretized into 90 uniform time steps. Citarella et al. (2015) evaluated the longitudinal and transverse dispersivities of the spherical beads, resulting in values of 0.16 cm and 0.048 cm, respectively. The flow and transport parameters are collected in Table 1.

The release happens at coordinates (18.5 cm, 30.5 cm), with a concentration of 20 mg/l and an injection rate of 0.95 cm^3/s. To start the ensemble Kalman filter 800 8-tuples of the source and plate parameters are generated from uniform distributions (not centered at the true values). The true values of the parameters to identify and the suspect range of the uniform distributions used to generate the initial ensemble are collected in Table 2.

4. Application

The objective of this work is to demonstrate the capacity of the r-EnKF for the identification of contaminant source information, including contaminant source location (Xs, Zs), injection information (Ic, Ir) and release time (Ts, Te) together with the position and length of the vertical plate (Xb, Zb), using concentration observations collected in a laboratory experiment. As a prior test, we analyze a synthetic case, in which the concentration data are generated by the numerical model of the sandbox, therefore removing any modeling error since the forward model used to forecast by the r-EnKF will coincide with the model used to generate the observations. In the next section, we will redo the analysis using the laboratory data, we will analyze the problems found and propose some solutions.

4.1. Synthetic Sandbox Test

In this case, we design two scenarios ($S1$, $S2$) with different number of observation wells to evaluate the performance of the r-EnKF: scenario $S1$ with 20 observation wells, and scenario $S2$ with 24 observation wells containing 4 additional wells (#21, #22, #23, #24) located at the four corners of the suspect release area (see Fig. 1). **The rationale for the second scenario is that, after analyzing the first scenario, we felt that additional information about the plume evolution was needed, and thus we decided to add four wells around the suspect release zone. Such an addition will, indeed, improve the characterization.** In both scenarios, model error is neglected and we assume that observation errors are uncorrelated with mean zero, and standard deviation of 0.1 mg/l.

Figure 2 and 3 show the time evolution of the ensemble mean and the ensemble variance, respectively, of the updated state parameters for the two scenarios. Figure 4 shows the evolution in time of the boxplots computed from the 800 ensemble members. **After time step 60, the convergence rate of the means and variances of the parameters are less than 1% and 5%, respectively, all the parameters get close to the final estimation and become stable. Notice also the sudden drop of the variance at a**

given time step for most of the parameters. This drop is related to the activation of new observation wells as time progresses, what implies that the amount of information assimilated by the filter does not vary continuously in time, but rather it increases stepwise, with steps occurring when new wells observe, for the first time, the arrival of the solute plume. We can distinguish between the parameters that are perfectly identified by an ensemble mean equal to the true value, and practically zero variance, and those that are approximated closely but which are not exact and present some residual uncertainty. In the first group, there are the position parameters for the plate, Xb and Zb, plus the vertical location of the release source Zs, independently of whether 20 or 24 data are used during the assimilation steps; in the second group are the remaining parameters, which become more precise (mean closer to the true value) and less uncertain (smaller variability) for $S2$ than for $S1$. The horizontal source location Xs is less sensitive to the concentration data, and only when the four additional data points in the corners of the suspect release location are added the algorithm is able to provide a good estimate for this parameter; similar comment can be made about the beginning Ts and end Te times of the release. The injection concentration Ic and injection rate Ir are well identified by their median values, with smallest uncertainty for S2. These results are consistent with the sensitivity of concentrations at the observation locations to changes in the parameter values: concentration distributions are most sensitive to the position of the plate, which affects the flow field, and the vertical release location, which affects the main trajectory of the contaminant plume, but are less sensitive to the other parameters, for which variations within the identified uncertainty ranges induce concentration changes of the same order of magnitude as the observation errors. Also notice that the horizontal coordinate of the release and the starting and ending release times are correlated for the purpose of identifying their values (a displacement of the horizontal coordinate of the release could be compensated with a displacement of its starting time), what also explains their larger uncertainties.

These results prove that the r-EnKF could work for the identification of a contaminant source and of some parameters defining the geometry of the aquifer. The next step is to test the algorithm under more realistic conditions using observations obtained from a laboratory experiment.

4.2. Laboratory Sandbox Test

The sandbox experiment was carried out as described previously. Figure 5 shows a picture of the fluorescein plume at the 48th time step (840 s since the beginning of the release) already transformed into concentration values and the position of the observation points. The deflection of the flowlines induced by the vertical plate is clearly seen. Notice that only a few observation piezometers will actually detect the plume breakthrough. Before testing the r-EnKF, we performed a simulation of the concentration evolution using the known release parameters and compared the predictions with the observed data. Figure 6 shows a comparison between observed and numerically predicted concentrations at five observation locations (wells #7, #9, #10, #13, #22) through which the plume passes. As can be seen, the reproduction is very good for the closest well #22, and it deteriorates with the distance from the source, but not dramatically, except for well #9. For this well, the beginning and ending times of the breakthrough curve are the same for predictions and observations, but the mismatch in concentrations indicates either some error in the model parameters or faulty observations. The predicted breakthrough curve in the farthest well, though, is quite close to the observed one. In the application that follows we will analyze different observation error distributions in an attempt to identify the source parameters by the r-EnKF.

