Data Assimilation under Geological Constraints

Proefschrift

ter verkrijging van de graad van doctor
aan de technische Universiteit Delft,
op gezag van de Rector Magnificus prof. ir. K.C.A.M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op
donderdag **18 december 2014** om *10.00 uur*

door

Bogdan Marius SEBACHER

Master of Science in Mathematics
geboren te Bucuresti, Romania.

Dit proefschrift is goedgekeurd door de promotor:
Prof. dr. ir. A. W. Heemink

Samenstelling promotiecomissie:

Rector Magnificus	voorzitter
Prof. dr. ir. A. W. Heemink,	Technische Universiteit Delft, promotor
Prof. dr. ir. J. D. Jansen,	Technische Universiteit Delft
Prof. dr. S. M. Luthi,	Technische Universiteit Delft
Prof. dr. F .H. J. Redig,	Technische Universiteit Delft
Prof. dr. R. Trandafir,	Universitatea Tehnica de Constructii Bucuresti, Romania
Associate Prof. dr. I. Hoteit,	KAUST, Saudi Arabia
Dr. R. G. Hanea,	Statoil, Norway
Prof. dr. ir. C. Vuik,	Technische Universiteit Delft, reservelid

ISBN: 978-94-6186-405-5
Typeset by the author with the LaTeXDocumentation System.
On the cover: drawing by Maria Sebacher.
Printed in The Netherlands by Gildeprint, Enschede.

To my family

Contents

Acknowledgements

This thesis is the outcome of the joint effort of many people that contributed either by playing an important role and taking active part in it or by generously sharing their knowledge and sparkling ideas or by simply inspiring me during our conversations.

Firstly, I am at a loss for words to express my gratitude to my supervisor Arnold Heemink for the help, guidance and support he offered me during my PhD research. Without his assistance and active participation in every step of the way, this thesis may have never been completed. Thank you very much for your encouragement and understanding over these past four years.

Secondly, I am infinitely grateful to my "angel", Remus Hanea. He was close to me during important moments, offering me opportunities, finding resources and above all investing time, smoothing my way through life. He has fought for my ideas, like I have never did. For all these and for many, many other wonderful things thank you, my dear friend.

My thoughts goes now to a person whose support my project relies on. His name Cris te Stroet. Now, that I finished, I would like to thank him for trusting me, for giving me the chance to have the journey of my life.

Working on my PhD project was a great experience thanks to the people surrounding me; for this reason, I would like to thank all my friends in the department of Applied Mathematics, TNO and Statoil for the constructive discussions, friendly atmosphere and all their help. I would like to thank, especially to Alin Chitu for his friendship, help and support offered during my Phd study. In addition, this acknowledgement would not be complete if I did not mention two people that unconditionally helped me: Evelyn Sharabi and Dorothee Engering. Without their help and support I would have remained

tangled in all the administrative challenges.

I am also glad to have worked with a wonderful friend Andreas Stordal for the last two years of my research. Thank you for all great moments that we have spent together and my only hope is that was only "the beginning of a beautiful friendship".

Last, but certainly not the least, I would like to thank the two girls that completed my life: my wife Florentina and my daughter Maria. I am grateful to them for all the support, care, patience, endless love and belief in me. I am again at a loss for words to express my love for them.

Bogdan Sebacher,
Delft, 2014.

Chapter 1

Introduction

A reservoir represents a natural accumulation of hydrocarbons captive within lithological structures. The hydrocarbons are formed by decomposition of the organic matter that accumulates from the deposition of marine microorganisms and vegetation in ancient sea basins. The transformation in buried sediments of organic matter to oil and gas takes millions of years depending on the condition of temperature and pressure in the subsurface. At the surface condition, the hydrocarbons can be found in gaseous state (natural gas), liquid state (oil) or even solid state (natural bitumen).

1.1 Petroleum Systems

An efficient recovery of the hydrocarbons trapped within a reservoir requires a prior definition of a conceptual model that enables spatial visualization of the reservoir. Initially, the conceptual model is built based on geological information collected during the reservoir exploration phase.

The first geologic information refers to the external geometry of the reservoir which is defined by lithological barriers that block the hydrocarbons displacement. Usually, these barriers consist of a low permeable or even impermeable rocks (seals, cap rocks) that stop the hydrocarbons movement. The buoyancy force produced by the difference in density between water and hydrocarbons generates the hydrocarbon migration. Within a reservoir, the hydrocarbon migration can propagate in two directions: the *primary migration* e.g. the movement within rock source (the rock formed after sedimentation of the organic matter) and the *secondary migration* e.g. the hydrocarbon movement from the rock source towards the reservoir permeable rock. When migration stops, a hydrocarbon accumulation forms, only where hydrocarbons encounter a trap. The trap is a

tectonic or lithologic arrangement that blocks the hydrocarbon migration, keeping them in equilibrium (Figure 1.1).

The second geologic information is the internal reservoir lithological architecture. The architecture includes the geometry and the topology of the rock types present in the reservoir. A petroleum reservoir is composed of various rock types whose spatial distribution obeys stratigraphic and diagenetic principles. The diagenesis refers the sediment modification occurring after deposition, whilst stratigraphy is the geological study of the form, geographic distribution, chronological succession, classification, and correlation of rocks [*].

The spatial distribution of the lithological formations has two components: one is the lateral distribution and the other is the vertical stacking of the geological deposits. These two components are interconnected, being the results of the evolution through time of the depositional system.

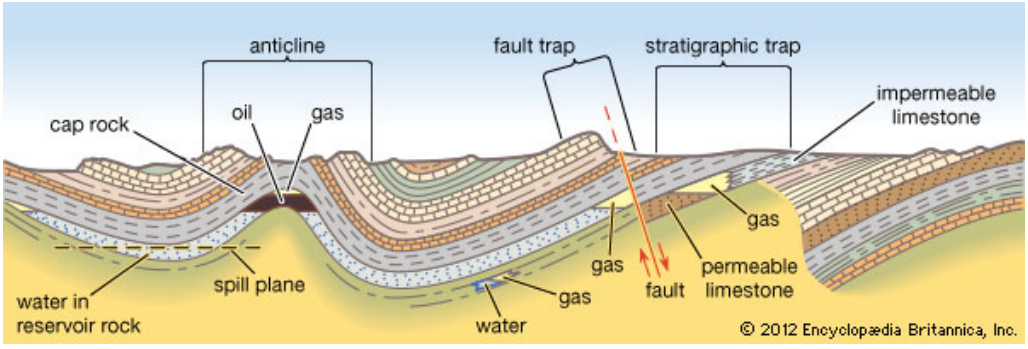


Figure 1.1: Types of hydrocarbon traps [†]

The oil recovery starts by drilling wells that perturb the equilibrium of the pressure system in the subsurface. The petroleum reservoir is a hydrodynamic unit, so any change in its pressure in time influences the entire pressure system. Consequently, when a pressure modification occurs, the hydrocarbons migrate towards areas that allow their release. The high pressure difference between surface and subsurface generates a flow that transports the hydrocarbons towards the surface. This is the *primary phase of oil recovery*, during which the hydrocarbon production reaches about $20\% - 30\%$ of the total reserves. With time, the pressure equilibrium is reestablished and the oil recovery gradually diminishes. Therefore, the oil companies proceed to a new stage named *secondary*

[*] Reservoir Engineering and Petrophysics-Petroleum Engineering Handbook, Vol V

[†] Image courtesy of Encylopaedia Britannica, in accordance with section 1.Terms of Use for Everyone subsection Use of Content of the terms at http://corporate.britannica.com/termsofuse.html, accessed on 28 February 2014.

recovery. During this phase the pressure system is stressed by injection of water or gas in some of the drilled wells (the injection wells). This injection generates hydrocarbon displacement and, consequently, their recovery through the other wells (the production wells). This procedure increases the amount of recovered hydrocarbons, but still not enough, half of the reserves being still trapped in the formation. After a while, the injected water reaches the production wells, which decreases the amount of hydrocarbons recovered. When the water injection is no longer economically profitable (the amount of hydrocarbons does not cover the costs with water injection), the oil companies have two options: either close the production, or pass to a further phase named *tertiary oil recovery* or enhanced oil recovery (EOR). During this phase, complex chemicals or steam are injected in order to change the properties of the fluids and the rocks for an easy flow of the hydrocarbons.
The hydrocarbon displacement, during exploitation, generates a flow which follows trajectories that depend on the spatial distribution of the petrophysical properties (e.g. porosity, permeability, water saturation, relative permeability) of the lithological structures present in reservoir. The spatial distribution of the petrophysical properties is related to the spatial distribution of the rock bodies that form the reservoir geology, different rock types, usually have different petrophysical properties. Only locally, certain lithological units may have comparable petrophysical properties. For example, crevasse splays may have, locally, comparable petrophysical properties with channel belts. These rock bodies, distinguished by petrophysical properties and mineralogy are called facies. Consequently, any production optimization plan or field development plan must take into account the reservoir geology. A better description of the reservoir geological architecture improves the geological simulation model, which is the foundation of the decision-making process.

1.2 Data assimilation in reservoir engineering

Even though the geological simulation model is well calibrated to generate geologically plausible realizations, theirs predictions (e.g. the simulated production outcome) are not necessary, close to the real production data. A field development plan, as part of the decision-making process, is based on plausible geological realizations whose simulated measurements predicts "well enough" the reality. These predictions are obtained coupling the geological model with a numerical model that simulates the reservoir behavior under established conditions (given geology, wells positions, well controls etc.). Under given well controls, the predictions depend on the initial state of the dynamical variables (pressures and saturations), petrophysical parameters (permeability, porosity, relative permeability, etc.) and properties of the fluids and/or gas present in reservoir (viscosity, density). A reservoir is located thousands of meters underground and its measurements

are sparse and either contains little spatial information or does not contain spatial information at all. This causes a high uncertainty during simulation using the geological model. Consequently, the initial conditions of the numerical model lie in a sea of uncertainty, which makes it almost impossible to compute estimation of future production close to reality. In order to obtain plausible geological realizations with high predictive capacity, a solution would be to use the data assimilation techniques. Data assimilation is a methodology that incorporates measurements into the mathematical model of the system. The purpose of this methodology is to calibrate the state of the system such that the model can simulate observations in the closest proximity of the measurements. In data assimilation, the period of time when the measurements are used for calibration is called the assimilation period. When the methodology is used for the dynamical variable estimation, we name the data assimilation procedure *state estimation*; when the methodology is used for the calibration of model parameters (static variables), we call it *inverse modeling* procedure (or parameter estimation).

In reservoir engineering the data assimilation is called **history matching** (HM) and it has been traditionally used to calibrate the model parameters, such that the simulated production response closely reproduces reservoir past behaviour. Initially, within the history matching framework, the model parameters were adjusted manually and, consequently the methodology was named **manually history matching**. In this case, the model parameters considered were the petrophysical properties of the rock types present in the reservoir. Nowadays, for the parameter calibration, complex procedures that involve developing of automatic routines are used and the methodology is named **Assisted History Matching** (AHM).

In the last two decades, many AHM algorithms were proposed (Aanonsen et al 2009, Oliver and Chen 2011, Oliver et al. 2008). Depending on the mathematical procedures used, one may distinguish two directions: *gradient based* and *gradient free* methodologies. The gradient based methodology defines an analytical function that is to be optimized. The simplest way to introduce a parameter calibration within a such procedure is achieved through a function that describes the square of the Euclidian distance between the simulated measurements and the reservoir observations (Oliver et al 2008). However, the measurements are contaminated with errors, so the pure euclidian distance is traditionally replaced with a weighted regularized least-square function. The minimization involves the use of gradients. Various types of algorithms were proposed to solve the minimization problem (basically, the gradient calculus), either involving the adjoints (Chen et al.(1974), Zhang and Reynolds (2002)) or approximations of the gradients using Newton-like methods (Oliver et al. 2008). The gradient free methodologies can be categorized according to Sarma et al. (2007) in *stochastic algorithms* (gradual deformation method Roggero and Hu (1998), Hu and Ravalec (2004), Gao et al. (2007), probability conditioning method Hoffman and Caers (2006)), *streamline based techniques* (Vasco

and Datta-Gupta (1997), (1999)) and *Kalman filter approaches*. In Chapter 3 we offer a detailed description of the Kalman filter methodology, together with an overview of the ensemble based methods introduced in the Kalman filtering framework.

Irrespective of the assisted history matching algorithm used, the result of the estimation process must satisfy a few requirements:

- To predict the future behaviour of the reservoir in existing and new wells with increased confidence.
- To maintain geological acceptability.
- To incorporate (condition on) all quantifiable information (seismic, gravity, well logs etc.)

