UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Stochastic Identification of Pollutant Sources in Aquifers by the Ensemble Kalman Filter

PhD Thesis submitted by

Zi Chen

Advisors:

J. Jaime Gómez-Hernández

Teng Xu

October 2020

Stochastic Identification of Pollutant Sources in Aquifers by the Ensemble Kalman Filter

PhD Thesis submitted by
Zi Chen

Advisors:
J. Jaime Gómez-Hernández
Teng Xu

Department:
Ingeniería Hidráulica y Medio Ambiente

Universitat Politècnica de València

October 2020, Valencia, Spain

UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

grupo de HIDROGEOLOGIA

Abstract

As part of the data assimilation methods, the ensemble-based methods have gained popularity in hydrogeology given their ability to deal with huge amounts of observed data simultaneously. More recently, researchers have started to employ these method to deduce contamination source information in synthetic cases (Xu and Gómez-Hernández, 2016b; Xu and Jaime, 2018). Based on these previous work, we take a step further to evaluate their performance in sandbox experiments. The main objective of this thesis is to verify the capacity of the ensemble-based methods in identifying contaminant source problem and complex geological heterogeneity.

The thesis could be divided into four parts. In the first part, the restart ensemble Kalman filter (r-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. The results show that the r-EnKF is capable of identifying both contaminant source information and aquifer-geometry-related parameters.

The second part shows an application of the restart normal-score ensemble Kalman filter (NS-EnKF) with covariance inflation in a heterogenous condcutivity laboratory experiment. The method is first tested using a synthetic case that mimics the sandbox experiment to establish the minimum number of ensemble members and the best technique to prevent filter collapse. Then, its application to the sandbox data shows that the restart NS-EnKF can benefit from Bauser's inflation to reduce the ensemble size and to arrive to a good joint identification of both the contaminant source and the spatial heterogeneity of conductivities.

In the third part, the ensemble smoother with multiple data assimilation (ES-MDA) is employed for the simultaneous identification of a contaminant source and the spatial distribution of hydraulic conductivity while using the r-EnKF as a benchmark. The outcome shows that the ES-MDA is able to outperform the r-EnKF, marginally, for the specific synthetic case analyzed with almost the same CPU consumption, and it can perform far better than the r-EnKF just with a cost of larger CPU usage.

The forth and last part investigates the performance of the ES-MDA in a time-varying release history identification problem. The influence of different observation intervals and inflation factor schemes on the determination

of the release curve are discussed. The outcome shows that the ES-MDA performs great in recovering release history when the history curve is discretized in not too many steps, and that it fails when the discretization is large. The frequency at which observation data are sampled is an influential factor in this application, while the number of iterations or the inflation schemes have less effect.

Resumen

Como parte de los métodos de asimilación de datos, los métodos basados en conjuntos han ganado popularidad en hidrogeología dada su capacidad para manejar grandes cantidades de datos observados simultáneamente. Recientemente, se ha comenzado a emplear este método para la identificación de fuentes de contaminación en casos sintéticos (Xu and Gómez-Hernández, 2016b; Xu and Jaime, 2018). Basándonos en estos trabajos anteriores, hemos dado un paso adelante evaluando su rendimiento en experimentos de tanque de laboratorio.

La tesis se puede dividir en cuatro partes. En la primera parte, el filtro de Kalman de conjuntos con reinicio (r-EnKF) se utiliza para la identificación espacio-temporal de una fuente puntual de contaminantes en un experimento en tanque de laboratorio, junto con la identificación de la posición y longitud de una placa vertical insertada en el tanque que modifica la geometría del sistema. Los resultados muestran que el r-EnKF es capaz de identificar tanto la fuente como los parámetros relacionados con la geometría del acuífero.

La segunda parte muestra una aplicación del filtro de Kalman de conjuntos con anamorfosis normal y reinicio (NS-EnKF) y con inflación de la covarianza en un experimento de laboratorio con conductividad heterogénea. El método se prueba primero utilizando un caso sintético que imita el experimento del tanque para establecer el número mínimo de miembros del conjunto y la mejor técnica para evitar el colapso del filtro. Luego, su aplicación a los datos del tanque muestra que el NS-EnKF con reinicio puede beneficiarse de la inflación de Bauser para reducir el tamaño del conjunto y llegar a una buena identificación conjunta tanto de la fuente de contaminantes como de la heterogeneidad espacial de las conductividades.

En la tercera parte, el filtro de Kalman de conjuntos suavizado con asimilación múltiple de datos (ES-MDA) se emplea para la identificación simultánea de una fuente de contaminantes y la distribución espacial de la conductividad hidráulica utilizando el r-EnKF como punto de referencia. El resultado muestra que el ES-MDA puede superar al r-EnKF, marginalmente, para el caso sintético específico analizado con el mismo consumo de CPU, y puede funcionar mucho mejor que el r-EnKF a cambio de un mayor costo de CPU.

La cuarta y última parte investiga el rendimiento del ES-MDA en un problema de identificación de una inyección de contaminante que varía en

el tiempo. Se analiza la influencia de diferentes intervalos de observación y esquemas de inflación de la covarianza en la determinación de la curva de inyección. El resultado muestra que el ES-MDA funciona muy bien en la identificación de la curva de inyección cuando la discretización de la misma no es muy alta, pero encuentra problemas de fluctuación en los casos con discretizaciones altas. La frecuencia con la que se muestrean los datos de observación es un factor influyente, mientras que el número de iteraciones o los métodos de inflación de la covarianza tienen menos efecto.

Resum

Com a part dels mètodes d'assimilació de dades, els mètodes basats en conjunts han guanyat popularitat en hidrogeologia donada la seua capacitat per a manejar grans quantitats de dades observades simultàniament. Recentment, s'ha començat a emprar aquest mètode per a la identificació de fonts de contaminació en casos sintètics (Xu and Gómez-Hernández, 2016b; Xu and Jaime, 2018). Basant-nos en aquests treballs anteriors, hem fet un pas avant avaluant el seu rendiment en experiments de tanc de laboratori.

La tesi es pot dividir en quatre parts. En la primera part, el filtre de Kalman de conjunts amb reinici (r-EnKF) s'utilitza per a la identificació espaciotemporal d'una font puntual de contaminants en un experiment en tanc de laboratori, juntament amb la identificació de la posició i longitud d'una placa vertical inserida en el tanc que modifica la geometria del sistema. Els resultats mostren que el r-EnKF és capaç d'identificar tant la font com els paràmetres relacionats amb la geometria de l'aqüífer.

La segona part mostra una aplicació del filtre de Kalman de conjunts amb anamorfosis normal i reinici (NS-EnKF) i amb inflació de la covariància en un experiment de laboratori amb conductivitat heterogènia. El mètode es prova primer utilitzant un cas sintètic que imita l'experiment del tanc per a establir el nombre mínim de membres del conjunt i la millor tècnica per a evitar el col·lapse del filtre. Després, la seua aplicació a les dades del tanc mostra que el NS-EnKF amb reinici pot beneficiar-se de la inflació de Bauser per a reduir la grandària del conjunt i arribar a una bona identificació conjunta tant de la font de contaminants com de l'heterogeneïtat espacial de les conductivitats.

En la tercera part, el filtre de Kalman de conjunts suavitzat amb assimilació múltiple de dades (ES-MDA) s'empra per a la identificació simultània d'una font de contaminants i la distribució espacial de la conductivitat hidràulica utilitzant el r-EnKF com a punt de referència. El resultat mostra que l'ES-MDA pot superar al r-EnKF, marginalment, per al cas sintètic específic analitzat amb el mateix consum de CPU, i pot funcionar molt millor que el r-EnKF a canvi d'un major cost de CPU.

La quarta i última part investiga el rendiment de l'ES-MDA en un problema d'identificació d'una injecció de contaminant que varia en el temps. S'analitza la influència de diferents intervals d'observació i esquemes de inflació de la covariància en la determinació de la corba d'injecció. El resul-

tat mostra que l'ES-MDA funciona molt bé en la identificació de la corba d'injecció quan la discretització no és massa alta, però troba problemes de fluctuació amb discretitzacions massa fines. La freqüència amb la qual es mostregen les dades d'observació és un factor influent en aquesta aplicació, mentre que el nombre d'iteracions o els mètodes d'inflació de la covariància tenen menys efecte.

Acknowledgements

This thesis would not come into being if I had not received the help and support from my supervisors, colleagues and family. Their comments and encouragements contributed a lot to the accomplishment of my PhD study.

First and foremost, I would like to extend my sincerest gratitudes to my supervisor Professor J. Jaime Gómez-Hernández for his insightful guidance, valuable suggestions and constant care both in my study and life. No matter what kind of problem I encountered, he is always willing to help. He has walked me through all my works and without his consistent and illuminating instruction, this thesis could not have reached its present form.

Also, my heartiest thanks flow to the institutions that financed my studies. The support to carry out my 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 my supervisor, reference PRX17/00150.

Then, particular thanks go to my co-supervisor and friend, Teng Xu. When I arrived in Valencia four years ago, it was him who showed me how live is in this foreign country and helped me get used to it. When I faced complex geostatistical theory, it was also he who provided me the necessary materials. What's more, his classic Chinese cooking was also a special bridge for all the overseas Chinese students and made us not to feel homesick anymore.

Besides, I'm also very grateful to all colleagues in the IIAMA group. We together worked in the Geostats2016 and Interpore2019 conference, discussed about the Ensemble methods, enjoyed Jaime's homemade pallea, entertained ourselves in Vanessa's house. Everyone of you made my life in Spain gorgeous and I'll cherish these moments for my whole life.

Last but not least, my thanks would go to my beloved parents. Their unfailing love and support are the source of my strength. I'll never be this far without them.

Contents

List of Figures

List of Tables

1 Introduction

1.1 Motivation and Objectives

In groundwater contamination issues, source information is generally difficult to obtain during the processes of environmental risk assessment, response accountability and further restoration. In general, the first sign of a contamination is normally some concentration measurements downgradient from the source while the source remains unknown. Attempting to identify the source from these limited downgradient observation data becomes a difficult inverse problem. Moreover in strong heterogeneus fields, where the problem of source identification is known to be ill-posed (Skaggs and Kabala, 1994; Carrera and Neuman, 1986).

Dozens of works have focused on this topic and many methods have been developed (Atmadja and Bagtzoglou, 2001b; Michalak and Kitanidis, 2004; Bagtzoglou and Atmadja, 2005; Sun et al., 2006a, e.g.,). They can be grouped into two main categories: optimization approaches and probabilistic approaches. The optimization approaches cast the problem as a deterministic one in which parameters are found that minimize a given objective function, such as least-squares regression, maximum likelihood, hybrid heuristic approach (e.g., Gorelick et al., 1983; Wagner, 1992; Aral et al., 2001; Yeh et al., 2007; Mirghani et al., 2009; Amirabdollahian and Datta, 2014; Ayvaz, 2016); the probabilistic approaches cast the problem in a stochastic framework and the parameters to estimate become random variables, such as minimum relative entropy method, adjoint state method (e.g., Bagtzoglou et al., 1992; Woodbury and Ulrych, 1996; Neupauer and Wilson, 1999; Butera et al., 2013; Koch and Nowak, 2016).

Among these researches, only a few works applied their methods to real cases (e.g., Woodbury et al., 1998; Michalak and Kitanidis, 2004; Cupola et al., 2015a; Zanini and Woodbury, 2016). The main reason why is the inherent heterogeneity of aquifer properties (e.g., Gómez-Hernández and Wen, 1998; Knudby and Carrera, 2005; Zinn and Harvey, 2003). Finding a reliable method to identify the contaminant source and aquifer properties jointly is a challenging task.

In the last decades, data assimilation methods are routinely used for identification purposes because of their ability to deal with various kinds of observed data simultaneously. The ensemble Kalman filter (EnKF) is the data assimilation method most used nowadays. This efficient method was first proposed by Evensen (2003) and has gained popularity in hydrogeology (Chen and Zhang, 2006; Huang et al., 2009; Kurtz et al., 2014; Li et al., 2012a; Franssen and Kinzelbach, 2009). More recently, researchers started to employ EnKF variants to deduce contamination source information in groundwater aquifer. Xu and Gómez-Hernández (2016b) 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 and Jaime, 2018). In the meanwhile, the ensemble smoother (ES), which was firstly introduced by van Leeuwen and Evensen (1996), and later developed into a variant named ensemble smoother with Multiple Data Assimilation (ES-MDA) Emerick and Reynolds (2013a), has proven its ability in dealing with history-matching problems. Based on these previous works, we take a step further to evaluate their performance in sandbox experiments for the purpose of contaminant source identification.

The sandbox experiments are carried out in the hydraulic laboratory of the Department of Engineering and Architecture of the University of Parma. The whole equipment is built with Polymethyl methacrylate (PMMA) plates of dimensions 120 cm $\times$ 14 cm $\times$ 70 cm. Inside the box, two reservoirs at the edges allow us to define prescribed constant boundary conditions. And by using different size of glass beads and a vertical plastic plate, we can vary the heterogeneity of the sanbox and its geometry. One moveable injector was installed at the upstream part and able to release tracer as desired. With the help of this device, we then were able to conduct several laboratory experiments to evaluate and analyse the application of the EnKF and the ES-MDA for contaminant source identification.

The main objectives of this thesis can be summarized as follows: first, to verify the capacity of the EnKF to identify a contaminant source together with geological heterogeneity, such as the presence of a vertical impermeable barrier or the heterogeneous spatial variability of the glass beads; second, to contrast the new ES-MDA method with the EnKF for the solution of the same problem; and third, to identifiy complex release histories using the ES-MDA.

1.2 Thesis Organization

The chapters have been written so that they could be published as independent papers. For this reason, there is some repetition. When this happens, it is indicated in the text so that the reader that has read the entire thesis does not have to go through it again, but the reader interested only in a specific chapter does not have to go through the whole thesis to find some of the material already presented. The next four chapters are independent papers which are published or under review or to be submitted in refereed international journals. Chapter 2 and Chapter 4 are published in Journal of Hydrology, Chapter 3 has been submitted to Mathematical Geosciences and under revision. And the last chapter is the summary of the whole work.

Chapter 2 presents an application of the restart ensemble Kalman filter in a contaminant source identification problem. The work 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.

Chapter 3 illustrates the performance of the restart normal-score ensemble Kalman filter (Ns-EnKF) for the joint identification of a contaminant source and a heterogeneous hydraulic conductivity distribution in a laboratory sandbox experiment.

Chapter 4 shows the capacity of the ensemble smoother with multiple data assimilation (ES-MDA) for the simultaneous identification of a contaminant source and the spatial distribution of hydraulic conductivity by assimilating both piezometric head and concentration observations in a synthetic aquifer while using the restart ensemble Kalman filter as a benchmark.

Chapter 5 applies the ES-MDA method to identify a time-varying contaminant injection. The impacts of different inflation schemes, number of iterations, and observation intervals are evaluated.

Chapter 6 summarizes all the works in this thesis and shows some suggestions for future research.

2

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

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.

2.1 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, 2001b; Michalak and Kitanidis, 2004; Bagtzoglou and Atmadja, 2005; Sun et al., 2006a, 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 differ-

ent 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; Xu and Gómez-Hernández, 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 Jaime (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.

2.2 Methodology

2.2.1 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:

$$S_s \frac{\partial h}{\partial t} = \nabla \cdot (K \nabla h) + w, \tag{2.1}$$

$$\frac{\partial (\theta C)}{\partial t} = \nabla \cdot (\theta D \cdot \nabla C) - \nabla \cdot (\theta v C) - q_s C_s \tag{2.2}$$

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 $[T^{-1}]$. In the transport governing equation, 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 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).

2.2.2 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; Xu and Jaime, 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, and they do not have to necessarily contain the "real" value, they are simply used to initialize the algorithm. We build a matrix S with the eight parameters plus N_m concentration data for N_e realizations:

$$S = \begin{pmatrix} (Xs, Zs, Xb, Zb, Ic, Ir, Ts, Te)^T \\ (C_1, C_2, \ldots, C_{N_m})^T \end{pmatrix} \tag{2.3}$$

   where N_m is the number of model nodes 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 the augment matrix is given by

$$S_t^f = \psi\left[C_0, A_{t-1}^a\right], \tag{2.4}$$

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; S_t^f is an $(8+N_m)\times N_e$ matrix containing the updated parameters and forecasted concentrations for all realization; A_{t-1}^a is the matrix 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

$$A_t^f = A_{t-1}^a. \tag{2.5}$$

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

$$\mathbf{P}_t^f = \frac{1}{N_e-1}\sum_{i=1}^{N_e}\{[S_{i,t}^f - \overline{S_t^f}][S_{i,t}^f - \overline{S_t^f}]^T\} \tag{2.6}$$

with

$$\overline{S_t^f} = \frac{1}{N_e-1}\sum_{i=1}^{N_e} S_{i,t}^f \tag{2.7}$$

where $\mathbf{P}_t^f$ is an $(8+N_m)\times(8+N_m)$ matrix of augment parameter covariances and $\overline{S_t^f}$ is an $(8+N_m)\times 1$ column vector of parameter averages.

Then, compute the Kalman gain matrix,

$$\mathbf{K}_t = \mathbf{P}_t^f\mathbf{H}^T[\mathbf{H}\mathbf{P}_t^f\mathbf{H}^T + \mathbf{R}_t]^{-1} \tag{2.8}$$

where $\mathbf{H}$ is the observation matrix that extracts out of the whole augmented state vector the elements at which N_o observations where taken, $\mathbf{R}_t$ is an $N_o\times N_o$ observation error covariance matrix, which will

be assumed to follow a Gaussian distribution of zero mean, given variance, and no correlation between observations, and proceed to update the parameter values, realization by realization by

$$S_t^a = S_t^f + \mathbf{K}_t \left[y_t^{obs} + \varepsilon_i - \mathbf{H} S_t^f \right], \tag{2.9}$$

where y_t^{obs} is an $N_o \times 1$ vector of observed concentrations at time step t, ε_i stands for an observation error with zero mean and covariance $\mathbf{R}_t$.

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

2.3 Experimental Case

2.3.1 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 Figure 2.1. 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 cm 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 Figure 2.1. 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 Jaime (2018)).

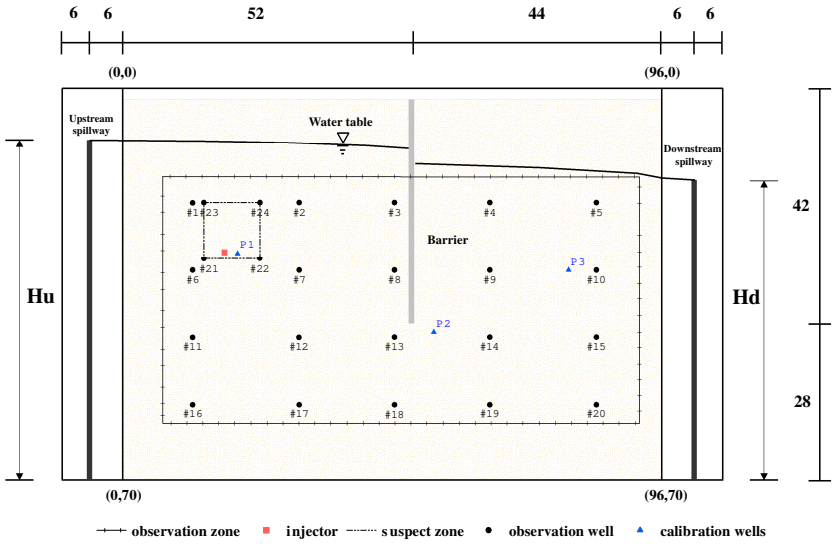

Figure 2.1. Sketch of the experimental device with indication of the upstream (Hu) and downstream (Hd) constant head boundaries. The ticked rectangle corresponds to the area captured by the camera in which concentrations will be monitored. Red dot is the release location. Dashed line around red dot indicates the release suspect location. Dimensions are in cm. Coordinates of the four corners of the flow and transport models are also shown.

2.3.2 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 Figure 2.1, 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, 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.

Table 2.1. 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 2.2. 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) - x-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	$960 - 1140$

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 2.1.

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 2.2.

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.

2.4.1 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 and the sensitivity of the observation wells near to the contaminant source: 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 Figure 2.1). In both scenarios, model error is neglected and we assume that observation errors are uncorrelated and follow a Gaussian distribution with mean zero, and standard deviation of 0.1 mg/l.

Figure 2.2 and 2.3 show the time evolution of the ensemble mean and the ensemble variance, respectively, of the updated state parameters for the two scenarios. Figure 2.4 and Figure 2.5 show 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. 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

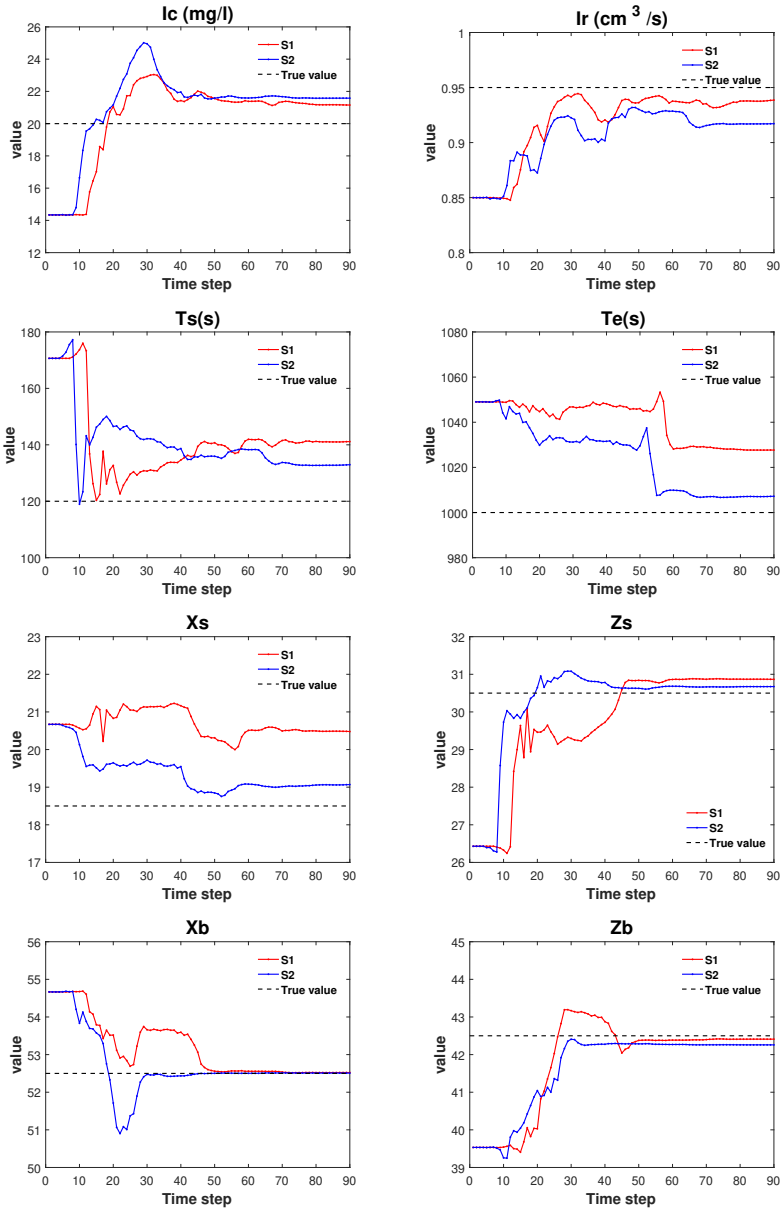

Figure 2.2. Time evolution of the ensemble mean of the 8 updated parameters, including contaminant source location (Xs, Zs), plate position (Xb, Zb), injection information (Ic, Ir) and release time interval (Ts, Te) in scenarios $S1$ and $S2$.

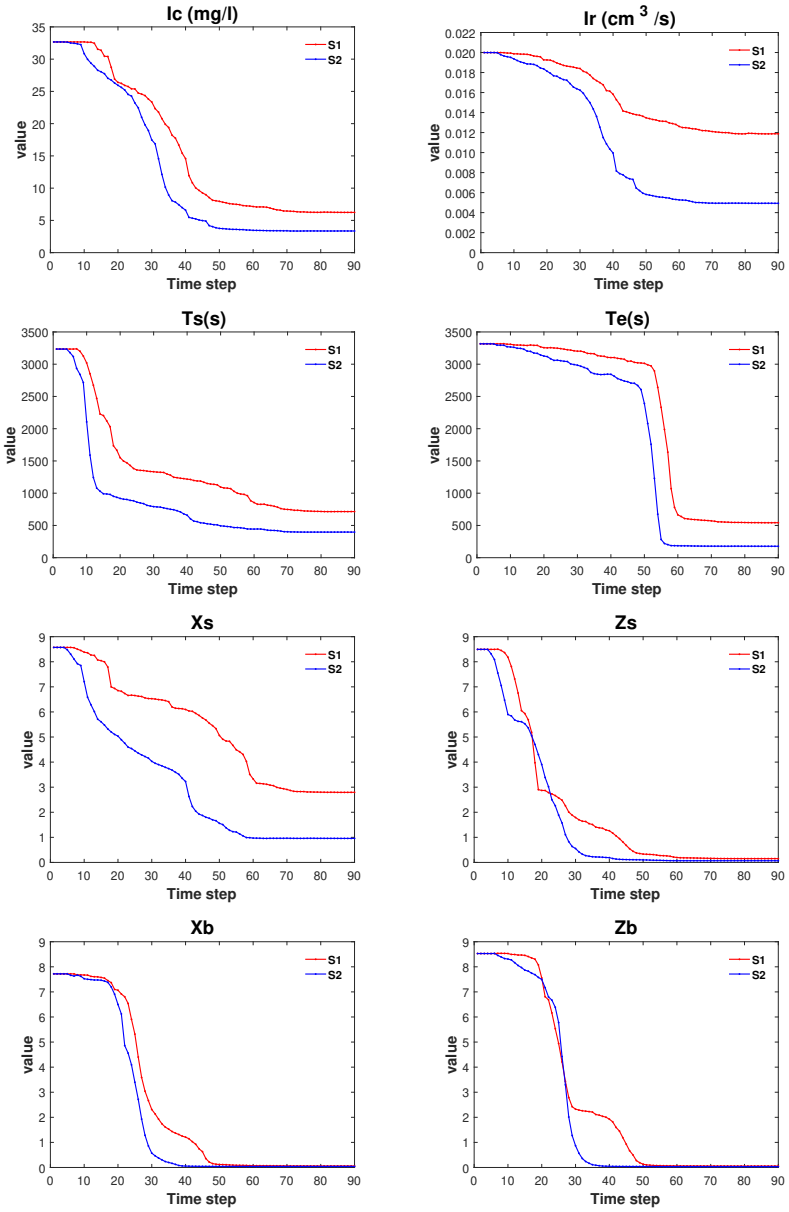

Figure 2.3. Time evolution of the ensemble variance for the same parameters and scenarios as in the previous figure.

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.

2.4.2 Laboratory Sandbox Test

The sandbox experiment was carried out as described previously. Figure 2.6 shows a picture of the fluorescein plume at the 48th time step (840s 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 concentration points 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 2.7 shows a comparison between observed and numerically predicted concentrations at five observation locations (wells #7, #7, #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 $R3$. Model error is introduced through an uncertainty in the hydraulic conductivity value, which is considered homogeneous in each realization and drawn from a Gaussian distribution with a mean of 0.58 cm/s and a standard deviation of 0.05 cm/s. The Gaussian description of these uncertainty instead of a certain value is an effective way to mimic the fluorescein calibration system and the hydraulic conductivity field, in the meanwhile, r-EnKF is optimal for the case where the parameters and state variables are multiGaussian (Aanonsen et al., 2009), and the relationship between the state variables and the observation/model error is linear. Therefore, to keep the Gaussinity, it is necessary to keep observation/model error Gaussian distributed.

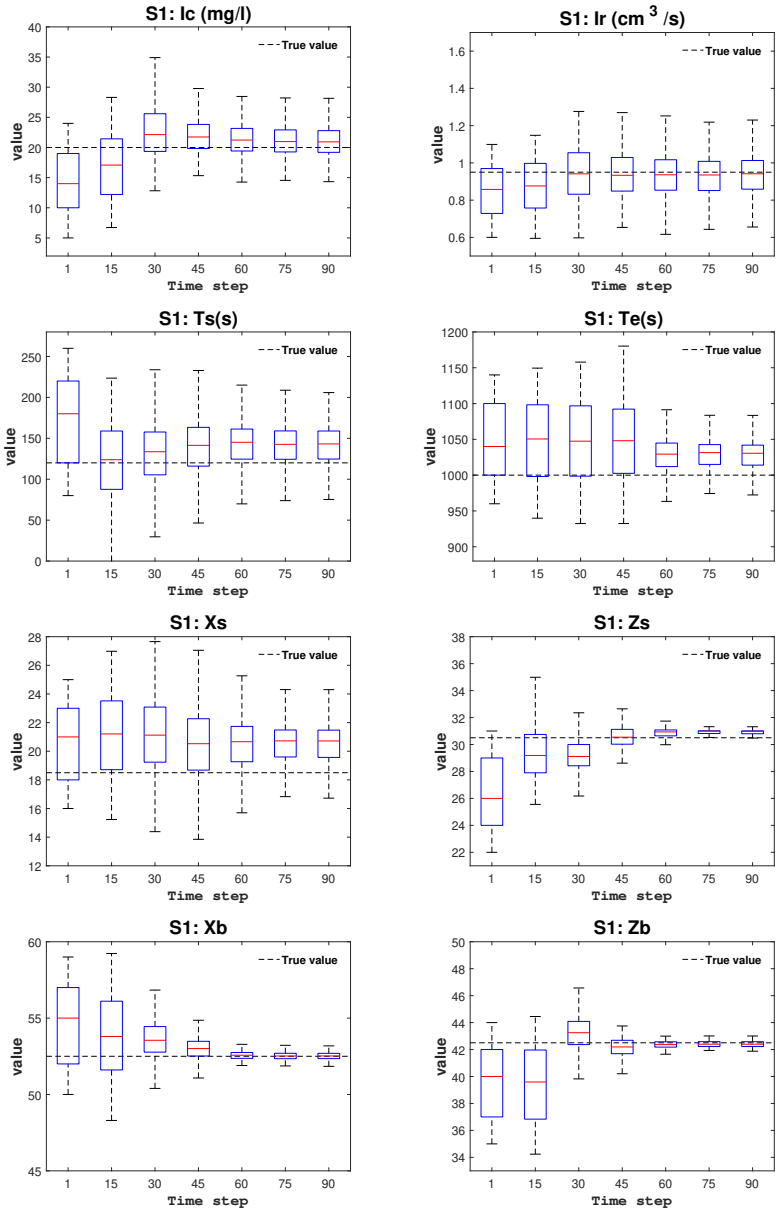

Figure 2.4. Boxplot of the 8 updated parameters at different time steps (1, 15, 30, 45, 60, 75, 90) for scenario $S1$.