We have run the r-EnKF with three different magnitudes of the observation error, which will be referred to as $R1$, $R2$, and $R3$. In all three cases, the error mean is zero and its standard deviation is 0.5 mg/l for $R1$, 1.0 mg/l for $R2$, and 3.0 mg/l for $R3$. **We must notice that in previous experiments, Cupola et al. (2015a) report an observation**

error with a standard deviation around 1 mg/l.

The hydraulic conductivity value of the beads, which is considered homogeneous in each realization, is considered uncertain and drawn from a Gaussian distribution with a mean of 0.58 cm/s and a standard deviation of 0.05 cm/s. **We have decided to introduce some uncertainty on the beads hydraulic conductivity as a surrogate to model error. The choice of a Gaussian distribution centered at the calibrated conductivity value was arbitrary, any other distribution could have been used. Considering that the differences in the results between including or not such an uncertainty are minimal (and not reported here), we believe that the choice of the specific distribution has little effect in the final outcome.**

Fig. 7 shows the boxplots of the updated parameters at different time steps for the three scenarios $R1$, $R2$, and $R3$. The results are not as good as for the synthetic case, for which the observed concentrations were generated with the same numerical model used for the forecast step in the Kalman filter. The first thing to note is that for scenario $R1$, the use of a small observation error makes the r-EnKF to seek for source parameter values that can be far from the true ones in order to produce concentrations that are close to the observed values, and, particularly, the injection concentration and injection rate do not seem to converge to a stable value after 90 time steps. The other parameters do reach a stable median, not as close to the true values as for the synthetic case but close enough except for the horizontal position of the vertical plate.

When the observation error is increased (scenario $R2$), the two main findings are that the two injection parameters now seem to reach a stable estimate (albeit with large uncertainty) with a median close to the true value, and that all parameters have a wider uncertainty range. The median estimate of the initial and ending release times is also closer to the true ones than in $R1$. The horizontal position of the vertical plate continues to be underestimated, as well as the length of the plate.

When the observation error is increased even more (scenario $R3$) the main effect is that

the final estimates have wide uncertainty estimates, and for some of the parameters it seems as if the concentration observations do not bring any added value since the boxplot width remains unaltered through the assimilation steps. The estimates of the parameters by their median is comparable to the results in $R2$, but their uncertainty is larger.

The predicted concentrations at three observation wells that were not used during the assimilation step computed using the initial 8-tuples of parameters, and using the 8-tuples obtained at the end of the three scenarios are shown in Figure 8. The figure shows the true concentrations in the sandbox as a dotted blue line, each one of the 800 predicted concentration breakthrough curves computed with the 8-tuples of the ensemble, along with their median, as a red line, and their 90% confidence interval, as dashed lines. It can be observed that, prior to assimilation (top row), concentration predictions were very scattered, and that after the assimilation (bottom three rows, one for each scenario) the breakthrough curves change substantially (compare, for instance, the median curves). For scenario $R1$, the scatter of prediction curves is the smallest but recall that these wells were not used during the assimilation, the updated parameters were biased because the algorithm tried to fit the observed concentrations too closely and as a result, at the control wells, the prediction of the true curves by the ensemble median is also biased, up to the point that the true curves are outside the 90% confidence interval. For scenarios $R2$ and $R3$ the median curves for the three wells have a smaller bias than for $R1$, and the main difference between $R2$ and $R3$ is the same as for parameter prediction, the uncertainty is the widest for $R3$. The true curve is in both cases within the 90% confidence interval of the predictions.

At this point, it seems that an observation error with a standard deviation of 1 mg/l was the most consistent with our observations and model. **As mentioned above, this conclusion fits the findings by (Cupola et al., 2015b).** Yet, we were concerned with the big discrepancy between predictions and observations at well #9, so we decided to rerun scenario $R2$ without using the data from this well. The results for this scenario, called $R2b$, are shown in Figure 9. When comparing this figure to the middle two rows in Figure 7 we

can notice that there is some overall improvement in the estimation of the true parameters —particularly for the position parameters— by the median values of the ensemble without a significant change on their uncertainty. This improvement reinforces our suspicion that there could have been some problems in the data collection at well #9.

We also considered that there could be a problem with the tightness of the vertical plate after its insertion in the sandbox. **The plate was supposed to represent a perfectly impermeable barrier, and no evidence of the contrary was observed during the experiments, yet the contact between the plate and sandbox walls could have had some small gaps, making the plate slightly permeable.** Therefore, we decided to rerun scenario $R2$ but assuming that the plate is slightly permeable, more precisely, with a conductivity of two orders of magnitude smaller than the beads, **this value was chosen arbitrarily low since nothing was actually observed in the laboratory**. The results for the new scenario, referred to as $R2c$ are shown in Figure 10. (Note that well #9 was kept in this scenario.) The main difference of this run is that the estimate of the size of the vertical plate by the median of the ensemble jumps from 40.5 cm to 44.2 cm (true value is 42.0 cm) indicating that possibly the plate conductivity used in this scenario was too large and, as a consequence, the algorithm enlarges the plate to reproduce the observed concentrations. **This result, while does not serve to justify that the tightness of the plate explains the numerical model misfit, shows the impact that such permeability would have in the estimation of the remaining parameters defining the plate.**

We can conclude that the r-EnKF can be applied to a more realistic case of a homogeneous aquifer in a sandbox for the identification of a contaminant source and some geometry parameters. A proper evaluation of the observation errors is paramount, since attempting to match too closely the data may result in biased estimates of the parameters.