In addition, when possible, the AHM method must offer an uncertainty quantification.

The preservation of the geological plausibility during the AHM process is still a challenge and it is the main goal of this work. The geology is very complex, and its acceptance preservation, when optimization procedures are performed, has been obtained only in few particular cases for synthetic reservoirs, and under particular assumptions regarding the geology simulation. A general AHM method, that is capable to preserve the geology requirements (geometry and topology) when applied, does not yet exist.

1.3 Geologically consistent data assimilation

One of the main problems still associated with the use of data assimilation methods for history matching of reservoir models is the lack of geological realism in updates. The assumption is that reservoir models which are geologically plausible and match all available data are better (i.e. give better predictions) than models which are not geologically realistic or do not match all data. In the normal workflow, the changes to the properties of reservoir models (like porosity and permeability) introduced by the history matching process are usually not consistent with geological knowledge about the reservoir, so that the optimal reservoir model is not achieved. Roughly speaking, there are three possible ways to solve these problems.

1. First, the parameters of the models used to generate the geological models are adjusted instead of the properties themselves. An example is the probability perturbation method (Hoffman and Caers, 2004) where the settings of the geological model used for generating realizations are estimated. Another example is a method implemented by Hu et al. 2012 where the parameters incorporated in the multi-point geological simulation model are calibrated within ensemble

based AHM process. These methods can guarantee realistic geology, but generally achieve a worse fit to the production data.

2. The second option is to estimate the properties while retaining the relevant geological input, for example by retaining multi-point statistics. More fundamental geological input, like the topological ordering of facies or layers, may still be lost in such an approach. Examples for the second approach are the Gaussian mixture models (Dovera and Della Rossa, 2011) for retaining multi-point statistics, discrete cosine transform (DCT, Jafarpour and McLaughlin 2007) for preserving patterns similar to image processing, wavelets (Jafarpour 2010) or Kernel PCA (Caers 2004, Sarma et al 2008, Sarma and Chen 2009) based on a projection of the log-permeability (the property of interest) into a high dimensional space based on kernel functions.

3. The third option is an intermediate solution in which the distribution of the facies is estimated explicitly, using appropriate parameterizations of the facies fields. Also here arises a question whether geological acceptability is preserved. Examples for the third approach are the truncated pluri-Gaussian method (Liu and Oliver, 2005), the level set method (David Moreno 2009), the gradual deformation method (Hu et al. 1998, 2004), parameterizations using distance fields (Lorentzen et al 2012) or using level-set type functions (Chang et al 2010).

Although the first method is intuitively appealing, some major problems are associated with it. The main problem is the strong non-linearity and lack of sensitivity between the parameters of the models generating the geological instances and the production data. This means that an approach where the geological model parameters are adjusted based on a mismatch between model predictions and observations are difficult to apply.
An important question for the second method is what should be preserved in the assimilation step. It is not straightforward which parameters are important in describing geology and, more specifically, the geological features important for reservoir simulation. The third method has the disadvantage that the production data can be rather insensitive to boundaries between facies, which are thus hard to estimate. The question also arises whether the properties of the facies need to be estimated jointly with the facies distribution. Finally, it may not always be possible to describe the reservoir using a facies distribution, especially in the case of gradual transitions between facies or complex fine-scale distributions of the facies.

1.4 Research objectives

The goal of this project is to investigate methods currently being used in the third category and identify the most promising one. One of these methods should be further developed and tested on a geologically realistic reservoir model. The question which method preserves the prior geological information in the best way, depends on two parts: how the geology is better described and how well the chosen parameters are calibrated by the data assimilation method. The first part is case dependent. Therefore, two different geological simulation models are used for the construction of the internal geometry and topology of the reservoir. For both geological simulation models, appropriate parameterizations should be found such that the AHM method chosen to provide geologically plausible updates.

1. The first geological simulation model is based on two point statistics and is defined in the context of plurigaussian simulation, either by a truncation scheme or by a pure simulation using spatially correlated Gaussian variables. In this case the parametrization is made straightforward by the Gaussian fields themselves. We aim to introduce appropriate settings of the geological model so that the simulated facies instances are geologically plausible and conditioned to all available measurements.

2. The second geological simulation model is based on multi-point geostatistical simulation (MPS). Here, complex curvilinear structures (as channels) can be successfully generated from conceptual geological models (training images). The MPS model has the advantage to generate facies instances with increased geological acceptability, but for their direct estimation a suitable parametrization is needed. Consequently, we aim to introduce a new parametrization of the facies fields coupled with an appropriate AHM method to drive the estimation process towards the requirements presented.

Besides the estimation purpose, by using ensemble based methods for data assimilation, we aim to offer an uncertainty quantification of the facies distribution. This is crucial when reservoir optimization production plans are investigated taking into account the geological uncertainties. The estimation of uncertainties in the structural model (the boundaries between layers, faults, external reservoir geometry) and the introduction of these uncertainties as updatable parameters in the AHM method are not addressed in this project.

1.5 Thesis outline

The thesis is organized as follows:

- In Chapter 2 we describe the geological simulation models used in this project. We present also the mathematical instruments required for defining the methodology of each geological simulation model.
- Chapter 3 is dedicated to the data assimilation methodology. We describe the ensemble based methods used for history matching.
- In Chapter 4 we present the probabilistic parametrization of the facies fields. Based on this parametrization, we build a geological simulation model which is further coupled with the ensemble Kalman filter (EnKF) as data assimilation method. We place this geological simulation model in the large family of plurigaussian simulation models. We provide an estimator of the reference facies field together with its associated uncertainty quantification. This chapter is based on Sebacher et al. 2013.
- In Chapter 5 we describe a geological simulation model by means of which the plurigaussian simulation is conditioned to facies probability maps. The name of the model is the Adaptive Plurigaussian Simulation (APS). We couple the geological simulation model with EnKF, as history matching method, for the geology uncertainty quantification. This chapter is based on an article submitted to Computational Geosciences.
- Chapter 6 is dedicated to channelized reservoirs. In this chapter we define a new parametrization of the facies fields, in the context of MPS geological simulation model, which is coupled with the EnKF and the iterative adaptive Gaussian mixture (IAGM). We perform the experiments on two reservoir models with different complexity. This chapter is based on the article submitted to Computational Geosciences.
- In Chapter 7 we present the performance of the APS model in the case of a real field application for a reservoir located in the North Sea. This chapter is based on an article presented at the 76th EAGE Conference and Exhibition 2014 (Hanea et al. 2014).
- In Chapter 8 we present the conclusions of this study and some recommendations regarding a possible continuation of this work.

Chapter 2

Geological Simulation Models

2.1 Truncated plurigaussian simulation (*TPS*)

The main ingredient for the plurigaussian truncation methodology consists of the Gaussian random field. Consequently, before presenting the (geological) simulation model, we will first give an introduction to random field theory with a focus on a particular shape e.g. the Gaussian random field. The second ingredient of the simulation model is the truncation scheme (map). This is introduced in the second subsection, where the simulation model is presented together with some illustrative examples.

2.1.1 The Gaussian random fields

Definition: Let (Ω, F, P) be a probability space and $T \subseteq \mathbf{R}^d$ ($d \geq 1$, integer), a parameter set. A *random field* is a function $Y : T \times \Omega \to \mathbf{R}^m$, such that $Y_t = Y(t, \cdot)$ (e.g. $Y_t(\omega) = Y(t, \omega)$, for $\omega \in \Omega$) is a $\mathbf{R}^m$-valued random variable on (Ω, F, P), for every $t \in T$.
In the following, we develop the theory for the particular case $m = 1$, which corresponds to real valued random fields. Consequently, for $T \subseteq \mathbf{R}^d$ we can define a random field Y as a function such that Y_t are random variables for any $t \in T$. The dimension d of the real space which includes T gives the dimension of the random field. A one-dimensional random field is usually called a stochastic process. The random fields in two or three dimensions are often encountered in earth sciences, such hydrology or geology, where spatial correlations of the values of some parameters are needed. Additionally, random fields are often used in meteorology to model space-time dependent variables.

Let be Y a real valued random field with parameter set $T \subseteq \mathbf{R}^d$ ($d \geq 1$). Usually, this random field is written as $\{Y_t, t \in T\}$. For any choice of the positive integer k and

of the finite set of parameters $\{t_1, t_2, \ldots, t_k\}$ from T, we define the function, $F_{t_1,t_2,\ldots,t_k} : \overline{\mathbf{R}}^k \to [0,1]$, where $\overline{R} = R \cup \{-\infty, \infty\}$

$$F_{t_1,t_2,\ldots,t_k}(x_1, x_2, \ldots, x_k) = P(Y_{t_1} \leq x_1, Y_{t_2} \leq x_2, \ldots, Y_{t_k} \leq x_k) \tag{2.1}$$

such that, $F_{t_1,\ldots,t_k}(x_1, \ldots, -\infty, \ldots, x_k) = 0$ and $F_{t_1,\ldots,t_k}(\infty, \ldots, \infty) = 1$
We call these functions the finite-dimensional (cumulative) distribution functions (f.d.d. functions) of the random field Y. These functions have two important properties, symmetry and compatibility.

1. **Symmetry**: For any permutation σ of the set $\{1, 2, \ldots, k\}$, and for any elements $t_i \in T, i \in \overline{1,k}$ we have

$$F_{t_1,t_2,\ldots,t_k} = F_{t_{\sigma(1)},t_{\sigma(2)},\ldots,t_{\sigma(k)}} \tag{2.2}$$

2. **Compatibility**: For any $k-1$ elements $t_i \in T, i \in \overline{1,k-1}$ and for any real numbers $x_i, i \in \overline{1,k-1}$ we have

$$F_{t_1,t_2,\ldots,t_{k-1}}(x_1, x_2, \ldots, x_{k-1}) = F_{t_1,t_2,\ldots,t_{k-1},t_k}(x_1, x_2, \ldots, x_{k-1}, \infty) \tag{2.3}$$

These properties hold because the values of the f.d.d. functions are probabilities calculated for the same events of the probability space (Ω, F, P). Related with these two properties is the Kolmogorov theorem which emphasise the importance of the f.d.d. functions in the context of the existence of the random fields.

Kolmogorov existence theorem[*]

If a system of probability measures denoted $F_{t_1,t_2,\ldots,t_k}$ ($t_i \in T \subseteq \mathbf{R}^d$) satisfies the symmetry conditions and the compatibility conditions, there exists a probability space (Ω, F, P) and a random field defined on it, having $F_{t_1,t_2,\ldots,t_k}$ as its finite-dimensional distribution functions.

This theorem gives the necessary and sufficient conditions for the random fields existence. Based on this theorem, we will introduce Gaussian random fields. Before that, we introduce the first two moments associated with the random fields.

1. **The expectation (mean) function**: $m : T \to \mathbf{R}$

$$m(t) = \int_{\mathbf{R}} x \, dF_t(x) = E(Y_t) \tag{2.4}$$

[*]Abrahamsen P., A review of Gaussian Random Fields and Correlation Functions, 1997

2. **The covariance function**: $C : T^2 \to \mathbf{R}$

$$\begin{aligned} C(t_1, t_2) &= Cov(Y_{t_1}, Y_{t_2}) = E((Y_{t_1} - m(t_1))(Y_{t_2} - m(t_2))) \\ &= \iint_{\mathbf{R}^2} xy \, dF_{t_1,t_2}(x, y) - m(t_1)m(t_2) \end{aligned} \tag{2.5}$$

For the random fields characterized by continuous variables, the f.d.d. functions can be defined by the probability density functions (*pdf*). These functions can be calculated as partial derivatives of the f.d.d functions.

$$p_{t_1,\ldots,t_k}(x_{t_1}, \ldots, x_{t_k}) = \frac{\partial_k F_{t_1,\ldots,t_k}(x_1, \ldots, x_k)}{\partial x_1 \ldots \partial x_k} \tag{2.6}$$

We can write the mean function and the covariance function with respect to the probability density functions. $m(t) = \int_{\mathbf{R}} x \, p_t(x) \, dx$ and $C(t_1, t_2) = \iint_{\mathbf{R}^2} xy \, p_{t_1,t_2}(x, y) \, dx \, dy$. In addition, if we denote by $\sigma(t) = \sqrt{C(t, t)}$ we can define **the (auto) correlation function**:

$$\rho(t_1, t_2) = \frac{C(t_1, t_2)}{\sigma(t_1)\sigma(t_2)} \tag{2.7}$$

Note that the definition of the correlation function is correct, since $C(t, t) > 0$.

