Reconstructing the release history of a contaminant source with different precision via the ensemble smoother with multiple data assimilation

Reconstructing the release history of a contaminant source with different precision via the ensemble smoother with multiple data assimilation

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

[a] *Nanjing Geological Survey Center, China Geological Survey, Nanjing, China*
[b] *State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing, China*
[c] *Institute of Water and Environmental Engineering, Universitat Politècnica de València, Valencia, Spain*
[d] *Department of Engineering and Architecture, Università degli Studi di Parma, Parma, Italy*

Abstract

Identifying a contaminant time-varying release history is an ill-posed problem but crucial for groundwater contamination issues. A precise inversed release history offers a promising estimation of contaminant movement and is of great importance for environmental monitoring and further management. In this paper, a recent emerging data assimilation method, the ensemble smoother with multiple data assimilation (ES-MDA) is employed to handle this conundrum. The study starts with some synthetic cases in which several factors are analyzed, such as the observation data frequency, covariance inflation schemes, iteration numbers used in the ES-MDA for the purpose of identifying a time-varying contaminant injection event with different precision. The results show that the ES-MDA performs well in recovering the release history when the injection is discretized into 50 or 100-time steps but encounters fluctuation problems in the cases with 300-time steps. Further comparison reveals that the observation data frequency is a very influential factor, while the number of iterations or the kind of covariance inflation used has a lesser effect.

Keywords: Inverse modeling, Source identification, inflation factor, Data assimilation,

*Corresponding author
Email address: teng.xu@hhu.edu.cn (Teng Xu)

1. Introduction

Groundwater contamination has gained extensive attention over the last several decades (e.g., Feyen et al., 2003b; Li et al., 2011; Feyen et al., 2003a; Gómez-Hernández et al., 2003) since it is becoming a huge threat to our ecosystem. Determining the responsible for the pollution is a forensic hydrogeology task needed to ensure the accountability of those responsible. This is not an easy task, since, in general, only a few observations downstream from the source are available when the contamination is first detected. Even with the help of advanced groundwater models, and with assumptions such as knowing the release location, identifying the release history, and, therefore, the total amount of pollutants injected into the aquifer, has proven to be a complicated endeavour. A challenge that faces the problem of ill-posedness (Skaggs & Kabala, 1994; Carrera & Neuman, 1986) common to all inverse problems (Franssen & Gómez-Hernández, 2002; Capilla et al., 1998; Wen et al., 1999) . Various methods have been devised to address this problem and several reviews have been published in the subject (e.g., Atmadja & Bagtzoglou, 2001; Michalak & Kitanidis, 2004; Bagtzoglou & Atmadja, 2005; Sun et al., 2006; Gómez-Hernández & Xu, 2021).

Among all these methods, one branch, data assimilation methods, comes out ahead because of its ability to deal with huge amounts of observed data simultaneously. Data assimilation methods are versatile, efficient, and simple to understand and implement (Zhou et al., 2014). Among the data assimilation methods, the ensemble Kalman filter (EnKF) stands out. It was first proposed by Evensen (2003) in order to deal with the nonlinear relationship between parameters and state variables in inverse problems and has gained popularity in multidisciplinary fields such as oceanography, meteorology, and geology (e.g., Houtekamer & Mitchell, 2001; Bertino et al., 2003; Chen & Zhang, 2006; Aanonsen et al., 2009). Specifically, in hydrogeology, the EnKF method has proven the ability to inverse identify aquifer parameters, such as hydraulic conductivity (Chen & Zhang, 2006; Huang et al., 2009; Kurtz et al., 2014), porosity (Li et al., 2012), recharge rates (Franssen & Kinzelbach, 2009), boundary conditions (Chen & Zhang, 2006) and also transport-related parameters (Lan et al., 2018). More recently, researchers have started to employ EnKF variants to identify the parameters describing a contaminant source in aquifers. Butera et al. (2013) employ a geostatistical approach with some weak hypotheses to identify the pollutant release history and the source location. Xu & Gómez-Hernández (2016) use the restart normal-score Ensemble Kalman filter (Ns-EnKF) for contaminant source identification in a synthetic deterministic aquifer and later extended this method to jointly identify hydraulic conductivity and source information (Xu & Gómez-Hernández, 2018). Then, Chen et al. (2018) move one step further, to identify contaminant source information plus the position and length of a vertical barrier in a sandbox experiment via the restart Ensemble Kalman filter. Chen et al. (2018) also discuss the influence of different inflation methods in the application of the restart Ns-EnKF and prove its ability for the joint identification of hydraulic conductivities and contaminant source information in a laboratory sandbox experiment. Li et al. (2019) used Kalman filtering combined with a mixed-integer nonlinear programming optimization model to deduce the accurate location and release history of a contaminant source. The aforementioned works are a strong demonstration that the EnKF and its variants are valid methods for contaminant source identification. However, except for the work by Butera et al. (2013), the release history identified in these works only focuses on a constant pulse, the magnitude of which is independent of time.

As an alternative to the EnKF, the ensemble smoother (ES), which was first introduced by van Leeuwen & Evensen (1996), assimilates all available data in one single step instead of updating the state variable sequentially. Thus, it is expected that it should be able to identify time-varying parameters better than the EnKF (and at a cheaper price). The EnKF and the ES produce the same results when they deal with linear state-transfer functions since they are based on the same covariance-based formulation (Evensen, 2004). However, in studying process with strong nonlinearities, such as in the case of inverting the groundwater flow and mass transport equations, the EnKF outperformed the ES (Evensen & van Leeuwen, 2000), until an iterative variant of the ES was proposed, the ES with Multiple Data Assimilation (ES-MDA), by Emerick & Reynolds (2013). Evensen (2018) compared the ES-MDA with other iterative ensemble smoothers to solve history matching problems. Ranazzi & Sampaio (2019) investigated the influence of the ensemble size on the use of an adaptive ES-MDA for history matching. Todaro et al. (2019) use the ES-MDA to find a solution to the reverse flow routing problem. Bao et al. (2020) coupled Generative Adversarial Networks and ES-MDA methods, then use them to reconstruct the channel structures and reduce the uncertainty of hydraulic head and contaminant concentration predictions. Xu et al. (2021) employed the ES-MDA to identify contaminant source parameters and heterogeneous hydraulic conductivity jointly with the comparison with restart EnKF. Todaro et al. (2021) employed ES-MDA for the simultaneous identification of the source location and the release history of a groundwater contamination event. These works are all good examples of ES-MDA dealing with time-varying input parameters. However, most of the aforementioned work approximated the time-varying parameters by a multiple step function without analyzing the impact that the step size had in the results.

In this work, the ES-MDA is employed to identify a time-varying release history in both synthetic and real cases, the capacity of ES-MDA in identifying a release history function as a function of the discretization used to approximate it is analyzed first. Then, a synthetic case is studied where the influence of observation data frequency and number of ES-MDA iterations are discussed. The synthetic also served to analyze two covariance inflation procedures (e.g., Le et al., 2016; Rafiee & Reynolds, 2017) to prevent smoother collapsing. Next, the ES-MDA is applied to the identification of release history functions in two sandbox experiments. The paper is organized as follows: in section 2, we describe the methodology; in section 3, the synthetic and the real sandbox experiment are presented, followed by the setup of different scenarios and evaluation criteria. Finally, in section 4, we discuss the results and draw some conclusions.

2. Methodology

2.1. Groundwater flow and solute transport equations

In this work, the contaminant is injected with a given flow rate into a transient groundwater flow system. Thus, the governing equations includes both the transient groundwater flow equation (Bear, 1972) and the solute transport equation (Zheng & 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 represents the specific storage $[L^{-1}]$; h is the hydraulic head $[L]$; t denotes time $[T]$; $\nabla\cdot$ is the divergence operator, while ∇ represents the gradient operator; K denotes the hydraulic conductivity $[LT^{-1}]$ and w represents distributed sources or sinks $[T^{-1}]$, θ represents the porosity of the medium [-]; C is dissolved concentration $[ML^{-3}]$; 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 equation; 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 $[ML^{-3}]$.

2.2. Ensemble Smoother with Multiple Data Assimilation(ES-MDA)

As we mentioned before, the ES-MDA is an improvement of the ES made by Emerick & Reynolds (2013) for handling nonlinear models. It is an iterative version of the ES where the number of iterations is predefined. The method is easy to understand and to implement and has been referred many times in the literature (Emerick & Reynolds, 2013; Evensen, 2018; Xu et al., 2021). A brief recall of the three steps that conform the method are described next.

1. Initialization step.

An ensemble of N_e realizations of the n parameters to identify is generated. In this case, the parameters are the mass loadings in time representing the discretized injection curve; their initial values are drawn from predefined uniform distributions. (Each ensemble member is a different release history function.) At this stage, we also need to set the number of iterations N_a (also referred to as assimilation steps), and the inflation factors α_j; the meaning of which are described later.