5. Summary and Conclusion

The main purpose of this paper was to test whether the restart ensemble Kalman filter, which had been successfully applied in synthetic experiments, could be applied to a more realistic case based on a sandbox experiment. The test focuses on the identification of the parameters defining a finite-pulse point injection of a solute, together with the position of a vertical plate that modifies the initial rectangular geometry of the sandbox.

As a preliminary step, we tested the r-EnKF in a synthetic case mimicking the sandbox. Under these very controlled conditions, the algorithm performs well, as expected. The main difference with previous synthetic analyses is that no piezometric head data were used during the assimilation step of the filter.

Then, the r-EnKF is tested using the data coming from the laboratory experiment. In this case, the observations were not generated by a computer code nor we knew the observation error magnitude. The analysis of the results show that using a too small observation error variance results in more or less precise but biased estimates, both for the parameters subject to identification and for the concentrations at control locations. When a larger observation error (with a standard deviation of 1 mg/l) is introduced, estimates and predictions improve, although with larger uncertainty. And finally, when the observation error is large, the results worsen considerably. The removal of a suspicious observation well, the concentration of which is always underestimated by our forecast model, improves the results, indicating that the measurements from such well may need to be reconsidered. The changes observed after making the vertical plate slightly permeable do not appear to justify the hypothesis that the plate leaks.

The r-EnKF appears as a good algorithm for source identification in aquifers, yet it still needs further tests in closer-to-reality conditions. Currently, the sandbox has been replaced with a heterogeneous distribution of glass beads, and the challenge is to test the method in this new sandbox.

Acknowledgements Financial support to carry out this work was received from the Spanish Ministry of Economy and Competitiveness through project CGL2014-59841-P, and from the Spanish Ministry of Education, Culture and Sports through a fellowship for the mobility of professors in foreign research and higher education institutions to the second author, reference PRX17/00150. The authors also would like to thank Università degli Studi di Parma for providing the experimental equipment and data.

References

Atmadja, J., Bagtzoglou, A.C., 2001. State of the Art Report on Mathematical Methods for Groundwater Pollution Source Identification. Environmental Forensics 2, 205–214.

Ayvaz, M.T., 2010. A linked simulation-optimization model for solving the unknown groundwater pollution source identification problems. Journal of Contaminant Hydrology 117, 46–59.

Bagtzoglou, A.C., Atmadja, J., 2005. Mathematical Methods for Hydrologic Inversion: The Case of Pollution Source Identification. Water Pollution 5, 65–96.

Bagtzoglou, A.C., Dougherty, D.E., Tompson, A.F.B., 1992. Application of particle methods to reliable identification of groundwater pollution sources. Water Resources Management 6, 15–23.

Butera, I., Tanda, M.G., Zanini, A., 2013. Simultaneous identification of the pollutant release history and the source location in groundwater by means of a geostatistical approach. Stochastic Environmental Research and Risk Assessment 27, 1269–1280.

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

Citarella, D., Cupola, F., Tanda, M.G., Zanini, A., 2015. Evaluation of dispersivity coefficients by means of a laboratory image analysis. Journal of Contaminant Hydrology 172, 10–23.

Cupola, F., Tanda, M.G., Zanini, A., 2015a. Contaminant release history identification in 2-d heterogeneous aquifers through a minimum relative entropy approach. SpringerPlus 4, 656.

Cupola, F., Tanda, M.G., Zanini, A., 2015b. Laboratory sandbox validation of pollutant source location methods. Stochastic Environmental Research and Risk Assessment 29, 169–182.

Evensen, G., 2003. The Ensemble Kalman Filter: Theoretical formulation and practical implementation. Ocean Dynamics 53, 343–367.

Gorelick, S.M., Evans, B., Remson, I., 1983. Identifying sources of groundwater pollution: An optimization approach. Water Resources Research 19, 779–790.

Houtekamer, P.L., Mitchell, H.L., 2001. A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation. 0203058.

Koch, J., Nowak, W., 2016. Identification of contaminant source architectures - A statistical inversion that emulates multiphase physics in a computationally practicable manner. Water Resources Research 52, 1009–1025. 2014WR016527.

Li, L., Zhou, H., Gómez-Hernández, J.J., Hendricks Franssen, H.J., 2012a. Jointly mapping hydraulic conductivity and porosity by assimilating concentration data via ensemble Kalman filter. Journal of Hydrology 428-429, 152–169.

Li, L., Zhou, H., Hendricks Franssen, H.J., Gómez-Hernández, J.J., 2012b. Groundwater flow inverse modeling in non-MultiGaussian media: Performance assessment of the normal-score Ensemble Kalman Filter. Hydrology and Earth System Sciences 16, 573–590.

Liu, C., Ball, W.P., 1999. Application of inverse methods to contaminant source identification from aquitard diffusion profiles at dover afb, delaware. Water Resources Research 35, 1975–1985.

Mahar, P.S., Datta, B., 2000. Identification of Pollution Sources in Transient Groundwater Systems. Water Resources Management 14, 209–227.

McDonald, J.M., Harbaugh, A.W., 1988. A modular three-dimensional finite-difference flow model. Techniques of Water Resources Investigations of the U.S. Geological Survey, Book 6 , 586.

Michalak, A.M., 2003. A method for enforcing parameter nonnegativity in Bayesian inverse problems with an application to contaminant source identification. Water Resources Research 39, 1–14.

Michalak, A.M., Kitanidis, P.K., 2004. Estimation of historical groundwater contaminant distribution using the adjoint state method applied to geostatistical inverse modeling. Water Resources Research 40.