Random fields properties

We have defined the random fields with no restrictions for the parameter space T. However $T \subseteq \mathbf{R}^d$, and $\mathbf{R}^d$ has a structure of an Euclidian space. When T borrows properties from $\mathbf{R}^d$ (the linear structure or the Euclidian structure) we can describe special classes of random fields. These classes are defined through properties that the f.d.d. or mean and covariance functions may have. These properties refer to the invariance of f.d.d. functions to some transformations of the space T.

- **Stationarity (Homogeneity)**

  Definition: A random field Y is said to be *strictly stationary (or homogeneous)* if its finite-dimensional distributions are invariant under a space translation of the linear space T.

  That is, if we consider the translation $\phi_\tau : T \to T$, $\phi_\tau(t) = t + \tau$ then the stationarity property can be written as

$$F_{t_1,t_2,\ldots,t_k}(x_1, x_2, \ldots, x_k) = F_{\phi_\tau(t_1),\phi_\tau(t_2),\ldots,\phi_\tau(t_k)}(x_1, x_2, \ldots, x_k) \tag{2.8}$$

  Consequently, for stationary random fields the mean function is constant on the space T ($m(t) = m$ for every $t \in T$), whereas the covariance function can be defined one-dimensional as $C(t_1, t_2) = C(\tau)$, where $\tau = t_2 - t_1$. This holds because

  $m(t) = \int_{\mathbf{R}} x \, dF_t(x) = \int_{\mathbf{R}} x \, dF_{0+t}(x) = \int_{\mathbf{R}} x \, dF_0(x) = m(0)$ and

  $\iint_{\mathbf{R}^2} xy \, dF_{t_1,t_2}(x, y) = \iint_{\mathbf{R}^2} xy \, dF_{0+t_1,(t_2-t_1)+t_1}(x, y) = \iint_{\mathbf{R}^2} xy \, dF_{0,t_2-t_1}(x, y)$

In addition, the correlation function can be written as

$$\rho(t_1, t_2) = \frac{C(t_2 - t_1)}{\sigma(t_1)\sigma(t_2)} = \frac{C(t_2 - t_1)}{\sigma(0)^2} = \rho(t_2 - t_1) \tag{2.9}$$

hence, $C(t_1, t_2) = C(t_2 - t_1) = \sigma(0)^2\rho(t_2 - t_1)$
When only the mean function and the covariance function are invariant to any space translation the random field Y is called *weakly stationary*.

- **Isotropy**
  For the definition of the isotropy, besides the linear characterization of the space parameter $T \subseteq \mathbf{R}^d$ we consider the Euclidian structure of the space $\mathbf{R}^d$. This means that we can calculate the Euclidian distance between any two point t_1 and t_2 from T, using the formula:

$$d(t_1, t_2) = \sqrt{\sum_{j=1}^{d} (t_1^j - t_2^j)^2}$$

  Definition: A stationary random field Y is said to be *isotropic* if its covariance function depends only on the euclidian distance. That is, $C(t_1, t_2) = C(d(t_1, t_2))$. One can observe that the isotropic random fields are invariant to the translations and rotations of the parameter space T, and therefore to its isometries.

- **Anisotropy**
  Let us consider a positive semi-definite symmetric matrix $B \in M_d(\mathbf{R})$. Then the function $\| \cdot \|_B$: $\mathbf{R}^d \to \mathbf{R}^d$, $\| t \|_B = \sqrt{t^T B\, t}$ is a norm in the euclidian space $\mathbf{R}^d$, named B-norm.
  Definition: A stationary random field Y is said to be *anisotropic* if its covariance function depends on the distance defined based on the B-norm.
  That is, $C(t_1, t_2) = C(\| t_2 - t_1 \|_B)$.
  We extend the anisotropy property for the bi-dimensional case (d=2). Using an orthogonal transformation of the space $\mathbf{R}^2$ the semi-positive defined matrix B can be adjusted to its diagonal form. That is, $B = \Lambda^T \tilde{B} \Lambda$ where $\tilde{B} = diag(a, b)$ is the diagonal form of the matrix B. Because of the semi-positive definiteness of the symmetric matrix B, a and b are non-negative real numbers. The directions of the two eigenvectors of the matrix B give the principal and secondary anisotropy directions. The principal direction is given by the eigenvector that corresponds to the higher eigenvalue, whereas the secondary direction is given by the eigenvector that correspond to the lower eigenvalue. One can notice that the two eigenvectors are orthogonal to each other due the symmetry of the matrix B.

Gaussian random fields

Definition: A *Gaussian random field* is a random field for which finite-dimensional distributions $F_{t_1,t_2,\ldots,t_k}$ are multivariate Gaussian distributions for any choice of the number k and the parameters $t_1, t_2, \ldots, t_k$ from the space T.
If the f.d.d. functions are consistently defined (satisfying the conditions of the probability measure), the Kolmogorov's hypotheses are fulfilled and, consequently the definition of the Gaussian random field is correct. A cumulative distribution function of a multivariate Gaussian variables is defined by its probability density function (*pdf*). In the multivariate Gaussian case the *pdf* is written as:

$$p_{t_1,t_2,\ldots,t_k}(x_1, x_2, \ldots, x_k) = \frac{1}{\sqrt{2\pi \mid \Sigma_{t_1,t_2,\ldots,t_k} \mid}} e^{-\frac{1}{2}(x-m)^T \Sigma^{-1}_{t_1,t_2,\ldots,t_k}(x-m)} \quad (2.10)$$

where, $x = (x_1, x_2, \ldots, x_k) \in \mathbf{R}$ and $m = (m(t_1), m(t_2), \ldots, m(t_k)$, $m(t_i) = E(Y_{t_i})$ are the expectations. The components of the matrix Σ give the covariances of the marginal variables, e.g. $\Sigma_{t_1,t_2,\ldots,t_k}(i,j) = Cov(Y_{t_i}, Y_{t_j})$, for every $t_i, t_j \in T$. Reciprocally, a function $p_{t_1,t_2,\ldots,t_k} : \mathbf{R}^k \to \mathbf{R}$ could be a probability density function of a multivariate Gaussian random variable if and only if the matrix $\Sigma_{t_1,t_2,\ldots,t_k}$ is symmetric and positive definite.

Consequently, the existence of the Gaussian fields is related to the consistent definition of the matrices $\Sigma_{t_1,t_2,\ldots,t_k}$, for any choice of the elements $t_1, t_2, \ldots, t_k$ from T. Since the components of matrices $\Sigma_{t_1,t_2,\ldots,t_k}$ are the covariances $Cov(Y_{t_i}, Y_{t_j})$, the Gaussian random fields are well defined by two functions $m : T \to \mathbf{R}$ and $C : T \times T \to \mathbf{R}$. If the function m has no restrictions the function C must be symmetric ($C(t_1, t_2) = C(t_2, t_1)$) and all the square matrices $\Sigma_{t_1,t_2,\ldots,t_k}$ of which components are defined as $\Sigma_{t_1,t_2,\ldots,t_k}(i,j) = Cov(X_{t_i}, X_{t_j})$ must be positive definite. The positive definiteness of the matrices Σ introduces a property for the function C called, as well, the positive definiteness. Consequently, a function $C : T \times T \to \mathbf{R}$ is said to be positive defined, if for every positive integer k and for every real numbers α_i $(1 \leq i \leq k)$ and elements $t_1, t_2, \ldots, t_k$ from T, we have

$$\sum_{i,j=1}^{k} \alpha_i \alpha_j C(t_i, t_j) > 0 \quad (2.11)$$

We will say that, the function C gives the covariance model of the Gaussian random field. To prove that a function is positive defined is a tedious process that involves the use of the Fourier transform (see Christoakos 1984,1992) and is not discussed in this introductory section. However, the finite sum or product of acceptable covariance models is an acceptable covariance model (satisfies the positive definiteness property, Christoakos 1992).
In the following, we present some of the underlying covariance models that are used

for defining stationary Gaussian random fields (isotropic or anisotropic). We will use these models in the next section where the truncated plurigaussian simulation model is presented.

Examples of covariance models:

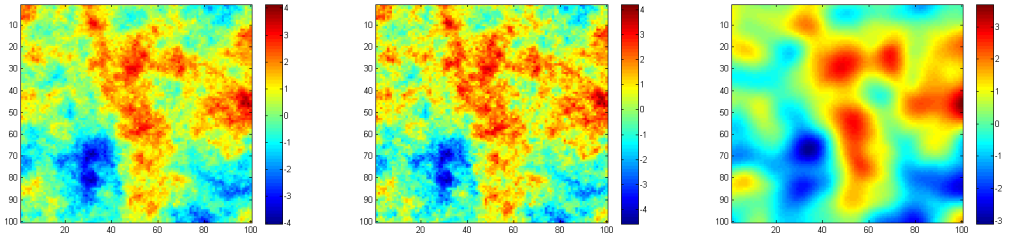
(1) Spherical model: $C(\tau) = \begin{cases} \sigma_0^2(1 - \frac{3}{2}(\frac{\tau}{h}) + \frac{1}{2}(\frac{\tau}{h})^3) & \text{if } 0 \leq \tau < h \\ 0 & \text{if } \tau \geq h \end{cases}$

(2) Exponential model: $C(\tau) = \sigma_0^2 e^{-\frac{3\tau}{h}}$, for $\tau \geq 0$

(3) Gaussian model: $C(\tau) = \sigma_0^2 e^{-\frac{3\tau^2}{h^2}}$, for $\tau \geq 0$

In geosciences, the term variogram model is commonly used instead of covariance model. The variogram of a random field Y is defined as $\gamma(t_1, t_2) = \frac{1}{2}Var(Y_{t_1} - Y_{t_2})$. For a stationary random field one may write $\gamma(t_1, t_2) = \frac{1}{2}E((Y_{t_1} - Y_{t_2})^2)$ and it can be shown that $\gamma(t_1, t_2) = C(0) - C(t_2 - t_1)$, where $C(t_1, t_2) = C(t_1 - t_2)$ is the covariance model. Consequently, for a stationary random field the variogram depends on the difference $t_2 - t_1$, $\gamma(t_1, t_2) = \gamma(t_2 - t_1)$ like the covariance does, and $\gamma(\tau) = C(0) - C(\tau)$ (Kelkar and Perez 2002). The variogram behaves complementary to the covariance. For example, when we use decreasing covariance models (like the ones presented above) for increasing lag-distance τ the variogram increases from zero towards σ_0^2. The parameter h that occurs in the definition of the covariance models presented before is named "the range" of the covariance model (or variogram model). From the covariance model perspective, when the distance τ exceeds value h the covariance is zero (for spherical model) or almost zero (for the other models).

The plurigaussian truncation model uses Gaussian fields defined on discrete parameter sets. To generate samples (values) of the Gaussian random fields, when the set of parameters is finite, one may use, for instance the sequential Gaussian simulation method (Kelkar and Perez 2002) or the moving averages method (Oliver 1995). For that, we have to set up the *geostatistical properties* of the Gaussian field e.g. the mean function, the variance function, the covariance (variogram) model type, the correlation direction, and the directional correlation ranges. In the Figure 2.1 we present three samples of stationary Gaussian random fields defined on the finite set $\{1, 2, \ldots, 100\}^2$. The mean model is equal to 0 and the variance model is equal to 1. The Gaussian random fields were chosen to have the isotropic property with a range of 20. The variogram models used in this example were the spherical (sub-figure (a)), exponential (sub-figure (b)) and Gaussian(sub-figure (b)). The samples were carried out using the sequential Gaussian simulation method implemented in S-GeMS (The Stanford Geostatistical Modeling Software).



(a) Spherical variogram model (b) Exponential variogram model (c) Gaussian variogram model

Figure 2.1: The Gaussian random fields samples

2.1.2 The truncated plurigaussian simulation model (*TPS*)

The truncated plurigaussian simulation model is a methodology that has been proposed to simulate the lithological architecture of the subsurface. Initially, it has been developed for the simulation of the facies distributions in petroleum reservoirs. The idea was to relate the spatial correlation of one or more Gaussian fields, defined on the reservoir domain, to the spatial distribution of the facies types. The facies simulation is carried out through a projection from a continuous space (defined by the space of the Gaussian fields values) into a discrete space (the facies fields space). The projection is defined by a truncation map of the Gaussian fields values. The truncation map is a decomposition of the Euclidian space $\mathbf{R}^n$ (n is the number of the Gaussian fields used in the model) in subdomains each having assigned a facies type. Depending on the number of the Gaussian fields involved, the method is called *the truncated Gaussian simulation (TGS)*, for a single Gaussian field and, respectively, *truncated plurigaussian simulation (TPS)* when more than two Gaussian fields are involved in the simulation. The truncated Gaussian simulation method have been used first by Matheron et al. 1987 for simulating facies fields for which the topology exhibits a sequential ordering of the facies types. The truncated plurigaussian simulation model was introduced later, in 1994 by Galli et al. (1994) and Le Loc'h (1994) as a generalization of the first model.