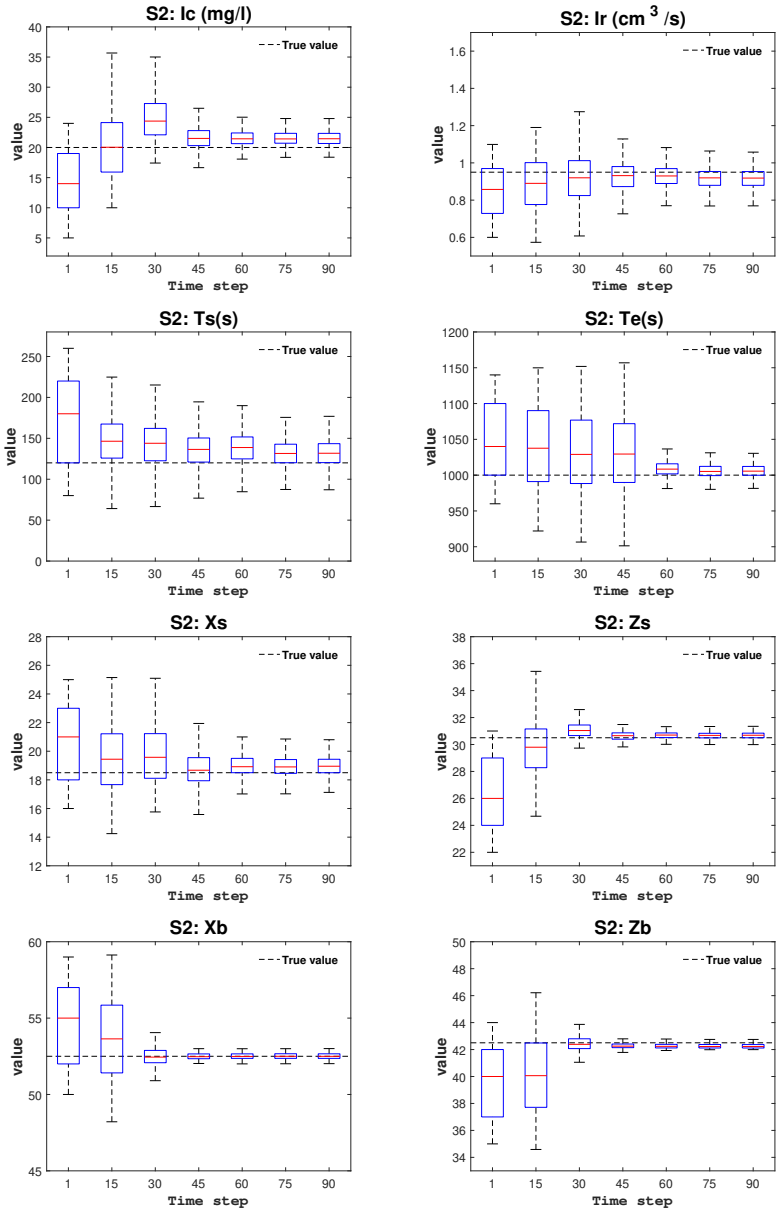

Figure 2.5. Boxplot of the 8 updated parameters at different time steps (1, 15, 30, 45, 60, 75, 90) for scenario $S2$.

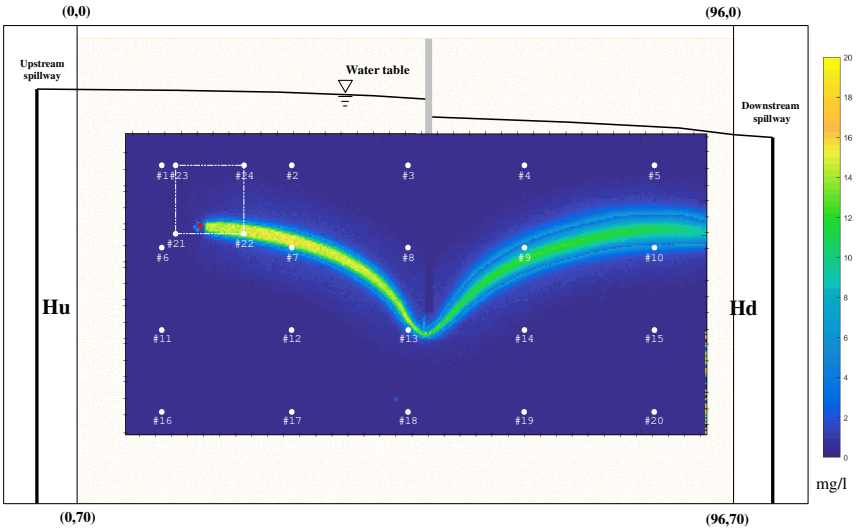

Figure 2.6. Fluorescein concentration field in the sandbox at the $48th$ time step. The area shown corresponds to the observation zone indicated in Figure 2.1. The dash line shows the suspect zone for the injection and the white dots indicate the observation wells.

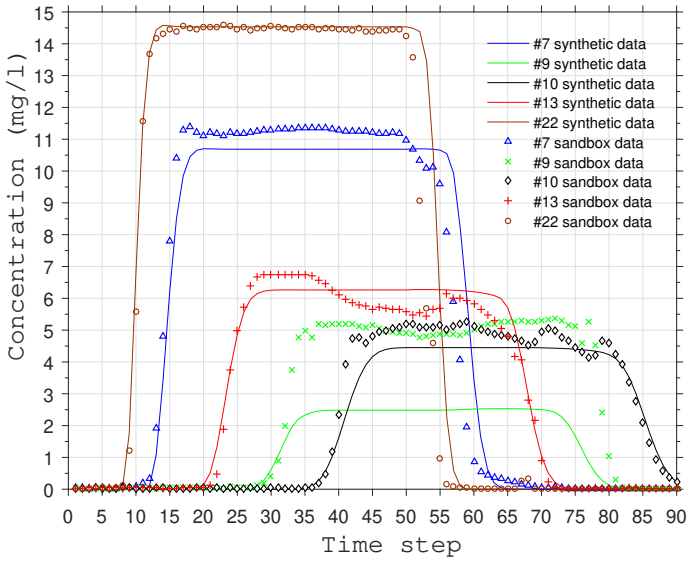

Figure 2.7. Fluorescein observed breakthrough curves at the observation wells located inside the plume and the curves computed from the numerical model.

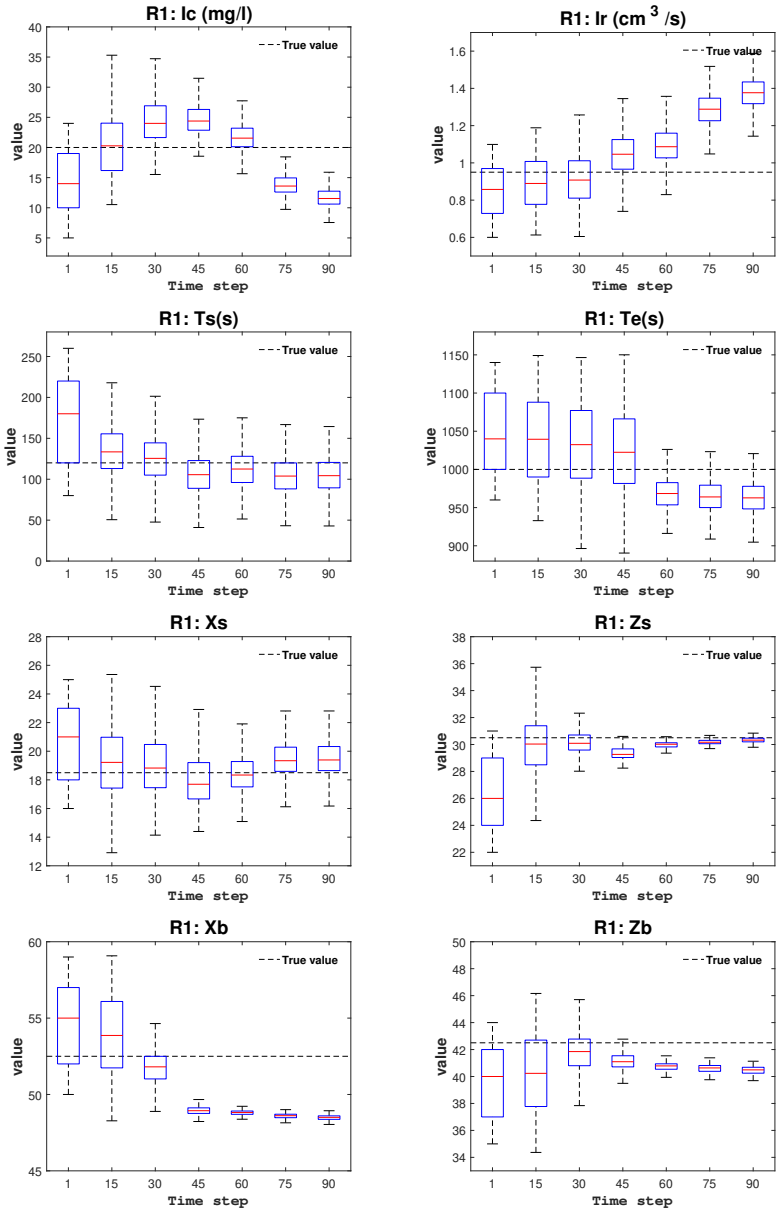

Figure 2.8. Boxplot of of the 8 updated parameters at time steps 1, 15, 30, 45, 60, 75 and 90 for scenario $R1$.

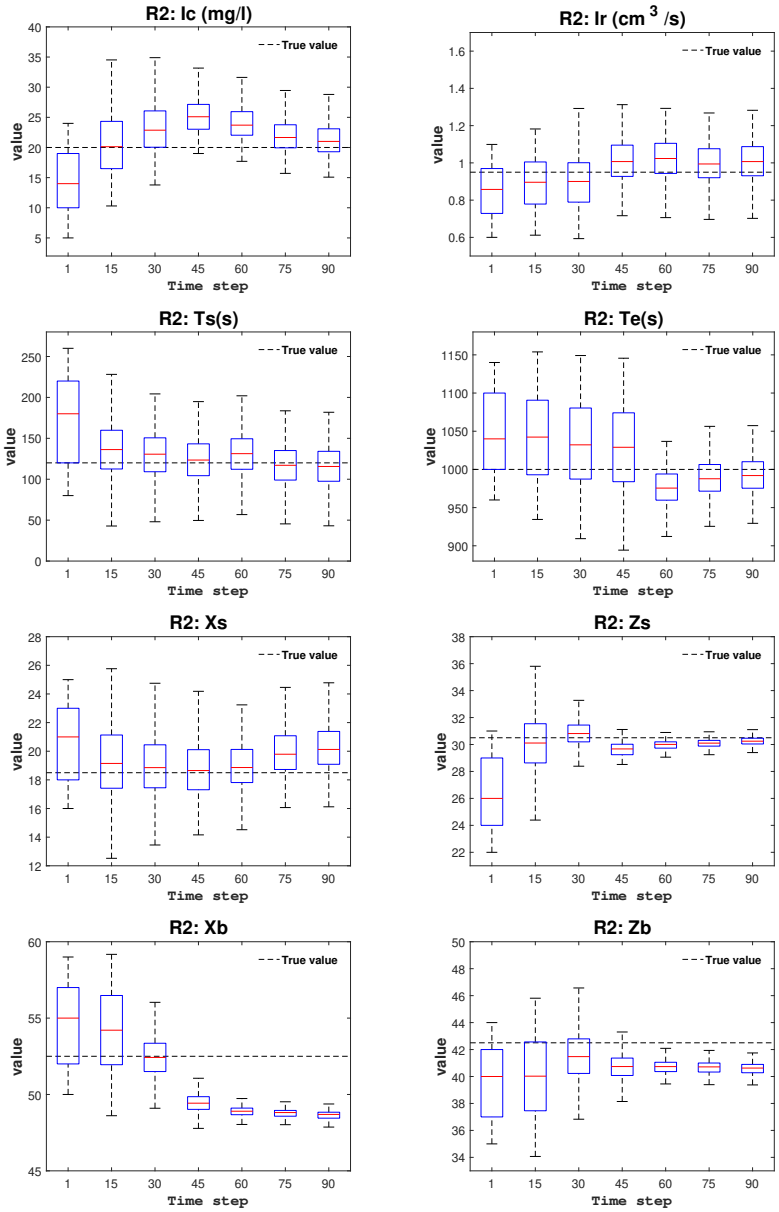

Figure 2.9. Boxplot of of the 8 updated parameters at time steps 1, 15, 30, 45, 60, 75 and 90 for scenario $R2$.

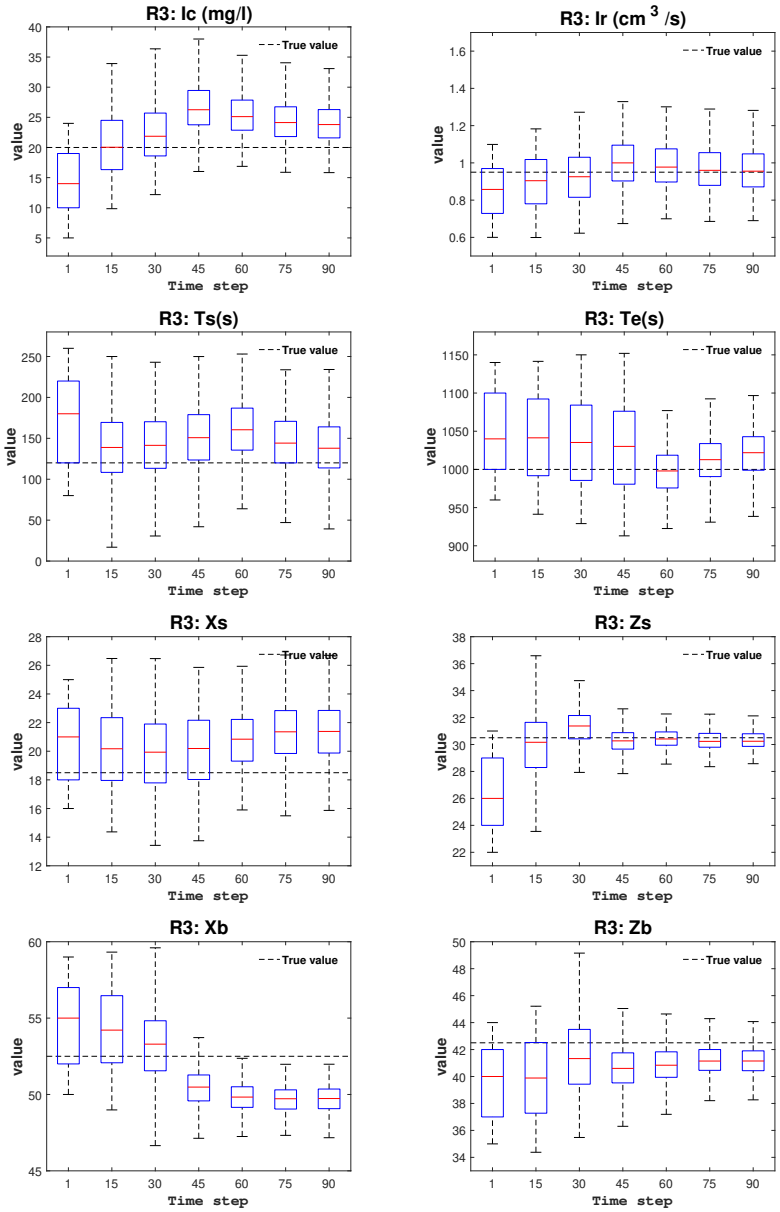

Figure 2.10. Boxplot of of the 8 updated parameters at time steps 1, 15, 30, 45, 60, 75 and 90 for scenarios $R3$.

Figure 2.11, Figure 2.12 and Figure 2.13 show 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 2.14. 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

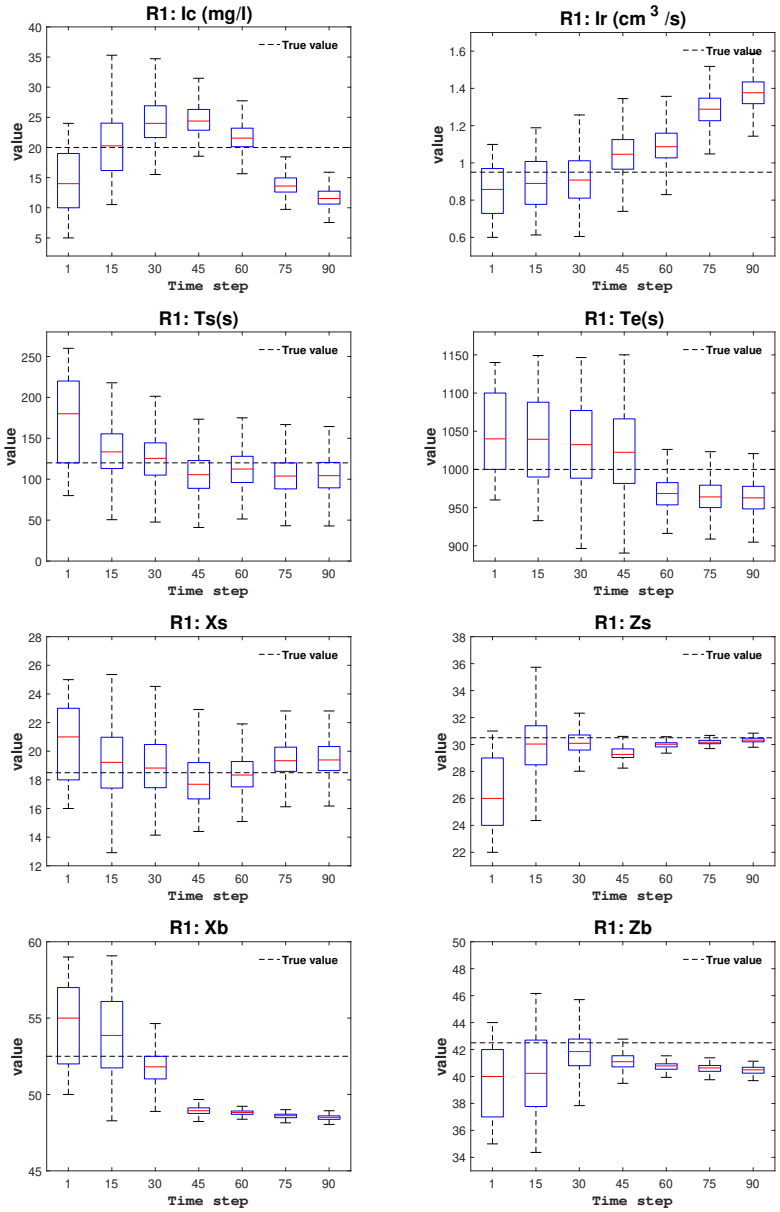

Figure 2.11. Boxplot of of the 8 updated parameters at time steps 1, 15, 30, 45, 60, 75 and 90 for scenario $R1$.

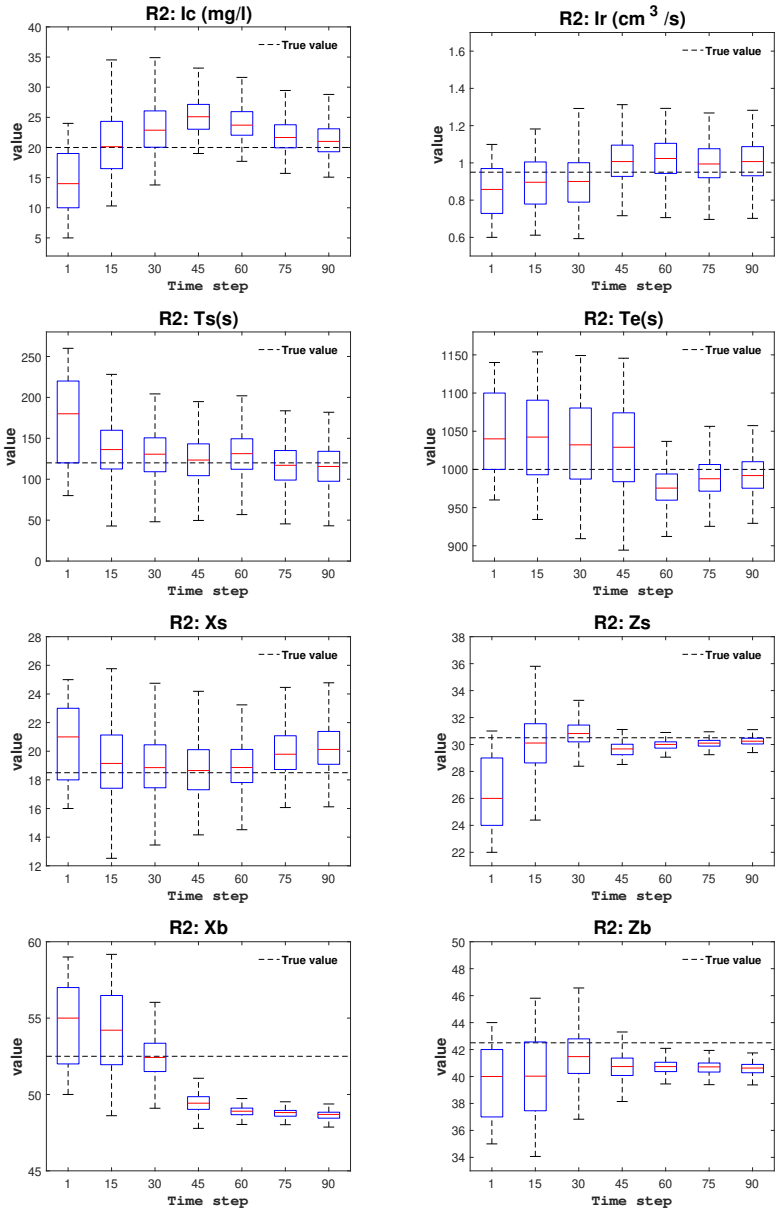

Figure 2.12. Boxplot of of the 8 updated parameters at time steps 1, 15, 30, 45, 60, 75 and 90 for scenario $R2$.

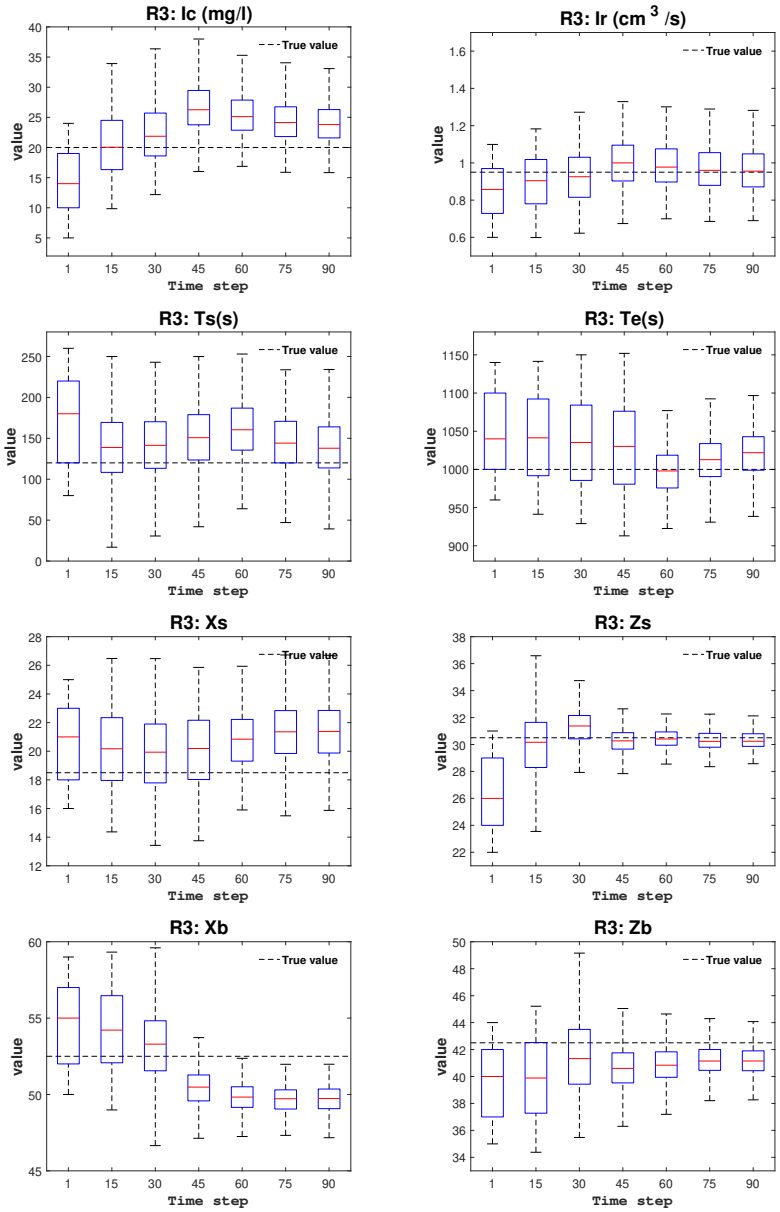

Figure 2.13. Boxplot of of the 8 updated parameters at time steps 1, 15, 30, 45, 60, 75 and 90 for scenarios $R3$.

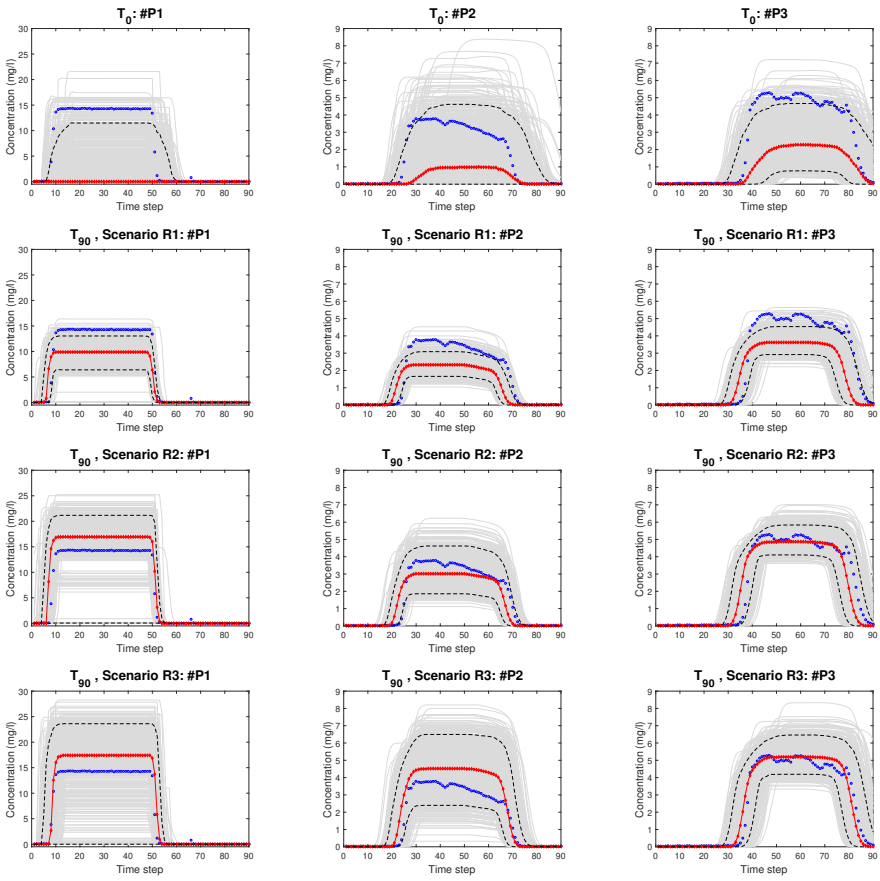

Figure 2.14. Breakthrough curves at control wells. The blue dots correspond to the curves in the sandbox experiment. The thin gray lines are the curves for all 800 realizations; they are summarized by their median (red diamond lines) and their 5 and 95 percentiles (black dash lines).

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. Actually, this conclusion fits the description of calibration of the same device in another experiment very well, in which the maximum measurement error was estimated as less than 3 mg/l and the standard deviation is around 1 mg/l (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 2.15. When comparing this figure to Figure 2.12 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, since the vertical plate was inserted into the sandbox afterwards, it could not match the sandbox without any gap in a practical way, not being completely impermeable as implemented in the numerical model. 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. The results for the new scenario, referred to as $R2c$ are shown in Figure 2.16. (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. The result proves the thought that the gap will influence the performance of the r-EnKF and it's better to take these model error into account in a real case.

These results show 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.