Neupauer, R.M., Wilson, J.L., 1999. Adjoint method for obtaining backward-in-time location and travel time probabilities of a conservative groundwater contaminant. Water Resources Research 35, 3389–3398.

Sun, A.Y., Painter, S.L., Wittmeyer, G.W., 2006. A constrained robust least squares approach for contaminant release history identification. Water Resources Research 42, 1–13.

Wagner, B.J., 1992. Simultaneous parameter estimation and contaminant source characterization for coupled groundwater flow and contaminant transport modelling. Journal of Hydrology 135, 275–303.

Woodbury, A., Sudicky, E., Ulrych, T.J., Ludwig, R., 1998. Three-dimensional plume source

reconstruction using minimum relative entropy inversion. Journal of Contaminant Hydrology 32, 131–158.

Woodbury, A.D., Ulrych, T.J., 1996. Minimum relative entropy inversion: Theory and application to recovering the release history of a groundwater contaminant. Water Resources Research 32, 2671–2681.

Xu, T., Gómez-Hernández, J.J., 2015. Inverse sequential simulation: A new approach for the characterization of hydraulic conductivities demonstrated on a non-Gaussian field. Water Resources Research 51, 2227–2242. 2014WR016527.

Xu, T., Gómez-Hernández, J.J., 2016a. Characterization of non-Gaussian conductivities and porosities with hydraulic heads, solute concentrations, and water temperatures. Water Resources Research 52, 6111–6136.

Xu, T., Gómez-Hernández, J.J., 2016b. Joint identification of contaminant source location, initial release time, and initial solute concentration in an aquifer via ensemble Kalman filtering. Water Resources Research .

Xu, T., Gómez-Hernández, J.J., 2018. Simultaneous identification of a contaminant source and hydraulic conductivity via the restart normal-score ensemble Kalman filter. Advances in Water Resources 112, 106–123.

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

Xu, T., Jaime Gómez-Hernández, J., Li, L., Zhou, H., 2013b. Parallelized ensemble Kalman filter for hydraulic conductivity characterization. Computers and Geosciences 52, 42–49.

Yeh, H.D., Chang, T.H., Lin, Y.C., 2007. Groundwater contaminant source identification by a hybrid heuristic approach. Water Resources Research 43, 1–16.

Zanini, A., Woodbury, A.D., 2016. Contaminant source reconstruction by empirical Bayes and Akaike's Bayesian Information Criterion. Journal of Contaminant Hydrology 185-186, 74–86.

Zheng, C., Wang, P.P., 1999. MT3DMS: A Modular Three-Dimensional Multispecies Transport Model , 219.

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

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

Table 1: Parameters of the groundwater flow and transport model

Hydr. conduct., K	0.58 cm/s
Porosity, ϕ	0.37
Long. disp., α_L	0.16 cm
Transv. disp., α_T	0.048 cm

Table 2: Source and geometry parameters. True values and suspect ranges for the generation of the initial ensemble of realizations

Parameter	Actual Value	Suspect Range
Xs (cm) - x-coordinate of source	18.5	$16 - 25$
Zs (cm) - z-coordinate of source	30.5	$23 - 32$
Xb (cm) - x-coordinate of plate	52.5	$50 - 59$
Zb (cm) - plate length	42.5	$35 - 43$
Ir (cm^3/s) - injection rate	0.95	$0.6 - 1.1$
Ic (mg/l) - injection load	20	$5 - 24$
Ts (s) - starting release time	120	$80 - 260$
Te (s) - ending release time	1000	$960 - 1140$

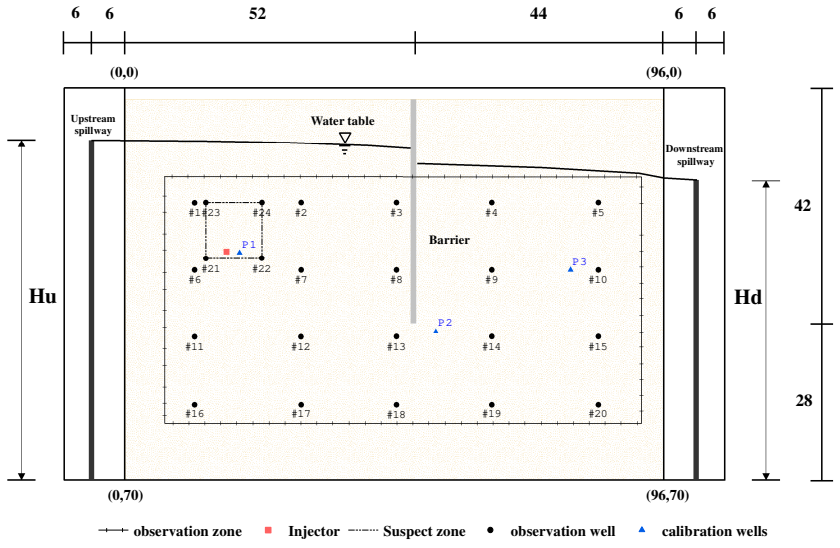

Figure 1: Sketch of the experimental device with indication of the upstream (Hu) and downstream (Hd) constant head boundaries. The ticked rectangle corresponds to the area captured by the camera in which concentrations will be monitored. Red dot is the release location. Dashed line around red dot indicates the release suspect location. Dimensions are in cm. Coordinates of the four corners of the flow and transport models are also shown.