2. Assimilation.

Once the number of iterations and the inflation coefficients are determined, it is time for the assimilation procedure, which consists of two steps, a forecast step, and an update step. These two steps are repeated for each iteration.

a. Forecast step

In this step, the groundwater flow and contaminant transport models, MODFLOW (McDonald & Harbaugh, 1988) and MT3DS (e.g., Zheng, 2010; Ma et al., 2012), are run for each member of the ensemble; in our case, for each different release history,

$$C^f_{i,j} = \psi[C_0, A_{i,j}], \tag{3}$$

where ψ represents the forward numerical model, $C^f_{i,j}$ are the predicted concentrations (in space and time) at assimilation iteration j for the last estimate of the release function i of the ensemble, $A_{i,j}$. The size of A depends on the number of time steps used to discretize it.

b. Update step

Then, the model parameters are updated as follows,

$$A_{i,j+1} = A_{i,j} + \Delta A_j (\Delta C_j^f)^T [\Delta C_j^f (\Delta C_j^f)^T + \alpha_j R]^{-1} [y_{obs} + \sqrt{\alpha_j}\varepsilon - C_{o,i,j}^f], \tag{4}$$

where y_{obs} is a column vector with dimensions $N_o \cdot N_t$ containing all observed concentrations at all locations and all time steps (N_o is the number of locations, and N_t the number of observation steps); ε stands for the observation error, while R is the covariance matrix of the observation error; $C_{o,i,j}^f$ is the vector of forecasted concentrations for the ensemble parameter set $A_{i,j}$ at the same locations and times where and when observations y_{obs} are made; ΔA_j and ΔC_j are matrices defined as

$$\Delta A_j = \frac{1}{\sqrt{N_e - 1}} [A_{1,j} - \overline{A}_j, A_{2,j} - \overline{A}_j, \ldots, A_{N_e,j} - \overline{A}_j], \tag{5}$$

$$\Delta C_j^f = \frac{1}{\sqrt{N_e - 1}} [C_{1,j}^f - \overline{C^f}_j, C_{2,j}^f - \overline{C^f}_j, \ldots, C_{N_e,j}^f - \overline{C^f}_j], \tag{6}$$

where $\overline{A}_j$ and $\overline{C^f}_j$ are the ensemble means of source release history parameters and forecasted concentrations at the j_{th} iteration, respectively. The products $\Delta C_j^f (\Delta C_j^f)^T$ and $\Delta A_j (\Delta C_j^f)^T$ are the concentration covariance and the concentration-release function parameters cross-covariance, respectively.

These forecast and update steps will be repeated until the predefined iterations are completed. One more thing needs to be pointed out: in our study, since the number of measurements is larger than the ensemble size, it is necessary to employ the truncated singular value decomposition (TSVD) method to compute a pseudo-inverse in Eq. (4).

2.3. The inflation factors α_j

The iteration number (N_a) and the inflation factor (α_j) are two influential parameters in the performance of the ES-MDA, which are related to one another. Emerick & Reynolds (2013) have proven that the ES-MDA could sample the posterior probability distribution function of the parameters precisely only in a linear model and only if the inflation factors α_j satisfy the following equation,

$$\sum_{j=1}^{N_a} \frac{1}{\alpha_j} = 1, \tag{7}$$

There are still many options on how to choose the α_j parameters satisfying the previous equation. Apparently, choosing a decreasing series may be the most appropriate, but some authors claim that using uniform values gives similar results, and that choosing these values arbitrarily may lead to filter collapse (Le et al., 2016). We have decided to explore two methods to select the inflation factors, one proposed by Rafiee & Reynolds (2017), and the other one proposed by Evensen (2018).

Rafiee & Reynolds (2017) propose that the inflation factor for the first iteration is computed as

$$\alpha_1 = (\frac{1}{N}\sum_{i=1}^{N} \lambda_i)^2, \tag{8}$$

where N is the minimum of N_e and $N_o \cdot N_t$, and λ_i are the singular values of matrix D_j given by

$$D_j = R^{-\frac{1}{2}} \triangle C_j^f. \tag{9}$$

The subsequent inflation factors are chosen in a geometrical decreasing progression,

$$\alpha_j = \beta^{j-1}\alpha_1, \tag{10}$$

where β is the ratio that fulfills that the sum of the inverse of the inflation factors equals one (Eq. (7))

$$\frac{1-(1/\beta)^{N_a-1}}{1-1/\beta} = \alpha_1. \tag{11}$$

Evensen (2018) define the inflation factors on the basis of two numbers, a nonzero value $\alpha_1^{'}$ and a geometrical ratio α_{geo}; with these two numbers, a sequence is built according to the following procedure

$$\alpha'_{j+1} = \frac{\alpha'_j}{\alpha_{geo}}, \quad (12)$$

which is then normalized to provide the α_j values that satisfy Eq. (7)

$$\alpha_j = \alpha'_j(\sum_{j=1}^{N_a} \frac{1}{\alpha'_j}) \quad (13)$$

This scheme has the capacity of defining the inflation factors as uniform, in an increasing sequence or in a decreasing one by choosing an α_{geo} equal, below or above one, respectively. Here, we define α_{geo} and α'_1 with the values of 2 and 1, respectively.

In this work, these two different schemes of generating the inflation factors are employed, and their impact is discussed.

3. Applications

A numerical model based on real sandbox experiments is used to demonstrate the proposed method. This sandbox equipment was built up by the Engineering and Architecture Department at the University of Parma, and has been employed in several groundwater contamination studies (Citarella et al., 2015; Cupola et al., 2015; Zanini & Woodbury, 2016). In this work, first, we generated synthetic data using this numerical model to test the ES-MDA method for the identification of a time-varying release history curve. In the synthetic case, we also analyze the impact of the choice of the method to choose the inflation factors, the number of iterations, the size of the observation time intervals, and the degree of discretization with which the release curve is represented in the numerical model. Then, we tested the ES-MDA with real observation data and analyzed the impact of the observation error magnitude.

3.1. Sandbox Set-up

The sandbox has an internal volume of 95 cm by 10 cm by 70 cm and is discretized into 95 columns, 1 row, and 70 layers. The reference hydraulic field inside the sandbox is shown in Figure 1. The reservoirs upstream and downstream are set up as constant piezometric boundaries with a water level of 62.5 cm and 60.6 cm, respectively. The bottom of the sandbox is regarded as a no-flow boundary while the top of the sandbox is a phreatic surface. An injector was installed inside the glass beads that discharges fluorescein during the experiment. Contaminant concentrations are observed in 25 observation points. The details about the acquisition of the concentration data could be found in Citarella et al. (2015); Cupola et al. (2015). The total experiment time is 3000 s and the injection starts at time zero. The main hydraulic parameters used for the simulation are listed in Table 1.

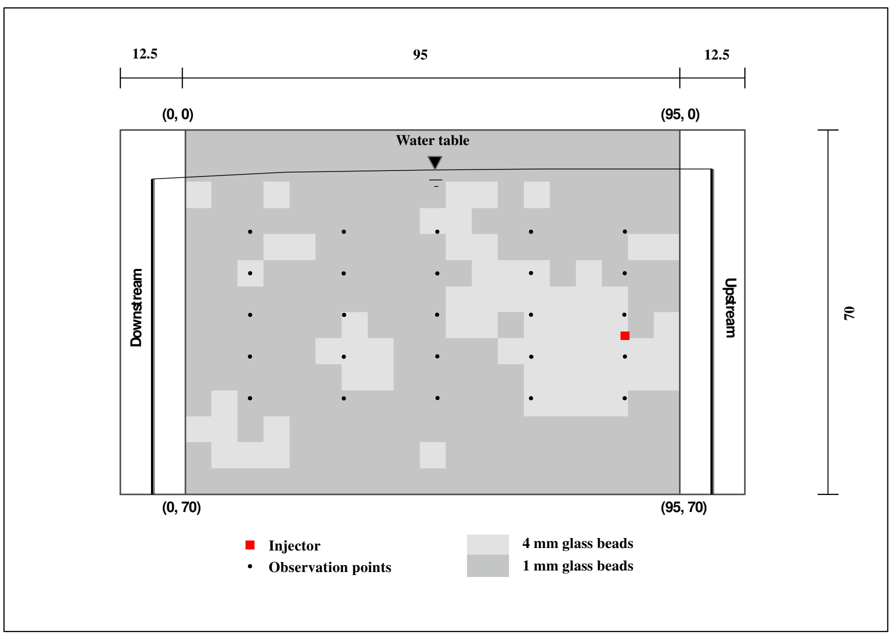

Figure 1: Sketch of the experimental device (lateral view). Length unit is cm.

Table 1: Parameters of the groundwater flow and transport models

	1 mm glass beads	4 mm glass beads
Hydraulic conductivity (cm/s)	0.65	10.4
Longitudinal dispersivity, α_T (cm)	0.106	0.2
Porosity	0.37	0.37
TRVT, α_T/α_L	0.45	0.45

3.2. Performance Assessment

The use of an ensemble-based method allows to analyze the performance of the method using the root mean square error (RMSE) and the relative RMSE:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(A^{ref} - \overline{A}_i)^2}, \tag{14}$$

$$\text{relative } RMSE = \frac{RMSE}{\text{intial } RMSE}, \tag{15}$$

where n is the number of points used to discretize the release history curve, A^{ref} is the reference release history while $\overline{A}$ stands for the ensemble mean of the updated release history, initial $RMSE$ refers to the $RMSE$ of the initial ensemble of realizations.

Based on the definition of RMSE, the smaller the value, the better. The relative RMSE is able to show the reduction of the uncertainty. Both parameters serve to evaluate quantitatively the outcome of ES-MDA.

3.3. Synthetic Case

The first set of analyses is based on the synthetic simulation of a time-varying release into the sandbox digital twin. The release function adopted is based on a proposal by Skaggs

& Kabala (1994):

$$
\begin{aligned}
S(t) = {} & 2.6 \cdot \exp(-\frac{(\frac{t}{10} - 20)^2}{50}) \\
& + 0.78 \cdot \exp(-\frac{(\frac{t}{10} - 50)^2}{200}) \\
& + 1.3 \cdot \exp(-\frac{(\frac{t}{10} - 90)^2}{98}) \qquad 0 \leq t \leq 3000.
\end{aligned} \tag{16}
$$

This function is shown in Figure 2. We run three sets of scenarios with different time discretizations while identifying the release history. More precisely, we chose to identify a release function over the 3000 s experiment duration using 50, 100, and 300 time steps. For each discretization, two sampling frequencies were considered: samples were taken every other time step or every ten time steps. Also, the number of assimilation iterations was varied between 4 and 8, and both the Rafiee and Evensen inflation schemes were tested. In total 24 scenarios were analyzed as reported in Table 2. And in all scenarios, the model error is neglected while we assume the observation errors follow Gaussian distribution with a mean of 0 and standard deviation of 0.1 mg/l.

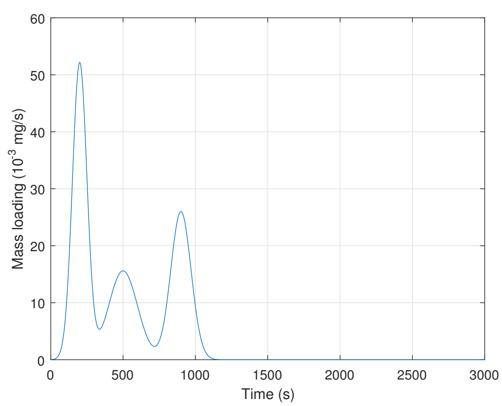

Figure 2: Release curve of a synthetic contaminant source.