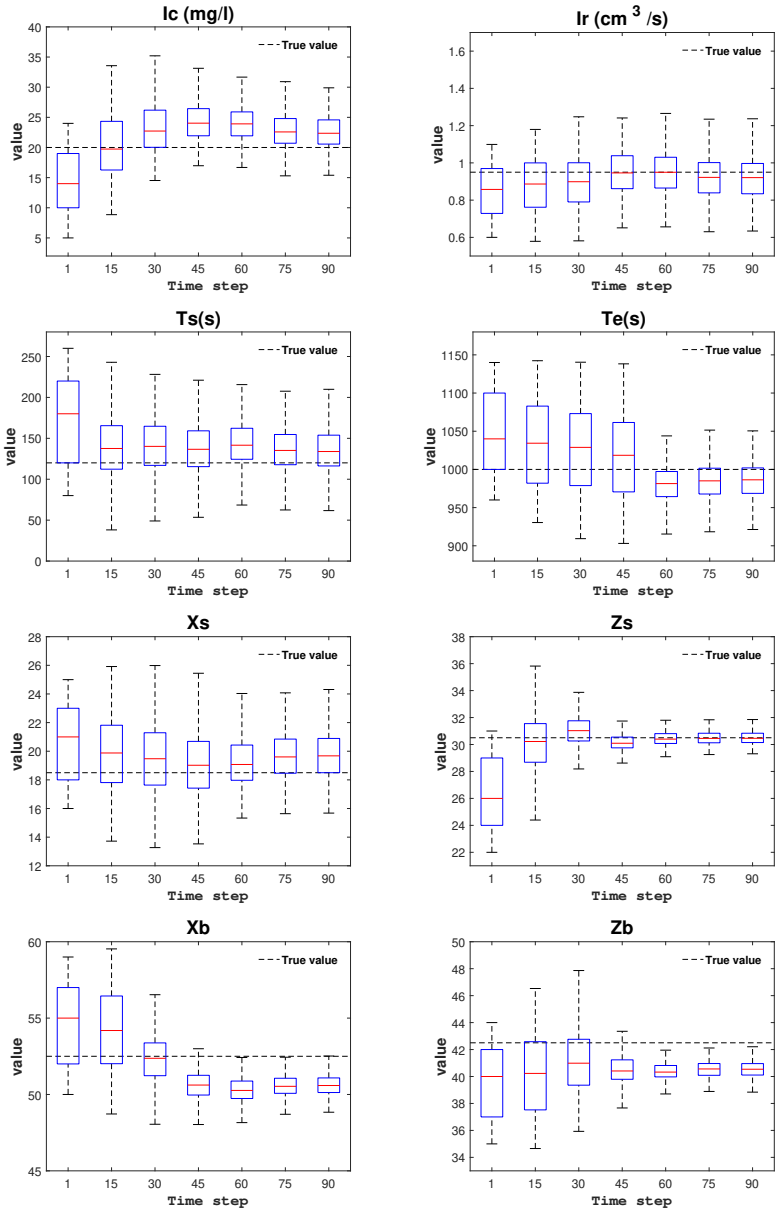

Figure 2.15. Boxplot of the 8 updated parameters in scenario *R2b* at different time steps (1, 15, 30, 45, 60, 75, 90).

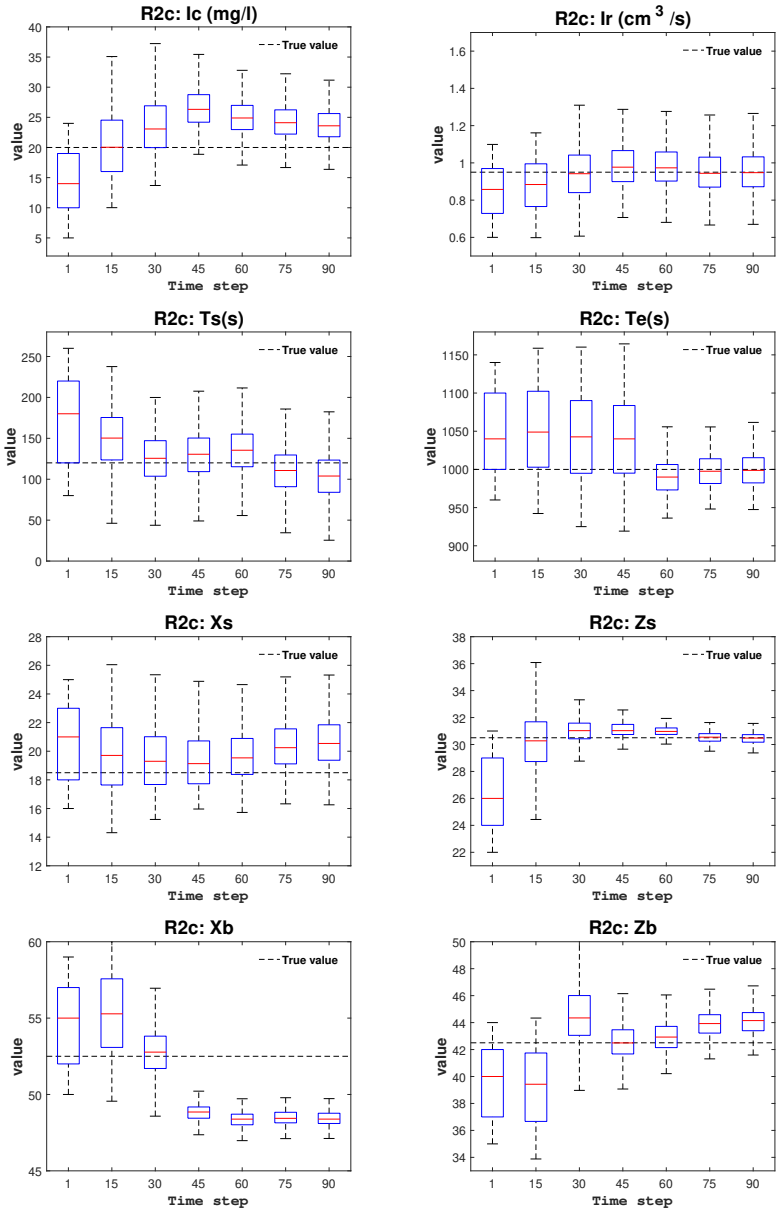

Figure 2.16. Boxplot of the 8 updated parameters in scenario *R2c* at different time steps (1, 15, 30, 45, 60, 75, 90).

2.5 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 in 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 covariance 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.

3

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

Abstract

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

inflation, the Bauser's covariance inflation method works best in preventing filter collapse. Its application to the sandbox data shows that the rNS-EnKF can benefit from Bauser's inflation to reduce the number of ensemble realizations substantially in comparison with a filter without inflation; yet, arriving to a good joint identification of both the contaminant source and the spatial heterogeneity of the conductivities.

3.1 Introduction

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

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

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

Most contaminant source identification approaches can be classified into two main categories: optimization ones and probabilistic ones. In the optimization approaches, an objective function is built and the algorithm tries to

minimize the discrepancies between simulated and measured concentrations using an optimization approach such as least-squares regression or maximum likelihood (e.g., Amirabdollahian and Datta, 2014; Aral et al., 2001; Ayvaz, 2016; Gorelick et al., 1983; Mirghani et al., 2009; Wagner, 1992; Yeh et al., 2007). In the probabilistic approaches, the problem is cast in a stochastic framework and the algorithm tries to maximize the posterior probabilities of the simulated concentrations conditioned on the observed values using techniques such as those based on minimum relative entropy or the use of adjoint states (e.g., Bagtzoglou et al., 1992; Butera et al., 2013; Koch and Nowak, 2016; Neupauer and Wilson, 1999; Woodbury and Ulrych, 1996).

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

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

important source of information for conductivity heterogeneity identification will be missing.

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

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

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

3.2 Methodology

3.2.1 Groundwater Flow and Solute Transport Equations

Water flow and contaminant transport in the sandbox are modeled using the corresponding governing equations for groundwater flow (Bear, 1972) and contaminant transport (Zheng and Wang, 1999), the equations were already introduced in chapter 2.2.1, but they are repeated here for a matter of completeness:

$$S_s \frac{\partial h}{\partial t} = \nabla \cdot (K \nabla h) + w \tag{3.1}$$

$$\frac{\partial (\theta C)}{\partial t} = \nabla \cdot (\theta D \cdot \nabla C) - \nabla \cdot (\theta v C) - q_s C_s \tag{3.2}$$

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

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

3.2.2 The Ensemble Kalman Filter

The ensemble Kalman filter (EnKF) was developed by Evensen (1994) as an extension to the Kalman filter (KF). The main difference between the EnKF and the KF is that, in the KF, the state covariance matrix is propagated in time using an explicit expression based on a linear transition equation, while, in the EnKF, this covariance matrix is derived from the statistical analysis of an ensemble of state realizations obtained after the solution of the state equations in each realization of the ensemble. The advantage of the EnKF over the KF is for systems in which the state transition equation is not linear; in such a case, the linear transition equation used by the KF is only an approximation and the resulting covariance deteriorates in time; whereas, in the EnKF, since the covariance is directly calculated from actual state spatial distributions, its value is more accurate, with the only limitation that the covariance is computed from a finite ensemble of realizations (if the number of realizations is small, the resulting estimate may be also inaccurate).

Although the EnKF was initially developed to update only the state of the system as observations are gathered, it has been shown that it can be also used for the update of the parameters using what is called an augmented state that includes both state variables and the parameters that control them (e.g., Chen and Zhang, 2006; Houtekamer and Mitchell, 2001; Li et al., 2012a,c). In summary, the EnKF has been proven to be an efficient algorithm for parameter identification, for strongly non-linear state-transfer equations, (Franssen and Kinzelbach, 2009), and has received much attention in the last decades. Next, the algorithm is described for the case study at hand, that is, the identification of the parameters defining a contaminant source together with the identification of the conductivities in a sandbox experiment for which only concentration data are available. The equations of r-EnKF were already introduced in chapter 2.2.2, but they are repeated here for a matter of completeness.

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

$$S = \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} (X_s, Z_s, I_c, I_r, T_e)^T \\ (lnK_1, lnK_2, \ldots, lnK_{N_m})^T \\ (C_1, C_2, \ldots, C_{N_m})^T \end{pmatrix} \tag{3.3}$$

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

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

$$S_t^f = \psi(B_{t-1}^a, A_{t-1}^a) \tag{3.4}$$

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

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

$$S_t^a = S_t^f + \mathbf{K}_t \left[y_t^{obs} + \varepsilon_i - \mathbf{H} S_t^f \right] \tag{3.5}$$

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

$$\mathbf{K}_t = \mathbf{P}_t^f \mathbf{H}^T [\mathbf{H} \mathbf{P}_t^f \mathbf{H}^T + \mathbf{R}_t]^{-1} \tag{3.6}$$

$$\mathbf{P}_t^f = \frac{1}{N_e - 1} \sum_{i=1}^{N_e} \{ [S_{i,t}^f - \overline{S_t^f}][S_{i,t}^f - \overline{S_t^f}]^T \} \tag{3.7}$$

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

The normal-score EnKF

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

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

The restart NS-EnKF

The EnKF was designed to update both parameters and state variables at each assimilation step. That is, the discrepancy between forecasted and observed variables is used to update the whole augmented state (see Eq. (3.5)). However, in general in the case of subsurface flow and transport, and in particular in the case at hand of contaminant source identification, the updated states could be inconsistent with the updated parameters, either because the mass conservation laws are not longer abided, or because the updated state is not coherent with the updated contaminant source location. For this reason, the forecast of the augmented state to the next observation time is not done based on the updated augmented state at the previous time state, but it is preferable to perform a forecast from time zero with the latest updated parameters (Camporese et al., 2011; Crestani et al., 2012). This approach is called, for this reason, the restart ensemble Kalman filter, or, in our case, the restart normal-score ensemble Kalman filter (rNS-EnKF).

The forecast function in Eq. (3.4) changes into:

$$S_t^f = \psi[C_0, A_{t-1}^a] = \begin{pmatrix} A_{t-1}^a \\ C_t \end{pmatrix} \tag{3.8}$$

where C_0 stands for the initial contaminant concentration in the domain. The restart EnKF has been applied before, for instance, by Camporese et al. (2011) and Crestani et al. (2012).

Damping

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

$$S_t^a = S_t^f + \alpha \mathbf{K}_t \left[y_t^{obs} + \varepsilon_i - \mathbf{H} S_t^f \right] \tag{3.9}$$

Inflation Methods

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

$$S_{i,t}^{inf,f} = \sqrt{\lambda_t}(S_{i,t}^f - \overline{S_t^f}) + \overline{S_t^f} \tag{3.10}$$

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

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

$$\lambda_t = \frac{(\mathbf{R}_t^{-\frac{1}{2}} d_t)^T \mathbf{R}_t^{-\frac{1}{2}} d_t - n_b}{trace\{\mathbf{R}_t^{-\frac{1}{2}} \mathbf{H} \mathbf{P}_t^f (\mathbf{R}_t^{-\frac{1}{2}} \mathbf{H})^T\}} \tag{3.11}$$

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

$$d_t = y_t^{obs} - \mathbf{H} * \overline{S_t^f} \tag{3.12}$$

Then, the updated augmented state vector is calculated as:

$$S_{i,t}^a = S_{i,t}^{inf,f} + \lambda_t \mathbf{P}_t^f \mathbf{H}^T [\mathbf{H} \lambda_t \mathbf{P}_t^f \mathbf{H}^T + \mathbf{R}_t]^{-1} \left[y_t^{obs} + \varepsilon - \mathbf{H} S_t^{inf,f} \right] \tag{3.13}$$

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

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

$$\lambda_t^f = \lambda_{t-1}^a \tag{3.14}$$

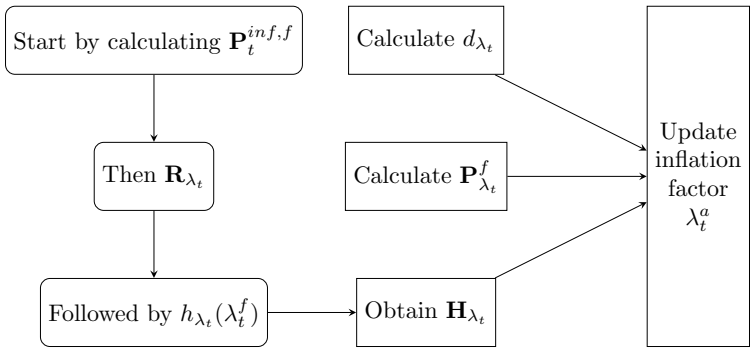

Figure 3.1. A flowchart of Bauser's method to update the inflation factors, λ_t^a.

$$\lambda_t^a = \lambda_t^f + \mathbf{K}_{\lambda_t}[d_{\lambda_t} - h_\lambda(\lambda_t^f)] \tag{3.15}$$

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

$$\mathbf{K}_{\lambda_t} = \mathbf{P}_{\lambda_t}^f \mathbf{H}_{\lambda_t}^T [\mathbf{H}_{\lambda_t} \mathbf{P}_{\lambda_t}^f \mathbf{H}_{\lambda_t}^T + \mathbf{R}_{\lambda_t}]^{-1} \tag{3.16}$$

$$(h_{\lambda_t}(\lambda_t^f))_i = [(\mathbf{R}_{\lambda_t})_{ii}]^{\frac{1}{2}} \tag{3.17}$$

The covariance of the inflation parameter, $\mathbf{P}_{\lambda_t}^f$, the observation matrix $\mathbf{H}_{\lambda_t}$ and the inflation parameter observation error $\mathbf{R}_{\lambda_t}$ can be obtained from the state covariance matrix $\mathbf{P}_t^f$, the observation matrix $\mathbf{H}$ and the observation error covariance matrix $\mathbf{R}$ of the augmented state vector $\mathbf{S}$ by:

$$(\mathbf{P}_{\lambda_t}^f)_{ij} = \sigma_\lambda^2 |(\mathbf{P}_t^f)_{ij}| [(\mathbf{P}_t^f)_{ii} (\mathbf{P}_t^f)_{jj}]^{-\frac{1}{2}} \tag{3.18}$$

$$(\mathbf{H}_{\lambda_t})_{ij} = [2[(\lambda_t^f)_j]^{\frac{1}{2}} (h_{\lambda_t}(\lambda_t^f))_i]^{-1} \sum_m (\mathbf{H})_{ij} (\mathbf{H})_{im} (\mathbf{P}_t^f)_{jm} [(\lambda_t^f)_m]^{\frac{1}{2}} \tag{3.19}$$

$$(\mathbf{R}_{\lambda_t})_{ij} = |(\mathbf{R})_{ij} + (\mathbf{H}\mathbf{P}_t^{inf,f}\mathbf{H}^T)_{ij}| \tag{3.20}$$

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

$$\mathbf{P}_t^{inf,f} = (\sqrt{\lambda_t^f} \sqrt{\lambda_t^f}^T) \cdot \mathbf{P}_t^f \tag{3.21}$$

A workflow summarizing how to apply Bauser's inflation method is shown in Figure 3.1.

Finally, the updated augmented state vector is computed as:

$$S_{i,t}^a = S_{i,t}^{inf,f} + \mathbf{P}_t^{inf,f} \mathbf{H}^T [\mathbf{H}\mathbf{P}_t^{inf,f} \mathbf{H}^T + \mathbf{R}_t]^{-1} \left[y_t^{obs} + \varepsilon - \mathbf{H} S_t^{inf,f} \right] \tag{3.22}$$

Table 3.1. Parameters used in the groundwater flow and transport models

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

3.3 Sandbox Experiment

The lab set up is the same as in chapter 2.3.1, it is repeated here for a matter of completeness. A contaminant experiment was carried out in a sandbox with sodium fluorescein as the tracer. The size of the sandbox is 120 cm by 14 cm by 70 cm. Two reservoirs with constant water levels at 62.5 cm and 60.6 cm with respect to the bottom of the sandbox are set at the upstream and downstream boundaries, respectively. (Notice that the experiment was performed with the upstream boundary on the right side of the sandbox, and all figures are represented in this way.) These two tanks define prescribed head boundaries, the bottom of the sandbox is impermeable and the top boundary is the phreatic surface. Between the upstream and downstream tanks, the area filled with glass beads has a size of 95 cm by 10 cm by 70 cm, which, for the purpose of modeling, is discretized into 95 columns, 1 row, and 70 layers of equal-sized cells of 1 cm by 10 cm by 1 cm. The sandbox is filled with glass beads of two different diameters, 1 mm and 4 mm, according to a spatial arrangement generated using a truncated Gaussian simulation (Journel and Isaaks, 1984) with the first quartile as the truncation threshold, resulting in a large-bead proportion of 0.25. The spatial distribution of the glass beads in the sandbox can also be seen in Figure 3.2. An injector is located at column 86, layer 40, at the position identified with a red dot in the figure. The whole sandbox was placed in a darkroom with a blue light source that was used to excite the injected fluorescein. Pictures of the plume, as it evolved in time, were taken and luminosity values were converted into concentration after a calibration procedure following Citarella et al. (2015).

The hydraulic properties of the beads (Table 3.1) had been characterized before with the same sandbox equipment (e.g., Cupola et al., 2015b; Citarella et al., 2015). The hydraulic conductivity of the large beads was estimated as 10.4 cm/s, and that of the small beads as 0.65 cm/s. The porosity is constant, independent of the bead size, and equal to 0.37. The longitudinal dispersivity within the large beads was estimated as 0.25 cm, and within the small beads as 0.106 cm. The ratio of transverse to longitudinal dispersivity is constant and equal to 0.45.

Although after processing the pictures the spatial distribution of concentration is fully known within the entire central area of the sandbox (dashed

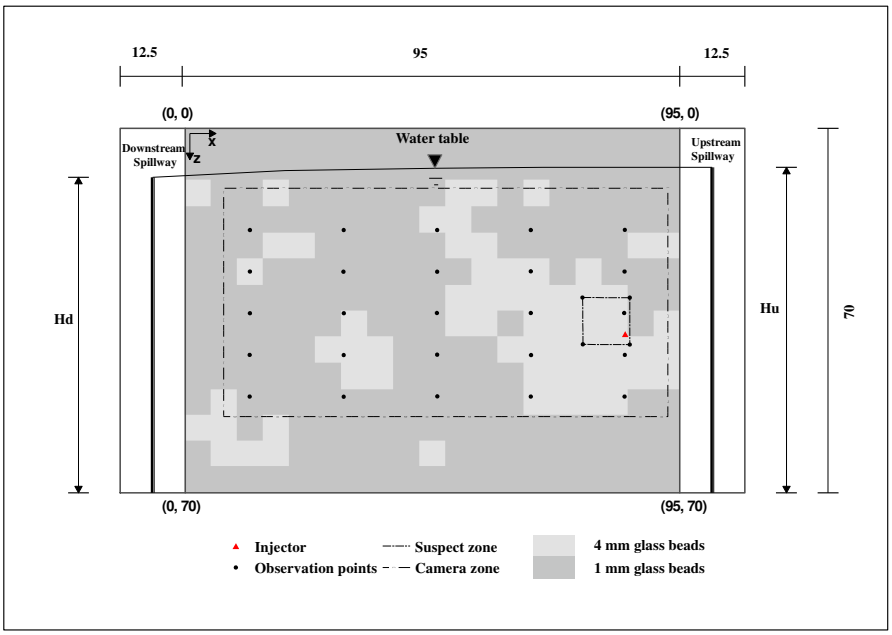

Figure 3.2. Sketch of the experimental device (view from the camera side inside the darkroom). H_u and H_d stand for the constant head boundaries, the dashed rectangle corresponds to the area captured by the camera in which concentrations will be monitored, the red triangle is the release location, and the small square around the red dot indicates the release suspect location during the identification process. Units are in cm. Pairs of numbers in parenthesis refer to row and column pairs in the numerical model.

rectangle in Figure 3.2), in order to mimic a potential sampling campaign in the field, only the concentrations observed at the twenty nine dots identified as observation points in the figure will be used for the purpose of identifying both the hydraulic conductivity and the contaminant source parameters. The release lasted 1200 s, the fluorescein concentration was 20 mg/l and the injection rate 2.60 cm^3/s. Observations were taken until after 3000 s from the beginning of the injection, every 30 s for a total of 100 observations at each observation point.

3.4 Definition of Scenarios and Ensemble Initialization

On a first attempt to apply the rNS-EnKF directly with the observed sandbox concentrations, some difficulties were found mostly related with the filter collapsing. These difficulties lead to perform a synthetic experiment prior to the application to the real data to analyze the impact of the number of ensemble realizations and the use of different approaches to prevent the filter to collapse. For this purpose, a reference set of synthetic concentrations was generated by solving, numerically, the flow and transport equations in a field with the same spatial distribution of conductivities as the sandbox, the same boundary conditions, and the same solute injection pulse. Then, 6 scenarios ($S1 - S6$) were analyzed with different ensemble sizes and different damping and inflation methods, more precisely, two ensemble sizes were tested (500 and 1000), two values for the damping coefficient (damping with a factor of 0.1 and with a factor of 0.5) and two covariance inflation methods (Wang's method and Bauser's method). After the analysis of the results using the synthetic reference, the conclusion was reached, as discussed below, that Bauer's inflation method was the best method to prevent filter collapse, thus two additional scenarios ($R1 - R2$) were run using the experimental data to test Bauer's inflation approach. The combination of ensemble sizes and inflation methods for the different scenarios is shown in Table 3.2.

The initial ensembles of log-conductivity realizations are the same for all scenarios (for the scenarios of 500 realizations only the first 500 of a total of 1000 realizations are retained). They are generated using a Gaussian random function with a mean equal to the weighted mean of the bead log-conductivities, 1.07 ln cm/s, and a variance equal to the variance of a binary Gaussian mixture of two facies with the means and proportions of the sandbox and an internal variance of one within each facies, i.e., 1.55 $(\ln cm/s)^2$. The correlation range of the log-conductivities is isotropic and equal to 15 cm. Previous studies (Xu et al., 2013a), in which no conditioning conductivity values had been used —as in this case—, have shown that the initial ensemble of log-conductivities is not as important as a sufficient number of observations of the state of the aquifer.

Table 3.2. Definition of scenarios

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

Table 3.3. Suspect ranges of source parameters for the generation of the initial ensemble of realizations and their true values

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

Similarly, the initial ensembles of source locations and pulses are the same for all scenarios. They are generated within suspect ranges that are defined using uniform distributions. The suspect source location (X_s, Z_s), in cm, ranges in U[78, 86] $\times$ U[38, 47] (see Figure 3.2), the suspect injection rate ranges in U[2, 3] cm^3/s, the suspect injection concentration ranges in U[5, 25] mg/l and the suspect final release time ranges in U[1050, 1250] s (see Table 3.3). These parameters are generated independently among them and of the log-conductivities. These ranges are used exclusively for the generation of the initial ensembles; afterwards, the updated parameter values are not restricted by any bounds.

3.5 Performance Evaluation

The rNS-EnKF is applied to each scenario assimilating the observed concentrations at the points indicated in Figure 3.2 at each time step. No log-conductivity or piezometric head data are observed at any time. After

assimilating the concentration data at the end of each time step, the filter provides an ensemble of updated parameters, which are analyzed in different ways:

1. Computing the ensemble mean and variance of the contaminant source parameters at the end of each time step. The ensemble mean can be interpreted as a parameter estimate and the variance as a measure of the estimation uncertainty,

2. Visually analyzing the spatial variability of the cell ensemble mean and ensemble variance of log-conductivities with respect to the reference log-conductivity spatial distribution,

3. Computing the root mean-squared error (RMSE), the ensemble spread (ES), and the ratio RMSE/SE of log-conductivities as given by

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\ln K_i^{ref} - \overline{\ln K_i})^2}, \tag{3.23}$$

$$\text{ES} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\sigma^2_{\ln K_i}}, \tag{3.24}$$

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

3.6 Results

As mentioned above, two analyses have been performed, a preliminary one using synthetic data to decide on the number of realizations and on a method to prevent filter collapse, followed by a specific analysis of the data collected at the sandbox experiment.

3.6.1 Analysis of the Synthetic Data

The synthetic analysis is performed on six scenarios with combinations between two numbers of realizations and five alternatives to prevent filter collapse as given in Table 3.2. Recall that the reference for the synthetic

case comes from a numerical simulation of flow and transport with the same characteristics as the sandbox experiment.

Figures 3.3 and 3.4 focus on the source parameters, they provide the ensemble mean and the ensemble variance, respectively, of all five source parameters, after the update at each time step for all six scenarios. The ranges of the ensemble variances were very different for each parameter and for this reason the results are displayed after standardization by the ensemble variances of the initial ensembles. It is hard to argue which is the scenario that performs best. Scenario $S3$, the one with a damping factor of 0.1, can be discarded since it is the one that ends with the highest variances for most of the parameters. Scenario $S5$, the one with Wang's inflation method, should also be discarded because it collapses the ensemble after a few time steps as shown by the rapid decrease of the ensemble variance to zero for almost all parameters. Scenario $S2$, with no inflation, but 1000 realizations –double than the rest of the scenarios– performs well in that it provides an estimate close to the true values and the variance decreases in time consistently and similarly to the rest of the scenarios. Scenario $S1$, with no inflation and 500 realizations shows some filter collapse, which does not happen as quickly as for $S5$ but ends with similar magnitudes for the ensemble variances. Scenario $S4$, with a damping factor of 0.5, does a good job in the estimation of the source parameters, except for Ic but the final uncertainties are the largest after $S3$ for most of the parameters. Finally, scenario $S6$, with Bauser's inflation method, could be considered as the one with the best performance, since it provides very good estimates for all parameters, except for Ir, and it has low final uncertainties without filter collapse. All methods estimate the vertical position Zs of the release point lower in the sandbox than its real position, this behavior can be produced by local velocity variations induced by the proximity of the injection to the boundary between two cells with different glass bead diameters which are not resolved by the observations.

Figure 3.5 shows the ensemble mean and Figure 3.6 the ensemble variance of the initial $\ln K$ realizations and of the updated ones computed at the 90th time step for all synthetic scenarios. The ensemble mean and ensemble variance of the initial $\ln K$ are almost homogeneous and equal to their prior values since no conditional data of $\ln K$ is employed. After assimilating all concentration data during 90 time steps, the ensemble mean of the updated $\ln K$ conductivities can capture the main patterns of variability of the glass bead distribution with a substantial reduction of the ensemble variance in most of the sandbox. A comparison among the different scenarios shows that, again, $S3$ performs worst, with the worst estimation of $\ln K$ and the largest estimation variances and $S5$ shows filter collapse at most locations. Of the remaining scenarios, $S2$ and $S6$ give the best results, with $S2$ being slightly better in $\ln K$ pattern estimation thanks to the larger number of ensemble members. For a more quantitative evaluation of the identification

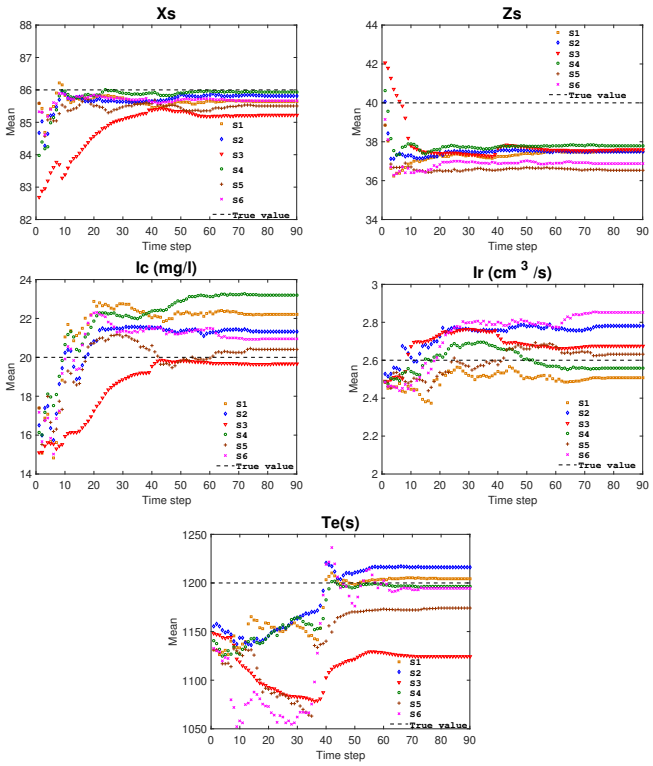

Figure 3.3. Time evolution of the ensemble means of the updated contaminant source parameters for all the synthetic scenarios ($S1 - S6$).