The truncated Gaussian simulation method has two ingredients:

- *The Gaussian random field* defined on a region of interest. The region of interest in reservoir engineering is the reservoir domain, but with a discrete structure (grid cells). Consequently, the parameter space T, on which the Gaussian field is defined is a discrete set. However, throughout this study we refer at Gaussian fields defined on the reservoir domain.

- *A truncation map* defined on the real axis. The truncation map is defined by some

thresholds that divide the real axis into intervals, each having assigned a facies type.

If we consider the reservoir domain D, in two or three dimensions, and a number of k facies types that occur in the domain, then the regions occupied by the facies type j in the domain D could be described as:

$$F_j = \{u \in D | s_{j-1} \leq Y(u) < s_j\} \tag{2.12}$$

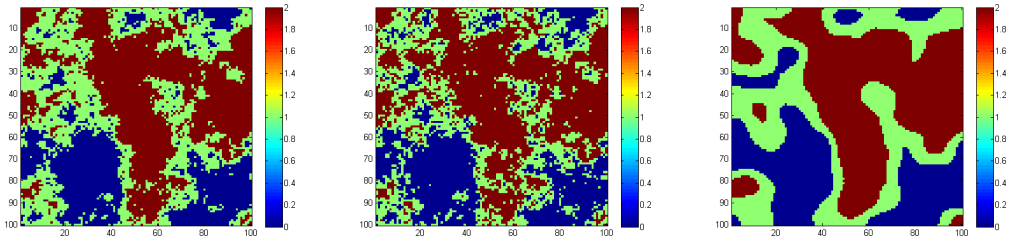
where Y is a sample of the Gaussian field, $Y(u)$ is the value of of the Gaussian field at the location $u \in D$, and $\{s_j, j = \overline{0,k}\}$ are the thresholds that define the truncation map. We use the convention $s_0 = -\infty$ and $s_k = \infty$. The choice of the thresholds that define the one-dimensional truncation map is based on the best knowledge about the proportions of the facies types. If we consider $prop_j$ as being the proportion of the facies type j, $(j = \overline{1,k})$ then we have the relation between the proportions and the thresholds

$$prop_j = \int_{s_{j-1}}^{s_j} pdf(x)dx \ = cdf(s_j) - cdf(s_{j-1}) \tag{2.13}$$

where pdf and cdf represents the probability density function and, respectively, the cumulative distribution function of the normal distribution used for definition of the (stationary) Gaussian field. Using these equations, we find the thresholds as functions of proportions

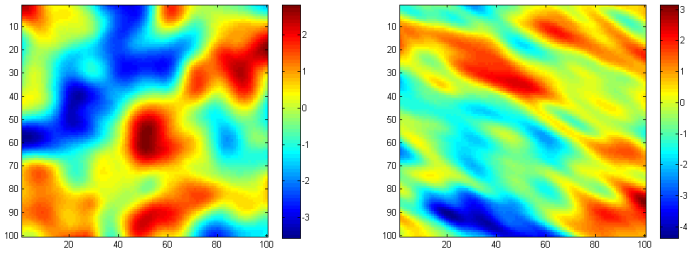
$$s_j = cdf^{-1}(\sum_{i=1}^{j} prop_i) \tag{2.14}$$

In practice, the facies proportions are estimated using data collected at the reservoir exploration phase. However, the indicator facies proportions has two components. One is the global indicator facies proportions and represents the proportion of each facies type in entire reservoir domain. Based on the values of this indicator one may define the thresholds described by the equations 2.14. The second meaning refers to a kind of spatial distribution of this indicator. The studies related with the lateral distribution of the geological deposits may conclude that in different regions of the reservoir domain the facies types may have different proportions. This means that we deal with spatial non-stationarity in facies proportions. In this case the values of the thresholds vary with the location in the domain (Galli et al 1997).
In Figure 2.2 we present three examples of facies fields obtained after truncation of the Gaussian fields presented in the Figure 2.1. For that, we have defined two thresholds $s_1 = -0.4$ and $s_2 = 0.8$ on the real axis and assign to each interval, a facies type. The blue facies type is assigned for the Gaussian field values less than -0.4, the green facies type for the values between -0.4 and 0.8 and the red facies type for the values greater than 0.8.

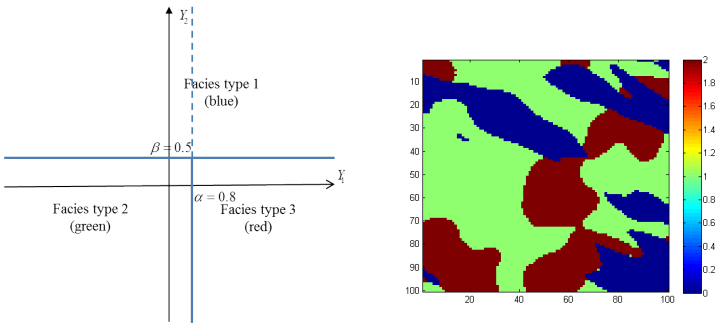


(a) Spherical variogram model (b) Exponential variogram model (c) Gaussian variogram model

Figure 2.2: The facies distributions in truncated Gaussian scheme



(a) The first gaussian random field (b) The second Gaussian random field



(c) The truncation map (d) The facies field

Figure 2.3: The truncated plurigaussian scheme

One can observe that using a single Gaussian field with an one-dimensional truncation map, the simulation results are restrictive regarding the facies topology. In our example, we refer to the constraint that the red facies type has no contact with the blue facies type. When the stratigraphic studies indicate the possibility of the contact among

many facies types or there is no clear transition among the facies types, the facies simulations with truncated Gaussian model are not reliable. In order to enlarge the facies simulation space, the methodology has been extended, to work with many Gaussian fields and with a truncation map defined in a multi-dimensional space. This is called *the truncated plurigaussian simulation (TPS)* method. For a easy understanding, we present the methodology for the case when two Gaussian random fields are involved in truncation.

In this case, the plurigaussian simulation technique has two ingredients: two Gaussian random fields defined on the reservoir domain and a truncation map designed in a bi-dimensional real space. The truncation map is defined by a partition of real plane in regions, each having assigned a facies type. This division of the plane could be done by intersection of various curves types. For an easy implementation of the model, the curves were chosen as horizontal and vertical lines (Galli 1994).
In Figure 2.3 we give an example of truncated plurigaussian methodology. For that, we sample from two uncorrelated stationary Gaussian random fields, denoted Y_1 and Y_2, defined on a rectangular domain of 100*100 grid cells. We set the Gaussian variogram type for both random fields. The first Gaussian field is isotropic with the correlation range of 20 grid cells (Figure 2.3, top left), whilst the second is anisotropic with principal direction of 120°, long correlation range of 30 grid cells and short correlation range of 10 grid cells (Figure 2.3, top right). We simulate the distribution of three facies types, any two of them possibly having contact. For that, we design the truncation map on the bi-dimensional real plane Y_1OY_2 by intersection of two lines, one horizontal and the other vertical. These lines divide the plane in four regions. We assign a facies type for each region, but having three facies types, two regions will have the same assignment. Consequently, we can consider that the plane partition contains only three regions each having assigned a facies type. The truncation map is presented at the bottom left of the Figure 2.3. The assignment of the facies type, at a certain cell u of the domain, is carried out representing the point $(Y_1(u), Y_2(u))$ in the cartesian system Y_1OY_2. At this location, we assign the facies type of the zone where the point $(Y_1(u), Y_2(u))$ belongs. This projection from a continuous space into a discrete space is also called the "rock type rule". In our example, "the rock type rule" works as follows. At the grid cells u where $Y_2(u) > 0.5$ we assign the blue facies type, where $Y_2(u) \leq 0.5$ and $Y_1(u) < 0.8$ we assign the green facies type and where $Y_2(u) \leq 0.5$ and $Y_1(u) \geq 0.8$ we assign the red facies type. The facies map obtained with this procedure is presented at the bottom right of the Figure 2.3. One can observe that, there is contact among all facies types, request that could not be fulfilled with the truncated Gaussian procedure. The existence or not of the contact among the facies types is solved in the truncated plurigaussian procedure by the truncation map. Two facies types have contact if theirs regions from the truncation map are neighbors.

If in the truncated Gaussian case the truncation map is defined by the choice of the thresholds, in the plurigaussian case the truncation map is uniquely defined by the parameters of the curves that describe it. For the rectangular truncation map used in the example, these parameters consist of two values: one situated on the horizontal axis (α) and the other on the vertical axis (β). They are calculated, like in the first case, based on the best knowledge about the facies proportions. The difference is that, now, two Gaussian random fields are involved; therefore the problem related with the correlation between them occurs. If we denote by D_j the region from the truncation map assigned to the facies type j, and by $prop_j$ the proportion of the same facies type, then we have the relation:

$$prop_j = \iint_{D_j} pdf_{(Z1,Z2)}(z_1, z_2)dz_1dz_2 \tag{2.15}$$

where $pdf_{(Z_1,Z_2)}$ is the probability density function of the multivariate Gaussian variable (Z_1, Z_2). The marginal variables Z_1 and Z_2 are Gaussian variables with the distribution defined by the mean function and variance function of the Gaussian fields Y_1 and Y_2. The multivariate Gaussian variable (Z_1, Z_2) expresses the correlation between the Gaussian fields Y_1 and Y_2. Finding the parameters of the truncation map requires to solve the system of equations given by the relations 2.15 (Armstrong et al 2011).

However, the most complex procedure, when one wants to apply the *TPS*, is to establish the geostatistical properties of the Gaussian fields and to design the truncation map such that the simulations reflect as better as possible the reality. This means that the facies simulations are geologically acceptable. The geological acceptance refers to obtaining of realistic topology (the relative position among the facies) and geometry (shape, dimension, number of the facies). When the geostatistical properties of the Gaussian fields are set, a realistic topology of the facies simulation is achieved with a reliable choice of the truncation map (Lantuejoul 2002). The geometry of the facies is a property controlled by the geostatistical properties of the Gaussian fields. For a given truncation map, the correlation directions (which give the isotropy or anisotropy) and the directional correlation ranges of the Gaussian fields can be estimated knowing the mathematical relation between the indicator variogram of the facies types and the variogram of the Gaussian fields (Le Loc'h and Galli 1997, Armstrong et al. 2011). A much simpler approach is presented by Chang and Zhang (2014). Here the authors proposed a trial and error procedure for the estimation of the geostatistical properties of the Gaussian fields.

2.2 Multiple-point geostatistical simulation (*MPS*)

The goal of any geological (geostatistical) simulation model is to provide instances of facies maps (reservoir models) that are able to reflect the best knowledge about the subsurface geology. When the information gathered indicates the existence of geo-bodies with defined geometrical shapes (such curvilinear or elliptical structures), the use of geological models based on two-points statistics (variogram) (such *TPS*) is not suitable. The reason resides on what a variogram represents. The variogram is a statistical instrument that measures the dissimilarity between the same (or different) variable(s) values at two spatial locations. Consequently, a variogram model is limited in capturing geological features with defined shapes, it can capture trends but not concepts. For instance, cannot capture with a simple mathematical formula the geometry of curvilinear features (e.g. channels), although can establish a high correlation in its propagation direction. For that reason, the *TPS* model provides various textures of the facies maps, but it cannot keeps a fixed geometry of facies shapes.

To be able to model features with defined shapes, a solution is to use the multiple-point geostatistical simulation models (*MPS*, Guardiano and Srivastava 1993, Journel 1993). The multi-point geostatistical simulation models are geostatistical methodologies that takes into account correlations between many spatial locations at the same time. These correlations are inferred from conceptual models e.g. training images. Depending on the variable that has to be modeled, the training image can be discrete (or categorial, for facies distributions) or continuous (for variables with continuous values such as porosity or permeability).

2.2.1 The (discrete) training image

A (discrete) training image is a conceptual image (in two or three dimensions) designed to reproduce the topology, geometry and the connectivity of the lithological units (facies) from the subsurface e.g. the geological heterogeneity. A training image has a powerful visual impact, reflecting a prior geological model, which design is carried out based on geological interpretation from all available sources (outcrops, sample data, stratigraphic studies etc.), but without conditioning on any hard or soft data. A training image can be viewed as library of geometrical patterns that we believe it could be present in the subsurface. A pattern is a geometrical configuration, extracted from the training image, identifying a possible structure of the spatial continuity. These patterns are incorporated in the training image and reciprocally the training image is an assemblage of the patterns.

In the traditional *MPS* methodology, not any image that describe a type of geological heterogeneity could be a training image. To be used as training image, a conceptual image must satisfy some requirements.