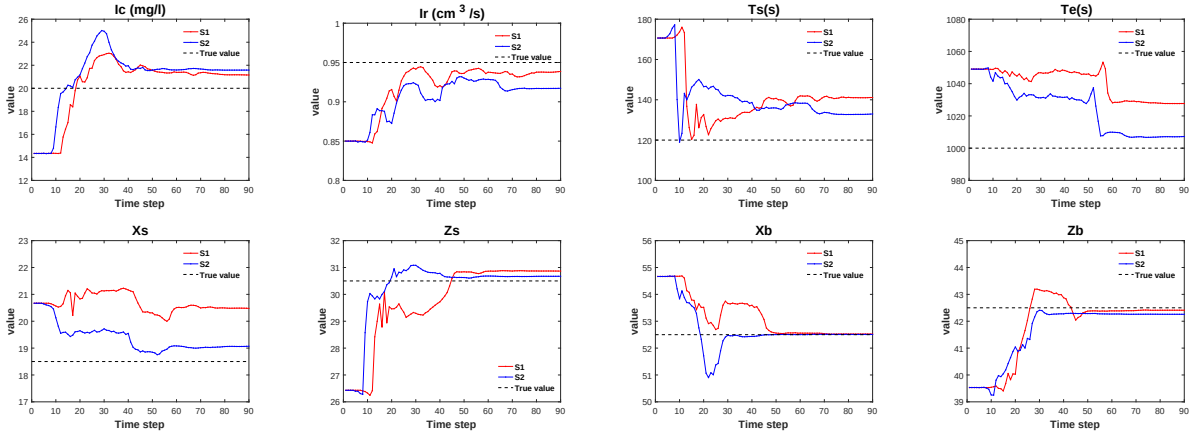

Figure 2: Time evolution of the ensemble mean of the 8 updated parameters, contaminant source location (Xs, Zs), plate position (Xb, Zb), injection information (Ic, Ir) and release time interval (Ts, Te) for scenarios $S1$ and $S2$

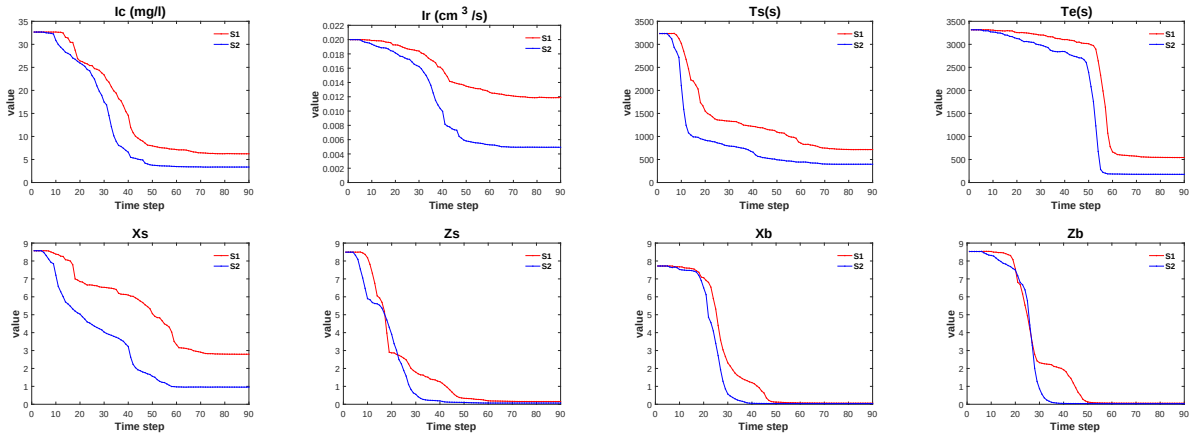

Figure 3: Time evolution of the ensemble variance for the same parameters and scenarios as in the previous figure.