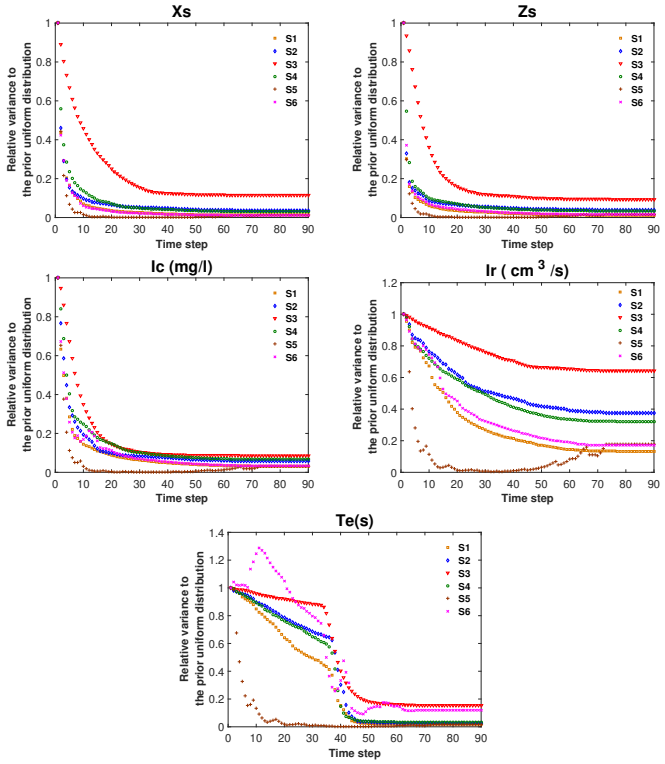

Figure 3.4. Time evolution of ensemble variances of the updated contaminant source parameters for all synthetic scenarios($S1 - S6$). Each variance plot has been standardized by the variance of the initial ensemble.

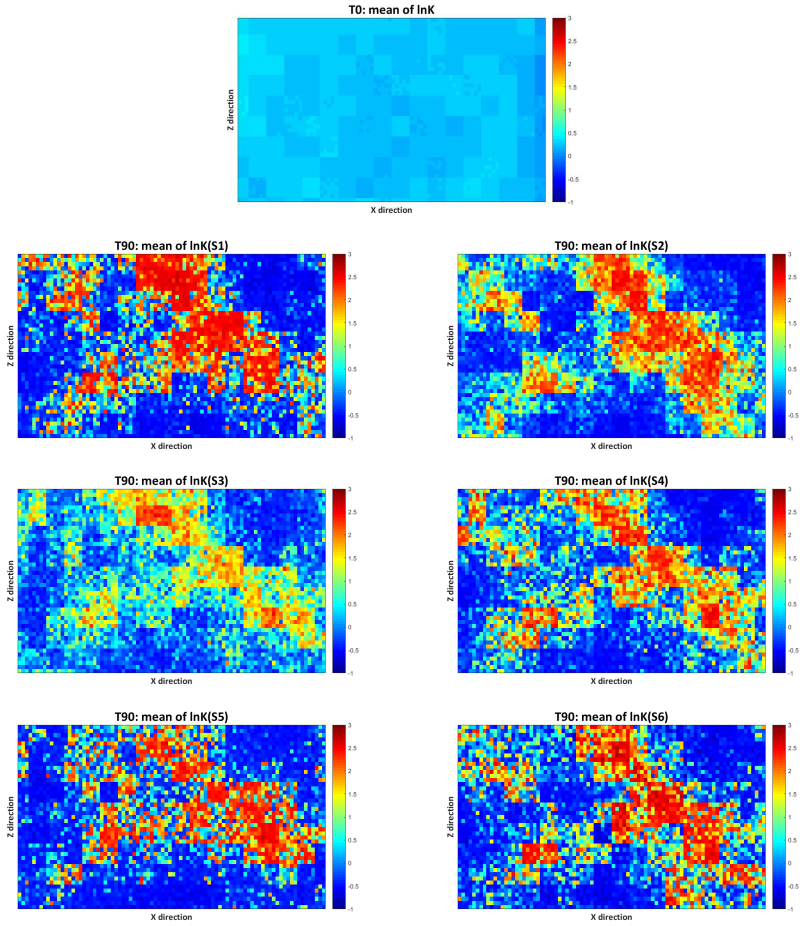

Figure 3.5. Ensemble mean of the initial ln$K$ realizations and the updated ln$K$ realizations of all synthetic scenarios($S1 - S6$) at the $90th$ time step.

of ln$K$, Figure 3.7 shows how the three statistics RMSE, ES and RMSE/ES evolve in time as the data assimilation proceeds. The best performance would be for the lowest values of RMSE and ES and the closest-to-one RMSE/ES ratio. The two best scenarios are $S2$ and $S6$, with $S6$ having the RMSE/ES ratio closest to one.

Taking into consideration the performance of the rNS-EnKF for the different synthetic scenarios, the two scenarios that will be analyzed with the experimental data are the non-inflation method with 1000 realizations, referred to as $R1$, and the Bauser's inflation method with 500 realizations, referred to as $R2$.

3.6.2 Analysis of the Sandbox Data

The difficulties found on the first attempt to apply the rNS-EnKF to the sandbox data must be due to observation errors in the concentrations. According to earlier work (Chen et al., 2018), an underestimation of the observation error will force the filter to fit the concentrations too closely producing biased estimates of the parameters, and an overestimation of the observation error will allow too loose a fit producing estimates with large uncertainty. Since the same sandbox equipment as Cupola et al. (2015b) and Chen et al.

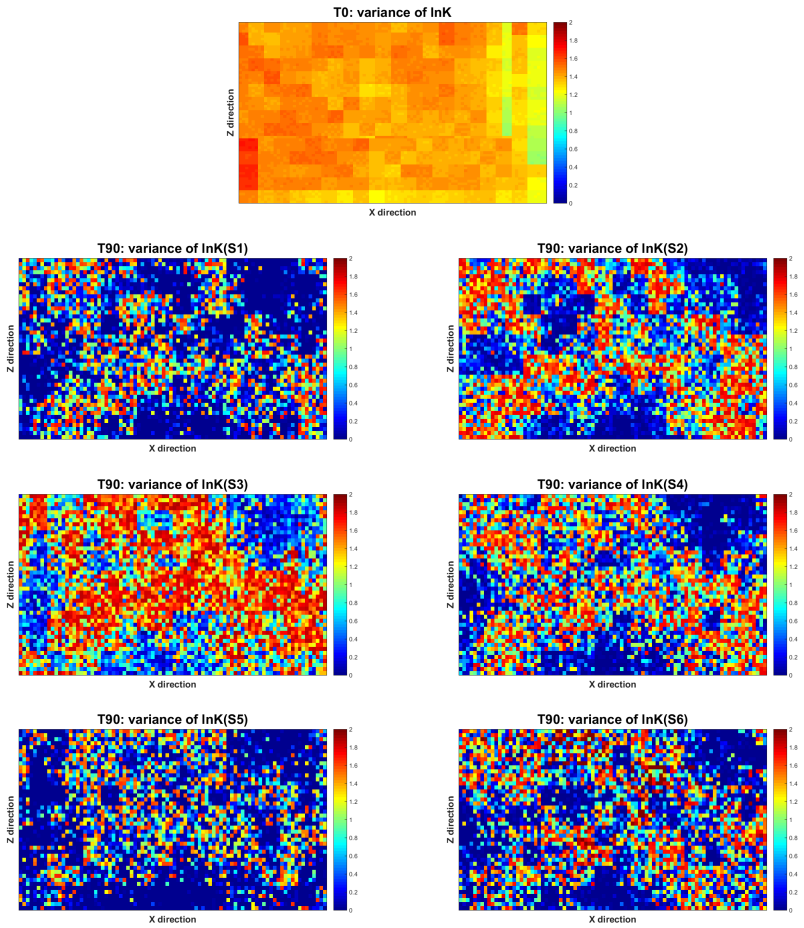

Figure 3.6. Ensemble variance of the initial $\ln K$ realizations and the updated $\ln K$ realizations of all synthetic scenarios$(S1 - S6)$ at the $90th$ time step.

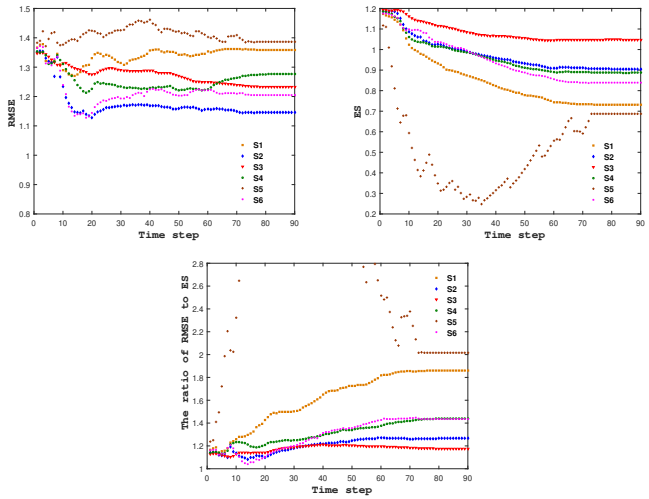

Figure 3.7. Time evolution of ln$K$ RMSE, ES and the ratio of RMSE to ES for all synthetic scenarios($S1 - S6$).

(2018) is used, the same observation error distribution with a mean of 0 mg/l and a standard deviation of 1 mg/l is retained for this analysis.

Figures 3.8 and 3.9 show the evolution of the ensemble mean and the ensemble variance, respectively, of the contaminant source parameters for the two sandbox scenarios ($R1, R2$). Both approaches perform well with mean estimates close to the true values and estimation variances that go down close to zero for all parameters. It seems that the injection concentration and the injection rate are more difficult to identify, they have the largest estimation error and the largest estimation variance; however, if the mass loading rate is computed, that is the product of the injection rate times the injection concentration, its mean and variance is similar to those of the other contaminant parameters. This result seems to indicate that there may be some indetermination in the identification of parameters Ic and Ir that disappears when the subject of identification is its product. Disregarding parameters Ic and Ir, it can be concluded that both scenarios perform equally and, therefore, that Bauser's inflation method can make up for the reduction from 1000 realizations to 500 realizations with similar performance.

Figure 3.10 shows the ensemble mean and variance of ln$K$ for scenarios $R1$ and $R2$ at the 90th time step. Figure 3.11 shows the ensemble mean of the absolute differences between the reference and updated ln$K$ maps at the 90th time step. Both scenarios capture the main patterns of variability of ln$K$ and the ensemble variance is substantially reduced in the areas of low conductivity. Comparing the two scenarios, variance reduction is larger for scenario $R2$ and the absolute deviations between reference and estimated conductivities are smaller for $R2$, implying again that Bauser's inflation method is a valuable approach to reduce ensemble size and achieve similar

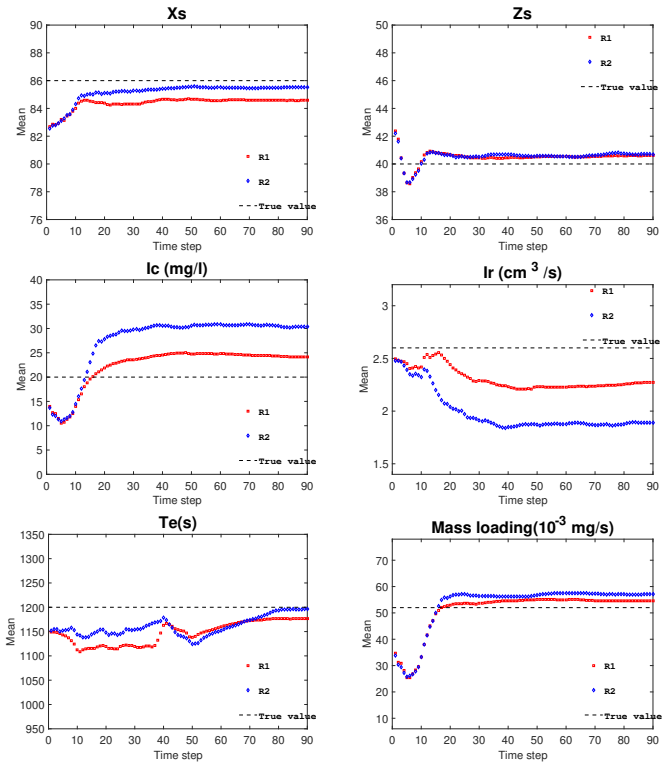

Figure 3.8. Time evolution of the ensemble means of the updated contaminant source parameters for the two sandbox scenario $(R1, R2)$. Also shown the mass loading rate $Ic \cdot Ir$.

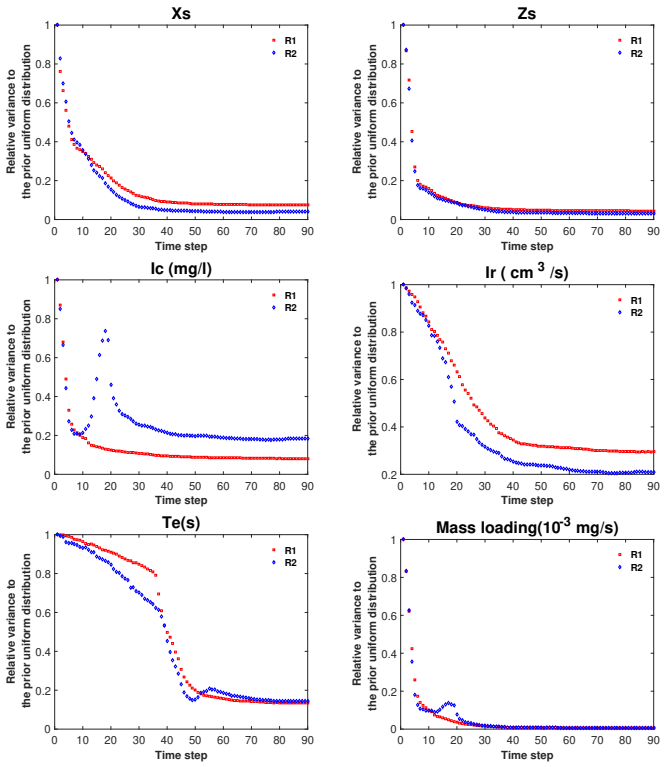

Figure 3.9. Time evolution of the ensemble variances of the updated contaminant source parameters for the two sandbox scenario $(R1, R2)$. Also shown the mass loading rate $Ic \cdot Ir$. Notice that each ensemble variance has been normalized by their values at time zero.

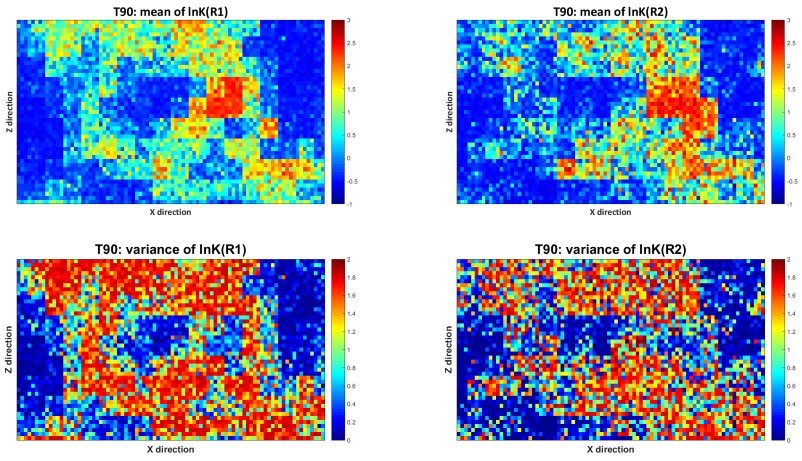

Figure 3.10. Ensemble mean (top row) and ensemble variance (bottom row) of updated $\ln K$ of scenarios $R1$ and $R2$ at the $90th$ time step.

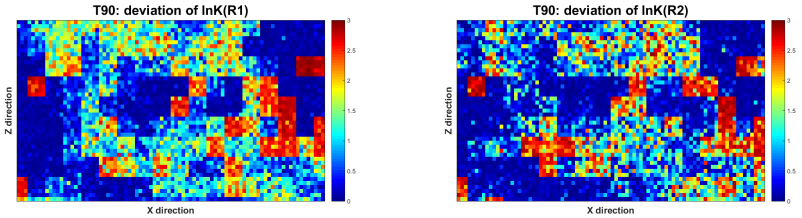

Figure 3.11. Ensemble mean of the absolute deviation between reference and updated $\ln K$ in scenarios $R1$ and $R2$ at the $90th$ time step.

(or better) results as when a larger ensemble is used. Figure 3.12 shows the evolution in time of the $\ln K$ RMSE, ES and RMSE/ES ratio for scenarios $R1$ and $R2$. Again, scenario $R2$ performs remarkably well as compared to scenario $R1$, with a similar RMSE, smaller ES and a ratio RMSE/ES not too far from one.

Figure 3.13 shows the evolution of the contaminant plume in the sandbox at the 10th, 40th, 60th and 90th time steps. Figures 3.14 and 3.15 show the ensemble mean of the contaminant plumes for scenarios $R1$ and $R2$, respectively, at the same time steps as in Figure 3.13 computed with all the parameters updated at the 90th time step. The comparison of the simulated plumes with the observed ones is very favorable, demonstrating that the estimated parameters are conditioned on the observed concentrations, and that they are capable of giving a good prediction of contaminant movement.

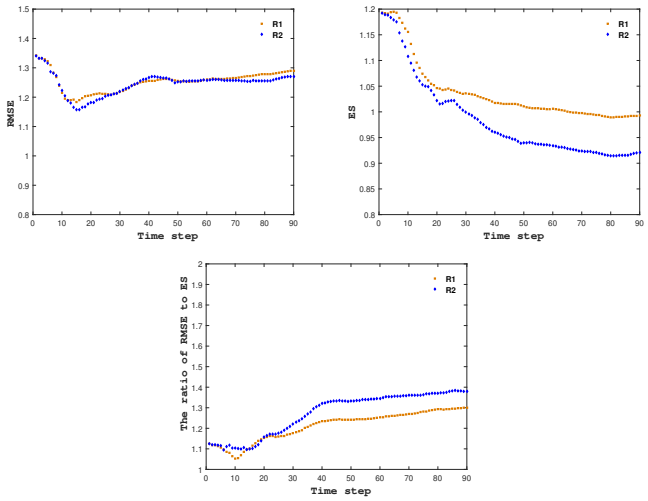

Figure 3.12. Time evolution of ln$K$ RMSE, ES and the ratio of RMSE to ES for scenarios $R1$ and $R2$.

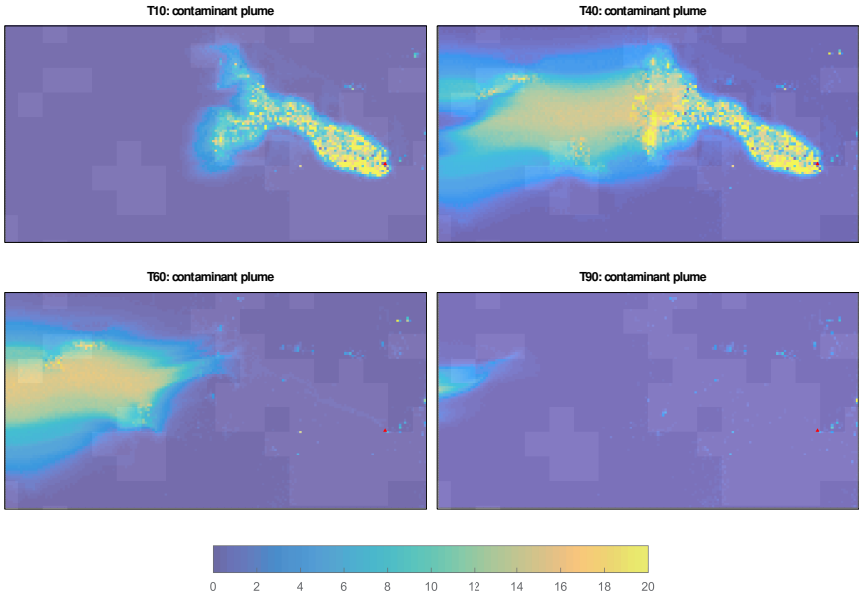

Figure 3.13. Reference contaminant plume evolution at the 10th, 40th, 60th and 90th time steps in the sandbox. Red triangle denotes the real injector.

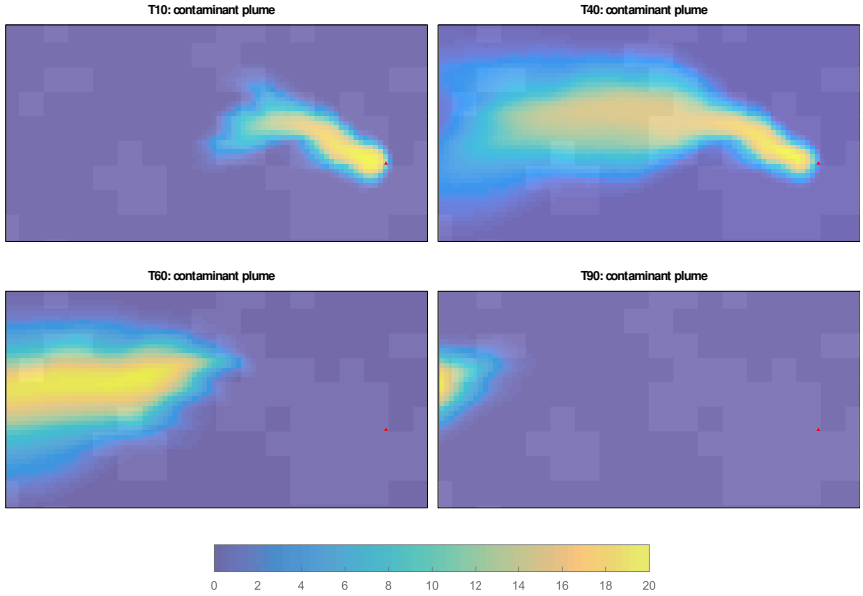

Figure 3.14. Ensemble mean of contaminant plume evolution of scenario $R1$ at the 10th, 40th, 60th and 90th time steps with all parameters updated after the 90th time step. Red triangle denotes the real injector.

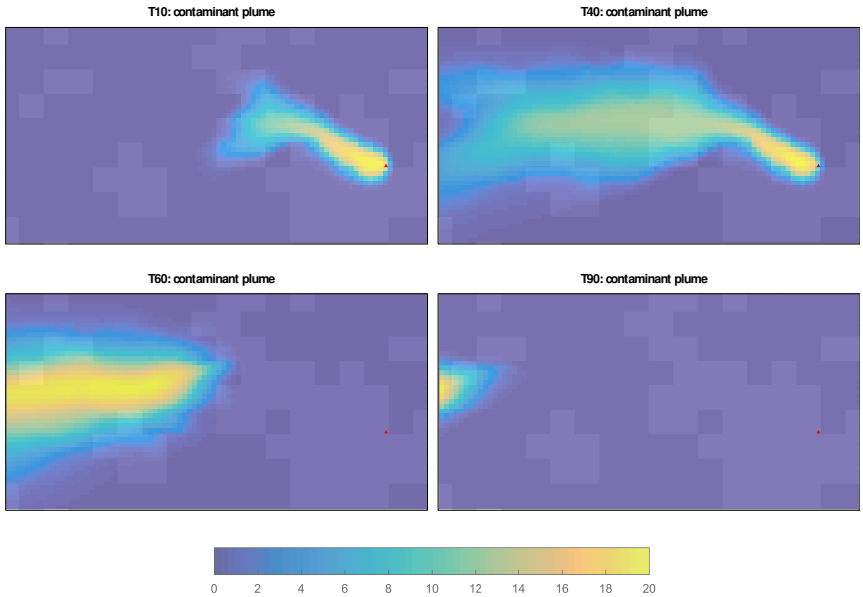

Figure 3.15. Ensemble mean of contaminant plume evolution of scenario $R2$ at the 10th, 40th, 60th and 90th time steps with all parameters updated after the 90th time step. Red triangle denotes the real injector.

3.7 Discussion and Conclusions

Xu and Jaime (2018) showed the capabilities of the the restart normal-score ensemble Kalman filter (rNS-EnKF) for the simultaneous identification of source parameters and hydraulic conductivities in synthetic aquifers. This work presents the first attempt to apply it to a non-synthetic exercise. An aquifer is mimicked by a laboratory sandbox in which geometry, initial and boundary conditions are known. The first finding was that it was not straightforward to apply the approach to the collected data; working under laboratory conditions does not preclude measurement and other errors, what prevented the filter to work properly on first attempts. The filter would collapse, even for large ensemble sizes, what led to an analysis of a synthetic case using solute concentrations generated by a numerical model, thus getting rid of model or measurement errors. In this synthetic exercise, six scenarios were compared showing the importance of a good selection of an approach to prevent filter collapse. Of the four alternative approaches, Bauser's covariance inflation method appeared as the most appropriate, allowing to reduce the ensemble size from 1000 members (without inflation) to 500 (with inflation) to yield similar results. In these synthetic scenarios, it could be observed also that the horizontal coordinate of the source was well identified, but that the vertical one was estimated a little bit downwards from the original position. The explanation must be due to the closeness of the source to a boundary between the large glass beads and the small ones. The synthetic results also showed that it is difficult to identify a binary conductivity field starting from a continuous distribution of log-conductivities, yet, the two main zones of high and low conductivities are well captured in the different scenarios, with the scenario having 1000 realizations performing best, followed by the scenario with 500 realizations and using Bauser's covariance inflation method.

The application of Bauser's inflation and 500 realization to the data observed in the sandbox was compared with a non-inflated filter and 1000 realizations, with comparable results. The identification of the source parameters is good in both cases, even for the vertical coordinate of the injection. A better identification of the source vertical position in the sandbox than in the synthetic exercises could be explained by the larger measurement error variance used in the sandbox observations than in the synthetic scenarios. A larger measurement error gives the filter more flexibility to update the parameters to fit the observations while resulting in a larger variance on the ensemble of final parameters. It was also evident that the estimation of both injection rate and injection concentration were biased; a further analysis showed that there is a degree of indetermination in the estimation of these two parameters since the parameter that really matters is their product, the mass loading rate. The mass loading rate is well estimated with no bias and little uncertainty. As in the synthetic case, the estimation of a

binary conductivity field by a continuous one is almost impossible, but the final ensemble of log-conductivities displays enough spatial heterogeneity to distinguish two main areas of high and low conductivities, and, more importantly, the solution of the mass transport equation in the final conductivity fields yields a contaminant plume that moves in space and time in a very similar pattern as the one observed in the sandbox.

It is important to notice that, in the sandbox experiment, the only available data was concentration data; no observations of either conductivities or piezometric heads were available. In a practical case, both conductivity and piezometric head data could also be assimilated resulting in an improved estimation of all parameters being identified.

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

4

A comparison between ES-MDA and restart EnKF for the purpose of the simultaneous identification of a contaminant source and hydraulic conductivity

Abstract

Understanding a contaminant source may help in a better management and risk assessment of a polluted aquifer. However, contaminant source information may not be available when a pollutant is detected in a drinking well. The restart ensemble Kalman filter (r-EnKF) has been demonstrated in synthetic and laboratory experiments as an efficient solution for the identification of a contaminant source. Recently, the ensemble smoother with multiple data assimilation (ES-MDA) has been proposed as an alternative to the r-EnKF as a more efficient solution given that the r-EnKF needs to restart the simulation of the state equation from time zero after each data assimilation step. An analysis, in a synthetic aquifer, of the performance of the ES-MDA for the simultaneous identification of a contaminant source and the spatial distribution of hydraulic conductivity by assimilating both

piezometric head and concentration observations is carried out using the r-EnKF as a benchmark. The conclusion is that for the ES-MDA to reach the same level of accuracy as the r-EnKF the number of multiple data assimilations must be large, and therefore, the apparent advantage of the ensemble smoother, i.e., the assimilation of all observational data at once, vanishes. The ES-MDA is able to outperform the r-EnKF, marginally, for the specific synthetic case analyzed, only for a sufficiently large number of iterations at a cost of larger CPU consumption than the r-EnKF, and it can perform far better than the r-EnKF just with a cost of larger CPU consumption.

4.1 Introduction

When contaminant is released into the subsurface, it will jeopardize not only human health but also damage the local ecosphere, especially if the contaminant is hazardous. When contamination happens inadvertently or is purposely hidden, it may be difficult to trace it back from concentration observations taken downstream from the source.

Yet, knowledge of the contaminant source is vital for groundwater contamination management, contamination control, contamination risk assessment and remediation.

How to identify a contaminant source once contamination has been detected has attracted much attention in the last decades. It is a difficult problem that has been addressed using inverse modeling. According to their characteristics, the inverse modeling approaches for contaminant source identification could be classified into three categories: optimization approaches, probabilistic approaches and deterministic approaches. The reader is referred to the reviews by Sun et al. (2006b); Atmadja and Bagtzoglou (2001b) for further information.

In the optimization approaches, the objective is to minimize an objective function that measures the differences between simulated concentrations and measurement observations and that is written in terms of the parameters defining the contaminant source. Some of the approaches used are least-squares regression and linear programming (Gorelick et al., 1983), maximization of correlation coefficients (Sidauruk et al., 1998), constrained robust least squares (CRLS) (Sun et al., 2006a), CRLS estimator combined with a branch-and-bound global optimization (Sun et al., 2006b), evolutionary search algorithms (Mirghani et al., 2009), or hybrid simulation-optimization (Ayvaz, 2016).

