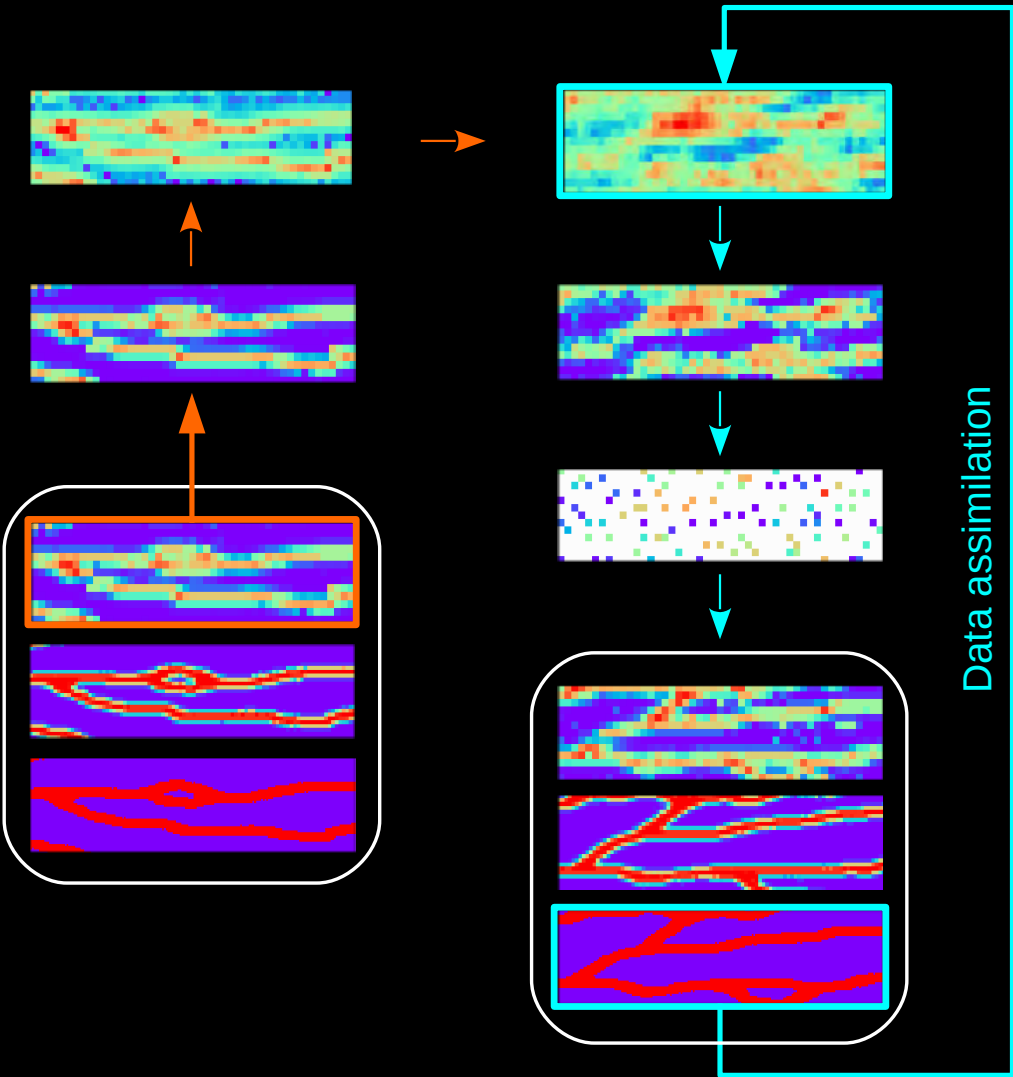




Conditioning groundwater flow parameters with iterative ensemble smoothers: Analysis and approaches in the continuous and the discrete cases

UNIVERSITY OF NEUCHÂTEL, SWITZERLAND
FACULTY OF SCIENCE
CENTER FOR HYDROGEOLOGY AND GEOTHERMICS
(CHYN)

Conditioning groundwater flow
parameters with iterative
ensemble smoothers: Analysis
and approaches in the continuous
and the discrete cases

A thesis presented for the degree of

DOCTOR OF SCIENCES

by

Dan-Thuy LAM

accepted on the recommendation of

Prof. Philippe RENARD

Dr. Jaouher KERROU

Dr. Mickaele LE RAVALEC

Dr. Anahita ABADPOUR

Prof. Jaime GOMEZ-HERNANDEZ

Dr. Hakim BENABDERRAHMANE

Prof. Pierre PERROCHET

DEFENDED ON OCTOBER 17th, 2019

IMPRIMATUR POUR THESE DE DOCTORAT

**La Faculté des sciences de l'Université de Neuchâtel
autorise l'impression de la présente thèse soutenue par**

Madame Dan-Thuy LAM

Titre:

**“Conditioning groundwater flow parameters
with iterative ensemble smoothers:
Analysis and approaches in the continuous
and the discrete cases”**

sur le rapport des membres du jury composé comme suit:

- Prof. associé Philippe Renard, directeur de thèse, Université de Neuchâtel, Suisse
- Dr Jaouher Kerrou, co-directeur de thèse, Université de Neuchâtel, Suisse
- Prof. Pierre Perrochet, Université de Neuchâtel
- Prof. Jaime Gomez-Hernandez, Université de Valencia, Espagne
- Dr Anahita Abadpour, Total, France
- Dr Mickaele Le Ravalec, IFP Energies nouvelles, Rueil-Malmaison, France
- Dr Hakim Benabderrahmane, Andra, France

Neuchâtel, le 28 novembre 2019

Le Doyen, Prof. P. Felber



*The idea comes to me from
outside of me - and is like
a gift. I then take the idea
and make it my own - that
is where the skill lies.*

—Johannes Brahms

Acknowledgements

This work was fully funded by the University of Neuchâtel and achieved at the Center for Hydrogeology and Geothermics (CHYN). I would like to thank the French National Radioactive Waste Management Agency (ANDRA) for motivating the topic of this thesis, the Faculty of science of the University of Neuchâtel for providing the computational resources, and the RandLab research group for providing a version of DeeSse dedicated to the development of the methodology presented in this thesis. I would also like to thank all the members of the jury for assessing this work and for their helpful remarks.

Remerciements

Je souhaiterais commencer par remercier l'ensemble des personnes qui ont contribué à la réalisation de cette thèse et à son évaluation.

Merci en premier lieu à mon co-directeur de thèse Jaouher Kerrou pour avoir vu en moi un potentiel afin de réaliser cette thèse. Merci à lui et Pierre Perrochet pour leur confiance, leur disponibilité, et de m'avoir donné une grande liberté d'entreprendre dès le début. J'ai eu la chance d'acquérir une autonomie enrichissante tout en bénéficiant de vos lumières sur la modélisation numérique des écoulements souterrains. J'ai aussi eu la chance d'effectuer un travail en lien avec le contexte réel du projet de l'Andra. Merci à Hakim Benabderrahmane pour les informations apportées lors de ces réunions de projet où j'ai pu régulièrement présenter l'avancée de mon travail.

Au-delà du projet initial, je tiens à exprimer ma profonde gratitude envers Philippe Renard, devenu directeur de ma thèse et à qui je dois la partie la plus innovante de celle-ci ainsi que la cohérence de l'ensemble. Merci d'avoir cru en ce travail et de m'avoir aidée à suivre la direction que je souhaitais à cette transition de la thèse. L'idée que tu as eue m'aura tenue en haleine de long mois avant de me donner confiance jusqu'à la toute fin. Merci pour cette expérience grisante et pour tes remarques qui m'ont aidée à aller plus loin dans mon raisonnement et à consolider le tout. Je tiens aussi à remercier Julien Straubhaar qui m'a été d'une grande aide durant cette phase plus inconnue de la thèse. En particulier, je ne serais pas arrivée à la méthode proposée sans son implication sur le logiciel ou ses précieux retours lors des développements intermédiaires.

La réussite de cette thèse, je la dois donc à mes encadrants ainsi qu'à toutes les personnes autour qui m'ont aidée et encouragée.

Merci également aux rapporteur(e)s de mon jury qui m'ont fait l'honneur d'accepter d'évaluer ma thèse : Mickaele Le Ravalec, Anahita Abadpour, et Jaime Gomez-Hernandez. La soutenance a été pour moi une excellente opportunité de présenter le travail réalisé auprès de personnes extérieures au projet. Je suis en particulier reconnaissante envers les personnes qui ont permis que celle-ci se déroule de la meilleure des façons possibles. Merci pour l'intérêt que vous avez témoigné pour ce travail à travers vos remarques, questions et suggestions.

Je veux aussi remercier mes collègues Laurent Fasnacht et Przemyslaw Juda sans qui j'aurais sans l'ombre d'un doute fini au chômage technique. Merci à eux d'avoir consacré beaucoup de leur temps de doctorat afin d'assurer la maintenance du cluster de la Faculté des sciences. Ce cluster a en effet animé mon quasi-quotidien, non sans pénibles complications qui m'ont marquée au fer rouge. Merci pour votre pédagogie et tous les conseils qui ont grandement facilité mon travail.

Un grand merci à toutes les personnes dont j'ai croisé la route et qui ont participé aux souvenirs mémorables de mes années de thèse. En particulier, merci Asmae pour tous les moments de convivialité et de rigolade qui ont égayé mon quotidien pendant quatre ans. Merci Pauline d'avoir été la meilleure collègue de bureau qui soit : pour tout le bon karma à chacune de tes venues et ton soutien indéfectible. Merci Aline pour ta compagnie réconfortante et tes conseils tous domaines confondus. Merci Jonas et les fidèles biologistes pour tous les moments partagés au début de ma thèse dont je garde d'excellents souvenirs. Merci Martin, Axa de m'avoir fait découvrir les apéros interlabos et pour votre recul sur la thèse. Merci Laurent pour nos discussions interminables et tes précieux conseils quand j'étais dans l'impasse. Merci Gunnar pour ta perspicacité et ton sens de l'éthique qui forcent l'admiration. Merci Kalliopi pour ta chaleureuse amitié et de m'avoir fait autant rire durant ma fin de thèse. Merci Przemek pour ton inspirante sérénité et ton expertise qui m'a éclairée tant de fois. Merci Batoul pour ta gentillesse inégalable et ta grande compréhension durant mes intenses analyses de *la* série. Merci Lucile pour ton enthousiasme contagieux et de raviver l'esprit d'équipe. Merci Reza pour ta bonne humeur et les discussions animées de la cafétéria. Merci Simone, Jeremy et Alvaro pour les innombrables bouffes en votre agréable compagnie. Merci Fabien pour tes encouragements malgré tes blagues que je ne saurais qualifier. Enfin, un grand merci à Roxane, Hannah, Simon d'avoir été d'aussi enthousiastes partenaires de musique

de chambre durant ces récrés mozartiennes, beethoveniennes, schumannniennes, brahmsiennes, hélas non-mendelssohniennes. Je remercie au passage le décanat de m'avoir laissée jouer presque tous les jours dès 17h15 sur le piano ressuscité de l'UniMail. Un grand merci aussi à l'Académie de Musique de Neuchâtel de m'avoir permis de jouer sur leurs beaux Petrof à demi-queue.

Ma vie à Neuchâtel n'aurait sans doute pas été la même sans ma superbe coloc'. Un immense merci à Nathalie, les monstres (chats) Loki et Roby pour vos bonnes ondes à l'appart' et pour tous les instants passés ensemble. Merci à mes amis ex-STUE/EOST de longue date Eleonore, Pauline, Rodolphe, Liliane, Malwina, Audrey, Léa et Marine pour nos retrouvailles épiques en Alsace, sur Skye, en Suisse, au milieu des champs de lavande, à Broadway, dans le Doubs... qui m'ont permis de couper régulièrement. Merci aussi à mes proches cousines pour tous les loisirs et fous-rires partagés au cours de ces années.

Je termine en remerciant mes parents et mon frère qui ont toujours cru en moi plus que moi-même. Merci de m'avoir communiqué toute votre persévérance et patience qui sont ma plus grande force depuis le début. Merci pour ces précieux moments en famille et aussi pour tous ces repas qui continuent à faire des envieux. Avec le recul, je mesure l'immense chance de vous avoir eu aussi proches durant cette étape.

Abstract

Data assimilation (DA) consists in combining observations and predictions of a numerical model to produce an optimal estimate of the evolving state of a system. Over the last decade, DA methods based on the Ensemble Kalman Filter (EnKF) have been particularly explored in various geoscience fields for inverse modelling. Although this type of ensemble methods can handle high-dimensional systems, they assume that the errors coming from whether the observations or the numerical model are multi-Gaussian. To handle potential nonlinearities between the observations and the state or parameter variables to estimate, iterative variants have been proposed. In this thesis, we first focus on two main iterative ensemble smoother methods for the calibration of a synthetic 2D groundwater model. Using the same set of sparse and transient flow data, we analyse each method when employing them to condition an ensemble of multi-Gaussian hydraulic conductivity fields. We then further explore the application of one iterative ensemble smoother algorithm (ES-MDA) in situations of variable complexity, depending on the geostatistical simulation method used to simulate the prior geological information. The applicability of a parameterization based on the normal-score transform is first investigated. The robustness of the method against nonlinearities is then further explored in the case of discrete facies realizations obtained with a truncated Gaussian technique and updated via their underlying variables. Based on the observed limitations and benefits of the forementioned parameterizations, we finally propose a new methodology for the conditioning of categorical multiple-point statistics (MPS) simulations to dynamic data with a state-of-the-art ensemble Kalman method by taking the example of the Ensemble Smoother with Multiple Data Assimilation (ES-MDA). Our methodology relies on a novel multi-resolution parameterization of the categorical MPS simulation. The ensemble of latent parameters is initially defined on the basis of the coarsest-resolution simulations of an

ensemble of multi-resolution MPS simulations. Because this ensemble is non-multi-Gaussian, additional steps prior to the computation of the first update are proposed. In particular, the parameters are updated at predefined locations at the coarsest scale and integrated as hard data to generate a new multi-resolution MPS simulation. The results on the synthetic problem illustrate that the method converges towards a set of final categorical realizations that are consistent with the initial categorical ensemble. The convergence is reliable in the sense that it is fully controlled by the integration of the ES-MDA update into the new conditional multi-resolution MPS simulations. Moreover, thanks to the proposed parameterization, the identification of the geological structures during the data assimilation is particularly efficient for this example. The comparison between the estimated uncertainty and a reference estimate obtained with a Monte Carlo method shows that the uncertainty is not severely reduced during the assimilation as is often the case. The connectivity is successfully reproduced during the iterative procedure despite the rather large distance between the observation points.

Key words

ensemble Kalman methods; data assimilation; iterative ensemble smoother algorithms; inverse modelling; parameter identification; uncertainty; groundwater; multi-Gaussian variables; non-Gaussian variables; categorical variables; parameterization; normal-score transform; multi-resolution multiple-point statistics

Contents

1	Introduction	1
1.1	Context and motivation	1
1.2	Scope of the current research	5
1.3	Structure of the thesis	6
2	Background on ensemble-based data assimilation	9
2.1	A brief statement of the situation	9
2.2	Background on ensemble-based data assimilation	11
2.2.1	Understanding the Kalman filter	11
2.2.2	The multivariate Kalman filtering equations . . .	13
2.2.3	The ensemble strategy	14
2.2.4	Inverse modelling with the ensemble Kalman filter	16
2.2.5	Practical issues with ensemble methods	18
3	Conditioning multi-Gaussian groundwater flow parameters to transient hydraulic head and flowrate data with iterative ensemble smoothers: a synthetic case study	25
3.1	Introduction	25
3.2	General background on the ensemble smoother and iterative	27
3.2.1	Levenberg-Marquardt Ensemble Randomized Likelihood (LM-EnRML)	29
3.2.2	Ensemble Smoother with Multiple Data Assimilation (ES-MDA)	31
3.2.3	Normal-score transform of state variables with ES-MDA	32
3.3	A synthetic inverse problem inspired by the Andra's Meuse/Haute-Marne site	34
3.3.1	Model set up	35
3.3.2	Synthetic data set	36

3.3.3	Initial ensemble of parameters and assumptions for the update step	37
3.4	Performance criteria	39
3.5	Results and discussion	42
3.5.1	Ensemble smoother and benefit of data transfor- mation	42
3.5.2	Comparing the accuracy of LM-EnRML and ES- MDA estimates	48
3.5.3	Assimilating both hydraulic head and flowrate data with LM-EnRML and ES-MDA	53
3.6	Conclusion	60
4	Updating non-multi-Gaussian groundwater flow param- eters of a channelized aquifer with iterative ensemble smoothing	63
4.1	Introduction	63
4.2	Methods	65
4.2.1	ES-MDA	65
4.2.2	Iterative normal-score ensemble smoother	67
4.3	Synthetic test case	69
4.3.1	Synthetic data set	69
4.3.2	Initial ensemble	71
4.4	Performance criteria	71
4.5	Results and discussion	72
4.6	Conclusion	80
5	Conditioning truncated Gaussian realizations to tran- sient hydraulic head and flowrate data with iterative en- semble smoothing	81
5.1	Introduction	81
5.2	Methods	83
5.2.1	Generation of facies realizations with the trun- cated Gaussian or bi-Gaussian simulation technique	83
5.2.2	Application of ES-MDA	84
5.2.3	Synthetic test cases	85
5.2.4	Performance criteria in the discrete case	88
5.3	Results and discussion	92
5.3.1	Truncated Gaussian case	92
5.3.2	Bi-Gaussian case	96
5.4	Conclusion	101

6	Multi-resolution approach to condition categorical multiple-point realizations to dynamic data with iterative ensemble smoothing	105
6.1	Introduction	105
6.2	Multi-resolution multiple-point geostatistical simulations with DeeSse	108
6.3	Ensemble data assimilation with ES-MDA	111
6.3.1	General background	112
6.3.2	Localization and normal-score transform	115
6.4	ES-MDA with multi-resolution MPS simulations	117
6.4.1	Obtaining the latent variable	117
6.4.2	Conditioning the categorical field	118
6.4.3	Normal-score transform	119
6.4.4	Sampling and final parameterization of the multi-resolution MPS simulations	121
6.4.5	Complete procedure	124
6.5	Synthetic test case	125
6.5.1	Model set up	125
6.5.2	Synthetic data set	127
6.5.3	Generation of the multi-resolution training images with DeeSse	129
6.5.4	Initial ensemble of parameters and assumptions for the ES-MDA update	130
6.5.5	DeeSse parameters for the multi-resolution MPS simulations	131
6.5.6	Performance assessment criteria	132
6.6	Results and discussion	134
6.6.1	Evolution of the data mismatch	134
6.6.2	Sensitivity to DeeSse parameters of the integration of the ES-MDA update	137
6.6.3	Identification of the categorical fields	139
6.6.4	Analysis of case B	144
6.6.5	Comparison with other approaches	146
6.7	Conclusion	153
7	Conclusions	155
7.1	Summary of main contributions	156
7.1.1	Critical ensemble size in the multi-Gaussian case	156
7.1.2	Clarification on the role of the normal-score transform	157

7.1.3	A novel parameterization of multiple-point real-izations	159
7.1.4	Constraining the reduced parameter space	160
7.1.5	Benefit of the multi-resolution	161
7.2	Perspectives for future research	161
Appendix		165

Chapter 1

Introduction

1.1 Context and motivation

In groundwater modelling, the term “calibration” is often used to refer to the process of adjusting model inputs (e.g. hydraulic parameters, boundary conditions, initial conditions, source terms) so that the simulated state variables (e.g. hydraulic heads, groundwater flow rates, solute concentrations, etc.) reasonably match available observations. In this context, state observations can be leveraged in order to identify uncertain inputs of the system being modelled. This problem, more generally known as the “inverse problem”, is particularly challenging in geosciences as the number of available observations is often much smaller than the number of unknown. In fact, the number of variables to be inferred is directly related to the model resolution. With the increasing complexity of the numerical models developed nowadays, this discrepancy between the number of observations and the number of unknown also tends to increase. In these situations, the inverse problem is under-determined and does not have a unique solution. Depending on the context, the inverse problems can suffer as well from the absence of solution or extreme sensitivities to small variations in measurements of the state variables. All these properties make the inverse problem mathematically ill-posed. A first step in mitigating the ill-posedness of the problem consists in restricting the space within which the unknown may vary by considering some prior information or so-called “expert knowledge”. Even then, many solutions of the inverse problem can exist. A main challenge when solving inverse problems is therefore to be able to address the uncertainty associated to the variables to estimate. In a stochastic framework, the initial uncertainty can be represented by a

presumed probability distribution. Then, inverse methods derived from a probabilistic framework can be employed in order to estimate a new distribution which allows a proper fit of the state observations.

This thesis was originally motivated by a real inverse problem as part of the investigations conducted by the French National Agency for Radioactive Waste Management (ANDRA) to study the conditions of a deep geological repository in northeastern France (Figure 1.1). In the framework of this project, a transient variably-saturated groundwater flow model of a local multi-layered aquifer system was developed in order to better understand the local flow behavior at the Meuse/Haute-Marne site (Figure 1.1) (Benabderrahmane et al., 2014; Kerrou et al., 2017). Since the excavation of the access shafts of about 500 m depth to the Underground Research Laboratory (URL) in the target low-permeable Callovo-Oxfordian formation, local hydraulic head time series (Figure 1.2) as well as flowrates produced at different depths in the drained shafts have been recorded in the permeable Oxfordian formations located above (Delay et al., 2007). The inverse problem consists in using these local flow data covering a large period of time (17 years) in order to improve the characterization of the lithologic structures and heterogeneities of the Oxfordian formations. Specifically, the goal is to condition the uncertain hydraulic conductivity and specific storage parameters of the numerical groundwater model.

Given the nonlinearity and high dimensionality of the numerical groundwater model developed by ANDRA, ensemble data assimilation methods have been considered because their large-scale applicability has been illustrated more particularly in numerical weather prediction (Anderson, 2009), oceanography (Bertino et al., 2003) and reservoir modelling (Gu and Oliver, 2006). The particular challenges that the data assimilation methods will have to face for this case are two-fold:

1. Considering the presence of geological structures of higher permeabilities in the conceptual model (Figure 1.3), an important goal will be to generate conditional fields that are consistent with the prior geological information.
2. Apart from the long time series of hydraulic head and flowrate observations collected in the near vicinity of the two shafts, there is no additional monitoring of the transient piezometric heads to support the model calibration in the region of interest (Figure 1.1).

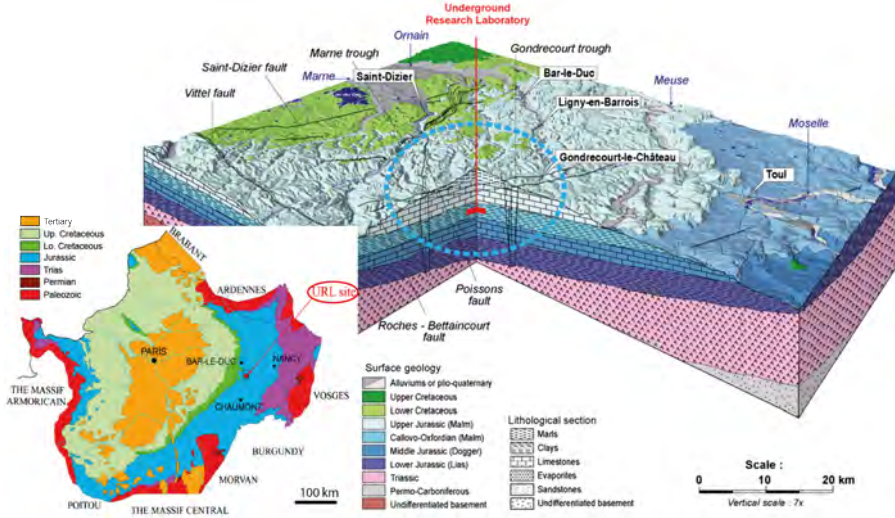


Figure 1.1: Location of the Meuse/Haute-Marne Underground Research Laboratory (URL) in the Paris Basin with the geological settings, and the region of interest for the model calibration circled in blue. (Modified after Delay et al. (2007); Linard et al. (2011); Kerrou et al. (2017).)

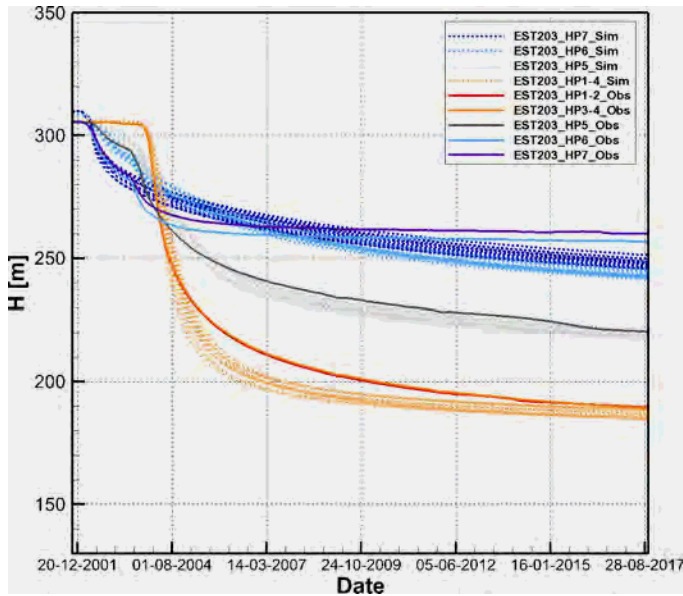
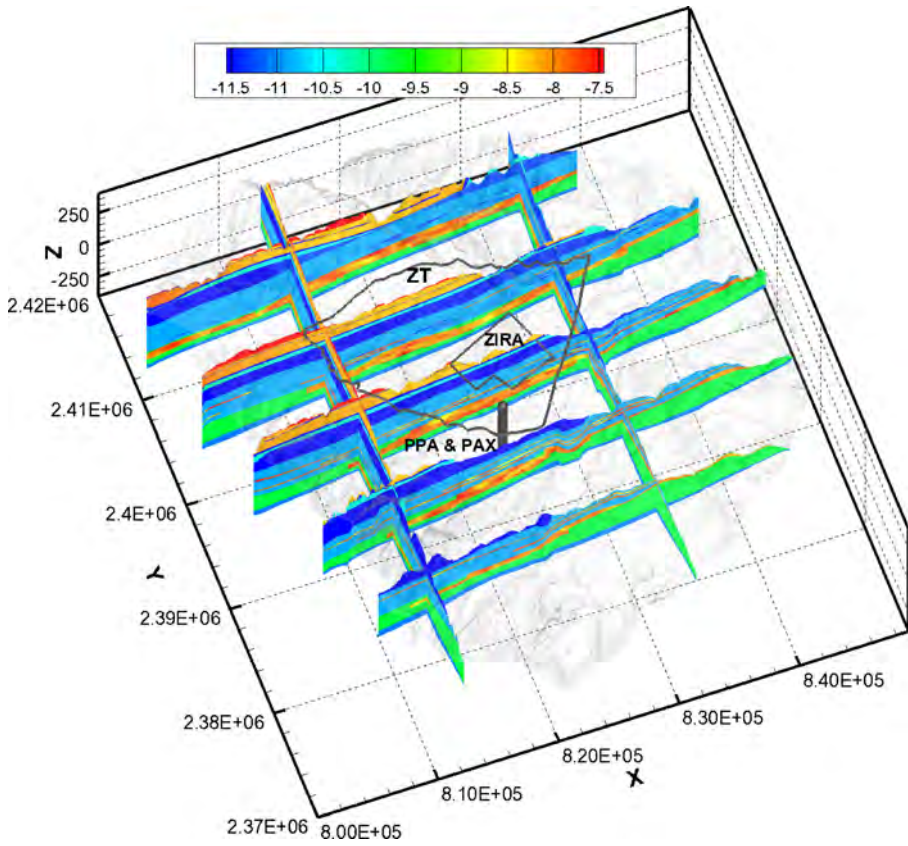
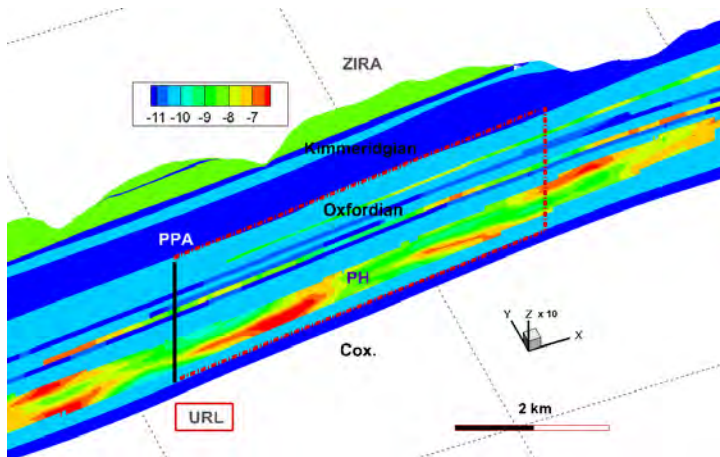


Figure 1.2: Example of hydraulic head observations (solid lines) monitored at different locations and depths in the near vicinity of the shafts.



(a) 2D cross sections of a 3D log hydraulic conductivity field with the locations of the two access shafts (PPA & PAX) and the area of interest (ZT and ZIRA).



(b) Zoom on one cross section in the vicinity of the PPA shaft.

Figure 1.3: Overview of the 3D groundwater model of the Andra's Underground Research Laboratory (URL) site in Meuse/Haute-Marne.

1.2 Scope of the current research

This thesis is focused on the analysis of a group of ensemble Kalman methods, namely the iterative ensemble smoothers, for the inverse modelling of groundwater flow parameters. Given the underlying Gaussian and linear assumptions of this type of data assimilation methods, favorable and less favorable synthetic cases have been investigated.

In many fields, we observe that there is a growing interest for the estimation of bounded or categorical variables (Bocquet et al., 2010). Within the framework of inverse modelling and data assimilation more generally, this interest motivates the development of novel methods which can deal with non-multi-Gaussian distributions (Liu and Oliver, 2005; Chen et al., 2009; Zhou et al., 2011; Li et al., 2012; Hu et al., 2013; Jafarpour and Khodabakhshi, 2011; Cao et al., 2018; Ma and Jafarpour, 2018). Hence after first having explored the capability of the iterative ensemble smoother in a proper multivariate Gaussian case, the research work aimed to progressively depart from this standard case and ultimately address the less favorable discrete case.

For all cases, a synthetic groundwater flow model inspired by the real field case of the ANDRA's project and previous studies (Bourgeat et al., 2004; Deman et al., 2015) was used to conduct the numerical experiments. The idea was to be able to perform systematic tests on a manageable model in terms of computational resource and running time. Although this model is synthetic, it is worth underlining that it still considers the particularities of the real field case. In particular, one specificity of this research is the sparse conditioning from a few observations points. This sparsity is not addressed to that extent in most synthetic studies, and actually represented a major constraint for the successful application of the method in the cases that we considered. Also, in order to consider a case that would be close to the real practical situation, we did not consider hard conditioning data prior to the data assimilation. Consequently, the issue of preserving the hard data conditioning throughout the data assimilation was not addressed in this thesis.

In the end, the research focused on the analysis of several variants of a synthetic inverse problem for which a proper implementation of the ensemble method was feasible despite well known practical issues of ensemble Kalman methods. We intend to give in this thesis a walk-through of the different aspects which have been learned through experiments, and which overall contributed to the development of a new methodol-

ogy for the conditioning of categorical fields generated by multiple-point statistics (MPS). Thanks to the realistic constraints that were set up in the synthetic inverse problem, we believe that the proposed data assimilation procedure offers promising perspectives for the application to real cases of ensemble Kalman methods in the discrete case.

Code and implementation

All the ensemble Kalman methods that were tested and studied in this thesis were implemented in Python from scratch. The detailed description of the algorithms can be found for example in the appendix of (Chen and Oliver, 2013; Emerick, 2016). To perform our tests efficiently, the code was developed such that it could run in parallel the ensemble of prediction and update steps on a Linux cluster. The implementation of the explored methodologies and the visualization of the test case results will soon be available on GitHub (<https://github.com/lamd91>). In addition, the two following codes have been employed:

- The finite element code GroundWater (GW) (Fabien Cornaton, 2014) for solving the groundwater flow equation in our synthetic case.
- The direct sampling code DeeSse (Straubhaar, 2017) for generating the categorical multiple-point simulations in our bimodal case of log hydraulic conductivity fields, and more importantly in our discrete MPS case with the multi-resolution approach based on the Gaussian pyramid technique (Straubhaar et al., 2020).

1.3 Structure of the thesis

We start in Chapter 2 by introducing the ensemble data assimilation framework which was more particularly considered in this thesis for the inverse modelling of parameter fields.

Chapter 3 addresses the application of the iterative ensemble smoother method in a multivariate Gaussian setting which is considered to be favorable for the method. For this first case, two conceptually different algorithms, but developed for the same purpose, were compared in order to get a better sense of the capability and existing variations of performance of the method.

Chapter 4 investigates the potential of an approach to handle non-Gaussian parameter and state variables which was proposed for the application of the ensemble Kalman filter (EnKF) to an inverse problem in

hydrogeology (Zhou et al., 2011). The approach is based on a nonlinear normal-score transform of the variables which has been used extensively in geosciences (Bertino et al., 2003; Gu and Oliver, 2006; Zhao et al., 2008; Simon and Bertino, 2009; Béal et al., 2010; Zhou et al., 2011; Li et al., 2012; Schoniger et al., 2012; Crestani et al., 2013; Zovi et al., 2017). By using this approach with an iterative ensemble smoother algorithm, we aim to address the non-Gaussian parameter case with an example of bimodal distribution of log hydraulic conductivities. Although this case is considered continuous in regards to the log hydraulic conductivity values, it addresses implicitly the discrete case since the underlying patterns were generated initially with a discrete geostatistical simulation technique. Will the iterative ensemble smoother succeed in recovering those discrete patterns?

Chapter 5 aims to learn from an example of application of ensemble methods in the discrete case based on the truncated Gaussian simulation technique. As shown by Liu and Oliver (2005), a parameterization based on the multi-Gaussian underlying variables is particularly appropriate in that case. However, the proposed parameterization may affect the reliability of the estimates as a result of the increased nonlinearity between the defined parameters and the observations. Through two simple truncated Gaussian and pluri-Gaussian examples, the nonlinear effects of the parameterization on the accuracy of the estimated uncertainty and the quality of the data match are discussed.

Chapter 6 presents a new methodology for the conditioning of categorical realizations generated by multiple-point statistics with an ensemble Kalman method. Similarly to the application with truncated Gaussian realizations, it relies on a parameterization based on a set of continuous underlying variables which control the categorical simulation. However, because the control variables are not multivariate Gaussian, additional steps to mitigate the effects of the non-multi-Gaussian prior ensemble are also integrated in the data assimilation procedure. This chapter constitutes the base of a paper that was accepted for publication in *Water Resources Research*¹.

Finally, we summarize the main contributions of this thesis and give some perspectives for future research in Chapter 7.

¹Lam, D.-T., Renard, P., Straubhaar, J., Kerrou, J. (2020). Multi-resolution approach to condition categorical multiple-point realizations to dynamic data with iterative ensemble smoothing. *Water Resources Research*, 56, 1-29. <https://doi.org/10.1029/2019WR025875>

Chapter 2

Background on ensemble-based data assimilation

2.1 A brief statement of the situation

In hydrogeology, it is often observed that the largest source of uncertainty for the flow model is associated with the poorly known aquifer hydraulic properties. The accuracy of the flow predictions will hence mostly depend on the accuracy of the defined aquifer hydraulic parameters. As a result, the inverse problem (De Marsily et al., 1999; Le Ravalec, 2005; Carrera et al., 2005; Oliver et al., 2008; Zhou et al., 2014) usually aims to use available state observations (e.g., hydraulic heads, flow rates, concentrations) in order to improve the identification of aquifer property values of the numerical model (e.g. hydraulic conductivities, specific storages).

The inverse problem was often formulated as a minimization problem. In this framework, the problem then consists in finding the set of parameters which gives the best match between the model output and available flow data. Such an optimization problem is usually solved by first defining an appropriate cost function and then by solving for the minimum using a gradient method. The difficulties however arise when the relationship between the parameters and observations is nonlinear. In this situation, the cost function is likely to contain local minima and the global minimum may be difficult to find. For this reason, these methods may not be suitable for the conditioning of discrete parameters since their relationship with the flow data can be strongly nonlinear

(Oliver and Chen, 2011). Inverse methods based on minimization methods are often combined with parameterization techniques allowing to reduce the number of parameters, thereby mitigating the ill-posedness of the inverse problem (Zhou et al., 2014). A general drawback of inverse methods based on minimization is that, unless performed in a stochastic framework using an ensemble of random realizations of parameter fields, they will lead to a single solution and hence limit the possibility of quantifying an uncertainty associated to that estimation (Hendricks Franssen et al., 2009; Zhou et al., 2014).

An alternative is to formulate the inverse problem in a Bayesian framework and use sampling methods, such as the rejection sampling or Markov Chain Monte Carlo methods. These methods aim to draw samples from the posterior probability distribution, i.e. the parameter distribution conditioned by all the prior information about plausible parameter values and the available state information (Linde et al., 2015). The advantage of sampling methods over optimization methods is that they can handle more complex parameter distributions which may be in closer agreement with the observed reality. Also, they can allow a more rigorous uncertainty quantification thanks to the Bayesian framework. However, a main drawback of this type of approaches is that their computational cost may be too high for practical applications when a large number of parameters is considered (Oliver and Chen, 2011).

Still in the Bayesian framework, another type of methods is the sequential Monte Carlo methods also known as particle filters. These methods were designed to give an approximate solution of the recursive Bayesian estimation problem using sequential importance sampling (Gustafsson, 2010). Although such methods can approximate any type of distribution, the number of samples to obtain a good approximation of the posterior distribution grows exponentially with the size of the problem (van Leeuwen, 1998). Similarly to particle filters, the ensemble Kalman filter (EnKF) (Evensen, 1994) is also a sequential Monte Carlo technique. But, unlike particle filters, EnKF assumes a linear Gaussian state-space model for the computation of the update of the samples used to approximate the posterior distribution. Despite this restrictive multivariate Gaussian assumption, EnKF methods have the ability to handle very high-dimensional space and in fact also nonlinear estimation problems (Evensen, 2009b). In the remaining of this chapter, we will focus more particularly on this type of ensemble methods. Although they were originally developed for state estimation problems, they are currently particularly applied in the context of inverse modelling in var-

ious fields of geosciences (Moradkhani et al., 2005; Chen and Zhang, 2006; Gu and Oliver, 2006; Hendricks Franssen and Kinzelbach, 2008; Evensen, 2009b; Zhou et al., 2011; Oliver and Chen, 2011).

2.2 Background on ensemble-based data assimilation

This section aims to give some background on Kalman filtering for the state estimation problem before introducing more particularly the ensemble Kalman filter (EnKF) of Evensen (1994). This sequential data assimilation method was designed to address high-dimensional and possibly nonlinear estimation problems, and has since led or inspired the derivation of a number of other ensemble methods including the iterative ensemble smoothers that we will discuss more particularly in the next chapters.

2.2.1 Understanding the Kalman filter

In estimation theory, Kalman filtering (Kalman, 1960) is a technique commonly used to estimate the current state of a linear dynamical system from uncertain observations. Given an imperfect model predicting the evolution of the state of the system and available measurements, the Kalman filter computes an update of the predicted state in order to produce an optimal estimate of the current state (Figure 2.1). The Kalman filter is recursive, as it will sequentially process the measurements collected at each time step. Using the result from the previous iteration only, the algorithm iteratively repeats itself for each new measurements.

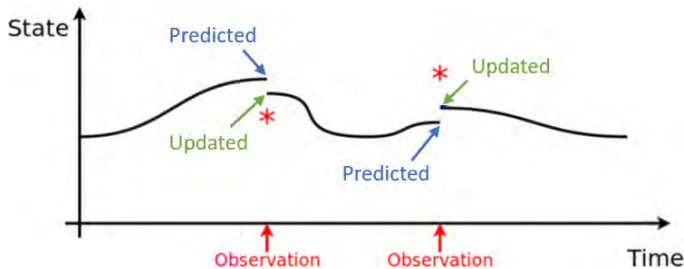


Figure 2.1: Illustration of the recursive prediction and update steps in Kalman filtering.

Let us consider at a time step h , when some measurements (y_h) are available, the vector of states (x_h^f) predicted by the model, and its explicit error covariance matrix (P_h^f) describing the magnitude of the uncertainty of each variable on its diagonal and the covariance between each pair off its diagonal. The Kalman filter computes the weights which will be used to linearly combine the predicted estimate with the measurement vector in order to produce an updated estimate (x_h^a) (Figure 2.2). These weights, or the so-called “Kalman gain”, will minimize the updated error covariance or equivalently the mean squared estimation error. They can be derived analytically if we consider that the model is linear and that all the errors are multi-Gaussian. The weights are based on the ratio between the error covariance of the predicted estimate (P_h^f) and the error covariance of the measurement vector (R_h). If the Kalman gain is large, it means that the uncertainty of the measurements is smaller than the uncertainty of the predicted estimate, therefore the measurement’s weight will be high and the predicted estimate’s will be low. Inversely, if the Kalman gain is small, it means that the uncertainty of the measurements is higher than the uncertainty of the predicted estimate, hence the measurement’s weight will be low and the predicted estimate’s will be high. After the state vector is updated, the Kalman filter updates the error covariance in order to reflect the new information from the measurements which have been assimilated.

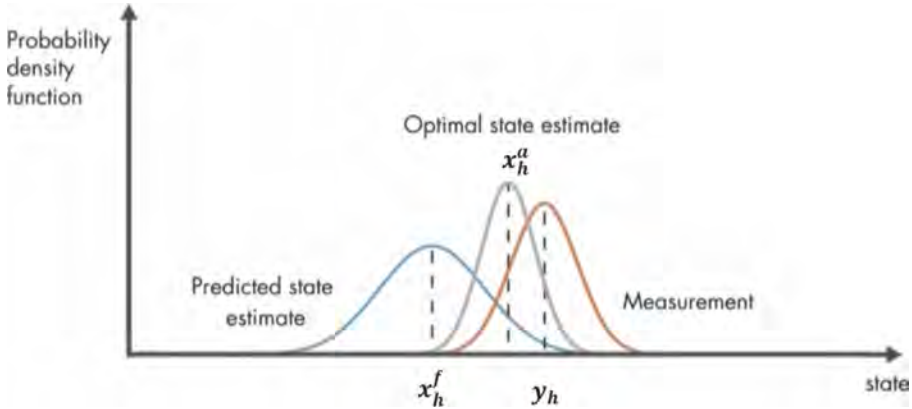


Figure 2.2: Illustration of the Kalman filter update step in one dimension at a time step h . Assuming a linear model and Gaussian errors, the new state x_h^a is estimated by linearly combining the predicted state x_h^f and the measurement y_h .

2.2.2 The multivariate Kalman filtering equations

Suppose that the state vector x_{h-1}^a and its error covariance matrix P_{h-1}^a are known at time step $h - 1$.

During the prediction step, the Kalman filter predicts the state vector from time step $h - 1$ to time step h according to the solution of the dynamic model describing the system

$$x_h^f = M_{h-1}x_{h-1}^a \quad (2.1)$$

where M_{h-1} is the state transition matrix relating the previous state x_{h-1}^a to the current state x_h^f . The associated covariance matrix is propagated as

$$P_h^f = M_{h-1}P_{h-1}^aM_{h-1}^T + Q_h \quad (2.2)$$

where Q_h is the covariance matrix of the model error.

If observations y_h are available at time step h , the Kalman filter performs an update of the predicted results x_h^f and P_h^f by computing an updated estimate x_h^a (also known as the “analysis”) and its updated error covariance P_h^a as follows

$$x_h^a = x_h^f + K_h(y_h - H_hx_h^f) \quad (2.3)$$

$$P_h^a = (I - K_hH_h)P_h^f \quad (2.4)$$

where H is a linear observation operator which maps the state space into the observation space and K a gain matrix also known as the Kalman gain matrix. K is derived so as to minimize the variance of the estimation error

$$K_h = P_h^f H_h^T (H_h P_h^f H_h^T + R_h)^{-1} \quad (2.5)$$

where R_h denotes the measurement error covariance matrix. Once the Kalman gain is calculated, the updated estimate x_h^a is obtained by combining the measurements y_h with the predicted estimate x_h^f . The estimation error is decreased after the update, as expressed in the updated error covariance P^a .

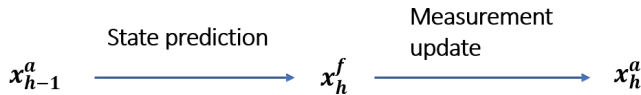


Figure 2.3: Overview of the two stages of Kalman filtering

The Kalman filter optimal estimation is derived in the linear case by implicitly assuming that all the errors statistics, i.e. for the state, the model and the measurements, are multivariate Gaussian. It is worth noting that in such a case, the updated state mean and updated error covariance calculated by the Kalman filter is the same as the estimation that would be obtained by Bayesian inference.

In practice, the Kalman filter (KF) is also used in nonlinear cases, namely when the model and/or the observation operator is nonlinear. However, optimality is lost in this “Extended Kalman filter” (EKF) as the probability density function of the state is probably not multivariate Gaussian anymore even if the initial errors were.

If M is nonlinear, the analysis step is the same as in KF while the forecast step uses the nonlinear model

$$x_{h+1}^f = \mathcal{M}_h(x_h^a) \quad (2.6)$$

The main issue comes from the propagation of the error covariance according to (2.2) which now relies on the linearization, and hence approximation, of the model with the Jacobian of the model evaluated at x^a

$$M_h = \left. \frac{d\mathcal{M}_h}{dx} \right|_{x_h^a} \quad (2.7)$$

If the observation operator is nonlinear, the nonlinear operator is used in the analysis step

$$x_h^a = x_h^f + K_h(y_h - \mathcal{H}_h(x_h^f)) \quad (2.8)$$

and the linearized observation operator

$$H_h = \left. \frac{d\mathcal{H}_h}{dx} \right|_{x_h^f} \quad (2.9)$$

is used to evaluate the Kalman gain (2.5) and the updated error covariance (2.4).

2.2.3 The ensemble strategy

One main limitation of Kalman filtering is the explicit computation and storage of the error covariance matrix of size $n \times n$ with n the size of the state vector. As a result, the Kalman filter is computationally intractable for high-dimensional systems as often encountered in geosciences.

To address high-dimensional estimation problems, an ensemble approximation of the Kalman filter named the ensemble Kalman filter (EnKF) was proposed by Evensen (1994). Unlike the Kalman filter, EnKF does not predict (nor update) directly the n -dimensional mean state and the $n \times n$ error covariance. Instead, it propagates $N < n$ realizations $x_0^{(i)f}$ with $i = 1, \dots, N$ and computes an approximative covariance matrix using the ensemble of realizations (Figure 2.4).

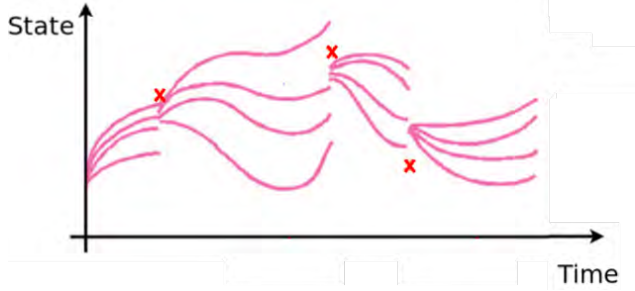


Figure 2.4: Schematic of the ensemble Kalman filter (EnKF). The lines represent individual ensemble members and the red crosses indicate observations.

In this sense, EnKF can be considered as a Monte Carlo implementation of the Kalman filter. The covariance is estimated from the ensemble as follows

$$x_h^f = \frac{1}{N} \sum_{i=1}^N x_h^{(i)f} \quad (2.10)$$

$$P_h^f = \frac{1}{N-1} \sum_{i=1}^N (x_h^{(i)f} - x_h^f)(x_h^{(i)f} - x_h^f)^T \quad (2.11)$$

In this manner, the explicit computation of the error covariance matrix of the Kalman filter is replaced by a low rank approximation which now only requires the storage of $n \times N$ values. Besides, according to how the error covariance is approximated in (2.11), the ensemble mean is considered to be the best estimate of the true state and the spreading of the ensemble around the mean provides an estimation of the uncertainty on the mean estimate (Evensen, 1994).

If observations y_h are available at time step h , then each predicted state (or forecast) of the ensemble is updated according to the same

Kalman filter equation used to update the mean forecast (2.3),

$$x_h^{(i)a} = x_h^{(i)f} + K_h(y_h - \mathcal{H}_h(x_h^{(i)f})) \quad (2.12)$$

Instead of the Kalman gain computed using the full covariance in the Kalman filter (2.5), an approximation of the Kalman gain is here calculated based on the approximated covariance P^f . As a matter of fact, P^f is never directly calculated in practice. The two following terms required in the Kalman gain are approximated from the ensemble (Evensen, 2003) as follows

$$P_h^f H_h^T = \frac{1}{N-1} \sum_{i=1}^N (x_h^{(i)f} - x_h^f) \left(\mathcal{H}_h(x_h^{(i)f}) - \frac{1}{N} \sum_{i=1}^N \mathcal{H}_h(x_h^{(i)f}) \right)^T \quad (2.13)$$

$$H_h P_h^f H_h^T = \frac{1}{N-1} \sum_{i=1}^N \left(\mathcal{H}_h(x_h^{(i)f}) - \frac{1}{N} \sum_{i=1}^N \mathcal{H}_h(x_h^{(i)f}) \right) \times \left(\mathcal{H}_h(x_h^{(i)f}) - \frac{1}{N} \sum_{i=1}^N \mathcal{H}_h(x_h^{(i)f}) \right)^T \quad (2.14)$$

Hence the linearized observation operator used in the extended Kalman filter, and which as a matter of fact causes instabilities in the propagation of the error, is not needed here.

The forecast or prediction to time step $h+1$ is then applied as

$$x_{h+1}^{(i)f} = \mathcal{M}_h(x_h^{(i)a}) \quad (2.15)$$

In the EnKF algorithm proposed by Evensen (1994), also known as the stochastic EnKF, each assimilated measurement y_h is in fact perturbed by a random noise vector $\epsilon_h^{(i)} \sim \mathcal{N}(0, R_h)$ for each ensemble member in the update step. The rationale for those perturbed observations is that the variance of the ensemble of updated EnKF estimates would be otherwise underestimated compared to the updated covariance calculated by the Kalman filter (Houtekamer and Mitchell, 1998).

2.2.4 Inverse modelling with the ensemble Kalman filter

Although the Kalman filter and its ensemble approximation are mainly used to estimate state variables which evolve in time, it can

also be used to estimate static input parameters m to a dynamic model in addition or instead of the states. A common way of introducing the parameter estimation problem with EnKF is with the so-called “augmented state” approach for the combined estimation of states and parameters. Each augmented state vector now also includes a sampled set of parameters $m^{(i)f}$

$$X^{(i)f} = \begin{bmatrix} x^{(i)f} \\ m^{(i)f} \end{bmatrix} \quad (2.16)$$

The analysis or update step can hence be expressed in terms of the augmented state vector at time step h

$$X_h^{(i)a} = X_h^{(i)f} + C_{X,h}^f H^T (H_h C_{X,h}^f H_h^T + R_h)^{-1} (y_h + \epsilon^{(i)} - H_h X_h^{(i)f}) \quad (2.17)$$

where H is a linear observation operator relating the augmented state vector to the observations. Unlike in the previous linear Gaussian state case, it contains more zeros because of the addition of the parameters in the state vector. C_X corresponds to the augmented covariance associated to the augmented state vector which can be approximated by

$$C_X = \begin{bmatrix} C_{xx} & C_{mx}^T \\ C_{mx} & C_{mm} \end{bmatrix} \quad (2.18)$$

where C_{xx} , C_{mx} , C_{mm} are respectively the auto-covariance of the state variables, the cross-covariance between the model parameters and the states, and the auto-covariance of the parameters. It is worth mentioning that if the operator relating the parameters to the state vector is nonlinear, then this augmented covariance implies a linearization of the operator at the ensemble mean $\bar{m}$ (Gu and Oliver, 2007).