1. *Stationarity*
   This property refers to the stationarity of the geometrical patterns that compose the training image. The goal of the *MPS* method is to generate facies maps with patterns borrowed from the training image. In the traditional *MPS* methodology, the simulation is carried out based on sampling from the empirical multivariate conditional probability density function (*cpdf*) of the geometrical templates (multi point statistics) calculated from the training image. Consequently, reliable facies maps are obtained when consistent *cpdf* are inferred from the training image. The consistency of the *cpdf* is ensured by the repeatability of the geometrical patterns within training image, coupled with stationarity in the geometry of the patterns (e.g. size and geometry of the elliptical shapes or width when refer to channels).

2. *The size of the training image*
   The size of the training image should be correlated with the size of the largest pattern that one would simulate within a given reservoir domain. For instance, when channels have to be simulated, the dimension of the training image should be at least twice larger than the dimension of the reservoir domain, in the direction of the channels continuity (Caers and Zhang 2004).

In addition, the number of the categorial variables from the training images should be restricted to maximum five. This is because the numbers of the geometrical templates increases exponentially with the number of categories present in the training image and a high number of categories may causes a huge computational effort in calculation and storage of the *cpdf*.

In Figure 2.4, we present three conceptual images showing three types of possible

(a) Conceptual image 1 (b) Conceptual image 2 (c) Conceptual image 3

Figure 2.4: Three possible candidates for a training image. Figure after Caers and Zhang 2004.

geological heterogeneities; the fluvial type (Image 1), the deltaic type (Image 2) and a conceptual image defined by elliptical shapes (Image 3). From these three images, only the first image fulfills the stationarity request and could be used as training image (for traditional *MPS* methodologies). One can observe that the deltaic image lacks the stationarity with respect to the width and anisotropy direction of the local patterns. The conceptual image with elliptical shapes does not keeps the stationarity in the dimension of the local patterns in space; its patterns are stationary regarding only the shape.

However, in the recent research on the *MPS*, the concept "training image" has been relaxed. New *MPS* methodologies have been developed to fit to training images that are not constraint to size and stationarity. Hu et al 2014 uses as training image an existent reservoir, and a new *MPS* algorithm is developed to simulate facies maps from "non-stationary" training images. The "non-stationary" training image can be built using any type of reservoir modeling approach. A non-stationary MPS methodology have been developed by Honarkhah and Caers (2012), for creating multiple-point geostatistical models, based on a distance-based modeling and simulation of patterns. The authors uses a deltaic training image (like Image 2) to present the methodology.

2.2.2 The simulation models

Two of the most commonly used *MPS* methodologies to simulate categories e.g. facies distributions are the SNESIM (single normal equation simulation, Strebelle 2002) and FILTERSIM (filter-based simulation, Zhang et al 2006,).

The SNESIM algorithm is an enhancement of the pioneered *MPS* algorithm proposed by Guardiano and Srivastava in 1993. It consists of two procedures. The first calculates, from the training image, the *cpdf* of the geometrical templates found within a user-defined window search. The second is a procedure that sequentially simulates a facies type at each grid cell of the reservoir domain. The simulation is based on a sample from the *cpdf* of a geometrical template found in the window centered at that location. The template is formed taking into account hard data (facies observations), if available, and the values at the previous simulated cells from the window. The grid cell of the domain are visited only once, based on a random path apriori given. The novelty introduced by Strebelle consists on a procedure for the calculation and the storage of the *cpdf* of the geometrical templates. In the original methodology, the training image had to be scanned each time when simulated at unsampled grid cell, which requires an extremely high CPU demanding. The SNESIM calculates the *cpdf* within a dynamical data structure called "search three", by which the training image has to be scanned only once. The dimension of the "search three" depends on the dimension of the window that scan the training image at the beginning of simulation.

In the later years the SNESIM procedure has been enhanced, with procedures that enable the conditioning of the facies simulation not only on the training image and the hard data, but also on probability maps coming from seismic interpretations. This has been done incorporating a probabilistic model named "tau model" in the simulation of a facies type at unsampled locations (Journel 2002, Krishnan et al. 2005). In addition, various procedures have been incorporated, by which non-stationary facies fields (with non-stationary geometrical patterns) can be generated started from training images with stationary characteristics (Strebelle and Zhang 2005). However, the SNESIM is limited regarding to the number of the facies types and is suitable only for discrete training images.

The *MPS* algorithm FILTERSIM has been proposed to overcome these issues. The FILTERSIM algorithm is less memory demanding with reasonable CPU cost and can handle with both discrete and continuous training images (Zhang et al 2006). FILTERSIM utilizes a set of filters to classify training (geometrical) patterns in a small real space of which dimension is given by the number of the filters used (called the filter space). A filter is a set of weights associated with all the cells of a geometrical template. A template is a local moving window used to scan the training image providing the training patterns. For a given training pattern each filter gives a score, and consequently each pattern is associated with a point in a filter space. By adequate partitioning of the filter space, the patterns are grouped in classes. An average of each class is called prototype for the patterns. The simulation with the FILTERSIM algorithm is performed in a sequential manner, using a random path to visit each cell. During simulation, the prototype closest to the conditioning data event (which comprises all the informed cells from the template with the center in the visited cell) is determined and a pattern randomly sampled from that prototype is pasted onto the simulation cell. As SNESIM, the FILTERSIM can be conditioned on probability maps inferred from seismic interpretations.

Another MPS algorithm is introduced by Arpat and Caers 2004. This methodology does not uses a grid-cell based simulation (as SNESIM) of the categorial variables, but a simulation based on patterns. For that reason the MPS methodology has been named pattern-based simulation. First, a database of patterns is extracted from a training image. Secondly, the simulation is carried out by pasting at each visited location along a random path a pattern that is compatible with the available local data and any previously simulated patterns. At each step, the reliable pattern is chosen by using a distance function that measure the similarity between patterns.

Chapter 3

Ensemble methods

3.1 Data assimilation basics

Data assimilation is a novel methodology for estimating variables representing certain state of nature. Estimation of a quantity of interest via data assimilation involves combining measurements with the underlying dynamical principles governing the system under observation. Different problems require knowledge of the distribution and evolution in space and time of the characteristics of the systems involved in each of them. The functions of space and time (state variables or model states) are the ones which characterize the state of the system under observation. A dynamical model to describe the nature consists of a set of coupled nonlinear equations for each state variable of interest. The discrete model for an evolution of dynamical system from time t_{k-1} to time t_k can be described by the equation of the form:

$$x(t_k) = \mathcal{M}(x(t_{k-1}), \theta) \tag{3.1}$$

where $x(t_k) \in \mathbf{R}^{n_m}$ denotes the vector of dynamical variables at time t_k and θ denotes the vector of poorly known parameters. Usually, the dynamic operator $\mathcal{M} : \mathbf{R}^{n_m+n_\theta} \rightarrow \mathbf{R}^{n_m}$ is nonlinear and deterministic (n_m and n_θ are the dimensions of the spaces where the variables takes values). So, we are working under the assumption of a perfect model. At each time step t_k, the relationship between measured data $d^{obs}(t_k)$ and state variables $x(t_k)$ can be described by a nonlinear operator $\mathcal{H}_k : \mathbf{R}^{n_m+n_\theta} \rightarrow \mathbf{R}^{n_{d^{obs}}}$. If we assume that observations are imperfect the simulated measurements are described by

$$d^{obs}(t_k) = \mathcal{H}_k(x(t_k), \theta) + v(t_k) \tag{3.2}$$

where $v(t_k)$ is the observation error with $v(t_k) \sim \mathcal{N}(0, \mathbf{R}(t_k))$. The most important properties of the system appear in the model equations as parameters (static variables).

In theory parameters of the system can be estimated directly from measurements. In practice, direct measuring the parameters of a complex system is difficult because of sampling, technical and resource requirements. Data assimilation however provides a powerful methodology for parameter estimation.

Given a dynamical model with initial and boundary conditions and a set of measurements that can be related to the model state, the *state estimation* problem is defined as finding the estimate of the model state that fits the best (under a certain criterion) the model equations, the initial and boundary conditions, and the observed data. The *parameter estimation* problem is different from the state estimation problem. Traditionally, in parameter estimation we want to improve estimates of a set of poorly known model parameters leading to a better model solution that is close to the measurements. Thus, in this case we assume that all errors in the model equations are associated with uncertainties in the selected model parameters. The model initial conditions, boundary conditions, and the model structure are all considered to be known. Hence, for any set of model parameters the corresponding solution is found from a single forward integration of the model. One way to solve both problems is to define a cost function that measures the distance between the model prediction and the observations plus a term measuring the deviation of the parameter values from a prior estimate of the parameter values. The relative weight between these two terms is determined by the prior error statistics for the measurements and the prior parameter estimate.

$$\begin{aligned} J(x) = \frac{1}{2}\sum_{k}[d_{obs}(t_k) - \mathcal{H}_k(x(t_k))]^T \mathbf{R}(t_k)^{-1}[d_{obs}(t_k) - \mathcal{H}_k(x(t_k))] \\ + \frac{1}{2}[x_p(t_0) - x(t_0)]^T \mathbf{C}_0^{-1}[x_p(t_0) - x(t_0)] \end{aligned} \tag{3.3}$$

These estimation problems are hard to solve due to nonlinear dynamics of the system as well as the observational model and due to the presence of multiple local minima in the cost function.

The schemes for solving the state and parameter estimation have different backgrounds. They often belong to either control theory or estimation theory. The control theory uses variational assimilation approaches to perform a global time and space adjustment of the model solution to all observations. The goal is to minimize the cost function (3.3) penalizing the time-space misfits between the observed data and predicted data, with the model equations and their parameters as constraints (Talagrand and Courtier 1987). Results depend on the a priori control weights and penalties added to the cost function. The dynamical model can be either considered as a strong (perfect model) or weak constraint (in the presence of model error). The variational data assimilation is a constrained minimization problem, which is solved iteratively with a gradient-based optimization method. First, the problem is reformulated as an unconstrained minimization

problem by adding the model equations and constraints to the objective function. The gradients are then obtained using a so-called adjoint method that allows us to calculate all the sensitivities by only two simulations: one backward using the adjoint and one forward in time.
The estimation theory uses a statistical approach to estimate the state of a system by combining all available reliable knowledge about the system including measurements and models. This falls in the Bayesian estimation territory, where we find the Kalman filter approach (Kalman [1960]) as a simplification for the case of linear systems. For linear models the Kalman Filter is the sequential, unbiased, minimum error variance estimate based on a linear combination of all past measurements and dynamics. Kalman filtering is formulated as sequential estimation procedure, i.e. such that the data are assimilated whenever they become available. The end results will be

- $x^a(t_k)$ is the optimal estimate of $x^t(t_k)$ using $d^{obs}(t_1) \dots d^{obs}(t_k)$
- $\mathbf{C}^a(t_k)$ is the covariance matrix of the estimation error,

where the superscript "a" denotes the analyzed state and covariance. These two are obtained by following the classical Kalman filter equations

- The first step, ***initialization***, is specification of an initial distribution for the true state

$$x^t(t_0) \sim \mathcal{N}(x^f(t_0), \mathbf{C}^f(t_0)) \tag{3.4}$$

- The second step, ***forecast step***, is to specify the error between the true state $x^t(t_{k+1})$ and the model forecast $M(t_k)x^t(t_k)$ which should be described in terms of Gaussian distribution

$$x^t(t_{k+1}) = \mathbf{M}(t_k)x^t(t_k) + \eta(t_k) \tag{3.5}$$

  where $\eta(t_k) \sim \mathcal{N}(0, \mathbf{Q}(t_k))$ and $\mathbf{M}$ represents the Tangent Linear Model (TLM) of $\mathcal{M}$.
  Giving the stochastic model 3.5 and the initial condition 3.4, the Kalman filter is able to compute the true state at every time in the future

$$\begin{aligned} x^f(t_{k+1}) &= E[x^t(t_{k+1})] = \\ &= \mathbf{M}(t_k)x^f(t_k) \\ \mathbf{C}^f(t_{k+1}) &= E[(x^t(t_{k+1}) - x^f(t_{k+1}))(x^t(t_{k+1}) - x^f(t_{k+1}))^\top] = \\ &= \mathbf{M}(t_k)\mathbf{C}^f(t_k)\mathbf{M}(t_k)^\top + \mathbf{Q}(t_k) \end{aligned} \tag{3.6}$$

- The third step is ***the analysis step***.

$$
\begin{aligned}
x^a(t_k) &= x^f(t_k) + \mathbf{K}(t_k)(d^{obs}(t_k) - \mathbf{H}(t_k)x^f(t_k)) && (3.7)\\
\mathbf{C}^a(t_k) &= E[(x^t(t_k) - x^a(t_k))(x^t(t_k) - x^a(t_k))^\top] = \\
&= (\mathbf{I} - \mathbf{K}(t_k)\mathbf{H}(t_k))\mathbf{C}^f(t_k)(\mathbf{I} - \mathbf{K}(t_k)\mathbf{H}(t_k))^\top \\
&+ \mathbf{K}(t_k)\mathbf{R}(t_k)\mathbf{K}(t_k)^\top && (3.8)
\end{aligned}
$$

where the choice for K is the minimal-variance gain and $\mathbf{H}$ represents the Tangent Linear Model (TLM) of $\mathcal{H}$

$$\mathbf{K}(t_k) = \mathbf{C}^f(t_k)\mathbf{H}(t_k)^\top[\mathbf{H}(t_k)\mathbf{C}^f(t_k)\mathbf{H}(t_k)^\top + \mathbf{R}(t_k)]^{-1} \tag{3.9}$$

It can be shown that for linear systems, and assuming Gaussian measurement and model noise, this sequential approach results in exactly the same answers as the variational method. That means that $x^a = \min_x J(x)$ (see eq. 3.3) is the same as the Best Linear Unbiased Estimator (BLUE) from eq.3.7 calculated in the update step of the Kalman filter equations.