An ensemble of 500 realizations was used. The initial release history curve of every realization is generated using a uniform distribution with ranges [0, 52] 10^{-3} mg/l.

Figure 3 shows the recovered release history for the set of scenarios with the coarsest discretization of the release function: 50 time steps. In each plot, the blue curve corresponds

Table 2: Definition of the synthetic scenarios

Scenario	Number of discr. time steps	Number of time observations	Number of iterations	Inflation factor
S1	50	5	4	Rafiee's scheme
S2	50	5	4	Evensen's scheme
S3	50	5	8	Rafiee's scheme
S4	50	5	8	Evensen's scheme
S5	50	25	4	Rafiee's scheme
S6	50	25	4	Evensen's scheme
S7	50	25	8	Rafiee's scheme
S8	50	25	8	Evensen's scheme
S9	100	10	4	Rafiee's scheme
S10	100	10	4	Evensen's scheme
S11	100	10	8	Rafiee's scheme
S12	100	10	8	Evensen's scheme
S13	100	50	4	Rafiee's scheme
S14	100	50	4	Evensen's scheme
S15	100	50	8	Rafiee's scheme
S16	100	50	8	Evensen's scheme
S17	300	30	4	Rafiee's scheme
S18	300	30	4	Evensen's scheme
S19	300	30	8	Rafiee's scheme
S20	300	30	8	Evensen's scheme
S21	300	150	4	Rafiee's scheme
S22	300	150	4	Evensen's scheme
S23	300	150	8	Rafiee's scheme
S24	300	150	8	Evensen's scheme

to the actual release history, the gray lines are the recovered release history curves for all 500 realizations, the red dotted line is the median of the ensemble and the black dashed lines mark the 5 and 95 percentiles. The first column uses Rafiee's inflation and the second column Evensen's inflation. The first two rows use samples every ten time steps (5 snapshots), and the last two rows samples every other time step (25 snapshots). The first and third rows use four iterations and the second and fourth rows use eight iterations. It can be observed that the median of the recovered release history curves is a good estimate of the actual release history for almost all cases (scenarios S2 and S4 being the exception), while the uncertainty estimate given by the spread of the curves is larger for the scenarios with the smallest sampling frequency (scenarios S1 to S4). Also, it can be noticed that Rafiee's inflation method always yields a smaller spread than Evensen's one. It is hard to argue about an improvement with the largest number of iterations since the results with four and eight iterations are almost the same.

Figure 4 shows the recovered release history of the set of scenarios with intermediate discretization of the release function: 100 time steps. The organization of the plots in the figure is the same as in the previous one. The impact of the inflation scheme, the observation data frequency, and the number of iterations is more or less the same as for the 50-time step case. However, the median of the recovered release history curves cannot capture the actual release history as precisely as in the previous set of realizations, more notably in the set of scenarios with samples every 10 time steps (scenarios S9 to S12). For all scenarios, there is clearly an excess of fluctuations in the recovered release curves, noticeable in the individual curves and also in the ensemble median and percentile curves. This fluctuation is more noticeable when the observation sampling frequency is smaller (scenarios S9 to S12). The fluctuations must be due to the inherent ill-posedness of the problem since we are trying to estimate a large number of parameters that, initially, are assumed to be independent. This problem could be alleviated by introducing some smoothing factor that forces that all updated curves after updating display a certain smoothness. It is also important to notice the poor estimation of the release curve at the end of the experiment, with a clear non-zero estimation for the final steps. This overestimation, which is less patent in the previous set of scenarios, must be due to the little or no sensitivity that observations have to release at the end of the simulation.

The deterioration in the estimation of the release curves becomes exacerbated when the number of discretization steps is increased up to 300. Figure 5 shows the results for scenarios S17 to S24, and their arrangement follows the same pattern as the previous two figures. The original release curves are only hinted at by the final ensemble of realizations or their median values, the main three peaks are well identified, but several other peaks appear, the spread of the realizations is very wide and the fluctuations in time are also quite noticeable. As in the previous set of scenarios, using a different parameterization of the release curve enforcing some kind of regularization might have helped in removing these artifacts. The only positive conclusion from this set of realizations is that, as in the previous two sets, the best results are always obtained when using Rafiee's inflation scheme, eight iterations, and the highest sampling frequency.

For a more quantitative evaluation of the performance of the ES-MDA to recover the time-varying release history, Table 3 and Figure 6 illustrates the RMSE and the relative RMSE of all 24 scenarios. Based on the RMSE at the last iteration step, we can conclude that the ES-MDA with Rafiee's scheme has a better performance in most scenarios for our case, especially when the observation data frequency is low. It is striking to see how the RMSE jumps to up to four times the RMSE of the initial ensemble on the first iteration for the fine discretization scenarios (last row of Figure 6), a distinct mark of ill-posedness in the formulation of the problem.

Based on this analysis, we decide to apply the ES-MDA to the sandbox experiment using Rafiee's inflation scheme, discretizing the release history into 50 or 100 time steps, and with

Table 3: RMSE of the synthetic scenarios at the last iteration step

Scenario	RMSE	Scenario	RMSE	Scenario	RMSE
S1	2.295	S9	3.621	S17	12.585
S2	2.136	S10	5.057	S18	15.181
S3	1.979	S11	4.222	S19	8.839
S4	1.818	S12	5.671	S20	12.103
S5	1.120	S13	1.711	S21	9.221
S6	1.178	S14	2.475	S22	8.963
S7	1.321	S15	1.891	S23	7.321
S8	1.182	S16	1.959	S24	6.853

Table 4: Definition of the sandbox scenarios for the train of pulses

Scenario	Number of discr. steps	Number of observ. time steps
R1	50	5
R2	50	25
R3	100	10
R4	100	50

8 assimilation iterations.

3.4. Laboratory Case

We performed two sandbox experiments with two release history curves. The first curve displays a train of four pulses lasting the entire duration of the experiment and the second curve consists of two pulses at the beginning of the experiment (Figure 7). In this experiment, we will not attempt to identify simultaneously the release and the conductivities, but rather, we will use the identified distribution of conductivities and observation errors from a previous work (Chen et al., 2021) shown in Figure 8. The observation errors follow a Gaussian distribution with zero mean and a standard deviation of 1 mg/l. Several scenarios will be analyzed that are described in Tables 4 and 4.

Figure 9 shows the recovered release history curves for the first sandbox experiment, the train of pulses. The observed performance is quite similar to the one observed for the synthetic experiments; the scenario with the smaller number of discretization steps and the

Table 5: Definition of the sandbox scenarios for the two pulses

Scenario	Number of discr. steps	Number of observ. time steps
R5	50	5
R6	50	25
R7	100	10
R8	100	50

highest frequency for that discretization is the one performing best. The same fluctuation as in the synthetic cases is observed about the four peaks of the release curve and the same uncertainty spread, which is smaller for scenario R2. Looking closer to this scenario, we can notice that the identification of the four pulses has a shift in time of a couple of time steps as if the injection had started a little bit later than in reality.

Figure 10 shows the recovered release curves for the second experiment, the two pulses. The same behavior as before is appreciated here. Large fluctuations about the two main peaks of the injection, with the best estimation by the median of the scenario with the smallest number of discretization steps and the largest frequency of observation for that discretization. Yet, there is a major failure in this test case in that the method is not able to capture the fact that the injection stops slightly before the middle of the experiment (at about 1200 s). In all scenarios, most injection curves for the individual members of the ensemble display positive values, and their median is still a relatively large positive value for the second half of the experiment, clearly overestimating the total mass injected into the system. The increase of values towards the end of the experiment is also quite noticeable. The main explanation for this overall behavior is the magnitude of the concentration observation error variance.

Table 6 and Figure 11 show the evolution of the RMSE and relative RMSE for the two sandbox scenarios. The results prove that the ES-MDA with Raffee's inflation scheme is an effective method in recovering releasing history in both sandbox experiments. The RMSE for all 8 scenarios is reduced after assimilating the observation data. A comparison among the

Table 6: The RMSE of sandbox scenarios at final iteration step

Scenario	RMSE	Scenario	RMSE
R1	24.073	R5	18.706
R2	22.419	R6	18.289
R3	24.318	R7	19.123
R4	24.878	R8	20.461

different sandbox scenarios also shows that a high observation data frequency has a lesser impact on the outcome than in the synthetic cases. This observation is particularly obvious for scenarios R3, R4, and R7, R8, which are the cases with 100 time steps; we believe it is mainly caused by the uncertainty about the observation data.

For a further evaluation of the results, we use the updated release history to generate contaminant plume evolution to visually analyze the capacity of the ES-DMA with Raffee's inflation scheme to reproduce the real plumes. Figure 12 and Figure 13 show the ensemble means of the contaminant plumes at time steps 600 s, 1200 s, 1800 s, and 2400 s for scenarios R2, R4, R6 and R8. We can observe that the simulated plume always spreads more than the reference, a consequence of the overestimation of the non-injection period. However, compared with the reference contaminant plume in the left column, the simulated plumes from the updated release history for the 4 scenarios are all acceptable reproductions of the reference.