Ultimately as $H = [H_x \ 0]$, the update of each of the subvectors of state variables and model parameters can be written

$$\begin{bmatrix} x^{(i)a} \\ m^{(i)a} \end{bmatrix} = \begin{bmatrix} x^{(i)f} \\ m^{(i)f} \end{bmatrix} + \begin{bmatrix} C_{xx}^f H_x^T (H_x C_{xx}^f H_x^T + R)^{-1} \times (y + \epsilon^{(i)} - H_x x^{(i)f}) \\ C_{mx}^f H_x^T (H_x C_{xx}^f H_x^T + R)^{-1} \times (y + \epsilon^{(i)} - H_x x^{(i)f}) \end{bmatrix} \quad (2.19)$$

Alternatively, the update equation of the parameters can also be equivalently expressed in terms of a model operator g which predicts the observations d from the parameters m , and hence includes the measurement operator H , as follows

$$m^{(i)a} = m^{(i)f} + C_{md}^f (C_{dd}^f + R)^{-1} (y_h + \epsilon^{(i)} - g(m^{(i)f})) \quad (2.20)$$

A prediction or forecast step to time step $h + 1$ then follows. Since the parameters do not evolve in time according to some dynamics, only the states are updated given the updated states and parameters

$$\begin{bmatrix} x_{h+1}^{(i)f} \\ m_{h+1}^{(i)f} \end{bmatrix} = \begin{bmatrix} \mathcal{M}_h(x_h^{(i)a}, m_h^{(i)a}) \\ m_h^{(i)a} \end{bmatrix} \quad (2.21)$$

Although the joint estimation of both states and parameters is possible using this augmented state approach, it can result in physical inconsistencies between the updated states and the updated parameters when the relationship is nonlinear, and also because the degree of freedom in the system is increased (Moradkhani et al., 2005; Gu and Oliver, 2007; Chen et al., 2009). As an alternative, (Moradkhani et al., 2005) proposed for a hydrological application a dual EnKF approach where parameters and states are updated one after the other at each iteration. Thus, only the parameters are first updated with EnKF by assimilation of some current data, and the updated parameters are used to predict, restarting from the old state calculated from the previous assimilation, the current states which are in turn updated with EnKF using again the same data. Another option to completely avoid this consistency issue is to update only the parameters when assimilating the measurements at a time step, and then perform the prediction step using the updated parameters from time zero until the new assimilation time step in order to obtain the predicted states (Gu and Oliver, 2007).

Moreover, the nonlinearity between the parameters and states can yield estimates which do not sufficiently honor the data (Gu and Oliver, 2007). To improve the accuracy of the estimates obtained with EnKF, iterative updating schemes of ensemble methods, such as the iterative ensemble Kalman filter originally proposed by Gu and Oliver (2007) with the sequential EnRML algorithm, have been developed among other iterative methods (Chen and Oliver, 2012; Emerick and Reynolds, 2012b,a; Chen and Oliver, 2013; Bocquet and Sakov, 2014).

2.2.5 Practical issues with ensemble methods

The low rank approximation of the error covariance matrix which actually makes ensemble Kalman methods feasible for high-dimensional problems is also known to introduce sampling errors. Although these sampling errors will decrease proportionally to $1/\sqrt{N}$ with N the ensemble size (Evensen, 2003), the size of the ensemble used is generally restricted by the computational cost. For high-dimensional systems, the

limited number of ensemble members, usually in the order of 10^2 , can hence be smaller than the number of variables to estimate. However, as pointed by Evensen (2009a), the degree of freedom of the state or parameter space can be much less than the actual number of gridblocks of the model as the variables may not be independent from each other given the prior correlations. Still, in many applications of ensemble methods, the ensemble may not span the model subspace adequately and the resulting under-sampled system will likely cause problems that will affect the performance of the ensemble method. That is mainly why extra measures are almost always required in real applications in order to obtain an acceptable performance of ensemble Kalman methods. Hereinafter, we give an overview of the issues associated to the under-sampling from a limited number of ensemble members, and some practical measures commonly employed to mitigate them. We then conclude with some computational aspects associated to the application of ensemble methods.

Issue n° 1: Inbreeding

Inbreeding is a term used by Houtekamer and Mitchell (1998) to describe the systematic underestimation of the uncertainty observed as a result of using the same ensemble to both approximate the Kalman gain and to estimate the error using that gain. The ensemble members are becoming more and more coupled after each update step, thereby potentially underestimating the uncertainty. van Leeuwen (1998) has however argued that this observed inbreeding was also due to the under-sampling resulting from the use of small ensembles. In the end, the combined effects of the under-sampling and the coupling effects from the Kalman gain will decrease with the increase of the ensemble size.

Issue n° 2: Spurious correlations

Because of the limited size N of the ensemble, the ensemble can only transport limited information with an approximated sampled error covariance of at most rank $N - 1$. As a result, existing zero entries of the full error covariance matrix are difficult to reproduce and spurious correlations can appear in the sampled covariance. In particular, components of the state vector distant from observations can be erroneously updated as a result of an error in the covariance which increases with the distance (Hamill et al., 2001). Although the spurious updates may cancel out and the drift in the estimated mean may be negligible,

the spurious updates will over time or over the iterations lead to an underestimation of the uncertainty (Evensen, 2009a).

Mitigation methods

Localization and inflation are extra measures of many applications of EnKF methods. These ad hoc methods have been mainly developed in atmospheric and oceanographic applications with the aim of mitigating the forementioned consequences of using small ensembles.

Localization aims to address the issue of spurious correlations. The idea of localization is to update only a subset of variables of the state vector with each measurement. In other words, localization allows the reduction of the dimension of the update step. There are in fact two classes of localization methods that are currently used, namely covariance localization and local updating (Evensen, 2009a). Both relies on prior knowledge of the true covariance structure. In particular, given the underlying dynamics in geo-science applications, a zero correlation is often assumed between two distant state components or between one state component and a distant observation.

Covariance localization consists in postprocessing the noisy $n \times n$ ensemble-based covariance C_X^f , and hence the resulting Kalman gain, prior to the computation of each update. The covariance is multiplied element by element with a distance-dependent correlation function which varies from 1.0 at the variable location to 0 at some predefined radial distance (Figure 2.5).

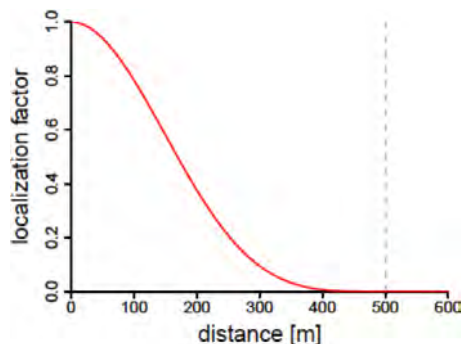


Figure 2.5: Example of a distance-based covariance localization function.

Emerick and Reynolds (2011) suggested setting this “critical distance”

based on the principle correlation lengths of the model and sensitivity matrices. A necessary condition on the chosen correlation function is that the resulting $n \times n$ localization matrix, noted ρ , should be positive definite in order to ensure a valid localized covariance $C_{X,loc}^f = \rho \circ C_X^f$ (where $\circ$ represents here the element-wise multiplication or Schur product). A standard choice of localization function is the Gaussian-like fifth order polynomial with compact support of Gaspari and Cohn (1999). Once the covariance is localized or “tapered”, the long range spurious correlations are effectively truncated in $C_{X,loc}^f$ and the resulting Kalman gain

$$K_h = \left[(\rho \circ C_{X,h}^f) H^T \right] \left[H_h (\rho \circ C_{X,h}^f) H_h^T + R_h \right]^{-1} \quad (2.22)$$

will allow to substantially reduce the errors in the updated estimates (Figure 2.6).

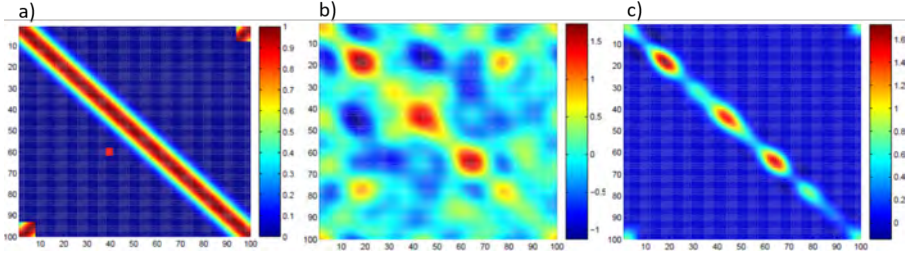


Figure 2.6: Illustration from Petrie and Dance (2010) of the motivation behind covariance localization. In order to better approximate the true covariance matrix P_X^f (a), the noisy spurious covariances in the sample covariance C_X^f (b) are filtered from the localized covariance $C_{X,loc}^f$ (c).

As a matter of fact, there are other variants of implementation of covariance localization (Chen and Oliver, 2017). For example, the following formulation of the Kalman gain

$$K_h = \left[\rho \circ (C_{X,h}^f H^T) \right] \left[\rho \circ (H_h C_{X,h}^f H_h^T) + R_h \right]^{-1} \quad (2.23)$$

has been proposed and allows to replace the computation of Schur products between $n \times n$ matrices with less expensive products of matrices. For the case of parameter estimation according to (2.19) or (2.20), the first Schur product in (2.23) will involve a matrix ρ_{md} of size $N_m \times N_d$, with N_m and N_d the number of parameters and the number of observations respectively, while the second Schur product will involve a matrix

ρ_{dd} of size $N_d \times N_d$. As mentioned in Chen and Oliver (2017), another option for covariance localization is to apply the localization on the whole Kalman gain using ρ_{md} in (2.20).

Local updating is a localization technique particularly used in atmospheric and oceanography applications. It allows to perform the update gridblock by gridblock by assimilating only a subset of observations located in the vicinity of the gridblock to be updated (Figure 2.7). In this manner, the contribution of distant observations which are likely to be dominated with noise are removed (Chen and Oliver, 2017). However, this range of influence requires tuning as it should be large enough to include the relevant measurements, but small enough to remove the spurious impact of remote observations.

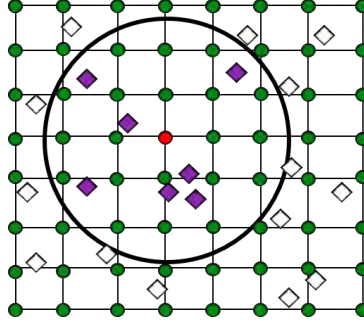


Figure 2.7: Illustration of a local update performed at the red node. Only observation within a critical distance are used to compute the update.

In practice, each gridblock is associated to a subset of variables of the state vector to be updated using the same subset of data. The resulting local state vectors can be of smaller dimension than the ensemble size N , thereby mitigating the issues related to the update of a state vector of larger size than N in high-dimensional problems. Another advantage of local updating compared to covariance localization is that it allows the updated ensemble to reach solutions that were not originally spanned by the ensemble (Evensen, 2009a) thanks to the different local Kalman gain matrices, and hence different combination of ensemble members for each local update (Chen and Oliver, 2017). Note that, these local Kalman gain matrices can also be tapered using a correlation localization function as used for covariance localization. For more details about the implementation of local updating and a comparison

of localization methods, the reader is referred to Chen and Oliver (2017).

Covariance inflation is a method used to counteract the tendency of ensemble Kalman methods to underestimate the uncertainty because of either undersampling, inbreeding, or spurious correlations. The spread of the ensemble is artificially increased before the assimilation of the observations, the most frequently according to

$$X_h^{(i)f} \leftarrow r \left(X_h^{(i)f} - \frac{1}{N} \sum_{i=1}^N X_h^{(i)f} \right) + \frac{1}{N} \sum_{i=1}^N X_h^{(i)f} \quad (2.24)$$

where r is factor slightly greater than 1.0 and the left arrow denotes the replacement of the previous value $X_h^{(i)f}$. By improving the stability of the ensemble method, the application of covariance inflation has been shown to improve the accuracy of the estimates and also to account partly for model errors. It was however noted that covariance inflation does not change the subspace spanned by the ensemble, and hence this method may not be effective in order to account for model errors which mostly project into a very different subspace (Hamill, 2006). Moreover, its application is not straightforward in models where observations are not evenly distributed. Indeed, in regions with nearly nonexistent data, inflation will lead to an unwanted unbounded increase of the ensemble variance because there will be nearly no reduction of the spread. More generally, the use of an artificial inflation of the ensemble spread, and hence uncertainty, will prevent an analysis of the uncertainty reduction which would have resulted from the data assimilation.

Computational cost

The computational cost of ensemble methods depends mainly on the cost of the N forward model runs used to predict the state variables for all N ensemble members before the new update. As mentioned in 2.2.4, this duration can be especially long in the case of nonlinear parameter estimation where it might be necessary to perform the N simulations from the beginning. As a result, the computational cost will also depend on the size N of the ensemble. However, because these simulations are independent from each other, they can be performed in parallel instead of one at a time. We remind that the update step cannot be performed before the prediction step is complete for all ensemble members because the required Kalman gain is calculated based on ensemble-based covariances.

During the update step, the computational cost will particularly depend on the number of observations to assimilate given the inversion of the matrix $C = H_h P_h^f H_h^T + R_h$ in the Kalman gain. Because in EnKF the number of data is usually small at each time point, a pseudo-inverse of C using a truncated singular value decomposition (SVD) can be computed efficiently. However, when the number of observations N_d is very large ($N_d \gg N$), as is often the case with ensemble smoother methods, it is computationally more efficient to calculate the pseudo-inverse according to the subspace inversion procedure of Evensen (2004). In order to better take into account data of different magnitudes, Emerick and Reynolds suggested in these procedures to rescale the components of the matrix prior to the calculation of the truncated SVD (Emerick and Reynolds, 2012b).

Chapter 3

Conditioning multi-Gaussian groundwater flow parameters to transient hydraulic head and flowrate data with iterative ensemble smoothers: a synthetic case study

3.1 Introduction

Since the ensemble Kalman filter (EnKF) (Evensen, 1994) has been introduced as a computationally efficient Monte Carlo approximation of the Kalman filter (Kalman, 1960; Anderson, 2003), ensemble methods for data assimilation have been widely used for high-dimensional estimation problems in geosciences (Evensen, 2009b). In all these methods, an initial ensemble of realizations which should capture the initial uncertainty of the state or parameter variables of interest is first generated. Then, thanks to the assimilation of available uncertain observations, an updated ensemble of realizations that are conditioned by the observations is obtained. However, a main limitation is that ensemble Kalman methods assume multivariate Gaussian error statistics in the distributions involved in the computation of the update. As a result, departures from this multi-Gaussian assumption can lead to an important loss of optimality in the estimated ensemble mean and variance.

EnKF is for example extensively applied in meteorology in order to estimate the current state of the atmosphere in real time (Anderson, 2009). In such atmospheric applications, newly obtained observations of the atmosphere are assimilated sequentially in order to update the initial conditions of weather predictions models. In reservoir modelling, ensemble methods have also become a standard tool with the use, more particularly, of smoother algorithms for inverse modelling (Evensen, 2018). Still with the aim of improving model forecasts, time series of state observations collected during the production of a reservoir are with smoother methods processed all simultaneously in order to update the static parameters of reservoir simulation models.

EnKF has also been applied in surface and groundwater hydrology for the estimation of both the parameter and state variables of a system (Moradkhani et al., 2005; Hendricks Franssen and Kinzelbach, 2008). In particular, it has been shown that the increased degree of freedom introduced by the larger number of unknowns can make the estimation of EnKF particularly unstable, especially in the presence of nonlinear dynamics (Moradkhani et al., 2005). When the problem is nonlinear, such joint estimation can also result in inconsistent predicted data after the update and physical inconsistencies between the updated states and parameters (Gu and Oliver, 2007; Chen et al., 2009). These issues have particularly motivated the development of iterative EnKF methods based on the iterative minimization of a cost function for each iteration of the standard EnKF (Gu and Oliver, 2007; Emerick and Reynolds, 2012b).

Because of the need of restarting the dynamic model multiple times in the context of nonlinear parameter estimation with EnKF, the simultaneous assimilation of all the data set in the ensemble smoother method (van Leeuwen and Evensen, 1996) has been considered a suitable alternative to EnKF in reservoir applications. Instead of having to update the variables at each assimilation time step, the ensemble smoother can process all the data of the time series in one single update step. Similarly to the iterative EnKF, successive updates can also be applied using iterative forms of the ensemble smoother in order to improve the data fit in nonlinear problems (Chen and Oliver, 2012; Emerick and Reynolds, 2012a).

This chapter focuses on the performance of two existing iterative forms of ensemble smoother for a synthetic groundwater flow application. Although the model considered is synthetic, it is inspired by

the real hydraulic perturbation observed at the Andra's Meuse/Haute-Marne site since the construction and operation of the Underground Research Laboratory (Kerrou et al., 2017). In particular, the problem considers a transient flow induced by one vertical shaft, and transient observation points at points located in a restricted region of the model.

The main objective of this synthetic study will be to assess the benefit of assimilating different types of flow data for the parameter identification of multi-Gaussian log hydraulic conductivity ($\log K$) fields. The effects of increasing the ensemble size using both methods will also be compared. Similarly to the data transformation proposed by Schöniger et al. for an EnKF application (Schöniger et al., 2012), the benefit of a normal-score transform on the state variables prior to the update will also be assessed for one of the tested algorithms.

Hereinafter, we first present in Section 3.2 the ensemble smoother and the two iterative ensemble smoother considered in this study: LM-EnRML (Chen and Oliver, 2013) and ES-MDA (Emerick and Reynolds, 2012a). Both are the main iterative variants currently used for inverse modelling in reservoir applications (Evensen, 2018). The synthetic case including the model set up, the initial ensemble, and the performance criteria is presented in Section 3.3. Finally, the results of the analysis of the synthetic case are discussed in Section 3.5.

3.2 General background on the ensemble smoother and iterative

The ensemble smoother (ES) introduced by van Leeuwen and Evensen (1996) is an extension of the ensemble Kalman filter. Both are similar in that a set of N realizations $\{m_i^{pr}, \dots, m_N^{pr}\}$ is used to represent a presumed multi-Gaussian distribution, and is updated by the assimilation of measurements in order to form a new conditional distribution. When using either ES or EnKF for the inverse modelling of parameters based on observations of state variables, each conditioned parameter realization m_i^{post} of the ensemble is calculated from the unconditioned realization m_i^{pr} according to the following equation

$$m_i^{post} = m_i^{pr} + K(d_{obs,i} - g(m_i^{pr})) \quad (3.1)$$

with

$$K = C_{MD}^{pr}(C_{DD}^{pr} + C_{err})^{-1}$$

The matrix K represents an approximation of the so-called ‘‘Kalman gain’’ in the Kalman filter update equation, derived so as to minimize

the error covariance of the posterior estimate. It is here computed based on approximations from the ensemble of the cross-covariance matrix between the vector of prior parameters and the vector of predicted data, noted C_{MD}^{pr} , the auto-covariance matrix of predicted data C_{DD}^{pr} , and the covariance matrix of observed data measurement errors C_{err} . By evaluating the relative uncertainty of the measurements and prior estimate, the Kalman gain weights the contribution of each conditioning observation relatively to the prior estimate m_i^{pr} for the computation of the update. More precisely, it weights the contribution from each component of the mismatch between the vector of perturbed observations $d_{obs,i}$, i.e. the observations corrupted with noise $z_{obs,i} \sim N(0, C_{err})$, and the corresponding vector of predicted states $g(m_i^{pr})$ using the forward operator g .

Unlike EnKF however, ES does not assimilate the data sequentially in time. Instead, it assimilates all the available observations simultaneously in a single conditioning step. Hence the prediction step in ES prior to the single update will be longer than each recursive one in EnKF since the ensemble of prior realizations need to be forwarded in time until the time of the last conditioning observation. Evensen and van Leeuwen (2002) showed that when the prior realizations are multi-Gaussian and the forward model is linear, ES and EnKF at the last data assimilation will give the same result. In this special case, they will converge to the exact solution in the Bayesian sense as the ensemble size increases to infinity (hence the subscripts *pr* and *post* used in the previous equations to denote the unconditioned and conditioned realizations respectively). In nonlinear cases, EnKF has been shown to outperform ES (Crestani et al., 2013). Indeed, the sequential processing of fewer data in EnKF effectively allows the computation of smaller updates than the single global update of ES. This fact particularly allows EnKF to better match the measurements than ES in nonlinear problems.

Even so, if the whole data set for the parameter estimation is already acquired, the assimilation with ES of the whole set of data in a global update step may seem more convenient to implement. Indeed, the additional computations of intermediate conditional ensembles over time with EnKF can be avoided. For nonlinear problems, iterative versions of the ES have been especially developed in order to improve the insufficient data match obtained with ES. Like ES, these iterative variants assimilate the complete data set during the conditioning step. However, the assimilation is performed multiple times on the same data set in order to reach the final solution. In the following sections, we

introduce two existing iterative ensemble smoother algorithms which are particularly used in reservoir applications (Evensen, 2018), namely the simplified version of the the Levenberg-Marquardt Ensemble Randomized Maximum Likelihood (LM-EnRML) of Chen and Oliver (2013) and the Ensemble smoother with Multiple Data Assimilation (ES-MDA) of Emerick and Reynolds (2012a).

3.2.1 Levenberg-Marquardt Ensemble Randomized Likelihood (LM-EnRML)

LM-EnRML is an iterative ensemble smoother based on a modified form of the Levenberg-Marquardt algorithm (Chen and Oliver, 2013). Assuming a prior multi-Gaussian distribution of realizations m_i^{pr} with $i = 1, \dots, N$, the algorithm aims to generate a posterior ensemble of N realizations m_i that each individually minimizes an objective function

$$O(m_i) = \frac{1}{2}(g(m_i) - d_{obs,i})^T C_{err}^{-1}(g(m_i) - d_{obs,i}) + \frac{1}{2}(m_i - m_i^{pr})^T C_M^{-1}(m_i - m_i^{pr}) \quad (3.2)$$

which measures the distance between m_i and the realization m_i^{pr} sampled from a prior distribution and the distance between the noisy observations $d_{obs,i}$ and the corresponding vector of predictions, noted $g(m_i)$, which results from the application of the forward operator g to m_i .

Each minimization of the ensemble of objective functions is performed iteratively, so that for each ensemble member, the updated realization at iteration $k + 1$ is computed using the results of iteration k as follows

$$m_i^{k+1} = m_i^k - C_{sc}^{1/2} \Delta m_k^e V_D^{P_D} W_D^{P_D} ((1 + \lambda_k) I_{P_D} + W_D^{P_D})^{-1} \times U_D^{P_D T} C_{err}^{-1/2} (g(m_i^k) - d_{obs,i}^k) \quad (3.3)$$

where C_{err} is the covariance of measurement errors, C_{sc} is a scaling matrix for the model parameters. Δm^e represents an ensemble of deviations from the mean of the parameters vectors, computed as

$$\Delta m_k^e = -C_{sc}^{-1/2} (m_k^e - \overline{m_k^e}) / \sqrt{N - 1} \quad (3.4)$$

with m^e the ensemble of N parameter vectors and $\overline{m}^e$ the ensemble mean. As for the terms $U_D^{P_D}$, $W_D^{P_D}$ and $V_D^{P_D}$, they result from the truncated singular value decomposition, based on a number P_D of singular values, of the ensemble of deviations from the mean of the vectors of predictions, noted Δd^e , so that $\Delta d^e = U_D^{P_D} W_D^{P_D} V_D^{P_D T}$. This ensemble of predicted data deviations is calculated as

$$\Delta d_k^e = -C_D^{-1/2}(d_k^e - \overline{d}_k^e)/\sqrt{N-1} \quad (3.5)$$

with d^e the ensemble of N prediction vectors and $\overline{d}^e$ the ensemble mean. Finally, λ is the so-called Levenberg-Marquardt parameter. After each update, λ is adjusted by a constant user-defined factor according to whether the objective function is increasing or decreasing. If at the current iteration, the objective function calculated on the basis of the updated realization has increased compared to the value of the previous iteration, λ will be increased for the computation of the next update and the current realization, as well as its prediction, are rejected. Inversely, if the objective function has decreased, λ will be reduced and both the current realization and its prediction will be considered for the computation of the next update. LM-EnRML will stop iterating after a maximum number of iterations allowed is reached, or if the relative decrease of the objective function or the magnitude of the realization update fall respectively below a defined threshold.

Thanks to a modification of the Hessian term, LM-EnRML avoids the explicit computation of the sensitivity matrix of the predicted data to the model parameters using the ensemble at each iteration as in the original LM-EnRML formulation (Chen and Oliver, 2013). This allows LM-EnRML to reduce the numerical instability usually observed with the original method for large-scale problems where the ensemble size becomes smaller than the number of parameters to estimate. We will in fact use a variant of LM-EnRML referred as LM-EnRML (approx.) in Chen and Oliver (2013) which neglects the contribution to the update of the mismatch between the updated and the prior realization. Indeed, Chen and Oliver (2013) concluded that using this simpler variant did not significantly affect the results that they obtained using an ensemble of 104 realizations for their large-scale estimation problem with 165000 parameters and 4000 observations.

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

Thanks to its simple formulation, ES-MDA of Emerick and Reynolds (2012a) is perhaps the most used iterative form of the ensemble smoother in geoscience applications. The standard algorithm consists simply in repeating a predefined number of times the standard ensemble smoother (ES) (Emerick and Reynolds, 2012a). However, unlike in the ES update, the covariance of the measurement errors in ES-MDA is inflated so that each realization m_i of the ensemble of size N at iteration $k+1$ is updated as follows

$$m_i^{k+1} = m_i^k + C_{MD}^k (C_{DD}^k + \alpha_{k+1} C_{err})^{-1} (d_{obs,i}^k - g(m_i^k)) \quad (3.6)$$

where $d_{obs,i}^k = d_{obs} + \sqrt{\alpha_{k+1} C_{err}^{1/2}} z_{d,i}$ with $z_{d,i} \sim N(0, I_d)$.

The purpose of inflating the Gaussian noise, sampled at every iteration, via the coefficient $\alpha > 1$ is to limit the confidence given to the data as they will be assimilated multiple times (Emerick and Reynolds, 2012a). In so doing, the parameter covariance reduction which occurs after each data assimilation is also limited. The inflation factors used to inflate the covariance matrix of the measurement errors need to satisfy the following condition

$$\sum_{k=1}^{N_a} \frac{1}{\alpha_k} = 1 \quad (3.7)$$

where N_a is the number of times we repeat the data assimilation. Indeed, this condition has been derived in order to make the single update by ES and multiple data assimilation by ES-MDA equivalent for the linear Gaussian case. For all our tests using ES-MDA, we set for simplicity the inflation coefficients equal to the predefined number of assimilations, as varying them in a decreased order do not lead to a significant improvement of the data match (Emerick, 2016).

ES-MDA effectively improves the data fit obtained by ES in the nonlinear case because ES is in fact equivalent to one single iteration of the Gauss-Newton procedure to minimize the objective function (3.2) when using a full step length and an average sensitivity matrix calculated from the ensemble (Emerick and Reynolds, 2012a; Le et al., 2016). The motivation for applying ES-MDA in nonlinear problems is that it would be comparable to several Gauss-Newton iterations with an average sensitivity matrix which is updated after each new data assimilation. By

calculating smaller updates than one single potentially large ES update, ES-MDA is expected to lead to better results than ES.

The quality of the final data fit achieved with ES-MDA will particularly depend on the predefined number of data assimilations. The standard ES-MDA algorithm is therefore not an optimized procedure as it requires some amount of trial before finding a number of iterations which allows an acceptable match. Although we did not consider them in this synthetic study, implementations of ES-MDA which allow to adapt the inflation coefficients and the number of iterations as the history match proceeds have been proposed (Le et al., 2016; Emerick, 2016).

3.2.3 Normal-score transform of state variables with ES-MDA

Considering that the forward model g can be nonlinear due to the physical process being modelled and/or the influence of imposed boundary conditions on the predicted states, the assumption of multi-Gaussian dependence among state variables in ensemble Kalman methods is generally not justified in subsurface flow modelling even in the case of multi-Gaussian log hydraulic conductivity fields (Schöniger et al., 2012). In addition to the linearization around the ensemble mean introduced by the use of an ensemble gradient, Evensen (2018) demonstrated that the derivation of the ES update equation (3.6) particularly involves a linearization of the forward model around the local estimate m_i^k . Hence ES-MDA, which is based on the repeated application of ES using inflated measurement errors, also applies such a local linearization at each update step (Evensen, 2018).

Because of this forementioned local linearization in the update step, we are here interested in assessing the effects of a normal-score transform in order to apply the ES-MDA update to locally Gaussian state distributions. In a previous study, Schöniger et al. observed an improved performance of EnKF for the identification of multi-Gaussian log hydraulic conductivity fields when assimilating normal-score transformed state variables such as hydraulic heads (Schöniger et al., 2012). They attributed this improvement to an “implicit pseudolinearization” of the relationship between the multi-Gaussian log hydraulic fields and the transformed predicted data which benefited the “linear updating step” of EnKF. In addition to this statement, we conjecture here that it is more specifically the local linearization highlighted by Evensen (2018), rather than the linearization around the ensemble mean in-

troduced by the approximated gradient, which will benefit from such a normal-score transform performed locally at the data points. The “pseudolinearization” observed by Schoniger et al. (2012) would then allow a better approximation of the forward model from the local linearization applied at each update step, thereby yielding more accurate updates.

The effects of the ES-MDA update in the transformed space will be tested as follows

$$m_i^{k+1} = m_i^k + C_{M\hat{D}}^k (\widehat{C_{DD}^k} + \alpha_{k+1} \widehat{C_{err}})^{-1} (\widehat{d_{obs,i}^k} - \widehat{g(m_i^k)}) \quad (3.8)$$

where

$$d_{obs,i}^k = d_{obs} + \sqrt{\alpha_{k+1}} C_{err}^{1/2} z_{d,i}$$

with

$$z_{d,i} \sim N(0, I_d)$$

where the hat indicates either transformed variables or covariances calculated based on transformed variables. A normal-score transform (NST) function denoted ψ_k , for each data type, is calculated based on the time series of state variables predicted at each one of the k observation locations (Figure 3.1) which we perturb with noise beforehand. Indeed, we perturb the predicted data using the same inflated measurement errors added to the observations d_{obs} in (3.8). This NST function maps the p -quantile of the ensemble distribution of original perturbed predicted values at one location to the p -quantile of a standard normal distribution. In order to evaluate properly the data mismatch in the transformed space, the predicted data $g(m_i^k)$ without added noise and the perturbed observations $d_{obs,i}^k$ are transformed using the same NST functions defined depending on the location of the data. Note that $C_{M\hat{D}}$ is the cross-covariance between the parameters and the normal-score transform of the predicted data without added measurement errors.

To transform the term $C_{DD}^k + \alpha_{k+1} C_{err}$, Schoniger et al. (2012) suggested to calculate the covariance using directly the transformed perturbed predicted data from which we built the NST functions. In this way, the terms C_{DD}^k and C_{err} don't need to be transformed separately. However, unlike in EnKF where the inversion of the matrix $C_{DD}^k + \alpha_{k+1} C_{err}$ is fast thanks to the generally small number of data at each observation time (Liu and Oliver, 2005), the assimilation of a large data set here with ES-MDA will require the use of the subspace inversion procedure of Evensen (2004). Consequently, the calculation of the

transformed terms $\widehat{C_{DD}^k}$ and $\widehat{C_{err}}$ will be needed in any case to compute the pseudo-inverse of $\widehat{C_{DD}^k} + \alpha_{k+1}\widehat{C_{err}}$ in (3.8). In the end, $\widehat{C_{DD}^k}$ can be calculated simply on the basis of the normal-score transform of the predicted data without added measurement errors, while the transformed measurement error covariance $\widehat{C_{err}}$ will be determined so as to respect the following ratio

$$\widehat{C_{err}}(C_{err})^{-1} = \widehat{C_{D'D',diag}}(C_{D'D',diag})^{-1} \quad (3.9)$$

where C_{err} is the untransformed measurement error covariance which is often assumed diagonal, $\widehat{C_{D'D',diag}}$ is the diagonal matrix constructed from the diagonal elements of the covariance of the transformed perturbed predicted data, and $C_{D'D',diag}$ is the untransformed equivalent. It is worth pointing out that transforming properly the measurement error covariance is particularly critical in this proposed application of ES-MDA in the transformed space as the data mismatch reduction with ES-MDA will rely on the inflation of $\widehat{C_{err}}$. Moreover, to better taking into account the data of different orders of magnitude in the update, Emerick (2016) suggested using the measurement error covariance to rescale the predicted data before calculating the truncated singular value decomposition (SVD) in the subspace inversion procedure.

3.3 A synthetic inverse problem inspired by the Andra's Meuse/Haute-Marne site

A synthetic case inspired by the hydraulic situation encountered at the Andra's Meuse/Haute-Marne since the operation of the Underground Research Laboratory (Benabderrahmane et al., 2014) was created. The model represents a two-dimensional vertical cross-section of a multilayered aquifer system and is inspired by previous modelling studies on the Andra's site (Bourgeat et al., 2004; Deman et al., 2015). It will be used to analyze the performance of the ensemble smoother and the two iterative ensemble smoother algorithms presented previously for the identification of multi-Gaussian parameter fields. We present hereinafter the inverse problem set up, the assumptions for the application of the methods, and the performance criteria.

3.3.1 Model set up

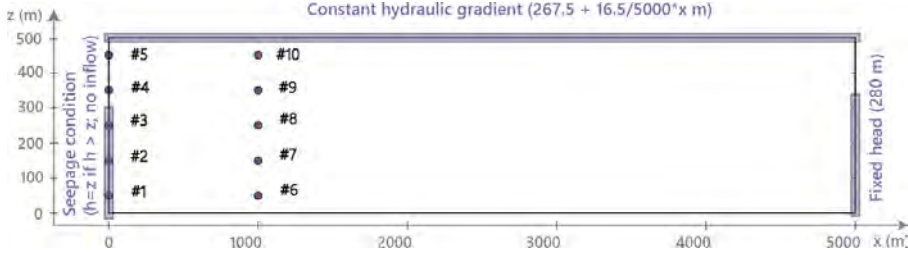


Figure 3.1: The two-dimensional vertical domain modelled of 5000 m length and 500 m depth. The thin rectangles indicates the location of nodes with a boundary condition applied. The dots indicates the head observation nodes. Not shown here, the seepage boundary is divided into 5 zones where the produced flowrates are measured.

The synthetic model is two-dimensional over a vertical domain of 5000 m along the west-east direction and 500 m in depth. In order to solve the groundwater flow equation, the domain is discretized by 50×500 square elements of 10 m wide. Along the top boundary, constant heads are imposed. The head values vary from 267.5 m on the west side to 284 m on the east side according to a small hydraulic gradient observed in the shallow aquifers which we did not represent in our model. At depth, between 150 m and 500 m, a constant head of 280 m is imposed on the east side to maintain a continuous east-west flow in the system (Figure 3.1). To mimic the hydraulic perturbation created by one vertical shaft to the underground research laboratory at the Andra's site, a subsequent transient flow simulation is calculated by adding from time zero a seepage condition over the nodes corresponding to the first 300 meters of the western boundary starting from the bottom (Figure 3.1). The nodes over the remaining 200 m towards the surface do not contribute to the production of water as they model the isolated upper part of the shaft as in the real case. Since the produced flowrates at the shaft are unknown prior to the computed head solution at each time step, using a seepage boundary condition instead of a pumping well with a prescribed extraction rate seems indeed more appropriate here. Note that in exploratory simulations, a sensitivity analysis was carried out on different levels of the finite element mesh refinement and the solver parameters, as well as on various types of flow boundary conditions and hydraulic property values. This was in order to ensure robust numerical

solutions and optimized CPU time, and also to avoid bias from boundary effects.

For simplicity, the groundwater flow equation was solved numerically under saturated conditions. All the groundwater flow simulations in the framework of this study were performed with the simulator GroundWater (GW) (Fabien Cornaton, 2014). This numerical code uses the standard Galerkin Finite Element and the Control Volume Finite Element methods, and has been validated on the basis of a series of standard benchmarks by comparison with analytical solutions as well as with commercial numerical simulators.

3.3.2 Synthetic data set

Since we intend in this synthetic study to apply ensemble Kalman methods to an initial ensemble of multi-Gaussian parameter realizations, a reference field of decimal logarithm of hydraulic conductivities ($\log K$) was generated with a multi-Gaussian simulation technique (Figure 3.2). To generate the synthetic data set of the inverse problem, a steady state groundwater flow was first simulated prior to the activation of the seepage boundary condition. The computed heads were then used as initial conditions for a subsequent transient flow simulation after the activation of the seepage condition at time zero. A constant specific storage value of 10^{-6} m^{-1} was assumed over the whole domain as they will not be considered in the inverse modelling. In the end, the data set for the inversion is composed of simulated heads collected every 1200 s from $t = 0$ to $t = 43\,200 \text{ s}$, hence at 37 time steps, at the ten observation points shown in Figure 3.1, and of the flowrates produced every 300 s in five defined zones of nodes spread along the seepage boundary. The final time step of the transient simulation was chosen so as to capture approximatively the drawdown before its stabilization, similarly to the data collected at the Andra's site in Meuse/Haute Marne (Kerrou et al., 2017).

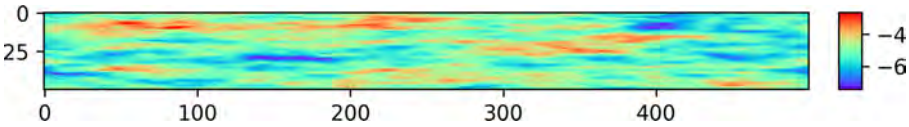


Figure 3.2: Reference field of decimal logarithm of hydraulic conductivities (m/s)

3.3.3 Initial ensemble of parameters and assumptions for the update step

In this synthetic case, an initial multi-Gaussian distribution of the variables to condition was considered in order to respect the underlying multivariate Gaussian assumption of ensemble Kalman methods. The initial ensemble is composed of multi-Gaussian realizations of mean value -5 of log hydraulic conductivity ($\log K$) fields generated, just as for the reference, using a fast Fourier transform method. The same variogram model as the one used in the reference was considered to describe the spatial variability of the log hydraulic conductivities, i.e. an exponential variogram of variance 0.49 and correlation lengths of 120 gridblocks in the horizontal direction and 10 gridblocks in the vertical direction. For this multi-Gaussian $\log K$ case, we assumed longer ranges in the horizontal direction to mimic the presence of the elongated porous horizons “HP” observed in the real field case. We assumed in the vertical direction the existence of five horizons, hence a correlation length was set to one fifth of the vertical dimension of the model. In the end, the realizations generated are considered heterogeneous since the correlation lengths are shorter with respect to the domain size, namely about one fifth of each dimension.

One practical issue of ensemble methods is that the finite number of members, usually no more than $O(100)$ in applications (Wen and Chen, 2005; Gillijns et al., 2006; Evensen, 2009a; Anderson, 2009), can cause spurious covariances between widely separated components of the parameter vector and between components of the vector of parameters and vector of predicted data. Since these long-range spurious correlations are non-physical, they can bring an undesirable response in the update computed. In general, they cause an underestimation of the ensemble variance which can potentially lead to “filter divergence” (Evensen, 2009a). This term refers to the situation where the assimilation of new measurements stops being effective because the spread of the ensemble has overly reduced or has “collapsed” to take them into account.

Considering that the observations are spatially restricted to the modelled domain (Figure 3.1) and that the hydraulic perturbation from the shaft is local relatively to the horizontal extension of the model, such long range correlations will be inevitable in this synthetic case. Therefore, it is necessary to try to filter them out as much as possible before computing the update. To this end, we will use a simple approach which consists in “localizing” the Kalman gain matrix so that only the parame-

ters located within a certain distance of an observation will be influenced by this observation during the update (Chen and Oliver, 2017). We will multiply the Kalman gain element-wisely with a “localization matrix” of the same size, i.e. number of parameters by number of data, which each entry will be a factor between 0 and 1 calculated by the correlation function defined by Gaspari and Cohn (1999)

$$\rho = \begin{cases} -\frac{1}{4}(\frac{\delta}{L})^5 + \frac{1}{2}(\frac{\delta}{L})^4 + \frac{5}{8}(\frac{\delta}{L})^3 - \frac{5}{3}(\frac{\delta}{L})^2 + 1, & \text{if } 0 \leq \delta \leq L \\ \frac{1}{12}(\frac{\delta}{L})^5 + \frac{1}{2}(\frac{\delta}{L})^4 + \frac{5}{8}(\frac{\delta}{L})^3 - \frac{5}{3}(\frac{\delta}{L})^2 - 5(\frac{\delta}{L}) + 4 - \frac{2}{3}(\frac{\delta}{L})^{-1}, & \text{if } L \leq \delta \leq 2L \\ 0, & \text{if } \delta > 2L \end{cases} \quad (3.10)$$

where δ denotes the distance between each couple of parameter and data variable and L is a predefined “critical length” beyond which the contribution of the data to the parameter update will be negligible. In this way, long-range spurious correlations will be partly removed from the Kalman gain and the performance of the ensemble method will be improved. Figure 3.3 illustrates the localization factors obtained by evaluating the Gaspari-Cohn function (3.10) for each couple of parameter of the model and the data located at the displayed point. Such factors will hence be used to localize the parameter update associated to each assimilated data at one of the observation locations. To avoid discontinuities in the localized parameter update between the two columns of data points (Figure 3.1), we set L to 1600 m, hence over a little larger distance than the distance separating the points horizontally, in all tests. Moreover, in this synthetic case, this distance is also consistent with the domain of the model that is identifiable given our data set, namely with the region influenced by the observed hydraulic perturbation.

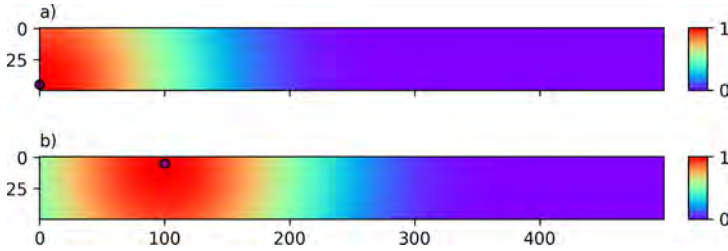


Figure 3.3: Examples of values calculated with the Gaspari-Cohn correlation function in order to localize the update of the parameter field when assimilating observations at the displayed point.

The ensemble smoother, LM-EnRML and ES-MDA all consider the vector of observations in their update equation as a random vector with the addition of a random noise vector sampled from the measurement error covariance C_{err} . For all tests performed, we assumed a diagonal covariance matrix of independent measurement errors of 0.05 m^2 for every head observation. When flowrate data were assimilated in addition to the hydraulic heads, C_{err} also included independent flowrate measurement errors set equal to the square of 20 % of the flowrate value. For the application of ES-MDA with transformed state variables, C_{err} was transformed as described in Section 3.2.3.

3.4 Performance criteria

Ideally, the application of ensemble Kalman methods for our parameter estimation problem aims to satisfy the following two criteria: (1) to reproduce the dynamic observations of state variables with the final ensemble of conditioned realizations, and (2) to obtain a final ensemble of conditional realizations which variations around the mean correctly quantify the uncertainty. As we assumed in this synthetic case no model errors or errors in the choice of the prior distribution used to sample the parameter space, meeting both these criteria should ensure that the ensemble-based uncertainty in the model predictions will be well represented.

To assess the quality of the fit between the simulated $g(m_i)$ and observed data d_{obs} using either LM-EnRML or ES-MDA with or without transformed data, we will consider the evolution of the sum of squared errors

$$D(m_i) = (g(m_i) - d_{obs})^T (g(m_i) - d_{obs}) \quad (3.11)$$

Note that this sum of squared errors does not correspond to the data mismatch term of the cost function which is actually minimized by each algorithm considered (cf. Section 3.2). Here, d_{obs} is the vector of original observations without added noise, hence no normalization by the variance of the measurement error is considered in the equation. We do not take into account these random perturbations and normalization factors simply so as to more easily compare the final mismatch values obtained with each algorithm, especially the application of ES-MDA using transformed state variables.

However, in order to assess the effects of using transformed compared to untransformed data with ES-MDA, we will compute for each case the data mismatch term of the different cost functions being minimized at each iteration. Indeed, Evensen (2018) noted that each updated estimate m_i^{k+1} of ES-MDA actually corresponds to an estimation of the minimum of a cost function written as

$$\begin{aligned} D(m_i^{k+1}) = & \frac{1}{2}(g(m_i^{k+1}) - d_{obs} - \sqrt{\alpha_k}z_{d,i})^T(\alpha_k C_{err})^{-1} \\ & \times (g(m_i^{k+1}) - d_{obs} - \sqrt{\alpha_k}z_{d,i}) \\ & + \frac{1}{2}(m_i^{k+1} - m_i^k)^T(C_M^k)^{-1}(m_i^{k+1} - m_i^k) \quad (3.12) \end{aligned}$$

where $m_i^{k=1} = m_i^{pr}$ and $C_M^{k=1} = C_M^{pr}$ with C_M^{pr} the prior model covariance approximated from the ensemble around the prior ensemble mean $m_i^{pr} = m_i^{pr}$.

Once an acceptable data match has been reached during the conditioning procedure, the uncertainty based on the variability of the final ensemble of conditioned realizations can be assessed. In practice, ensemble-based Kalman methods are known to particularly overestimate the uncertainty reduction when they are not applied in sufficiently optimal conditions. In particular, the use of a finite ensemble inevitably introduces sampling errors in the approximated covariances which can lead to an overly reduced variance of the conditioned ensemble. Moreover, the solution space spanned by the ensemble when computing the update will likely be under-sampled (Evensen, 2009a) and the estimated uncertainty will not be reliable. To reduce these sampling errors, a natural solution is to increase the ensemble size. However, because the Monte Carlo sampling errors decrease proportionally to $1/\sqrt{N}$, with N the ensemble size, the improved performance of the ensemble method can come at a significant computational cost. For example, Chen and Zhang (2006) noted that an ensemble size of 1000 allowed EnKF to obtain an accurate uncertainty estimation for their synthetic case, but ultimately concluded that an ensemble size of 200 was sufficient to achieve results with both accuracy and efficiency. When applying iterative ensemble smoothers, the choice of the ensemble size will be particularly constrained by the time needed to run all the forward predictions of the ensemble during one iteration. Consequently, larger ensemble sizes than $O(100)$ are generally not affordable if the available computational resources are not sufficient to efficiently parallelize the

prediction step of every ensemble member. For these practical reasons, we will discuss the performance of the different ensemble methods based on tests involving ensemble sizes of the order of $O(10^2)$ at most.

To assess the accuracy of the uncertainty captured by the final conditioned ensemble, a common approach consists in comparing the error between the ensemble mean of $\log K$ realizations and the known reference to the ensemble mean error that was obtained by the ensemble method, also called the “ensemble spread” (Houtekamer and Mitchell, 1998). For the first error mentioned, we will simply compute the root-mean-squared error (RMSE) as follows

$$RMSE = \sqrt{\frac{1}{N_m} \sum_{i=1}^{N_m} (m_{true} - m_e)^2} \quad (3.13)$$

where N_m is the number of gridblocks, m_{true} is the reference $\log K$ field, and m_e stands for the mean of the ensemble of $\log$ hydraulic conductivity fields.

The ensemble spread, here noted S_{ens} , corresponds to an average uncertainty on the $\log K$ property calculated from the ensemble

$$S_{ens} = \sqrt{\frac{1}{N_m} \sum_{i=1}^{N_m} \sigma_{ens,i}^2} \quad (3.14)$$

where $\sigma_{ens,i}^2$ corresponds to the variance estimated from the ensemble of $\log K$ realizations at one gridblock i . For our uncertainty analysis, we will systematically evaluate these average errors by considering the gridblocks of the model located within a certain distance from the seepage boundary which include all the data points. In particular, we will set this distance equal to the “critical length” used for the localization of the update (cf. Section 3.3.3). In this way, the comparison between the RMSE and the ensemble spread will be based on the set of parameters which are the most informed by the observations and hence which uncertainty will decrease the most during the assimilation.

One goal of this analysis will be to illustrate how the spread between ensemble members updated using either LM-EnRML, ES-MDA with or without transformed data is representative of the difference between the ensemble mean and the reference (RMSE) depending on the ensemble size. Indeed, as the size of the ensemble increases, the discrepancy between the final conditioned ensemble and the RMSE should

reduce thanks to the reduced sampling errors. However as previously mentioned, a trade-off between accuracy of the uncertainty and computational efficiency will need to be found. In the end, if the ensemble spread is not overly underestimating the RMSE, we will consider that the ensemble method has performed correctly given its intrinsic limitations. In addition, although the inverse problem is ill-posed, i.e. there are more parameters to infer than data to inform them in a unique way, we expect for this synthetic case that the information contained in the data set is sufficient in order to yield after the data assimilation estimates that are closer to the reference than initially in the most updated region of parameters. Therefore, the performance of the smoother will also be assessed in its ability to decrease the RMSE value.

3.5 Results and discussion

3.5.1 Ensemble smoother and benefit of data transformation

As expected in a nonlinear case, the application of a non-iterative ensemble method such as the ensemble smoother (ES) results in an insufficient match between the final ensemble of predicted data, i.e. computed from the ensemble of updated realizations, and the observations. Figure 3.4 shows for three locations at different depths 1 km away from the producing shaft the simulated heads before and after the update of 100 realizations with ES. By using the same initial ensemble, the evolution of the data mismatch obtained from the assimilation of the head data only and both the head and flowrate data simultaneously are shown respectively in Figure 3.6a and 3.6c. Because the represented data mismatch does not include the inflated measurement error in the observed data and hence the normalization by the error variance, as defined in Section 3.4, the additional contribution to the mismatch from the flowrates in Figure 3.6c is not visible. Even so, compared to Figure 3.6a), we can observe that the assimilation of the flowrates in addition to the heads allows the data mismatch to decrease thanks to a better match of the head data, as confirmed in Figure 3.7.