3.2 Ensemble methods for nonlinear filtering problems

The classical Kalman filter technique is optimal in case of linear models and Gaussian noise. In reality, the models describing complex physical phenomena tend not to be linear. Therefore, ensemble Kalman Filter (EnKF, Evensen 1997) was introduced and became one of the most promising History Matching methodology. This is a Monte Carlo technique where the probability density of the state is represented by an ensemble of possible realizations that are simultaneously propagated in time by the non-linear model and updated sequentially when observations become available. One of the issues that temper with the quality of the EnKF is the finite number of ensemble members. In the literature (Evensen 1997, 2003, 2006, Aanonsen et al. 2009) a unified opinion was formed, agreeing that 100 ensemble members are sufficiently enough to keep the EnKF computationally feasible, without sacrificing the quality of the updates. Therefore, EnKF represent a solution for bypassing the need for a linear model in the Kalman filter framework. However, it cannot overcome the constraint on Gaussian distribution for the errors. Regardless of the distribution of the prior uncertainties, EnKF has the tendency to provide approximations to the posterior that are too close to a Gaussian distribution.

Particle filters (PF) represent a solution for the non-Gaussian assumptions on the errors statistics. It belongs to the same ensemble based approaches. In comparison with the

EnKF, where the ensemble members are equally probable (equal weights), in case of PF there is a weight associated with each ensemble member (Doucet 2001). The sum of these weights is one. Even if the philosophy in the back of the PF is the same as in the EnKF approach (the ensemble is updated using the Bayes' rule), the update step differs. In case of PF's the weights are the ones updated by the observations sequentially. It is mathematically proven that the particle filter is the only data assimilation scheme that produces a sample from the exact posterior distribution. Nevertheless, there is a limitation of this methodology related with the computational time needed when applied to large-scale application. Moreover, for a high dimensional state vector, this methodology suffers from phenomenon known in literature as the curse of dimensionality, i.e. as the dimension of the system increases the largest of the sample weights converges to one in probability (Bengtsson 2008). The consequence is called filter degeneracy, where the posterior distribution is represented by a single point in the state space.

A logical step is to combine the strengths of the two above-mentioned approaches: the computationally feasibility of the EnKF with the sampling procedure of the PF. The result is a hybrid filter, which uses the Kalman filter update step and the weighted correction. Examples of these kinds of filters are EnKF-SIS (Mandel 2009) and Gaussian mixture filters (Bengtsson 2003, Hoteit 2008a). In the later ones, a mixture density (Silverman 1986) approximates the prior distribution where each ensemble member forms the center of a Gaussian density function (a kernel). The mixture densities together with the weights are propagated by the dynamical model and updated accordingly using the Bayes' rule. Therefore, the mixture filters in their basic form, are also prone to weight degeneration. Hoteit 2008b showed that the EnKF can be reformulated as a mixture filter, by using Gaussian kernels and forcing the weights to be uniform. The latter requirement can be made more flexible. Following this idea, Stordal et al. 2012 introduced a tuning parameter $\alpha \in [0, 1]$ such that when $\alpha = 0$ one obtains the EnKF equally distributed weights, and $\alpha = 1$ one obtains the weights of the Gaussian mixture filter. Hence, taking α to be small the weight degeneracy problem is reduced, but taking $\alpha > 0$ EnKF approximation of the posterior is improved. Consequently, we obtain a better preservation of the non-Gaussian features of the marginal distributions. The proposed approach is adaptive, in the sense that an optimal α is sought at each assimilation step resulting in an adaptive Gaussian mixture filter (AGM).

3.2.1 The nonlinear filtering problem

The nonlinear filtering problem consist of estimating sequentially in time the state, $\mathbf{x_t}$, of a nonlinear dynamical system conditioned on some noisy measurements taken on the state. For the simplicity of our notations in this section we will denote $x(t_k)$ by x_t, $x(t_{k-1})$ by x_{t-1}, $d^{obs}(t_k)$ by d_t^{obs}. Let us consider the following nonlinear problem:

$$\begin{aligned} x_t &= \mathcal{M}_t(x_{t-1}) + \eta_t, \\ d_t^{obs} &= \mathbf{H}_t x_t + v_t, \end{aligned} \tag{3.10}$$

We consider the model error η_t known, given by its probability density function, the initial distribution given by its probability density function $p(x_0)$ and the measurement error which is assumed Gaussian $\mathcal{N}(0, \mathbf{R}_t)$. The solution of the filtering problem is given by determining the posterior density $p(x_t|d_{1:t}^{obs})$. That is, the conditional distribution of the state x_t given the observations $d_{1:t}^{obs}$, where $d_{1:t}^{obs} = (d_1^{obs}, ..., d_t^{obs})$. At each moment of time t, the state of the system is defined by two conditional probability density functions: the prior density $p(x_t|d_{1:t-1}^{obs})$ and the posterior density $p(x_t|d_{1:t}^{obs})$. At the initial moment, where no measurements are available, the probability density function $p(x_0)$ is considered the posterior. Given the posterior density at time $t-1$, the prior density at time t can be calculated using the Chapman-Kolmogorov equation (Jazwinsky 1970)

$$p(x_t|d_{1:t-1}^{obs}) = \int p(x_t|x_{t-1}) p(x_{t-1}|d_{1:t-1}^{obs}) dx_{t-1}. \tag{3.11}$$

When a new observation d_t^{obs} occurs, the posterior density is calculates using the Bayes inversion scheme:

$$p(x_t|d_{1:t}^{obs}) = \frac{p(d_t^{obs}|x_t) p(x_t|d_{1:t-1}^{obs})}{p(d_t^{obs}|d_{1:t-1}^{obs})}, \tag{3.12}$$

where the term $p(d_t^{obs}|d_{1:t-1}^{obs})$ can be calculated using the relation:

$$p(d_t^{obs}|d_{1:t-1}^{obs}) = \int p(d_t^{obs}|x_t) p(x_t|d_{1:t-1}^{obs}) dx_t.$$

However, the probability density functions $p(x_t|d_{1:t}^{obs})$ (eq.3.12) cannot be computed analytically. Consequently, we need to find approximate solutions for these probability density functions. A simple approach is the use of a Monte Carlo methodology, where the prior and posterior densities are approximated with a random sample. Having to approximate two types of densities (the prior densities and the posterior densities), we use different notations for samples. We denote by the set of particles $\{x_t^i\}_{i=1}^N$ a sample approximating the prior density (3.11) and by the set of particles $\{\hat{x}_t^i\}_{i=1}^N$ a sample approximating the posterior density (3.12).

3.2.2 The ensemble Kalman filter

The ensemble Kalman filter EnKF (Evensen 1994) is a sequential Monte Carlo methodology introduced in the Kalman filter framework that aims to provide approximate solutions to the nonlinear filtering problem.
For a given a sample, $\{\hat{x}^i_{t-1}\}_{i=1}^N$, from the posterior density $p(x_{t-1}|d^{obs}_{1:t-1})$, a sample from the prior density $p(x_t|d^{obs}_{1:t-1})$ is obtained in the EnKF methodology by sampling from the forward density $p(x_t|\hat{x}^i_{t-1})$ using eq. (3.10), for each $i \in \{1 \dots N\}$.

$$x^i_t = \mathcal{M}_t(\hat{x}^i_{t-1}) + \eta^i_t, \quad i = 1, \dots, N. \tag{3.13}$$

where η^i_t, for $i \in \{1, \dots, N\}$ is a sample from $\mathcal{N}(0, Q_t)$.
If the density $p(x_t|d^{obs}_{1:t-1})$ is Gaussian with mean μ_t and covariance matrix $\mathbf{C}_t$, it follows that all the densities from equation 3.12 are Gaussian. First, the observation operator has Gaussian error, therefore the density
$p(d^{obs}_t|x_t) = p_{v_t}(d^{obs}_t - \mathbf{H}(t)x_t)$ (Jazwinsky 1970) is Gaussian. Secondly, the normalizing term can be written as:

$$d^{obs}_t|d^{obs}_{1:t-1} = \mathbf{H}(t)x_t|d^{obs}_{1:t-1} + v_t \tag{3.14}$$

being Gaussian.

Consequently, the mean $\hat{\mu}_t$ and covariance matrix $\hat{\mathbf{C}}_t$ of the conditional variable $x_t|d^{obs}_{1:t}$ are given by the standard Kalman filter update,

$$\begin{aligned} \hat{\mu}_t &= \mu_t + \mathbf{K}_t(d^{obs}_t - \mathbf{H}_t\mu_t), \\ \hat{\mathbf{C}}_t &= (\mathbf{I} - \mathbf{K}_t\mathbf{H}_t)\mathbf{C}_t, \\ \mathbf{K}_t &= \mathbf{C}_t\mathbf{H}_t^T(\mathbf{H}_t\mathbf{C}_t\mathbf{H}_t^T + \mathbf{R}_t)^{-1}. \end{aligned} \tag{3.15}$$

Using the relation 3.13 an ensemble $\{x^i_t|i \in \{1, \dots, N\}\}$ of samples from the prior density $p(x_t|d^{obs}_{1:t-1})$ is obtained. The mean calculated from this ensemble is an estimator for μ_t and is given by the following formula:

$$\mu_t \approx \overline{x}_t = \frac{1}{N}\sum_{i=1}^N x^i_t, \tag{3.16}$$

If $\{\hat{x}^i_t\}_{i=1}^N$ are samples from the posterior density $p(x_t|d^{obs}_{1:t})$, then

$$\hat{\mu}_t \approx \frac{1}{N}\sum_{i=1}^N \hat{x}^i_t \tag{3.17}$$

Combining eq. 3.15 and eq. 3.16, we obtain:

$$\hat{\mu}_t = \mu_t + \mathbf{K}_t(d^{obs}_t - \mathbf{H}_t\mu_t) \approx \frac{1}{N}\sum_{i=1}^N (x^i_t + \mathbf{K}_t(d^{obs}_t - \mathbf{H}_t x^i_t)) \tag{3.18}$$

If we replace the covariance $\mathbf{C}_t$ with the empirical covariance matrix $\overline{\mathbf{C}}_t$ (calculated from the ensemble $\{x^i_t|i \in \{1, \dots, N\}\}$) in the calculus of the Kalman gain $\mathbf{K}_t$ (eq.

3.15), from the equations 3.17 and 3.18 we find the update of the particle i from the original form of the EnKF (Evensen 2007).

$$\hat{x}_t^i = x_t^i + \mathbf{K}_t(d_t^{obs} - \mathbf{H}_t x_t^i), \quad i = 1, \ldots, N. \tag{3.19}$$

In the same time, considering the stochastic nature of the observation operator, we can modify the relation 3.18 by introducing an additive noise in the relation:

$$\hat{\mu}_t = \mu_t + \mathbf{K}_t(d_t^{obs} - \mathbf{H}_t\mu_t) \approx \frac{1}{N}\sum_{i=1}^{N}(x_t^i + \mathbf{K}_t(d_t^{obs} - \mathbf{H}_t x_t^i + \epsilon_t^i)) \tag{3.20}$$

where ϵ_t^i, for $i \in \{1, \ldots, N\}$ is a sample from $N(0, \mathbf{R}_t)$. The relation 3.20 stands because $\sum_{i=1}^{N} \epsilon_t^i \approx 0$. From the equations 3.17 and 3.20 we find the update of the particle i form the EnKF with correction introduced by Burgers et al (1998).

$$\hat{x}_t^i = x_t^i + \mathbf{K}_t(d_t^{obs} - \mathbf{H}_t x_t^i + \epsilon_t^i), \quad i = 1, \ldots, N. \tag{3.21}$$

This small modification increases the variability within ensemble, and is used now in the EnKF methodology.