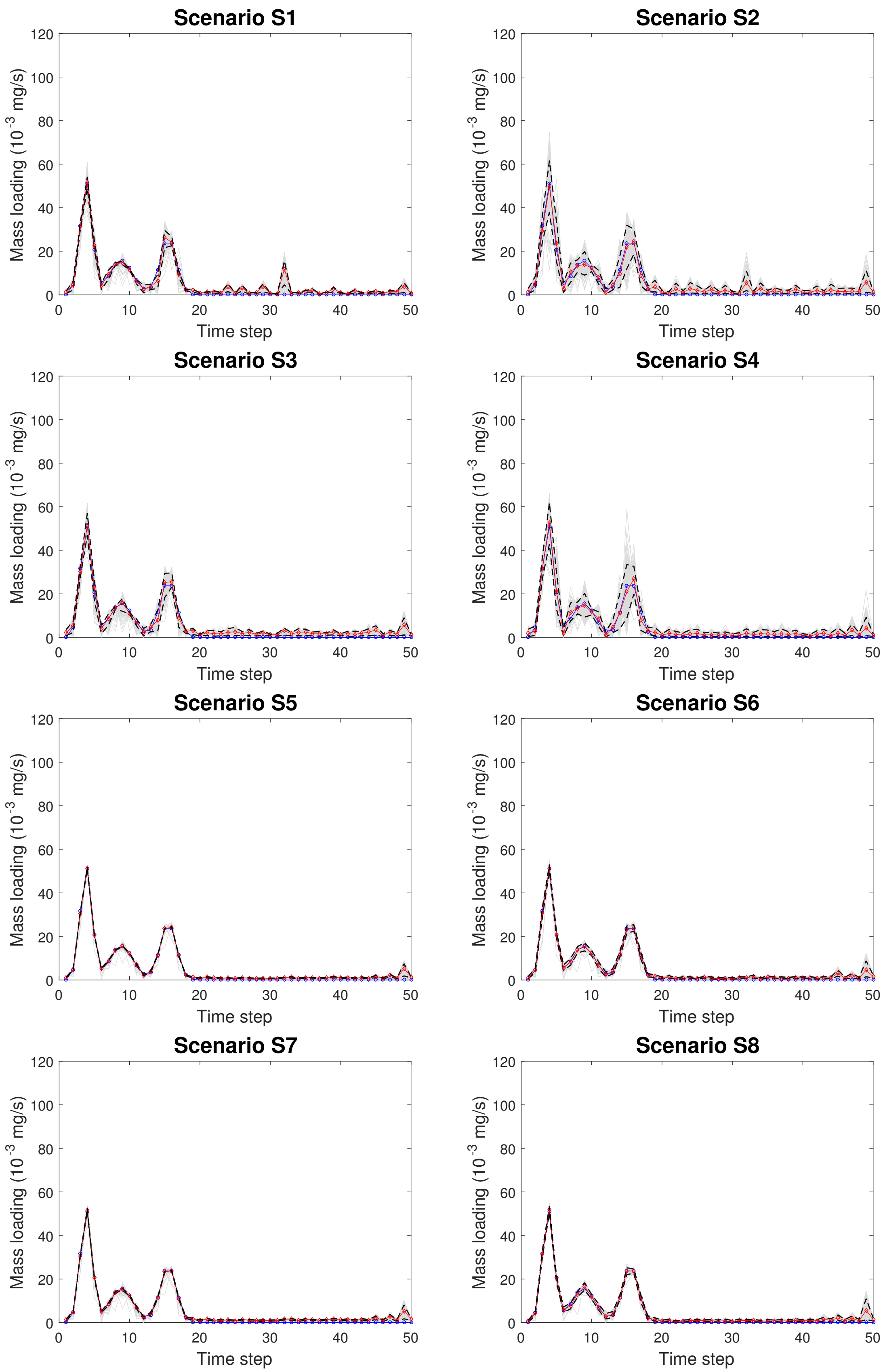

Figure 3: Recovered release histories for scenarios S1 to S8. The blue curve corresponds to the actual release history. The gray lines are the recovered release history curves for all 500 realizations, the red dotted lines is the median, and black dashed lines the 5 and 95 percentiles.

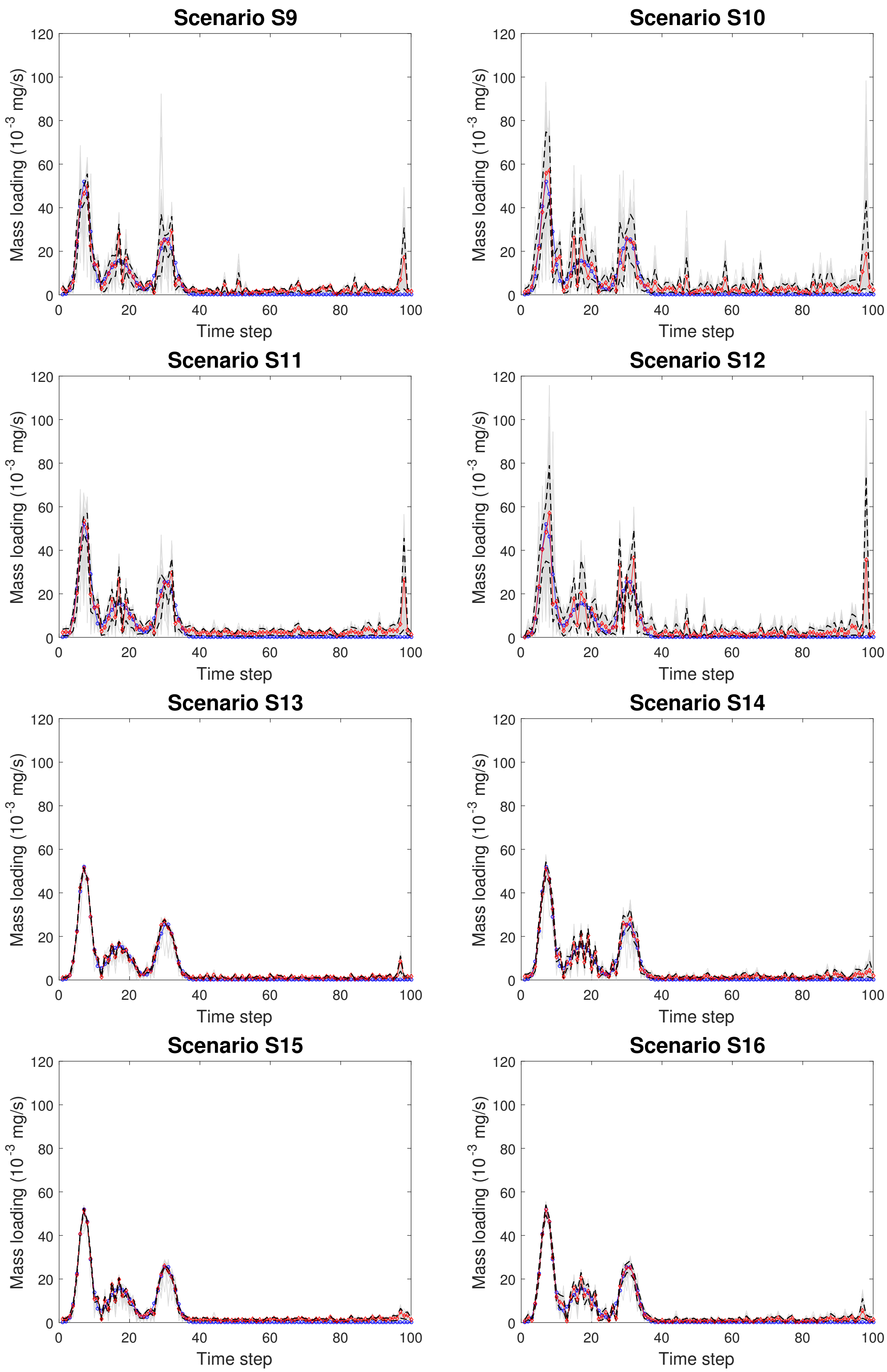

Figure 4: Recovered release histories for scenarios S9 to S16. The blue curve corresponds to the actual release history. The gray lines are the recovered release history curves for all 500 realizations, the red dotted lines is the median, and black dashed lines the 5 and 95 percentiles.