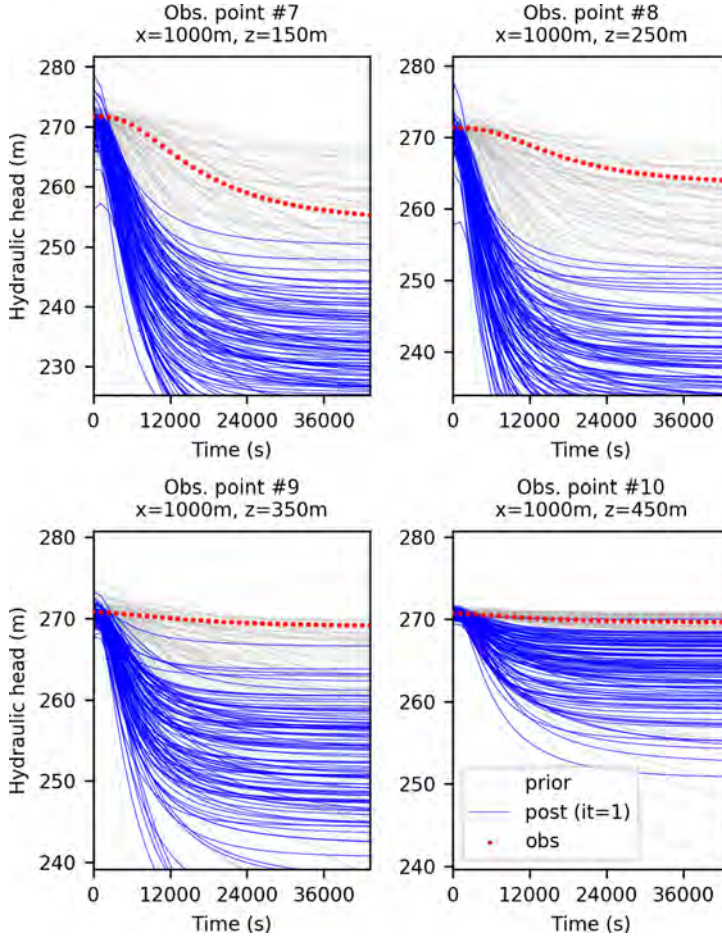


Figure 3.4: Predicted hydraulic head at three different locations, before (in grey) and after (in blue) the assimilation with ES of the head data without normal-score transform. The red dots are the observed data; The ensemble size is 100.

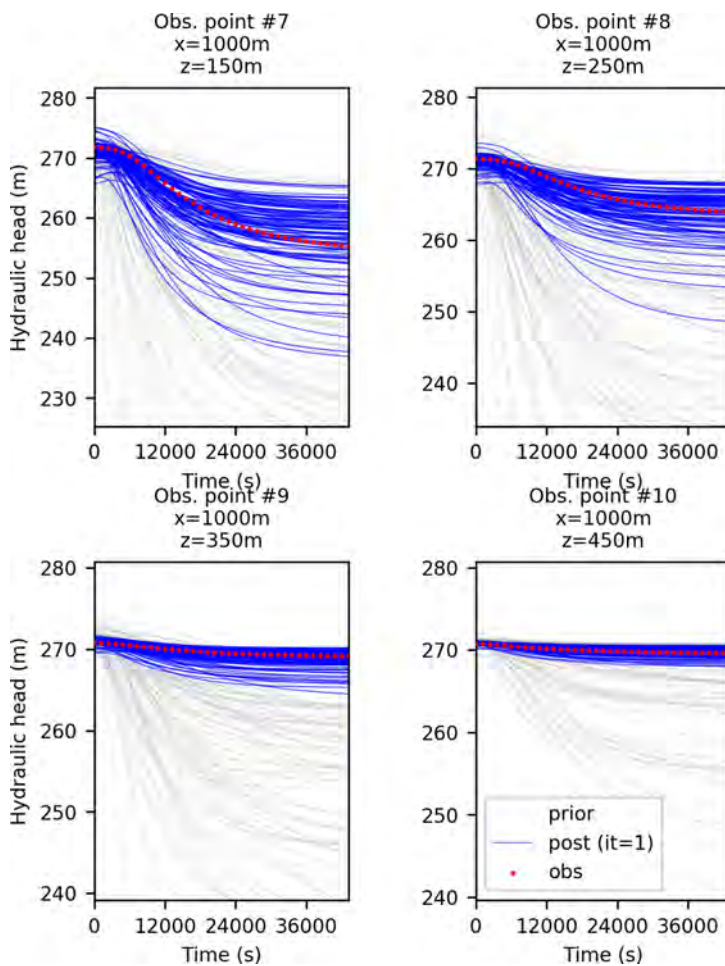


Figure 3.5: Predicted hydraulic head at three different locations, before (in grey) and after (in blue) the assimilation with ES of normal-score transformed head data. The red dots are the observed data; The ensemble size is 100.

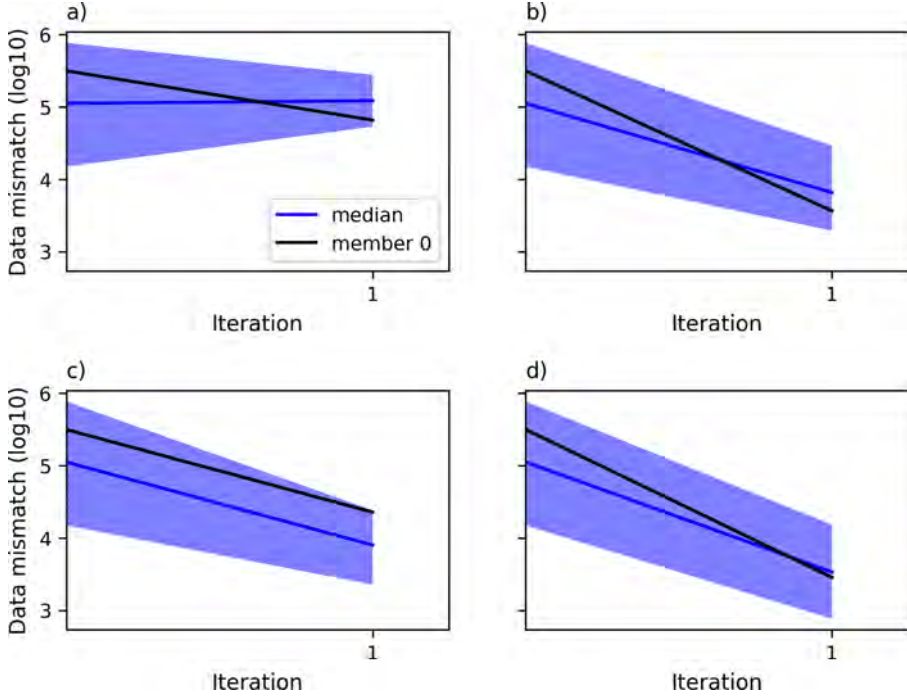


Figure 3.6: Evolution of the decimal logarithm of the data mismatch when assimilating with ES: a) the head data, b) the transformed head data, c) both the head and flowrate data, d) both the transformed head and transformed flowrate data. The blue curve is the median, the blue area corresponds to the region between the percentiles P5 and P95 of the ensemble of values, and the black curve is the values computed for the ensemble member no. 0.

The strategy of transforming the data prior to each update of ES-MDA, as described in Section 3.2.3, was also tested here with ES using the same initial ensemble. Indeed, ES is equivalent to ES-MDA with only one iteration and hence applies once a local linearization of the model around the initial parameter estimate to compute its single update. As shown in Figures 3.6b and 3.6d, whether only the head data are assimilated or both the head and flowrate data, the data mismatch is reduced in both cases with a slightly lower mismatch in the case mentioned last. Therefore, compared to the case where only untransformed head are assimilated (Figure 3.6a), the data transformation seems to have a positive impact on the decreased data mismatch which, we remind, considers the original data even for the transformed case. The improved

history match can be observed for example by comparing Figure 3.4 with Figure 3.5. In particular, we note that when data transformation is used, the ensemble of predicted values better tracks the observations, most certainly as a result of a more accurate update of each ensemble member. Indeed, the comparison of the RMSE and the ensemble spread in Figure 3.8 for each one of the cases considered in Figure 3.6 shows that the assimilation of the transformed heads instead of the original data (Figure 3.8b) results in a more accurate estimation of the ensemble mean and spread thanks to a decrease of the RMSE which reduces the discrepancy with the reduced ensemble spread. As for the assimilation of both the head and flowrate data simultaneously (Figure 3.8c), we can see that compared to when the data are transformed (Figure 3.8), the assimilation is less efficient with a less reduced ensemble spread and a RMSE that has almost not reduced. Based on both the evolution of the RMSE and the “coverage”, which we defined as the percentage of true values of the reference captured by the final ensemble spread in the most updated region of the model, we can conclude from Table 3.1 that transforming the data clearly benefits the performance of ES, although we also note that the assimilation of both the transformed head and flowrate data with ES does not allow to further reduce the RMSE compared to when using the transformed heads only.

Assimilated data	RMSE	Ens. spread	Coverage
h	0.72 - 1.00	0.70 - 0.42	57 %
h, q	0.72 - 0.69	0.70 - 0.53	93 %
transf. h	0.72 - 0.58	0.70 - 0.44	91 %
transf. h, transf. q	0.72 - 0.59	0.70 - 0.41	88 %

Table 3.1: Influence of different types of data assimilated (transformed or not) on the RMSE, ensemble spread, and coverage calculated with ES. The ensemble size was 100 for all the tests shown in this table.

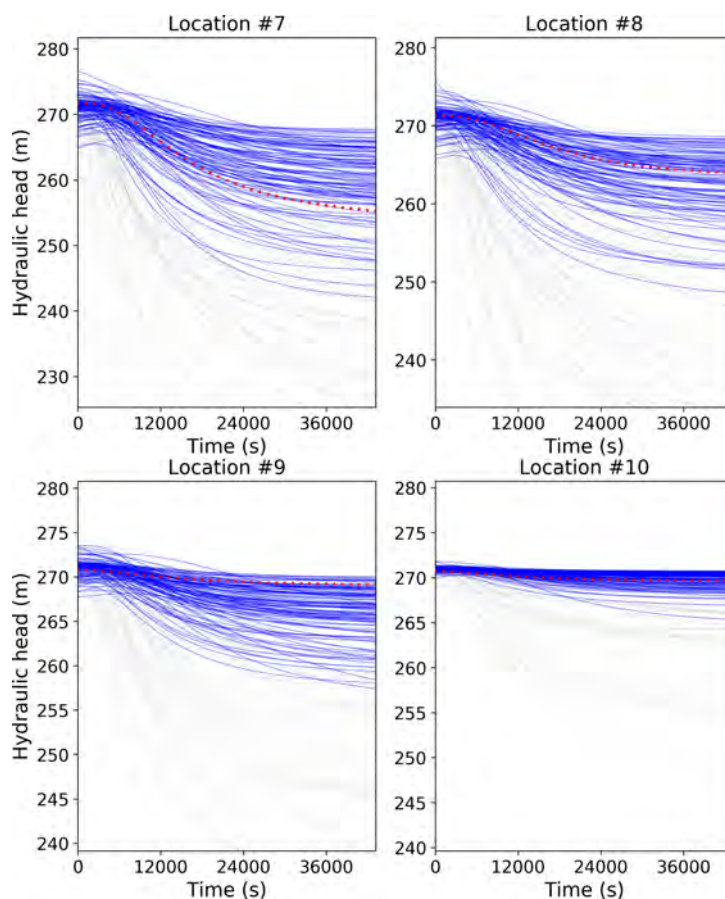


Figure 3.7: Predicted hydraulic head at three different locations, before (in grey) and after (in blue) the simultaneous assimilation of both the head and flowrate data. The red dots are the observed data; The ensemble size is 100.

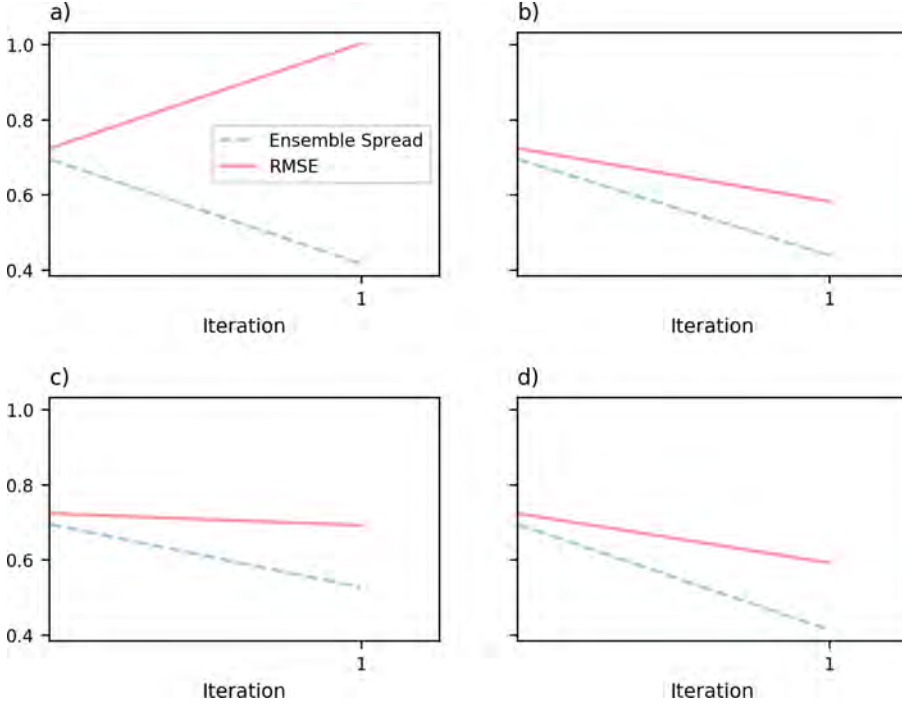


Figure 3.8: Evolution of the RMSE and ensemble spread in the updated region (up to 1600 m from the seepage boundary) before and after assimilating with ES: a) head data, b) transformed head data, c) both head and flowrate data, d) both transformed head and transformed flowrate data.

3.5.2 Comparing the accuracy of LM-EnRML and ES-MDA estimates

Although ES-MDA and the LM-EnRML were developed with the aim of improving the data match obtained with ES in the nonlinear case, the iterative updates will likely cause a more important underestimation of the uncertainty than with ES. Indeed, at each iteration, the sampling errors in the approximated covariances of the Kalman gain will affect the estimates of the method, and more particularly the ensemble uncertainty. Seeing as how the Monte Carlo sampling errors decrease according to $1/\sqrt{N}$ where N is the ensemble size (Evensen, 2009a), we tested for our synthetic case the effect of increasing N on the performance of ES-MDA and LM-EnRML when only assimilating the head data. Figure 3.9 shows the evolution of the RMSE and the ensemble

spread when applying ES-MDA with 4 arbitrarily defined iterations. Whether we consider an ensemble of 100, 200, or 400, hence of the order of $O(10^2)$ to remain computationally efficient, we note that starting from the first update, the ensemble spread measuring the variability between the ensemble members systematically underestimates the RMSE which measures the error between the reference and the ensemble mean. This discrepancy with the RMSE further increases with each iteration as the spread keeps decreasing while the RMSE remains higher. Although we do not show it here, not localizing the update resulted in a larger discrepancy between the RMSE and ensemble spread for the different ensemble sizes tested.

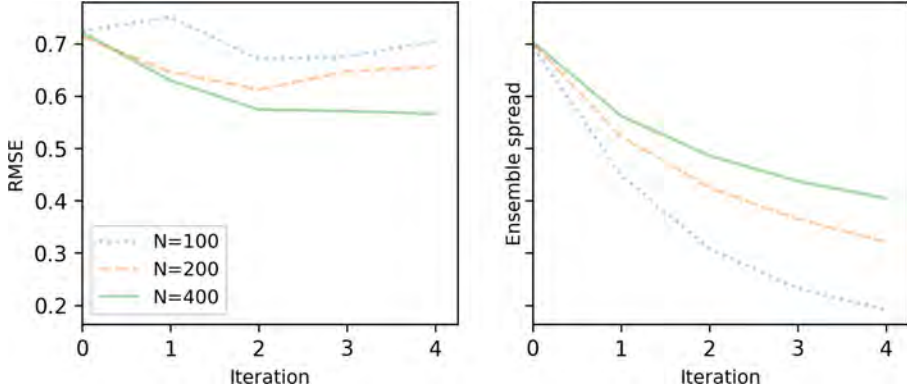


Figure 3.9: Sensitivity to the ensemble size of the RMSE and ensemble spread with ES-MDA. The dashed curves correspond to results obtained without localization.

Nevertheless, we note that the larger the ensemble size, the more reduced the gap between the RMSE and the ensemble spread becomes. As expected, the estimated uncertainty gains in accuracy with the ensemble size. We note however that the most significant improvement occurs when increasing the ensemble from 100 to 200 realizations for both algorithms (Figures 3.9 and 3.10). Indeed, as indicated for ES-MDA in Table 3.2, the coverage, defined as the proportion of reference values that are within the ensemble range in the most updated region of the model, has more than doubled with a final value of 77 % after the same number of iterations. In a similar manner, the coverage also doubled with LM-EnRML (Table 3.3). Moreover, when using an ensemble size of 100, we can see that the evolution of the RMSE is particularly unstable for both LM-EnRML and ES-MDA. The ensemble means esti-

mated with LM-EnRML even lead to a final RMSE value which is much higher than initially. In contrast, an ensemble size of 200 allows LM-EnRML to provide a more accurate estimation of the true uncertainty than with ES-MDA. The results shown in Table 3.3 were all calculated based on the estimates of LM-EnRML obtained after 8 iterations when the reduction of the objective function seemed to have decreased to a sufficiently low data mismatch value as shown in Figure 3.11a. We note also from Figure 3.11, however, that it seems to take more iterations for LM-EnRML to converge to the same level of data match achieved by ES-MDA using 6 predefined iterations.

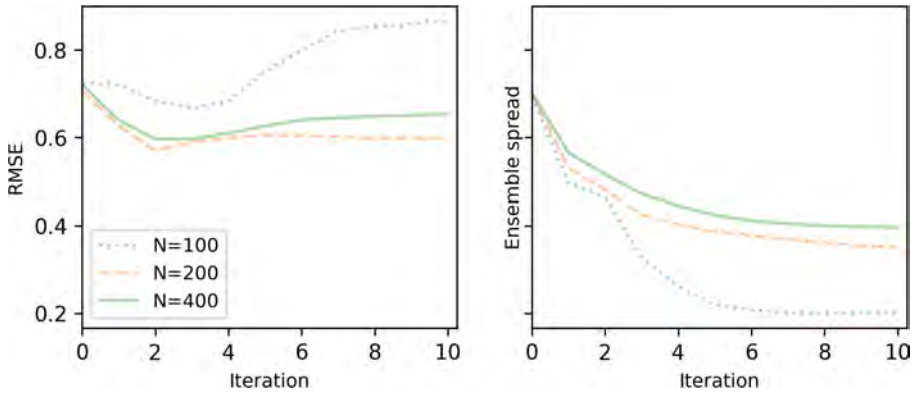


Figure 3.10: Sensitivity to the ensemble size of the RMSE and ensemble spread with LM-EnRML

Ens. size	Assimilated data	RMSE	Ens. spread	Coverage
100	h	0.72 - 0.71	0.70 - 0.19	37 %
200	h	0.71 - 0.66	0.70 - 0.32	77 %
400	h	0.72 - 0.57	0.70 - 0.40	96 %

Table 3.2: Effect of increasing the ensemble size on the RMSE, ensemble spread, and coverage after 4 predefined ES-MDA iterations.

Ens. size	Assimilated data	RMSE	Ens. spread	Coverage
100	h	0.72 - 0.88	0.70 - 0.19	49 %
200	h	0.70 - 0.60	0.70 - 0.34	93 %
400	h	0.72 - 0.66	0.70 - 0.39	94 %

Table 3.3: Effect of increasing the ensemble size on the RMSE, ensemble spread, and coverage with LM-EnRML at iteration 8.

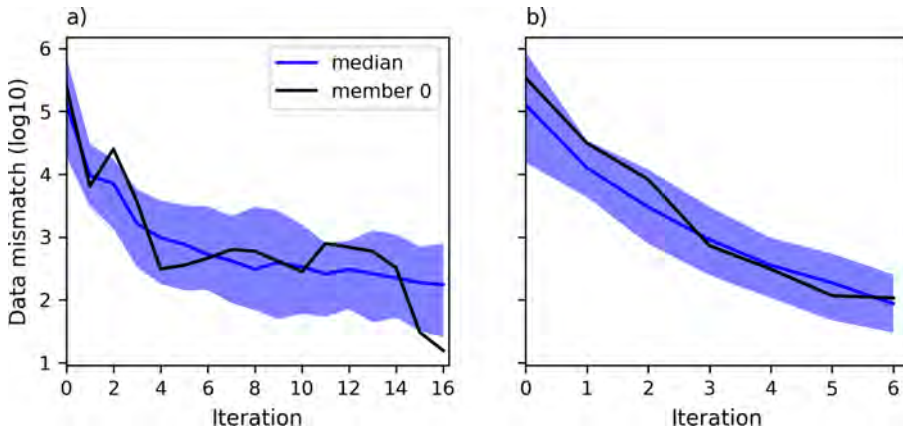


Figure 3.11: Evolution of the decimal logarithm of the data mismatch when assimilating head data with: a) LM-EnRML, b) ES-MDA ($\times 6$). The ensemble size is 200.

From all these tests, it is clear that the larger the size, the more accurate the uncertainty estimated by LM-EnRML or ES-MDA will be. Distributions of the ensembles of $\log K$ values per grid-block for the different ensemble sizes tested can be found in the Appendix. However, as can be observed in Figure 3.9, this improvement in the estimates comes at an increased computational cost. For our synthetic problem, we note that an ensemble of 400 realizations would still not allow the ensemble to represent correctly the true error on the ensemble mean. Since increasing the ensemble size does not seem to be a reasonable option computationally, it seems that the application of such ensemble methods necessarily requires to achieve a balance between accuracy and computational efficiency. This computational efficiency is of even more importance for large-scale problems since the biggest constraint would

often be the long flow simulation time required to run the prediction step for each ensemble member. Hence the larger the ensemble, the longer it will take to complete all the predictions of the ensemble before being able to start the update step. Indeed, we remind that the update of each member is based on the same ensemble-based Kalman gain matrix. It is worth mentioning that this matter of efficiency also concerns the case where the ensemble of predictions steps is parallelized. Indeed, it would still require to have enough resources to efficiently run the whole ensemble of forward simulations simultaneously after each update step.

From Tables 3.2 and 3.3, it is clear that the improvement on both the ensemble mean and spread is very significant when increasing the size from 100 to 200 with either LM-EnRML and ES-MDA. However, the differences in the estimated ensemble mean fields shown in Figure 3.12 suggest that the two methods converge towards different solutions. This is confirmed by the fact that when increasing the ensemble size from 200 to 400, the accuracy of both the error and uncertainty estimated with ES-MDA keeps increasing whereas for LM-EnMRL, no visible improvement is observed as if an optimal ensemble size has been reached around 200.

Although LM-EnRML provides the more accurate estimation of the uncertainty compared to ES-MDA for any of the tested ensemble sizes, ES-MDA seems to converge faster, as shown in Figure 3.11. Indeed, it takes more iterations with LM-EnRML to reach the same mismatch value which corresponds to a good match as obtained with ES-MDA using 6 iterations. In the end, although the estimates of ES-MDA definitely improve with an ensemble size of 400, we will consider that using an ensemble of size 200 with either ES-MDA or LM-EnRML in our case is acceptable in order to achieve both accuracy and efficiency.

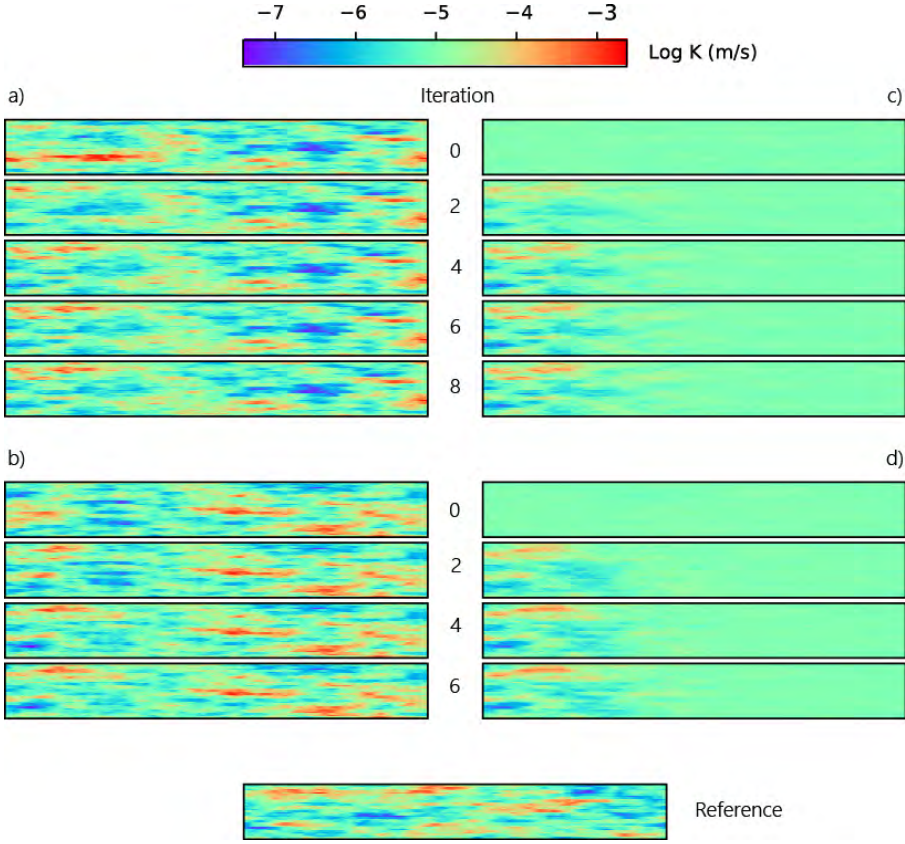


Figure 3.12: Two initial different realizations of log hydraulic conductivity being updated respectively with LM-EnRML (a) and ES-MDA (b) throughout the iterations. (c) and (d) show the corresponding ensemble mean fields at each iteration. The ensemble size is 200.

3.5.3 Assimilating both hydraulic head and flowrate data with LM-EnRML and ES-MDA

As shown in Table 3.4, the assimilation of both head and flowrate data with LM-EnRML yields estimates of similar accuracy to when only the heads are assimilated. Neither the larger number of observations nor the higher “weights” of the flowrates in the objective function, given their much lower magnitudes, seem to have affected the minimization process. Figure 3.13a shows in fact that the mismatch associated to the flowrate data in the objective function, i.e the sum of squared residuals weighted by the inverse of the measurement variance, is mostly reduced after

the first iteration so that ultimately the remaining of the optimization deals with the reduction of the mismatch associated to the head data (Figure 3.13b). Although the convergence of LM-EnRML seems effective when assimilating different data types, we note that the accuracy of the estimates does not improve with the additional information coming from the assimilated flowrates. This lack of improvement could be related to possible over-corrections of the parameters during the first iteration as often observed with gradient-based methods. One way to reduce such over-corrections would be to artificially increase the measurement errors. We did not try that however since the parameter estimates obtained seem sufficiently accurate in terms of the estimated uncertainty.

Algorithm	Assimilated data	RMSE	Ensemble spread	Coverage
LM-EnRML	h	0.72 - 0.60	0.70 - 0.34	93 %
LM-EnRML	h, q	0.71 - 0.60	0.69 - 0.34	93 %
ES-MDA($\times 6$)	h	0.72 - 0.65	0.69 - 0.29	72 %
ES-MDA($\times 8$)	h, q	0.72 - 0.70	0.70 - 0.20	58 %
ES-MDA($\times 8$)	transf. h	0.72 - 0.66	0.69 - 0.31	80 %
ES-MDA($\times 8$)	transf. h, q	0.70 - 0.63	0.70 - 0.30	82 %

Table 3.4: Effect of assimilating the flowrate data in addition to the hydraulic heads on the RMSE, ensemble spread, and coverage with LM-EnRML and ES-MDA with or without transformed data. The ensemble size was 200 for all the tests shown in this table. The corresponding number of iterations was chosen so as to obtain an acceptable and comparable data match for all the results.

For the assimilation of both data types, we first note that ES-MDA converged more rapidly, i.e. in 8 iterations, to the same level of data match achieved with LM-EnRML after 16 iterations (Figures 3.13a and 3.14a). The history match of the head and flowrate observations obtained with LM-EnRML and ES-MDA are shown respectively in Figures 3.15 and 3.16. However, unlike previously with LM-EnRML, the mean and variance estimates obtained with ES-MDA are less accurate than those obtained by the assimilation of head data only (Table 3.4). This illustrates that the different data types are not properly taken into account with ES-MDA during the multiple assimilations compared to LM-EnRML.

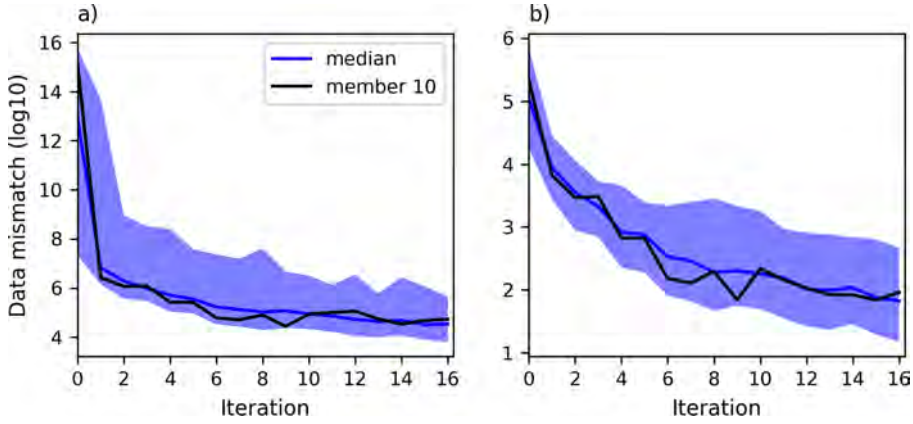


Figure 3.13: Evolution of the decimal logarithm of the data mismatch when assimilating both head and flowrate data with LM-EnRML: a) With the normalization of the residuals by their variance (as minimized by the objective function). b) Without normalization, hence the values reflect the head data mismatch reduction mainly. The ensemble size is 200.

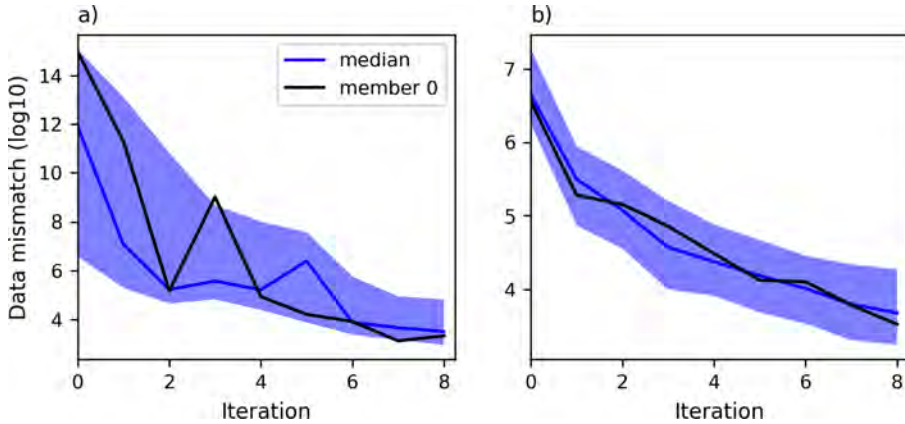


Figure 3.14: Evolution of the decimal logarithm of the data mismatch term of the cost function minimized at each iteration when assimilating both head and flowrate data with ES-MDA: a) Without transformed data. b) With transformed data. The ensemble size is 200.

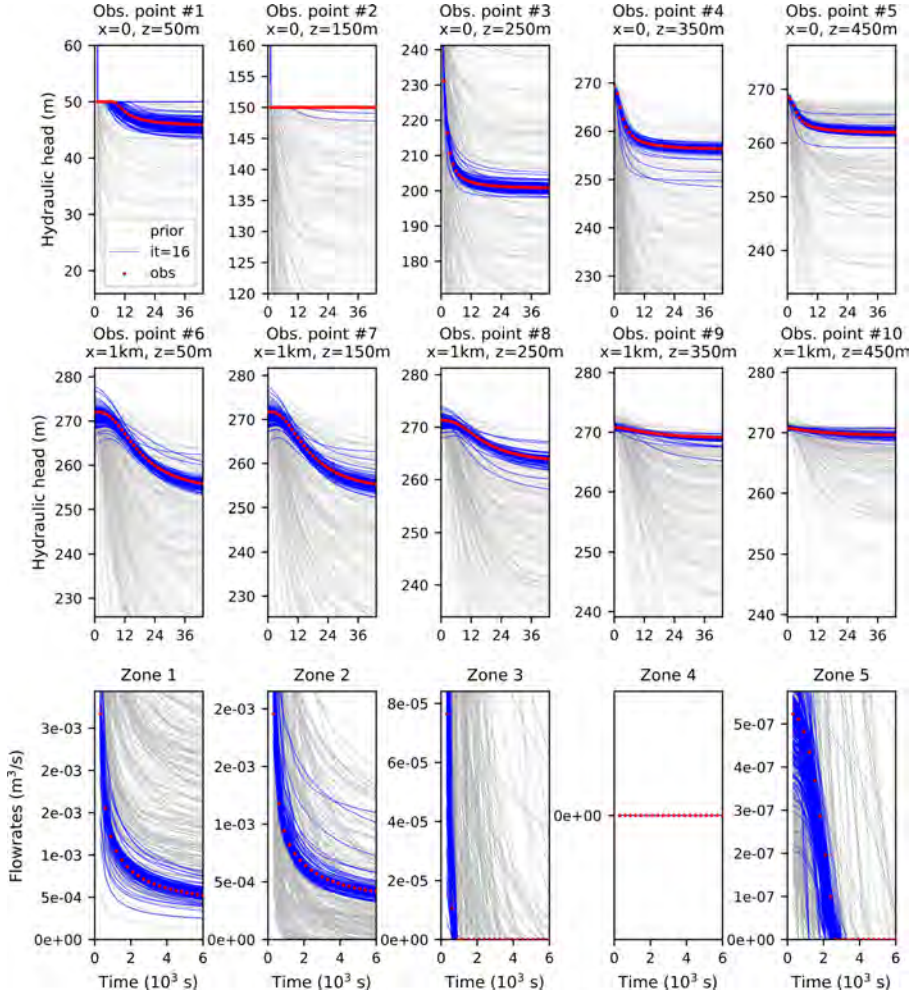


Figure 3.15: Predicted hydraulic head and flowrate data at the observation locations, before (in grey) and after (in blue) the assimilation of both the hydraulic heads and flowrate data with LM-EnRML. The red dots are the observed data; the ensemble size is 200.

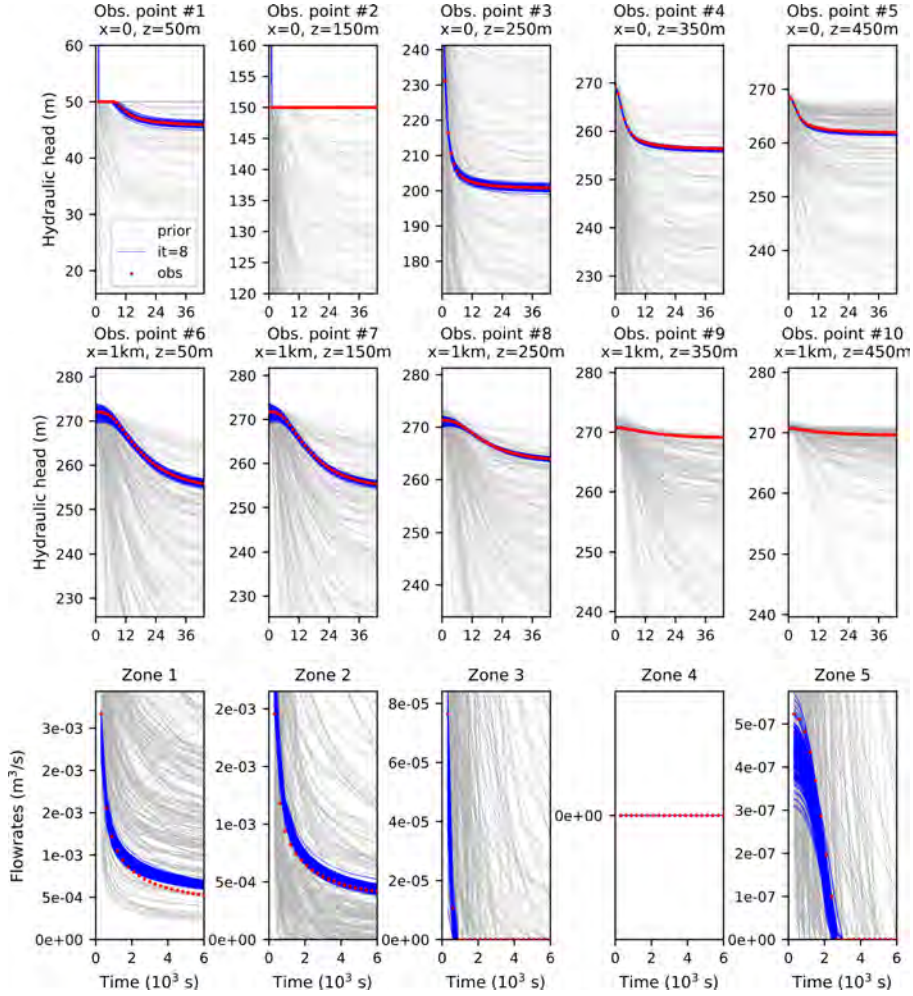


Figure 3.16: Predicted hydraulic head and flowrate data at the observation locations, before (in grey) and after (in blue) the assimilation of both the hydraulic heads and flowrate data with ES-MDA. The red dots are the observed data; the ensemble size is 200.

As done previously with ES in order to mitigate the effects of non-Gaussianity on the performance of the method, we applied a normal-score transform on the data so as to obtain ensemble distributions of predicted variables which are locally Gaussian, for each data type, before performing each update step of ES-MDA in the transformed data space. As shown in Figure 3.17, the proposed state transformation allows ES-MDA to achieve an acceptable match of the head and flowrate

observations after 8 iterations. Table 3.4 shows the estimates obtained after assimilating only the transformed heads resulted in a slight improvement. The coverage has slightly increased, which indicates that the estimated uncertainty, although still underestimated, has improved a little. A much more significant improvement, however, can be observed when both the head and flowrate data are assimilated in the transformed space. Compared to when no transformation is applied, the assimilation of both transformed data types is much more efficient given the larger variability of the ensemble and the smaller RMSE.

Figure 3.14 shows that when both data types are transformed, the decrease of the cost functions minimized for each iteration of ES-MDA (Evensen, 2018) is more stable than when no transformation is applied. In the case without transformation, the larger range of the initial ensemble of values reflects the differences of magnitude of the mismatch associated to the flowrate data on the one hand, and the one associated to the heads on the other hand. As shown in Figure 3.18a, the distribution is spread around two distinct modes. Hence the less stable reduction observed is probably related to the contribution of the flowrate data which dominates the cost function to minimize because of their much more smaller measurement errors. In contrast, the initial ensemble of mismatch values resulting from the transformed data (Figure 3.14b) is much less spread and resembles a Gaussian distribution (3.18). By making Gaussian the distribution of all mismatch values defining the first cost function to be minimized with ES-MDA, the normal-score transform applied to each data type seems to have benefited the data assimilation with ES-MDA. Since a new cost function is minimized after each ES-MDA iteration, we suppose that the initial Gaussian distribution of all the data mismatch values resulting from the transformation contributes to make the minimization more efficient by allowing to take equally account of each data type in the assimilation from the start of the assimilation. However, as discussed by Schoniger et al. (2012), it is important to note that such local transformation of the state variables would not have led to a successful application of ES-MDA if the multivariate dependence of the state variables was not sufficiently Gaussian in the first place. Indeed, since the proposed transformation only affects the marginal distribution, it is essential that the multivariate structure of the state variables is near-Gaussian in order to really observe an improvement of the performance of the ensemble Kalman method. Crestani et al. (2013) illustrated this point by observing that similar local transformation performed on concentration data can dete-

riorate the performance of EnKF because of an insufficient dependence between the multi-Gaussian log hydraulic conductivity parameters and the concentrations.

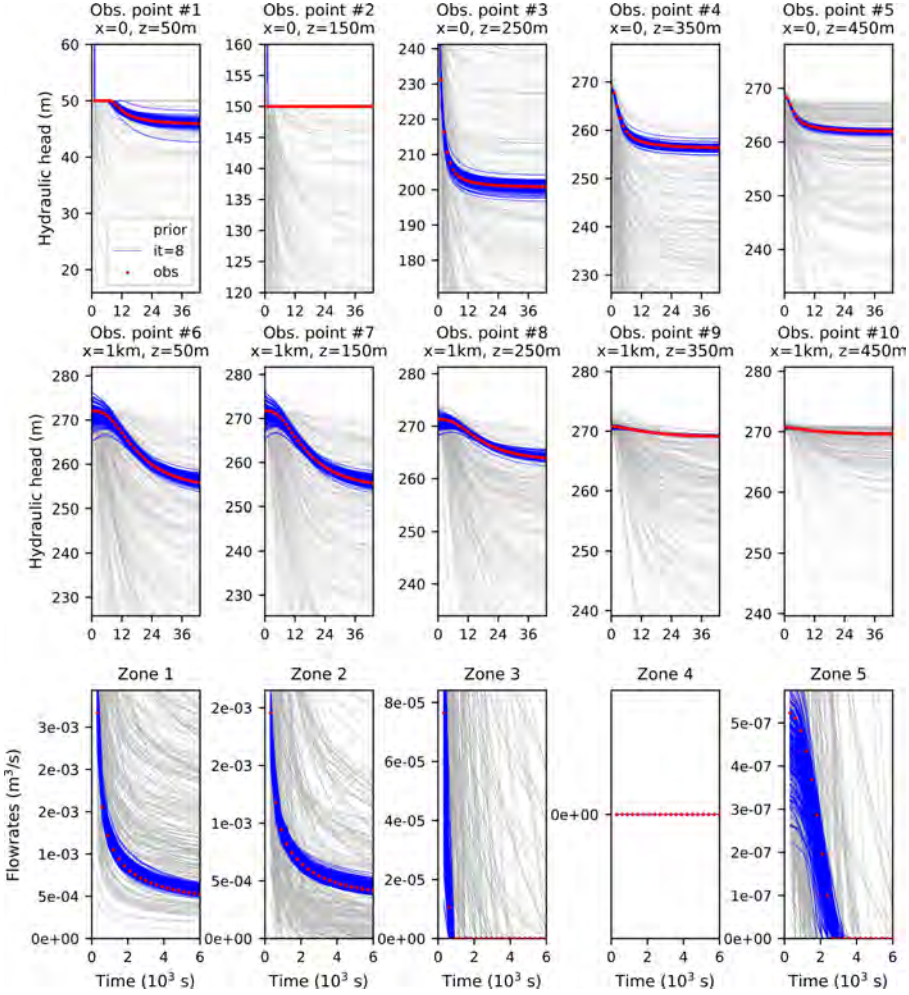


Figure 3.17: Predicted hydraulic head and flowrate data at the observation locations, before (in grey) and after (in blue) the assimilation of both the transformed hydraulic heads and transformed flowrate data with ES-MDA. The red dots are the observed data; the ensemble size is 200.

It could be argued that equal contributions from both the heads and flowrate data to the first cost function could also have been obtained in

the case without transformation by decreasing the measurement variance of the head measurements. However, the use of very small values of measurement errors can cause problems during the data assimilation as it may lead to severe underestimations of the parameter uncertainty due the overconfidence in the assimilated data.

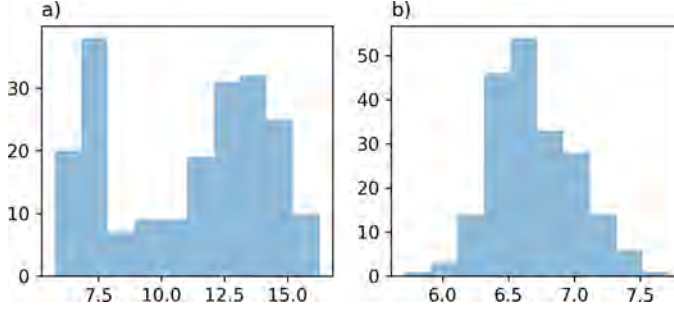


Figure 3.18: Distribution of the decimal logarithm of the sum of weighted squared errors prior to the assimilation of head and flowrate data with ES-MDA: a) Without transformation. b) With transformation.

3.6 Conclusion

In this synthetic study, two iterative forms of ensemble smoother, LM-EnRML (Chen and Oliver, 2013) and ES-MDA (Emerick and Reynolds, 2012a), were applied and compared. They were both used to condition an ensemble of multi-Gaussian log hydraulic conductivity fields to transient hydraulic data. Given the nonlinear dynamics of the groundwater flow problem inspired by a real hydraulic situation, the results show the necessity of using an iterative instead of non-iterative ensemble smoother in order to obtain an ensemble of hydraulic conductivity fields which all match properly the data.

Despite the multi-Gaussian log K distribution considered, the repeated assimilations of the data set have highlighted the known tendency of ensemble Kalman methods to underestimate the true error due to the finite ensemble size. A trade-off between accuracy of the estimation and computational efficiency hence needs to be found when applying such methods. Our uncertainty analysis based on the comparison between the ensemble spread and the true error, here accessible in this synthetic study, clearly indicated the existence of a threshold ensemble size below

which the updated mean and variance are not reliable. For our case, we found that the final estimates using an ensemble of 100 were not acceptable whereas increasing the ensemble to a still reasonable size of 200 improved significantly the accuracy of the updated mean and spread of the ensemble for both algorithms.

Although LM-EnRML outperforms ES-MDA in terms of the accuracy of the estimated uncertainty when using an ensemble size of 200, the performance of ES-MDA seems to improve steadily with the ensemble size while LM-EnRML does not. In particular, the RMSE of the ES-MDA estimates keeps decreasing. This particularly underlines that both algorithms do not converge to the same solution. It is not so surprising considering that LM-EnRML was derived so as to minimize an objective function using a gradient-based approach while ES-MDA was derived to minimize the variance of the error (Chen and Oliver, 2013; Evensen, 2018).

The benefit of transforming the state variables with ES-MDA was mainly observed when assimilating the hydraulic head and flowrate data simultaneously. Indeed, the normal-score transform of the state variables allowed a normalization of the magnitudes of all data types to values drawn from a standard normal distribution. Consequently, both data types could be more equally taken into account during the data assimilation as shown by the more stable reduction of the cost functions compared to the application of ES-MDA without transformation. However, as commented in previous studies (Schoniger et al., 2012; Crestani et al., 2013), the applicability of such state transformation could be questioned in cases where the multi-Gaussian dependence of the state variables is not as strong as in this multi-Gaussian log K case, either as a result of the nonlinearity of the model and/or because of an initial non-multi-Gaussian parameter distribution.

Chapter 4

Updating non-multi-Gaussian groundwater flow parameters of a channelized aquifer with iterative ensemble smoothing

4.1 Introduction

The ensemble Kalman filter (EnKF) and other ensemble-based variants have been introduced as an efficient alternative to the Kalman filter for high-dimensional state estimation problems (Evensen, 2009b). In the context of inverse modelling, ensemble-based Kalman methods can be employed for the assimilation of observations of the system state in order to condition the non-observed parameters of the model. In Chapter 3, we concluded that for our synthetic case, the application of iterative ensemble smoothers on an initial ensemble of multi-Gaussian parameter fields could yield reasonably accurate estimates within a few iterations by using an ensemble size of the order of $O(10^2)$. However, this multi-Gaussian assumption may not always be adapted for the problem considered. For instance, a multi-Gaussian distribution cannot represent certain types of heterogeneities associated to highly connected or disconnected structures while this complexity could very much affect the flow and transport processes (Zinn and Harvey, 2003; Kerrou et al., 2008).

Although the application of ensemble-based Kalman methods in

problems with nonlinear dynamics has shown to be quite robust especially with the development of iterative methods (Gu and Oliver, 2007; Chen and Oliver, 2013), the biggest challenge lies in the application of such methods in the suboptimal case of a non-multi-Gaussian initial ensemble. By construction, the moments of order higher than two cannot be captured during the linear update step of these methods (Evensen, 2009b). Even so, the ability of ensemble methods to handle high-dimensional estimation problems, combined with their convergence achieved at a lesser cost compared to other Monte-Carlo approaches, has continued to motivate parameterization strategies in order to apply them more properly to non-multi-Gaussian or discrete cases (Liu and Oliver, 2005; Jafarpour and McLaughlin, 2008; Chen et al., 2009; Jafarpour and Khodabakhshi, 2011; Zhou et al., 2011; Hu et al., 2013; Cao et al., 2018). In general, non-multi-Gaussian state variables can either result from the technique used to simulate the ensemble of stochastic parameter fields, or from a nonlinear process model and imposed boundary conditions even if the parameters are initially multi-Gaussian (Schoniger et al., 2012; Evensen, 2018).

To get closer to the multi-Gaussian assumption, Zhou et al. (2011) and Li et al. (2012) suggested to apply EnKF with a normal-score transform on the log hydraulic conductivity ($\log K$) parameters and the hydraulic head data. It is worth pointing out that unlike the normal-score approach on the state variables proposed by Schoniger et al. (2012) when considering multi-Gaussian $\log K$ fields, the motivation in this non-Gaussian parameter case is not to trigger a pseudo-linearization of the relationship between the parameters and the data, but simply to perform the EnKF update on Gaussian marginal distributions. In fact, such transformations will increase the non-linearity between the transformed parameter and state variables (Zhou et al., 2011). Even so, NS-EnKF of Zhou et al. (2011) was shown to outperform EnKF in their synthetic case in terms of the successfully retrieved connectivity of the $\log K$ fields and match of state observations. A drawback however was that an ensemble size of 1000 was required in order to avoid a collapse of the ensemble variance of final log hydraulic conductivities. Given that ensemble sizes of no more than $O(100)$ are used in most applications (Wen and Chen, 2005; Chen and Zhang, 2006; Gillijns et al., 2006; Evensen, 2009a; Anderson, 2009), the computational cost associated to ensembles of thousands of members can indeed be considered a limiting factor. Moreover, Zovi et al. (2017) showed in another synthetic case using an ensemble size of 2000 that the ability of NS-EnKF to retrieve

consistent structures also depended on the number of conditioning observations points in the modelled domain.

Mainly motivated by the NS-EnKF approach proposed by Zhou et al. (2011), this synthetic study aims to explore, in a non-Gaussian parameter case, the potential of using a similar normal-score approach for the application of the iterative ensemble smoother ES-MDA (Emerick and Reynolds, 2012a). Indeed, in the previous multi-Gaussian log K case, we have shown that the normal-score transform of state variables could especially benefit the simultaneous assimilation of hydraulic head and flowrate measurements with ES-MDA.

Using the same model set up presented in the previous multi-Gaussian log K case (cf. Figure 3.1) and in which sparse observation points are especially considered, the extent of the effects introduced by parameter and state transformations for the inversion of non-Gaussian log K fields will be assessed. Although previous EnKF applications considered the normal-score transform of hydraulic head data only (Zhou et al., 2011; Li et al., 2012; Schoniger et al., 2012; Zovi et al., 2017), we will consider in addition the feasibility of the normal-score approach when assimilating both head and flowrate measurements for our specific case.

4.2 Methods

This section first introduces ES-MDA, an iterative ensemble smoother proposed by Emerick and Reynolds (2012a), before explaining how the normal-score transform will be combined with the ensemble method to yield an “iterative normal-score ensemble smoother”.