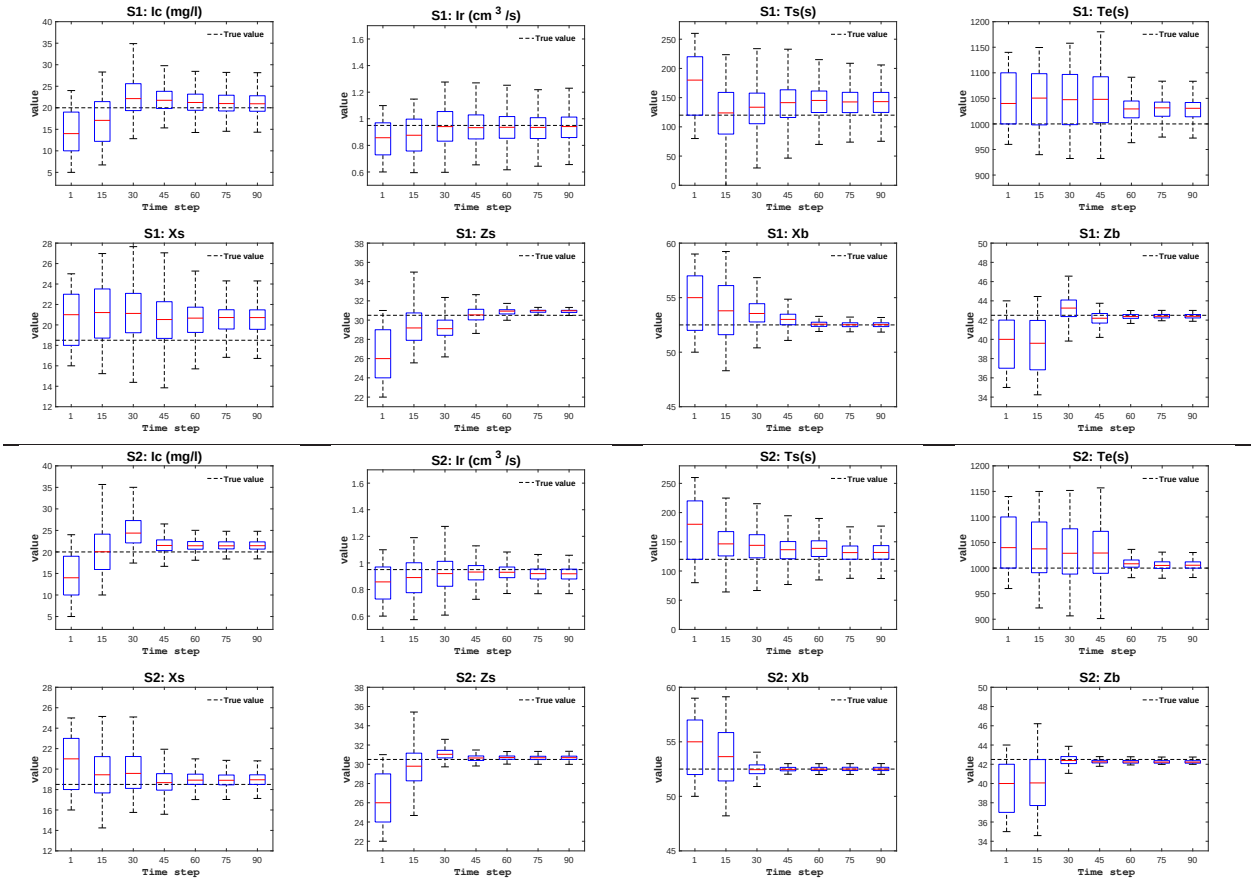

Figure 4: Boxplot of the 8 updated parameters at different time steps (1, 15, 30, 45, 60, 75, 90) for scenarios $S1$ and $S2$

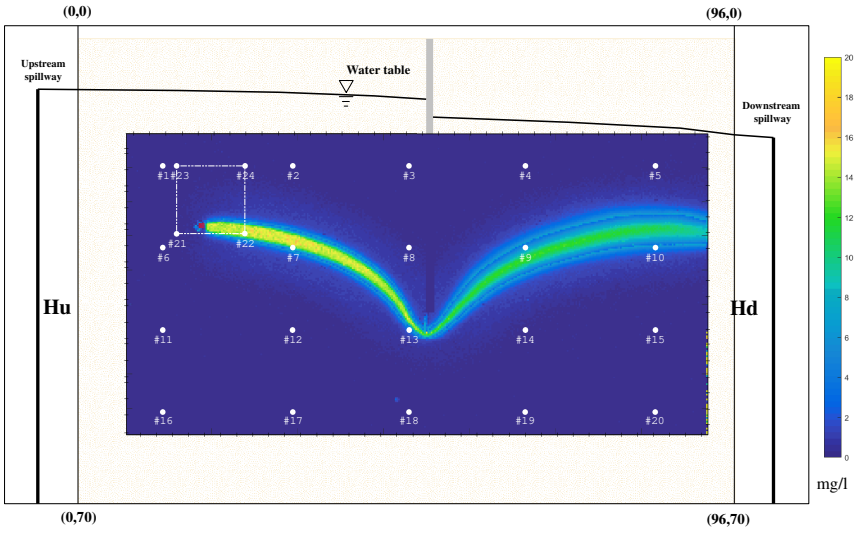

Figure 5: Fluorescein concentration field in the sandbox at the 48th time step. The area shown corresponds to the observation zone indicated in Fig. 1. The dash line shows the suspect zone for the injection and the white dots indicate the observation wells.

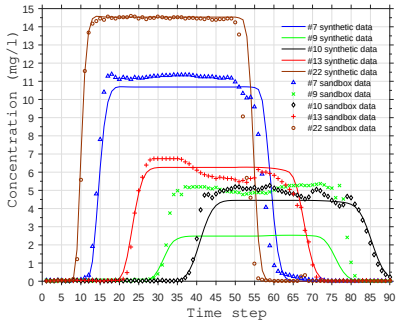

Figure 6: Fluorescein observed breakthrough curves at the observation wells located inside the plume and the curves computed from the numerical model