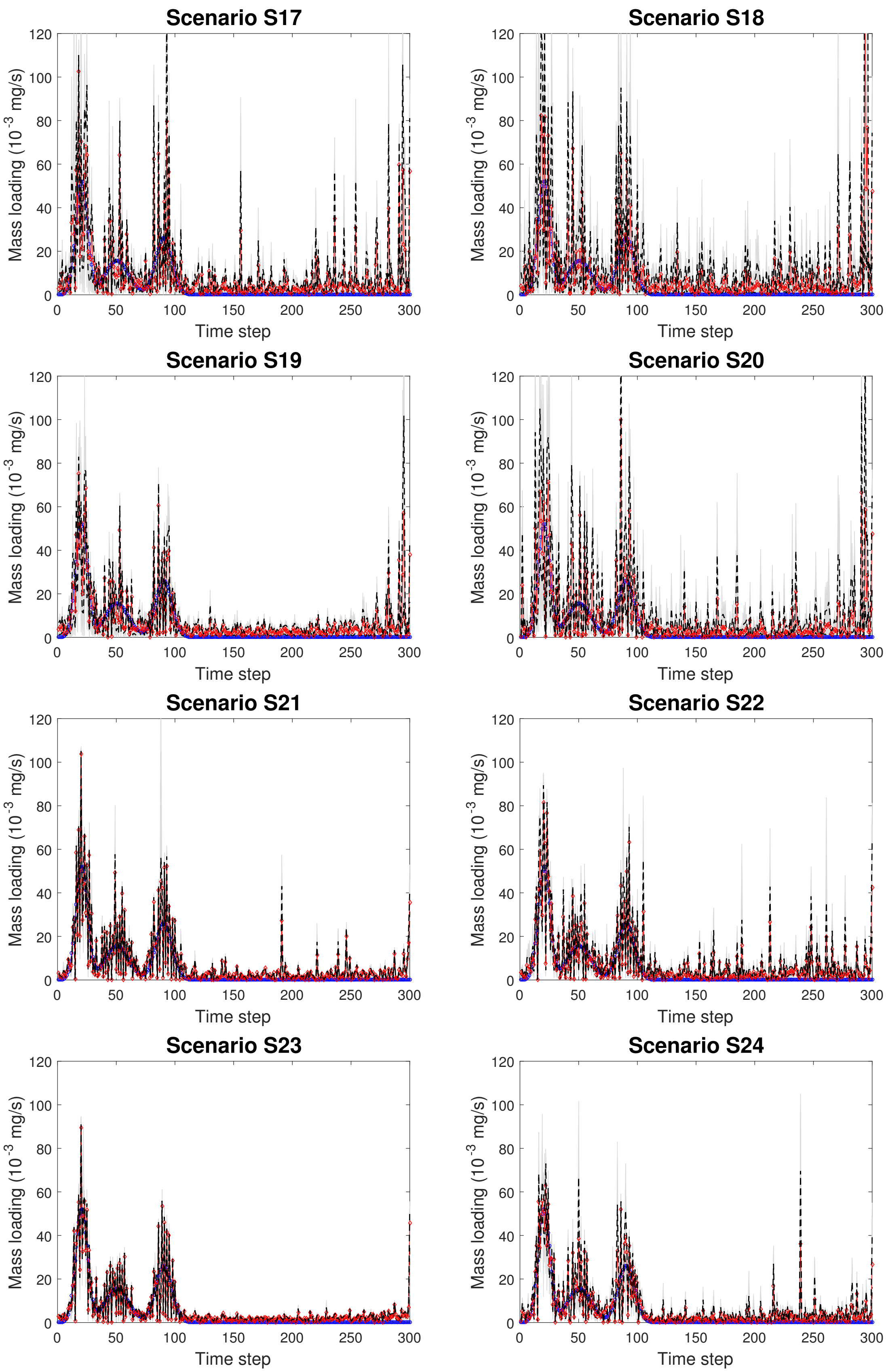

Figure 5: Recovered release histories for scenarios S9 to S16. The blue curve corresponds to the actual release history. The gray lines are the recovered release history curves for all 500 realizations, the red dotted lines is the median, and black dashed lines the 5 and 95 percentiles.

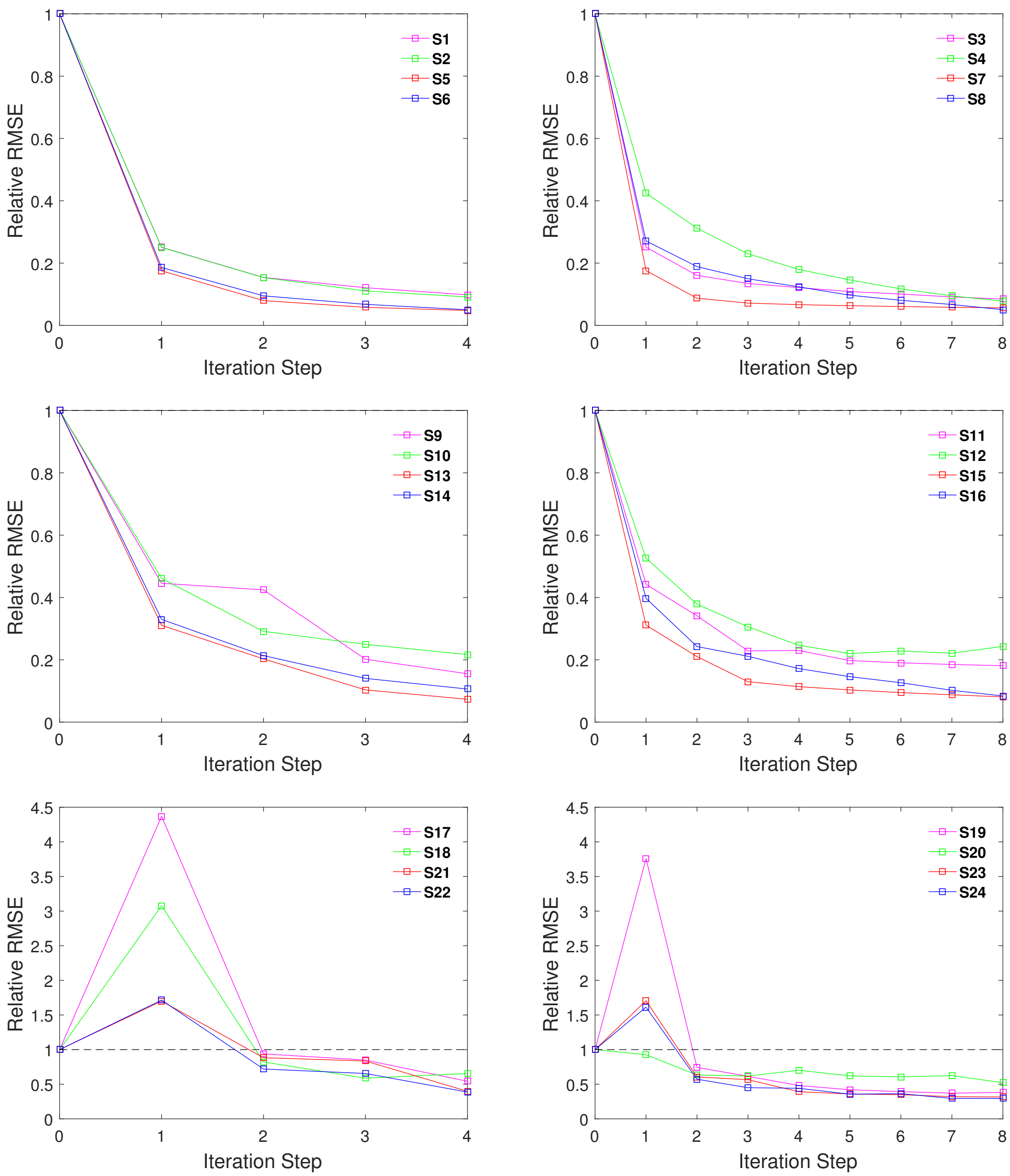

Figure 6: Evolution of the Relative RMSE for the synthetic scenarios as a function of the iteration step.

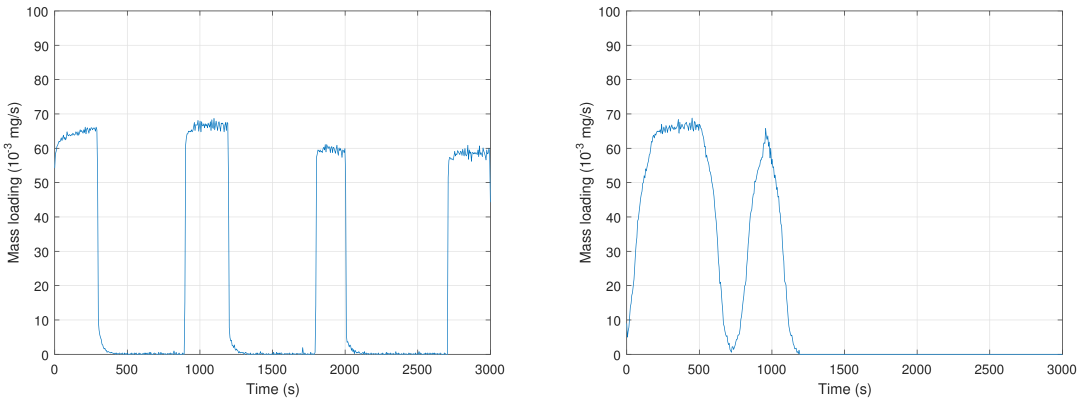

Figure 7: Release history curves for the two sandbox experiment.

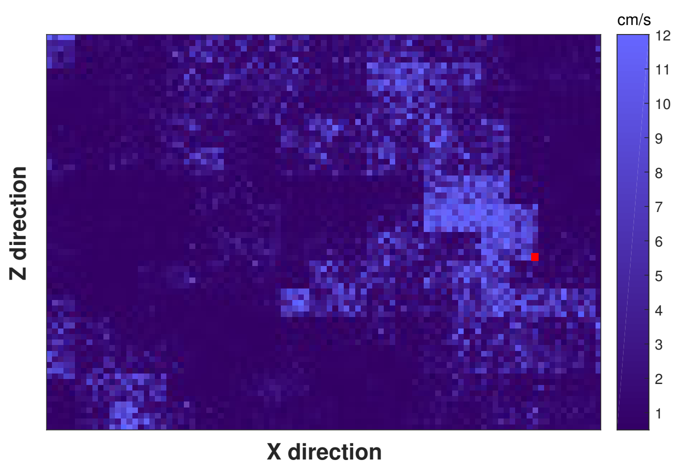

Figure 8: Hydraulic conductivity field. The red square denotes the source location. Flow and transport are from right to left.

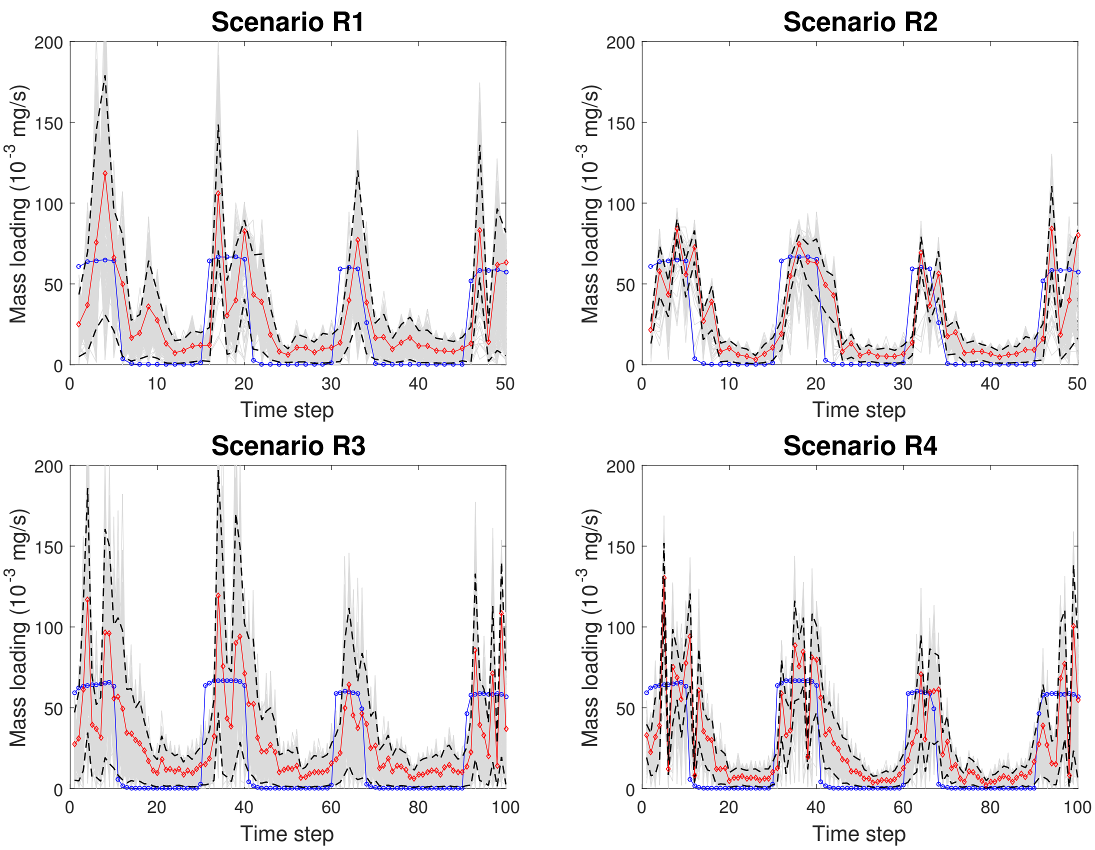

Figure 9: Recovered release history for first sandbox experiment, scenarios R1 to R4. The blue curve corresponds to the actual release history. The gray lines are the recovered release history curves for all 500 realizations, the red dotted lines is the median, and black dashed lines the 5 and 95 percentiles.