4.2.1 ES-MDA

The “Ensemble Smoother with Multiple Data Assimilation”, abbreviated ES-MDA, relies on the ensemble smoother (ES) update formula which can actually be viewed as an extension of the ensemble Kalman filter (EnKF). Indeed, the equation of the ES update is the same as the one used in EnKF, except that here the vector of observations considers the observations of all the time points (van Leeuwen and Evensen, 1996). Hence both ES and ES-MDA assimilate the whole set of transient observations during the update step. Unlike ES which performs

the update only once, ES-MDA performs the data assimilation a predefined number of times with a covariance of the measurement errors C_{err} in ES-MDA which is inflated by a factor α . Each realization m_i of the ensemble of size N is updated at iteration $k + 1$ as follows

$$m_i^{k+1} = m_i^k + C_{MD}^k (C_{DD}^k + \alpha_{k+1} C_{err})^{-1} (d_{obs,i}^k - g(m_i^k)) \quad (4.1)$$

where $d_{obs,i}^k$ is the vector of observations perturbed by a vector of inflated Gaussian noise sampled from $N(0, \alpha_{k+1} C_{err})$ at iteration $k + 1$, and C_{MD}^k and C_{DD}^k are respectively the cross-covariance matrix between the parameter vector m_i and the vector of predicted data $g(m_i)$ and the auto-covariance matrix of predicted data which are both approximated from the ensemble at iteration k . It is worth noting that except from the batch processing and inflated covariance of the measurement errors, the ES-MDA update has the expression as in EnKF.

Because 4.1 was derived based on the equivalence between single and multiple data assimilations in the linear Gaussian case (Emerick and Reynolds, 2012a), the application of ES-MDA requires to satisfy the following condition on the N_a inflation factors α_k

$$\sum_{k=1}^{N_a} \frac{1}{\alpha_k} = 1 \quad (4.2)$$

For simplicity, constant factors equal to the number of assimilations N_a will be considered. Note that several optimized versions of ES-MDA which adapt α_k based on the magnitude of the data mismatch have been proposed based on this condition (Emerick, 2016; Le et al., 2016). The reason why ES-MDA succeeds in improving the data match obtained with ES in the nonlinear case lies in the equivalence between ES, or equivalently ES-MDA with only one iteration, and one single iteration of the Gauss-Newton optimization procedure using a full step-length and an average sensitivity estimated from the prior ensemble (Emerick and Reynolds, 2012a). Multiple smaller updates resulting from several ES-MDA iterations will therefore lead to a more accurate estimation as would be the case when using the Gauss-Newton method with shorter step lengths.

4.2.2 Iterative normal-score ensemble smoother

Similarly to the benefit which has been observed with NS-EnKF over EnKF (Zhou et al., 2011; Li et al., 2012), the application of ES-MDA with normal-score transform aims to improve the conditioning to transient state observations of non-Gaussian realizations of log hydraulic conductivities ($\log K$). It is worth noting that here, compared to the previous multi-Gaussian case for the same synthetic groundwater flow problem addressed in Chapter 3, the transient state observations are non-Gaussian as a result of both the nonlinear flow dynamics and the non-Gaussian hydraulic parameters. Given a generated initial ensemble of non-Gaussian $\log K$ fields and a nonlinear forward model operator, the approach proceeds as follows:

1. Prediction (or forecast) step. An ensemble prediction is run from $t = 0$ to the time of the last observation in the time series using the hydraulic conductivity fields.
2. Normal-score of the original parameter variables. For each grid cell, a normal-score function ϕ_k is defined by mapping the p -quantile of the ensemble distribution of the initial $\log K$ values to the p -quantile of a standard normal distribution. In this manner, an ensemble of transformed $\log K$ fields with Gaussian marginal distributions is obtained.
3. Update (or analysis) step. The transformed parameters are updated with ES-MDA by considering in (4.1) the transformed parameters and transformed data instead of the original variables. For a detailed description of the computational implementation of the matrix operations in the ES-MDA update, especially when using the subspace inversion procedure with rescaling, the reader is referred to Appendix B of Emerick (2016).
 - (a) Definition of the normal-score transform functions. A normal-score transform (NST) function denoted ψ_k , for each data type, is calculated based on the time series of state variables predicted at each one of the k observation locations (Figure 3.1) which we perturb with noise beforehand. Indeed, we perturb the predicted data using the same inflated measurement errors added to the observations to assimilate. This NST function maps the p -quantile of the ensemble distribution of original perturbed predicted values at one lo-

cation to the p -quantile of a standard normal distribution. In this manner, at each observation location, the ensemble of transformed perturbed predicted head or flowrate data when considering all the points in the time series will be Gaussian. We will particularly discuss in the results of the non-convergence of the proposed approach when assimilating transformed flowrates in addition to the transformed heads.

- (b) Transformation of the predicted data (without perturbation). The data mismatch term in the ES-MDA update equation (4.1) involves the predicted data vector without the inflated noise. This predicted data vector is transformed according to the previously defined NST functions ψ_k depending on the location of the predicted data.
- (c) Transformation of the perturbed observations. Similarly to the transformation of the predicted data vector, the inflated observation vector is also transformed according to ψ_k depending on the location of the observed data. By applying the same local normal score functions as used to transform the predicted data vector, both transformed vectors can be compared properly when evaluating the data mismatch term in (4.1). Note that in NS-EnKF (Zhou et al., 2011), the predicted and observed data at a same observation point are both transformed using the same normal-score function as well. However, no inflated measurement errors need to be taken into account when defining the normal-score functions in NS-EnKF.
- (d) Covariance between the transformed parameters and transformed predicted data. This covariance matrix C_{MD} is calculated between the transformed parameters and the transformed predicted data vector without the inflated noise.
- (e) Calculation of the pseudo-inverse of the transformed matrix. The assimilation of a large data set compared to the ensemble size requires to calculate the pseudo-inverse of the matrix $C_{DD}^k + \alpha_{k+1}C_{err}$ using the subspace inversion procedure of Evensen (Evensen, 2004; Emerick, 2016). As a result, the terms C_{DD}^k and C_{err} needs to be transformed separately in order to be able to calculate the pseudo-inverse of the transformed matrix. C_{DD}^k is calculated simply based on the normal-score transform of the predicted data without perturbation, while the transformed measurement error covariance

$\widehat{C_{err}}$ is determined so as to respect the following ratio

$$\widehat{C_{err}}(C_{err})^{-1} = \widehat{C_{D'D',diag}}(C_{D'D',diag})^{-1} \quad (4.3)$$

where C_{err} is the untransformed measurement error covariance which is often assumed diagonal, $\widehat{C_{D'D',diag}}$ is the diagonal matrix constructed from the diagonal elements of the covariance of the transformed perturbed predicted data, and $C_{D'D',diag}$ is the untransformed equivalent.

4. Back transform the updated transformed parameters. The transformed parameters are transformed back to log K values using the inverse of the normal-score transform defined in Step 2, noted ϕ_k^{-1} .
5. The algorithm goes back to Step 1 until the last ES-MDA update and back-transform step. It is important to mention that the normal-score transform functions ϕ_k applied in Step 2 in the next iterations are still the one built from the initial ensemble of log K realizations. As pointed out by Zhou et al. (2011) for NS-EnKF, keeping the same function helps the updated log K distribution to not deviate too much from the initial distribution.

4.3 Synthetic test case

The performance of the proposed iterative normal-score ensemble smoother was evaluated on the synthetic inverse problem inspired by the real situation observed at the Andra's Underground Research Laboratory site in Meuse/Haute Marne (cf. Introduction in Chapter 1). The model set up including the model geometry, the boundary conditions, the number of observation points and their location are described in the previous chapter in Section 3.3.1.

4.3.1 Synthetic data set

The synthetic data set is composed of the transient heads observed at the ten observation points, as well as the transient flowrates produced at five different depths along the seepage boundary (Figure 3.1). They were simulated by a groundwater flow simulation in saturated conditions given the reference field of log hydraulic conductivities (log K) shown in

Figure 4.1b and a constant specific storage value of 10^{-6} m^{-1} over the whole domain. This log K reference field results from two simulation steps. As shown in Figure 4.1a., a categorical binary field was first simulated by multiple-point statistics (MPS) using as the conceptual model the training image (TI) shown in Figure 4.2. This TI of size 1000×100 was extracted from a shrunk version in the vertical dimension of the 2500×2500 hand drawn image by Zahner et al. (2016) inspired by Strebelle (2002). Then, continuous log K values extracted from two multi-Gaussian simulations generated by the Fast Fourier Transform method of respectively mean value -4 and -6 and variance 0.25 for each were assigned to each category of the binary field. As a result, the reference log K field contains two different scales of heterogeneity, one at the scale of the model with the presence of two facies, and another smaller one at the scale of the facies. In the real field case, the more connected facies, i.e. the channels, could represent the more permeable porous horizons “HP” within the less permeable layers.

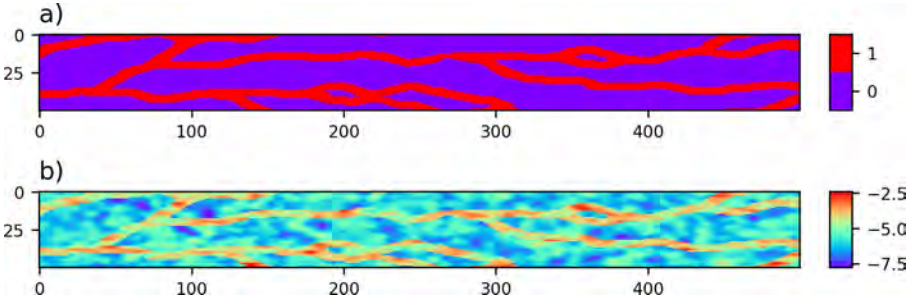


Figure 4.1: a) Binary multiple-point simulation used to create the log hydraulic conductivity reference. b) Reference field of log hydraulic conductivities.

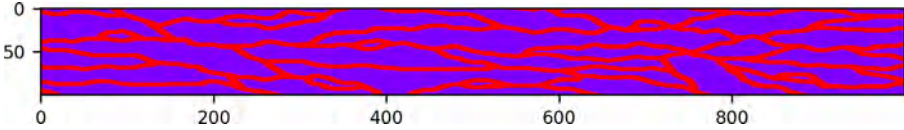


Figure 4.2: Training image with two categories used to generate the multiple-point simulations with DeeSse.

4.3.2 Initial ensemble

An ensemble of bimodal log K realizations was generated according to the same two simulation step used to generate the reference log K fields. Similarly to the reference, all log K realizations contain underlying categorical patterns from the pre-generated categorical multiple-point simulation obtained with DeeSse (Straubhaar, 2017).

Given the present problem set up, one objective of the synthetic numerical experiments will be to evaluate whether the proposed data assimilation approach can allow to retrieve or not structures that are consistent to the underlying structures of the initial ensemble.

4.4 Performance criteria

In this non-Gaussian parameter case, the fact that we are dealing with bimodal log hydraulic conductivity (log K) realizations showing curvilinear structures calls for a criterion to assess the plausibility of the conditioned fields. Indeed, neither the RMSE nor ensemble spread can inform about the shape of the log K distribution or the consistency of the structures after the assimilation. Unlike the multi-Gaussian case where the prior correlation lengths of the log K realizations would be essentially maintained throughout the assimilation (Emerick and Reynolds, 2011), the prior non-Gaussian structures and connectivity patterns considered here are bound to deteriorate from the first assimilation of ES-MDA if no mitigation strategies are employed.

We will measure the effects of the normal-score transform on the parameter and state variables on the accuracy of the ES-MDA estimates, as we did previously in the multi-Gaussian case. To evaluate the plausibility of the structures retrieved after the assimilation, we will compute the connectivity function (Renard and Allard, 2013) of the facies of value s_i , noted $\pi_i(h)$,

$$\pi_i(h) = Prob\{x \leftrightarrow x + h | s(x) = s(x + h) = s_i\} \quad (4.4)$$

which describes the probability that two points separated by a distance h of the same facies of value s_i are connected (4.4). We will compute it before and after the data assimilation for the facies of value 1 which represents the channels in our case.

Calculating these connectivity functions for each facies of the ini-

tial ensemble of $\log K$ realizations is straightforward if we consider the ensemble of categorical facies realizations which were simulated with DeeSse (Straubhaar, 2017) before the assignment of the continuous $\log K$ values. The problem arises when one wishes to calculate this connectivity after the data assimilation. Indeed, since the iterative ensemble smoother here is employed to directly update the $\log K$ values independently from the facies it comes from, the underlying categorical patterns of the initial $\log K$ realizations will not be preserved. It will therefore be necessary to categorize the inverted $\log K$ fields in order to calculate the new connectivity functions and compare them to the initial ones. The problem with such categorization is the threshold value which is required in order to associate each $\log K$ value to either one of the two facies. Depending on how close the number of modes and close they are, finding these thresholds may require the use of a partitioning technique such as k-means clustering. For this simple bimodal case with relatively distant modes, we will consider a threshold value of -5 which corresponds to the average of the two modes of the $\log K$ distribution in order to identify the channel from the matrix facies.

4.5 Results and discussion

The proposed approach was first applied for the assimilation of head observations with only one iteration of ES-MDA, hence using simply ES. As shown in Figure 4.3 which shows the parameter field at different stages of the proposed approach, the update applied on the transformed initial $\log K$ field (Figure 4.3b) breaks distinctly the underlying categorical patterns of the initial field (Figure 4.3c).

Also, the visible discontinuities between updated and non-updated parameters due to the localization (Figure 4.3c) raises the issue of how to preserve through the update the continuity of prior connected structures when going from an informed to a non-informed region. Indeed, such transitions will be encountered for any modelled situation where observations are nearly nonexistent in areas of the model, such as in this synthetic case (Figure 3.1).

We note however that the bimodality of the $\log K$ field could be successfully retrieved after the back normal-score transform (we do not show the distribution here for brevity). As mentioned in Section 4.2.2, the local back normal-score functions are defined based on the initial

ensemble of $\log K$ values, thereby helping to recover the bimodal distribution after each back transform of the updated transformed fields.

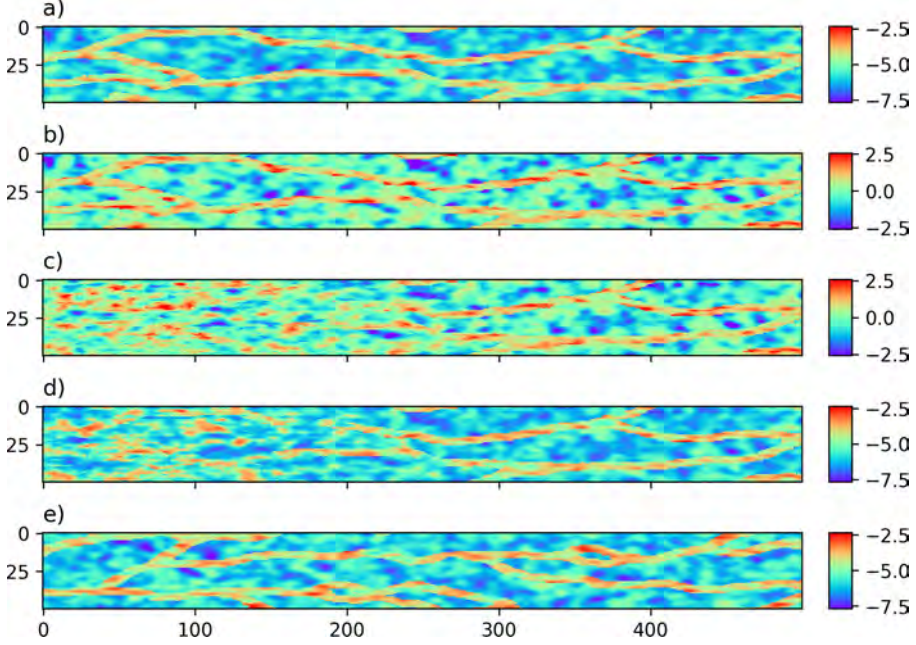


Figure 4.3: a) Initial log hydraulic conductivity ($\log K$) field no. 24 from the ensemble. b) Normal-score transformed $\log K$ field. c) Updated transformed field with ES (i.e. ES-MDA with one single iteration). d) $\log K$ field after back-transform of the updated transformed field. e) Reference field.

As expected by the already nonlinear groundwater flow dynamics in the multi-Gaussian $\log K$ case, ES-MDA needs to be iterated a larger number of times in order to reach an acceptable data match. In fact, the nonlinearity of the flow is even more important when using an ensemble of bimodal $\log K$ realizations. We can observe this nonlinearity by looking at the Q-Q plot of the ensemble distribution of the normalized bimodal $\log K$ values against the ensemble distribution of the normalized predicted states at different observation locations. As shown in Figure 4.4, both distributions don't have the same shape since the quantiles from each distribution plotted against each other are far from being aligned along a straight line (here the bissector more particularly since the distributions were normalized beforehand). This implies that the two distributions are not related by a linear transformation. The

same Q-Q plot in the less nonlinear multi-Gaussian log K case is shown in the Appendix for comparison.

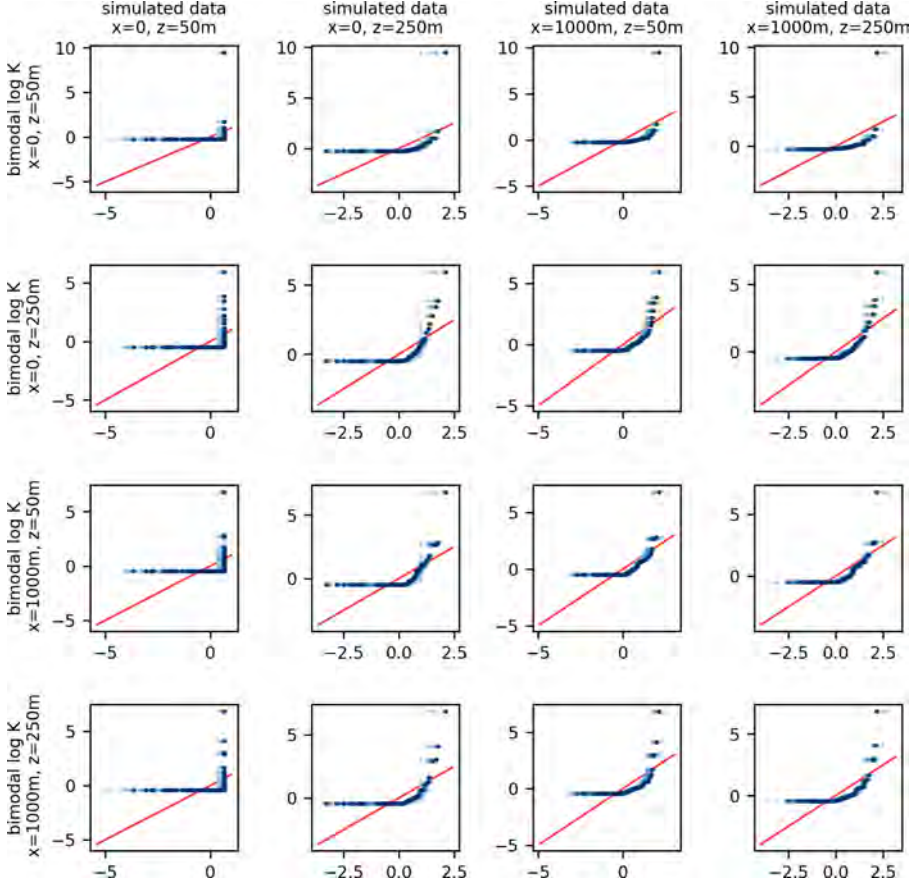


Figure 4.4: Q-Q plots of the bimodal log K values against the predicted state observations at several observation locations.

Still, after performing ES-MDA with 8 iterations, Figure 4.5 illustrates that the consistency of the back-transformed log K fields did not improve despite the reduced data mismatch (Figure 4.7). By assessing the variability of the ensemble of log K realizations in Table 4.1, we note that the ensemble underestimated significantly the true error which actually increased during the assimilation. Table 4.1 also includes results from the assimilation of head data with ES-MDA($\times 8$) without performing the normal-score on the parameters or on the data in order

to illustrate that the benefit on the ensemble spread comes mainly from the transformation of the parameters.

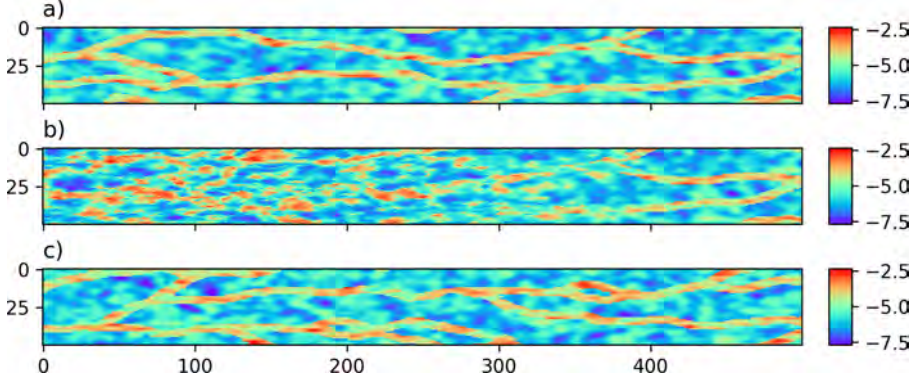


Figure 4.5: a) Initial log hydraulic conductivity field no. 24 from ensemble of 100. b) Back-transformed log hydraulic conductivity field after updating the normal-score transformed initial field with ES-MDA($\times 8$). c) Reference field.

Ens. size	Algorithm	Assimilated data	RMSE	Ens. spread
100	ES	transf. h	1.04 - 1.15	1.03 - 0.88
100	(*)ES-MDA($\times 8$)	h	1.04 - 1.72	1.03 - 0.07
100	ES-MDA($\times 8$)	h	1.04 - 1.58	1.03 - 0.20
100	ES-MDA($\times 8$)	h, q	1.04 - 1.87	1.03 - 0.12
100	ES-MDA($\times 8$)	transf. h	1.04 - 1.30	1.03 - 0.28
1000	ES-MDA($\times 8$)	transf. h	1.03 - 1.03	1.04 - 0.88

Table 4.1: Influence of the normal-score transform of parameters and/or data, on the RMSE, ensemble spread, and coverage calculated with ES-MDA. The indication (*) means that the parameters were not transformed for that test.

As noted by Zhou et al. (2011), NS-EnKF could achieve accurate estimations of the uncertainty in their synthetic case provided that an ensemble size of 1000 was considered. For this reason, it could be argued that our approach was not successful owing to the significantly smaller ensemble size of 100 used. Interestingly, when we increased the size to 1000 realizations, we observed that the algorithm did not manage to re-

duce the data mismatch as efficiently as when we used an ensemble size of 100 (Figure 4.8). As illustrated in Figure 4.6 with the mean log K field obtained after the assimilation using an ensemble size of 100 or 1000, the ensemble of 1000 conditioned log K fields, on average, did not allow to sufficiently identify the underlying structures compared to the results of the case of Zhou et al. (2011). This insufficient characterization is however consistent with the small decrease of the ensemble variability, as shown in Table 4.1. Given the resulting enlarged space spanned by the much larger ensemble size, this still large reduced uncertainty after the data assimilation is indicative of an insufficient constraint from the conditioning data. Restarting the ES-MDA by predefining a larger number of iterations may allow to improve the data match but such a test would be computationally expensive. Based on our current results, it seems that a larger number of data would be required in order to sufficiently constrain the identification of the structures in our non-multi-Gaussian case using a normal-score approach. We note however that the hinted structures seem more consistent than those yielded by the small ensemble. Although we do not show it here, the individual conditioned log K fields from the large ensemble were still not sufficiently consistent relatively to the initial ensemble.

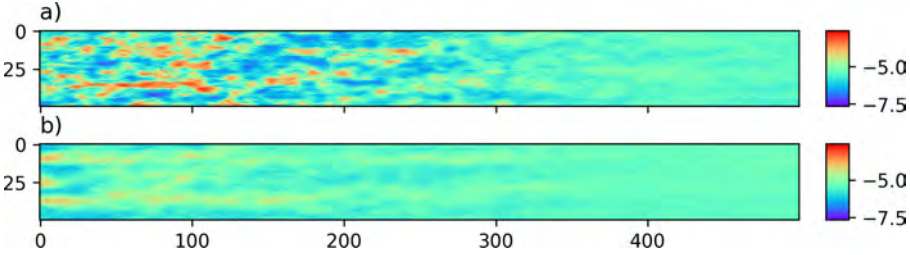


Figure 4.6: Mean fields obtained after ES-MDA($\times 8$) when using an ensemble size of 100 (a) and 1000 (b).

Figure 4.9 illustrates the visible loss of connectivity observed for the more permeable facies, i.e. the channels, after the assimilation of the head data whether using ES or ES-MDA($\times 8$) with an ensemble size of 100, or whether using ES-MDA($\times 8$) with an ensemble size of 100 or 1000. As we can observe, although the connectivity improved in the case of the large ensemble, the ensemble of connectivity values no longer includes the connectivity value of the reference or of most of the initial ensemble.

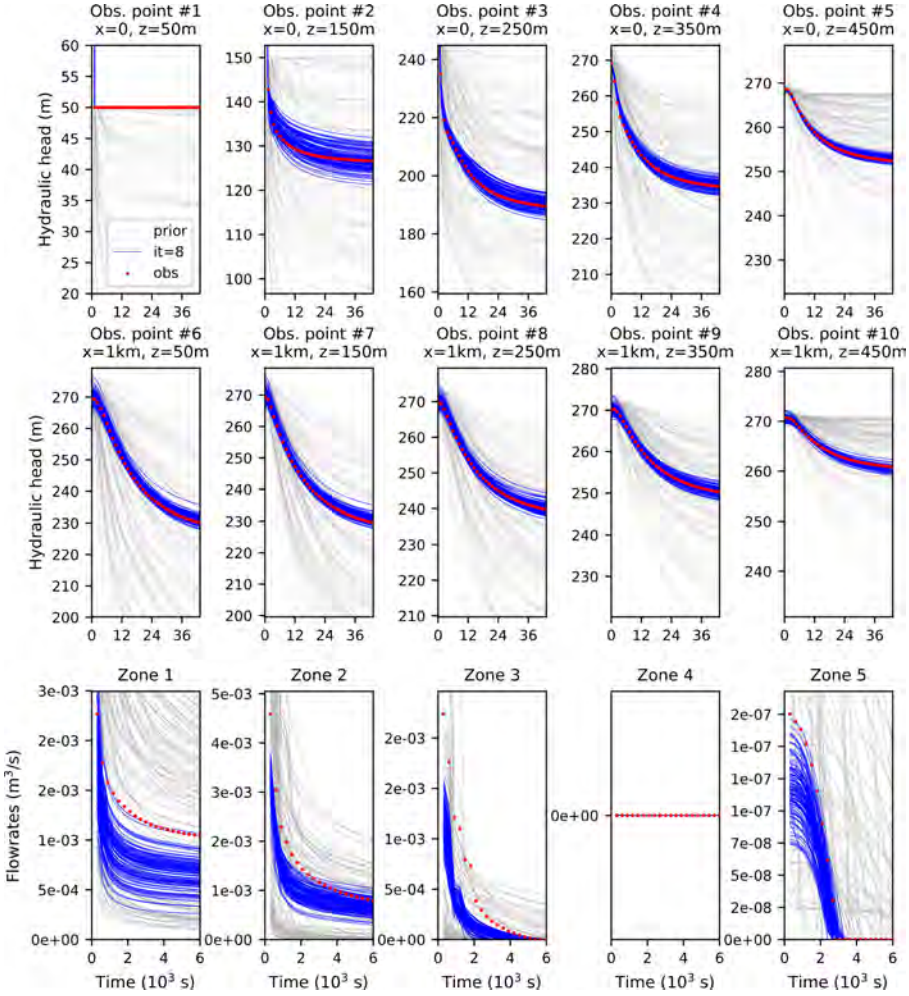


Figure 4.7: Predicted hydraulic head and flowrate data at the observation locations, before (in grey) and after (in blue) the assimilation of the transformed hydraulic heads with ES-MDA($\times 8$) using an ensemble size of 100. The red dots are the observed data.

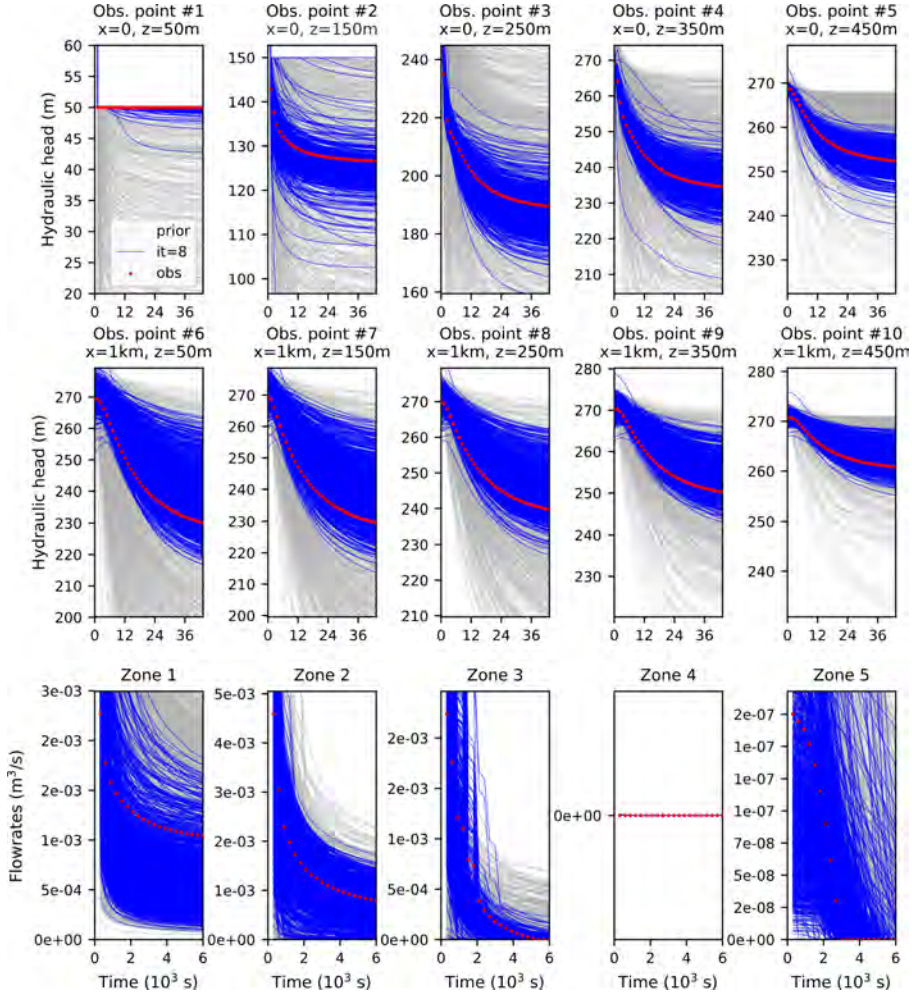


Figure 4.8: Predicted hydraulic head and flowrate data at the observation locations, before (in grey) and after (in blue) the assimilation of the transformed hydraulic heads with ES-MDA($\times 8$) using an ensemble size of 1000. The red dots are the observed data.

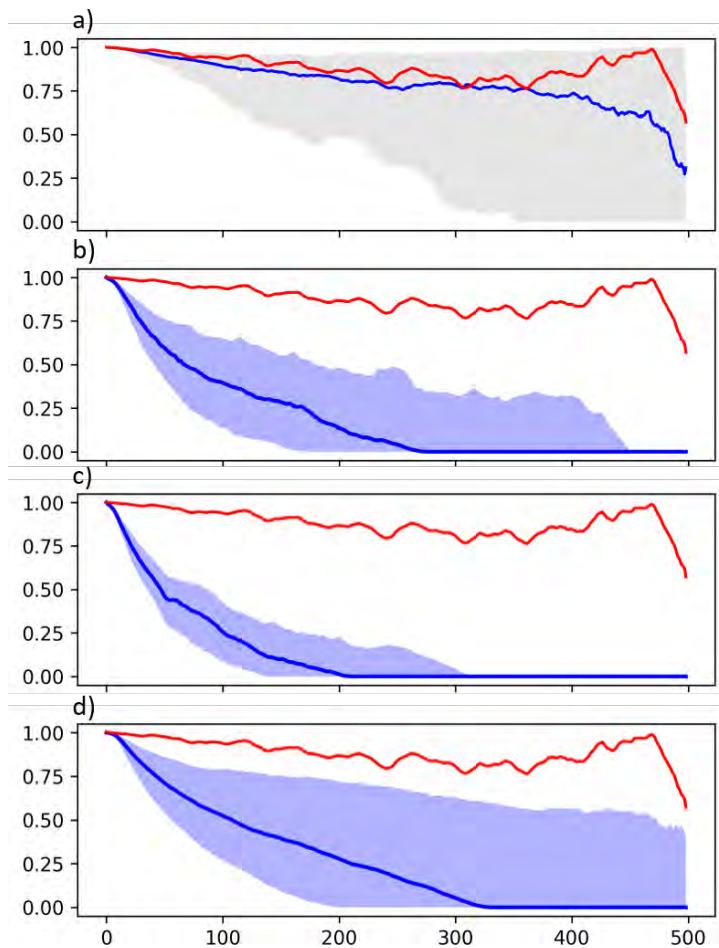


Figure 4.9: Connectivity function of the channel facies computed in the horizontal direction for the ensemble of realizations: (a) before the update (b) after ES using 100 realizations (c) after ES-MDA($\times 8$) using 100 realizations (d) after ES-MDA($\times 8$) using 1000 realizations. The grey and blue regions correspond to the values between the percentiles P5 and P95. The red curve is the connectivity function calculated with reference A and the blue curve is the median of the ensembles of values taken by the connectivity function.

4.6 Conclusion

In this chapter, we applied the iterative ensemble smoother ES-MDA (Emerick and Reynolds, 2012a) by considering a normal-score transform on both the parameter and state variables as proposed by Zhou et al. (2011) in the normal-score ensemble Kalman filter (NS-EnKF). The use of this normal-score approach was motivated by the fact that both EnKF and ES-MDA share the same update equation as the Kalman filter, thereby potentially yielding from ES-MDA the similar advantages observed with NS-EnKF in the non-Gaussian case (Zhou et al., 2011; Li et al., 2012; Zovi et al., 2017). When using an ensemble size of 100, the approach did not allow the retrieval of the initial underlying structures during the assimilation of the head observations. As a result, the error increased significantly. Therefore, we can first conclude that the Gaussian marginal distributions were not sufficient to alleviate the effects of the still non-multi-Gaussian prior distribution. We note, however, that the uncertainty was less underestimated with the use of this normal-score on both the parameter and state variables than without. Even after increasing the ensemble size to 1000, the improvement was not sufficient, most probably because the sparse conditioning data points in the horizontal direction are not sufficient to constrain the search of the enlarged parameter space. Besides, although the iterative normal-score ensemble smoother converged with the assimilation of the head observations, it did not when assimilating both the head and flowrate data. Since the same definition of the normal-score was used for the successful assimilation of both data types in the previous multi-Gaussian case, we conjecture that it is the increased nonlinearity resulting from the transformation of both the parameters and states, as also noted by Zhou et al. (2011), which prevented the successful simultaneous assimilation of both data types. As it turns out, when only the parameters are transformed, thereby removing the nonlinearity due to the transformed states, the assimilation of both data types effectively allows ES-MDA to converge. It therefore seems that in the non-Gaussian case, the normal-score should be applied only on the non-Gaussian parameter variables in order to not prevent the assimilation of certain data types such as flowrates for the parameter identification.

Chapter 5

Conditioning truncated Gaussian realizations to transient hydraulic head and flowrate data with iterative ensemble smoothing

5.1 Introduction

In order to apply ensemble Kalman methods in the non-Gaussian parameter case, Zhou et al. (2011) proposed the NS-EnKF approach which is based on a normal-score transform of the parameter and state variables prior to the EnKF update step. We previously concluded that for our synthetic test case, a similar normal-score approach based on the iterative ensemble smoother could not sufficiently mitigate the effects of the non-Gaussian joint distributions of the transformed log hydraulic conductivities. Indeed, the data assimilation did not result in sufficiently consistent fields given the initial geological information. The large distance separating some of the observations points, combined with the large ensemble size requirement seemed to be the main limiting factors to the successful application of the approach.

From Chapter 3, we remind that despite the nonlinear flow dynamics of the test problem, the iterative ensemble Kalman method could per-

form efficiently when the prior ensemble is multi-Gaussian. It therefore seems that ensemble Kalman methods are more robust against deviations from linearity of the forward model operator rather than deviations from the multi-Gaussian assumption. However, the assumption made in Chapter 3 that groundwater flow parameters such as the logarithm of hydraulic conductivities could be described by a multi-Gaussian distribution may be insufficient for a wide variety of subsurface modelling problems (Gómez-Hernández and Wen, 1998). Indeed, such multi-Gaussian realizations are not suitable to represent a petrophysical property with sudden transitions of values or with enough complexity in the spatial structure while such aspects could significantly affect the groundwater flow and transport predictions (Zinn and Harvey, 2003; Kerrou et al., 2008).

In general, the prior geological knowledge is first integrated in conceptual models in the form of categorical features, leading to what we commonly refer as “facies models”, before continuous properties are considered for each facies. As a matter of fact, NS-EnKF has been applied with the aim of retrieving structures initially generated by a discrete simulation technique that used the multiple-point statistics of a categorical training image (Zovi et al., 2017). Another type of geostatistical simulation techniques commonly used to generate discrete realizations is the Truncated Gaussian method (TG) or its variant, the Pluri-Gaussian method (PG) (Armstrong et al., 2003). Both methods are similar in the way they first simulate one or more multi-Gaussian random fields before performing a truncation which produces the categorical field.

In the context of an application of ensemble Kalman methods in the discrete case, such underlying Gaussian variables provide a practical way to indirectly condition the categorical realizations while still honoring the multi-Gaussian assumption. Previous reservoir applications have actually illustrated the potential of a parameterization based on these underlying multi-Gaussian fields for the conditioning of truncated Gaussian-based realizations using the ensemble Kalman filter or the iterative ensemble smoother (Liu and Oliver, 2005; Zhao et al., 2008; Astrakova and Oliver, 2015).

From basic examples using the truncated Gaussian and bi-Gaussian simulation techniques to represent the geological information in our synthetic test case, this chapter aims to learn about the effects of using multi-Gaussian underlying variables for the application of the iterative ensemble smoother in a discrete case. In particular, the applicability

of the normal-score transform in such nonlinear cases is addressed, as well as the effects of the nonlinear truncation step on the uncertainty of the final conditioned ensemble of facies fields. For a proper analysis of the results, we introduce in this chapter new performance criteria suitable for the discrete case which are in fact comparable to the standard criteria commonly used in the multi-Gaussian continuous case.

5.2 Methods

5.2.1 Generation of facies realizations with the truncated Gaussian or bi-Gaussian simulation technique

The truncated Gaussian method relies on the truncation of a Gaussian realization y of a standard Gaussian function $Y(x)$. The number of thresholds is equal to $N - 1$ with N the number of facies simulated. Two thresholds s_1 and s_2 would hence be needed to truncate a Gaussian realization into a discrete distribution with three facies. Assuming known stationary facies proportions, the thresholds are determined by using the inverse of the empirical cumulative density function G^{-1} of the Gaussian realization as a function of the facies proportions:

$$s_i = G^{-1} \left(\sum_{j=1}^{N-1} p_j \right) \quad (5.1)$$

where p_i is the proportion of facies i , s_i is the upper threshold for facies i with the convention that $s_0 = -\infty$ and $s_N = +\infty$.

One limitation of the truncated Gaussian method is that the simulated facies will necessarily be in an ordered sequence. An alternative in order to add more flexibility in the facies distribution would be to use the bi-Gaussian method. The principle is the same as the truncated Gaussian method except that the thresholds are applied to two Gaussian realizations instead of one. According to the pair of coordinates given by each Gaussian value $Y_1(x)$ and $Y_2(x)$ simulated at each cell of the simulation grid, a facies type will be assigned according to its position in a truncation threshold map to define based on information about the contacts between the facies and their proportions.

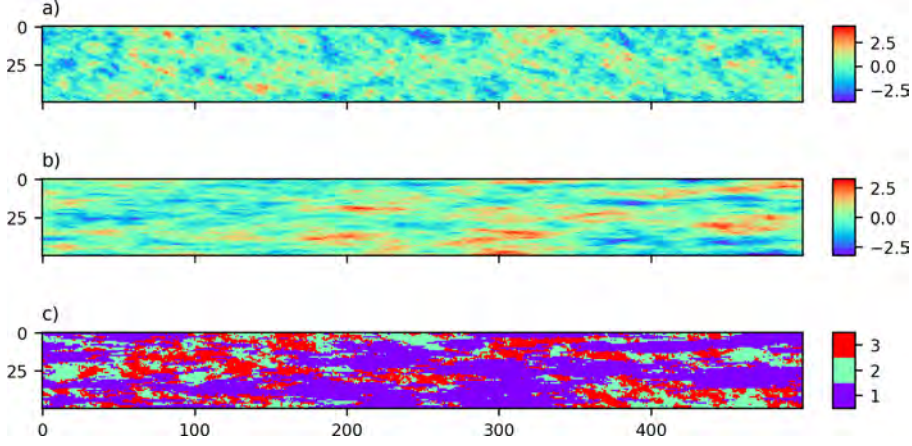


Figure 5.1: Illustration of a categorical pluri-Gaussian realization (c) based on the truncation of two multi-Gaussian realizations (a and b).

5.2.2 Application of ES-MDA

For the same synthetic test case described in Section 3.3.1, the iterative algorithm ES-MDA was considered in order to address the nonlinear dynamics of the transient groundwater flow problem. For a description of the algorithm, the reader is referred to Section 4.2.1.

Parameterization

As a reminder, at each iteration $k + 1$ of ES-MDA, the vector of model parameters m_i is updated according to

$$m_i^{k+1} = m_i^k + C_{MD}^k (C_{DD}^k + \alpha_{k+1} C_{err})^{-1} (d_{obs,i}^k - g(m_i^k)) \quad (5.2)$$

where $d_{obs,i}^k$ is the vector of observations perturbed by a vector of inflated Gaussian noise sampled from $N(0, \alpha_{k+1} C_{err})$ at iteration $k + 1$, and C_{MD}^k and C_{DD}^k are respectively the cross-covariance matrix between the parameter vector m_i and the vector of predicted data $g(m_i)$ and the auto-covariance matrix of predicted data which are both approximated from the ensemble at iteration k .

Given the multi-Gaussian assumption of ensemble Kalman methods, the underlying multi-Gaussian variables which define the facies distribution are the most suitable variables to be estimated. Hence m_i is composed of the initial variables or updated variables from the previous

step associated to the multi-Gaussian field Y_1 for the truncated Gaussian case, or associated to the two multi-Gaussian fields Y_1 and Y_2 for the bi-Gaussian case. If the underlying field(s) have the same resolution than the flow simulation grid of size n_{grid} , then the number of parameters to be estimated is n_{grid} or $2 \times n_{grid}$ respectively for the truncated Gaussian and bi-Gaussian cases.

Generation of the conditioned facies fields

Unlike in the previous multi-Gaussian and bimodal cases, the iterative ensemble smoother ES-MDA do not update the log hydraulic conductivity values in a post-processing procedure of the initial ensemble generated. In this application, it will correct the multi-Gaussian variables before the truncation step, hence before a new facies fields is generated. Thus a major difference compared to the application of ES-MDA on the hydraulic property fields directly is the need to regenerate the conditioned truncated Gaussian realizations after each update of the underlying variables.

If the proportions are known, the thresholds of the truncation diagram can be derived after each update based on the new empirical cumulative distribution obtained. If the proportions are not known, the thresholds can be chosen to be fixed and the proportions will be allowed to vary during the course of the assimilation.

Based on the facies distribution obtained, hydraulic parameters such as the hydraulic conductivities can be assigned according the facies type for every grid cell in order to perform the groundwater flow simulation.

Although the facies distribution or the log hydraulic conductivities inside the facies are not updated directly with the ensemble Kalman method, this parameterization based on the underlying variables allows ultimately to condition the facies distribution, even indirectly, from the flow data.

5.2.3 Synthetic test cases

Two synthetic tests based on an initial ensemble of respectively unconditional truncated Gaussian realizations and unconditional bi-Gaussians realizations were considered for the application of ES-MDA in a discrete framework. Although previous studies have addressed the challenge of preserving the conditioning of the initial ensemble to hard

observations of facies type through the ensemble-based data assimilation procedure (Liu and Oliver, 2005; Astrakova and Oliver, 2015), this question was beyond the scope of the following test cases.

Apart from the simulation technique employed to generate the categorical fields, the problem set up including the model geometry, boundary conditions, location of the observation points, is exactly the same as for the previous cases and is described in Section 3.3.1. In both tests, the initial ensemble of categorical fields was generated in the same manner as the reference in order to consider the variability of the underlying Gaussian field(s) as the only source of uncertainty in the parameter estimation problem.

Truncated Gaussian case

A reference truncated Gaussian realization with proportions $p_1 = 50\%$, $p_2 = 20\%$ and $p_3 = 30\%$ respectively for the facies 1, 2 and 3 was created (Figure 5.2a). To mimic horizontal structures, an exponential variogram model with horizontal and vertical correlation lengths of respectively 120 and 12 gridblocks was assumed for the underlying multi-Gaussian realization in order to yield horizontal structures after the truncation. The resulting facies 1, 2 and 3 were populated with constant log hydraulic values -6 , -5 , -4 respectively (Figure 5.2b) in order to generate after a flow simulation the synthetic data set of hydraulic heads and flowrates at the observation points (Figure 6.8). Given the horizontal extent of the drawdown shown in Figure 5.3, a critical length of 4000 m was defined in order to evaluate the Gaspari-Cohn function used to localize the Kalman gain matrix.

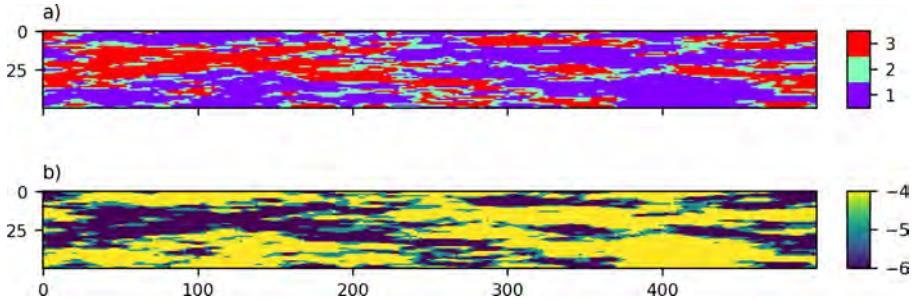


Figure 5.2: a) Truncated gaussian realization used as the categorical reference. b) Log hydraulic conductivity field (m/s) resulting from populating the categorical reference.

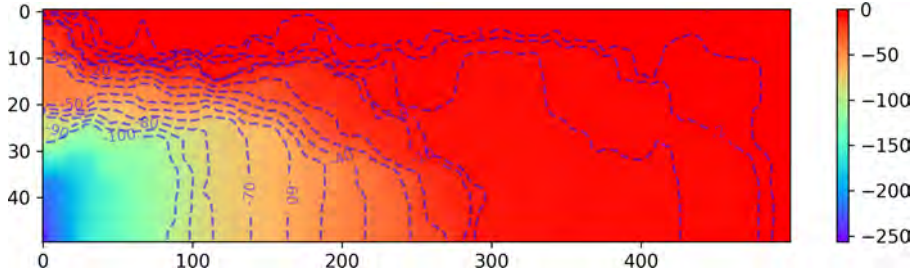


Figure 5.3: Drawdown (m) computed at the end of the transient groundwater flow simulation ($t=43\,200\text{ s}$) using the log K field obtained from the truncated Gaussian reference.

A reference bi-Gaussian realization with the same proportions $p_1 = 50\%$, $p_2 = 20\%$ and $p_3 = 30\%$ of respectively log hydraulic conductivity -6 , -5 and -4 was created by truncating two multi-Gaussian realizations Y_1 and Y_2 according to the truncation threshold map shown in Figure 5.4. An exponential variogram model with correlation lengths of 20 gridblocks in the principal direction 25° north-west and 10 gridblocks in the perpendicular direction was assumed for Y_1 . Another exponential variogram model with horizontal and vertical correlation lengths of respectively 120 and 12 gridblocks was assumed for Y_2 . As shown in Figure 5.5, connections between each pair of facies are possible unlike in the truncated Gaussian reference (Figure 5.2). Given the horizontal extent of the hydraulic perturbation shown in Figure 5.6, a critical length of 3000 m was defined in order to evaluate the Gaspari-Cohn function used to localize the Kalman gain matrix.

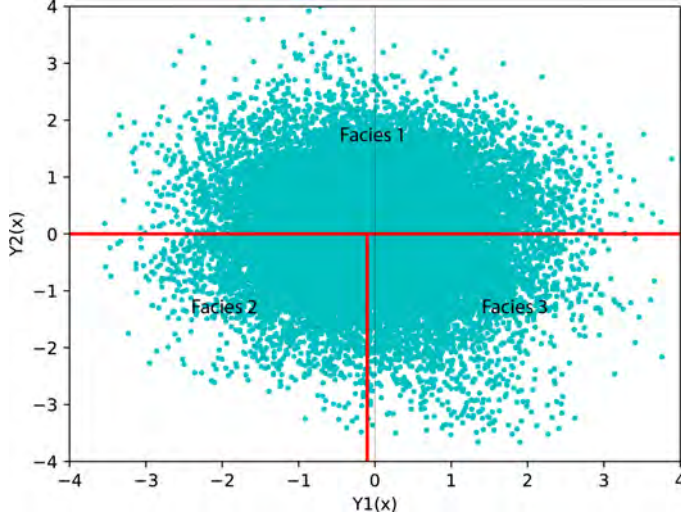


Figure 5.4: Truncation threshold map used to generate the reference bi-Gaussian field in Figure 5.5

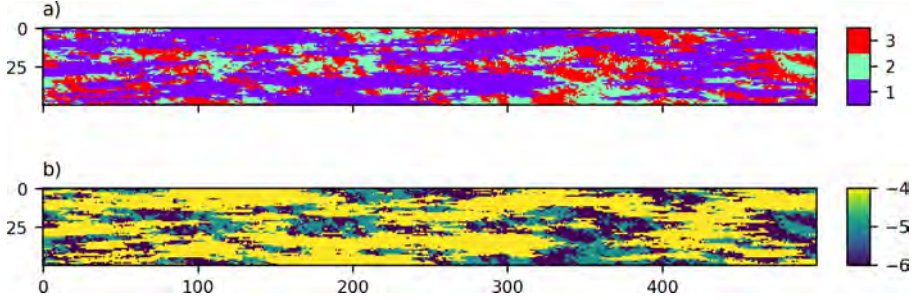


Figure 5.5: a) Bi-gaussian realization used as the categorical reference. b) Log hydraulic conductivity field (m/s) resulting from populating the categorical reference.