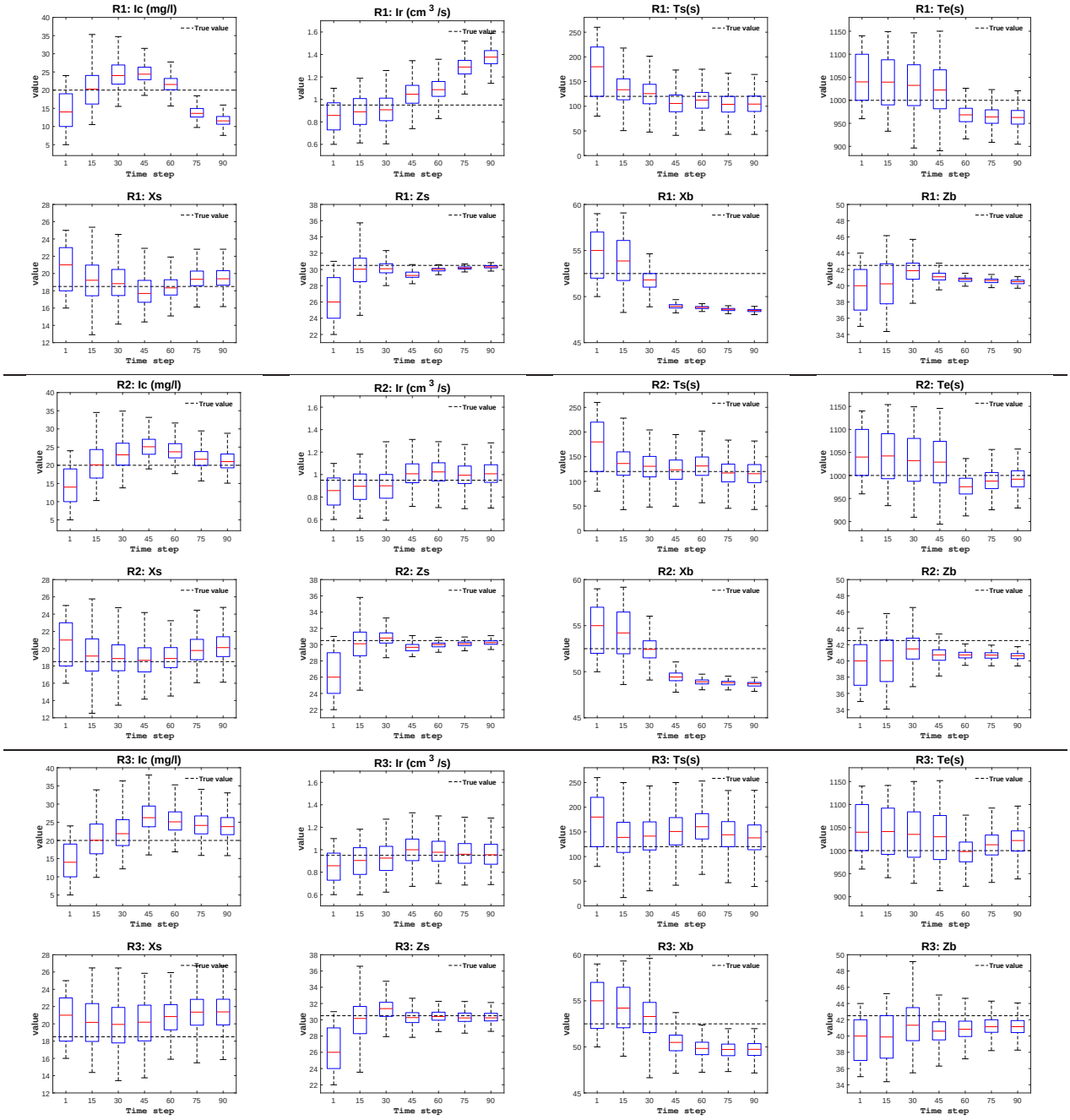

Figure 7: Boxplot of of the 8 updated parameters at time steps 1, 15, 30, 45, 60, 75 and 90 for scenarios $R1$, $R2$ and $R3$