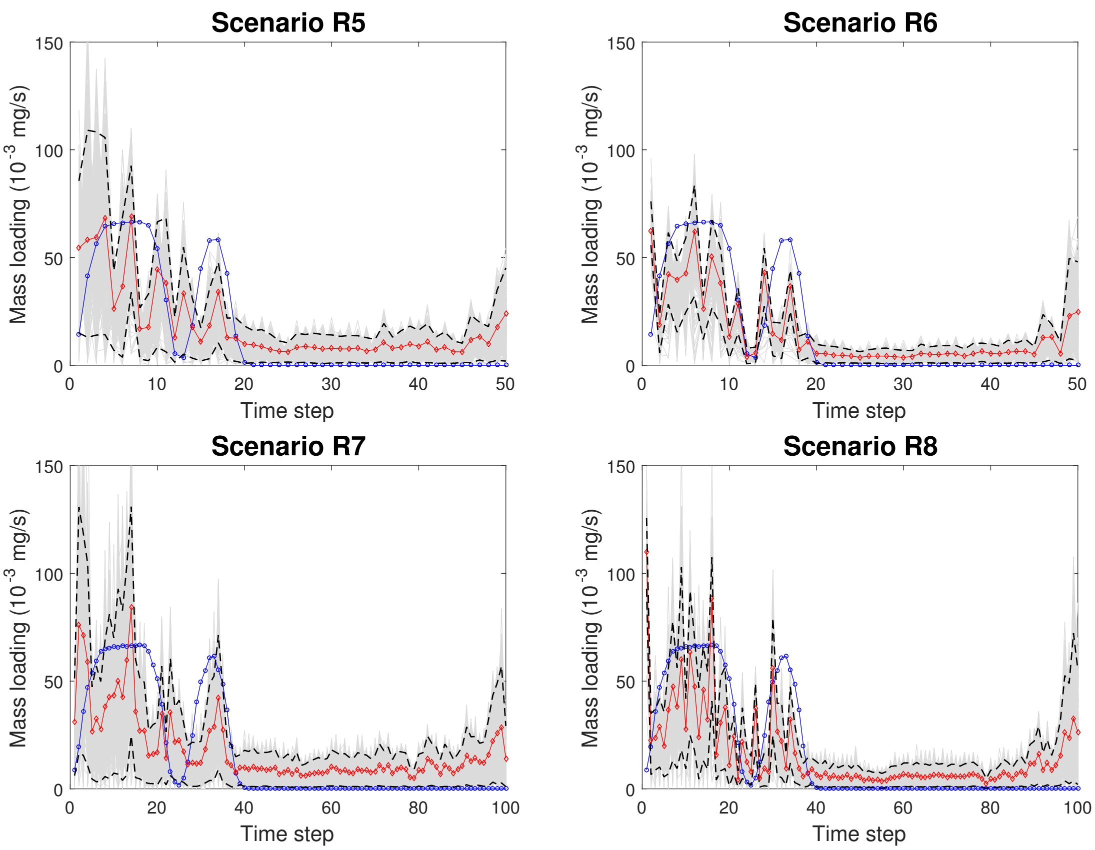

Figure 10: Recovered release history for the second sandbox experiment, scenarios R5 to R8. The blue curve corresponds to the actual release history. The gray lines are the recovered release history curves for all 500 realizations, the red dotted lines is the median, and black dashed lines the 5 and 95 percentiles.

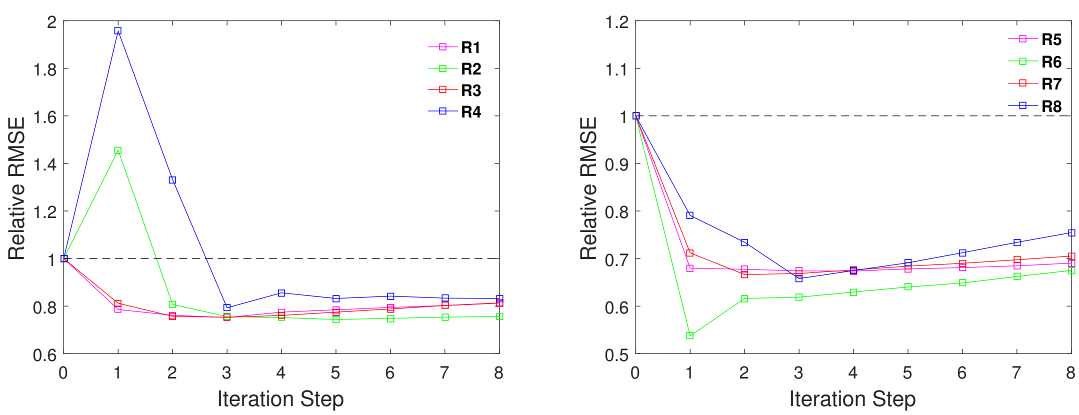

Figure 11: Relative RMSE of sandbox scenarios(R1-R8)

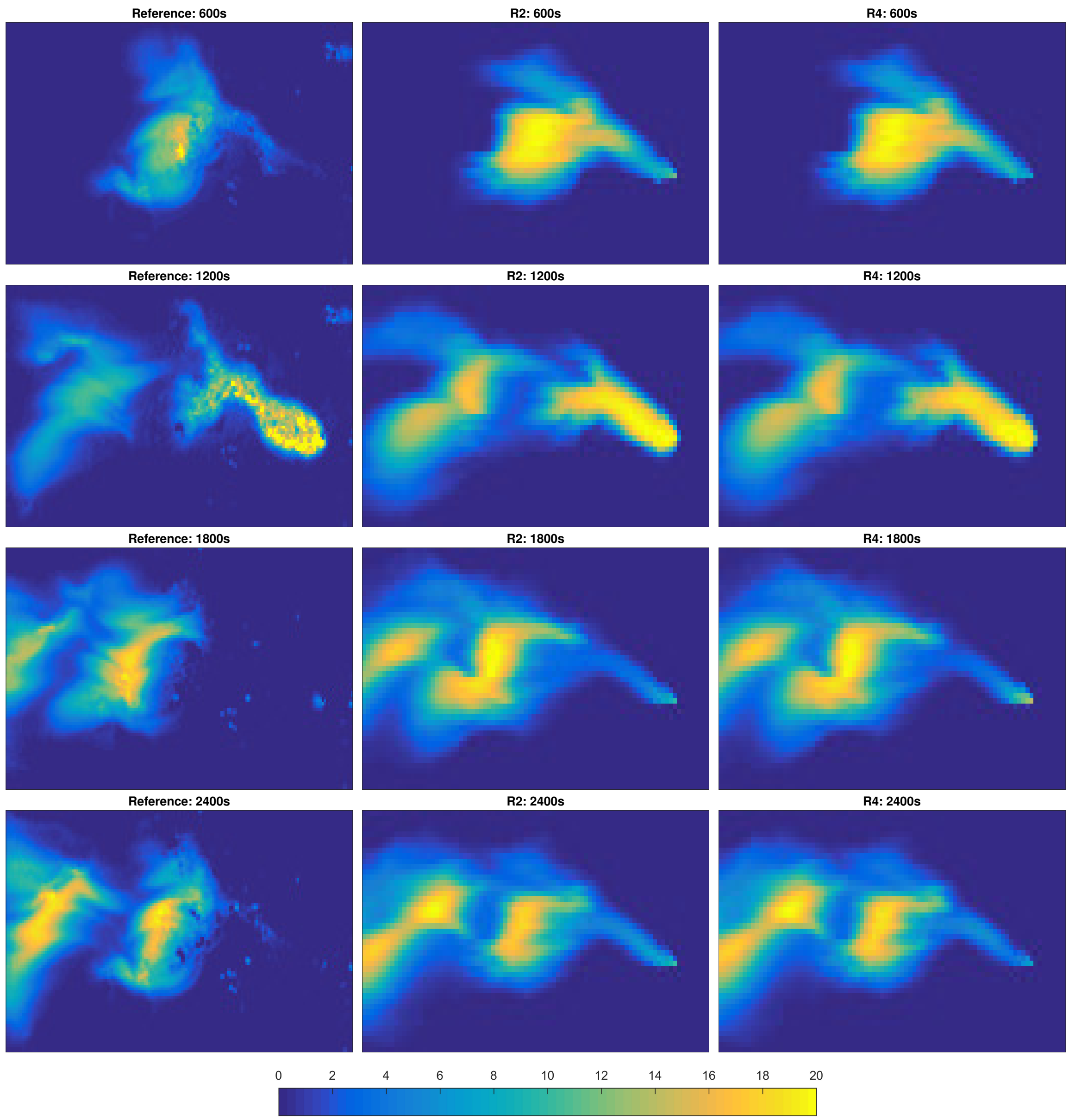

Figure 12: First sandbox experiment: train of pulses. Ensemble mean of the contaminant plume evolution obtained with the updated release functions for scenarios R2 and R4 at 600 s, 1200 s, 1800 s and 2400 s. The first column corresponds to the reference contaminant plume.