5.2.4 Performance criteria in the discrete case

As first defined in Section 3.4 of Chapter 3, the RMSE and the ensemble spread are the two statistical measures commonly used to assess the accuracy of the estimates of ensemble-based Kalman methods in synthetic studies (Houtekamer and Zhang, 2016). Both calculate a distance involving the mean of the ensemble of continuous estimates since

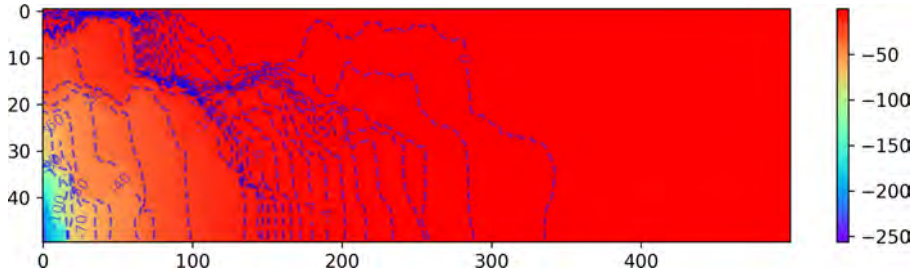


Figure 5.6: Drawdown (m) computed at the end of the transient groundwater flow simulation ($t=43\,200\text{ s}$) using the log K field obtained from the bi-Gaussian reference.

the mean can be interpreted as the best estimate in the ensemble-based framework (Evensen, 2003). In previous studies addressing the normal-score transform of parameters updated by an ensemble Kalman method (Zhou et al., 2011; Zovi et al., 2017; Li et al., 2018), it is worth noting that the RMSE and ensemble spread are not calculated using the ensemble of updated transformed variables, but using instead the ensemble of resulting log hydraulic conductivity (log K) realizations after the back-normal score transform. Even though such an analysis could be questioned given the nonlinearity of the back-transform function, it could in fact still be applied because the back-transformed fields were ultimately continuous. In the discrete example of the truncated Gaussian case, a similar analysis of the RMSE versus the ensemble spread using directly the categorical realizations is not possible. In a previous study, Astrakova and Oliver (2015) calculated these criteria using the continuous permeabilities assigned to the facies fields after the nonlinear truncation step. Indeed, the issue with an ensemble of categorical realizations is that the values of the categories will affect the calculated mean although they are assigned arbitrarily. Therefore, the calculations of the RMSE and the ensemble spread would be based on meaningless values in general in the categorical realization case.

There is actually one case where such an analysis could make sense, namely when the ensemble of realizations follows a Bernoulli distribution with a probability p of value 1 and probability $1 - p$ of value 0. In that case, the ensemble mean in the RMSE would correspond to the proportion p of value 1, while the ensemble spread would be a measure of the ensemble spread around p . Given this possible interpretation, one could consider applying the same analysis for every indicator func-

tion of the different categories. We would have to compare, on the one hand, an error between the probability distribution obtained from the ensemble and the true distribution; and on the other hand, a measure of uncertainty associated to the probability distribution of the ensemble. With this strategy, it is however not clear how each couple of calculated RMSE and ensemble spread related to one particular facies proportion could be combined together into one single global measure of error and of uncertainty associated to the whole probability distribution.

Instead of the RMSE and the ensemble spread, we suggest for the discrete case, to consider two other well-established measures which originate from information theory. These measures are interesting considering how they could be defined so as to obtain the same interpretation as respectively the RMSE and the ensemble spread in the continuous case. In place of the ensemble spread, as a measure of uncertainty, we propose here to calculate the entropy, also known as the Shannon entropy. It is estimated from the probability distribution obtained from the ensemble. The Shannon entropy provides a way of quantifying the average amount of information contained in a probability distribution. If p_i denotes the proportion of a facies category i calculated from the ensemble of categorical realizations at one gridblock, the entropy of the probability distribution is defined as:

$$H = - \sum_{i=1}^N p_i \cdot \log p_i \quad (5.3)$$

where N is the number of events and corresponds here to the number of facies categories. The entropy is hence a weighted average of the random variable $I = -\log(p_i)$, which actually defines in information theory the amount of information conveyed by the event of a category i of probability p_i . Note that the units of the entropy will depend on the base of the logarithm used. Also, we can see that compared to the ensemble spread, the entropy does not depend on the value of the random variable but only on the probabilities for every possible outcome. It is also worth noting that the uncertainty associated to the uniform probability distribution yields the highest entropy with a value equal to the logarithm of N .

As for the error relatively to the reference, it is important to define it in the same units as our measure of uncertainty, hence in terms of the units of information amount, if we want to be able to compare them. For this reason, we propose here to use the Kullback-Leibler divergence

(Kullback and A Leibler, 1951), often shortened KL divergence, or also known as the relative entropy. As we will see, its formulation only slightly differs from the Shannon entropy formula. The KL divergence is similar to the RMSE we calculated in the continuous case in that it quantifies a discrepancy between an observed reference and an estimation of that reference. But unlike the RMSE, it considers a discrepancy between two probability distributions rather than two outcomes. Given a reference probability distribution P and an approximating distribution Q of respectively proportions p_i and q_i for each of the N facies categories, the KL divergence of P and Q is

$$D_{KL}(P \parallel Q) = \sum_{i=1}^N p_i \cdot (\log p_i - \log q_i) \quad (5.4)$$

As expressed here, we can see that it actually corresponds to a weighted average of the difference between, on the one hand, the amount of information associated to the event of category i in the approximating distribution Q , and on the other hand, the amount of information of the same event in the reference distribution P . As a result, the KL divergence can be interpreted as the amount of information lost when using Q to approximate the reference P , which is very similar to the interpretation of the RMSE as the error when using the ensemble mean value to approximate the reference value. In this way, the KL divergence can be considered as our measure of error in the discrete case in terms of the average amount of lost information when using the probability distribution calculated from the ensemble to approximate the distribution directly given by our synthetic reference categorical field.

As long as the KL divergence is expressed in the same units as the Shannon entropy, i.e. by using the same base of logarithm, it will hence be possible to compare them using a similar interpretation as with the RMSE and the ensemble spread in the continuous case. For the analysis of our results, we will use the logarithm to base N , with N the number of facies, so that the Shannon entropy will be at most equal to 1.

Besides, note that in the results presented, we will compute an average of the forementioned measures based on the gridblocks located in the most informed region of the model, i.e. up to 1600 m from the seepage boundary, so as to consider almost twice the distance between the two series of observation points at $x = 0$ m and $x = 1000$ m.

5.3 Results and discussion

5.3.1 Truncated Gaussian case

As shown in Figure 5.7, the conditioning of the categorical realizations via the update of the underlying variables with ES-MDA led to an efficient reduction of the data mismatch. The oscillations observed for each ensemble member are indicative of the increased nonlinearity of the forward model operator due to the additional truncation step before the flow simulation. However, as indicated by the progressively decreasing median curve, the convergence of the ensemble is overall robust. Given the large initial variability of the ensemble of predicted data, the data match obtained after 8 iterations of ES-MDA can be considered reasonable at almost every observation point when assimilating both head and flowrate observations (Figure 5.8).

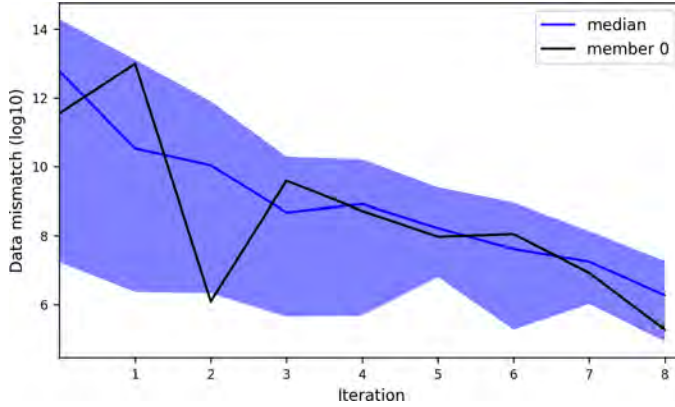


Figure 5.7: Evolution with the iterations of the decimal logarithm of the data mismatch for the truncated Gaussian case when assimilating both head and flowrate data. The blue curve is the median, the blue area corresponds to the region between the percentiles P5 and P95 of the ensemble of values, and the black curve is the values computed for the ensemble member no. 0.

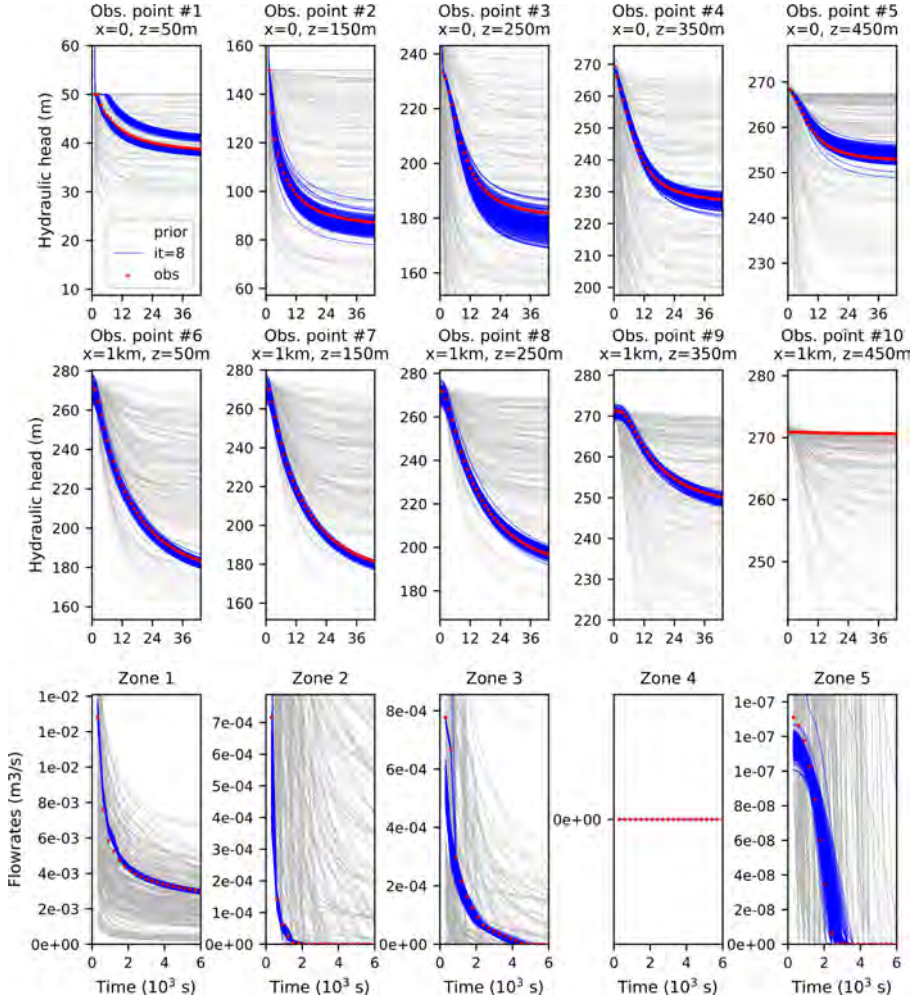


Figure 5.8: Predicted hydraulic head and flowrate data at the observation locations, before (in grey) and after (in blue) the assimilation of both hydraulic head and flowrate data in the truncated Gaussian case. The red dots are the observed data; the ensemble size is 200.

From Figure 5.9, we can observe that the conditioned facies fields are geologically consistent with the initial ensemble of facies fields generated. Both the structures and the proportions have been preserved thanks to the regeneration of the fields from the updated multi-Gaussian variables. It is worth noting that in the NS-EnKF approach of Zhou et al. (2011), the back normal-score transform of the updated transformed variables actually plays the same role as this facies regeneration from the trun-

cation step since it allows to get back to the property distribution and allows the ensemble integration forward in time with the flow simulation. However, the multi-Gaussianity of the variables being updated in this case are a determining factor in the consistency of the regenerated structures. Indeed, we remind from the previous bimodal case that the non-Gaussian joint distribution of the transformed variables led to a loss of the initial structures right after the first update step.

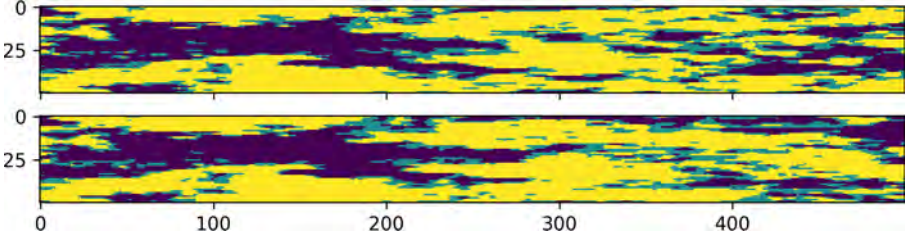


Figure 5.9: Examples of two truncated Gaussian realizations obtained after updating their respective underlying multi-Gaussian realization using ES-MDA($\times 8$)

However, Figure 5.9 also suggests that there is very little variability between the final facies fields in the updated region of the model. To quantify this reduction of variability, we can calculate the average ensemble spread of the underlying multi-Gaussian variables, similarly to Liu and Oliver (2005). In the region up to 1600 m from the seepage boundary, the average spread of the underlying multi-Gaussian ensemble has indeed been significantly reduced from 0.93 to 0.15, hence by a factor 6 which is significantly larger than the factor of around 2 observed in the less nonlinear multi-Gaussian log K case.

As in the previous multi-Gaussian log K and bimodal cases, we tested the effect of a normal-score on the data, following the idea proposed by Zhou et al. (2011) of applying the update on locally Gaussian ensemble distributions of predicted data. On the basis of the performance criteria proposed in Section 5.2.4, we could address properly the effect of such transformation in the present discrete case. In Figure 5.10, we can distinctly observe that the error relatively to the reference probability distribution, represented here by the KL divergence, decreases significantly when no transformation is applied on the data. This observed KL divergence reduction indicates that the yielded conditioned facies fields are closer to the reference distribution and hence are more accurate. Besides, given the discrepancy between the KL divergence

and the Shannon entropy which here represents the uncertainty of the categorical ensemble, we can also conclude that the uncertainty is less underestimated without the use of transformation.

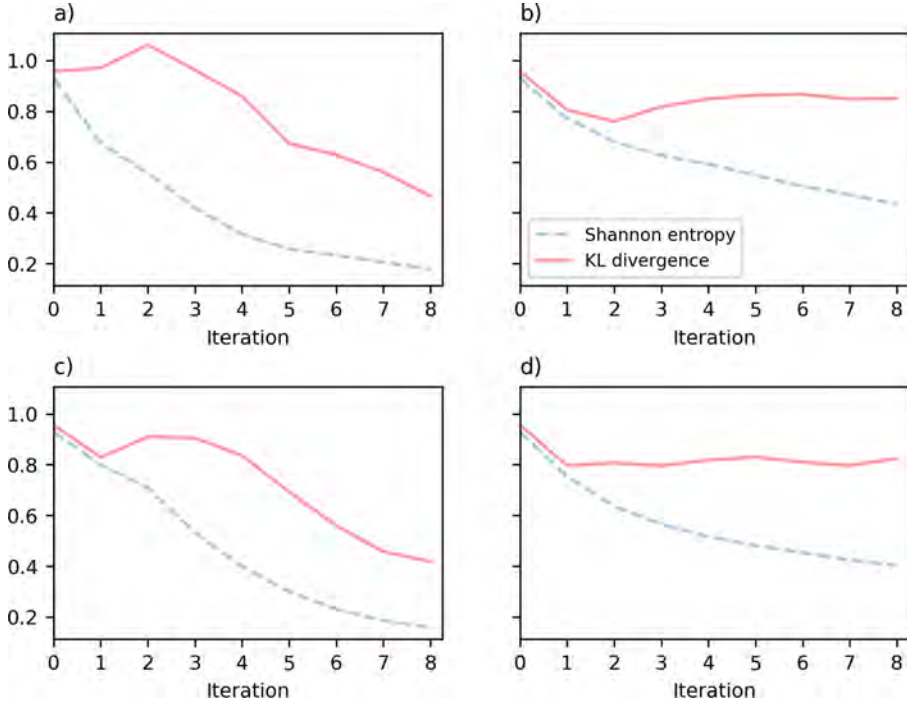


Figure 5.10: Evolution in the truncated Gaussian case of the Kullback-Leibler divergence and Shannon entropy (in the region up to 1600 m from the seepage boundary) during the assimilation with ES-MDA of: a) Head data. b) Transformed head data. c) Head and flowrate data. d) Transformed head and flowrate data.

Through this discrete case example, we showed the importance of considering other performance criteria besides the reduction of the data mismatch as the only indicator of a successful application of ensemble Kalman methods. Indeed, ES-MDA also converged when using transformed data in this truncated Gaussian case. Despite the small variability of the ensemble of conditioned facies fields obtained after the assimilation, those fields are actually very close to the reference (Figure 5.2) as a result of the efficient assimilation of the head and flowrate data. Indeed, the uncertainty of the categorical ensemble distribution was not

so overly underestimated compared to the true error as we may have expected.

Therefore, we can overall consider the application of ES-MDA in this discrete case efficient for the conditioning of truncated Gaussian fields via their underlying multi-Gaussian field. To improve the remaining data mismatch observed at certain observation points, a possibility would be to increase the current ensemble size of 200 in order to compensate for the uncertainty underestimation associated to the increased nonlinearity of the truncation step. In so doing, we might increase the insufficient spread of the underlying multi-Gaussian field and hence allow a better fit of the observations. In comparison to the also nonlinear but non-multi-Gaussian bimodal case, we reckon that thanks to the multi-Gaussianity of the underlying variables, increasing the ensemble size in the order of $O(10^2)$ still will be sufficient to benefit from the assimilation as we have observed in the multi-Gaussian log K case.

5.3.2 Bi-Gaussian case

Similarly to what we observed in the truncated Gaussian case, the conditioned bi-Gaussian fields obtained throughout the data assimilation with ES-MDA are consistent (Figure 5.11) thanks to the facies regeneration based on updated underlying multi-Gaussian variables.

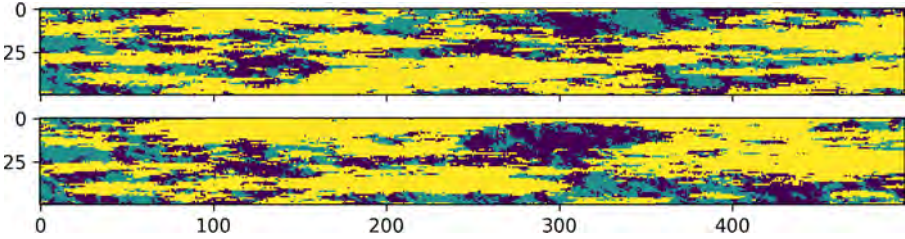


Figure 5.11: Example of two bi-Gaussian fields obtained after updating the underlying multi-Gaussian fields with ES-MDA($\times 8$)

Unlike the truncated Gaussian case, however, the ensemble distribution of the facies fields departs from the reference distribution during the assimilation. As shown in Figure 5.12, the KL divergence becomes larger and as a result, further increases the discrepancy with the Shannon entropy. It is worth noting that the uncertainty of the ensemble has less decreased compared to the truncated Gaussian case, and so have the

spread of the ensemble of underlying multi-Gaussian fields. If we look at the resulting data match obtained at the observation locations in Figure 5.13, it is therefore not surprising to observe a still large variability in the ensemble of predictions at several observation locations.

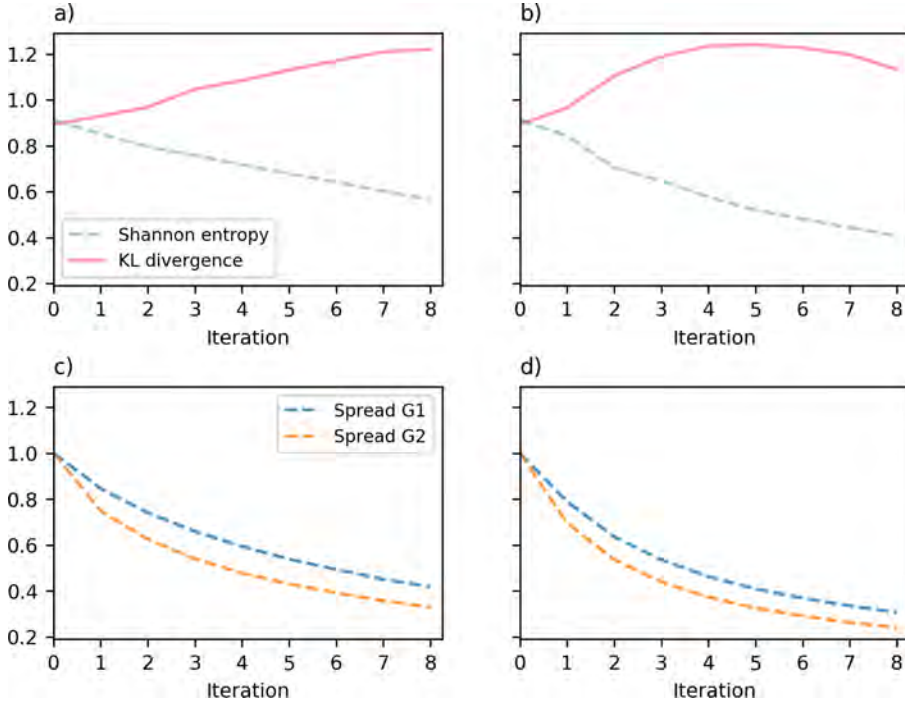


Figure 5.12: Evolution in the bi-Gaussian case of: the Kullback-Leibler divergence and Shannon entropy (in the region up to 1600 m from the seepage boundary) during the assimilation with ES-MDA of head data (a) or both head and flowrate data (b); and respectively for case a) and b) the ensemble spread of the underlying multi-Gaussian fields (G1 and G2) in c) and d).

What is of higher concern, however, is that even when the number of ES-MDA iterations is increased to 16, the data match does not improve sufficiently and even gets worse at other locations (Figure 5.14). As pointed out by Liu and Oliver (2005), the truncation step with the assignment of different facies types and hence distinct hydraulic property values makes the relationship between the multi-Gaussian variables $Y_1(x)$ and $Y_2(x)$ and the predicted flow data highly nonlinear in the bi-Gaussian case. Compared to the truncated Gaussian case, we assume

that this nonlinearity is significantly increased here given the more complex truncation rules. It therefore seems that the insufficient data match observed after the assimilation is partly related to the nonlinearity of the forward model operator. Note that, the normal-score of the data does not allow ES-MDA to converge in this case, thereby confirming the limitation of the transformation already observed in the truncated Gaussian case as a result of the nonlinear forward model operator.

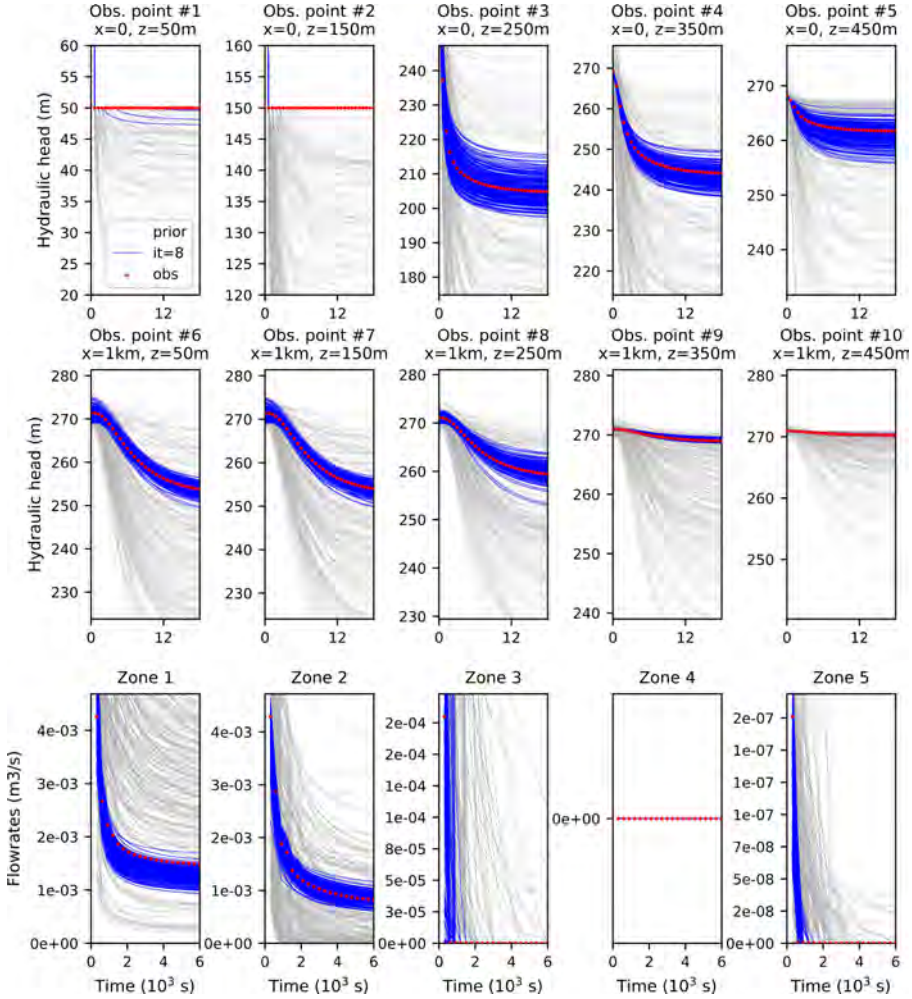


Figure 5.13: Predicted hydraulic head and flowrate data at the observation locations, before (in grey) and after (in blue) the assimilation of both hydraulic head and flowrate data in the bi-Gaussian case. The red dots are the observed data; the ensemble size is 200.

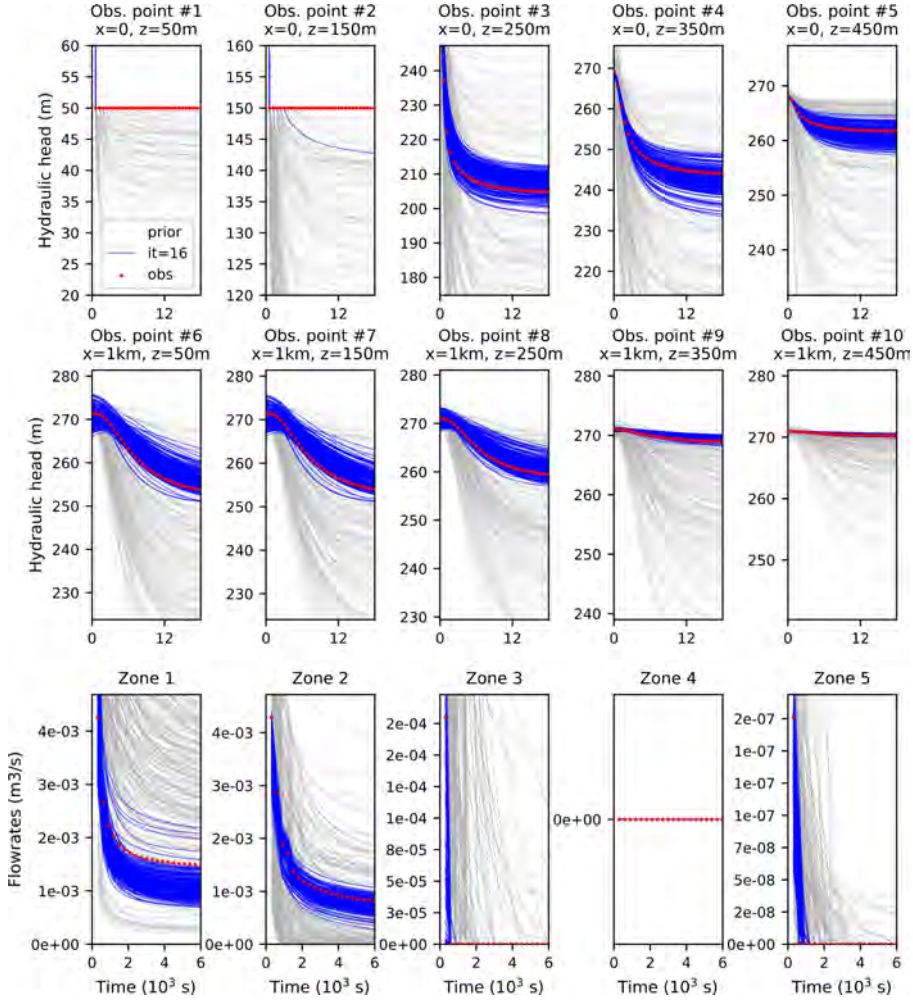


Figure 5.14: Predicted hydraulic head and flowrate data at the observation locations, before (in grey) and after (in blue) the assimilation with 16 ES-MDA iterations of both hydraulic head and flowrate data in the bi-Gaussian case. The red dots are the observed data; the ensemble size is 200.

We remind that after each ES-MDA update of every couple of underlying multi-Gaussian variables $Y_1(x)$ and $Y_2(x)$ in the truncation map, the thresholds need to be recalculated in order to respect the known proportions of the reference. However, by moving these thresholds after each update, we conjecture that it will also be more complicated for the smoother to properly condition the underlying variables so as to predict

the correct facies type and hence hydraulic property value at one location based on the covariances only. In order to test this hypothesis, we purposely considered the same thresholds as in the reference for all ensemble members, and kept them constant throughout the assimilation. As shown in Figure 5.15, the data match seems to be more efficient when considering constant thresholds. Even if this improvement could also be attributed to the proportions which are now allowed to vary according to the assimilation of the flow data, we could also argue that removing those constraints on the proportions can make the estimation of the true facies distribution more challenging and hence the history match as well. In fact, the KL divergence still increases even with this modified definition of the truncation. It is worth noting that a control on the proportions could still be possible with this defined truncation, namely by including them as conditioning data in the assimilation procedure.

It is also worth noting that although the two generated multi-Gaussian realizations can be independent initially, the estimation of the ensemble method will likely assume correlations between the couple of variables from respectively each realization at one location. Indeed, the truncation step defining the facies type at one location will depend on both the estimated variables from each underlying field at that location. Hence one possible explanation as to why the uncertainty is not so reduced compared to the truncated Gaussian case, especially given the increased number of parameters, could be the result of these additional correlations in the covariance matrices. To decrease the impact of these noisy correlations (which may be the reason for increased errors), one may consider reducing the number of correlated variables to be estimated in both underlying multi-Gaussian fields. The accuracy of the covariance matrices computed during the update, using the same ensemble size, may then increase.

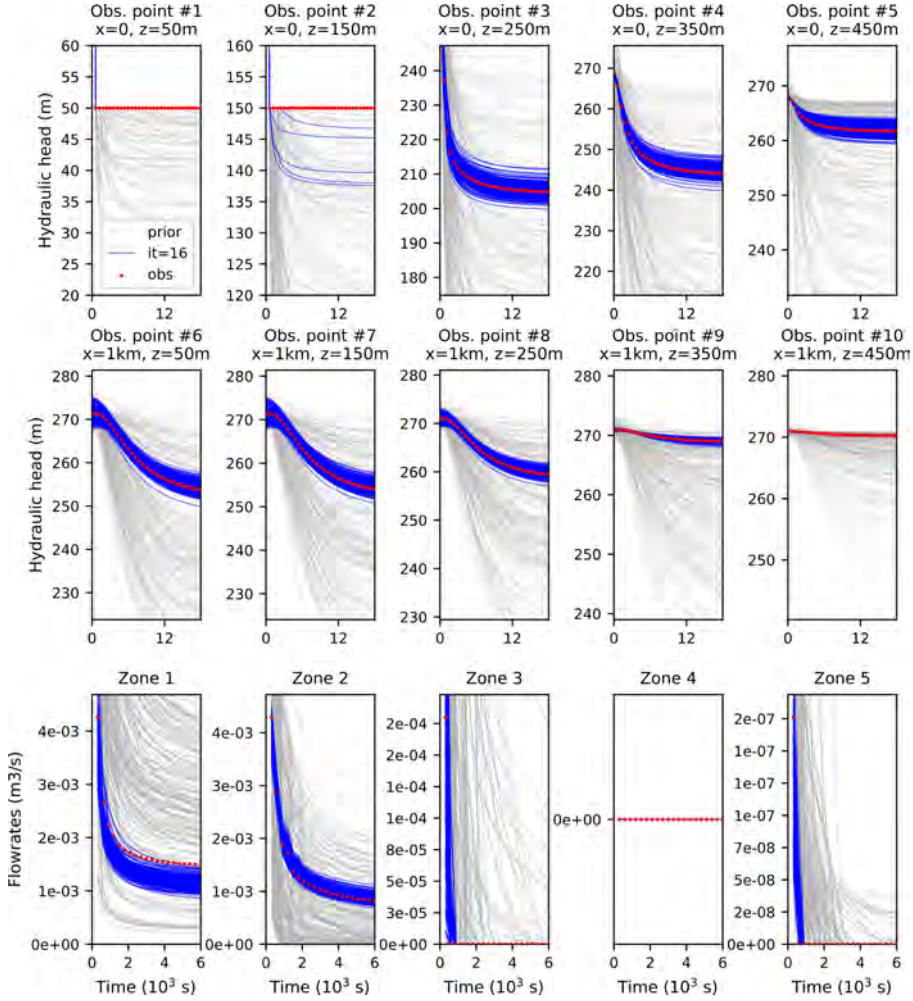


Figure 5.15: Predicted hydraulic head and flowrate data at the observation locations, before (in grey) and after (in blue) the assimilation with 16 ES-MDA iterations of both hydraulic head and flowrate data in the bi-Gaussian case. The same constant thresholds were applied to all ensemble members. The red dots are the observed data; the ensemble size is 200.

5.4 Conclusion

The truncated Gaussian and bi-Gaussian examples illustrate the benefit of an appropriate parameterization as investigated in previous

studies (Liu and Oliver, 2005; Astrakova and Oliver, 2015; Zhao et al., 2008). Compared to the iterative normal-score ensemble smoother approach proposed in the previous chapter, the update of the continuous underlying variables which define the facies distribution offers several advantages. The first and most important one is the Gaussian joint distribution of the underlying variables. As seen in both the truncated Gaussian and bi-Gaussian cases, the conditioned facies fields are in this case consistent with the initial fields prior to the assimilation thanks to the consistent updated underlying fields. Indeed, similarly to the multi-Gaussian log K case, the linear update of the ensemble Kalman method is also applied on a multi-Gaussian distribution. In the end, despite the increased nonlinearity due to the additional truncation step in the definition of the forward model operator, the application of the ensemble Kalman method could result in an efficient assimilation of both the head and flowrate data for the synthetic truncated Gaussian case. The second advantage is actually a consequence of the multi-Gaussian prior ensemble, namely the uncertainty is not underestimated as much in the truncated Gaussian case as observed in the bimodal case for the same order of ensemble size. Hence the need to compensate the underestimation of the uncertainty with a significantly larger ensemble size is also mitigated in this case.

However, the bi-Gaussian example (which uses a stronger nonlinear truncation) shows that not only the parameterization is important for the application of the ensemble method, but also the definition of the forward model operator. Indeed, in the multi-Gaussian log K case, we have seen that applying a normal-score on the data benefited the performance of ES-MDA as a result of a linearization of the relationship between the log hydraulic conductivities and the predicted flow data. In this bi-Gaussian case, such transform can no longer help linearize the relationship as a result of the already too nonlinear truncation step. As it turns out, this transform prevents the ensemble method to converge unlike in the truncated Gaussian case. We found, however, that the recalculation of the thresholds in order to honor the true proportions after each update may affect the convergence of the assimilation. Thus, with this same idea of making the forward model operator less nonlinear and hence more tractable for the computation of the update, we suggested to define differently the truncation step by keeping the thresholds constant and identical for all ensemble members throughout the assimilation. We observed that the resulting modified forward model operator effectively allowed a better fit of the observations. As for the relatively high un-

certainty of the ensemble remaining after the assimilation, we assume that it is related to the introduction of noisy correlations between the couples of variables from each underlying field during the assimilation and the increase of the degree of freedom. To decrease the error, the ensemble size could be increased, but at the expense of computational efficiency. An interesting alternative could be to decrease the number of variables to be estimated from both underlying fields, by using a pilot point parameterization for example, thereby reducing the correlations and improving the accuracy of the estimation.

Chapter 6

Multi-resolution approach to condition categorical multiple-point realizations to dynamic data with iterative ensemble smoothing

6.1 Introduction

The characterization of rock properties is necessary to make reliable numerical predictions of flow and mass transport in the subsurface. Given the sparsity of measurements, geostatistical simulation techniques can be used to obtain multiple realizations of the value of one or more properties over the whole discretized domain while honoring the available measurements. The ensemble of realizations can then be used to quantify the uncertainty of the property, assuming that the variability between the realizations is sufficiently well captured by the ensemble. When these realizations are conditioned to static data only, many realizations often do not reproduce the dynamic flow data. At this stage of the property characterization, the motivation for applying stochastic inverse approaches is to use the available state observations to condition the generated ensemble of geostatistical realizations and thereby reduce the uncertainty. However, as the current trend in hydrogeology is to increase the complexity of the conceptual models with the development

of advanced geostatistical techniques, the parameter identification problem using stochastic inverse methods has in parallel also become more challenging.

As underlined by Zhou et al. (2014) for initial non-Gaussian realizations, one drawback of a number of existing inverse approaches is that the reproduced states after the inversion, although in agreement with the observations, are obtained with a loss of consistency: the final realizations depart far too much from the prior geological model. The ensemble Kalman filter (EnKF) introduced by Evensen (1994) and its variants are data assimilation methods which have drawn particular attention because of their computational efficiency in high-dimensional and nonlinear state and/or parameter estimation problems (Evensen, 2009b). However, due to their strong assumption of linearity between the observations and the variables to estimate, these types of algorithms fail to preserve the non-Gaussian structures if they are applied directly to condition a non-multi-Gaussian distribution.

Thanks to an appropriate parameterization, Liu and Oliver (2005) showed that ensemble Kalman methods could be successfully applied even in the more challenging case of categorical simulations generated by truncated Gaussian-based methods. By updating the underlying multi-Gaussian variables instead of the non-Gaussian categorical ones directly, the multi-Gaussian assumption on the variables to update could still be fulfilled, allowing in this way an acceptable performance of the method both in terms of the consistency of the structures characterized and the quality of the data fit achieved.

Among approaches proposed to condition multiple-point statistics (MPS) simulations to dynamic data with ensemble-based Kalman methods, a probability conditioning method was introduced by Jafarpour and Khodabakhshi (2011) in order to derive at each update of EnKF a facies probability map from the ensemble of updated log hydraulic conductivity values which initially populated categorical MPS realizations of facies. Originally, Jafarpour and Khodabakhshi (2011) proposed to map the mean value computed at each gridblock from the ensemble of updated log hydraulic conductivity fields to a probability value by using a linear transformation. Recently, Ma and Jafarpour (2019) proposed an improved approach to derive these probability values by also taking into account the ensemble variance. Either way, the resulting facies probability map was then used as soft conditioning data for the next conditional multiple-point simulation using the MPS algorithm SNESIM

(Strebelle, 2002). An interesting outcome of the proposed methodology was that each regeneration of soft-data-conditioned multiple-point simulations following an iteration of EnKF allowed the introduction of some variability in the facies distribution due to the stochastic nature of the MPS method.

The probability conditioning method (Jafarpour and Khodabakhshi, 2011), however, does not attempt to mitigate the effects of the non-multi-Gaussian distribution of log hydraulic conductivities before each update of EnKF. Zhou et al. (2011) proposed a normal-score transform prior to each computation of the EnKF update so that the ensemble of initial or updated parameter variables is marginally Gaussian in order to reduce the consequences of violating the multi-Gaussianity assumption from the beginning. For their synthetic case, they concluded that such normal-score transform prior to the update allowed to outperform EnKF without any transformation in terms of the structures that could be retrieved by the end of the data assimilation. It is worth mentioning that their application of the ensemble method was based on an ensemble of 1000 realizations, which is quite large compared to the standard ensemble sizes of around a hundred of members, and also on an important number of observation points for the hydraulic head data used in the assimilation.

Similarly to the strategy of Liu and Oliver (2005) to condition categorical truncated Gaussian simulations with EnKF via their underlying multi-Gaussian variables, Hu et al. (2013) proposed to apply EnKF to update either the uniform fields underlying the ensemble of categorical MPS realizations generated with SNESIM (Strebelle, 2002), or the coefficients of combination of those underlying uniform fields. Both proposed parameterizations of the MPS simulations have the advantage of allowing the conditioned categorical fields to remain consistent with the prior geological model. Although both allow to deal with continuous instead of discrete parameters, the resulting increased nonlinearity between the parameters to update and the dynamic response ultimately prevented EnKF from reaching a sufficient data match in either case (Hu et al., 2013). Heidari et al. (2013) had also observed a similar lack of flexibility in terms of the final data match achieved when updating gradual deformation parameters with EnKF to perturb an ensemble of multi-Gaussian realizations of a petrophysical property.

Keeping this idea of a parameterization reducing the number of adjustable parameters, Gardet et al. (2014) proposed a multiscale param-

eterization on the basis of a multi-Gaussian simulation method. The goal was to allow to adapt the scale of the conditioning of parameters according to the data used. Knowing that the finer realizations would be conditioned to the realizations simulated at the coarse scale, an interesting perspective was the possibility of applying at this coarser scale the update from an inverse method. The potential of such strategy was illustrated with the gradual deformation technique.

Mainly motivated by the results of the forementioned studies, this paper introduces a new methodology to condition categorical MPS simulations of facies to dynamic data with an ensemble Kalman method. The workflow relies on a multi-resolution parameterization based on generated multi-resolution training images which allow MPS simulations at coarser resolutions than that of the original training image. The performance of the proposed approach is evaluated using a synthetic 2D vertical groundwater model and transient hydraulic data at particularly sparse locations in the horizontal direction, as would be the case in many real problems. In this work, the code DeeSse (Straubhaar, 2017) which offers a multiscale parameterization of the direct sampling algorithm (Straubhaar et al., 2020) was used. Among existing ensemble Kalman methods, the ES-MDA algorithm (Emerick and Reynolds, 2012a) was chosen for its simple formulation and demonstrated efficiency in a real reservoir history-matching problem (Emerick, 2016).

In the next two sections, we first introduce the concepts that are at the heart of the proposed methodology. Section 6.2 aims to give some background on how the multi-resolution MPS simulations are performed in DeeSse. The update algorithm ES-MDA is then presented in Section 6.3. We finally give a walkthrough of the proposed methodology in Section 6.4 based on the previously introduced concepts. The synthetic case including the model set up, the definition of the DeeSse parameters, and the performance criteria is presented in Section 6.5. Finally, the results of the synthetic case are presented and discussed in Section 6.6.

6.2 Multi-resolution multiple-point geostatistical simulations with DeeSse

Since a few decades, the interest for multiple-point statistics (MPS) simulations has grown considerably with applications in various fields of geosciences (Mariethoz and Caers, 2015). Given a “training image” (TI),

an image provided by the user as the conceptual model of the variable to simulate, the goal of MPS is to generate random fields that reproduce the spatial statistics of the TI. Depending on the TI, MPS can render a large variety of geological heterogeneities, including complex connectivity patterns. Among existing MPS algorithms, the direct sampling method (DS) (Mariethoz et al., 2010) consists in simulating the nodes of a simulation grid sequentially by directly sampling the TI. For each node to be simulated, the pattern formed by a number of neighboring nodes, either informed by data or which have been previously simulated, is searched in the TI. As soon as the distance measuring the mismatch between the conditioning pattern and the one being currently scanned in the TI is below a given threshold, the corresponding TI value is pasted at the simulation grid location. If no compatible pattern is found after a certain fraction of the TI has been scanned, the search is stopped and the best candidate found so far is used. Hence the maximal number of neighbors, the distance threshold, and the maximal scanned fraction are the key parameters of an MPS simulation with DS. Besides being computationally efficient, DS offers two advantages of general interest: (1) the ability of handling categorical and continuous variables, and (2) the possibility of performing simulations of multiple properties given the multiple-point dependence between given properties (Mariethoz et al., 2012).

DeeSse (Straubhaar, 2017) is an implementation of the DS algorithm with some features allowing to improve the quality of the MPS simulations. Among the recently added features, the “Gaussian pyramid” technique allows to control the simulation at multiple resolutions. The motivation behind this feature was originally to better capture the different scales of heterogeneities that might be present in the TI (Straubhaar et al., 2020). The technique consists in computing coarser versions of a given TI. The resulting set of multi-resolution training images can then be used to perform a series of MPS simulations at different scales, starting from the coarsest scale up to the original scale of the training images. Those multi-resolution MPS simulations establish the multi-resolution framework of our proposed data assimilation approach to condition categorical MPS simulations with an ensemble Kalman method presented in Section 6.4. We provide here some background on how such multiscale MPS simulations are performed by DeeSse in the case of a categorical TI.

For a given categorical TI with N categories, the Gaussian pyramid technique creates a pyramid of multi-resolution training images for

$N - 1$ indicator functions of the different categories. The remaining category is deduced from all the previously simulated ones. Starting from the provided TI, the upscaled images result from the recursive application of a reduce operator on the last image created. As currently implemented in DeeSse, this reduce operator corresponds to the convolution product of the image with a set of Gaussian-like kernels of either 5-pixel width distributed every two nodes in each direction, or 7-pixel width distributed every three nodes, or a combination of both of these possibilities. Consequently, the spacing of two or three nodes between each kernel will correspond to the reduction factor applied in each direction to the image. By iteratively applying this reduce operator to the last reduced image, one thereby obtains a so-called “Gaussian pyramid” of images. In addition, each reduced image is expanded onto the grid of the next finer scale by assigning first the values in the grid nodes corresponding to the center nodes of the kernel that was used for the reduce operator (one node over two or three depending on the kernel width), and then by computing a moving average, in every node, by using the same kernel with weights normalized over the informed nodes (and ignoring uninformed locations) (Burt and Adelson, 1983). As will be described further below, each pair of expanded and reduced training images of the same level will allow to perform multivariate MPS simulations (Mariethoz et al., 2012) which are guided by the expansion of the simulation result in the previous coarser scale grid.

For example, Figure 6.1 shows for a binary TI the complete set of multi-resolution training images which will be used at some stage of a series of multiscale MPS simulations performed by DeeSse. As illustrated, this set includes the Gaussian pyramid of training images obtained for the indicator function of the category of value 1, i.e. the channel facies, with two levels of coarse resolutions with a reduction factor of two applied in each direction between each level, as well as the pyramid formed by the two expanded images resulting from the respective expansion of the two reduced images. Hence given both these pyramids of reduced and expanded training images, an unconditional univariate MPS simulation can first be performed at the coarsest level of the TI pyramid by using the corresponding TI. Then, iteratively, the last coarse simulation is expanded to the next finer scale in order to guide, by way of a bivariate MPS simulation, the MPS simulation at that new fine scale using the bivariate TI composed of both the reduced and expanded images of the corresponding level. Based on the multi-resolution training images shown in Figure 6.1, Figure 6.2 shows an

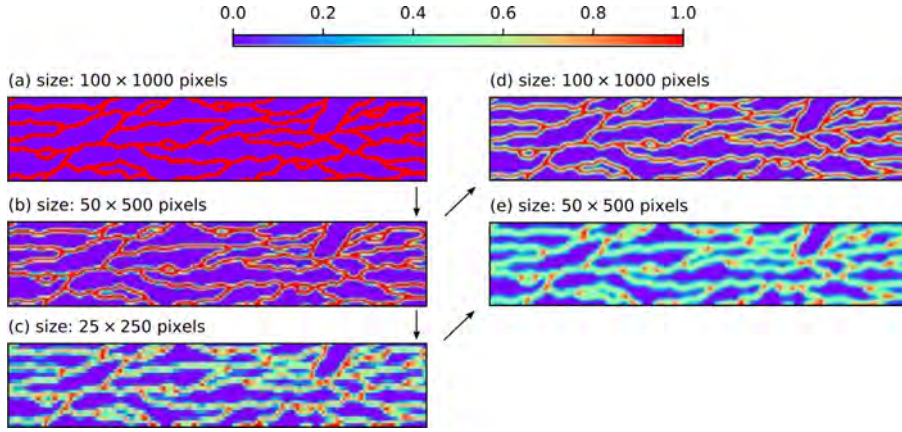


Figure 6.1: Example of a set of multi-resolution training images with two additional coarse levels obtained with the Gaussian pyramid technique. A reduction factor of two between each level was applied in each direction. Given a categorical training image (a), a first reduced image (b) is obtained. A second reduced image (c) is then obtained from (b). Finally, the images (e) and (d) are created based on the expansion of (c) and (b) respectively.

example of pyramids of multiscale MPS simulations obtained, as well as the expanded images resulting from the coarser-scale simulations which were used. It is also worth noting that the additional MPS simulations at the coarse scales required with this pyramid option are not necessarily much more time consuming, considering that the number of pixels is reduced from one level of the pyramid to the next coarser one, and also as they are guided by the expanded image from the coarser level in the next finer scale MPS simulation (Straubhaar et al., 2020).

6.3 Ensemble data assimilation with ES-MDA

The purpose of this section is to provide some background on the ensemble Kalman method which was used for the proposed methodology although another variant could be considered. The data assimilation algorithm will allow to update the previously introduced multi-resolution MPS simulations and ultimately to condition the categorical facies fields to the observed dynamic data.

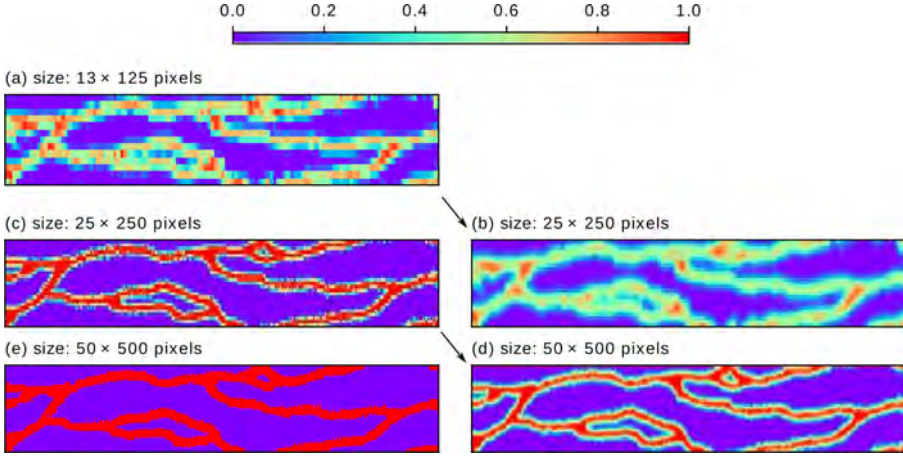


Figure 6.2: Illustration of the multi-resolution MPS simulation procedure in DeeSse using the training data set of Figure 6.1. At the coarsest resolution, a first univariate MPS simulation is generated (a). The simulation is then expanded to the next finer scale of the TI pyramid (b) to guide the next MPS simulation at the finer scale (c) using the corresponding bivariate TI. Again, the last simulation is expanded to the next finer scale (d) and used to guide the categorical MPS simulation (e) using the corresponding bivariate TI.

6.3.1 General background

The ensemble smoother with multiple data assimilation (ES-MDA) developed by Emerick and Reynolds (2012a) is currently one of the most used ensemble-based Kalman algorithms for parameter estimation in reservoir applications (Evensen, 2018; Abadpour et al., 2018). It can be viewed as an iterative form of the ensemble smoother method (ES) which aims to improve the insufficient match between the observed and simulated data that usually remains after a single-iteration ES update in nonlinear problems. Indeed, when the whole data set is assimilated in one single step with ES, the update step which should remain linear in ensemble Kalman methods tend to be nonlinear, resulting in the computation of a suboptimal update. The purpose of iterative ensemble smoother algorithms in general is to iteratively handle the non-linearities between the parameters and the assimilated data, so that ultimately a better fit of the data can be achieved.