For a general linear system (linearity for both operators $\mathcal{M}$ and $\mathcal{H}$) with Gaussian statistics, at the limit (when the number of the ensemble members goes to infinite), the particles from the analyzed state are samples from the true posterior distribution $p(x_t|d_{1:t}^{obs})$ (Evensen 2007). For a general nonlinear system, the updated ensemble members will not represent a sample from the true posterior distribution even in the limit of an infinite number of ensemble members. However, despite all non-linearities, the EnKF is able to capture aspects from the posterior distribution.

3.2.3 The EnKF methodology for state and parameter estimation

In practice, the relation between the dynamical variables and the observations are far from being linear. Moreover, the system may depends on some uncertain parameters (usually denoted by θ). Consequently, let us consider the dynamical system described by

$$x_t = \mathcal{M}(x_{t-1}, \theta) + \eta_t, \tag{3.22}$$

where $\eta_t \sim \mathcal{N}(0, Q_t)$. In reservoir engineering the model is considered perfect, so the term η_t vanishes. We further consider a nonlinear relation between the observed data and the dynamical variables and parameters:

$$d_t^{obs} = \mathcal{H}_t(x_t, \theta) + v_t, \tag{3.23}$$

where $v_t \sim \mathcal{N}(0, \mathbf{R}_t)$. To be able to apply the EnKF methodology, first we need to set a linear observation operator for the observed data. For that, a trick is made by augmenting the state the dynamical variables with the parameters and the simulated data (d^{sim}). Here

simulated data is the response of the observation operator given the dynamical variables and the parameters: $d_t^{sim} = \mathcal{H}_k(x_t, \theta)$. Consequently, the augmented state variable denoted by X_t is defined as:

$$X_t = \begin{bmatrix} x_t \\ \theta \\ d_t^{sim} \end{bmatrix} \in \mathbf{R}^{n_m+n_\theta+n_d} \tag{3.24}$$

For this "new" variable, the stochastic dynamical model can be written as

$$X_t = \overline{\mathcal{M}}(X_{t-1}) + \overline{\eta_t} = \begin{bmatrix} \mathcal{M}(x_{t-1}, \theta) \\ \theta \\ \mathcal{H}_t(x_t, \theta) \end{bmatrix} + \begin{bmatrix} \eta_t \\ 0 \\ 0 \end{bmatrix} \tag{3.25}$$

The relation between the state variable and the measurements becomes linear

$$d_t^{obs} = \mathbf{H}_t X_t + v_t = \begin{bmatrix} 0 & 0 & \mathbf{I} \end{bmatrix} X_t + v_t = \begin{bmatrix} 0 & 0 & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathcal{M}(x_{t-1}, \theta) + \eta_t \\ \theta \\ \mathcal{H}_t(x_t, \theta) \end{bmatrix} + v_t \tag{3.26}$$

The methodology can be summarized as follows:

- *Initialization*
  At the initial phase we generate N independent samples, denoted θ_0^i, from the prior distribution of the parameters and N independent samples, denoted x_0^i, from the initial distribution of the dynamical variable ($i \in \{1, \ldots, N\}$). These samples generate the initial ensemble $\{X_1, X_2, \ldots, X_N\}$. For a consistent description we define $\theta_0 = \theta^a$ and $x_0 = x^a$

- *Forecast step*
  Each particle (ensemble member) is updated by the stochastic dynamical model towards the next assimilation step using eq. 3.25. For that, we generate N independent samples, denoted η_t^i from the Gaussian distribution $\mathcal{N}(0, Q_t)$, and each ensemble member is changed according to

$$X_t^{f,i} = \begin{bmatrix} x_t^{f,i} \\ \theta_t^{f,i} \\ d_t^{sim,f,i} \end{bmatrix} = \begin{bmatrix} \mathcal{M}(x_{t-1}^{a,i}, \theta_{t-1}^{a,i}) + \eta_t^i \\ \theta_{t-1}^{a,i-1} \\ \mathcal{H}_t(x_t^{f,i}, \theta_{t-1}^{a,i}) \end{bmatrix} \tag{3.27}$$

  One may observe that at this stage the parameter values are not changed by the forward model.

- *Update (analysis) step*
  Each ensemble member is modified, due the assimilation of the observed data according to the Kalman equations. The observed data is perturbed with noise

sampled from the Gaussian distribution $\mathcal{N}(0, \mathbf{R}_t)$. This extra perturbation is necessary to avoid the ensemble collapse (Burgers et al 1998).

$$X_t^{a,i} = X_t^{f,i} + \mathbf{K}_t(d_t^{obs} - \mathbf{H}_t X_t^{f,i} + \epsilon_t^i), \quad i = 1, \ldots, N. \tag{3.28}$$

In the Kalman gain $\mathbf{K}_t = \mathbf{C}_t^f \mathbf{H}_t^T (\mathbf{H}_t \mathbf{C}_t^f \mathbf{H}_t^T + \mathbf{R}_t)^{-1}$, the covariance matrix of the forecasted state, $\mathbf{C_t^f}$, is calculated directly, using the particles from the forecasted ensemble. For that, all the ensemble members are augmented within a matrix, with $n_m + n_\theta + n_d$ rows and N columns, denoted X:

$$X_t^f = \left[\begin{array}{cccc} X_t^{f,1} & X_t^{f,2} & \ldots & X_t^{f,N} \end{array} \right] \tag{3.29}$$

Then, the covariance matrix of the forecasted state can be calculated as

$$\mathbf{C}_t^f = \frac{(X_t^f - \overline{X_t}^f \mathbb{1}_{1,N})(X_t^f - \overline{X_t}^f \mathbb{1}_{1,N})^T}{N-1} \tag{3.30}$$

where, $\overline{X_t}^f = \frac{1}{N} \sum_{i=1}^N X_t^{f,i}$ and $\mathbb{1}_{1,N}$ is a row matrix with N columns with each element equal to 1.

3.2.4 The adaptive Gaussian mixture (AGM) and its iterative variant (IAGM)

The quality of the updated ensemble obtained after the use of the EnKF methodology depends on the type of initial distribution of the dynamical variables and the prior distribution of the parameters. Referring to quality, we mean how good the updated ensemble captures the posterior distribution. In the same time, the non-linearity strongly affects this quality, especially when the prior is non-Gaussian (Zafary and Reynolds 2007). One ensemble based method that has shown promise on nonlinear models is the adaptive Gaussian mixture filter (AGM, Stordal et al. 2011).

The AGM replaces the Gaussian assumption of the prior distribution $p(x_t|d_{1:t-1}^{obs})$ with the assumption that this distribution is described by a Gaussian mixture of type:

$$p(x_t|d_{1:t-1}^{obs}) \sim \sum_{i=1}^N w_{t-1}^i \mathcal{N}(x_t - x_t^i, \mathbf{C}_t) \tag{3.31}$$

The weights $\{w_i | i \in \{1, \ldots, N\}\}$ are positive and sum to 1. The model carry the initial ensemble towards the first assimilation time, when the Gaussian mixture is defined by equal weights of $\frac{1}{N}$, where N is the number of the ensemble members. The AGM introduces a tuning parameter, denoted $h \in [0, 1]$, by which the covariance matrix from the Gaussian mixture (3.31) is defined as $\mathbf{C}_t = h^2 \overline{\mathbf{C}}_t$. The matrix $\overline{\mathbf{C}}_{t_1}$ is the empirical covariance matrix calculated from the ensemble $\{x_{t_1}^1, \ldots, x_{t_1}^N\}$. Following the same approach as in the EnKF case, if we assume the Gaussian mixture distribution for

$p(x_t|d_{1:t-1}^{obs})$, taking into account the Gaussian nature of the measurement error, it comes out that the posterior $p(x_t|d_{1:t}^{obs})$ has a Gaussian mixture distribution of type:

$$p(x_t|d_{1:t}^{obs}) \sim \sum_{i=1}^{N} w_t^i \mathcal{N}(x_t - \hat{x}_t^i, \hat{\mathbf{C}}_t) \tag{3.32}$$

where the particles are updated according to:

$$\begin{aligned} \hat{x}_t^i &= x_t^i + \mathbf{K}_t(d_t^{obs} - \mathbf{H}_t x_t^i + \epsilon_t^i), \\ \mathbf{K}_t &= \mathbf{C}_t\mathbf{H}_t^T(\mathbf{H}_t\mathbf{C}_t\mathbf{H}_t^T + \mathbf{R}_t)^{-1}, \\ \hat{\mathbf{C}}_t &= (\mathbf{I} - \mathbf{K}_t\mathbf{H}_t)\mathbf{C}_t. \end{aligned} \tag{3.33}$$

The weights are updated according to

$$\begin{aligned} \overline{w}_t^i &= w_{t-1}\Phi(d_t^{obs} - \mathbf{H}_t x_t^i, \mathbf{H}_t\mathbf{C}_t\mathbf{H}_t^T + \mathbf{R}_t), \\ w_t^i &= \frac{\overline{w}_t^i}{\sum_{i=1}^N \hat{w}_t^i}, \end{aligned} \tag{3.34}$$

Here, the function $\Phi(x - \mu, \mathbf{C})$ represents multivariate Gaussian density with mean μ and covariance matrix $\mathbf{C}$. To avoid filter degeneracy that occurs in high dimension and complex systems a weight interpolation is introduced

$$\begin{aligned} \hat{w}_t^i &= \alpha_t w_t^i + (1 - \alpha_t)N^{-1}, \quad \alpha_t \in [0, 1], \\ \alpha_t &= N^{-1}(\sum_{i=1}^{N}(w_t^i)^2)^{-1}. \end{aligned} \tag{3.35}$$

The implementation of AGM is very simple especially if EnKF is already implemented. Let us consider the state and parameter estimation problem of the model defined by eq. 3.22 and eq. 3.23. We denote the ensemble of parameter fields by $\{\theta_i\}_{i=1}^N$. Like in the EnKF methodology, we define an augmented state vector denoted X_t defined as:

$$X_t = \begin{bmatrix} x_t \\ \theta \\ d_t^{sim} \end{bmatrix} \in \mathbf{R}^{n_m + n_\theta + n_d} \tag{3.36}$$

The linear measurement operator is defined by the eq. 3.26. As in the EnKF, we denote by $\mathbf{C}_t^f$ the sample covariance matrix of $\{X_t^{f,i}\}_{i=1}^N$, but calculated with respect to the weights. That is, for its calculation we use eq. 3.30, but with the mean calculated with respect to the weights:

$$\overline{X_t}^f = \sum_{i=1}^{N} w_{t-1}^i X_t^{f,i} \tag{3.37}$$

Initially, the weights are considered having uniform values of $1/N$ and afterwards they are updated using eq 3.35. The *Initialization* and the *Forecast step* are the same as for

the EnKF. At each assimilation time, the augmented state vector is updated in the AGM (and IAGM) for each $i = 1, \ldots, N$ as

$$X_t^{a,i} = X_t^{f,i} + \mathbf{C}_t^f \mathbf{H}_t^T (\mathbf{H}_t \mathbf{C}_t^f \mathbf{H}_t^T + h^{-2}\mathbf{R}_t)^{-1} \left(y_t - \mathbf{H}_t X_t^{f,i} + \epsilon_t^i \right), \tag{3.38}$$

where ϵ_t^i is a sample from the Gaussian measurement error distribution $\mathcal{N}(0, \mathbf{R})$. We notice that the eq. 3.38 can be written as

$$X_t^{a,i} = X_t^{f,i} + (h^2 \mathbf{C}_t^f) \mathbf{H}_t^T (\mathbf{H}_t (h^2 \mathbf{C}_t^f) \mathbf{H}_t^T + \mathbf{R}_t)^{-1} \left(y_t - \mathbf{H}_t X_t^{f,i} + \epsilon_t^i \right), \tag{3.39}$$

which is exact the first relation from eq. 3.33. Note that the only difference with a standard EnKF update is the scaling h^{-2} of the measurement error covariance matrix $\mathbf{R}$. In other words the linear update is dampened where the dampening factor $h \in [0, 1]$ is the bandwidth of the Gaussian mixture (see Stordal et al 2011).