In the probabilistic approaches, the objective is, generally speaking, to maximize some posterior probability of the source parameters given the observations. Some approaches used for this purpose are minimum relative entropy (Woodbury and Ulrych, 1996; Woodbury et al., 1998; Cupola et al., 2015a), the geostatistical approach (Sun, 2007; Gzyl et al., 2014; Butera

et al., 2013), Markov chain Monte Carlo (Wang and Jin, 2013), or a Bayesian approach (Zeng et al., 2012; Zhang et al., 2015; Zanini and Woodbury, 2016).

In the deterministic approaches, the main objective is to solve the advection-dispersion equation backward in time. Some of the approaches employ the marching-jury backward beam equation method (Atmadja and Bagtzoglou, 2001a; Bagtzoglou and Atmadja, 2003), Tikhonov regularization (Skaggs and Kabala, 1994; Neupauer et al., 2000), or a quasi-reversibility method together with minimum relative entropy (Skaggs and Kabala, 1995).

In addition to the three types of approaches mentioned above, recently, the use of the restart ensemble Kalman filter (r-EnKF), was proposed by Xu and Gómez-Hernández (2016b) to identify a contaminant source by assimilating concentration observations. The work was inspired by the EnKF, which can give the efficient results obtained in the solution of standard inversion problems(e.g., Franssen and Kinzelbach, 2009; Xu et al., 2013a; Xu and Gómez-Hernández, 2015). Later, (Xu and Jaime, 2018) extended their work to jointly identify the source information and the underlying hydraulic conductivity field in a synthetic aquifer, and they also have successfully tested it in a tank experiment (Chen et al., 2018). These works have proven the capability of the EnKF for contaminant source identification. However, given the nature of the ensemble Kalman filter, with an extended state vector including the parameters controlling the state equation, it was impossible to forecast the state (concentration distribution) from the updated concentrations and to account for the updated parameters, which, in this case, are the ones describing the source. To consider the updated parameters (say, the updated location of the release) there is a need to restart the simulation from time zero after each updating step, since the contaminant source parameters refer to the source at time zero, but this restart makes it very time consuming as the number of observation steps increases, as well as some updated source parameters still with high uncertainty (e.g., Xu and Gómez-Hernández, 2016b; Xu and Jaime, 2018; Chen et al., 2018).

The ensemble smoother (ES), first proposed by van Leeuwen and Evensen (1996), is an alternative that could alleviate the computational burden of the EnKF, because it assimilates all data for all time steps at once. This avoids the restart of the simulation at every time step and makes the ES faster and easier to implement than EnKF (Emerick and Reynolds, 2013a). However, the performance of the ES for the case of non-linear state equations is not good (e.g., Evensen and van Leeuwen, 2000; Crestani et al., 2012), the main reason being the lack of multiple updating inherent to the EnKF.

A detailed explanation why the EnKF outperforms the ES in dealing with non-linear problems can be referred to the work Evensen (2018). Here, a brief explanation is given as follows. Both the EnKF and the ES rely heavily on covariances, which can only capture linear relationships. The EnKF recursively updates the parameters of interest by assimilating observation information in time to refine it close to the reference solution.The ES makes

a single update using all the data from all time steps. That is, the EnKF is equivalent to many linear approximations to the state equation followed by incremental updates along the linear approximation, whereas the ES is equivalent to a single linear approximation to the state equation and a single large update along the linear approximation. The EnKF is equivalent to a non-linear optimization based on local linear approximations, whereas the ES is a linear minimization, which may be very far from optimal if the state equation is highly nonlinear. Unless, iteration is also introduced into the ES. This is what Emerick and Reynolds (2013a) propose with their ensemble smoother with multiple data assimilation (ES-MDA). The basic idea is to assimilate all data from all time steps several times, progressively updating the parameters after each assimilation.

Several successful applications of the ES-MDA are reported in the reservoir history-matching literature (e.g., Emerick et al., 2013; Emerick and Reynolds, 2013b; Le et al., 2015, 2016; Lee et al., 2013; Fokker et al., 2016). In these works, the reservoir state equations are nonlinear, and the ES-MDA results outperforms the EnKF for both synthetic and real field problems. Recently, a few applications have been reported in the hydrogeology literature (Li et al., 2018b,a) for the characterization of hydraulic conductivities by assimilating piezometric heads.

In this paper, the ES-MDA is used, for the first time to the best of our knowledge, to jointly identify a heterogeneous hydraulic conductivity field and contaminant source information on a synthetic aquifer. As a benchmark, the performance of the ES-MDA will be compared with the restart EnKF (r-EnKF). Note that the main aim of this work is to evaluate the capability of the ES-MDA and compare the performance difference between the ES-MDA and r-EnKF in the joint identification of conductivity field and contaminant source information.

The paper is organized as follows. First, we introduce the algorithmic description of the r-EnKF and the ES-MDA, and then test and compare the ES-MDA with the r-EnKF on a synthetic aquifer. The paper ends with a summary discussion of whether the ES-MDA actually outperforms the r-EnKF or not.

4.2 Methodology

4.2.1 Ensemble Kalman filter

The EnKF was developed based on the Kalman filter proposed by Kalman et al. (1960) to better tackle nonlinear state-transfer equations. The main difference between the EnKF and the Kalman filter is on how the covariance matrices are calculated. In the original filter, the covariances were propagated in time using a linear state-transfer function (or a linear approximation in case the function is non-linear), while in the EnKF, the covariances are

calculated from the states obtained after solving the state-transfer function on an ensemble of realizations (Evensen, 2003, 2009; Chen and Zhang, 2006; Xu et al., 2013a; Xu and Gómez-Hernández, 2015). Like the Kalman filter, the EnKF consists of two steps: forecast and analysis. The first one is to forecast the state variables —in our case, piezometric heads and solute concentrations —at the t^{th} time step (B_t^f) from the state variables (B_{t-1}^a) and model parameters —in our case, hydraulic conductivities and contaminant source parameters —updated after the last time step (A_{t-1}^a).

However, as already discussed in Xu and Gómez-Hernández (2016b), it is impossible to take into account the updated parameters (A_{t-1}^a) in the forecast step when these parameters define the spatiotemporal position of a contaminant source unless the forecast is restarted from time zero. The details of r-EnKF was already introduced in chapter 2.2.2, but they are repeated here for a matter of completeness. The forecast equation,

$$S_t^f = \psi(B_0, A_{t-1}^a). \tag{4.1}$$

where S_t^f is an augmented matrix containing the state variables and model parameters , ψ represents the state-transfer function, and B_0 represents the state variables at time zero. The update step modifies the parameter values from the previous time step (A_{t-1}^a) as a function of the discrepancy between forecasted and observed state variables at observation locations

$$S_t^a = S_t^f + \mathbf{K}_t \left[y_t^{obs} + \varepsilon_i - \mathbf{H} S_t^f \right] \tag{4.2}$$

with

$$\mathbf{K}_t = \mathbf{P}_{AB,t}^f (\mathbf{P}_{BB,t}^f + \mathbf{R}_t)^{-1}, \tag{4.3}$$

where $y_t^{obs} + \varepsilon_i$ is the vector of observed concentrations and piezometric heads (composed of the sum of the true head or concentration y_t^{obs} plus an observation error ε_i of zero mean and covariance $\mathbf{R}_t$), $\mathbf{K}_t$ is the Kalman gain, $\mathbf{P}_{AB,t}^f$ is the cross-covariance between parameters and forecasted state variables at observation locations, and $\mathbf{P}_{BB,t}^f$ is the auto-covariance between the forecasted state variables at the observation locations.

If we assume there are N_e realizations in the ensemble and each realization has N_m elements. The state variable vector B contains piezometric heads H and concentrations C at all aquifer model cells

$$B = \begin{bmatrix} H \\ C \end{bmatrix}, \tag{4.4}$$

with N_e realizations of $2N_m$ variables, and the model parameter vector A contains hydraulic log-conductivity $\ln K$ in all aquifer model cells and the contaminant source parameters, which are source location, X for the x-

coordinate, and Y for the y-coordinate, initial release time T, release duration ΔT, and mass-loading rate M

$$A = \begin{bmatrix} \ln K \\ X \\ Y \\ T \\ \Delta T \\ M \end{bmatrix}. \tag{4.5}$$

with N_e realizations of $(N_m + 5)$ variables.

4.2.2 Ensemble smoother with multiple data assimilation

The ensemble smoother is, conceptually, the same as the r-EnKF but limited to one forecast step (for all the time steps for which observations are available) and a single update step (based on the discrepancies between observations and predictions at all time steps).

The equations that describe the ES are almost the same as those for the r-EnKF above, with some differences. The forecast step is given by

$$B^f = \psi(B_0, A_0). \tag{4.6}$$

where now B^f contains the state forecasted at all time steps —computed from the initial state B_0 and the initial ensemble of parameters A_0. And the update step is given by

$$A^a = A_0 + \mathbf{K}(y^{obs} + \varepsilon - B_o^f), \tag{4.7}$$

with

$$\mathbf{K} = \mathbf{P}_{AB}^f(\mathbf{P}_{BB}^f + R)^{-1}, \tag{4.8}$$

where y^{obs} are all of the observations at observation locations, ε are the observation errors, and B_o^f are the forecasts at observation locations. The covariances appearing in Eq. (4.8), $\mathbf{P}_{AB}^f$ and $\mathbf{P}_{BB}^f$ are computed for all time steps; these covariance matrices include the cross-covariances between time steps, an aspect not accounted for in the r-EnKF that it was thought could render the ES superior to the EnKF. From a computational point of view, if there are N_o observations locations sampled N_t times, the sizes of the matrices involved in the r-EnKF are proportional to N_o, whereas in the ES they are proportional to the product $N_o \cdot N_t$; Hence, the size of the cross-covariance $\mathbf{P}_{AB,t}^f$ is $(N_m + 5) \times 2N_o$, the size of the covariance matrix $\mathbf{P}_{BB,t}^f$ and R_t in Eq. (4.3) is $2N_o \times 2N_o$; whereas the size of the cross-covariance $\mathbf{P}_{AB}^f$ is $(N_m + 5) \times (2N_o \cdot N_t)$, the size of $\mathbf{P}_{BB}^f$ and R in Eq. (4.8) is $(2N_o \cdot N_t) \times (2N_o \cdot N_t)$.

The solution provided by Emerick and Reynolds (2013a) to improve the performance of the ES for non-linear state-transfer equations is to iterate, what is called multiple data assimilation (because the same data is assimilated multiple times) on the basis that each iteration of the ES is similar to a Gauss-Newton iteration (Reynolds et al., 2006; Gu and Oliver, 2007). Basically, Eq. (4.6) is iteratively applied using the latest updated parameters as the initial parameters for the next iteration.

However, since all data are assimilated multiple times, there is a need to inflate the observation error for each assimilation step. For this purpose, a non-increasing sequence of error variance inflation coefficients $\{a_j, j = 1, ..., N_a\}$ is used in the updating equations, with N_a being the number of assimilation iterations, and satisfying that $\sum_{j=1}^{N_a} \frac{1}{a_j} = 1$.

The ES-MDA equations display the following differences. The forecast step is given by

$$B_j^f = \psi(B_0, A_j). \tag{4.9}$$

where j is the iteration counter, and for each iteration the forecast uses the last updated parameters from the previous iteration. And the update equation is given by

$$A_{j+1} = A_j + \mathbf{K}_j(y^{obs} + \varepsilon - B_{o,j}^f) \tag{4.10}$$

with

$$\mathbf{K}_j = \mathbf{P}_{AB,j}^f(\mathbf{P}_{BB,j}^f + a_j R)^{-1}, \tag{4.11}$$

In Eq. (4.10) and Eq. (4.11), we can see how the observation variance is amplified by factor a_j and the observation error is amplified by $\sqrt{a_i}$.

Please notice that, in case that there may be a loss of rank in the ensemble when solving $(\mathbf{P}_{BB,j}^f + a_j R)^{-1}$ when the setting of the case is $N_e < 2N_o$ for the EnKF or $N_e < 2N_o * N_t$ for the ES-MDA, the subspace inversion introduced by Evensen (2004) is used to solve the problem. The detailed explanation can be referred to the work by Evensen (2004); Emerick and Reynolds (2013a).

4.3 Application

A synthetic confined aquifer is designed and constructed on a 1000 [L] by 1000 [L] by 50 [L] cube discretized into 50 by 50 by 1 cells, where each cell is 20 [L] by 20 [L] by 50 [L]. Please note that no specific units are given, only their dimensional analysis, any set of consistent units will yield the same results. The reference log-conductivity field is drawn from a multivariate Gaussian random process defined by the parameters in Table 4.1 using the GCOSIM3D —a sequential Gaussian simulation program (Gómez-Hernández and Journel, 1993). The resulting reference log-conductivity field is shown in Figure 4.1.

Table 4.1. Parameters of the random functions used to generate the ln$K$ realizations. Spherical variogram with anisotropic spatial correlation defined by λ_{max} and λ_{min}, which are the ranges in the maximum and minimum directions of continuity. The angle corresponds to the maximum continuity direction and it is measured clockwise from the North direction

	Mean	Std. dev.	Variogram	λ_{max}	λ_{min}	Angle
ln$K$	-1	1	Spherical	300	200	135

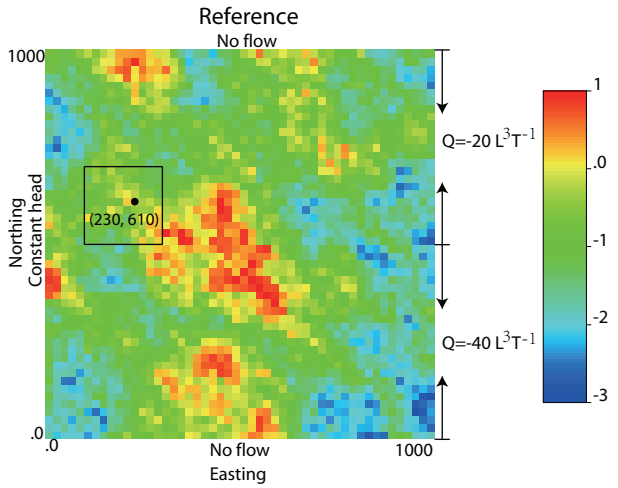

Figure 4.1. Reference ln$K$ and boundary conditions. The source location is marked with a dark dot. The inner square indicates the suspect contaminant source.

The model boundaries, as indicated in Figure 4.1, are set as follows: north and south boundaries are impermeable; west boundary is a prescribed head condition with a constant value of 50 [L]; east boundary is a prescribed flow boundary divided into two equal-length segments: the north segment with a total prescribed flow extraction rate of 20 [L^3T^{-1}] and the south segment with a total extraction prescribed flow rate of 40 [L^3T^{-1}]. Figure 4.2 shows the location of the 25 observation wells (red triangles) and the two control wells (blue diamonds).

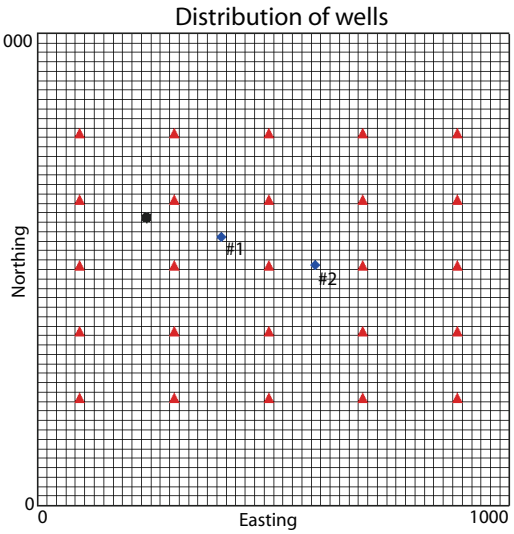

Figure 4.2. Location of wells. Red triangles mark observation wells; blue diamonds mark verification wells. The black circle is the contaminant source location.

The initial concentration is zero [ML^{-3}] and the initial head for the whole domain is 58 [L], except the west constant boundary. Other groundwater flow and contaminant transport parameters are assumed known and set as homogeneous: porosity of 0.3 [−], longitudinal dispersivity of 2 [L], transverse to longitudinal dispersivity ratio of 0.1.

We assume the contaminants are inert. Only advection and dispersion are considered as transport mechanisms. Both groundwater flow and contaminant transport are under transient conditions. The groundwater flow simulator MODFLOW (McDonald and Harbaugh, 1988) and the transport simulator MT3DMS (e.g., Zheng, 2010; Ma et al., 2012) are used as forward models to solve the groundwater flow and contaminant transport problems, respectively.

The total simulation time is 10000 [T] and is discretized into 100 time steps with increasing size following a geometric series with ratio 1.01 (The first time step is 58.66 [T]). The observations of both piezometric head and

concentration from the first 60 time steps (around 4790 [T]) are assimilated for the purpose of parameter identification.

The contaminant is released at location $(X, Y) = (230, 610)$ [L] with a mass-loading rate of 1000 [MT^{-1}], starting at time 613 [T] (around the 10th time step) and ending at time 2867 [T] (around the 40th time step), with a release duration of 2254 [T].

Figure 4.3 shows three snapshots of piezometric head and solute concentration taken on the reference aquifer at the 10th simulation time step (beginning of contaminant injection), 40th time step (end of contaminant injection), and at 60th time step (end of assimilation period). This figure also shows the location where both piezometric heads and concentration are sampled for the purpose of their assimilation in the different scenarios described next.

Seven scenarios will be evaluated. The first one, used as a benchmark to evaluate the efficiency of the ES-MDA is the r-EnKF, which has already proven its ability for the identification of contaminant source parameters and hydraulic conductivity characterization; it will be referred to as S0. The second one is the ES in its original implementation, that is, without any iteration. Then, to evaluate the effect of the number of iterations for the identification, the ES-MDA is run for five different scenarios, the difference between them is the number of iterations (or data assimilations) performed; they will be labeled S2 to S6 with 2, 4, 6, 8 and 10 iterations, respectively. Notice that the observation error inflation coefficients a_i will, in all cases, be equal to the number of iterations, in terms of the work by Emerick and Reynolds (2013a) in which the use of decreasing inflation coefficients leads to only small improvements with respect to using the inflation coefficients equal to the number of data assimilations.

An ensemble of 400 initial log-conductivity realizations is generated using the same random function model and parameters as for the reference log-conductivity field. Notice that there are no conditioning log-conductivity data, thus the ensemble mean and ensemble variance of the initial log-conductivity realizations are flat and equal to their marginal values. As already discussed by Xu et al. (2013a) the use of the same random function parameters for the generation of the initial realizations as for the generation of the reference case is only a marginal advantage given that there are no conditioning conductivities. Indeed, Xu et al. (2013a) demonstrate the effectiveness of the r-EnKF using a totally uninformative prior random function for the generation of the initial ensemble with similar results as when the "true" random function is used. Also, an ensemble of 400 5-tuplets for the source parameters is generated, each 5-tuplet contains five values drawn independently from the following uniform distributions: initial release time $T \in \mathcal{U}[550, 750]$, release duration $\Delta T \in \mathcal{U}[2100, 2300]$, mass-loading rate $M \in \mathcal{U}[900, 1100]$, and source location $(X, Y) \in (\mathcal{U}[100, 300] \times \mathcal{U}[500, 700])$.

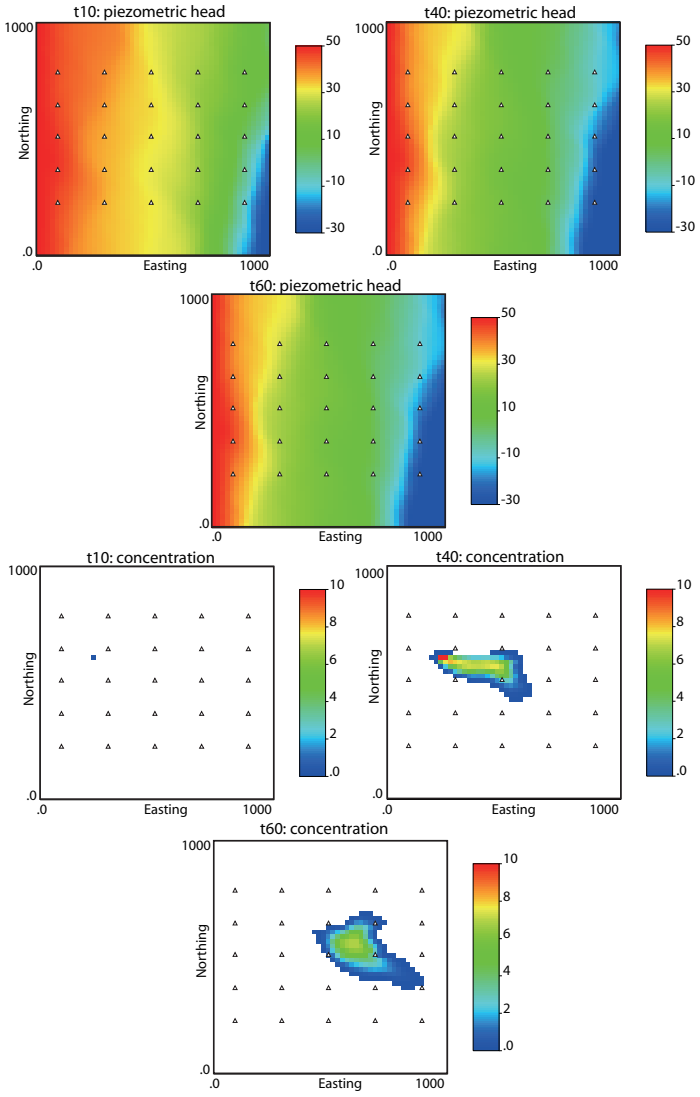

Figure 4.3. Reference. Piezometric head (top row) and contaminant plume (bottom row) at the 10th (beginning of solute injection), 40th (end of solute injection), and 60th (end of assimilation) time steps in the reference aquifer. White triangles mark the observation wells.

4.4 Results

Before starting the analysis of the results, Table 4.2 shows the CPU consumption for all scenarios. Recall that in the r-EnKF (S0) there are 60 assimilation plus updating steps, and 60 runs to solve the state equations, each run strating from time zero up to the assimilation step, but in each updating step only 25 observations are assimilated; whereas, in the ES-MDA the number of assimilation plus updating steps and of model runs is equal to the number of data assimilations, but in each updating step 1500 observations (25 observation locations times 60 time steps) are assimilated at once. For the current model and setup, the ES-MDA is cheaper to run than the r-EnKF up until data are assimilated four times, with an ES-MDA cost higher than two and half times that of the r-EnKF when data are assimilated ten times.

The r-EnKF, the ES and the ES-MDA will be used to assimilate the piezometric head and concentration data at the 25 observation locations. This assimilation will result in an ensemble of updated parameters (for the spatial distribution of hydraulic conductivity and for the parameters defining the contaminant source) that are used to produce an ensemble of piezometric heads and concentrations past the assimilation period (60th time step) for 40 time steps more. The performance of the different scenarios will be evaluated by comparing the different final ensembles to their corresponding counterparts in the reference aquifer.

Figure 4.4 shows the ensemble average and the ensemble variance of the updated logconductivities for scenarios S0 to S3 and S6, and the corresponding maps for S4 and S5 for this and following figures are shown in the appendix of this chapter. The ensemble mean shows how the main patterns of variability of the reference are captured by the updated ensemble, and the ensemble variance shows the local variability of the updated logconductivities. From a purely qualitative point of view it is clear that the r-EnKF does a good job in capturing the reference patterns with a small local uncertainty where the ensemble variance is close to zero, that the ES is able to extract patterns which are, overall, similar to the reference but still far from them and with a substantial local variability, and that the ES-MDA gets better the more times data are assimilated, with scenario S6 —for which data are assimilated 10 times —giving the best results.

The above analysis can be quantified by computing the average absolute bias (AAB) and the ensemble spread (ESp). The AAB is used to measure the average absolute deviation between the updated values and the reference ones. The ESp measures the precision of the ensemble of updated realizations by calculating the root square of the ensemble variance. Their

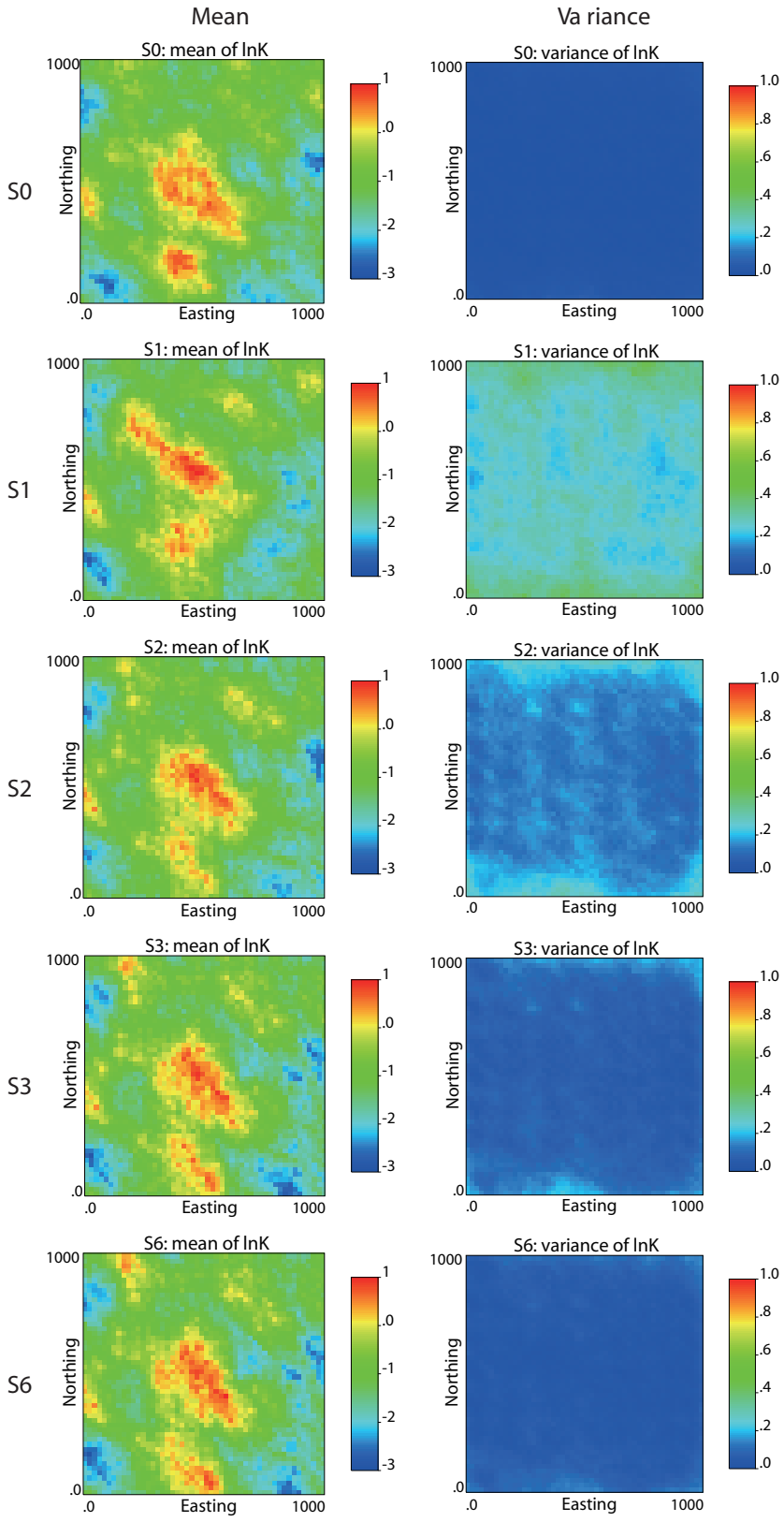

Figure 4.4. Scenarios S0-S3 and S6. Ensemble mean (left column) and ensemble variance (right column) of updated log-conductivity realizations.

Table 4.2. Definition of scenarios and CPU time consumption. The number in parenthesis refers to the number of data assimilation steps used in the ES-MDA. (ES would be equivalent to ES-MDA(1))

Method	Scenario	CPU in s	CPU in % of S0
restart EnKF	S0	16366	100%
ES	S1	4981	30%
ES-MDA(2)	S2	9526	58%
ES-MDA(4)	S3	17937	110%
ES-MDA(6)	S4	27432	149%
ES-MDA(8)	S5	34936	210%
ES-MDA(10)	S6	42422	259%

expressions are the following:

$$\mathrm{AAB} = \frac{1}{N_e}\sum_{i=1}^{N_e}\frac{1}{N_r}\sum_{j=1}^{N_r}|\ln K_{i,j} - \ln K_{i,ref}|, \tag{4.12}$$

$$\mathrm{ESp} = \sqrt{\frac{1}{N_e}\sum_{i=1}^{N_e}\sigma_i^2}, \tag{4.13}$$

where N_e is the number of model elements, N_r is the number of realizations, $\ln K_{i,ref}$ is the reference logconductivity value at node i, $\ln K_{i,j}$ is the logconductivity at node i for realization j and σ_i is the logconductivity ensemble variance at node i.

Figure 4.5 shows the AAB and ESp of lnK and of the parameters defining the contaminant source for all scenarios computed before any data assimilation and after data assimilation over the first 60 time steps. The values are high when comparing the initial ensemble with the reference (no assimilation has been performed yet), the AAB and ESp is reduced considerable for the r-EnKF except for that of ΔT and M, and the values for the smoother keeps decreasing with the number of assimilation steps. Specifically, the AAB and ESp of updated lnK, and Y of scenarios S3-S6 is close to that of scenario S0, and the AAB and ESp of updated T of scenario S6 is close to that of scenario S0, while, the AAB and ESp of updated X, ΔT and M of scenarios S3-S6 is smaller than that of scenario S0, namely that, after 4 times assimilation steps, the ES-MDA starts to perform better than the r-EnKF.