ES-MDA was motivated by the similarity between ES and one single

iteration of the Gauss-Newton method when it is applied to an ensemble of vectors of model parameters m_i with a full step and an average sensitivity estimated from the ensemble (Emerick and Reynolds, 2012a). For each vector of model parameters m_i , this Gauss-Newton method minimizes the following objective function

$$O(m_i) = \frac{1}{2}(g(m_i) - d_{obs,i})^T C_{err}^{-1}(g(m_i) - d_{obs,i}) + \frac{1}{2}(m_i - m_{pr,i})^T C_M^{-1}(m_i - m_{pr,i}) \quad (6.1)$$

where $m_{pr,i}$ is a vector of model parameters sampled from the multivariate Gaussian distribution with mean m_{pr} and covariance C_M , $d_{obs,i}$ is a vector of perturbed observations sampled from $N(d_{obs}, C_{err})$, and $g(m)$ is the vector of data predicted from the nonlinear forward model $g()$. From this analogy, the ES-MDA algorithm was developed so that it would be comparable to several Gauss-Newton iterations with an average sensitivity matrix updated from the ensemble after each iteration. For nonlinear problems, these multiple iterations especially allow to compute smaller and more accurate updates than one single potentially large and less accurate Gaussian-Newton iteration, and hence equivalent ES update. This is why a more efficient reduction is expected throughout the iterations of ES-MDA than with ES in a nonlinear case, and also why it is expected that each ensemble member updated with ES-MDA will minimize an objective function that is similar to the one minimized by the iterative Gaussian-Newton optimization procedure (Evensen, 2018).

As suggested by the name, the ES-MDA algorithm consists in repeating the ES update equation, for each parameter realization m_i where $i = 1, 2, \dots, N$ with N the ensemble size, a predefined number of times k

$$m_i^{k+1} = m_i^k + C_{MD}^k (C_{DD}^k + \alpha_{k+1} C_{err})^{-1} (d_{obs,i}^k - g(m_i^k)) \quad (6.2)$$

where C_{MD}^k , the cross-covariance matrix between the parameter vector m_i and the vector of predicted data $g(m_i)$, and C_{DD}^k , the auto-covariance matrix of predicted data, are approximated from the ensemble at iteration k .

The only difference compared to ES is that in ES-MDA (6.2), the covariance matrix of the measurement errors C_{err} is inflated by a factor $\alpha > 1$. Hence at each new iteration $k + 1$, each updated realization

m_i^{k+1} is computed as the linear combination of the updated realization m_i^k of the previous iteration k and each component of the vector of residuals between $d_{obs,i}^k$, the vector of observed data to assimilate, and $g(m_i^k)$, the vector of corresponding predicted data from the nonlinear forward model $g()$. For each component of the ensemble of parameter vector realizations m_i^k , each residual is then weighted on the basis of the Kalman gain matrix K computed from the ensemble of realizations m_i as

$$K = C_{MD}^k (C_{DD}^k + \alpha_{k+1} C_{err})^{-1} \quad (6.3)$$

Note that, for simplicity, it is often assumed that the covariance matrix of measurement errors C_{err} is diagonal, i.e. the measurements are uncorrelated, and the variance assigned for each measurement will only depend on the data type. It is worth reminding that the reason why the Kalman gain matrix is derived according to (6.3) is to minimize the trace of the covariance matrix of updated parameters m_i^{k+1} (Burgers et al., 1998). As a matter of fact in ES-MDA, the observed data used to compute the residuals in the update equation (6.2) are the actual measurements $d_{obs,i}$ which have been perturbed with a Gaussian noise sampled at each iteration from an inflated covariance of measurement errors $\alpha_{k+1} C_{err}$ as follows

$$d_{obs,i}^k = d_{obs} + \sqrt{\alpha_{k+1}} C_{err}^{1/2} z_{d,i} \quad (6.4)$$

where z_i is a vector with standard random normal deviates, i.e. $z_i \sim N(0, I)$. The purpose of inflating the covariance of measurement errors with a factor α_{k+1} is in fact to limit the confidence given to the data as they will be processed multiple times. By doing that, the parameter covariance reduction which occurs after each data assimilation will not be overestimated. Still, because of this variance-minimizing scheme underlying the computation of the Kalman gain-based update, it requires some amount of trial to determine the number of ES-MDA iterations which allows to both reach an acceptable goodness of fit, and not end up with an unnecessarily overreduced parameter uncertainty. In this sense, ES-MDA in its standard implementation is less flexible than methods which adapt the number of iterations in the course of the optimization. Recently, Evensen suggested that for a limited ensemble size of 100 realizations, applying the ES-MDA update for more than 16 iterations should not lead to improved results as the estimated mean would contain sampling errors that are larger than the uncertainty reduction (Evensen, 2018). Several modified ES-MDA algorithms have actually been proposed in order to adapt the inflation coefficients, and hence

the update step sizes, as the history match proceeds (Le et al., 2016). No matter the implementation used, all the factors α_k used to inflate the observations assimilated by ES-MDA should satisfy the following condition (Emerick and Reynolds, 2012a)

$$\sum_{k=1}^{N_a} \frac{1}{\alpha_k} = 1 \quad (6.5)$$

where N_a is the number of times we repeat the data assimilation. This condition was derived to ensure the equivalence between the single ES update and the iterative ES-MDA for a linear Gaussian case. To test the proposed methodology to condition categorical MPS realizations presented hereinafter in Section 6.4, we will use the standard implementation of ES-MDA (Emerick and Reynolds, 2012a) and set, for simplicity, constant inflation coefficients equal to the predefined number of assimilations. As mentioned by Emerick and Reynolds (2012a), varying these inflation factors in a decreased manner to reduce the magnitude of the initial large updates only leads to minor improvements.

6.3.2 Localization and normal-score transform

Ensemble Kalman methods are generally implemented along with ad hoc methods designed to mitigate their intrinsic deficiencies, thereby getting the best possible performance out of them. Hereinafter, the purpose of two mitigation methods are stated since they will both play a significant role in the proposed methodology to condition categorical multiple-point realizations presented in Section 6.4.

A major limitation of EnKF and other variants such as ES-MDA is that the finite size of the ensemble can introduce undersampling errors in the covariance matrices approximated from the ensemble (Houtekamer and Mitchell, 2006). In particular, “spurious” long-range correlations between components of the parameter vector and the predicted data can occur and bring an undesirable response in both the updated mean and covariance. At worst, they can reduce the parameter variance to the point where the covariance of the updated ensemble can collapse, thereby wrongly increasing the confidence in certain parameters and, as a result, preventing the successful assimilation of data that could contribute significantly to the parameter identification.

The tapering of the sample covariances, also referred as “covariance localization”, is a method widely used to filter out those undesirable

correlations. For our application of ES-MDA, we will apply a standard approach which consists in restricting the magnitude of the update computed close to the observation locations by using a tapering function. The Kalman gain matrix (6.3) will be attenuated by multiplying each one of its entries by a localization factor between 0 and 1. This localization is often defined as a function of the physical distance between each couple of parameter and data variables. An example of common forms used for this function is the fifth-order correlation function with compact support as defined by Gaspari and Cohn (1999). We will use this function since it has been used for many applications where localization in the horizontal direction is important (Anderson, 2012), as in our synthetic case presented in Section 6.5. Note that to evaluate this function, a “critical length” which corresponds to the distance beyond which the localization factor will rapidly decrease to zero needs to be defined initially.

Another important and more general limitation of ensemble Kalman methods is their underlying assumption of multivariate normal distributions. Because the Kalman gain is based only on the evaluation of covariance matrices, the update performed by all these types of methods will only be able to take into consideration the linear dependence between the variables. As a result, the statistics of order higher than two of an initial ensemble of realizations generated by MPS, for example, would not be preserved after the update, and neither the potential initial non-Gaussian structures.

To deal with non-multi-Gaussian ensembles, Zhou et al. (2011) proposed to perform the update in such a way that every marginal distribution is Gaussian. As illustrated in Zhou et al. (2011), for each parameter variable at every location, a normal-score transform function is defined such that the p -quantile of the ensemble distribution of original values is mapped to the p -quantile of a standard normal distribution. After the update is performed on these transformed parameter variables, they need to be back-transformed to be able to run the forward model and predict the corresponding new observations. Although this approach has shown to improve the estimates of the ensemble Kalman method if supported by enough spatially distributed data (Li et al., 2012; Zovi et al., 2017), it is important to note that this parameter transformation only mitigates, and hence does not completely solve the problem related to updating non-multi-Gaussian variables. Indeed, dealing with Gaussian marginal distributions does not necessarily imply a joint Gaussian distribution for all the variables. It is also worth reminding that the nor-

mal score transform approach could be applied until now, in the context of conditioning of MPS simulations with ensemble Kalman methods, because the parameters were values of a continuous property, such as log hydraulic conductivities, and not categorical values, such as for geological facies. For categorical variables, it would not be possible to define such normal-score transform. As the proposed methodology presented in the next section aims to condition categorical MPS simulations, the idea will be to update, in the same manner as Zhou et al. (2011), continuous parameters. As will be further explained, those continuous parameters will be defined at a coarser resolution than the resolution of the categorical MPS simulation. Moreover, those parameters will be more directly related to the categorical variables thanks to the multiscale framework for data assimilation established by the multi-resolution MPS simulations.

6.4 ES-MDA with multi-resolution MPS simulations

This section aims to describe how the multi-resolution MPS simulations with DeeSse and the ES-MDA algorithm are combined together to yield the proposed methodology.

6.4.1 Obtaining the latent variable

As previously introduced in Section 6.2, a multi-resolution MPS simulation provides a straightforward way to relate the categorical MPS simulation to a latent continuous variable at a coarse scale. The ES-MDA update can then be applied on the latent continuous variable instead of the categorical variable. Therefore, the first step of the proposed methodology consists in extracting the coarsest-resolution MPS simulation for each member of an ensemble of multi-resolution MPS simulations performed with DeeSse. It is worth mentioning that if hard conditioning data are provided for the categorical MPS simulation, these data would first be propagated through the different coarse levels of the pyramid before performing the series of MPS simulations (Straubhaar et al., 2020).

An immediate consequence of choosing the coarse image for the parameterization is a massive reduction of the number of parameters to

update with ES-MDA. The number of parameters will ultimately depend on the number of pyramid levels and the size of the Gaussian kernels applied in each direction for each level. A second consequence of the proposed parameterization comes from the process of the training image reduction itself. The coarse MPS simulation will be a coarse version of the fine categorical MPS simulation, even though the former is simulated first as described in Section 6.2. As shown in Figure 6.2, the coarse simulation will present smooth instead of sharp transitions between the facies. Considering that ensemble-based Kalman methods such as ES-MDA are not suitable for directly updating fields containing discrete features, it is expected that the blurring effect affecting the simulated structures at coarse resolutions will benefit the performance of the method.

6.4.2 Conditioning the categorical field

To assimilate the state variable observations, the computation of the ES-MDA update requires to compute the flow model response associated to the parameters which are going to be, or are already updated. For this purpose, the same hard conditioning procedure which is at the heart of the multi-resolution MPS simulation, as described in Section 6.4.5, can be employed to derive a categorical MPS simulation conditioned to the coarsest resolution parameters. The categorical field can then be conveniently used for the flow simulation, after assigning petrophysical property values to each category of facies. Once all the flow simulations of the ensemble are completed, the update of the coarsest-resolution variables can be computed based on the covariance matrices estimated from both the ensembles of coarse parameters and simulated flow data. An overview of the conditioning procedure is illustrated in Figure 6.3.

Note that, although they are not shown in Figure 6.3, two additional pre-processing steps are included in the final methodology proposed and are described in the next sections. They aim to mitigate the limitations which inevitably arise when the multi-Gaussian requirement on the initial ensemble is not met.

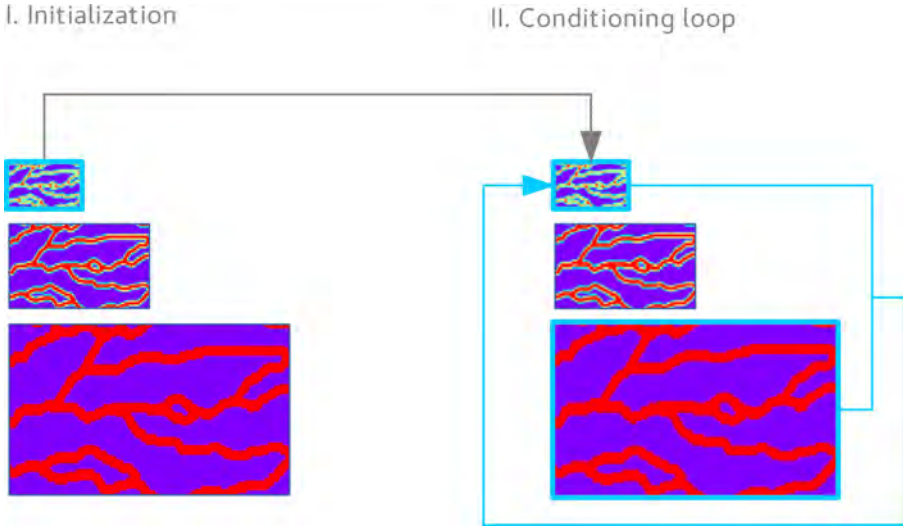


Figure 6.3: Overview of the procedure to condition categorical multiple-point simulations with ES-MDA using a multi-resolution parameterization. For each ensemble member, (I) a multi-resolution MPS simulation is generated and the image of the coarsest level is extracted. (II) The coarsest-resolution parameters are used to condition the categorical MPS simulation. Both the ensemble of coarsest-resolution parameters and categorical fields are involved in the computation of the update for the current iteration.

6.4.3 Normal-score transform

We compute a normal-score (NS) transform function for each pixel of the coarsest resolution fields of the initial ensemble. Indeed, Figure 6.4 shows that the distribution for the ensembles of coarsest-scale MPS simulations generated using different settings of pyramid parameters have marginal distributions that are not Gaussian. This figure also shows that the local distributions are getting closer to normal when the number of levels increases. By applying each NS transform locally, we ensure that every marginal distribution becomes Gaussian. The ES-MDA update will then be applied on the transformed variable rather than on the images coming directly from the pyramid of MPS simulations, as illustrated in Figure 6.5. Note that the computation of local normal-score transforms is performed only once before the whole iterative process with ES-MDA.

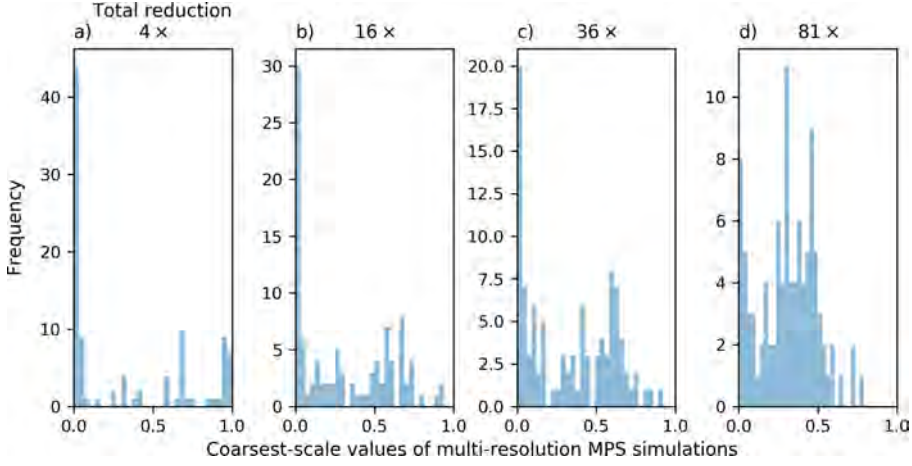


Figure 6.4: Distribution of values at a random location from an ensemble of coarsest-resolution MPS simulations for 4 different resolutions a, b, c, d. For all cases, the size of the original (finest) binary simulation is 50×500 pixels. a) 25×250 (reduced once with a factor 2 in each direction), b) 13×125 (reduced twice with a factor 2 in each direction), c) 9×84 (reduced first with a factor 2 in each direction, reduced again with a factor 3 in each direction), d) 6×56 (reduced twice with a factor 3 in each direction). The distribution becomes increasingly smoother as the resolution decreases.

As mentioned in the previous section, the conditioned categorical fields will be obtained from the coarsest resolution according to the same hard conditioning procedure allowing to generate a multi-resolution MPS simulation. Therefore, the updated transformed variables will need to be back-transformed in order for the values to be consistent with those of the coarsest-resolution TI, i.e. between 0 and 1. For this reason, we will apply on the ensemble, after each update, a back normal-score transform (BNS) function at every location. This function corresponds to the inverse function of the NS function previously used. When applying the BNS function, we also ensure that the back-transformed values will be bounded between 0 and 1 by not allowing extrapolation for values falling outside the domain of the function.

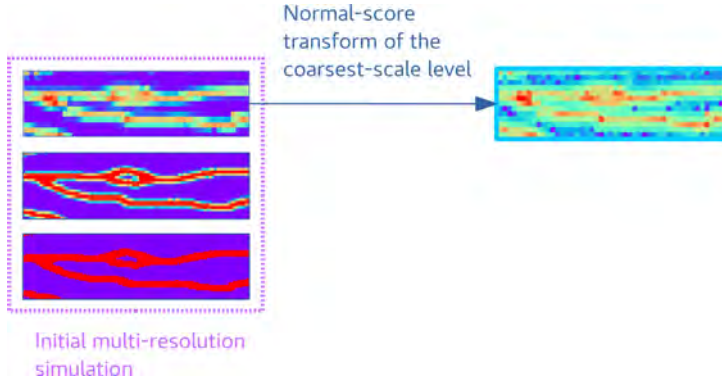


Figure 6.5: Illustration of the normal-score transform step performed for each member of an initial ensemble of multi-resolution MPS simulations before the conditioning.

6.4.4 Sampling and final parameterization of the multi-resolution MPS simulations

Given that ensemble Kalman methods are known to break the prior structures after an update, we do not fully inform the coarsest resolution of the multi-resolution MPS simulation with the complete back-transformed image as hard data. Instead, only a set of values at selected locations of the back-transformed image will be considered as hard conditioning points. Those locations should be as distributed as possible over the whole model domain so that DeeSse can still simulate in between them and in this way recover more consistent structures. The highest the number of non-adjacent locations, the better in order to efficiently constrain the data match. Ultimately, DeeSse will simulate a complete pyramid starting by a conditional simulation at the coarsest scale.

As can be observed in Figure 6.6, using conditioning points instead of a conditioning image will effectively prevent DeeSse to pass on the finer scale MPS simulations the degraded structures of the whole updated field. Therefore, the update of ES-MDA is integrated only locally at the coarsest resolution of the MPS simulation, and then indirectly to the subsequent finer ones via the expansion of the last MPS simulation and its use as a fully informed hard conditioning image for the following scale. Figure 6.6 illustrates the benefit of the sampling process and its integration at the coarsest scale of the multi-resolution MPS simulation.

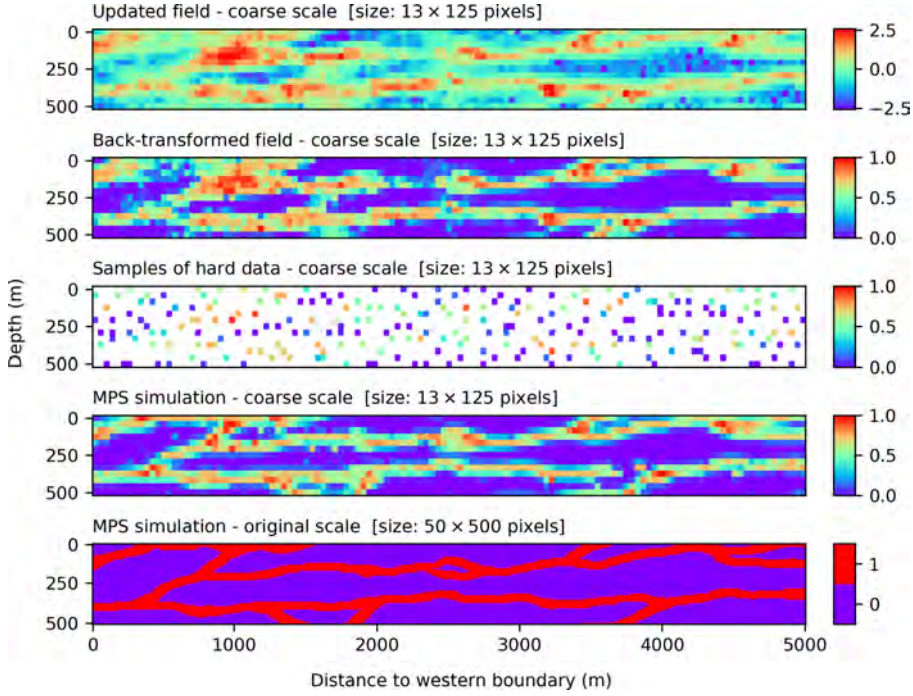


Figure 6.6: Conditioning a categorical MPS realization (bottom image) from the updated normal-scored image at the coarsest pyramid scale (ensemble member no. 11) at iteration 8 out of 16 of ES-MDA (top image) via the fixedly sampled back-transformed image used initially as hard conditioning points for the multi-resolution MPS simulation.

In addition, the sampled hard conditioning points over the whole domain offers a way to prevent expected discontinuous transitions at every scale of the MPS simulations when localizing the update of the parameters close to the observations. This is in fact illustrated in Figure 6.6 where only the region located up to around $x = 1600$ m from the western boundary is being updated due to the locations of the observations only in this region of the model. As can be observed, the structures simulated by MPS after each update, first at the coarsest and ultimately at the original scale, remain continuous at the transition between the updated and non-updated parameters at $x = 1600$ m.

The conditioning of the categorical field in the proposed methodology relies on a series of MPS simulations performed at every scale of a generated set of multi-resolution training images. This implies that a high degree of randomness will be introduced in the relationship between

the coarse parameters and the flow data calculated from the conditioned categorical field.

For the application of the ensemble Kalman method, it is critical to establish a forward operator that ensures (1) the derivation of a unique categorical field given a set of coarse parameters, and (2) a derivation which is obtained in a systematic manner for all members of the ensemble. Otherwise, the estimated covariance matrices will not be reliable for the computation of the updates. For both these above-mentioned reasons, an important first prerequisite is that the hard conditioning points should be sampled from the same locations for all ensemble members and for all iterations of ES-MDA. Then, the second and most decisive prerequisite is that the same seed number for the multi-resolution MPS simulations should be used for all iterations and for all ensemble members.

For one ensemble member and given the same locations of hard conditioning points for all iterations, a same seed value allows the reproducibility of the multi-resolution MPS simulation if the values at the hard conditioning points are the same. By extending for all ensemble members and for all iterations, both these conditions on the same locations of the hard conditioning points and the same seed value, we allow the ensemble of multi-resolution MPS simulations to be performed in a systematic manner. Indeed, provided that the locations of the hard conditioning points at the coarsest scale are the same for all members, using the same seed will make DeeSse simulate the ensemble of coarsest-scale MPS simulations by using the same randomly generated simulation path. A same path for all ensemble members will also be used for the subsequent ensemble of finer scale simulations. In this manner, the forward operator established will be sufficiently systematic for the ensemble Kalman method to capture and exploit the relationship between the parameters and the data, and this in spite of the intrinsic randomness of the parameterization with multi-resolution MPS simulations.

Moreover, the use of the same seed value for all the ensemble members means that the variability between the ensemble members will entirely be controlled by the variability of the values at hard conditioning points which are updated by ES-MDA. Therefore, we will be certain that a successful convergence of the proposed methodology will be mostly on account of the performance of the update algorithm rather than the randomness of the DeeSse simulations.

The question of the number of hard conditioning points to use is

case-dependent, and hence we cannot draw general conclusions based only on our synthetic case study. However, the optimal number will depend on the multi-resolution parameterization defined, i.e. the number of pyramid levels and the reduction factor applied in each direction of the TI for each level. In general, the pyramid parameters should depend on the largest size of the structures we wish to simulate. In the context of the proposed methodology, the idea is to ensure that the structures at the coarsest scale will be discernible with the lowest resolution possible so as to mitigate the expected discontinuities after each ES-MDA update of a non-multi-Gaussian distribution. Indeed, such a multi-resolution parameterization, alone, will allow to alleviate those consequences thanks to the resulting more continuous distribution at the coarsest scale. Since the continuity of the structures will be, in this manner, less compromised after an ES-MDA update, the optimal number of points will be allowed to be significantly higher than in a case without such a multi-resolution parameterization. However, it is important that the hard conditioning points are distributed evenly instead of being clustered. This way, DeeSse can play its role of retrieving the connectivity between the points according to the compatible patterns found in the TI, as shown in Figure 6.6. We will further address the benefit of the multi-resolution parameterization on the number of hard conditioning points when discussing the results obtained on our synthetic case in Section 6.6.

6.4.5 Complete procedure

To summarize, the proposed algorithm proceeds as follows. We start from the normal-score transformed variables obtained from the initial ensemble of coarsest-resolution MPS simulations, as illustrated in Figure 6.5. Then, Figure 6.7 illustrates the complete conditioning procedure for each ensemble member. The method includes the back-transform and sampling steps before the local integration of the ES-MDA update at the coarsest scale of the multi-resolution MPS simulation.

It is worth noting that the sampled hard conditioning points are ultimately the only points affecting the simulated categorical fields. Therefore, the sampling could be in fact performed before the normal-score transform step. In this manner, only the sampled transformed variables would be updated before being back-transformed as hard data for the multi-resolution MPS simulation. However, for the sake of illustrating

the motivation behind the sampling step, we purposely kept the description of the complete algorithm in this order.

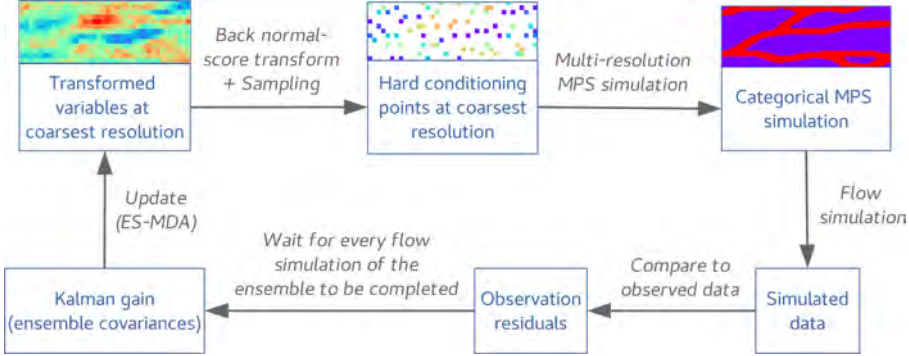


Figure 6.7: Workflow for each ensemble member to condition the categorical MPS simulations with ES-MDA. The multi-resolution simulations are regenerated based on updated parameters at the coarsest scale.

6.5 Synthetic test case

A synthetic numerical experiment is conducted to evaluate the performance of the proposed method. The model is inspired by the artificially induced flow resulting from the excavation of vertical shafts. The hydraulic impact of such perturbation has for example been monitored and studied at the Andra's Meuse/Haute-Marne since the operation of the Underground Research Laboratory (Benabderrahmane et al., 2014; Deman et al., 2015; Kerrou et al., 2017).

6.5.1 Model set up

The domain is two-dimensional and covers a vertical section of rectangular shape of 5000 m along the east-west direction and 500 m in depth. It represents a section of the Oxfordian limestone multi-layered aquifer above the Callovo-Oxfordian clay host formation which is not included in the model. For groundwater flow computation, the domain is discretized by 50×500 square elements of 10 m wide.

Along the top boundary, constant heads are imposed. The head values vary linearly between 267.5 m on the west side and 284 m on the east side. At depth, between 150 m and 500 m, a prescribed head of 280 m is

imposed on the east side to maintain a continuous east-west flow in the system (Figure 6.8). In order to mimic the type of hydraulic perturbation created by a vertical shaft, a subsequent transient flow simulation is calculated by adding from time zero, in addition to the boundary conditions previously listed, a seepage condition over the nodes corresponding to the first 300 meters of the western boundary starting from the bottom (Figure 6.8). We assume that the nodes over the remaining 200 meters towards the surface are not affected as we suppose that they are associated to an isolated upper part of the shaft. Note that in exploratory simulations, a sensitivity analysis was carried out on different levels of the finite element mesh refinement and the solver parameters, as well as on various types of flow boundary conditions and hydraulic property values. This was done in order to ensure robust numerical solutions and also to avoid bias from boundary effects.

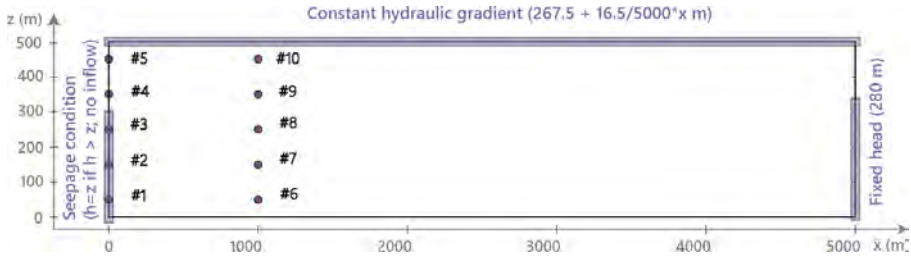


Figure 6.8: The two-dimensional vertical domain modelled of 5000 m length and 500 m depth. The thin rectangles indicates the location of nodes with a boundary condition applied. The dots indicates the head observation nodes. Not shown here, the left boundary is divided into 5 zones where the flowrates are observed.

For simplicity, the groundwater flow equation was solved numerically under saturated conditions. All the groundwater flow simulations in the framework of this study were performed with the simulator GroundWater (Fabien Cornaton, 2014). This numerical code uses the standard Galerkin finite element and the control volume finite element methods, and has been validated on the basis of a series of standard benchmarks by comparison with analytical solutions as well as with commercial numerical simulators.

6.5.2 Synthetic data set

The reference parameter field is a categorical MPS realization generated with DeeSse using the 100×1000 pixel binary training image (TI) shown in Figure 6.9. It represents the distribution of two geological facies in the domain. The TI was obtained using the 2500×2500 hand drawn image by Zahner et al. (2016) inspired by Strebelle (2002). Before extracting our 100×1000 TI from this forementioned image, we first resized its first dimension so to have more than one channel simulated on the vertical direction of our 50×500 simulation grid. The TI for our test case was chosen four times bigger than the size of the simulation grid used to simulate the categorical MPS realizations in order to ensure sufficient variability between the initial unconditional realizations.

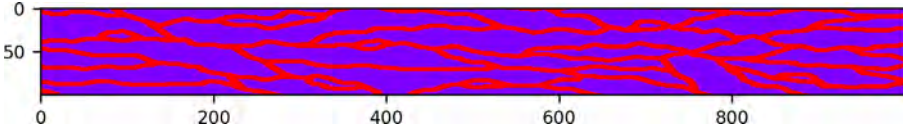


Figure 6.9: Binary training image of size 100×1000 pixels used to generate the multiple-point simulations with DeeSse. It is the same categorical TI for which an example of set of multi-resolution training images was shown previously in Figure 6.1.

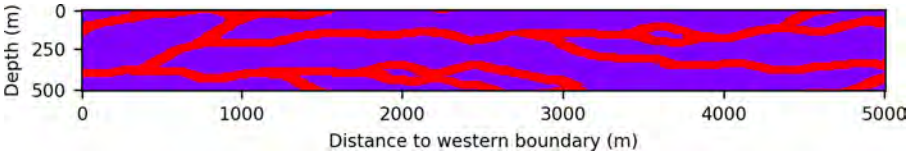


Figure 6.10: Reference field A of size 50×500 pixels corresponding to a multiple-point realization of the training image.

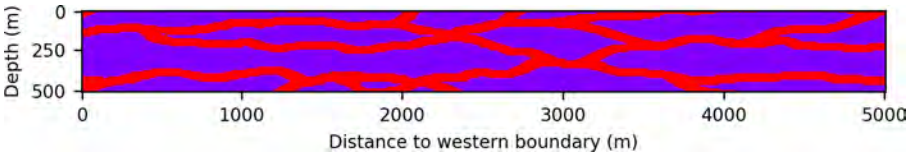


Figure 6.11: Reference field B size 50×500 pixels corresponding to a multiple-point realization of the training image.

To simulate the groundwater flow, each category of facies was pop-

ulated with hydraulic properties. An initial steady state flow before the activation of the seepage boundary condition was first computed by assuming constant values of hydraulic conductivity of 10^{-4} m/s for the channels and of 10^{-6} m/s for the less permeable matrix. For the subsequent transient groundwater flow simulation, the specific storage was set to a constant value of 10^{-6} m⁻¹ over the whole domain. The transient flow was computed over a duration of 43 200 s so that the expected decrease of hydraulic heads at the observation locations shown in Figure 6.8 can be well captured. Hence the drawdown is locally monitored in 10 locations at five different depths every 100 m. The first points are located along the seepage vertical boundary and the second ones at 1000 m from the first. The simulated heads were collected every 1200 s from $t = 0$ s to $t = 43\,200$ s, hence at 37 time steps. The resulting flowrates produced at the seepage western boundary were also measured, but at a smaller constant time step of 300 s for five zones of nodes spread along the boundary. To test the methodology proposed in Section 6.4, we will consider the ten time series of simulated hydraulic head as the data for the assimilation. The five flowrate data time series will then be used in addition in order to test the flow predictions obtained. For this inverse problem, we assume that the uncertainty related to the built groundwater flow model is negligible compared to the uncertainty on the facies distribution.

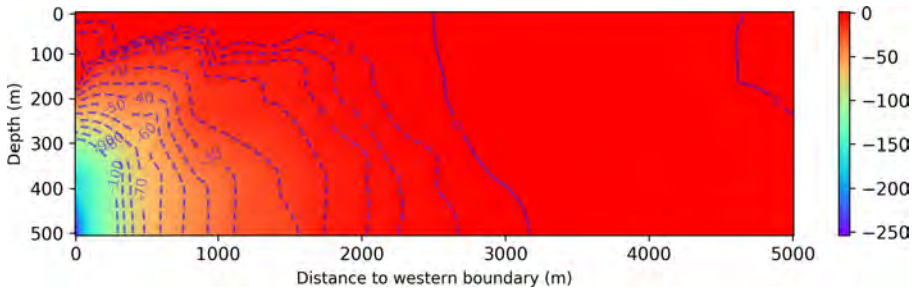


Figure 6.12: Drawdown (m) computed at the end of the transient groundwater flow simulation ($t=43\,200$ s) using the reference field A.

To test the performance of the methodology, we considered in this study two different reference fields of the facies distribution named ‘A’ (Figure 6.10) and ‘B’ (Figure 6.11). They differ by the extent of the drawdown region obtained by the end of the transient flow simulation ($t = 43\,200$ s) as illustrated in Figures 6.12 and 6.13. Indeed, the drawdown simulated in the case B goes around a hundred of meters further

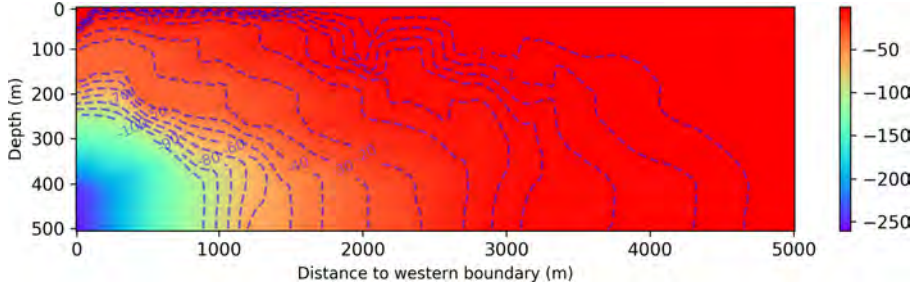


Figure 6.13: Drawdown (m) computed at the end of the transient groundwater flow simulation ($t=43\,200\text{ s}$) using the reference field B.

away from the western seepage boundary than the drawdown simulated with the reference A. This implies that the synthetic data set relative to the reference B is sensitive to a broader region of parameters as compared to the reference A.

6.5.3 Generation of the multi-resolution training images with DeeSse

To generate the multi-resolution training images, the parameters of the Gaussian pyramid technique, i.e. the number of levels and the number of nodes between each Gaussian-like kernel applications, were defined in such a way that the edges of the structures at the coarsest resolution of the pyramid would be as much blurred as possible while still being discernible. As shown by Figure 6.4, the almost uniform marginal distribution of the MPS simulations below a certain coarse resolution prevents to still detect the structures.

For the categorical TI in our synthetic case, if the most reduced image is only four times smaller than the original scale, as shown by the first level in Figure 6.1b, the image still appears more discrete than continuous. As a result of this still high resolution compared to the size of the structures in the resulting coarse simulations, the strong deviation from the multi-Gaussian assumption (Figure 6.4a) will likely cause the ES-MDA update to break the initial structures even after having applied beforehand the normal-score transform. In contrast, if the reduced image is 36 or 81 times smaller, the size of the pixels becomes too large to still detect the structures. In the end, a TI pyramid with two coarse levels and Gaussian-like kernels applied every two nodes was

generated (Figure 6.1). The coarsest scale of the multi-resolution MPS simulations is 13×125 pixels, hence 16 times smaller than the original size of 50×500 .

6.5.4 Initial ensemble of parameters and assumptions for the ES-MDA update

The initial ensemble of normal-score transformed parameters was obtained from the coarsest-scale simulations of an ensemble of 100 multi-resolution MPS simulations. Since only a subset of the 1625 gridblocks at the coarsest resolution will be taken into account to condition the multi-resolution MPS simulation for each ensemble member, using an ensemble of 100 members seems acceptable to approximate the covariance matrices involved in the computation the ES-MDA update. Even though a categorical MPS simulation was initially obtained from the coarsest-resolution MPS simulation, we remind that this categorical field is not the one used to compute the first update of ES-MDA. It is the one simulated from the initial member of parameters according to the forward model operator established in the conditioning procedure (Figure 6.7) that will be used to ensure a consistency between all the updates.

A diagonal covariance matrix of independent measurement errors of 0.05 m^2 for every head observation was defined. A critical length of 1600 m, hence almost twice the distance between the two series of observation points at $x = 0 \text{ m}$ and $x = 1000 \text{ m}$, was defined to localize the Kalman gain matrix on the basis of the Gaspari-Cohn function. By doing that, we are assuming that the estimated covariance between one parameter and an observation located beyond the defined critical length is too noisy to be trusted for the computation of the parameter update. As mentioned in Section 6.3.2, the contribution from these observations will not be considered in the parameter update by applying the forementioned distance-based localization. As a matter of fact, we defined this distance in such a way that the parameters located outside the perturbed area would not be updated (Figure 6.12). Note that this perturbed area is known precisely here only because we could simulate the flow over the whole domain knowing the reference field. In a real application, we would expect to have a similar knowledge of the extent of the hydraulic perturbation, but with much higher uncertainty.

6.5.5 DeeSse parameters for the multi-resolution MPS simulations

DeeSse parameters for multiscale MPS simulations	Initialization	Conditioning loop
Seed	From 777 to 876	999
Number of neighbours	64	64
Distance threshold	0.05	0.05
Maximal scan fraction at level 2 (coarsest)	0.25	0.8
Maximal scan fraction at level 1 (intermediate)	0.25	0.8
Maximal scan fraction at level 0 (original)	0.25	0.25

Table 6.1: DeeSse parameters used for multiscale MPS simulations in the proposed conditioning approach for categorical MPS realizations with ES-MDA.

We show in Table 6.1 the parameters of DeeSse used to generate the multi-resolution MPS simulations for our synthetic case. For the “Initialization” step, i.e. when the initial ensemble of multi-resolution MPS simulations is created, the seed value was set to be different for each ensemble member. This way, the MPS simulations are ensured to be different from each other. The maximal scan fraction was also set to 0.25 in order to ensure a sufficient variability in the initial ensemble of coarsest-resolution MPS simulations before the data assimilation. During the conditioning with ES-MDA, the seed was set to be the same value for all ensemble members and for all iterations in order to have a properly defined forward operator especially during the multi-resolution MPS simulation as explained in Section 6.4. As a matter of fact, it is possible to adjust the key parameters of DS, i.e. number of neighbours, distance threshold and maximal scan fraction, according to the level of the pyramid in a multi-resolution MPS simulation. For every level, we set the number of neighbours to 64 and the distance threshold to 0.05, both of which being in agreement with recommended parameters for MPS simulations with DeeSse (Meerschman et al., 2013). As for the

maximal scanned fraction, knowing that the size of coarser simulation grids is much smaller, we could afford to set at both coarse levels a high value of 0.8, hence three to four times greater than the more standard value of 0.25 defined for the original-scale MPS simulations.

6.5.6 Performance assessment criteria

In this section, we present the quantitative criteria used for assessing the performance of the proposed methodology on our synthetic case. In particular, these criteria aim to inform about how the approach deals with the issues often encountered when applying ensemble Kalman methods to condition non-multi-Gaussian realizations. They will allow us to address the quality of the data fit achieved, the plausibility of the categorical fields, and the accuracy of the estimated uncertainty.

Considering that the multi-resolution parameterization involves the regeneration of a multi-resolution MPS simulation after each iteration of ES-MDA (Figure 6.3), it is expected that the variability introduced by the MPS simulations at every scale will make the iterative reduction of the data mismatch more challenging. Checking the convergence of the reduction of the data mismatch over the ES-MDA iterations should therefore be the first indication as to the success of the proposed methodology. To evaluate the mismatch between the simulated and observed data, we will compute the sum of the squared deviations between the simulated data and observations normalized by the measurement error standard deviation.

$$D(m_i) = (g(m_i) - d_{\text{obs},i})^T C_{\text{err}}^{-1} (g(m_i) - d_{\text{obs},i}) \quad (6.6)$$

Even if the reduction of the data mismatch over the iterations turns out to converge to a sufficiently low final value for every final conditioned categorical MPS realization, the geological realism of the facies fields is not guaranteed throughout the inversion procedure using ES-MDA. Given their assumption of multi-Gaussianity, ensemble Kalman methods are indeed expected to “Gaussianize” any type of prior distribution. For this reason, perhaps the most important performance criteria for the assessment of the approach will be those checking the consistency of the final categorical MPS realizations of facies in comparison to the ones obtained initially before the conditioning. For this purpose, we will

compute the connectivity function (Renard and Allard, 2013) of the facies of value s_i , noted $\pi_i(h)$,

$$\pi_i(h) = \text{Prob}\{x \leftrightarrow x + h | s(x) = s(x + h) = s_i\} \quad (6.7)$$

which describes the probability that two points separated by a distance h of the same facies of value s_i are connected (6.7). We will compute it before and after the data assimilation for the facies of value 1 which represents the channels in our case.

In most applications, ensemble Kalman methods may significantly underestimate the true uncertainty of the estimated variables (Chen and Zhang, 2006). We assume that this underestimation would be even greater if the prior distribution is not multi-Gaussian. A main concern when applying an ensemble Kalman method is therefore to obtain a proper estimate of the reduced uncertainty, or in other words, one which is not too much overreduced. Based on the initial and updated ensembles of facies realizations, the uncertainty reduction can be quantified for each facies by observing the variability of the ensemble of facies indicator fields during the assimilation. This variability can be measured by calculating the spatially-averaged uncertainty estimated from the ensemble of indicator fields, also known as the “ensemble spread” as follows

$$S_{\text{ens}} = \sqrt{\frac{1}{N_m} \sum_{i=1}^{N_m} \sigma_{\text{ens},i}^2} \quad (6.8)$$

where $\sigma_{\text{ens},i}^2$ denotes the variance at one gridblock i from the ensemble of indicator realizations.

To check the accuracy of the uncertainty reduction estimated by ES-MDA in the proposed methodology, we will compare the ensemble spread to the root-mean-square error (RMSE) which measures the discrepancy between the ensemble mean of the facies indicator realizations and the reference facies indicator field as follows

$$RMSE = \sqrt{\frac{1}{N_m} \sum_{j=1}^{N_m} (m_{\text{true}}^j - m_e^j)^2} \quad (6.9)$$

where N_m is the number of gridblocks, m_{true}^j is the value at node j in the reference facies indicator field, and m_e^j is the value at node j of the mean

of the ensemble of facies realizations. The RMSE can then be calculated based on the initial ensemble of facies realizations generated initially, and based on the ensembles of conditioned categorical fields obtained after each ES-MDA update. If the uncertainty is properly estimated during the data assimilation, the ensemble spread should be close to the RMSE (Chen and Zhang, 2006). Indeed, an uncertainty reduction estimated by the ensemble should normally translate into a reduction, of the same magnitude, of the discrepancy between the ensemble mean of facies indicator realizations and the reference. Note that, both the ensemble spread and RMSE are both unitless since we are considering either the indicator variables, their mean or their variance.

Assuming that the amount of information provided by the data is sufficient, the performance of the proposed methodology could also be assessed by its ability to obtain facies realizations that are closer to the reference than before the assimilation. To measure this, we can consider a root-mean-square error defined differently from previously, noted $RMSE_2$, which would here consider the discrepancy, on average, of the reference with each facies indicator realization as follows

$$RMSE_2 = \sum_{i=1}^{N_m} \sqrt{\frac{1}{N} \sum_{j=1}^N (m_{\text{true}}^i - m_j^i)^2} \quad (6.10)$$

where N is the ensemble size, N_m is the number of gridblocks, m_{true}^i is the value at node i in the reference field, and m_j^i is the value at node i in the j^{th} facies indicator realization of the ensemble.

6.6 Results and discussion

6.6.1 Evolution of the data mismatch

The proposed conditioning approach was first tested by assimilating the transient hydraulic head data generated from the reference ‘A’ (Figure 6.10) at the ten observation points (Figure 6.8). Based on 6×41 hard conditioning points located approximately every two cells vertically and every three cells horizontally, hence a sampling rate of 15 % at the coarsest level, a convergence of the reduction of the head data mismatch was observed (Figure 6.14). Figure 6.15 confirms that the data match at every head observation point was effectively improved after 16 ES-MDA iterations. In addition, the flowrate observations were well predicted

by the final categorical fields although they were not considered for the assimilation.

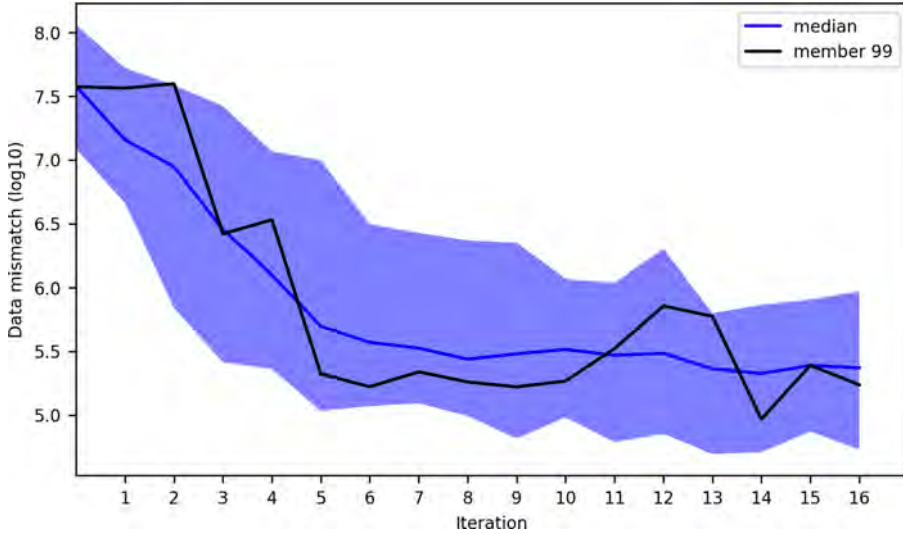


Figure 6.14: Evolution with the iterations of the decimal logarithm of the data mismatch for the test case with reference A. The solid blue curve is the median, the blue area corresponds to the region between the percentiles P5 and P95 of the ensemble of values, and the dashed black curve is the values computed for the ensemble member no. 99.

As explained in Section 6.2, the same seed value and same hard conditioning locations for all ensemble members during the conditioning loop ensure that the convergence can be fully attributed to the integration of the ES-MDA update at those points. We also repeated the test using other unique seed values than the one indicated in Table 6.1 and observed each time a convergence of the approach.

In comparison to applications of ensemble Kalman methods in the multi-Gaussian case or in MPS cases without facies resimulation within the update procedure, it is worth mentioning that the decrease of the data mismatch observed here is not as smooth. Similarly to the updating of truncated Gaussian realizations via the underlying multi-Gaussian variables, the oscillations during the reduction of the mismatch are indicative of nonlinearities in the forward model operator relating the parameter variables and the dynamic data to assimilate. In this categorical MPS case by using the proposed approach, the oscillations can be large because the process allowing to obtain the categorical field from

the coarsest-scale parameters, i.e. the multi-resolution MPS simulation, is nonlinear in the same manner as the truncation step is nonlinear in the truncated Gaussian case (Liu and Oliver, 2005).

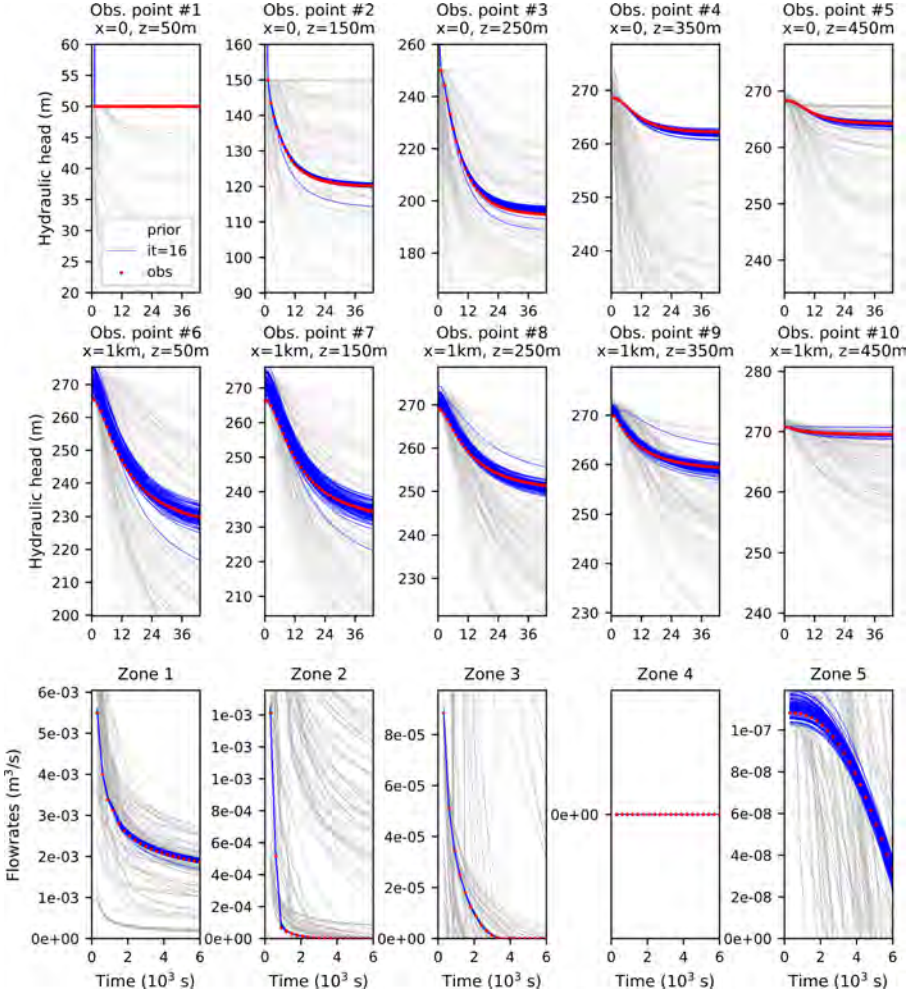


Figure 6.15: Predicted hydraulic head and flowrate data at the observation locations, before (in gray) and after (in blue) the assimilation of the hydraulic heads only. The red dots are the observed data; the ensemble size is 100.