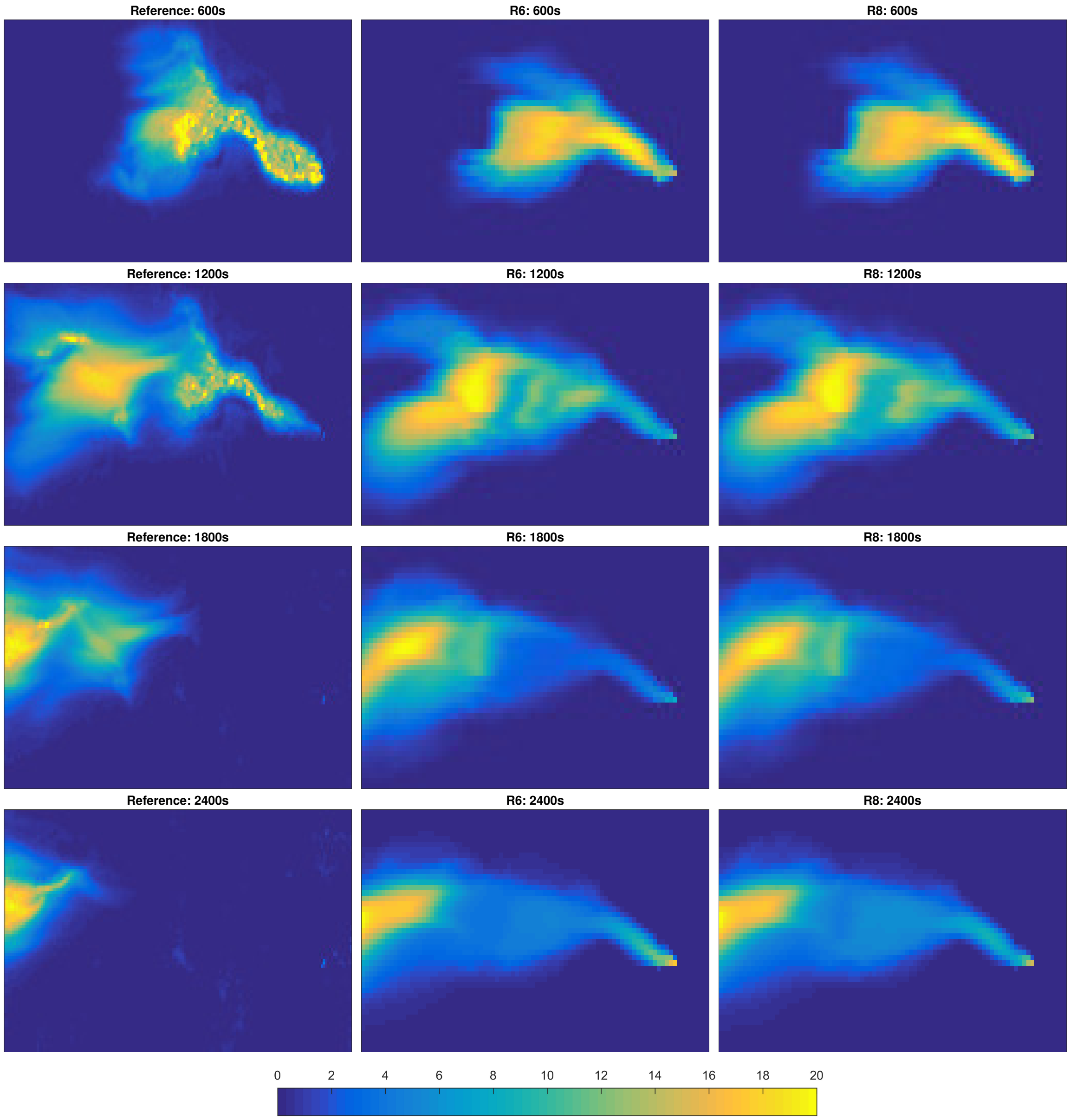

Figure 13: Second sandbox experiment: two pulses. Ensemble mean of the contaminant plume evolution obtained with the updated release functions for scenarios R2 and R4 at 600 s, 1200 s, 1800 s and 2400 s. The first column corresponds to the reference contaminant plume.

4. Summary and Conclusions

In this paper, we employ the ES-MDA to identify a time-varying release history with different precision in both synthetic and laboratory cases. In the synthetic cases, we examined the capacity of the ES-MDA to identify the release function for (i) different levels of time discretizations, with a ratio of 1 to 6 between the coarsest and finest discretizations; (ii) the impact of the observation data frequency (every other time step versus one step every ten time steps); (iii) the choice of the inflation factors (between Rafiee's and Evensen's proposals); and (iv) the impact of the number of iterations in the ES-MDA formulation (between four and eight). In total, 24 scenarios with combinations of the aforementioned features were generated and compared. The results show that the ES-MDA with Rafiee's scheme has a better performance in most scenarios in our case. Also, in all scenarios, increasing the observation data frequency always improves the identification of the recovered release history curve. The number of iterations, whether four or eight, does not have an important effect on the performance of the ES-MDA. In general, the ES-MDA performs well in recovering the release history, when the discretization is equal to 50 or 100 time steps but displays large fluctuations in the scenarios with 300 time steps. We believe this problem could be alleviated by choosing a different parameterization of the release curve, rather than using uncorrelated uniform random numbers to generate the initial ensemble of realizations.

Then, we apply the ES-MDA (using Rafiee's inflation scheme and eight iterations) to two sandbox experiments using different release history curves, a train of four pulses, and two pulses during the first half of the experiment. The results show that the ES-MDA works well for the train of pulses, but overestimates the injection concentrations for the second experiment after the two pulses have ended. We believe that this poor behavior could be explained again by the parameterization of the injection curves and the magnitude of the concentration observation errors.

In conclusion, the ES-MDA is a method capable to identify a time-varying release history in both synthetic and real cases. Better results than the ones presented here could have been obtained with a more elaborated parameterization of the time functions to be identified.

5. Acknowledgments

Financial support to carry out this work was received from grants PID2019-109131RB-I00 and PRX17/00150 funded by MCIN/AEI/10.13039/501100011033. Teng Xu also acknowledges the financial support from the Fundamental Research Funds for the Central Universities (B200201015) and Jiangsu Specially-Appointed Professor Program (B19052). The authors would like to thank University of Parma for providing the experimental equipment. Part of the work was performed during a stay of the third author at the University of Parma under the TeachInParma initiative, co-funded by Fondazione Cariparma and University of Parma.

References

Aanonsen, S. I., Nævdal, G., Oliver, D. S., Reynolds, A. C., & Vallès, B. (2009). The Ensemble Kalman Filter in Reservoir Engineering–a Review. *SPE Journal*, *14*, 393–412. URL: `http://www.onepetro.org/doi/10.2118/117274-PA`. doi:`10.2118/117274-PA`.

Atmadja, J., & Bagtzoglou, A. C. (2001). State of the Art Report on Mathematical Methods for Groundwater Pollution Source Identification. *Environmental Forensics*, *2*, 205–214. URL: `http://www.sciencedirect.com/science/article/pii/S1527592201900552`. doi:`http://dx.doi.org/10.1006/enfo.2001.0055`.

Bagtzoglou, A. C., & Atmadja, J. (2005). Mathematical Methods for Hydrologic Inversion: The Case of Pollution Source Identification. *Water Pollution*, *5*, 65–96. URL: `http://www.springerlink.com/index/10.1007/b11442`. doi:`10.1007/b11442`.

Bao, J., Li, L., & Redoloza, F. (2020). Coupling ensemble smoother and deep learning with generative adversarial networks to deal with non-Gaussianity in flow and transport data assimilation. *Journal of Hydrology*, *590*, 125443. URL: https://doi.org/10.1016/j.jhydrol.2020.125443. doi:10.1016/j.jhydrol.2020.125443.

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

Bertino, L., Evensen, G., & Wackernagel, H. (2003). Sequential Data Assimilation Techniques in Oceanography. *International Statistical Review*, *71*, 223–241. URL: http://doi.wiley.com/10.1111/j.1751-5823.2003.tb00194.x. doi:10.1111/j.1751-5823.2003.tb00194.x.

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. doi:10.1007/s00477-012-0662-1.

Capilla, J. E., Gömez-Hernández, J. J., & Sahuquillo, A. (1998). Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric head data—3. application to the culebra formation at the waste isolation pilot plan (wipp), new mexico, usa. *Journal of Hydrology*, *207*, 254–269.

Carrera, J., & Neuman, S. P. (1986). Estimation of Aquifer Parameters Under Transient and Steady State Conditions: 1. Maximum Likelihood Method Incorporating Prior Information. *Water Resources Research*, *22*, 199–210. doi:10.1029/WR022i002p00199.

Chen, Y., & Zhang, D. (2006). Data assimilation for transient flow in geologic formations via ensemble Kalman filter. *Advances in Water Resources*, *29*, 1107–1122. doi:10.1016/j.advwatres.2005.09.007.

Chen, Z., Gómez-Hernández, J. J., Xu, T., & Zanini, A. (2018). Joint identification of contaminant source and aquifer geometry in a sandbox experiment with the restart ensemble Kalman filter. *Journal of Hydrology*, *564*, 1074–1084. doi:10.1016/j.jhydrol.2018.07.073.

Chen, Z., Xu, T., Gómez-Hernández, J. J., & Zanini, A. (2021). Contaminant Spill in a Sandbox with Non-Gaussian Conductivities: Simultaneous Identification by the Restart Normal-Score Ensemble Kalman Filter. *Mathematical Geosciences*, *53*, 1587–1615. URL: https://doi.org/10.1007/s11004-021-09928-y. doi:10.1007/s11004-021-09928-y.