Figure 4.6 shows the piezometric head distribution at the 60th time step computed with the final updated parameters for scenarios S0 to S3 and S6. The maps show, in the left column, the piezometric head distributions for an individual ensemble member (realization #300), in the center column,

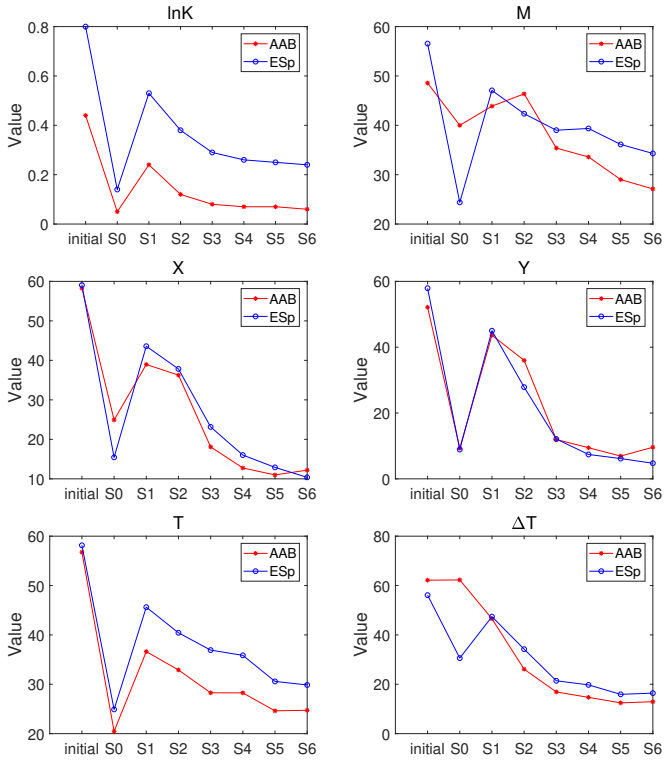

Figure 4.5. Scenarios S0-S6. Average absolute bias (AAB) and ensemble spread (ESp) of updated log-conductivity realizations (lnK), the source location (X and Y), initial release time (T), release duration (ΔT), and mass-loading rate (M) computed with the initial parameters and with the updated parameters after 60 time steps.

the ensemble average obtained as the local mean of the piezometric head at each node through the 400 realizations, and in the right column the ensemble variance. Please, notice that the middle row with the ensemble mean piezometric heads is not the solution of the state equations in the ensemble log-conductivity average of Figure 4.4. An analysis of these maps shows the robustness of the r-EnKF (S0) that produces an ensemble average map quite close to the reference one (upper right corner in Figure 4.3) and with little variability everywhere. The smoother performs well when comparing the average ensemble with the reference map, but the uncertainties associated are quite large, especially in scenarios S1 and S2 (with only one or two data assimilations); there is a need to assimilate at least four times the data (S3) to get a variance reduction that approximates that of the r-EnKF.

Figure 4.7 shows the concentration plume computed with the parameters updated using observations at 60 time steps. In the left column, the plume in realization #300, in the center, the ensemble average of the 400 plumes computed in the 400 realizations with updated parameters, and in the right column the local concentration variance computed at each node through the ensemble of realizations. Please, notice that, as with piezometric heads, the middle row with the ensemble average concentrations is not the solution of the state equations in the ensemble logconductivity average of Figure 4.4. An analysis of these maps reaches the same conclusions as for the piezometric heads, the r-EnKF is quite robust producing an ensemble average plume quite close to the reference (lower right corner in Figure 4.3) and with lower variability. The smoother performs well only when the number of assimilations is large (S3 and S6); for the cases of 1, and 2 assimilations (S1 and S2, respectively), the ensemble mean plume is quite spread, the local variance is large, and the plume in the single selected realization shown in the left column of the figure can be quite far from the reference one.

Figure 4.8 shows the time evolution of piezometric heads and solute concentrations at the two verification wells (#1 and #2, see Figure 4.2) computed using the initial ensembles of contaminant source parameters and logconductivities. The spread of predicted values is quite large since no observation has been assimilated yet. Figure 4.9 and 4.10 show the time evolution of piezometric heads and solute concentrations computed with the updated source parameters and logconductivity fields after the assimilation of the observations during the first 60 time steps, respectively. The spread of the curves after the assimilation is considerably reduced, especially for scenarios S0, S3 and S6. Although these two wells were not used during the assimilation, the reproduction of piezometric heads, even after the assimilation period ends is very good both for the r-EnKF (S0) and for the ES-MDA with 4 and 10 assimilations (S3 and S6), with the former performing slightly better than the latter.

Up to here, regarding the characterization of the logconductivity field and the reproduction of the state variables, the r-EnKF seems to outperform

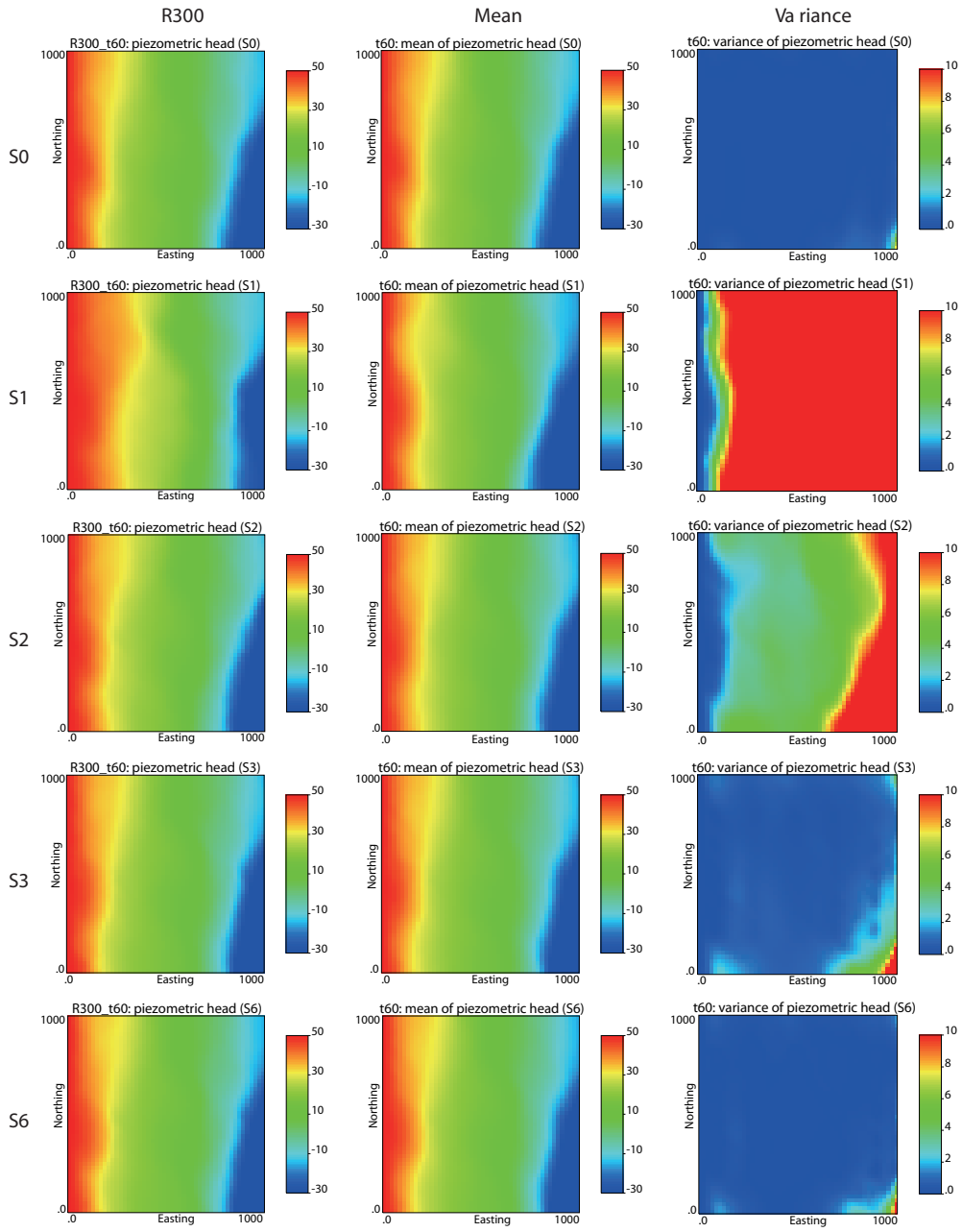

Figure 4.6. Scenarios S0-S3 and S6.$_1$ Piezometric heads as computed with the updated parameters at the end of the 60th time step. From left to right, heads in realization #300; ensemble mean, and ensemble variance.

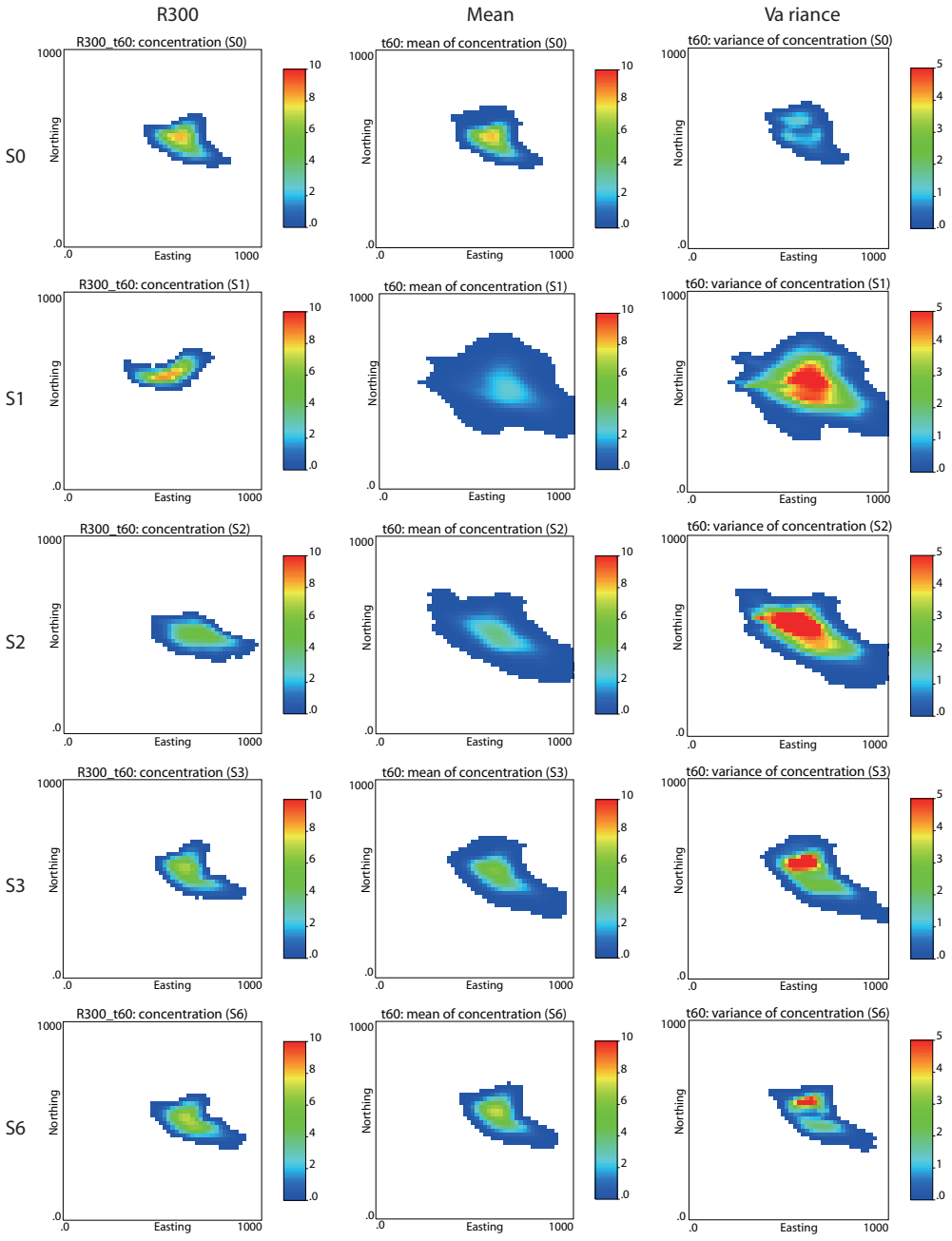

Figure 4.7. Scenarios S0-S3 and S6.[1] Contaminant plume as computed with the updated parameters at the end of the 60th time step. From left to right, Contaminant plume in realization #300; ensemble mean of all contaminant plumes, and ensemble variance of all contaminant plumes.

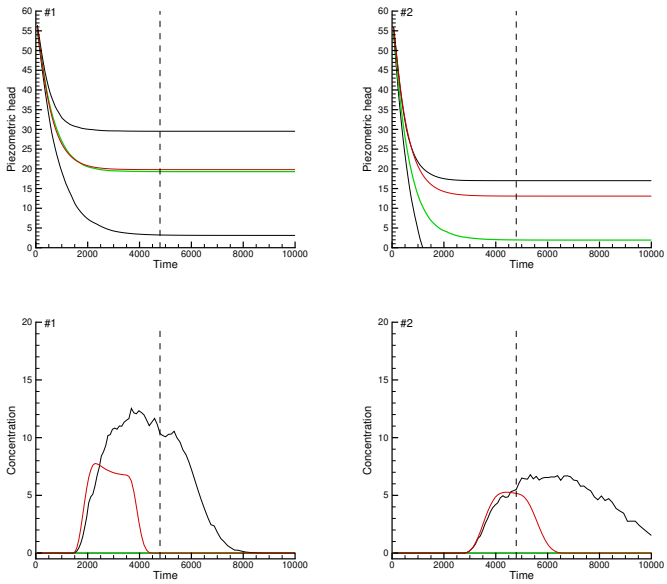

Figure 4.8. Time evolution of piezometric heads (top row) and solute concentrations (bottom row) at the two verification wells #1, and #2 computed on the initial ensemble of source information parameters and $\ln K$. The red line corresponds to the the reference field. The black lines correspond to the 5 and 95 percentiles of all realizations, and the green line corresponds to the median. The vertical dashed lines mark the end of the assimilation period.

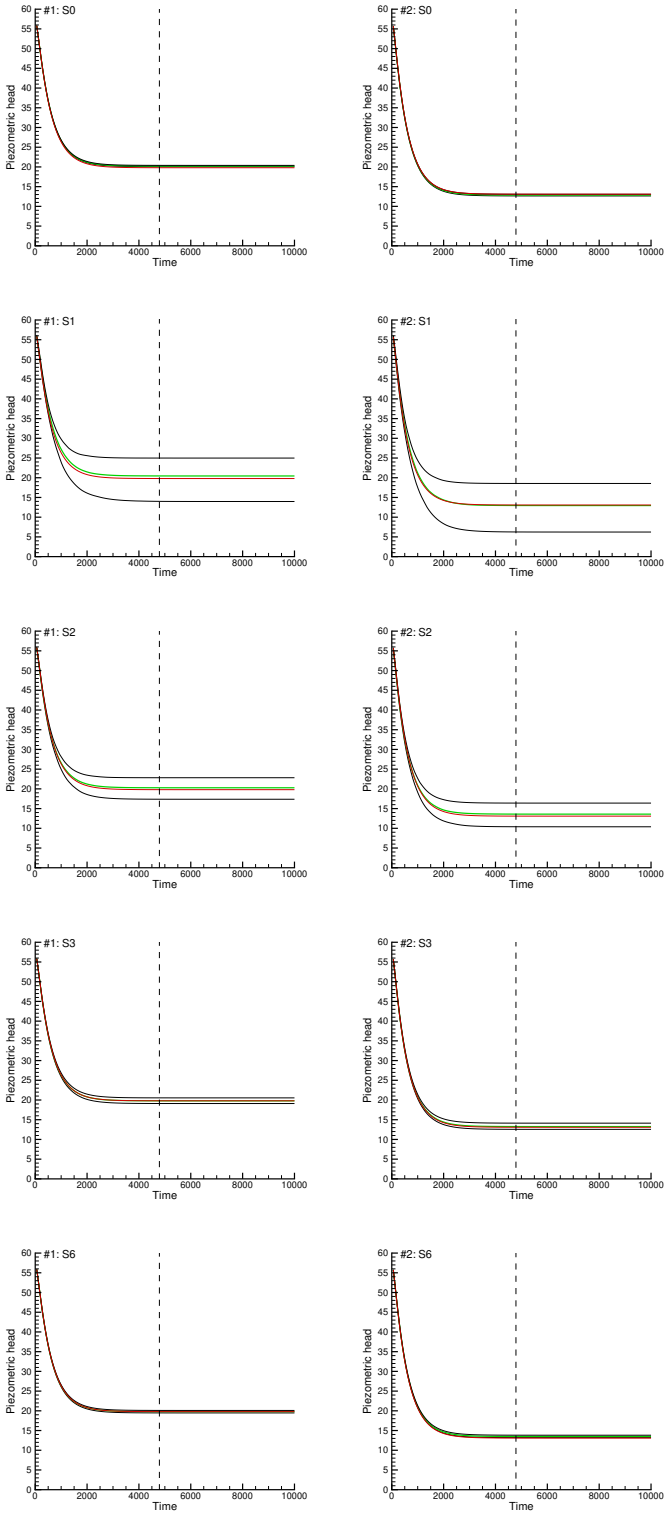

Figure 4.9. Scenarios S0-S3 and S6. Time evolution of the piezometric heads at the two verification wells #1, and #2 computed with the updated ensemble of source information parameters and $\ln K$ after the assimilation of the observations of the first 60 time steps. The red line is the evolution of the piezometric head in the reference. The black lines correspond to the 5 and 95 percentiles of all realizations, and the green line corresponds to the median. The vertical dashed lines mark the end of the assimilation period.

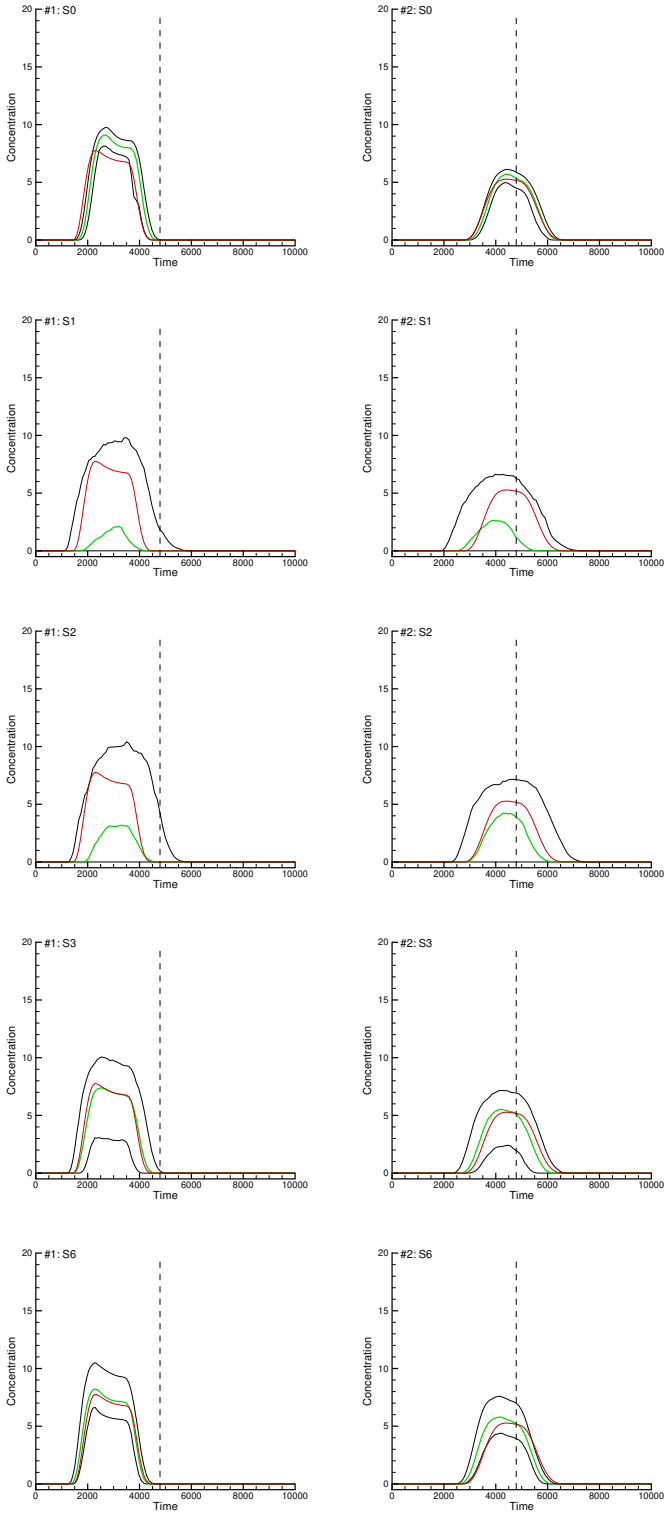

Figure 4.10. Scenarios S0-S3 and S6. Time evolution of the solute observations at the two verification wells #1, and #2 computed with the updated ensemble of source information parameters and $\ln K$ after the assimilation of the solute observations of the first 60 time steps. The red line is the evolution of the concentration in the reference. The black lines correspond to the 5 and 95 percentiles of all realizations, and the green line corresponds to the median. The vertical dashed lines mark the end of the assimilation period.

the ES-MDA. The AAB($\ln K$) and ESp($\ln K$) are the smallest for S0 (r-EnKF), and the piezometric head and concentration predictions are also the best for S0. Only the ES-MDA with 10 assimilation steps (S6) gives comparable results, although at a CPU cost 2.6 times larger than the r-EnKF.

However, when we analyze the reproduction of the contaminant source parameters, the ES-MDA is superior to the r-EnKF. Figure 4.5 shows the AAB and ESp for the contaminant source parameters computed with expressions similar to Eq. (4.12) and (4.13). An analysis of this figure shows how the AAB and ESp for the contaminant source parameters decrease with the number of assimilation steps, reaching the smallest values for the ES-MDA with 10 assimilations (S6). This observation is complemented by the results shown in Figure 4.11, in which boxplots of the initial ensemble of the source parameters and of the updated ensemble of the source parameters for the six scenarios are shown. Some observations that can be derived from this figure are: the r-EnKF (S0) produces good estimates for X, Y and T with a considerable reduction of uncertainty with respect to the initial ensemble, while the estimates for ΔT and M are somehow biased without a large reduction of uncertainty, which is due to that the r-EnKF needs to restart the simulation of the state equation from time zero without updating state variables, and a few of observation at an assimilation step, enlarges the possibility to reach the same "observations" by different combinations of ΔT and M; the ES (S1) is not effective, the spreads of the ensemble is almost the same as for the initial ensemble prior to assimilation for all parameters; the ES-MDA starts to work after 4 data assimilations by performing multiple smaller corrections in the ensemble, and gives the best results for 10 assimilations, and, in this case, better than for the r-EnKF, especially for parameters X, ΔT and M.

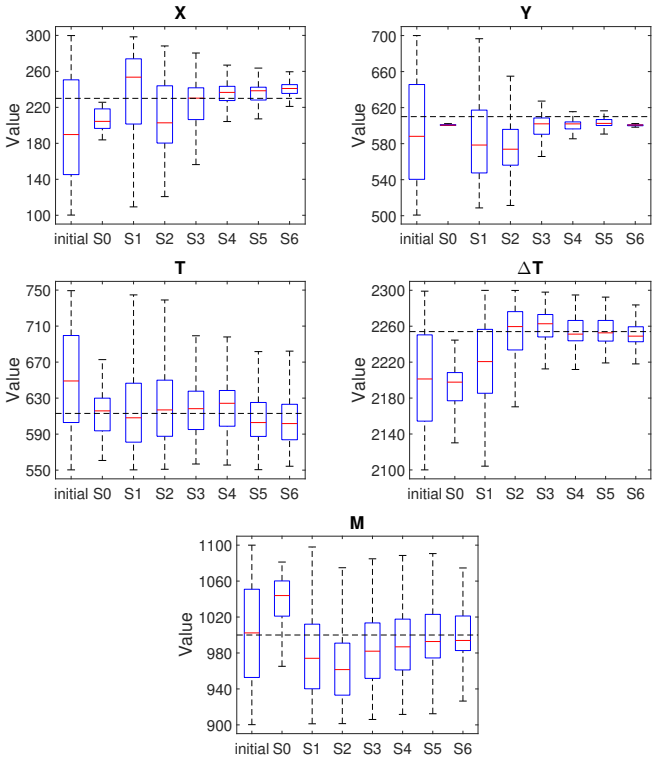

Figure 4.11. Scenarios S0-S6. Boxplots of the source location (X and Y), initial release time (T), release duration (ΔT), and mass-loading rate (M) computed with the initial parameters and with the updated parameters after 60 time steps. The dashed horizontal black line corresponds to the reference value.

4.5 Summary and Discussion

The purpose of this paper is to analyze the ability of the ES-MDA for the identification of contaminant source parameters together with a spatially heterogeneous hydraulic conductivity field in comparison with the r-EnKF. The results show that the ES-MDA has the ability to estimate hydraulic conductivity field and identify the contaminant source parameters —including source location, initial release time, release duration and mass-loading rate– with a proper number of data assimilation steps, besides, the results also indicate that these estimate parameters are good enough to provide good forecasts of solute concentrations and piezometric heads.

Furthermore, with the comparison of the performance of the source parameters identification between the r-EnKF and the ES-MDA (also including the performance of scenarios S4 and S5 in the appendix of this chapter), the ES-MDA proves it performs better than the r-EnKF, especially for the identification of contaminant source parameters when using enough number of data assimilation steps. For this specific test done here, the ES-MDA starts to outperform the r-EnKF after 4 time assimilation steps, and spends almost the same computational cost as that for r-EnKF at the 4th assimilation step. Plus, the ES-MDA can perform even better, at a cost of more computational time. Specifically, for the r-Enkf, the updated mass-loading and release duration still with large uncertainty is attributed to far less observations employed at an assimilation step compared with that of the ES-MDA, which result in a larger possibility to reach the same "observations" by different combination of mass-loading and release duration.

4.6 Appendix. Results of scenarios S4 and S5

Results for scenarios S4 and S5 are displayed in Figures 4.12 to 4.15. The details are as follows: Figure 4.12 shows the ensemble mean and ensemble variance of the updated $\ln K$; Figures 4.13 and 4.14 show the 300th realization, ensemble mean and ensemble variance of piezometric heads and of the contaminant plume at the end of the 60th time step, respectively; Figures 4.15 and 4.16 show the time evolution of the piezometric heads and of solute concentrations at the two verification wells #1, and #2 computed with the updated source parameters and hydraulic conductivities.

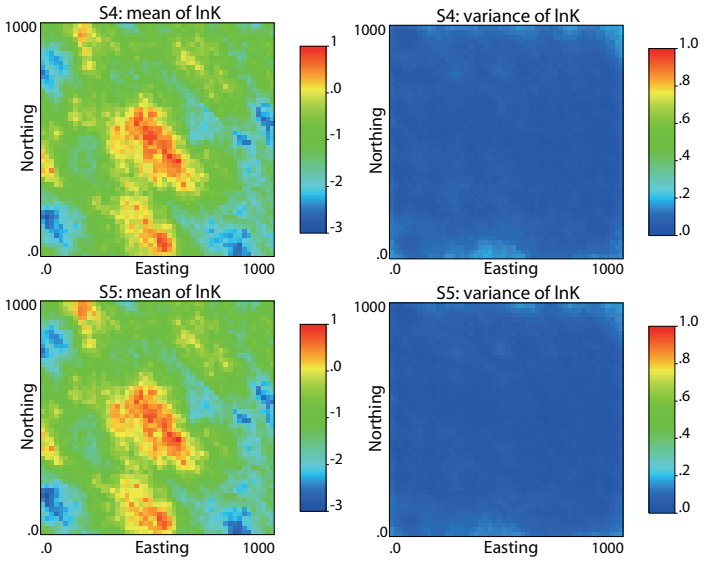

Figure 4.12. Scenarios S3-S4. Ensemble mean (left column) and ensemble variance (right column) of updated log-conductivity realizations. (This figure complements Figure 4.4.)

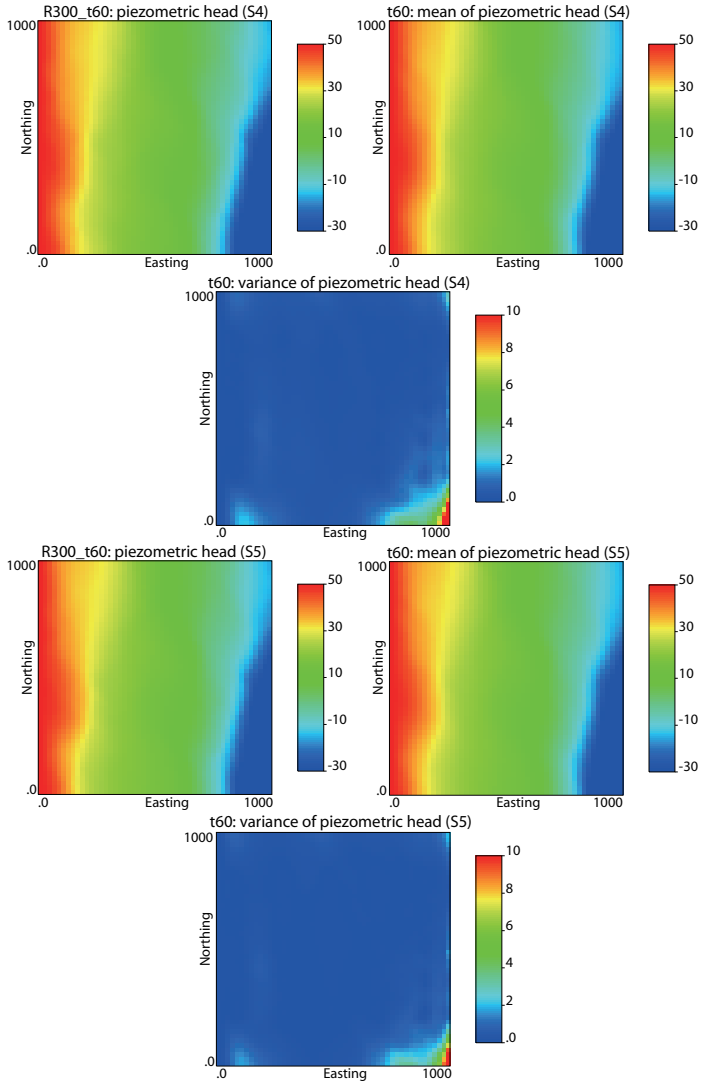

Figure 4.13. Scenarios S4-S5. Piezometric heads computed with the updated parameters at the end of the 60th time step. From left to right, heads in realization #300; ensemble mean, and ensemble variance. (This figure complements Figure 4.6.)