6.6.2 Sensitivity to DeeSse parameters of the integration of the ES-MDA update

As shown in Table 6.1, we allowed in the tests to further scan the TI at both coarse levels including the one where the ES-MDA update is integrated. In the context of the proposed methodology, the maximal scanned fraction at the coarse scales can indeed influence the convergence of the algorithm. It makes sense since by increasing the scanned fraction of the coarsest TI, the more likely a more compatible pattern which include the hard data derived from the ES-MDA update will be found, and hence the better the integration of the ES-MDA update will be. It is worth noting that this constraint on the maximal scanned fraction at the coarse levels should be computationally affordable. Indeed, given the significant reduction by a factor 4 of the simulation grid at each coarse level, the search for patterns in the corresponding training image will also be faster.

By testing several distance thresholds, we observed that a value lower than 0.01 or greater than 0.1 would result in a less efficient convergence and an insufficient data match. Even if the history match had turned out to be acceptable, the biggest issue was the fact that these tested values did not allow the retrieval of the expected structures after the data assimilation. After some tuning, we found that a value around 0.05 was a good compromise to obtain simultaneously both a good history match and plausible structures. The consistency of the retrieved structures is confirmed by evaluating the connectivity function of the channel facies for each final categorical MPS realization. Figure 6.16 shows that the range of final connectivity values obtained is consistent with the range obtained with the initial facies fields generated before the conditioning. The median values before and after the assimilation indicate a global increase of the connectivity of the channel facies.

It is important to keep in mind that in the proposed approach, the hard conditioning points used at the coarsest scale of the multiscale MPS simulations do not represent actual data but are only the result of the back transformed parameters updated with ES-MDA at those points. Since ES-MDA updates the parameters iteratively and in doing so improves the estimates over the iterations, it also means that the values at the sampled hard conditioning points will correspond to less good estimates at the first iterations of ES-MDA. This is especially true in a non-multi-Gaussian case with the degradation of the structures as soon as the first ES-MDA update is applied. As a result, in order for

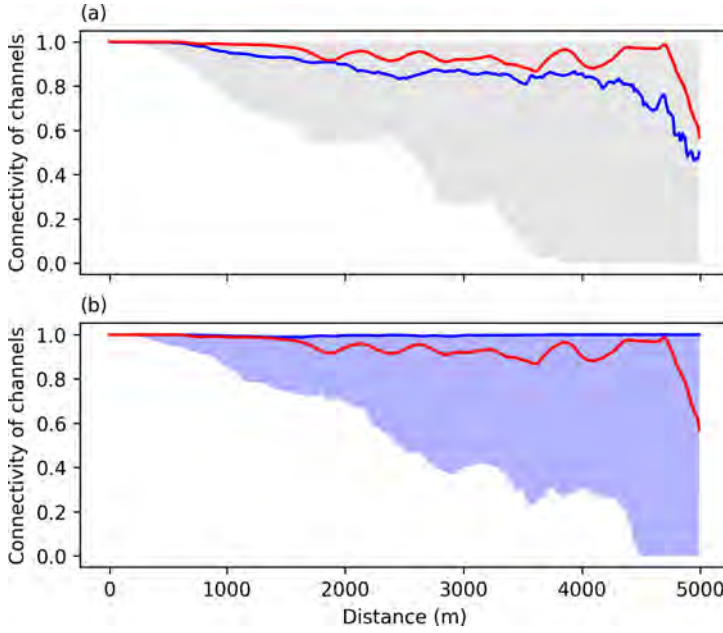


Figure 6.16: Connectivity function of the channel facies computed in the horizontal direction for the ensemble of categorical MPS simulations: (a) before the update (b) after all the ES-MDA updates. The gray and blue regions correspond to the values between the percentiles P5 and P95. The red curve is the connectivity function calculated with reference A and the blue curve is the median of the ensembles of values taken by the connectivity function.

our conditioning approach to be successful, DeeSse must take on the critical role of helping retrieve the structures especially during the early iterations of ES-MDA by not giving too much weight to the first estimates of the hard conditioning points. That is why defining a too low distance threshold value, such as 0.01, does not make sense here because DeeSse should not too rigorously search for matching patterns given the inaccurate first ES-MDA estimates as hard data. Instead, the threshold should be defined in order to allow some flexibility when accommodating the less reliable estimates of the first iterations. In this way, DeeSse will give preference to the simulation of patterns that are as consistent as possible to the TI given the hard conditioning points, rather than try to exactly honor them at the beginning of the data assimilation and simulate inconsistent structures. However, if this distance threshold is too large, for example 0.1, the ES-MDA updates at the later iterations will

be on the contrary not sufficiently integrated into the multi-resolution MPS simulations, thereby compromising the efficient convergence of the conditioning approach. Even though it was not tested, one may consider to adapt the threshold in a decreasing manner from one iteration to another to improve the integration of the update and hence the robustness of the convergence.

6.6.3 Identification of the categorical fields

Figure 6.17 shows the categorical realizations obtained at different stages of the conditioning procedure for four ensemble members. As expected, the simulated structures become more and more consistent over the ES-MDA iterations. Although not shown here for brevity, the $RMSE_2$ calculated in the updated region has decreased, thereby confirming that the conditioned fields resemble the reference. The probability maps of the channel facies, calculated before and during the conditioning (Figure 6.18), shows the progressive identification of the structures and retrieval of the connectivity during the assimilation. In between the two columns of observation points, the channels are characterized with a very low uncertainty and they look very much like the channels of the reference A in the same region. Upstream to the last locations informed by data, at around $x = 2500$ m, a progressive increase of uncertainty can also be observed. Overall, the identification of the parameters is consistent with the distribution of the data points and the distance-based localization that was applied to restrain the update to the vicinity of the data points.

As shown in Figure 6.19, the ensemble spread calculated from the facies realizations significantly decreased in the updated region during the data assimilation. In our case, this ensemble spread represents on average, in the updated region, the spread around the probability of the channel facies.

In the non-multi-Gaussian bimodal case of log hydraulic conductivity fields, Zhou et al. (2011) illustrated the influence of the ensemble size on the accuracy of the estimated uncertainty with the normal-score EnKF approach. They concluded that using an ensemble size of the order of 100, such as in our example, led to a severe uncertainty underestimation as quantified by the very large discrepancy between the ensemble spread and the RMSE at the end of the data assimilation.

While the RMSE is an appropriate measure of the true uncertainty

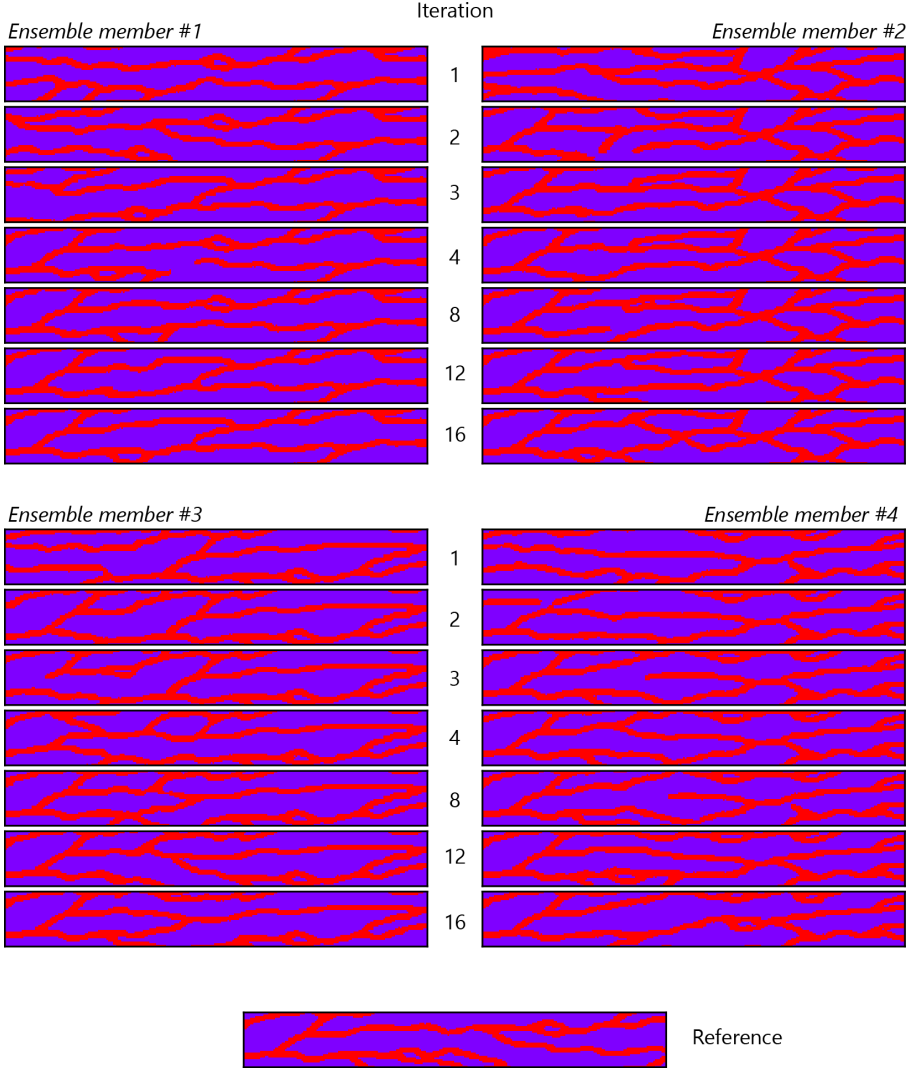


Figure 6.17: Conditioned categorical MPS simulations obtained after certain iterations of ES-MDA out of 16 in total, for 4 ensemble members. The image at the bottom is the reference field A.

in the multi-Gaussian case, we note that it is not a rigorous one in the discrete case. In the multi-Gaussian case, it is assumed that the reference field is the mean of a multi-Gaussian distribution of realizations which are all conditioned to the data. It therefore makes sense to calculate in that case the distance between the ensemble mean and

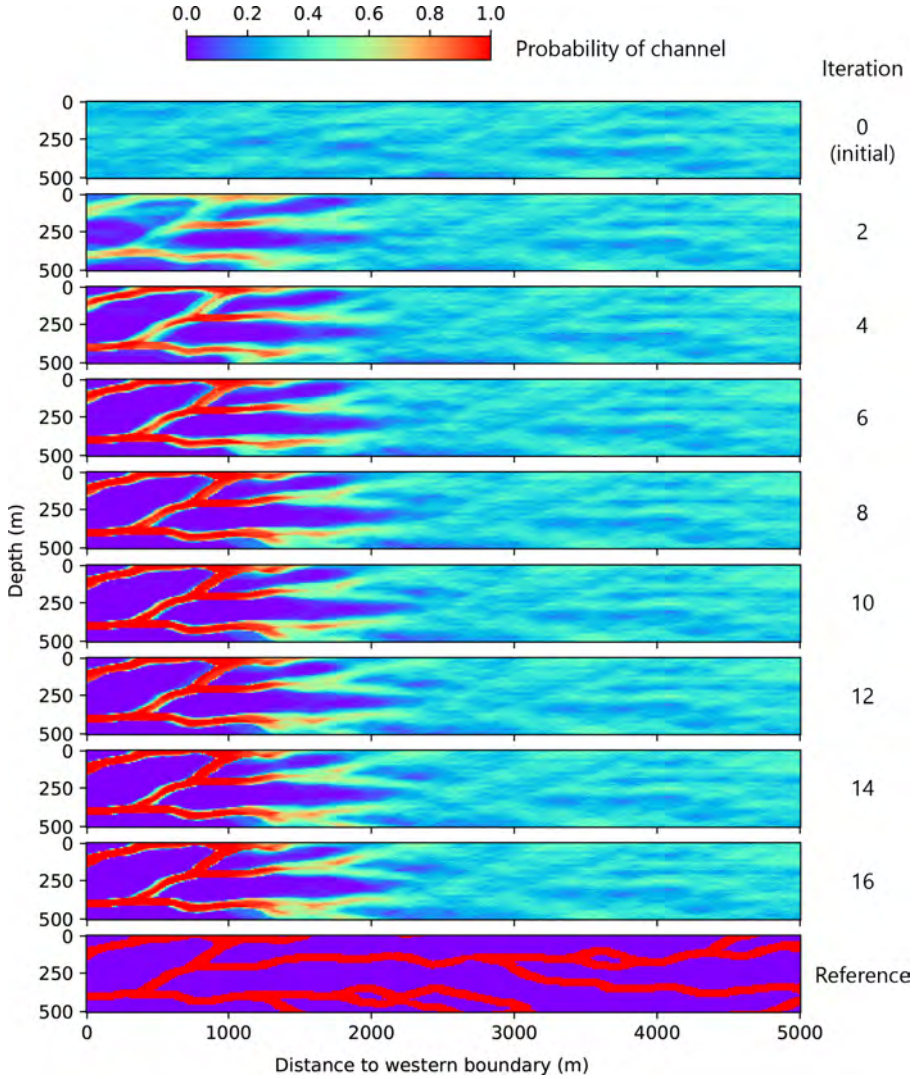


Figure 6.18: Probability map for the channelized facies calculated from the ensemble of categorical MPS simulations before conditioning and after every two iterations of ES-MDA using the hydraulic head observations from reference A (shown at the bottom).

the reference in order to obtain an estimation of the true error. In the discrete case, however, the reference field is a random sample from a distribution which can be much more complex. Hence the RMSE which measures the distance between this reference sample and the ensemble

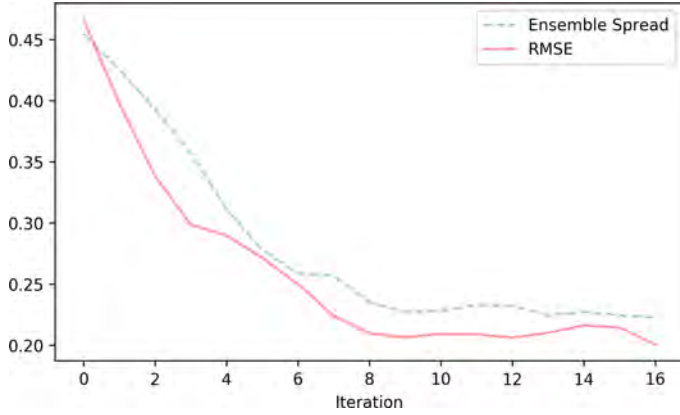


Figure 6.19: Evolution of the RMSE and the ensemble spread (both unitless) before and during the conditioning of the categorical MPS realizations in the updated region (up to 1600 m from the seepage boundary) for the ES-MDA update with the synthetic data of reference A.

mean could ultimately result in a large underestimation of the true uncertainty. Therefore, although it can be observed in Figure 6.19 that both the ensemble spread and RMSE decrease to a same value after all the updates thanks to the proposed methodology, this convergence ultimately does not tell us how accurate the estimated uncertainty is.

To test the uncertainty estimated with our approach, we used the rejection sampling method to generate samples from the posterior distribution and used it as the reference distribution. Rejection sampling consists in sampling from the prior distribution of unconditional categorical realizations and accepting each sample with a probability calculated based on the ratio between the likelihood of the sample and the highest likelihood value which can be obtained (Tarantola, 2005). The ensemble of accepted samples follows the posterior distribution. Figure 6.20a shows the probability map obtained based on 108 samples which have been accepted out of 185 000 unconditional realizations generated with DeeSse. By visually comparing this map to the one calculated based on the ensemble of 100 conditioned categorical fields obtained by data assimilation using our proposed methodology (Figure 6.20b), we can see clearly that the posterior distribution obtained by rejection sampling also identifies with very low uncertainty the location of the channels in the same region which was updated by data assimilation using our proposed approach, i.e. up to $x = 1600$ m from the western boundary.

Moreover, the structures identified by both approaches in this region are very similar to each other. To measure quantitatively how one distribution is different from another distribution, one could for example calculate the Kullback-Leibler divergence (Kullback and A Leibler, 1951). For this case, a low value close to zero was calculated and indicates that the two distributions are indeed very similar. Therefore, we can finally conclude that the uncertainty reduction which led to the identification of the categorical fields with very low uncertainty using our proposed methodology is mostly the result of an efficient assimilation of the data rather than of a suboptimal application of the ensemble Kalman method due to an insufficient ensemble size as pointed out for example in Zhou et al. (2011) or Chen and Zhang (2006).

Despite the use here of a small ensemble size of 100, we assume that the efficiency of the proposed data assimilation methodology results from the combined effect of: (1) the reduction of the number of parameters via the multi-resolution parameterization and the sampling, and (2) the role of DeeSse in finding consistent patterns. Indeed, although the whole coarsest-resolution field was updated in our case, the number of effective parameters was ultimately the number of sampled hard conditioning points, hence 246. Given the ensemble size of 100, we can assume that the covariance matrices were more correctly approximated from the ensemble.

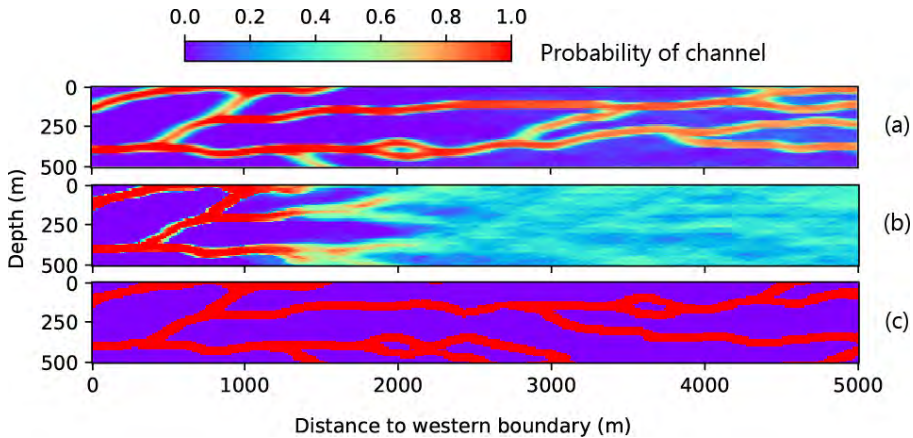


Figure 6.20: (a) Probability map obtained by rejection sampling. (b) Probability map obtained after the final update at iteration 16 with the proposed methodology. (c) Reference field.

6.6.4 Analysis of case B

To further assess the performance of the proposed methodology, we also performed tests using the synthetic set of head data generated from the reference field B (Figure 6.11). A longer critical length of 2600 m was defined in this case to take into account the larger extent of the drawdown at the last timestep of the synthetic data set (Figure 6.13). Therefore, the region of parameters to update by data assimilation in this case was also larger.

Using a similar density of sampled hard conditioning points as in the previous case with reference A, the approach converged successfully and towards a good final match of the transient head and flowrate data. (We do not illustrate those results here for the sake of brevity.) However, we note from Figure 6.21 that upstream from the second set of data points, between approximately $x = 1600$ m and $x = 2600$ m, the very low uncertainty on the updated channel locations is clearly inaccurate given how very few channels were identified and how the channels seem less connected compared to our prior knowledge from the TI. We can therefore conclude that the uncertainty in this area which is less informed by data is significantly underestimated. Figure 6.22 shows a much larger reduction of the ensemble spread compared to the RMSE in the updated region up to 2600 m from the seepage boundary, confirming that those very low probabilities of channels between 1600 m and 2600 m after the data assimilation in Figure 6.21 should not be relied upon.

Although we do not show it here for brevity, the RMSE actually decreases in the same manner as the ensemble spread if only the region up to 1600 m from the seepage boundary is considered. This indicates that the uncertainty is better estimated in this more informed area of the updated region. Also, the decrease of the RMSE suggests that the data assimilation was especially efficient with probabilities of channels getting closer to the reference which represents here the indicator values of the channel facies. Figure 6.21 shows clearly the much more consistent structures before $x = 1600$ m, and also that the identified structures look very much like those from the reference in the same region.

Considering that the number of parameters to update in this case B is larger than in the previous case A, we assume that the severe underestimation of the uncertainty observed beyond $x = 1600$ m in the update region is related to spurious correlations affecting the computed parameter updates. Indeed, the observation points at $x = 0$ m and $x =$

1000 m may be too distant from the parameters located in that region to inform them with sufficient accuracy by ensemble data assimilation. As mentioned in Section 6.3.2, the performance of ensemble Kalman methods can be severely reduced by an excessive underestimation of the uncertainty due to spurious correlations between parameters and observations far apart.

To avoid those inaccurate estimations and resulting inconsistencies in the conditional categorical fields, a solution could be to sample the predefined hard conditioning points only at locations of the updated region which are sufficiently informed by nearby observations. In this manner, locations that are likely to be affected by spurious correlations can be simulated with DeeSse instead, thereby preventing the unreliable updates from being taken into account in the multi-resolution MPS simulation. Although this strategy is ad-hoc, it is somewhat very similar to the way the widely used and also ad-hoc covariance inflation method must be implemented in practice. As pointed out by Hamill and Whitaker (2005), inflating the ensemble spread uniformly over the whole model domain will lead to an undesirable unbounded growth of the ensemble variance at locations with nearly nonexistent data. Hence while covariance inflation should be performed only in areas experiencing a significant uncertainty reduction, i.e. near the observation locations, the sampling of the updated conditioning points in our proposed methodology should be performed similarly only near the observation locations since those sampled points will not be affected too negatively by the unavoidable long range spurious correlations. Figure 6.23 illustrates this sampling strategy close to the observations located at $x = 0$ m and $x = 1000$ m, hence leaving out the uninformed part between $x = 1600$ m and $x = 2600$ m of the updated region.

As can be observed in Figure 6.23, it is important for the stability of the results to also have hard conditioning points sampled from the non-updated region, i.e. beyond $x = 2600$ m. Indeed, those invariant samples will ensure that the simulated structures do not noticeably change from one iteration to the other in this area, thereby allowing approximately the same contribution from this region to affect the predicted flow observations. We assume that such invariant contribution throughout the iterations affect positively the related covariance matrices by making the parameter estimation more tractable for the proposed methodology. As suggested by the probability maps obtained using this adaptive sampling strategy (Figure 6.24), the consistency of the identified structures is this time preserved during the data assimilation no matter the loca-

tions of the observations in the updated region. Moreover, despite the larger role given to DeeSse here in the uninformed regions, the proposed methodology still converged successfully towards a good final match of the transient head and flowrate data (Figure 6.25).

Ultimately, the case B illustrates the broader applicability of the proposed methodology in more realistic situations where the observation points may not be evenly distributed in the region of interest for the parameter identification.

6.6.5 Comparison with other approaches

We consider that the greatest advantage of the proposed methodology is the nature of the link between the continuous latent variables updated by the ensemble Kalman method and the conditioned facies fields. The Gaussian pyramid technique relates directly the value of the latent variable at a coarse resolution with the categorical patterns for a given realization at a finer scale. It allows a reduction of the dimension of the problem and gives a direct relation between the continuous values and the facies field.

Among the previously published methods, a classical technique is to use the local hydraulic properties as a latent variable and avoid using the facies during the update. For example, the normal-score EnKF (Zhou et al., 2011) updates the continuous hydraulic conductivities populating an initial ensemble of facies distribution in a full postprocessing procedure. We note that for such approach to work, the hydraulic conductivities must be defined in a continuous manner otherwise the normal-score transform function will not be defined. In the proposed approach, although the exact same normal-score transform is used, the advantage is that the latent variable is by construction necessarily continuous at every location of the coarsest resolution level (Figure 6.4).

Cao et al. (2018) proposed to update only the conductivity values at randomly selected pilot points. The updated values are then used as hard data for a conditional MPS simulation with a TI representing continuous log hydraulic conductivities that populate a categorical facies field. Their motivation for using pilot points was actually the same as ours. Indeed, they noticed that updating the whole field of log hydraulic conductivities with iterative ensemble smoothing do not allow to preserve the prior non-Gaussian spatial patterns. Following the same strategy, we do not update the complete coarsest resolution. But unlike

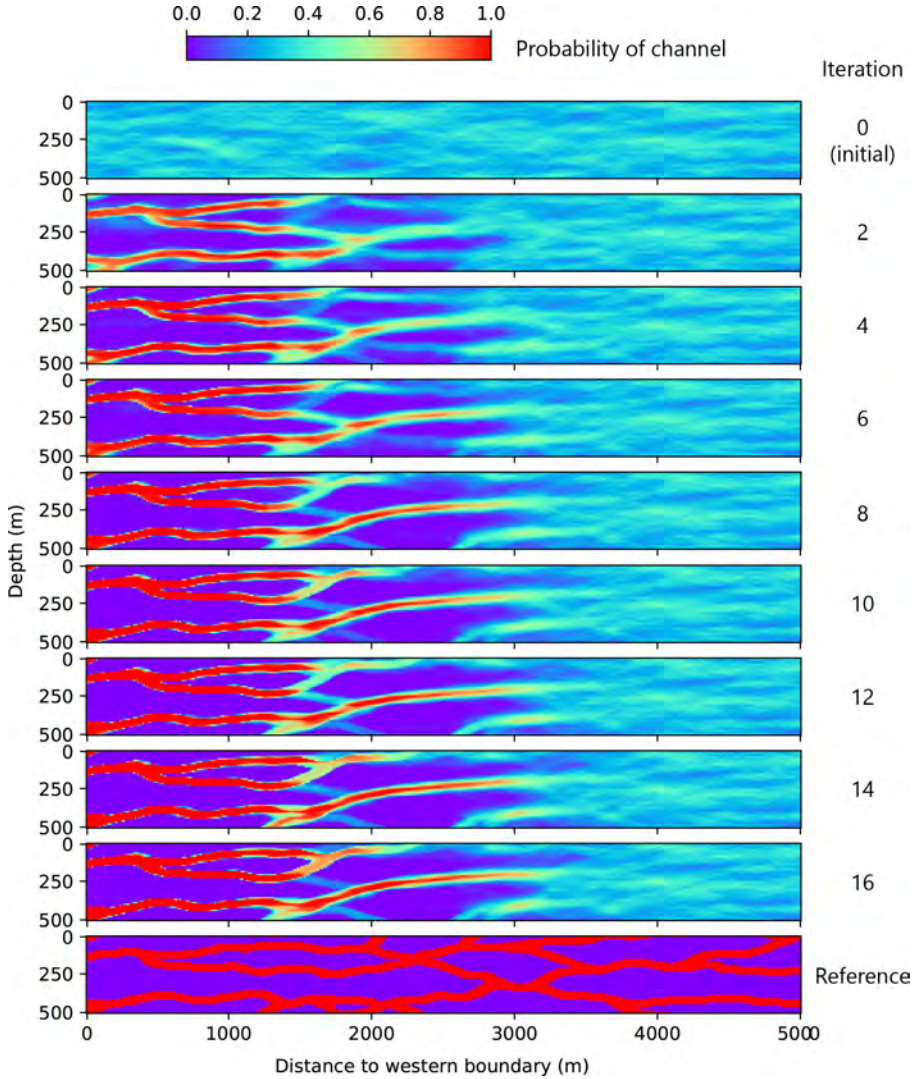


Figure 6.21: Probability map for the channelized facies calculated from the ensemble of categorical MPS simulations before conditioning and after every two iterations of ES-MDA using the hydraulic head observations from reference B (shown at the bottom).

Cao et al. (2018) who apply the updates at the fine resolution, our approach will better preserve the structures even if the updates are applied on the complete field at the coarsest scale because of the multi-resolution strategy.

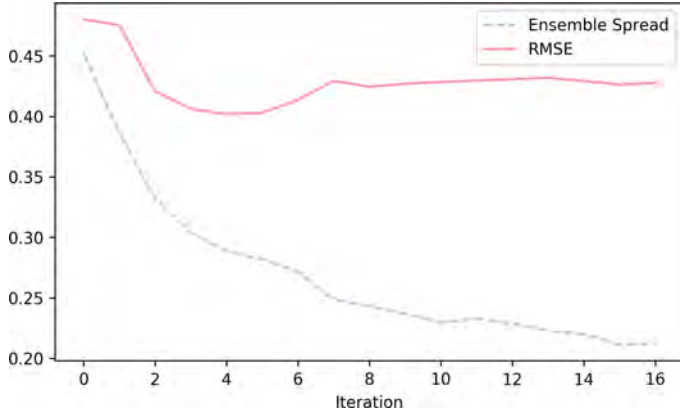


Figure 6.22: Evolution of the RMSE and the ensemble spread (both unitless) before and during the conditioning of the categorical MPS realizations in the updated region (up to 2600 m from the seepage boundary) for the ES-MDA update with the synthetic data of reference B.

Still considering the use of hydraulic conductivities as primary variables for the EnKF update is possible when the TI is continuous as used for example in Cao et al. (2018). Although it is beyond the scope of this study, it is worth mentioning that multi-resolution simulations with DeeSse based on a continuous TI originally can also be performed. Therefore, the proposed methodology is not limited to the categorical case described in the synthetic example treated in this paper but could also be applied in the continuous case. As compared to previous work, the gain would again be the reduction of problem dimensions through the use of the pyramids.

Ultimately, comparing the normal-score EnKF to our approach, we note that the reduction of the data mismatch seems to be more efficient for all members of the final ensemble with our proposed method. This could be attributed to the randomness of the conditional MPS simulations which is probably not sufficiently constrained as in our proposed methodology with both the use of the same seed and the same fixed hard conditioning points for all ensemble members. Another reason for the mismatch could be related to the fact that no other mitigation strategies than the pilot points are employed to reduce the loss of consistency of the simulated structures. Indeed, we believe that the multi-resolution parameterization alone in our approach allows to alleviate this problem

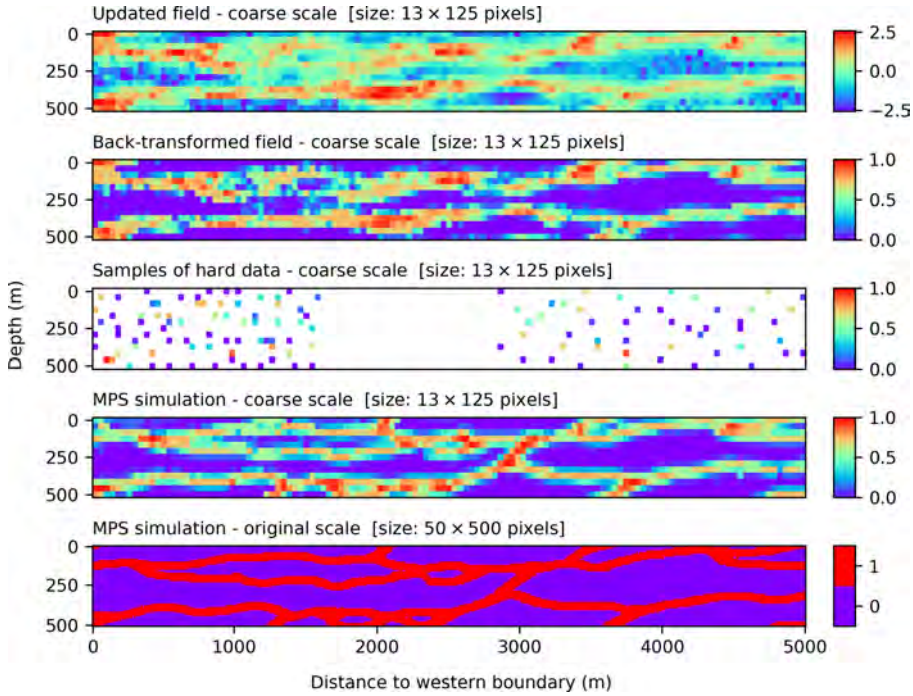


Figure 6.23: Conditioning a categorical MPS realization (bottom image) from the updated normal-scored image at the coarsest pyramid scale (ensemble member no. 11) at iteration 8 out of 16 of ES-MDA (top image) via fixed conditioning points sampled from the back-transformed image near the observation observations for the multi-resolution MPS simulation.

and will ultimately positively influence the correction applied at the hard conditioning points.

Another interesting approach is the probability conditioning method (PCM) of Jafarpour and Khodabakhshi (2011). It also uses the hydraulic conductivity as a continuous variable to update the fields, but here the link with the categorical MPS simulations is controlled using a probability map. The updated variables is related to the facies distribution in an indirect manner. At each iteration, the hydraulic conductivities are updated, and a probability map is derived from the mean of the ensemble of updated log hydraulic conductivity field using some assumed triangular probability density functions. Once the corrected probability map is obtained, a new set of MPS simulations is generated. All these simulations are conditioned to this unique probability map using

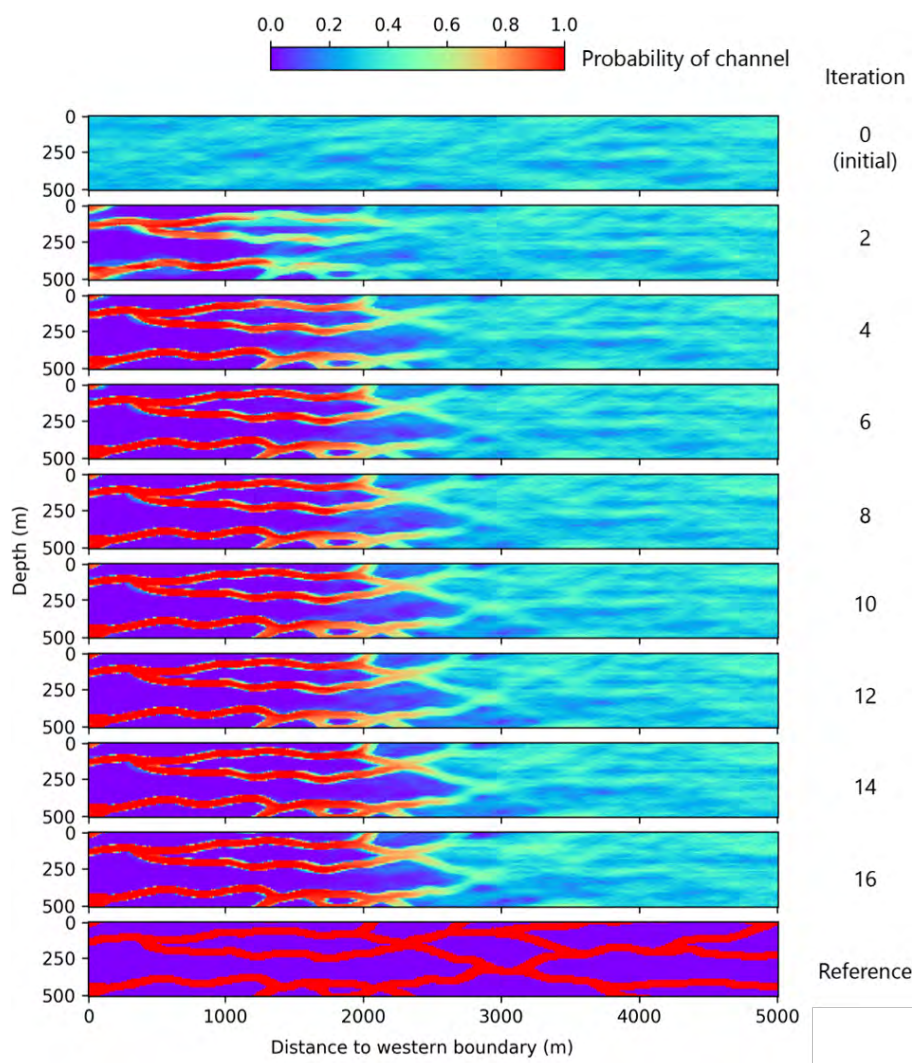


Figure 6.24: Probability map for the channelized facies calculated from the ensemble of categorical MPS simulations before conditioning and after every two iterations of ES-MDA using the hydraulic head data generated from reference B (bottom image). The hard conditioning point locations as shown in Figure 6.23 were used.

SNESIM and the τ -model. Although this step allows to control the conditional facies MPS realizations, the derived probability maps are prone to significant noise due to the initial non-multi-Gaussian distribution of hydraulic conductivities, and the structures are not well preserved.

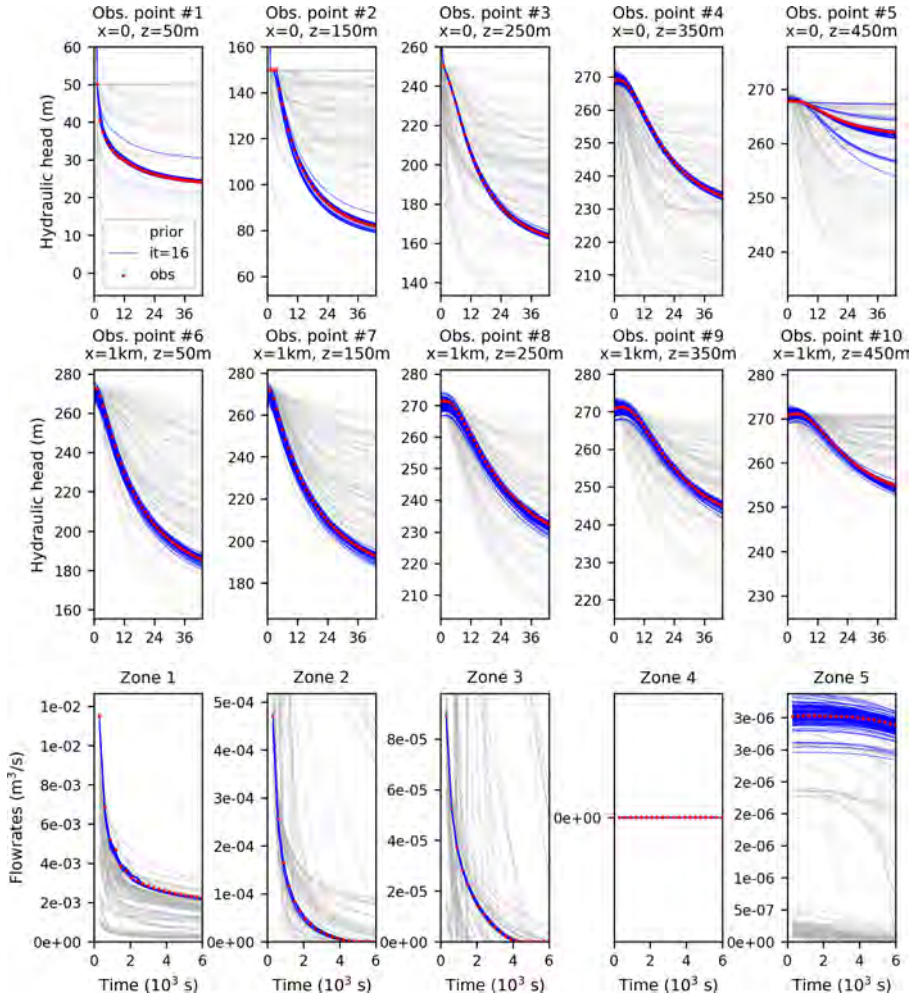


Figure 6.25: Predicted hydraulic head and flowrate data at the observation locations, before (in gray) and after (in blue) the assimilation of the hydraulic heads only. The red dots are the observed data; the ensemble size is 100. The hard conditioning point locations as shown in Figure 6.23 were used.

Another complication in the PCM comes directly from how soft data are considered in the MPS algorithm employed SNESIM. Indeed, the integration of a probability map requires the definition of a parameter, also known as the tau parameter, which will control the weight given to the probability map relatively to the conditional probability estimated from the TI. As a result, such integration of a probability map as soft

data in the MPS simulations is bound to not be fully taken into account and its impact on the final probability obtained will hence be somewhat arbitrary. This ultimately make the impact of the estimated soft data and hence the actual performance of the PCM difficult to assess compared to a less ambiguous hard data-based conditioning approach.

Ma and Jafarpour (2018) proposed recently a pilot point parameterization for the hard conditioning of categorical MPS simulations with an ensemble Kalman approach. Similarly to Cao et al. (2018), an ensemble of log hydraulic conductivities is updated without prior normal-score transform. Also, in the same manner as in the PCM, the values estimated at the pilot points are converted to facies probabilities. These probabilities are then used to sample a facies type, thereby allowing a conditioned categorical MPS simulation using the updated facies as hard data. However, unlike Cao et al. (2018) or in our proposed approach, they select the hard conditioning points based on the combined information of the sensitivity of the model response to the parameters and the variance of the parameters based on the ensemble of log hydraulic conductivity fields. As a result, one main difference with our approach is that when using ES-MDA, the locations of the pilot points may change from one iteration to another according to the updated sensitivity and parameter variance. As observed by the less efficient reduction of the data mismatch compared to our proposed methodology, the randomness introduced by each conditional MPS simulation with the proposed selection of pilot points may lead to a less tractable and hence less robust convergence.

Finally, as illustrated in our case with the reference A, provided that the ES-MDA update is applied with care to avoid underestimating the uncertainty in insufficiently informed areas, it is conceptually simpler to rely on ES-MDA to compute proper estimates of the parameter and variance no matter the location. In this way, we can consider that every location can potentially be a hard conditioning point. Then, the choice of only a subset of locations can be done in a simple, uniform manner enabling DeeSse to properly retrieve the structures inferred by ES-MDA in between the same fixed locations and by using the same seed for all ensemble members.

As an alternative to the standard EnKF, Zhou et al. (2012) also proposed the Ensemble PATtern matching method (EnPAT) which was then further applied and extended (Li et al., 2014; Lochbühler et al., 2014). Unlike the proposed methodology which still relies on the linear

update step of ensemble Kalman methods, the update step in EnPAT is fully replaced by the search of matching conditional patterns composed of both model parameters and state observations. Although this method ensures by construction the reproduction of curvilinear patterns at multiple scales, a main drawback however is the very large number of realizations required for the matches to be found (Li et al., 2014).

6.7 Conclusion

In this paper, a new approach to condition categorical MPS realizations with an iterative ensemble Kalman method was introduced. The methodology uses a multi-resolution parameterization with the code DeeSse based on the Gaussian pyramid technique which massively reduces the number of parameters to update. Initially, these parameters correspond to the normal-score transformed values of the coarsest level of a multi-resolution MPS simulation. After each update, the parameters can be integrated at the coarsest level of a multi-resolution MPS simulation to finally condition the categorical simulation performed at the scale of the original categorical TI. Although it was not in the scope of this study, it is worth pointing out the applicability of the same multi-resolution parameterization to originally continuous cases as well.

At the coarsest scale of the multi-resolution MPS simulation, only a subset of locations are updated with ES-MDA. As a result, the methodology also heavily relies on the ability of DeeSse to recover the structures by the integration, as hard data, of the result of each data assimilation. The more the number of hard conditioning points considered, the more control over the data match, but the less control over the consistency of the simulated structures. However, our results obtained for two synthetic examples show that the proposed multi-resolution parameterization has the potential to satisfy simultaneously both the requirements on the quality of the data match and the consistency of the updated fields with the prior geological information. Because the locations of the hard conditioning points are chosen to be fixed, and identical for all ensemble members throughout the assimilation, our specified parameterization of the ensemble of multi-resolution MPS simulations effectively ensures the stability of the proposed algorithm. Indeed, the robust convergence observed in our synthetic case, for any of the two tested references, can unambiguously be attributed to the improved estimated values at the hard conditioning points throughout the assimilation.

Our numerical experiments tend to show that both the multi-resolution framework and the sampled hard conditioning points contribute to mitigate the limitations resulting from the ensemble Kalman method in our non-multi-Gaussian case. As long as the amount of information given by the synthetic data is sufficient to identify the region of interest, our results show that the proposed methodology can successfully condition facies realizations that both honor the dynamic observations and are consistent with the prior conceptual model.

Moreover, the important uncertainty reduction during the conditioning procedure has been shown to be the consequence of a particularly efficient parameter identification, rather than an underestimation related to a suboptimal application of the ensemble Kalman method. We can assume that the improved performance of the ensemble Kalman method is mainly related to the massive parameter reduction which ultimately allowed to reduce the discrepancy with the limited ensemble size used, and hence to obtain more accurate updates. In addition, DeeSse plays an undeniable role in the significant observed decrease of the RMSE considering that its goal, ultimately, is to find compatible patterns in a TI that was used to simulate the reference fields of this synthetic case. The use of a unique TI to generate both the reference and the MPS simulations involved during the conditioning procedure is indeed a simplification made to evaluate the proposed approach in the situation where the conceptual geological model is consistent with the reality. For real cases, different TIs should be tested during the conditioning procedure.

In conclusion, although the results presented in this paper illustrate the performance of our methodology in a synthetic case with only two simulated facies, the same methodology can be applied to any number of simulated facies. The pyramid of training images has indeed been shown to work in the multicategorical case (Straubhaar et al., 2020). The ensemble Kalman method would then be applied to update the values of not only one initial coarse MPS simulation generated for each ensemble member, but several initial coarse MPS simulations generated from each facies indicator field. Since the number of gridblocks at the coarsest level is significantly reduced, and because only a subset of locations will be considered for the update ultimately, the increased number of parameters when considering a large number of facies could in practice still be much lower than the number of unknown in the categorical simulation.

Chapter 7

Conclusions

A synthetic groundwater flow model which mimics the hydraulic perturbation observed at the Andra's Underground Research Laboratory site in Meuse/Haute Marne, France, was developed with the aim of testing and developing inverse modelling strategies for the identification of groundwater parameter fields. Given the high-dimensionality of the numerical model developed for the real application, the computational efficiency of the inverse method to explore on the synthetic case and potentially apply on the real one was perhaps the most important selection criterion. For this main reason, ensemble methods for data assimilation have been rapidly considered since they have been efficiently applied for large-scale estimation problems in numerical weather prediction and reservoir simulation among other fields.

The work was focused on the application of an iterative ensemble smoother method using the Ensemble Smoother with Multiple Data Assimilations (ES-MDA) of Emerick and Reynolds (2012a) and the modified version of Levenberg-Marquardt Ensemble Randomized Maximum Likelihood (LM-EnRML) proposed by Chen and Oliver (2013). Both have been developed and especially used in the field of reservoir modelling (Evensen, 2018) and seemed suitable to address our groundwater flow inverse problem.

Various cases of variable complexity for the application of ensemble Kalman methods in general have been investigated. These cases considered situations with either continuous or discrete parameter fields. Although the problem considered was synthetic, its representation of a realistic situation highlighted limitations associated to the application

of those methods on real field cases with sparse observations. To answer some of those limitations, a new multi-resolution methodology combining multiple-point statistics (MPS) and data assimilation was proposed for the the categorical case.

In the following, we first summarize our main results and then propose a few directions for future research.

7.1 Summary of main contributions

7.1.1 Critical ensemble size in the multi-Gaussian case

The first case presented in this thesis (Chapter 3) addressed the performance of LM-EnRML and ES-MDA methods when conditioning an ensemble of multi-Gaussian realizations of a continuous field, namely the logarithm of hydraulic conductivities. Despite the favorable setting given the underlying multi-Gaussian assumption of ensemble Kalman methods, the tendency of these iterative methods to underestimate the uncertainty if not applied with care could be well illustrated in this example. Both LM-EnRML and ES-MDA performed poorly in terms of the accuracy in the mean and variance estimated from the final ensemble when the ensemble size used was insufficient. The comparison of the spread of the ensemble to the true error, here accessible via the distance between the ensemble mean and the known reference, allowed to determine a threshold ensemble size below which the performance of both methods was not reliable for this problem. With an ensemble size of 100, the error increased and the ensemble spread became too low using either LM-EnRML or ES-MDA. We note however that a good data match was achieved, thereby showing the importance of performing an uncertainty analysis when checking the reliability of the estimates obtained with ensemble Kalman methods. By simply doubling the ensemble size to 200 members, however, a clear improvement in the estimated uncertainty and ensemble mean was observed for both algorithms. A threshold ensemble size of 200 was ultimately considered for this case as it allows to achieve results with both accuracy and computational efficiency using either LM-EnRML or ES-MDA.

It also seems that LM-EnRML and ES-MDA do not converge towards the same solution. Compared to the estimates of LM-EnRML, the estimates of ES-MDA were shown to improve steadily with the ensemble size, especially with an error which keeps decreasing unlike

the estimates of LM-EnRML. This convergence of the error observed with ES-MDA is most likely due to its underlying variance-minimizing scheme which is known to be particularly efficient in Kalman filtering. Moreover, although LM-EnRML is designed to minimize an objective function which includes a data mismatch term, the conditioned parameter fields estimated by ES-MDA were shown to allow a better data fit than those corresponding to the most reduced data match achieved with LM-EnRML.

7.1.2 Clarification on the role of the normal-score transform

An interesting fact about the use of the normal-score transform in past applications of ensemble Kalman methods is that it can be employed for very different purposes depending on the problem addressed. In the multi-Gaussian case of log hydraulic conductivity ($\log K$) fields (Chapter 3), we applied a normal-score transform in order to improve locally the linearity of the forward model operator relating the parameter field to the transformed state observations. As shown by the increased accuracy of the estimates obtained and the more efficient reduction of the data mismatch, such transformation improved the performance of ES-MDA especially when assimilating head and flowrate data simultaneously. Considering that this normal-score transform did not allow LM-EnRML to converge, but was shown, similarly to ES-MDA, to benefit the ensemble Kalman filter (EnKF) update step in a previous study (Schoniger et al., 2012), we suppose that such local normal-score transformation can only be of use for ensemble methods which updating scheme implies a local linearization of the model around the parameter estimate, as Evensen mentioned about ES-MDA and EnKF in Evensen (2018).

In the non-multi-Gaussian cases explored in the rest of the thesis with ES-MDA (Chapters 4 and 5), the normal-score transform no longer aimed to trigger a local linearization of the forward model operator. Instead, it was used to mitigate the effects of no longer dealing with multi-Gaussian parameter or state variables by rendering them all marginally Gaussian. However, as mentioned by Zhou et al. (2011), the forward model operator relating the transformed parameters to the transformed states becomes even more nonlinear in this case than initially. Through the non-multi-Gaussian cases explored in the rest of the thesis, we observed experimentally the limitations of such local trans-

formations due to the increased nonlinearity of the forward model operator. In the bimodal case of hydraulic conductivity fields, although the assimilation of only transformed heads converged, the assimilation of transformed flowrates in addition to the transformed heads effectively prevented ES-MDA from converging. Since the same transformation of the states variables worked efficiently in the multi-Gaussian log K case, we attributed this non-convergence to the increased nonlinearity of the forward operator due to not only the nonlinear transformation of the parameters, but also due to the initially higher nonlinearity between the original bimodal parameters and predicted states. Another indication in that case that the nonlinearity can prevent the successful application of a normal-score transform is that transforming only the parameters variables effectively allowed ES-MDA to converge when assimilating both the untransformed data types. It therefore seems that the less nonlinear the dependence between the parameters to condition and the observations, the more likely such transformation will work. This observation is consistent with a previous study (Crestani et al., 2013) where such local normal-score transform was shown to lead to an alteration of the correlation structure between multi-Gaussian log hydraulic conductivities and concentrations, thereby effectively preventing the ensemble method from working properly.