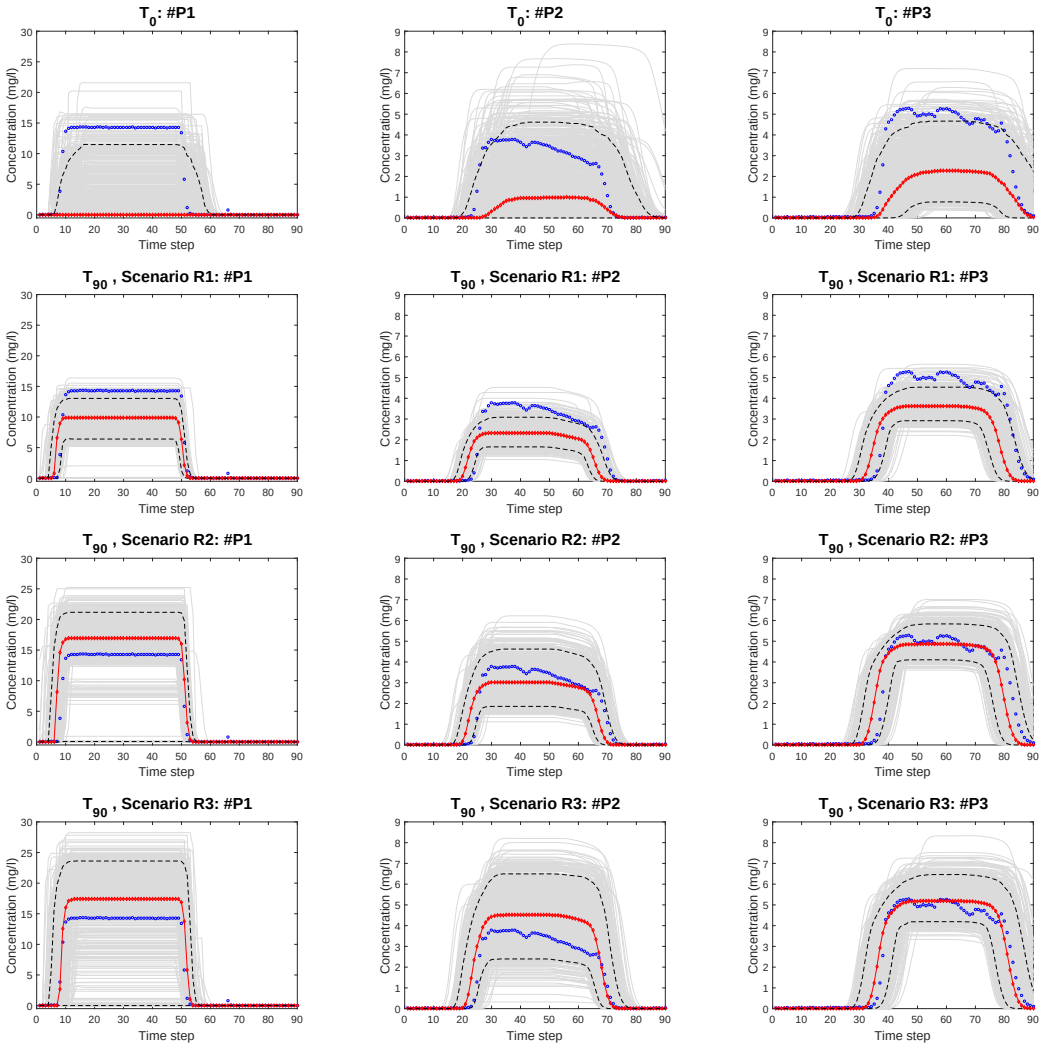

Figure 8: Breakthrough curves at control wells. The blue dots correspond to the curves in the sandbox experiment. The thin gray lines are the curves for all 800 realizations; they are summarized by their median (red diamond lines) and their 5 and 95 percentiles (black dash lines).

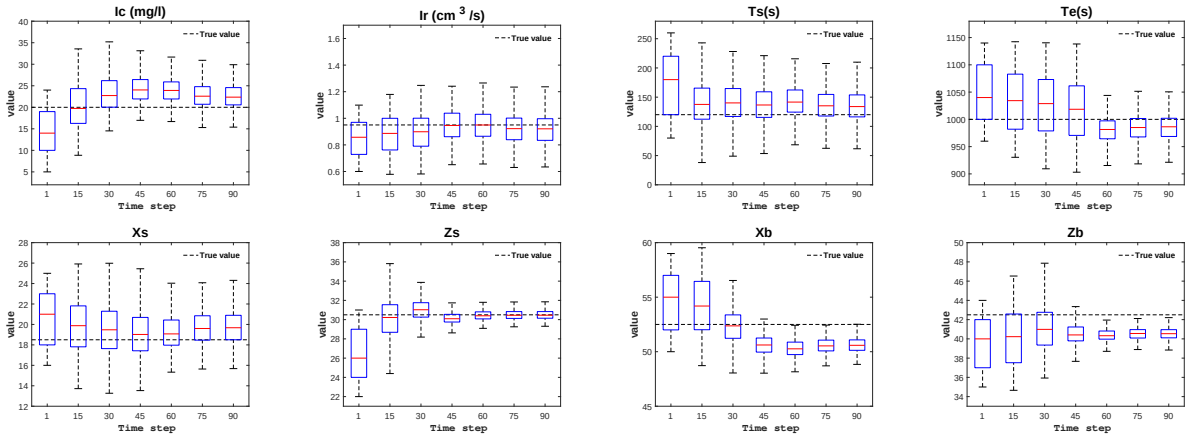

Figure 9: Boxplot of the 8 updated parameters in scenario *R2b* at different time steps (1, 15, 30, 45, 60, 75, 90)

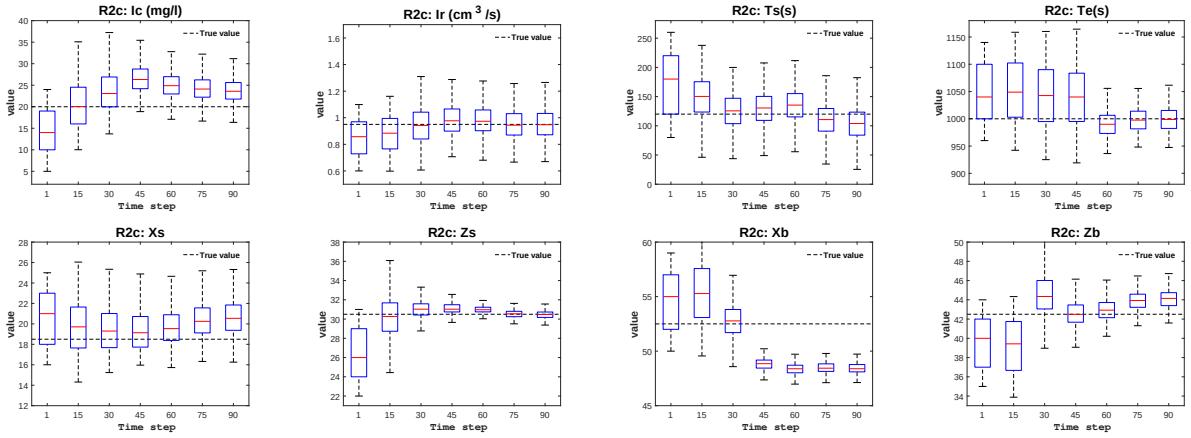

Figure 10: Boxplot of the 8 updated parameters in scenario *R2c* at different time steps (1, 15, 30, 45, 60, 75, 90)