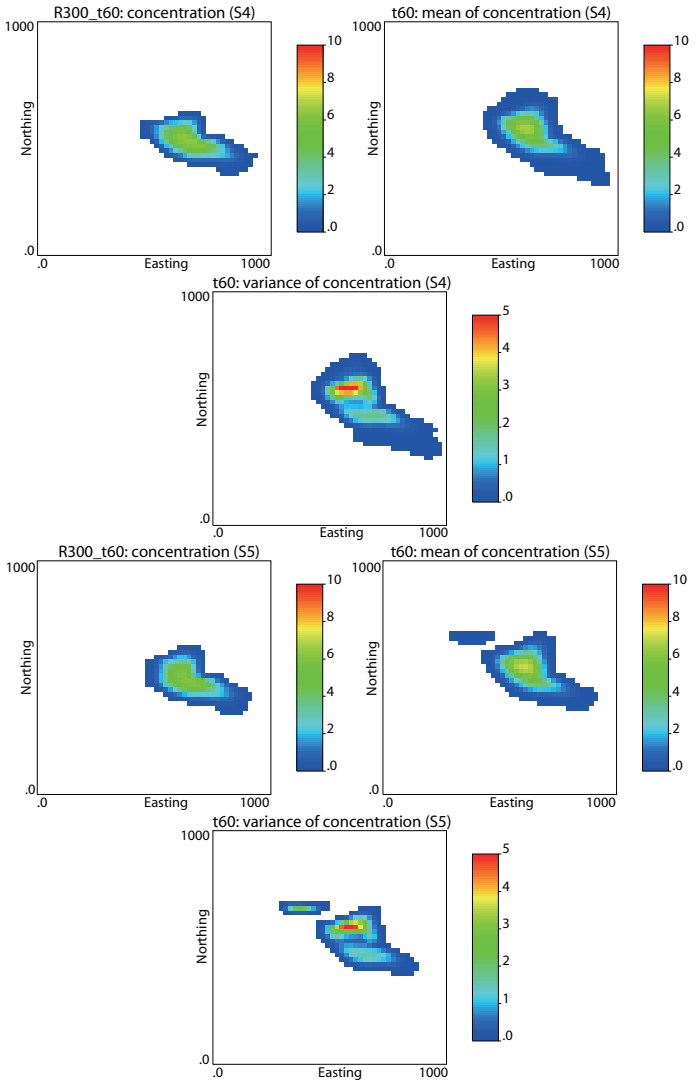

Figure 4.14. Scenarios S4-S5. Contaminant plume computed with the updated parameters at the end of the 60th time step. From left to right, Contaminant plume in realization #300; ensemble mean, and ensemble variance. (This figure complements Figure 4.7.)

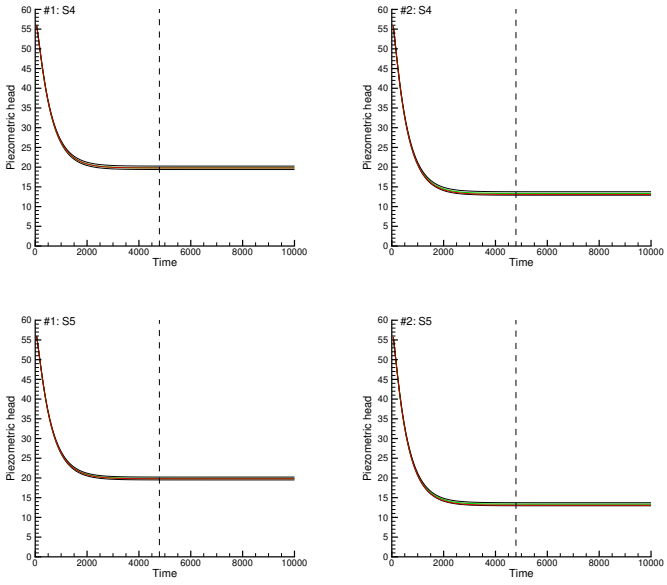

Figure 4.15. Scenarios S4-S5. Time evolution of the piezometric heads at the two verification wells #1, and #2 computed with the updated ensemble of source information parameters at the end of the 60th time step. The red line is the evolution of the piezometric head in the reference. The black lines correspond to the 5 and 95 percentiles of all realizations, and the green line corresponds to the median. The vertical dashed lines mark the end of the assimilation period. (This figure complements Figure 4.9.)

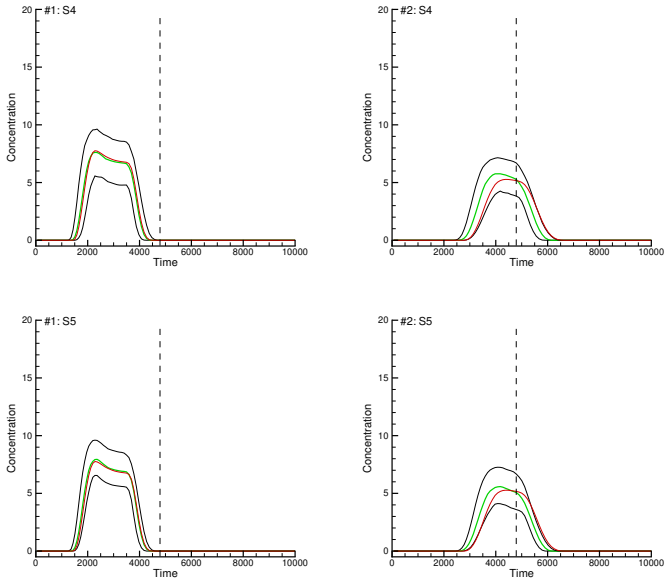

Figure 4.16. Scenarios S4-S5. Time evolution of the solute concentrations at the two verification wells #1, and #2 computed with the updated ensemble of source information parameters at the end of the 60th time step. The red line is the evolution of the solute concentration in the reference. The black lines correspond to the 5 and 95 percentiles of all realizations, and the green line corresponds to the median. The vertical dashed lines mark the end of the assimilation period. (This figure complements Figure 4.10.)

5

Reconstructing the release history of a contaminant source via Ensemble smoother with multiply data assimilation

Abstract

Identifying a contaminant time-varying release history is an ill-posed problem but crucial for groundwater contamination issues. In this paper, a recently 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, the use of covariance inflation, or the number of iteration used in the ES-MDA for the purpose of identifying a time-varying contaminant injection event discretized in several time steps. 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. As expected, 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. The application of the method to two sandbox experiments shows

that the ES-MDA has the ability to recover finite release injections, but performs disappointingly when the injection is continuous.

5.1 Introduction

Groundwater contamination has gained extensive attention over the last several decades since it is becoming a huge threat to our ecosystem. Determinig 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 and Kabala, 1994; Carrera and Neuman, 1986). Various methods have been devised to address this problem and several reviews have been published in the subject (Atmadja and Bagtzoglou, 2001b; Michalak and Kitanidis, 2004; Bagtzoglou and Atmadja, 2005; Sun et al., 2006a, e.g.,).

Among all these methods, one branch, data assimilation methods, comes out ahead because of their 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 (Houtekamer and Mitchell, 2001; Bertino et al., 2003; Chen and Zhang, 2006; Aanonsen et al., 2009, e.g.,). Specifically, in hydrogeology, the EnKF method has been proven able to invert aquifer parameters, such as hydraulic conductivity (Chen and Zhang, 2006; Huang et al., 2009; Kurtz et al., 2014), porosity (Li et al., 2012a), recharge rates (Franssen and Kinzelbach, 2009) and also boundary conditions (Chen and Zhang, 2006). More recently, researchers have started to employ EnKF variants to identify the parameters describing a contaminant source in aquifers. Xu and Gómez-Hernández (2016b) 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 and Jaime, 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 restart Ensemble Kalman filter. Chen et al. (2018) also discusses 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 filter combined with a mixed-integer nonlinear programming optimization model to deduce the accurate location and release history of a contaminant source. The aforementioned work are strong demonstrations that the EnKF and its variants are valid methods for contaminant source identification. However, the release history identified in these works only focus on a constant pulse, the magnitude of which is independent of time.

As an alternative to the EnKF, the ensemble smoother (ES), which was firstly introduced by van Leeuwen and Evensen (1996), assimilate 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 Bayesian formulation (Evensen, 2004). However, in studying process with strong nonlinearities, such as is the case of inverting the groundwater flow and mass transport equations, the EnKF outperformed the ES (Evensen and van Leeuwen, 2000), until an iterative variant of the ES was proposed, the ES with Multiple Data Assimilation (ES-MDA), by Emerick and Reynolds (2013a). Evensen (2018) compared the ES-MDA with other iterative ensemble smoothers to solve history matching problems. Ranazzi and 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 of the reverse flow routing problem. These works are all good examples about ES-MDA dealing with time-varying input parameters.

In this work, the ES-MDA is employed for the first time to identify a time-varying release history in both synthetic and real cases. The influence of observation data frequency is discussed in relation with the precision with which the release history can be identified. Also, two kinds of covariance inflaction procedures (e.g., Le et al., 2016; Rafiee and Reynolds, 2017) are analyzed, one predefined and the other one adaptative. In section 2, we describe the methodology; in section 3, the synthetic and the real sandbox experiment are presented, following by the setup of different scenarios and evaluation criteria. Finally, in section 4, we discuss the results and draw some conclusions.

5.2 Methodology

5.2.1 Groundwater flow and solute transport equations

The water flow and solute transport theory in a porous media under a Cartesian coordinate system was already introduced in chapter 2.2.1, but they are

repeated here for a matter of completeness (Bear, 1972; Zheng and Wang, 1999),

$$S_s \frac{\partial h}{\partial t} = \nabla \cdot (K \nabla h) + w, \tag{5.1}$$

$$\frac{\partial (\theta C)}{\partial t} = \nabla \cdot (\theta D \cdot \nabla C) - \nabla \cdot (\theta v C) - q_s C_s, \tag{5.2}$$

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 $[T^{-1}]$. In the transport governing equation, 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 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 $[ML^{-3}]$.

5.2.2 Ensemble Smoother with Multiple Data Assimilation(ES-MDA)

The ES-MDA was first introduced by Emerick and Reynolds (2013a) as an improvement of the ES for handling nonlinear models. The ES-MDA updates the model parameters using the same set of observations with a predefined number of iterations, it is easy to understand and to implement (Evensen, 2018). A brief introduction of ES-MDA has already been presented in chapter 4.2.2, here we will give a more detailed description. The ES-MDA algorithm can be described as follows:

1. Initialization.

The first step is to generate N_e realizations of n parameters. This will be the initial ensemble of realizations. There are several approaches to generate this values, in this case, we have chosen to draw the numbers from predifined uniform distributions. The second step is to choose the number of iterations (also referred to as assimilation steps), N_a, and the inflation factor α_j. How to choose the inflation factor is explained in detail below.

2. Assimilation.

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

a. Forecast step

In this step, the groundwater flow and contaminant transport models, MODFLOW and MT3DS, are run for each member of the ensemble; in our case, for each different release history,

$$B_{i,j}^f = \psi[B_0, A_{i,j}], \tag{5.3}$$

for i=1, 2, ..., N_e, and j=1, 2, ..., N_a, where B^f stands for the vector of forecasted concentrations; ψ represents the forward numerical model; A is the vector of source release history and its size is determined by the number of discretization steps chosen to describe the injection curve.

b. Update step

Then, the model parameters are updated as follows,

$$A_{i,j+1} = A_{i,j} + \Delta A_j (\Delta B_j^f)^T [\Delta B_j^f (\Delta B_j^f)^T + \alpha_j R]^{-1} [y_{obs} + \sqrt{\alpha_j}\varepsilon - B_{o,i,j}^f], \tag{5.4}$$

where y_{obs} is a column vector with dimensions $N_o \cdot N_t$ dimensional of real measurements (N_o is the number of locations, and N_t the number of observation time steps); ε stands for the observation error, while R is the covariance matrix of the observation error; $B_{o,i,j}^f$ stands for the vector of forecasted concentrations at the same locations and times where and when observations y_{obs} are made; ΔA_j and ΔB_j are square root matrices defined as

$$\Delta A_j = \frac{1}{\sqrt{N_e - 1}} [A_{1,j} - \overline{A}_j, A_{2,j} - \overline{A}_j, \ldots, A_{N_e,j} - \overline{A}_j], \tag{5.5}$$

$$\Delta B_j^f = \frac{1}{\sqrt{N_e - 1}} [B_{1,j}^f - \overline{B^f}_j, B_{2,j}^f - \overline{B^f}_j, \ldots, B_{N_e,j}^f - \overline{B^f}_j], \tag{5.6}$$

where $\overline{A}_j$ and $\overline{B^f}_j$ are the ensemble means of source release history parameters and forecasted concentrations at the j_{th} iteration, 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 the measurements is larger than the ensemble size, it is necessary to employ the truncated singular value decomposition (TSVD) method to do a pseudo-inversing in Eq. (5.4).

5.2.3 Schemes for 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 Emerick and Reynolds (2013a) 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 factor α_j satisfies the following equation,

$$\sum_{j=1}^{N_a} \frac{1}{\alpha_j} = 1, \tag{5.7}$$

There are still many options on how to choose the α_j parameters satisfying the previous equation. Apparently, chosing a decreasing series may be the most appropriate, but some authors claim that using uniform values give similar results, and that choosing these values arbitrarily may lead to filter collaps (Le et al., 2016). We have decided to explore two methods to

select the inflation factors, one proposed by Rafiee and Reynolds (2017), and the other one proposed by Evensen (2018).

Rafiee and Reynolds (2017) define the inflation factor at the first assimilation step according to the discrepancy principle, as follows

$$\alpha_1 = (\frac{1}{N}\sum_{i=1}^{N}\lambda_i)^2, \tag{5.8}$$

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

$$D_j = R^{-\frac{1}{2}}\triangle B_j^f, \tag{5.9}$$

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

$$\alpha_j = \beta^{j-1}\alpha_1, \tag{5.10}$$

where β is the ratio that fulfills that the sum of the inverse of the inflation factors equals one and is given by

$$\frac{1-(1/\beta)^{N_a-1}}{1-1/\beta} = \alpha_1, \tag{5.11}$$

Evensen (2018) defines a scheme simply selecting a nonzero value $\alpha_1^{'}$ and a geometrical ratio α_{geo}; with these two numbers defines a sequence

$$\alpha_{j+1}^{'} = \frac{\alpha_j^{'}}{\alpha_{geo}}, \tag{5.12}$$

which is then normalized to provide the α_j values that satisfy equation 5.7

$$\alpha_j = \alpha_j^{'}(\sum_{j=1}^{N_a}\frac{1}{\alpha_j^{'}}) \tag{5.13}$$

This scheme has the capacity of defining the inflation factor in either increasing or decreasing sequence by choosing an α_{geo} below or above one, respectively. Here, we define the α_{geo} and $\alpha_1^{'}$ with the value of 2 and 1, respectively.

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

5.3 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

Table 5.1. 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

has been employed in several groundwater contamination studies (Citarella et al., 2015; Cupola et al., 2015b; Zanini and Woodbury, 2016). In this work, first, we generated synthetic data with 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 numrerical model. Then, we tested the ES-MDA with real observation data and analyzed the impact of the observation error magnitude.

5.3.1 Sandbox Set-up

The lab set up is the same as in chapter 2.3.1, it is repeated here for a matter of completeness. 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 5.1. The reservoirs upstream and downstream are set up as constant piezometric boundary with a water level of 62.5 cm and 60.6 cm, respectively. The bottom of the sandbox is regarded as 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. There are 25 points in total where contaminant concentration is observed for the purpose of the application of the ES-MDA. The details about the acquisition of the concentration data could be found in Citarella et al. (2015); Cupola et al. (2015b). The total experiment time is 3000 s and the injection starts from the beginning. The main hydraulic parameters for simulation are listed in Table 5.1.

5.3.2 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

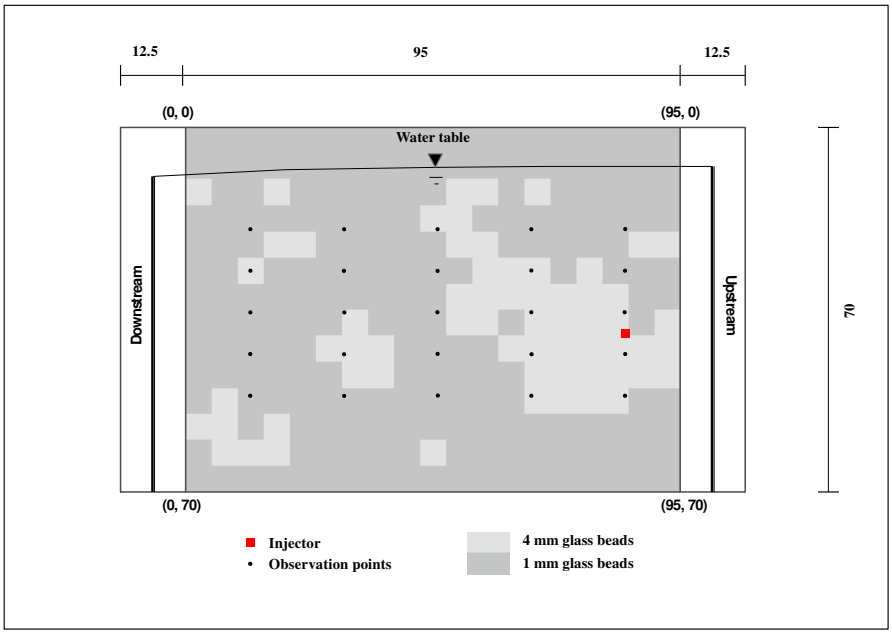

Figure 5.1. Sketch of the experimental device (lateral view). Length unit is cm.

RMSE:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(A^{ref} - \overline{A}_i)^2}, \tag{5.14}$$

$$\text{relative } RMSE = \frac{RMSE}{\text{intial } RMSE}, \tag{5.15}$$

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.

5.3.3 Synthetic Case

The first set of analyses are based on the synthetic simulation of a time-varying release into the numerical model that mimics the sandbox. The release function adopted is based on a proposal by Skaggs and Kabala (1994):

$$\begin{aligned} S(t) = 2.6 \cdot \exp(-\frac{(\frac{t}{10} - 20)^2}{50}) \\ + 0.78 \cdot \exp(-\frac{(\frac{t}{10} - 50)^2}{200}) \\ + 1.3 \cdot \exp(-\frac{(\frac{t}{10} - 90)^2}{98}) \qquad 0 \leq t \leq 3000 \end{aligned} \tag{5.16}$$

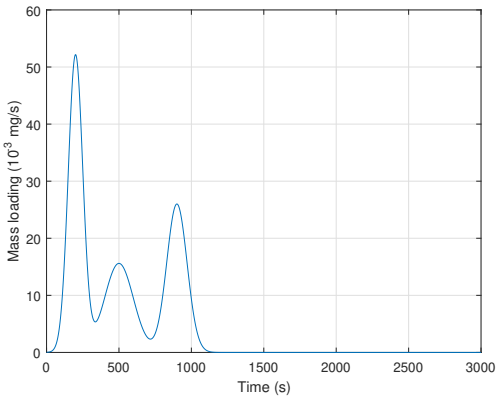

Figure 5.2. Release curve of a synthetic contaminant source.

This function is shown in Figure 5.2. For the purpose of identifying the source, we run three sets of scenarios as a function of how many steps we used to discretize the release history; we chose to divide the 3000 s duration of the experiment into 50, 100 and 300 time steps. Then for each discretization, two sampling frequency were considered, samples were taken every other time step or every ten time steps. Also, the number of 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 5.2. 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.

An ensemble of 500 realizations was used. The initial release history curve of every realization is generated using uniform distribution with ranges in U[0, 52] 10^{-3} mg/l. Figure 5.3 shows the recovered release history of the set of scenarios with 50 time steps. In each plot, the blue curve corresponds to the actual release history; the gray lines are the recovered release history curves for all 500 realizations, and they are summarized by their median (red dotted lines) and their 5 and 95 percentiles (black dashed lines). The first column uses Rafiee's inflation and the second column Evensen's inflation. The first two rows use samples every ten time steps, and the last two rows samples every other time step. The first and third row use four iterations and the second and fourth row 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 the Rafiee's inflation method always yields lesser realization spread that the Evensen's inflation method. It is

Table 5.2. Definition of the synthetic scenarios

Scenario	number of time steps	Observation frequency	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

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 5.4 shows the recovered release history of the set of scenarios with 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 is no longer able to capture the actual release history as precisely as in the previous set of realizations, more notably in the set of scenarios with samples every ten 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 (scenario S9 to S12). The fluctuations may be due to an ill-posedness of the problem and the fact that 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 so that the curves display a smoothness in time throughout the iterations. 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 sensitivity that most observations have to a release in the final stages 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 5.5 shows the results for scenarios S17 to S24, and their arrangement follows the same pattern as the previous two figures. The original release curves is only hinted 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 to be identified 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 5.3 and Figure 5.6 illustrates the RMSE and the relative RMSE of all 24 scenarios. Based on the RMSE at the final iteration step, we can draw the conclusion that the ES-MDA with Rafiee's scheme has a better performance in most scenarios in our case, especially when the observation data frequency is low. The iteration evolution of the relative RMSE of the last set of scenarios, the ones with a discretization of 300 time steps, may explain why these scenarios perform so

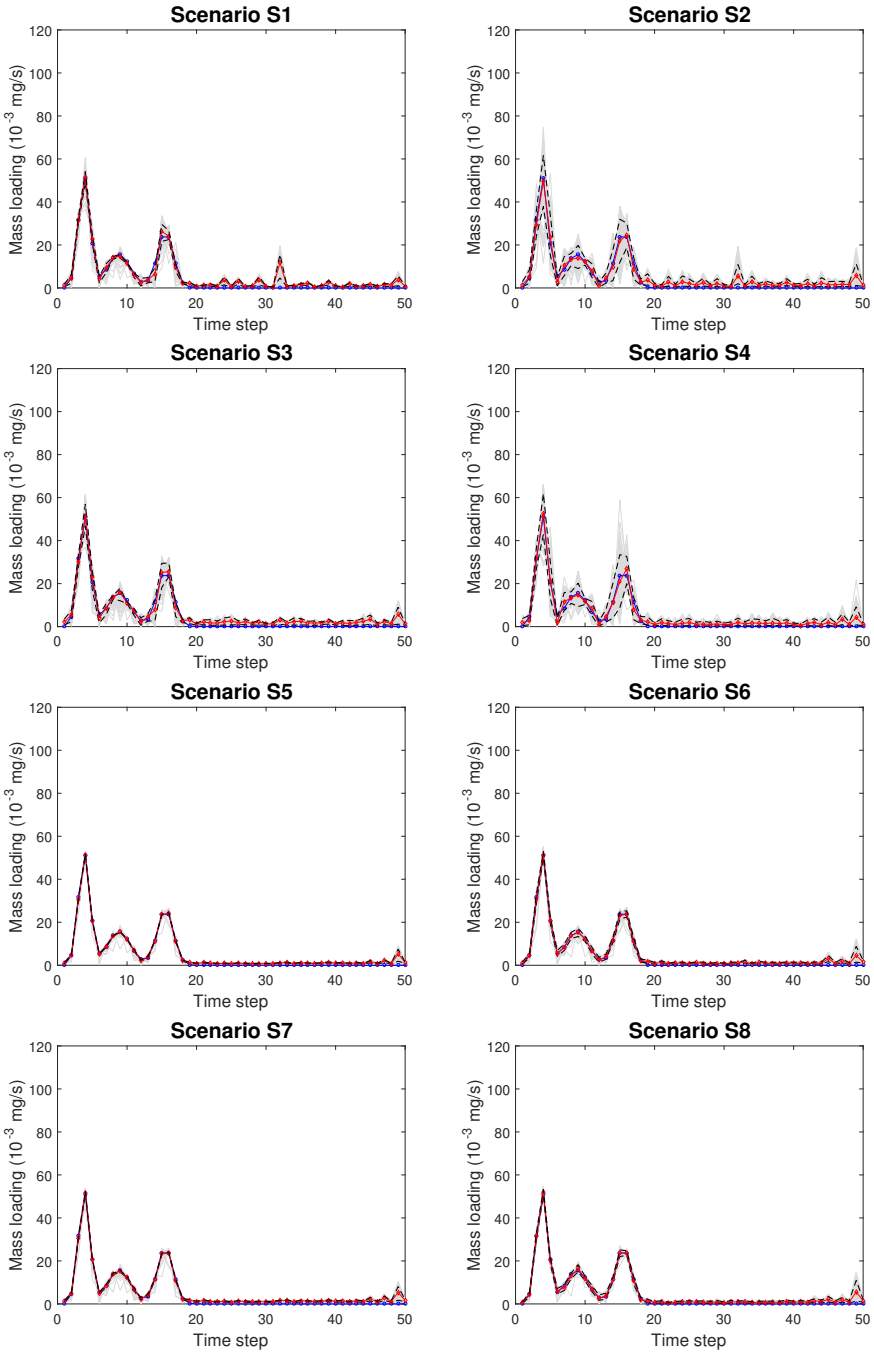

Figure 5.3. Recovered release histories for scenarios S1 to S8. The blue curve corresponds to the actual release history. The gray lines are the recovered release history curves for all 500 realizations, summarized by their median (red dotted lines) and their 5 and 95 percentiles (black dashed lines).

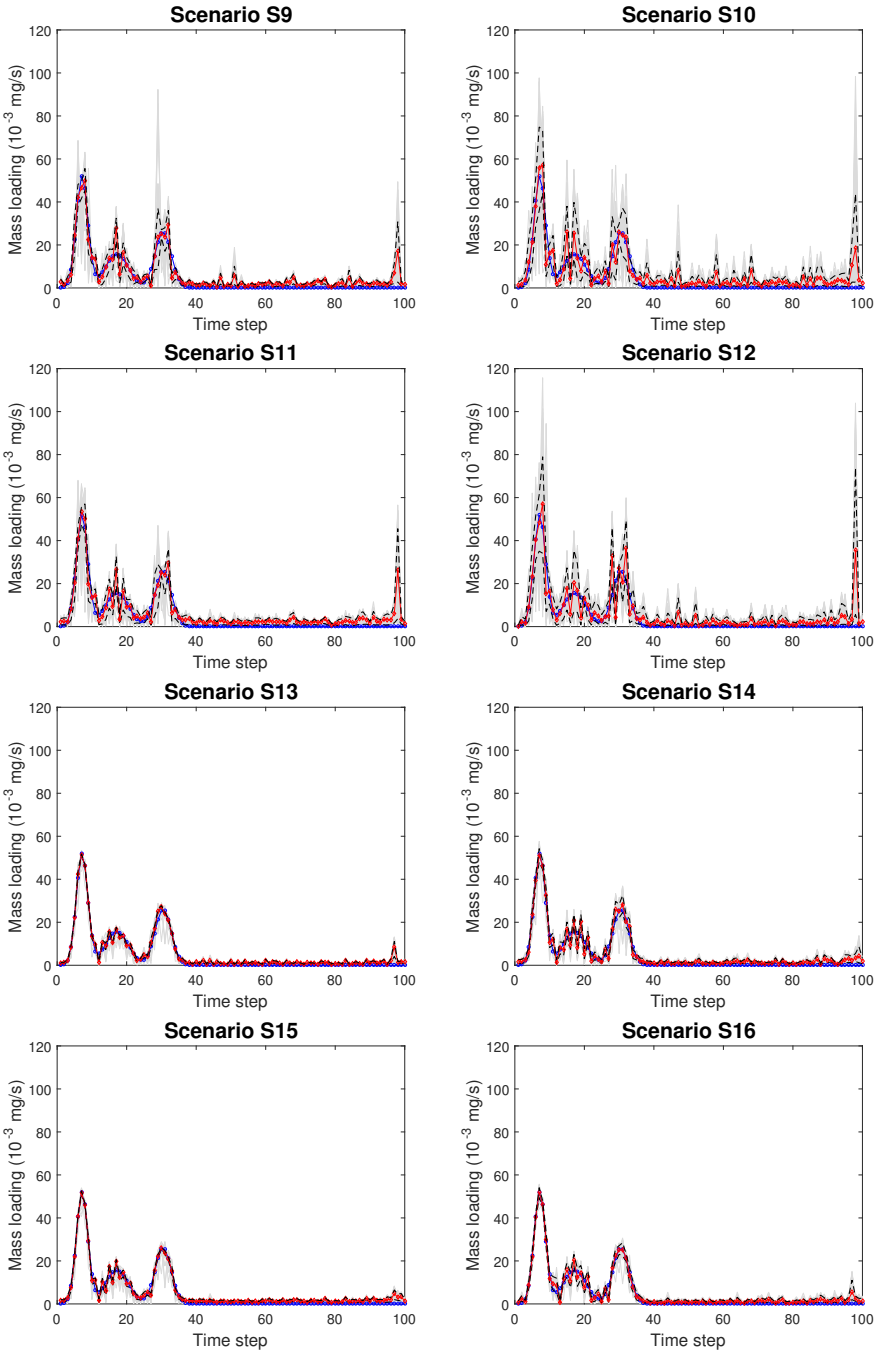

Figure 5.4. Recovered release histories for scenarios S9 to S16. The blue curve corresponds to the actual release history. The gray lines are the recovered release history curves for all 500 realizations, summarized by their median (red dotted lines) and their 5 and 95 percentiles (black dashed lines).