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. URL: http://dx.doi.org/10.1016/j.jconhyd.2014.11.001. doi:10.1016/j.jconhyd.2014.11.001.

Cupola, F., Tanda, M. G., & Zanini, A. (2015). Laboratory sandbox validation of pollutant source location methods. *Stochastic Environmental Research and Risk Assessment*, *29*, 169–182. doi:10.1007/s00477-014-0869-4.

Emerick, A. A., & Reynolds, A. C. (2013). Ensemble smoother with multiple data assimilation. *Computers and Geosciences*, *55*, 3–15. URL: http://dx.doi.org/10.1016/j.cageo.2012.03.011. doi:10.1016/j.cageo.2012.03.011.

Evensen, G. (2003). The Ensemble Kalman Filter: Theoretical formulation and practical implementation. *Ocean Dynamics*, *53*, 343–367. doi:10.1007/s10236-003-0036-9.

Evensen, G. (2004). Sampling strategies and square root analysis schemes for the EnKF. *Ocean Dynamics*, *54*, 539–560. doi:10.1007/s10236-004-0099-2.

Evensen, G. (2018). Analysis of iterative ensemble smoothers for solving inverse problems.

Computational Geosciences, *22*, 885–908. URL: http://link.springer.com/10.1007/s10596-018-9731-y. doi:10.1007/s10596-018-9731-y.

Evensen, G., & van Leeuwen, P. J. (2000). An Ensemble Kalman Smoother for Nonlinear Dynamics. *Monthly Weather Review*, *128*, 1852–1867. doi:10.1175/1520-0493(2000)128<1852:aeksfn>2.0.co;2.

Feyen, L., Gómez-Hernández, J., Ribeiro Jr, P., Beven, K. J., & De Smedt, F. (2003a). A bayesian approach to stochastic capture zone delineation incorporating tracer arrival times, conductivity measurements, and hydraulic head observations. *Water resources research*, *39*.

Feyen, L., Ribeiro Jr, P., Gomez-Hernandez, J., Beven, K. J., & De Smedt, F. (2003b). Bayesian methodology for stochastic capture zone delineation incorporating transmissivity measurements and hydraulic head observations. *Journal of hydrology*, *271*, 156–170.

Franssen, H., & Gómez-Hernández, J. (2002). 3d inverse modelling of groundwater flow at a fractured site using a stochastic continuum model with multiple statistical populations. *Stochastic Environmental Research and Risk Assessment*, *16*, 155–174.

Franssen, H. J., & Kinzelbach, W. (2009). Ensemble Kalman filtering versus sequential self-calibration for inverse modelling of dynamic groundwater flow systems. *Journal of Hydrology*, *365*, 261–274. URL: http://dx.doi.org/10.1016/j.jhydrol.2008.11.033. doi:10.1016/j.jhydrol.2008.11.033.

Gómez-Hernández, J., Franssen, H.-J. H., & Sahuquillo, A. (2003). Stochastic conditional inverse modeling of subsurface mass transport: a brief review and the self-calibrating method. *Stochastic Environmental Research and Risk Assessment*, *17*, 319–328.

Gómez-Hernández, J. J., & Xu, T. (2021). Contaminant source identification in aquifers: A critical view. *Mathematical Geosciences*, (pp. 1–22).

Houtekamer, P. L., & Mitchell, H. L. (2001). A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation. URL: http://journals.ametsoc.org/doi/abs/10.1175/1520-0493{\%}282001{\%}29129{\%}3C0123{\%}3AASEKFF{\%}3E2.0.CO{\%}3B2. doi:10.1175/1520-0493(2001)129<0123:ASEKFF>2.0.CO;2. arXiv:0203058.

Huang, C., Hu, B. X., Li, X., & Ye, M. (2009). Using data assimilation method to calibrate a heterogeneous conductivity field and improve solute transport prediction with an unknown contamination source. *Stochastic Environmental Research and Risk Assessment*, *23*, 1155--1167. doi:10.1007/s00477-008-0289-4.

Kurtz, W., Hendricks Franssen, H.-J., Kaiser, H.-P., & Vereecken, H. (2014). Joint assimilation of piezometric heads and groundwater temperatures for improved modeling of river-aquifer interactions. *Water Resources Research*, *50*, 1665--1688. URL: http://doi.wiley.com/10.1002/2013WR014823. doi:10.1002/2013WR014823.

Lan, T., Shi, X., Jiang, B., Sun, Y., & Wu, J. (2018). Joint inversion of physical and geochemical parameters in groundwater models by sequential ensemble-based optimal design. *Stochastic Environmental Research and Risk Assessment*, *32*, 1919--1937. URL: https://doi.org/10.1007/s00477-018-1521-5. doi:10.1007/s00477-018-1521-5.

Le, D. H., Emerick, A. A., & Reynolds, A. C. (2016). An Adaptive Ensemble Smoother With Multiple Data Assimilation for Assisted History Matching. *SPE Journal*, *21*, 2195--2207. URL: https://doi.org/10.2118/173214-PA. doi:10.2118/173214-PA.

van Leeuwen, P. J., & Evensen, G. (1996). Data Assimilation and Inverse

Methods in Terms of a Probabilistic Formulation. *Monthly Weather Review*, *124*, 2898--2913. URL: http://journals.ametsoc.org/doi/abs/10.1175/1520-0493{\%}281996{\%}29124{\%}3C2898{\%}3ADAAIMI{\%}3E2.0.CO{\%}3B2. doi:10.1175/1520-0493(1996)124<2898:DAAIMI>2.0.CO;2.

Li, J., Lu, W., Wang, H., & Fan, Y. (2019). Identification of groundwater contamination sources using a statistical algorithm based on an improved Kalman filter and simulation optimization. *Hydrogeology Journal*, *27*, 2919--2931. URL: http://link.springer.com/10.1007/s10040-019-02030-y. doi:10.1007/s10040-019-02030-y.

Li, L., Zhou, H., & Gómez-Hernández, J. J. (2011). A comparative study of three-dimensional hydraulic conductivity upscaling at the macro-dispersion experiment (made) site, columbus air force base, mississippi (usa). *Journal of Hydrology*, *404*, 278--293.

Li, L., Zhou, H., Gómez-Hernández, J. J., & Hendricks Franssen, H. J. (2012). Jointly mapping hydraulic conductivity and porosity by assimilating concentration data via ensemble Kalman filter. *Journal of Hydrology*, *428-429*, 152--169. URL: http://dx.doi.org/10.1016/j.jhydrol.2012.01.037. doi:10.1016/j.jhydrol.2012.01.037.

Ma, R., Zheng, C., Zachara, J. M., & Tonkin, M. (2012). Utility of bromide and heat tracers for aquifer characterization affected by highly transient flow conditions. *Water Resources Research*, *48*.

McDonald, M. G., & Harbaugh, A. W. (1988). *A modular three-dimensional finite-difference ground-water flow model* volume 6. US Geological Survey Reston, VA.

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*. doi:10.1029/2004WR003214.

Rafiee, J., & Reynolds, A. C. (2017). Theoretical and efficient practical procedures for the generation of inflation factors for ES-MDA. *Inverse Problems*, *33*. doi:10.1088/1361-6420/aa8cb2.

Ranazzi, P. H., & Sampaio, M. A. (2019). Ensemble size investigation in adaptive ES-MDA reservoir history matching. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, *41*, 413. URL: http://link.springer.com/10.1007/s40430-019-1935-0. doi:10.1007/s40430-019-1935-0.

Skaggs, T. H., & Kabala, Z. J. (1994). Recovering the release history of a groundwater contaminant. *Water Resources Research*, *30*, 71--79. URL: http://doi.wiley.com/10.1029/93WR02656. doi:10.1029/93WR02656.

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. doi:10.1029/2005WR004312.

Todaro, V., D'Oria, M., Tanda, M. G., & Gómez-Hernández, J. J. (2019). Ensemble smoother with multiple data assimilation for reverse flow routing. *Computers & Geosciences*, . URL: https://linkinghub.elsevier.com/retrieve/pii/S0098300419301992. doi:10.1016/j.cageo.2019.06.002.

Todaro, V., D'Oria, M., Tanda, M. G., & Gómez-Hernández, J. J. (2021). Ensemble smoother with multiple data assimilation to simultaneously estimate the source location and the release history of a contaminant spill in an aquifer. *Journal of Hydrology*, *598*. doi:10.1016/j.jhydrol.2021.126215.

Wen, X.-H., Capilla, J. E., Deutsch, C., Gómez-Hernández, J., & Cullick, A. (1999). A program to create permeability fields that honor single-phase flow rate and pressure data. *Computers & Geosciences*, *25*, 217--230.

Xu, T., & Gómez-Hernández, J. J. (2016). Joint identification of contaminant source location, initial release time, and initial solute concentration in an aquifer via ensemble Kalman filtering. *Water Resources Research*, . doi:10.1002/2014WR016618.Received.

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. URL: https://doi.org/10.1016/j.advwatres.2017.12.011. doi:10.1016/j.advwatres.2017.12.011.

Xu, T., Gómez-Hernández, J. J., Chen, Z., & Lu, C. (2021). A comparison between ES-MDA and restart EnKF for the purpose of the simultaneous identification of a contaminant source and hydraulic conductivity. *Journal of Hydrology*, *595*, 125681. URL: https://doi.org/10.1016/j.jhydrol.2020.125681. doi:10.1016/j.jhydrol.2020.125681.

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. URL: http://dx.doi.org/10.1016/j.jconhyd.2016.01.006. doi:10.1016/j.jconhyd.2016.01.006.

Zheng, C. (2010). *MT3DMS v5. 3Supplemental users guide: Tuscaloosa, Ala., University of Alabama Department of Geological Sciences*. Technical Report Technical Report to the US Army Engineer Research and Development Center.

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

Zhou, H., Gómez-Hernández, J. J., & Li, L. (2014). Inverse methods in hydrogeology: Evolution and recent trends. *Advances in Water Resources*, *63*, 22--37. URL: http://dx.doi.org/10.1016/j.advwatres.2013.10.014. doi:10.1016/j.advwatres.2013.10.014.