Although we did not discuss it in the discrete cases where ES-MDA was used to update continuous underlying variables (Chapters 5 and 6), the resulting increased nonlinearity of the forward model operator also prevented the successful application of the normal-score transform to the state variables, except in the truncated Gaussian case which we assumed less nonlinear. Although the data assimilation still converged in that case, we showed that the transformation of state variables in that case resulted in less accurate estimates of ES-MDA than in the untransformed case. For this analysis based on the ensemble of categorical realizations, we had to introduce new performance criteria as a replacement of the ensemble spread and RMSE (Chapter 5). Hence the Shannon entropy was introduced as the measure of uncertainty of the ensemble distribution, and the Kullback-Leibler divergence as measure of the error when using the ensemble distribution to approximate the distribution given directly by the reference field. Note that both criteria were also evaluated for the discrete MPS case addressed in Chapter 6 and are shown in the Appendix.

7.1.3 A novel parameterization of multiple-point realizations

A novel multi-resolution parameterization based on the Gaussian pyramid technique implemented in DeeSse (Straubhaar et al., 2020) was proposed to condition categorical multiple-point statistics (MPS) realizations with ES-MDA (Chapter 6). The parameterization was designed so as to counteract expected limitations of ensemble Kalman methods in the non-Gaussian case. The parameters to update correspond to underlying continuous variables, at a coarse scale, which are directly linked to categorical patterns simulated, at a finer scale, via a multi-resolution MPS simulation performed by DeeSse. Unlike the conditioning of truncated Gaussian realization using underlying multi-Gaussian variables, the distribution of the proposed parameters in this case is only marginally Gaussian thanks to a normal-score transform applied before the start of the iterative conditioning procedure.

As we saw in the bimodal case of hydraulic conductivity fields, the normal-score transform does not solve the problem of the still non-Gaussian joint distributions in the application of ensemble Kalman methods. However, the transformed variables used as parameters in the proposed methodology have the advantage of still being underlying variables to the categorical field to regenerate. This precisely offers the flexibility that we did not have so far, namely to anticipate the limitations of the ensemble-based Kalman update. As we saw in the bimodal case of hydraulic conductivity fields, the direct application of the normal-score approach to the groundwater flow parameters does not leave room for improvement in a situation where the approach is not supported by a sufficient density of data points to compensate for the expected loss of connectivity.

Furthermore, the conditioning of the categorical fields using only a reduced number of continuous variables defined at a coarser scale can massively reduce the number of unknown for the computation of the update, as seen in our synthetic case. By using similar ensemble sizes as in the previous cases where the number of parameters was significantly larger, such parameter reduction should alleviate to some extent the issue of undersampling by allowing the ensemble to span more adequately the parameter subspace.

7.1.4 Constraining the reduced parameter space

Despite the non-Gaussian joint distributions of the underlying continuous variables, the proposed methodology for the discrete MPS case (Chapter 6) allows to reproduce consistent structures. Indeed, only a subset of non-adjacent variables at the coarse scale are updated, allowing the MPS algorithm DeeSse to simulate structures which are hard conditioned by the local updates performed at the coarse scale.

However, this larger role assigned to DeeSse at the scale of the update also causes an increase of the nonlinearity of the forward model operator relating the parameters to update to the data in the proposed algorithm. The main challenge at that point was to establish the necessary constraints in the forecast step in order to allow the convergence of ES-MDA towards an acceptable data fit. In particular, two following conditions were established and are thought to be the reasons of the success of our methodology, namely (1) the subset of underlying variables should be fixed and the same for all ensemble members throughout the iterations, and (2) the same seed which controls the reproducibility of the DeeSse simulations should be used for all ensemble members when regenerating the categorical field after the first update of the coarse-scale parameters. In this way, provided that the density of predefined parameters is sufficient, the robust convergence of the proposed methodology is ensured and fully controlled by the updated variables at the coarse scale. Through two synthetic examples, we illustrated the potential of the proposed approach for the retrieval of consistent geological structures relatively to the prior information in the course of the assimilation with ES-MDA using an ensemble size of 100.

Ultimately, the novelty of the proposed approach lies in the fact that the unavoidable loss of discrete structures during the linear update of ensemble Kalman methods is compensated by the incorporation, at each forecast step, of a non-multi-Gaussian information through the MPS simulation performed at the coarse scale of the update. Indeed, the conditional MPS simulation at the coarse scale consists in merging the non-multi-Gaussian prior information from the training image with the local information obtained from the data assimilation. Such an approach is therefore particularly relevant in situations where the observation points are very sparse, such as in our inverse problem. Besides, this integration of the MPS simulations in the forecast step was shown to directly benefit the final ensemble variance compared to the severe ensemble collapse observed with the normal-score approach in the bi-

modal case of log hydraulic conductivity fields using the same ensemble size of 100.

7.1.5 Benefit of the multi-resolution

In general, we found that the multi-resolution framework introduced in Chapter 6 contributed to the efficiency of the data assimilation at different levels. At the coarsest level, where the variables to update are located, the sampling error resulting from the approximation of the Kalman gain was reduced thanks to the reduced number of parameters. At the same time, the finer resolution of the reference categorical field from which the synthetic state observations were derived could still be accessible. Indeed, the defined parameterization of the multi-resolution MPS simulations, followed by the flow simulator, successfully links the coarse scale parameters to the finer scale data in a same forward model operator. Although our synthetic case assumed no model structural errors, the possibility of having access to finer scales at which available observations should be predicted is particularly important for real-world applications where model errors are inevitable. Indeed, using only an ensemble of low-resolution models for both the update and the prediction will negatively affect the parameter estimation by not taking into account the uncertainty associated to the potentially large model errors. In the end, it seems that the multi-resolution framework offers an interesting flexibility to improve the data assimilation framework for the inverse modelling of high-dimensional parameter fields.

7.2 Perspectives for future research

Based on the previous results and experiments, a few future research directions seem to be promising.

First, it would be very interesting to apply the proposed multi-resolution parameterization of the MPS simulations in other cases in order to better understand its advantages and limitations compared to other inverse methods. Examples of tests could include applications to a case with three or more facies, a sensitivity analysis on the number of conditioning data used in the time series, especially in terms of the final variability in the ensemble. In general, it would be relevant for the uncertainty analysis, to compare more rigorously the estimated ensemble distribution with a distribution representing the exact posterior distri-

bution. Given our first promising results in the discrete case, it would also be worth testing the proposed parameterization using directly a training image of a continuous property such as the log hydraulic conductivities. Such an application could allow to take into account the uncertainties associated to both the facies location and the hydraulic property value in the same data assimilation procedure.

Then, although this thesis focused on the use of iterative ensemble smoothers, and more particularly on the ES-MDA algorithm, it could be interesting to improve the efficiency of the proposed methodology in the discrete MPS case (Chapter 6) by combining the multi-resolution parameterization with another data assimilation method instead of ES-MDA. With ES-MDA, it required 16 iterations to achieve a good data match for our two synthetic examples, hence approximately twice the number of iterations required in the multi-Gaussian case. This increased number of iterations was also observed in the bimodal log K case and the truncated Gaussian cases, and probably results from the stronger nonlinearity of the forward model operator relating the parameters to estimate and the observations.

Therefore, to reduce the number of iterations, one direction could be to try to reduce the nonlinearity of the defined forward model operator. Interestingly, Chen and Oliver (2012) commented on the benefits of a sequential adjustment of parameters using the sequential iterative ensemble filter (seq-EnRML) over the iterative ensemble smoother version (batch-EnRML) based on the same randomized maximum likelihood method (RML). They noted that the sequential processing of the data allows a steady reduction of the nonlinearity of the problem. Therefore, it could be worth comparing in the context of the proposed methodology, the capability of seq-EnRML, or EnKF, and ES-MDA to address the nonlinearities for the same synthetic case. Another possible alternative could be to use a particle filter method such as the Sequential Importance Resampling Particle Filter (SIR-PF). In general, the large number of ensemble members required in order to adequately sample high-dimensional space often prevents the application of such methods in practice. However, the massive reduction of the number of parameters achieved in our examples could allow an appropriate application of these methods. Although the forecast step is similar to the one in ensemble Kalman methods, the update step in particle filter methods does not assume a multi-Gaussian distribution of the errors. As a result, the non-Gaussian ensemble distribution of the underlying coarse MPS simulations could be directly updated without performing a normal-score

transform step beforehand. The associated back normal-score transform performed at each iteration would also be removed from the forward model operator. In addition, if such particle filter methods were computationally feasible, the sensitivity to certain DeeSse parameters such as the distance threshold and maximum scan fraction may also reduce.

Another direction to increase the efficiency of ensemble methods in general is to reduce the duration of the forecast step which mainly includes the groundwater flow modelling. Ultimately, the computation of the update of ensemble Kalman methods only involves the solution of the groundwater flow equation at the observation locations. Therefore, the development of surrogate models in order to emulate the flow response at the observations points only would be of great relevance to make the iterative conditioning procedure faster, especially in real case applications.

If we now come back to the general question of the applicability of ensemble Kalman methods to real and complex 3D models such as the one used by ANDRA, a general challenge is to account for sparse dynamic data while remaining consistent with the prior geological knowledge. To bypass the problem of the loss of continuity of the discrete geological structures during the assimilation, a solution is to condition only the continuous properties populating each modelled facies. Although such a strategy will allow to characterize the heterogeneities at the scale of the facies, it will however not allow to take into account the uncertainty associated to the location of the facies boundaries. This is where the application of the proposed multi-resolution parameterization could be the most relevant. As seen in the discrete case, both our synthetic examples have illustrated the strong potential of the proposed multi-resolution parameterization in recovering consistent structures between widely separated points. This fact alone is particularly encouraging for real field applications where the monitoring locations are generally sparse. In addition, by using a continuous training image which integrates the different scales of heterogeneities, the proposed multi-resolution method could allow a parameter identification not only at the different scales of the geological structures, but also at the scale of the continuous property within the facies. However, a remaining challenge to answer seems to be when there are regions of the model with nearly nonexistent data, but which are still significantly contributing to the observed flow data. In such a situation, one of our synthetic examples showed that there is a high risk for inconsistency in the identified structures, most likely

because of unwarranted uncertainty reductions due to long range spurious correlations. To avoid those undesirable inconsistencies in the conditional fields obtained during the assimilation, an option to further investigate would be to sample the predefined update locations only in the areas which are informed by observations. In this manner, regions of the model which were originally affected by unreliable structures could then be simulated with DeeSse instead, without taking into account the unreliable updates in the multi-resolution MPS simulations. This idea is somewhat very similar to the way covariance inflation must in practice be applied only in areas experiencing a significant uncertainty reduction, hence only close to the observations, in order to avoid this time unwarranted increases in the ensemble variance. Hence if such an additional constraint on the current parameterization could be successful in the more general case of unevenly distributed observations, the proposed multi-resolution parameterization in real modelling situations could be even more applicable. In this manner, a systematic consistency of the conditional fields based on the locations of the observation points, as a result of either mainly the data assimilation, or DeeSse, would be ensured.

Appendix

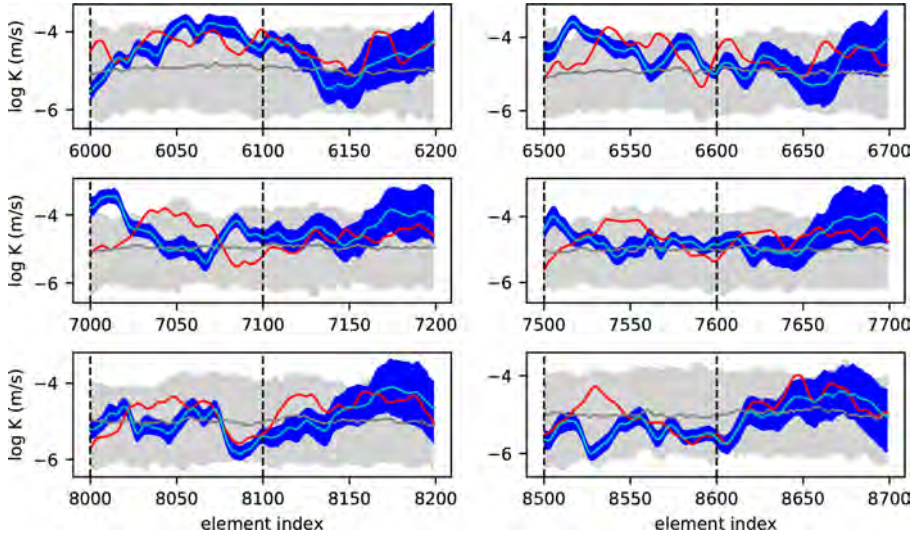


Figure 7.1: Distribution by gridblock of the ensemble of log hydraulic conductivities before (in grey) and after (in blue) ES-MDA ($\times 4$) using head data and 100 multi-Gaussian realizations. The regions between the percentiles P5 and P95 are represented. The red curve is the reference; the cyan curve is the calibrated ensemble mean. The vertical dashed lines denote the x -coordinate of the two sets of data points at respectively $x = 0$ and $x = 1000$ m.

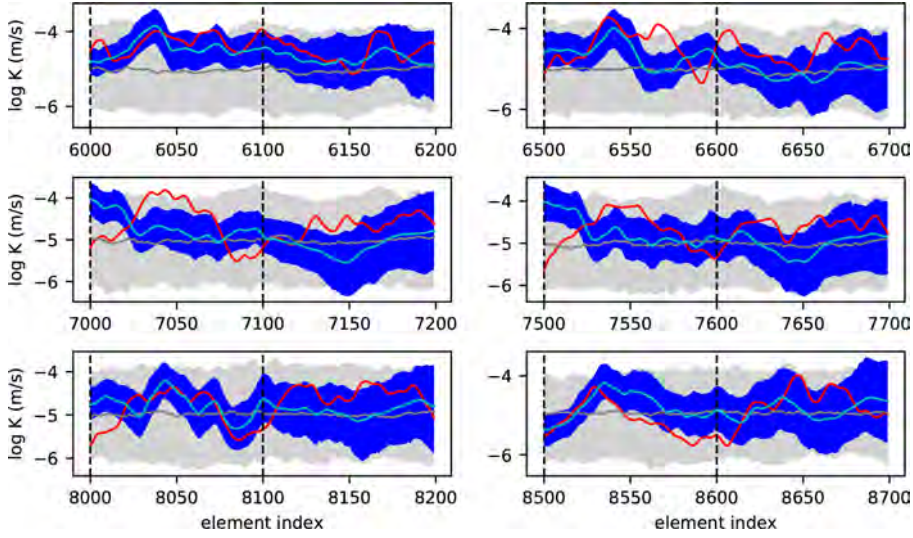


Figure 7.2: Distribution by gridblock of the ensemble of log hydraulic conductivities before (in grey) and after (in blue) ES-MDA ($\times 4$) using head data and 200 multi-Gaussian realizations. The regions between the percentiles P5 and P95 are represented. The red curve is the reference; the cyan curve is the calibrated ensemble mean. The vertical dashed lines denote the x -coordinate of the two sets of data points at respectively $x = 0$ and $x = 1000$ m.

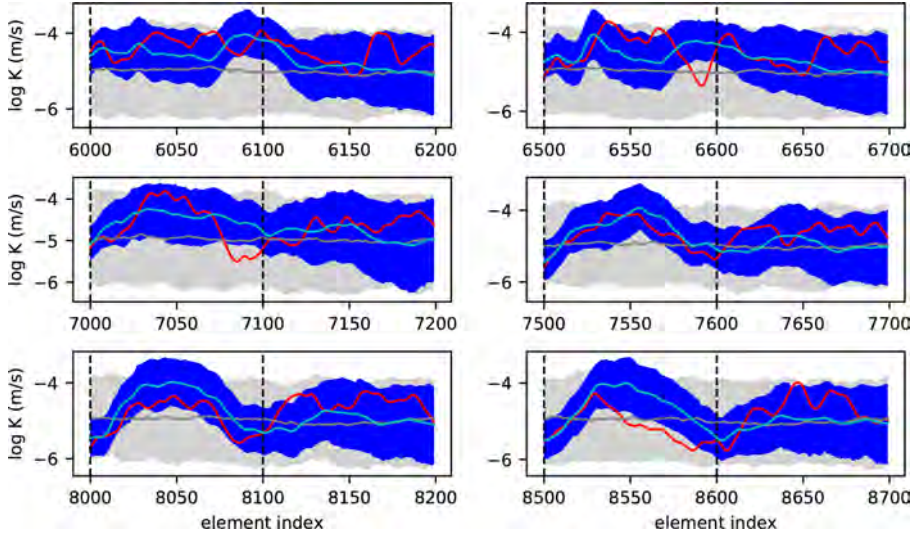


Figure 7.3: Distribution by gridblock of the ensemble of log hydraulic conductivities before (in grey) and after (in blue) ES-MDA ($\times 4$) using head data and 400 multi-Gaussian realizations. The regions between the percentiles P5 and P95 are represented. The red curve is the reference; the cyan curve is the calibrated ensemble mean. The vertical dashed lines denote the x -coordinate of the two sets of data points at respectively $x = 0$ and $x = 1000$ m.

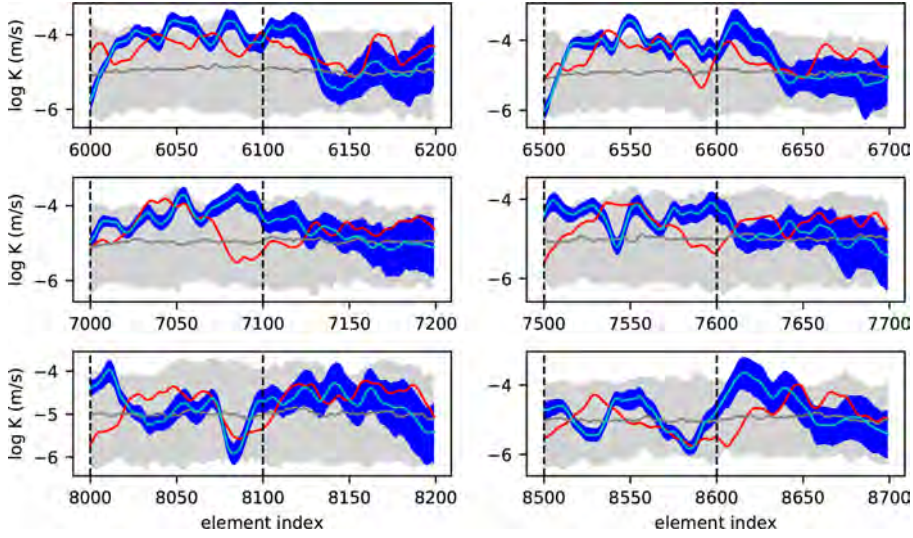


Figure 7.4: Distribution by gridblock of the ensemble of log hydraulic conductivities before (in grey) and after (in blue) LM-EnRML at iteration 8 using head data and 100 multi-Gaussian realizations. The regions between the percentiles P5 and P95 are represented. The red curve is the reference; the cyan curve is the calibrated ensemble mean. The vertical dashed lines denote the x -coordinate of the two sets of data points at respectively $x = 0$ and $x = 1000$ m.

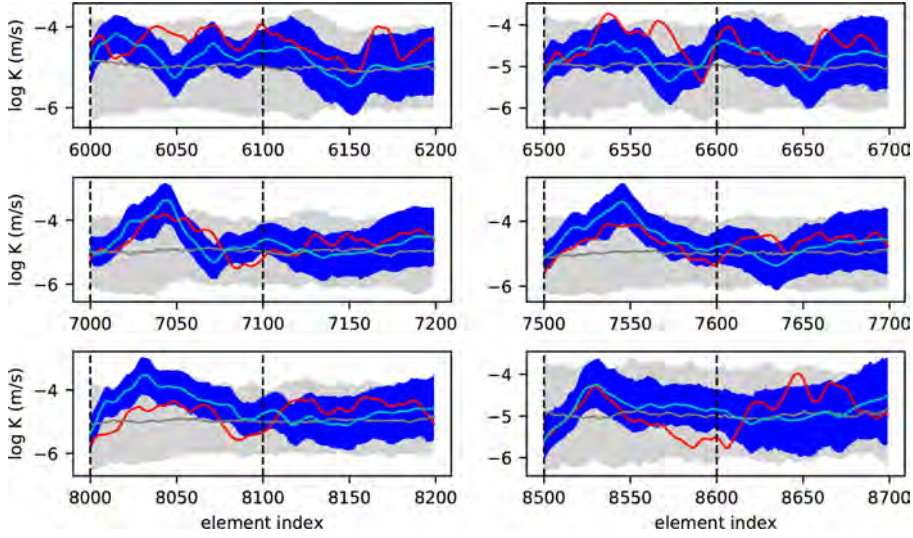


Figure 7.5: Distribution by gridblock of the ensemble of log hydraulic conductivities before (in grey) and after (in blue) LM-EnRML at iteration 8 using head data and 200 multi-Gaussian realizations. The regions between the percentiles P5 and P95 are represented. The red curve is the reference; the cyan curve is the calibrated ensemble mean. The vertical dashed lines denote the x -coordinate of the two sets of data points at respectively $x = 0$ and $x = 1000$ m.

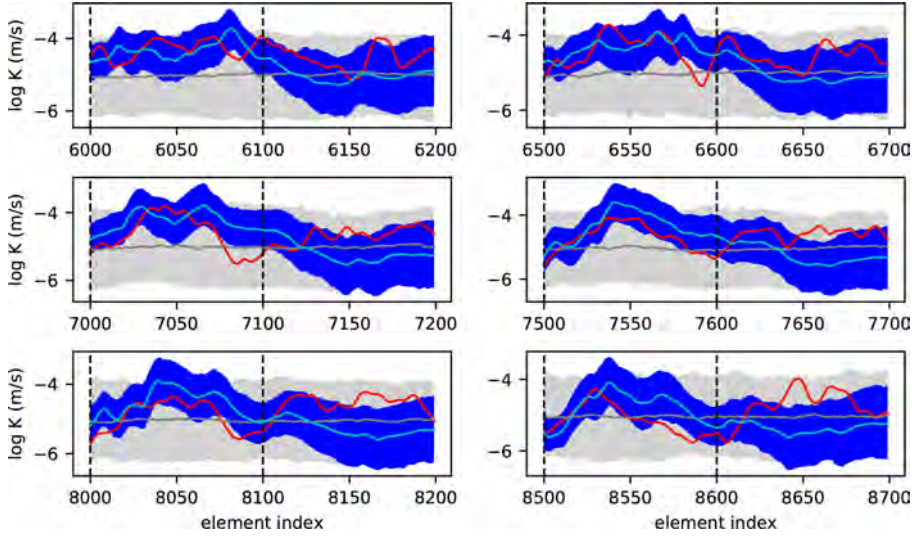


Figure 7.6: Distribution by gridblock of the ensemble of log hydraulic conductivities before (in grey) and after (in blue) LM-EnRML at iteration 8 using head data and 400 multi-Gaussian realizations. The regions between the percentiles P5 and P95 are represented. The red curve is the reference; the cyan curve is the calibrated ensemble mean. The vertical dashed lines denote the x -coordinate of the two sets of data points at respectively $x = 0$ and $x = 1000$ m.

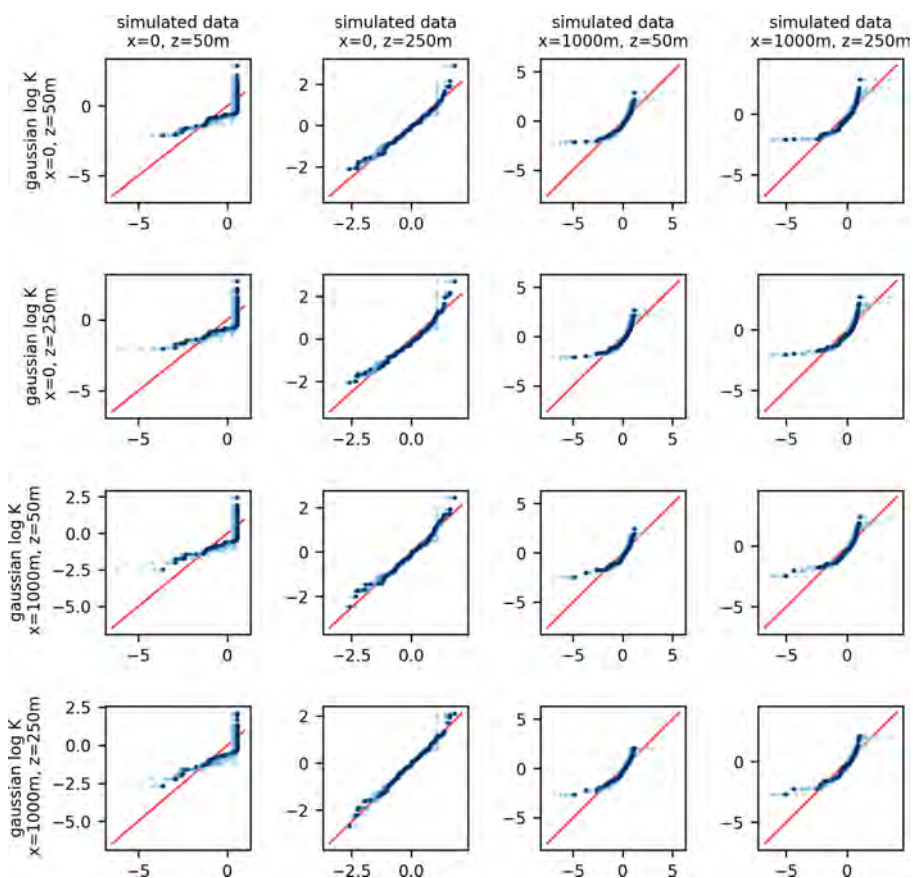


Figure 7.7: Q-Q plots of the multi-Gaussian log K values against the predicted state observations at several observation locations.

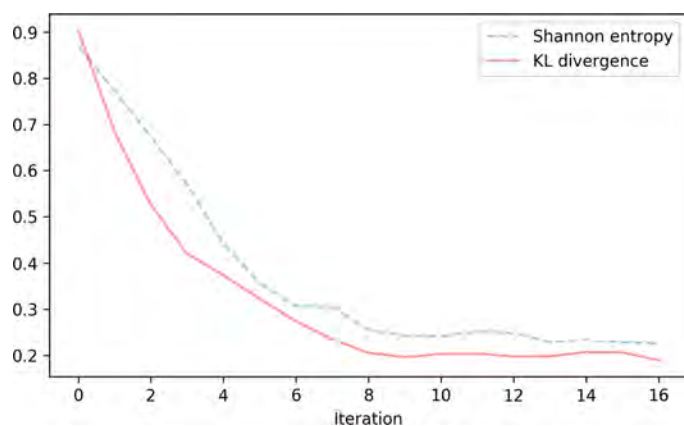


Figure 7.8: Evolution of the KL divergence and the Shannon entropy before and during the conditioning of the categorical MPS realizations in the updated region (up to 1600 m from the seepage boundary) for the ES-MDA update with the synthetic data of reference A.

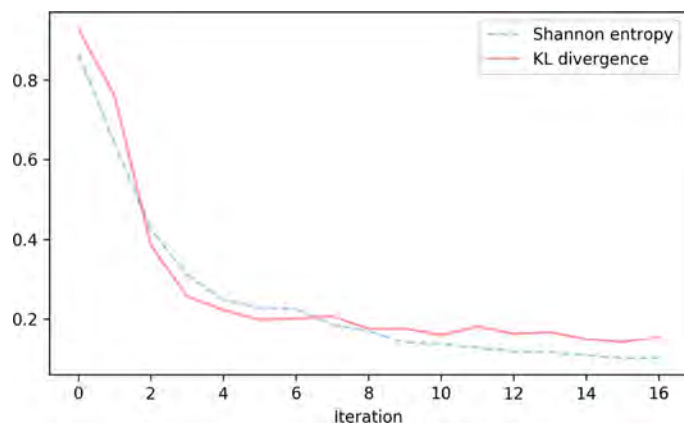


Figure 7.9: Evolution of the KL divergence and the Shannon entropy before and during the conditioning of the categorical MPS realizations in the updated region (up to 1500 m from the seepage boundary) for the ES-MDA update with the synthetic data of reference B.

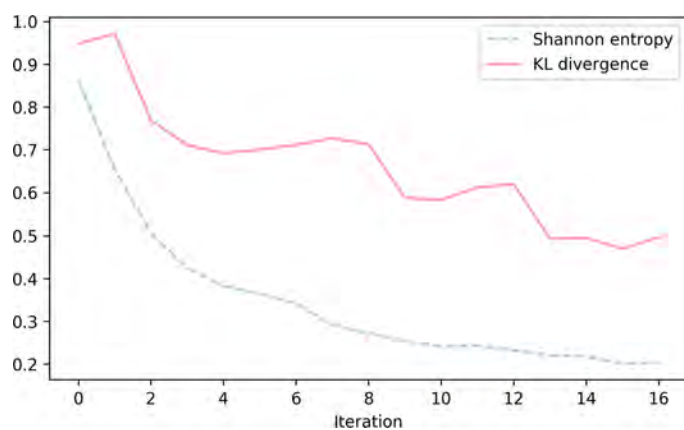


Figure 7.10: Evolution of the KL divergence and the Shannon entropy before and during the conditioning of the categorical MPS realizations in the updated region (up to 2600 m from the seepage boundary) for the ES-MDA update with the synthetic data of reference B.

Bibliography

A. Abadpour, M. Adejare, T. Chugunova, H. Mathieu, and N. Haller. Integrated Geo-Modeling and Ensemble History Matching of Complex Fractured Carbonate and Deep Offshore Turbidite Fields, Generation of Several Geologically Coherent Solutions Using Ensemble Methods. Society of Petroleum Engineers (SPE), 2018. doi: 10.2118/193028-ms.

J. L. Anderson. An Ensemble Adjustment Kalman Filter for Data Assimilation. *Monthly Weather Review*, 129(12):2894–2903, 2003.

J. L. Anderson. Ensemble Kalman Filters for Large Geophysical Applications: Adaptive algorithms for improving ensemble filter performance. *IEEE Control Systems*, 29(3):66–82, 2009. ISSN 1066033X. doi: 10.1109/MCS.2009.932222.

J. L. Anderson. Localization and Sampling Error Correction in Ensemble Kalman Filter Data Assimilation. *Monthly Weather Review*, 140(7): 2359–2371, 2012. doi: 10.1175/mwr-d-11-00013.1.

M. Armstrong, A. G. Galli, G. Le Loc’h, F. Geffroy, and R. Eschard. *Plurigaussian Simulations in Geosciences*. Springer Berlin Heidelberg, 2003. doi: 10.1007/978-3-662-12718-6.

A. Astrakova and D. S. Oliver. Conditioning Truncated Pluri-Gaussian Models to Facies Observations in Ensemble-Kalman-Based Data Assimilation. *Mathematical Geosciences*, 47(3):345–367, 2015. ISSN 18748953. doi: 10.1007/s11004-014-9532-3.

D. Béal, P. Brasseur, J. M. Brankart, Y. Ourmières, and J. Verron. Characterization of mixing errors in a coupled physical biogeochemical model of the North Atlantic: Implications for nonlinear estimation using Gaussian anamorphosis. *Ocean Science*, 6(1):247–262, 2010. ISSN 18120792. doi: 10.5194/os-6-247-2010.

H. Benabderrahmane, J. Kerrou, L. Tacher, G. Deman, and P. Perrochet. Modelling of Predictive Hydraulic Impacts of a Potential Radioactive Waste Geological Repository on the Meuse/Haute-Marne Multi-layered Aquifer System (France). *Journal of Applied Mathematics and Physics*, 02(12):1085–1090, 2014. doi: 10.4236/jamp.2014.212125.

L. Bertino, G. Evensen, and H. Wackernagel. Sequential Data Assimilation Techniques in Oceanography. *International Statistical Review*, 71(2):223–241, 2003. doi: 10.1111/j.1751-5823.2003.tb00194.x.

M. Bocquet and P. Sakov. An iterative ensemble Kalman smoother. *Quarterly Journal of the Royal Meteorological Society*, 140(682):1521–1535, 2014. ISSN 1477870X. doi: 10.1002/qj.2236.

M. Bocquet, C. Pires, L. Wu, and C. A. Pires. Beyond Gaussian Statistical Modeling in Geophysical Data Assimilation. *Monthly Weather Review*, 138:2997–3023, 2010. doi: 10.1175/2010MWR3164.1.

A. Bourgeat, M. Kern, S. Schumacher, and J. Talandier. The COUPLEX test cases: Nuclear waste disposal simulation. *Computational Geosciences*, 8:83–98, 2004.

G. Burgers, P. Jan Van Leeuwen, and G. Evensen. Analysis Scheme in the Ensemble Kalman Filter. *Monthly Weather Review*, 126:1719–1724, 1998.

P. Burt and E. Adelson. The laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, 31(4):532–540, 1983. doi: 10.1109/TCOM.1983.1095851.

Z. Cao, L. Li, and K. Chen. Bridging iterative Ensemble Smoother and multiple-point geostatistics for better flow and transport modeling. *Journal of Hydrology*, 565(August):411–421, 2018. doi: 10.1016/j.jhydrol.2018.08.023.

J. Carrera, A. Alcolea, A. Medina, J. Hidalgo, and L. J. Slooten. Inverse problem in hydrogeology. *Hydrogeology Journal*, 13(1):206–222, 2005. doi: 10.1007/s10040-004-0404-7.

Y. Chen and D. S. Oliver. Ensemble Randomized Maximum Likelihood Method as an Iterative Ensemble Smoother. *Mathematical Geosciences*, 44(1):1–26, 2012. ISSN 18748961. doi: 10.1007/s11004-011-9376-z.

- Y. Chen and D. S. Oliver. Levenberg-Marquardt forms of the iterative ensemble smoother for efficient history matching and uncertainty quantification. *Computational Geosciences*, 17(4):689–703, 2013. ISSN 14200597. doi: 10.1007/s10596-013-9351-5.
- Y. Chen and D. S. Oliver. Localization and regularization for iterative ensemble smoothers. *Computational Geosciences*, 21(1):13–30, 2017. doi: 10.1007/s10596-016-9599-7.
- Y. Chen and D. Zhang. Data assimilation for transient flow in geologic formations via ensemble Kalman filter. *Advances in Water Resources*, 29(8):1107–1122, 2006. ISSN 03091708. doi: 10.1016/j.advwatres.2005.09.007.
- Y. Chen, D. S. Oliver, and D. Zhang. Data assimilation for nonlinear problems by ensemble Kalman filter with reparameterization. *Journal of Petroleum Science and Engineering*, 66(1-2):1–14, 2009. doi: 10.1016/j.petrol.2008.12.002.
- E. Crestani, M. Camporese, D. Baú, and P. Salandin. Ensemble Kalman filter versus ensemble smoother for assessing hydraulic conductivity via tracer test data assimilation. *Hydrology and Earth System Sciences*, 17(4):1517–1531, 2013. ISSN 10275606. doi: 10.5194/hess-17-1517-2013.
- G. De Marsily, J.-P. Delhomme, F. Delay, and Buoro A. 40 years of inverse problems in hydrogeology. *Comptes Rendus de l'Académie des Sciences - Series IIA - Earth and Planetary Science*, 329(2):73–87, 1999.
- J. Delay, H. Rebours, A. Vinsot, and P. Robin. Scientific investigation in deep wells for nuclear waste disposal studies at the Meuse/Haute Marne underground research laboratory, Northeastern France. *Physics and Chemistry of the Earth*, 32(1-7):42–57, 2007. ISSN 14747065. doi: 10.1016/j.pce.2005.11.004.
- G. Deman, K. Konakli, B. Sudret, J. Kerrou, P. Perrochet, and H. Benabderrahmane. Using sparse polynomial chaos expansions for the global sensitivity analysis of groundwater lifetime expectancy in a multi-layered hydrogeological model. *Reliability Engineering and System Safety*, 147:156–169, 2015. doi: 10.1016/j.res.2015.11.005.
- A. Emerick and A. Reynolds. Combining sensitivities and prior information for covariance localization in the ensemble Kalman filter for petroleum reservoir applications. *Computational Geosciences*, 15(2):251–269, 2011. doi: 10.1007/s10596-010-9198-y.

A. A. Emerick. Analysis of the performance of ensemble-based assimilation of production and seismic data. *Journal of Petroleum Science and Engineering*, 139:219–239, 2016. doi: 10.1016/j.petrol.2016.01.029.

A. A. Emerick and A. C. Reynolds. Ensemble smoother with multiple data assimilation. *Computers and Geosciences*, 55:3–15, 2012a. doi: 10.1016/j.cageo.2012.03.011.

A. A. Emerick and A. C. Reynolds. History matching time-lapse seismic data using the ensemble Kalman filter with multiple data assimilations. *Computational Geosciences*, 16(3):639–659, 2012b. doi: 10.1007/s10596-012-9275-5.

G. Evensen. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research*, 99(C5):10143, 1994. ISSN 0148-0227. doi: 10.1029/94jc00572.

G. Evensen. The Ensemble Kalman Filter: Theoretical formulation and practical implementation. *Ocean Dynamics*, 53(4):343–367, 2003. ISSN 16167341. doi: 10.1007/s10236-003-0036-9.

G. Evensen. Sampling strategies and square root analysis schemes for the EnKF. *Ocean Dynamics*, 54(6):539–560, 2004. doi: 10.1007/s10236-004-0099-2.

G. Evensen. *Data Assimilation*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009a. doi: 10.1007/978-3-642-03711-5.

G. Evensen. The Ensemble Kalman Filter for Combined State and Parameter Estimation: Monte Carlo techniques for data assimilation in large systems. *IEEE Control Systems*, 29(3):83–104, 2009b. ISSN 1066033X. doi: 10.1109/MCS.2009.932223.

G. Evensen. Analysis of iterative ensemble smoothers for solving inverse problems. *Computational Geosciences*, 22(3):885–908, 2018. ISSN 15731499. doi: 10.1007/s10596-018-9731-y.

G. Evensen and P. J. van Leeuwen. An Ensemble Kalman Smoother for Nonlinear Dynamics. *Monthly Weather Review*, 128(6):1852–1867, 2002. doi: 10.1175/1520-0493(2000)128<1852:aeksfn>2.0.co;2.

Fabien Cornaton. Ground Water (GW) A 3-D Ground Water and Surface Water Flow, Mass Transport and Heat Transfer Finite Element Simulator. Technical report, University of Neuchâtel, 2014.

- C. Gardet, M. Le Ravalec, and E. Gloaguen. Multiscale Parameterization of Petrophysical Properties for Efficient History-Matching. *Mathematical Geosciences*, 46(3):315–336, 2014. ISSN 18748953. doi: 10.1007/s11004-013-9480-3.
- G. Gaspari and S. E. Cohn. Construction of correlation functions in two and three dimensions. *Quarterly Journal of the Royal Meteorological Society*, 125(554):723–757, 1999. ISSN 00359009. doi: 10.1256/smsqj.55416.
- S. Gillijns, O. Mendoza, J. Chandrasekar, B. De Moor, D. Bernstein, and A. Ridley. What is the ensemble Kalman filter and how well does it work? pages 4448–4453. Institute of Electrical and Electronics Engineers (IEEE), 2006. doi: 10.1109/acc.2006.1657419.
- J. J. Gómez-Hernández and X. H. Wen. To be or not to be multi-Gaussian? A reflection on stochastic hydrogeology. *Advances in Water Resources*, 21(1):47–61, 1998. ISSN 03091708. doi: 10.1016/S0309-1708(96)00031-0.
- Y. Gu and D. S. Oliver. The Ensemble Kalman Filter for Continuous Updating of Reservoir Simulation Models. *Journal of Energy Resources Technology*, 128(1):79, 2006. ISSN 01950738. doi: 10.1115/1.2134735.
- Y. Gu and D. S. Oliver. An Iterative Ensemble Kalman Filter for Multiphase Fluid Flow Data Assimilation. *SPE Journal*, 12(04):438–446, 2007. doi: 10.2118/108438-pa.
- F. Gustafsson. Particle filter theory and practice with positioning applications. *IEEE Aerospace and Electronic Systems Magazine*, 25(7): 53–82, 2010. doi: 10.1109/MAES.2010.5546308.
- T. M. Hamill. Ensemble-based atmospheric data assimilation. In *Predictability of Weather and Climate*, volume 9780521848824, pages 124–156. Cambridge University Press, 2006. doi: 10.1017/CBO9780521848824.007.
- T. M. Hamill and J. S. Whitaker. Accounting for the Error due to Unresolved Scales in Ensemble Data Assimilation: A Comparison of Different Approaches. Technical report, 2005.
- T. M. Hamill, J. S. Whitaker, and C. Snyder. Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Monthly Weather Review*, 129(11):2776–2790, 2001. ISSN 00270644. doi: 10.1175/1520-0493(2001)129<2776:DDFOBE>2.0.CO;2.

L. Heidari, V. Gervais, M. L. Ravalec, and H. Wackernagel. History matching of petroleum reservoir models by the Ensemble Kalman Filter and parameterization methods. *Computers and Geosciences*, 55:84–95, 2013. doi: 10.1016/j.cageo.2012.06.006.

H. J. Hendricks Franssen and W. Kinzelbach. 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(9), 2008. doi: 10.1029/2007WR006505.

H. J. Hendricks Franssen, A. Alcolea, M. Riva, M. Bakr, N. van der Wiel, F. Stauffer, and A. Guadagnini. A comparison of seven methods for the inverse modelling of groundwater flow. Application to the characterisation of well catchments. *Advances in Water Resources*, 32(6): 851–872, 2009. doi: 10.1016/j.advwatres.2009.02.011.

P. L. Houtekamer and H. L. Mitchell. Data assimilation using an ensemble Kalman filter technique. *Monthly Weather Review*, 126(3):796–811, 1998. ISSN 00270644. doi: 10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2.

P. L. Houtekamer and H. L. Mitchell. Ensemble Kalman filtering. *Quarterly Journal of the Royal Meteorological Society*, 131(613):3269–3289, 2006. doi: 10.1256/qj.05.135.

P. L. Houtekamer and F. Zhang. Review of the Ensemble Kalman Filter for Atmospheric Data Assimilation. *Monthly Weather Review*, 144(12): 4489–4532, 2016. doi: 10.1175/mwr-d-15-0440.1.

L. Y. Hu, Y. Zhao, Y. Liu, C. Scheepens, and A. Bouchard. Updating multipoint simulations using the ensemble Kalman filter. *Computers and Geosciences*, 51:7–15, 2013. ISSN 00983004. doi: 10.1016/j.cageo.2012.08.020.

B. Jafarpour and M. Khodabakhshi. A Probability Conditioning Method (PCM) for Nonlinear Flow Data Integration into Multipoint Statistical Facies Simulation. *Mathematical Geosciences*, 43(2):133–164, 2011. ISSN 18748961. doi: 10.1007/s11004-011-9316-y.

B. Jafarpour and D. B. McLaughlin. History matching with an ensemble Kalman filter and discrete cosine parameterization. *Computational Geosciences*, 12(2):227–244, 2008. ISSN 14200597. doi: 10.1007/s10596-008-9080-3.

R. E. Kalman. A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, 82(1):35, 1960. ISSN 00219223. doi: 10.1115/1.3662552.

J. Kerrou, P. Renard, H. J. Hendricks Franssen, and I. Lunati. Issues in characterizing heterogeneity and connectivity in non-multiGaussian media. *Advances in Water Resources*, (1):147–159, 2008. doi: 10.1016/j.advwatres.2007.07.002.

J. Kerrou, G. Deman, L. Tacher, H. Benabderrahmane, and P. Perrochet. Numerical and polynomial modelling to assess environmental and hydraulic impacts of the future geological radwaste repository in Meuse site (France). *Environmental Modelling and Software*, 97:157–170, 2017. doi: 10.1016/j.envsoft.2017.07.018.

S. Kullback and R. A. Leibler. On information and sufficiency. *The annals of mathematical statistics*, 22(1):79–86, 1951.

D. H. Le, A. A. Emerick, and A. C. Reynolds. An Adaptive Ensemble Smoother With Multiple Data Assimilation for Assisted History Matching. *SPE Journal*, 21(06):2195–2207, 2016. ISSN 1086-055X. doi: 10.2118/173214-pa.

M. Le Ravalec. *Inverse Stochastic Modeling of Flow in Porous Media: Applications to Reservoir Characterization*. Ophrys edition, 2005.

L. Li, H. Zhou, H. J. Hendricks Franssen, and J. J. Gómez-Hernández. Groundwater flow inverse modeling in non-MultiGaussian media: Performance assessment of the normal-score Ensemble Kalman Filter. *Hydrology and Earth System Sciences*, 16(2):573–590, 2012. ISSN 10275606. doi: 10.5194/hess-16-573-2012.

L. Li, S. Srinivasan, H. Zhou, and J. J. Gómez-Hernández. Simultaneous Estimation of Geologic and Reservoir State Variables Within an Ensemble-Based Multiple-Point Statistic Framework. *Mathematical Geosciences*, 46(5):597–623, 2014. ISSN 18748953. doi: 10.1007/s11004-013-9504-z.

L. Li, L. Stetler, Z. Cao, and A. Davis. An iterative normal-score ensemble smoother for dealing with non-Gaussianity in data assimilation. *Journal of Hydrology*, 567:759–766, 2018. doi: 10.1016/j.jhydrol.2018.01.038.

Y. Linard, A. Vinsot, B. Vincent, J. Delay, S. Wechner, R. De La Vaisière, E. Scholz, B. Garry, M. Lundy, M. Cruchaudet, S. Dewonck, and G. Vigneron. Water flow in the Oxfordian and Dogger limestone around the Meuse/Haute-Marne Underground Research Laboratory. *Physics and Chemistry of the Earth*, 36(17-18):1450–1468, 2011. ISSN 14747065. doi: 10.1016/j.pce.2011.07.012.

N. Linde, P. Renard, T. Mukerji, and J. Caers. Geological realism in hydrogeological and geophysical inverse modeling: A review. *Advances in Water Resources*, 86:86–101, 2015. doi: 10.1016/j.advwatres.2015.09.019.

N. Liu and D. S. Oliver. Ensemble Kalman filter for automatic history matching of geologic facies. *Journal of Petroleum Science and Engineering*, 47(3-4):147–161, 2005. ISSN 09204105. doi: 10.1016/j.petrol.2005.03.006.

T. Lochbühler, G. Pirot, J. Straubhaar, and N. Linde. Conditioning of Multiple-Point Statistics Facies Simulations to Tomographic Images. *Mathematical Geosciences*, 46(5):625–645, 2014. doi: 10.1007/s11004-013-9484-z.

W. Ma and B. Jafarpour. Pilot points method for conditioning multiple-point statistical facies simulation on flow data. *Advances in Water Resources*, 115:219–233, 2018. doi: 10.1016/j.advwatres.2018.01.021.

W. Ma and B. Jafarpour. Integration of soft data into multiple-point statistical simulation : re-assessing the probability conditioning method for facies model calibration. 2019.

G. Mariethoz and J. Caers. *Multiple-point Geostatistics: Stochastic Modeling with Training Images, First edition*. 2015.

G. Mariethoz, P. Renard, and J. Straubhaar. The direct sampling method to perform multiple-point geostatistical simulations. *Water Resources Research*, 46(11):1–14, 2010. ISSN 00431397. doi: 10.1029/2008WR007621.

G. Mariethoz, M. F. McCabe, and P. Renard. Spatiotemporal reconstruction of gaps in multivariate fields using the direct sampling approach. *Water Resources Research*, 48(10), 2012. ISSN 00431397. doi: 10.1029/2012WR012115.

- E. Meerschman, G. Pirot, G. Mariethoz, J. Straubhaar, M. Van Meirvenne, and P. Renard. A practical guide to performing multiple-point statistical simulations with the Direct Sampling algorithm. *Computers and Geosciences*, 52:307–324, 2013. doi: 10.1016/j.cageo.2012.09.019.
- H. Moradkhani, S. Sorooshian, H. V. Gupta, and P. R. Houser. Dual state-parameter estimation of hydrological models using ensemble Kalman filter. *Advances in Water Resources*, 28(2):135–147, 2005. doi: 10.1016/j.advwatres.2004.09.002.
- D. S. Oliver and Y. Chen. Recent progress on reservoir history matching: A review. *Computational Geosciences*, 15(1):185–221, 2011. doi: 10.1007/s10596-010-9194-2.
- D. S. Oliver, N. Liu, and A. C. Reynolds. *Inverse theory for petroleum reservoir characterization and history matching*. Cambridge University Press, 2008. ISBN 9780521881517.
- R. E. Petrie and S. L. Dance. Ensemble-based data assimilation and the localisation problem. *Weather*, 65:65–69, 2010.
- P. Renard and D. Allard. Connectivity metrics for subsurface flow and transport. *Advances in Water Resources*, 51:168–196, 2013. doi: 10.1016/j.advwatres.2011.12.001.
- A. Schoniger, W. Nowak, and H. J. Hendricks Franssen. Parameter estimation by ensemble Kalman filters with transformed data: Approach and application to hydraulic tomography. *Water Resources Research*, 48(4):1–18, 2012. ISSN 00431397. doi: 10.1029/2011WR010462.
- E. Simon and L. Bertino. Application of the Gaussian anamorphosis to assimilation in a 3-D coupled physical-ecosystem model of the North Atlantic with the EnKF: a twin experiment. *Ocean Science*, 5:495–510, 2009. doi: 10.5194/os-5-495-2009.
- J. Straubhaar. User ’s Guide May 2017 DeeSse Software patented by the University of Neuchâtel. 2017.
- J. Straubhaar, P. Renard, and T. Chugunova. Multiple-point statistics using multi-resolution images. *Stochastic Environmental Research and Risk Assessment*, 2020. doi: 10.1007/s00477-020-01770-8.
- S. Strebelle. Conditional simulation of complex geological structures using multiple-point statistics. *Mathematical geology*, 34(1):1–21, 2002. doi: 10.1109/CEC.2011.5949612.

A. Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation*. Society for Industrial & Applied Mathematics, Philadelphia, 2005.

P. J. van Leeuwen. Comment on "Data assimilation using an ensemble Kalman filter technique". *Monthly Weather Review*, 127:1374–1377, 1998. doi: 10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2.

P. J. van Leeuwen and G. Evensen. Data Assimilation and Inverse Methods in Terms of a Probabilistic Formulation. *Monthly Weather Review*, 124(12):2898–2913, 1996. doi: 10.1175/1520-0493(1996)124<2898:daaimi>2.0.co;2.

X. H. Wen and W. Chen. Real-time reservoir model updating using ensemble Kalman filter. *SPE Reservoir Simulation Symposium, Houston, TX, USA*, pages 1–14, 2005.

T. Zahner, T. Lochbühler, G. Mariethoz, and N. Linde. Image synthesis with graph cuts: A fast model proposal mechanism in probabilistic inversion. *Geophysical Journal International*, 204(2):1179–1190, 2016. ISSN 1365246X. doi: 10.1093/gji/ggv517.

Y. Zhao, A. C. Reynolds, and G. Li. Generating Facies Maps by Assimilating Production Data and Seismic Data With the Ensemble Kalman Filter. *Paper 113990 presented at the SPE/DOE Symposium on Improved Oil Recovery, Tulsa, Oklahoma, 20–23 April*, 2008. doi: 10.2118/113990-ms.

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

H. Zhou, J. J. Gómez-Hernández, and L. Li. A pattern-search-based inverse method. *Water Resources Research*, 48(3), 2012. ISSN 00431397. doi: 10.1029/2011WR011195.

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

B. Zinn and C. F. Harvey. 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(3):1–19, 2003. ISSN 00431397. doi: 10.1029/2001WR001146.

F. Zovi, M. Camporese, H.-J. Hendricks Franssen, J. A. Huisman, and P. Salandin. Identification of high-permeability subsurface structures with multiple point geostatistics and normal score ensemble Kalman filter. *Journal of Hydrology*, 548:208–224, 2017. doi: 10.1016/j.jhydrol.2017.02.056.