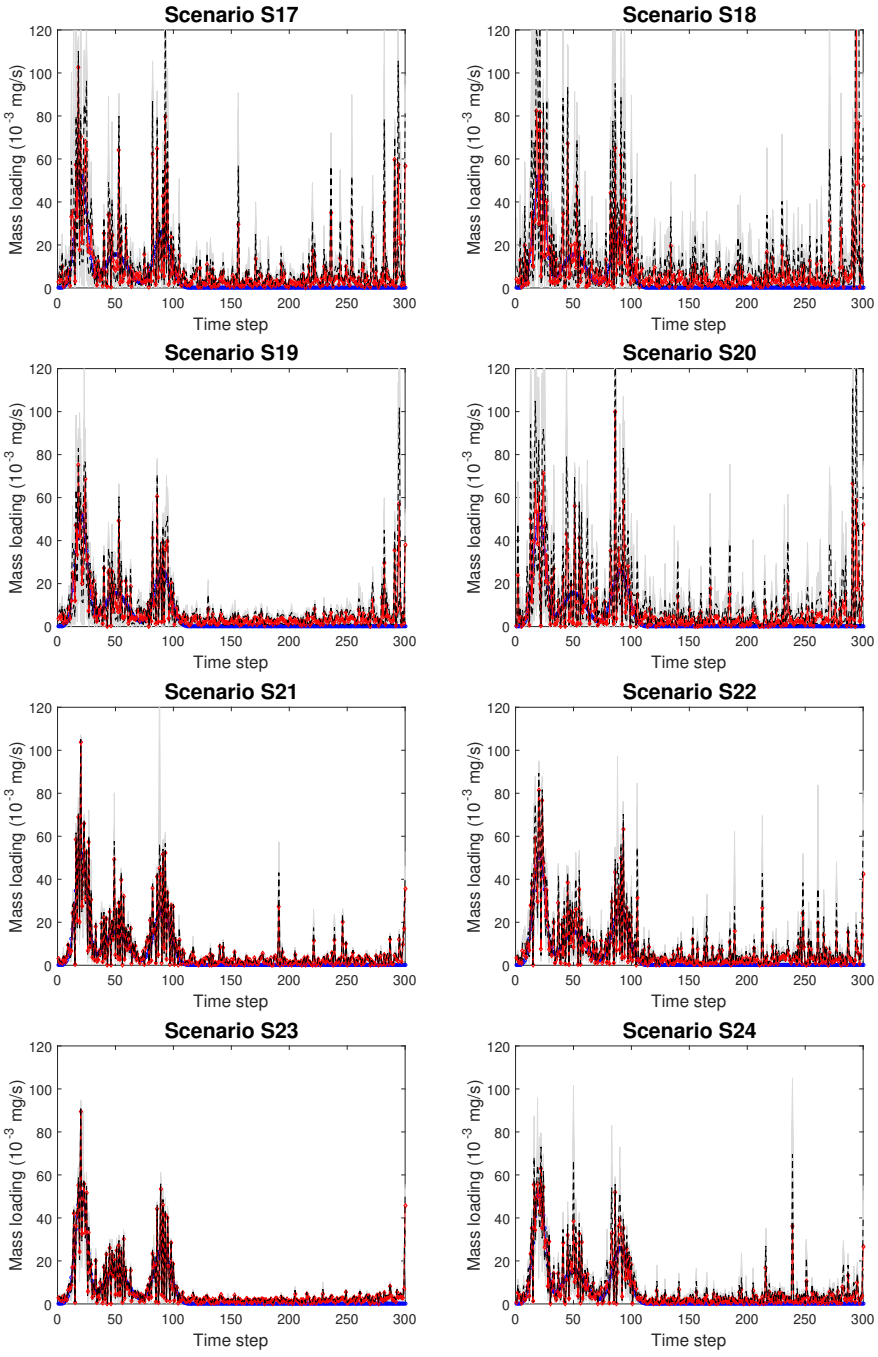

Figure 5.5. Recovered release histories for scenarios S17 to S24. The blue curve corresponds to the actual release history. The gray lines are the recovered release history curves for all 500 realizations, and they are summarized by their median(red dot lines) and their 5 and 95 percentiles(black dash lines).

Table 5.3. RMSE of the synthetic scenarios at the final 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 5.4. Definition of the sandbox scenarios

Scenario	time step	observation frequency	injection pulses
R1	50	5	train
R2	50	25	train
R3	100	10	train
R4	100	50	train
R5	50	5	two
R6	50	25	two
R7	100	10	two
R8	100	50	two

poorly. At the first iteration, the release history estimates are much worse than the totally random estimates used to initialize the ensemble. In the next realizations the method is capable to recover itself and to reduce the initial RMSE, although the absolute values shown in the table are by far the largest ones.

Based on this analysis, we decide to apply the ES-MDA with Rafiee's inflation scheme to the sandbox experiment. The release history is discretized into 50 or 100 time steps, and the number of iterations is set equal to eight.

5.3.4 Real Case

We performed two sandbox experiments with two kinds of release history curves, the first curve displays a train of four pulses lasting the entire duration of the experiment (Figure 5.7) and the second curve consists of two pulses at the beginning of the experiment (Figure 5.8). 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 chapter 3.6.2 for this two experiments. A number of scenarios will be analyzed that are described in Table 5.4.

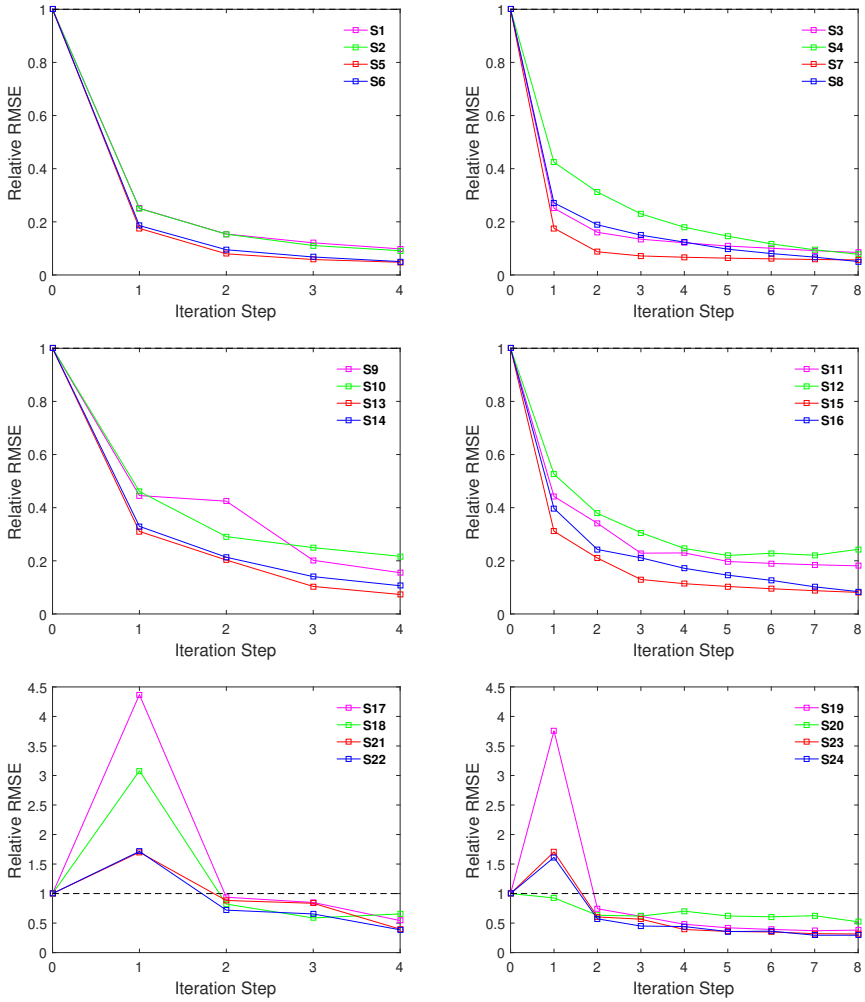

Figure 5.6. Evolution of the Relative RMSE of the synthetic scenarios as a function of the iteration step.

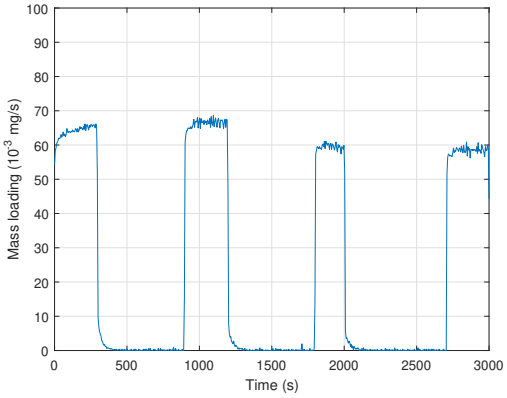

Figure 5.7. Release curve of the first sandbox experiment.

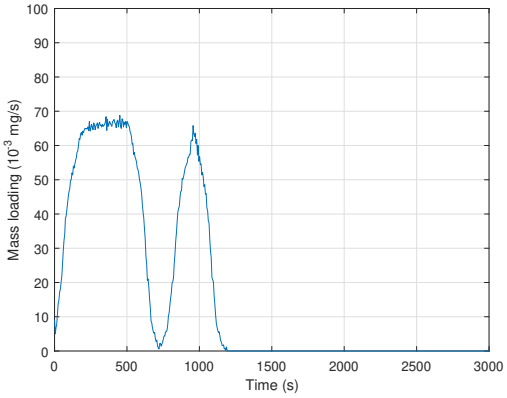

Figure 5.8. Release curve of the second sandbox experiment.

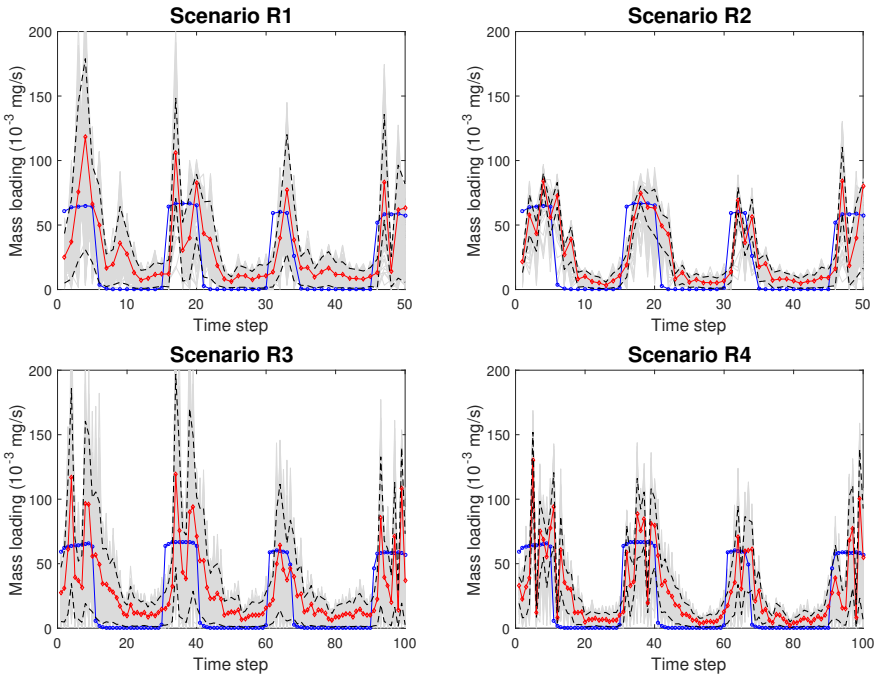

Figure 5.9. Recovered release history for first sandbox experiment, scenarios R1 to R4. The blue curve corresponds to the actual release history. The gray lines are the recovered release history curves for all 500 realizations, and they are summarized by their median (red dot lines) and their 5 and 95 percentiles (black dash lines).

Figure 5.9 shows the recovered release history curves for the first sandbox experiment. 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 highest frequency for that discretization is the one performing best. The same fluctuation as in the synthetic cases are 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 5.10 shows the recovered release curves for the second experiment. 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 clear pitfall 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

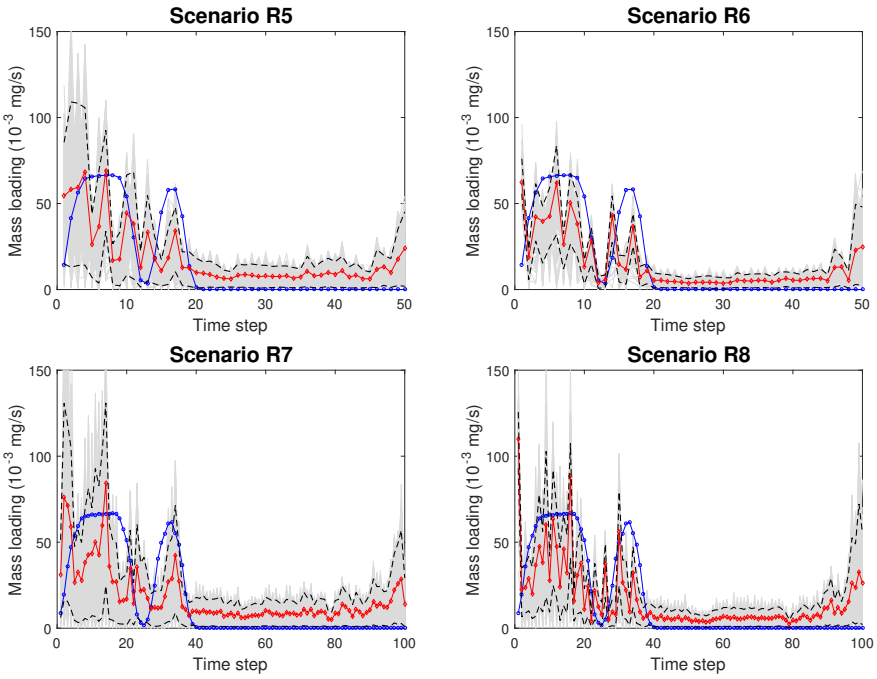

Figure 5.10. Recovered release history for the second sandbox experiment, scenarios R5 to R8. The blue curve corresponds to the actual release history. The gray lines are the recovered release history curves for all 500 realizations, and they are summarized by their median (red dotted lines) and their 5 and 95 percentiles (black dashed lines).

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 behavior is the magnitude of the concentration observation error variance.

5.4 Summary and Conclusion

In this paper, we employ the ES-MDA for the first time to identify a time varying release history in both synthetic and laboratory cases. In the synthetic cases, we examined the capacity of ES-MDA with different levels of discretization of the release curve, with a ratio of 1 to 6 between the coarsest and finest discretizations; the impact of the observation data frequency (every other time step versus one step every ten); the choice of an inflation factor scheme (between Rafiee's and Evensen's schemes); and the importance 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 of eight, does not have an important effect in 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 poorly random numbers for each time step with no temporal correlation at all.

Then, we apply the ES-MDA (using Rafiee's inflation scheme and eight iteration) 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 shows 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.

6 Conclusions

6.1 Summary

The work of this thesis is aimed to fulfill the main objectives described in Chapter 1.1. The main contributions or conclusions are the following:

- The restart EnKF is proven capable to identify a contaminant source together with some parameters defining the geometry of the aquifer in synthetic cases. Further application of the method on laboratory experiments shows the impact that observation errors may have in the estimation uncertainty. Using a too small observation error covariance results in more or less precise but biased estimates, while a too large observation error covariance results in a poor identification and too large uncertainty. Only with a proper evaluation of the observation errors, the r-EnKF could obtain a reliable result.

- Through the application of the restart NS-EnKF in synthetic case, we prove that this method with either a proper ensemble size or a suitable inflation method is able to jointly deduce contaminant source information and heterogenous conductivities by using only concentration data. Of all the test cases analyzed, Bauser's covariance inflation method appeared as the most appropriate, allowing to reduce the ensemble size from 1000 members (without inflation) to 500 (with inflation) and yielding similar results. However, if the ensemble size is too small and no inflation method is used, the filter will collapse. Further tests in sandbox experiments show that with a correct description of the observation error, the filter has more flexibility to update the parameters to

fit the observations while resulting in a larger variance on the ensemble of final parameters. The results also show a degree of indetermination in the estimation of the injection rate and the injection concentration; but, their product, the mass loading rate, is still well estimated with no bias and little uncertainty.

- Our work illustrates that the ES-MDA has the ability of identifying contaminant source parameters together with a spatially heterogeneous hydraulic conductivity field. The comparison of the performance between the r-EnKF and the ES-MDA shows that the ES-MDA performs equally or even better than the r-EnKF when using enough number of data assimilation steps, especially in identifying mass-loading and release duration parameters.

- The ES-MDA is a powerful method to identify a time-varying injection in both synthetic and real cases. In the synthetic cases, Rafiee's inflation scheme outperforms Evensen's inflation scheme. In all scenarios anaylzed, increasing the observation data frequency always improves the outcome of the recovered release injection curve. Using four or eight assimilation steps in the ES-MDA did not have a significant impact on the recovery of the release history in most scenarios.

6.2 Suggestions for Future Research

There are still some issues deserving further attention.

- **Analyzing different contamination events and their parameterization.**

  The outcome from the last chapter indicates that a more intelligent parameterization of the release curves could have led to better results. The study of alternative parameterization such as using basis functions or splines is worth investigating. Also, contamination events could be multiple and simultaneous in time, and also not be limited a point injection but spread over a larger area; the analysis of complex events is another promising line for further research.

- **Combining various observation data to update state parameters.**

  In this thesis, for the sandbox experiments, we used only the concentration data obtained after the luminosity-concentration transform described by Citarella et al. (2015). But, in reality, we will have access to other data, such as piezometric head data or electrical resistivity tomography data, which will be beneficial for identification purposes.

A promising research avenue is to combine different kind of observation data to jointly determine the contaminant source parameters and hydraulic conductivities.

- **Optimization of the observation network to improve the performance of ensemble based methods.**

  We have discussed the impact that different space (Chapter 2) and temporal (Chapter 5) arrangements of the observation network may have in the correct identification of the contaminant source; but we could go one step further and attempt to design the optimal spatio-temporal arrangement to achieve the same results at the smallest cost.

- **Demonstrating the r-EnKF, the restart NS-EnKF or the ES-MDA in field case studies.**

  The works we've done in this thesis have already move from synthetic cases to laboratory experiments. But these demonstrations are still far from field conditions, where boundary and initial conditions, forcing terms or geometry are not necessarily known. We need to consider the impact of these uncertainties in the application of ensemble-based algorithms, and find avenues to solve the difficulties found on field cases.

Bibliography

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.

Amirabdollahian, M., Datta, B., 2014. Identification of pollutant source characteristics under uncertainty in contaminated water resources systems using adaptive simulated anealing and fuzzy logic. International Journal of GEOMATE 6, 757–762.

Anderson, J.L., 2007. An adaptive covariance inflation error correction algorithm for ensemble filters. Tellus, Series A: Dynamic Meteorology and Oceanography 59, 210–224.

Aral, M.M., Guan, J., Maslia, M.L., 2001. Identification of Contaminant Source Location and Release History in Aquifers. Journal of Hydrologic Engineering 6, 225–234.

Atmadja, J., Bagtzoglou, A.C., 2001a. Pollution source identification in heterogeneous porous media. Water Resources Research 37, 2113–2125.

Atmadja, J., Bagtzoglou, A.C., 2001b. 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.

Ayvaz, M.T., 2016. A hybrid simulation-optimization approach for solving the areal groundwater pollution source identification problems. Journal of Hydrology 538, 161–176.

Bagtzoglou, A.C., Atmadja, J., 2003. Marching-jury backward beam equation and quasi-reversibility methods for hydrologic inversion: Application to contaminant plume spatial distribution recovery. Water Resources Research 39.

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.

Bauser, H.H., Berg, D., Klein, O., Roth, K., 2018. Inflation method for ensemble Kalman filter in soil hydrology. Hydrology and Earth System Sciences 22, 4921–4934.

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.

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.

Camporese, M., Cassiani, G., Deiana, R., Salandin, P., 2011. Assessment of local hydraulic properties from electrical resistivity tomography monitoring of a three-dimensional synthetic tracer test experiment. Water Resources Research 47, 1–15.

Capilla, J.E., Rodrigo, J., Gómez-Hernández, J.J., 1999. Simulation of non-gaussian transmissivity fields honoring piezometric data and integrating soft and secondary information. Mathematical Geology 31, 907–927.

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.

Chang, H., Zhang, D., Lu, Z., 2010. History matching of facies distribution with the EnKF and level set parameterization. Journal of Computational Physics 229, 8011–8030.

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

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.

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.

Crestani, E., Camporese, M., Baú, D., Salandin, P., 2012. Ensemble Kalman filter versus ensemble smoother for assessing hydraulic conductivity via tracer test data assimilation. Hydrology and Earth System Sciences Discussions 9, 13083–13115.

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.

Datta, B., Chakrabarty, D., Dhar, A., 2009. Simultaneous identification of unknown groundwater pollution sources and estimation of aquifer parameters. Journal of Hydrology 376, 48–57.

Emerick, A.A., Reynolds, A.C., 2013a. Ensemble smoother with multiple data assimilation. Computers and Geosciences 55, 3–15.

Emerick, A.A., Reynolds, A.C., 2013b. Investigation of the sampling performance of ensemble-based methods with a simple reservoir model. Computational Geosciences 17, 325.

Emerick, A.A., Reynolds, A.C., et al., 2013. History-matching production and seismic data in a real field case using the ensemble smoother with multiple data assimilation, in: SPE Reservoir Simulation Symposium, Society of Petroleum Engineers.

Evensen, G., 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. Journal of Geophysical Research 99, 10143.

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

Evensen, G., 2004. Sampling strategies and square root analysis schemes for the EnKF. Ocean Dynamics 54, 539–560.

Evensen, G., 2009. Data assimilation: the ensemble Kalman filter. Springer Science & Business Media.

Evensen, G., 2018. Analysis of iterative ensemble smoothers for solving inverse problems. Computational Geosciences 22, 885–908.

Evensen, G., van Leeuwen, P.J., 2000. An Ensemble Kalman Smoother for Nonlinear Dynamics. Monthly Weather Review 128, 1852–1867.

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.

Fokker, P., Wassing, B., van Leijen, F., Hanssen, R., Nieuwland, D., 2016. Application of an ensemble smoother with multiple data assimilation to the bergermeer gas field, using ps-insar. Geomechanics for Energy and the Environment 5, 16–28.

Franssen, H.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.

Gómez-Hernández, J.J., Journel, A.G., 1993. Joint sequential simulation of Multi-Gaussian fields, in: Soares, A. (Ed.), Geostatistics Tróia '92, Kluwer Academic Publishers, Dordrecht. pp. 85–94.

Gómez-Hernández, J.J., Wen, X.H., 1998. To be or not to be multi-Gaussian? A reflection on stochastic hydrogeology. Advances in Water Resources 21, 47–61.

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

Gu, Y., Oliver, D., 2007. An iterative ensemble kalman filter for multiphase fluid flow data assimilation. SPE Journal 12, 438–446.

Gzyl, G., Zanini, A., Fraczek, R., Kura, K., 2014. Contaminant source and release history identification in groundwater: A multi-step approach. Journal of Contaminant Hydrology 157, 59–72.

Hendricks Franssen, H.J., Kinzelbach, W., 2008. Real-time groundwater flow modeling with the Ensemble Kalman Filter: Joint estimation of states and

parameters and the filter inbreeding problem. Water Resources Research 44, 1–21.

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

Journel, A., Isaaks, E., 1984. Conditional indicator simulation: application to a saskatchewan uranium deposit. Journal of the International Association for Mathematical Geology 16, 685–718.

Journel, A.G., Gomez-Hernandez, J.J., et al., 1993. Stochastic imaging of the wilmington clastic sequence. SPE formation Evaluation 8, 33–40.

Kalman, R., et al., 1960. A new approach to linear filtering and prediction problems. Journal of basic Engineering 82, 35–45.

Knudby, C., Carrera, J., 2005. On the relationship between indicators of geostatistical, flow and transport connectivity. Advances in Water Resources 28, 405–421.

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.

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.

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.

Le, D.H., Younis, R., Reynolds, A.C., et al., 2015. A history matching procedure for non-gaussian facies based on es-mda, in: SPE Reservoir Simulation Symposium, Society of Petroleum Engineers.

Lee, K., Jeong, H., Jung, S., Choe, J., 2013. Improvement of ensemble smoother with clustered covariance for channelized reservoirs. Energy Exploration & Exploitation 31, 713–726.

van Leeuwen, P.J., Evensen, G., 1996. Data Assimilation and Inverse Methods in Terms of a Probabilistic Formulation. Monthly Weather Review 124, 2898–2913.

Li, H., Kalnay, E., Miyoshi, T., 2009. Simultaneous estimation of covariance inflation and observation errors within an ensemble Kalman filter. Quarterly Journal of the Royal Meteorological Society 135, 523–533.

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.

Li, L., Puzel, R., Davis, A., 2018a. Data assimilation in groundwater modelling: ensemble kalman filter versus ensemble smoothers. Hydrological Processes 32, 2020–2029.

Li, L., Stetler, L., Cao, Z., Davis, A., 2018b. An iterative normal-score ensemble smoother for dealing with non-gaussianity in data assimilation. Journal of Hydrology .

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.

Li, L., Zhou, H., Hendricks Franssen, H.J., Gómez-Hernández, J.J., 2012c. Modeling transient groundwater flow by coupling ensemble kalman filtering and upscaling. Water Resources Research 48.

Liang, X., Zheng, X., Zhang, S., Wu, G., Dai, Y., Li, Y., 2011. Maximum likelihood estimation of inflation factors on error covariance matrices for ensemble kalman filter assimilation. Quarterly Journal of the Royal Meteorological Society 138, 263–273.

Liang, X., Zheng, X., Zhang, S., Wu, G., Dai, Y., Li, Y., 2012. Maximum likelihood estimation of inflation factors on error covariance matrices for ensemble Kalman filter assimilation. Quarterly Journal of the Royal Meteorological Society 138, 263–273.

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.

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.

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.

Mirghani, B.Y., Mahinthakumar, K.G., Tryby, M.E., Ranjithan, R.S., Zechman, E.M., 2009. A parallel evolutionary strategy based simulation-optimization approach for solving groundwater source identification problems. Advances in Water Resources 32, 1373–1385.

Neupauer, R.M., Borchers, B., Wilson, J.L., 2000. Comparison of inverse methods for reconstructing the release history of a groundwater contamination source. Water Resources Research 36, 2469–2475.

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.

Rafiee, J., Reynolds, A.C., 2017. Theoretical and efficient practical procedures for the generation of inflation factors for ES-MDA. Inverse Problems 33.

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.

Reynolds, A.C., Zafari, M., Li, G., 2006. Iterative forms of the ensemble kalman filter, in: ECMOR X-10th European Conference on the Mathematics of Oil Recovery.

Sidauruk, P., Cheng, A.D., Ouazar, D., 1998. Ground water contaminant source and transport parameter identification by correlation coefficient optimization. Ground Water 36, 208–214.

Skaggs, T.H., Kabala, Z., 1995. Recovering the history of a groundwater contaminant plume: Method of quasi-reversibility. Water Resources Research 31, 2669–2673.

Skaggs, T.H., Kabala, Z.J., 1994. Recovering the release history of a groundwater contaminant. Water Resources Research 30, 71–79.

Sun, A.Y., 2007. A robust geostatistical approach to contaminant source identification. Water resources research 43.

Sun, A.Y., Morris, A.P., Mohanty, S., 2009. Sequential updating of multimodal hydrogeologic parameter fields using localization and clustering techniques. Water Resources Research 45, 1–15.

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

Sun, A.Y., Painter, S.L., Wittmeyer, G.W., 2006b. A robust approach for iterative contaminant source location and release history recovery. Journal of contaminant hydrology 88, 181–196.

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 .

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.

Wang, H., Jin, X., 2013. Characterization of groundwater contaminant source using bayesian method. Stochastic environmental research and risk assessment 27, 867–876.

Wang, X., Bishop, C.H., 2003. A Comparison of Breeding and Ensemble Transform Kalman Filter Ensemble Forecast Schemes. Journal of the Atmospheric Sciences 60, 1140–1158.

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. Probability fields revisited in the context of ensemble kalman filtering. Journal of Hydrology 531, 40–52.

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., 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, J.G., 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., 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.

Zeng, L., Shi, L., Zhang, D., Wu, L., 2012. A sparse grid based bayesian method for contaminant source identification. Advances in Water Resources 37, 1–9.

Zhang, J., Zeng, L., Chen, C., Chen, D., Wu, L., 2015. Efficient bayesian experimental design for contaminant source identification. Water Resources Research 51, 576–598.

Zheng, C., 2010. 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 , 219.

Zheng, X., 2009. An adaptive estimation of forecast error covariance parameters for Kalman filtering data assimilation. Advances in Atmospheric Sciences 26, 154–160.

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., 2012a. A pattern-search-based inverse method. Water Resources Research 48.

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

Zhou, H., Li, L., Franssen, H.J.H., Gómez-Hernández, J.J., 2012b. Pattern recognition in a bimodal aquifer using the normal-score ensemble kalman filter. Mathematical Geosciences 44, 169–185.

Zinn, B., Harvey, C.F., 2003. When good statistical models of aquifer heterogeneity go bad: A comparison of flow, dispersion, and mass transfer in connected and multivariate Gaussian hydraulic conductivity fields. Water Resources Research 39, 137–147.