To further improve the AGM on nonlinear models an iterative version, namely the iterative adaptive Gaussian mixture (IAGM) was introduced in Stordal et al 2012. The new approach IAGM runs AGM iteratively, with a tuning parameter $h \in [0, 1]$. At the end of an assimilation period, performed with the AGM, an new ensemble is generated by sampling from the Gaussian mixture defined by the samples at the end of previous iteration (analyzed state). That is, the sampling is performed from

$$\sum_{i=1}^{N} w_{j-1}^i \mathcal{N}(X - X_{j-1}^{a,i}, h^2 \mathbf{C}_{j-1}^a) \tag{3.40}$$

Consequently, the sampling has the form:

$$X_j^i = X_{j-1}^{i,a} + \frac{h}{\sqrt{N-1}} (X^a - \overline{X}^a \mathbb{1}_{1,N}) \xi_i, \quad i = 1, \ldots, N. \tag{3.41}$$

where ξ_i is multivariate Gaussian with zero mean and covariance matrix $\mathbf{I}_N$. To be completely rigorous, the ratio between the old and new prior for each ensemble member should then be evaluated, in order to correct the weights assigned to each ensemble member. However, this would require the estimation of the prior in addition to evaluation of weights in high dimensions. With the weight reduction leading to almost uniform weights due to the dimension, the work is a lot more than the gain. Consequently, we are using uniform weights after resampling. With the new initial ensemble the AGM is then performed. In most applications only a few iterations, typically two or three, are required.

Chapter 4

A probabilistic parametrization for geological uncertainty estimation using the Ensemble Kalman Filter (EnKF)

4.1 Introduction

The term facies was introduced in 1837 by the Swiss geologist Amanz Gressly and signifies a body of rock with specified characteristics, which formed under certain condition of sedimentation. The facies types are distinguished either by the petrophysical properties (permeability, porosity, grain size, mineralogy): **lithofacies**, or by the fossil content: **biofacies**.

The estimation of the (litho)facies distribution into a reservoir domain has two goals; one is related with the high predictive capacity of flow measurements of a reservoirs that are geologically realistic and the other one is related with the need of accurate properties to predict the consequences of changing conditions in the reservoir lifetime (e.g. the response to enhance recovery). The first step in the assessment of the geologic continuities in the subsurface is to define a geological model to simulate plausible facies distributions that are consistent with the prior knowledge about subsurface geology (numbers of facies type that occur, facies proportion, core information, type of patterns etc). The geologi-

This chapter is full reference of the article published in Computational Geosciences, Volume 17, Issue 5, pp 813-832, 2013

cal model can be constructed using various methods such as an object-based simulation technique (Deutsch and Journel 1998, Deutsch and Wang 1996) or a pixel-based simulation technique such as Multiple-Point Geostatistique (Guardiano and Srivastava 1993, Strebelle 2002, Caers and Zang 2004). Another approach is the parametrization of the facies maps with some uncertain variables (parameters), through which we are able to adjust them such that these will honor the underlying knowledge of subsurface geology. These parameters can be calibrated (hence, also the maps) using additional information about flow measurements within a process named in reservoir engineering, history matching (HM). A comprehensive description about HM methods is given by Oliver et al. (2008).

One of the most promising HM method is the Ensemble Kalman Filter (EnKF) (Evensen 1997, 2003, 2006, Aanonsen et al. 2009) which is a Monte Carlo technique introduced in the Kalman filtering framework where the probability density of the state is represented by an ensemble of possible realizations that are simultaneously updated. For the estimation of the spatial distribution of the facies using EnKF as HM method, various approaches were proposed, depending on the particularity of geological characterization (number of the occurring facies, type of transition between facies types).

Jafarpour and Khodabakhshi (2011) have used for the geological model simulation a Multi-point Geostatistical tool (snesim, Strebelle 2002, 2003) through which the initial ensemble is simulated from a training image. Then the EnKF is performed to update the log-permeability field and after that, a probability map is constructed for each facies type, based on a function defined in the log-permeability space that converts the mean log-permeability field to a probability map. At each assimilation time step a new ensemble is generated, using SNESIM, conditioning the training image to the probability map previously defined. For the case with three facies types, the study uses a training image in which the facies types with lowest and highest permeability do not have direct contact.

Lorentzen et al. (2011) proposed for the facies maps simulation a geological simulation model based on distances to facies boundaries defined with level set functions (Osher and Fedkiw (2003)). For each facies type a signed distance function is defined as the minimum euclidian distance from each grid cell to boundaries of that facies type. The sign is positive if the grid cell is in the facies type domain, is negative if the grid cell is outside and is zero if the grid is on the boundary of that facies type. After the EnKF update of the distances, the values obtained are no longer distances. An extra procedure is performed in order to transform these updated values to distances and, after that, the updated facies maps are constructed based on a maximization criterion applied to the transformed values.

Another application of the level set method applied for geometry with 3 facies is presented in the work of Chang et al. (2010), but here, the values of the level set functions

in some predefined nodes are treated as random variables with known distribution and updated in the EnKF process. Then, using a interpolation procedure the level set function values are calculated in every grid cell. For the case with 3 facies types, two level set functions are used and the facies maps are constructed based on sign combination of those function values.

In our study, for the simulation of the facies maps we use a truncation plurigaussian method. This method was introduced by Galli et al. (1994) and further developed by Le Loc'h (1994) and Le Loc'h and Galli (1997). This method is a projection from a continuous space (the space determined by the Gaussian random fields (GRF)) into a discrete space (the facies maps space) through a map designed in the GRF space and defined by intersection of some thresholds (values for one GRF, curves for two GRF, surfaces for three GRF, etc.). In the paper of Le Loc'h et al. (1994), two Gaussian random fields were truncated using a rectangular map. They showed various resulting maps changing the geostatistical characteristics of the GRF (variogram type, isotropy, anisotropy, range correlation, principal direction, etc).

Xu et al (2006) propose a "flexible true plurigaussian code for facies simulation", introducing a binary dynamic contact relation matrix (DCM), with which the authors define the contact relations among facies types. For a bi-dimensional GRF space (bi-Gaussian truncation) they have used rectangular truncation maps with different number of facies types occurring. For three-Gaussian truncation scheme, the maps were constructed based on intersection of some orthogonal planes defined by the partition of each Gaussian field.

In the truncation plurigaussian context, Liu and Oliver (2005) have used for the facies simulation model a truncation map (in a bi-Gaussian space), constructed using three lines, whose intersections separate the space in seven regions, each having a facies type assigned. In order to adjust the facies boundaries to the correct position they introduced the simulation model into a HM process using two methods. One method is a gradient-based method (with some transformations of the map for consistent computation of the sensitivities) and the other is the EnKF. The parameters that define the truncation lines could not by estimated in the HM process and the threshold lines were calibrated in an trial-error procedure. Also, the authors showed, in a 2005 article, that, for the facies estimations, the EnKF produce better results than the gradient-based method.

Agbalaka and Oliver (2008) have used the same simulation model, to generate facies maps for a 3D reservoir having three horizontal layers, vertically uncorrelated. They introduced the geological model in the EnKF process, together with a localization scheme, defined with the fifth-order compact function of Gaspari and Cohn (1999). The authors used an internal loop in the EnKF process in order to reposition each facies map of the ensemble to the correct facies type in the measurement locations. The reason for doing that is related to the fact that, after the assimilation of the production data, some of the simulated facies maps do not preserve the facies observations at the well locations. This

enforcement on facies constraints affects only a smaller state vector extracted from the initial one, state which contains only the variables that define the geological simulation model.

Zhao et al. (2008) integrate the truncated plurigaussian simulation model, introduced by Liu and Oliver, into an iterative EnKF process, assimilating besides production data, seismic data, for a 3D reservoir model having three layers vertically uncorrelated. For each layer the authors used different assignment of the facies type in the regions of the truncation map and different geostatistical properties for the Gaussian fields.

Agbalaka and Oliver (2009) applied EnKF to facies fields with uncertain petrophysical properties that exhibit non-stationarity in facies proportions. They truncated a single Gaussian field using two thresholds calculated with respect to three predefined facies proportion maps. This model is suitable for fields that contain three facies types that exhibit transition from a facies with smallest permeability to facies with highest permeability. This transition is made through another facies type with medium permeability.

In our paper we investigate the reservoirs having a geology defined by three facies types that occur, with the property that, each two could have a contact. We present a new approach of the truncated plurigaussian method for facies simulations, in the sense that, we define a truncation map which is not introduced ad-hoc, but which appears from the internal construction of the model. This means that, it is not the map which generates the model, but the model generates the map. We use a bi-Gaussian scheme in which the truncation map depends on two parameters that have the ability to control the facies proportions in the model. The change of the truncation parameters does not affect the geometry of the map, nor the assignment of each facies type in the demarcated areas of the map. This bijection between the map parameters and the map geometry enables their introduction in the state vector. The placement of the map parameters in the state ensures a better quantification of the uncertainty and has as a result, a better estimation of the facies distribution in the domain. For the Gaussian fields we use different geostatistical properties, in order to generate complex geologic instances for the facies distribution maps. The projection from a continuous space (the space where the Gaussian fields take values) into a discrete space (or categorial space, the facies maps space) is done through an intermediary space (space with probabilities). We need this intermediary space because we introduce a new observation operator of the facies types at the well locations. If prior to this work, the observation operator of the facies was a binary function (in the terms yes (0) or no (1)) that express the affiliation of the pairs of GRF's to predefined areas of the map, in this study the observation operator measures the probability of occurrence for each facies type in the grids with wells. Therefore, the Gaussian fields are not constrained directly to the map (as in the previous studies), but to the observations which are defined in probabilistic terms. The shape of the truncation map depends on the shape of the truncation function used for the projection of each Gaussian field into

a probabilities field. The function used is piece-wise linearly symmetric with compact support. Also, at the end of the assimilation period we are able to construct a field, as a result of the estimation process, that is an estimator for the truth field (the estimated field). The method is tested using a 2-D reservoir case.
In the next sections we present the description of the geological simulation model, the EnKF implementation for the facies update and three examples where we test the model. For the examples presented, we used different geostatistical properties for the Gaussian fields and different types of uncertainty in parameters and in the construction of the Gaussian fields.

4.2 The Method Description

4.2.1 The "probabilities fields"

In this study our reservoir has a rectangular domain and there is an initial prior knowledge about the characteristics of the subsurface form different sources (i.e. seismic data, geologist interpretation and core analysis). This information relates with the existence of three different types of facies in the reservoir, namely facies type 1, facies type 2, and facies type 3, having the property, that could be a contact between any two. In every well (production or injection) we are certain (perfect observations) about the type of facies present there, so, we can use that information for the whole grid where the well is situated. Let be (i, j) an arbitrary grid block of our reservoir domain where facies type 1 occurs. If we consider $A_k^{i,j}$ the event in which in the grid (i, j) the facies type k occurs, then, in terms of probabilities we have $P(A_1^{i,j}) = 1$, $P(A_2^{i,j}) = 0$, $P(A_3^{i,j}) = 0$ (i.e. the probability that facies type 1 occurs in the grid (i, j) is 1 and the probabilities that the facies type 2 and type 3 occur in the grid (i, j) are 0). Hence, every facies type generates a field defined on the reservoir domain whose values in grids cells are 0 or 1 depending on whether the facies type occurs or not in those locations. The regions occupied by these fields in the reservoir domain represent the true position of each facies type. These fields have binary values (discrete fields) and, in order to estimate them, we will use a new object named "probabilities field". We define a probabilities field as a random field with values in the interval [0,1] having spatial correlations. The reason of using the probabilities fields is related with the fact that, knowing the facies type in the grids cells where wells are located (and, of course, the probabilities in those grids), when we moving away from the wells, the information from the wells dissipates with a certain correlation and become more and more uncertain with the distance from the wells increase. The values of the probabilities fields in a certain grid could be interpreted as confidence levels in the occurrence of the facies types in that grid. In our geological model having three facies types, we need to define three probabilities fields, one for each

facies type. A value of 1 in a grid of one probabilities field means that certainly in this grid the associated facies type occurs, and, consequently, the value of 0 indicates that the facies type does not occurs in the specified location. We choose to model the probabilities fields using projection in [0,1] of Gaussian random fields defined on the reservoir domain, through a projection function. In this way the spatial correlation between the elements of the Gaussian fields are transferred to the elements of the probabilities field. Because in every grid the sum of the probabilities is 1 we will start in this study only with two Gaussian fields corresponding for two facies types, denoted y_1 and y_2 (with $y_k^{i,j}$ having normal distribution) of which projection generates two probabilities fields. The probabilities field associated to the third facies type is defined based on the knowledge of the other two using the fundamental rule of probability. In Figure 4.1 (left) we present a binary field that defines the shapes of a facies type in the reservoir domain (colored in red) and in Figure 4.1 (right) a probabilities field which can be an estimator of that field.

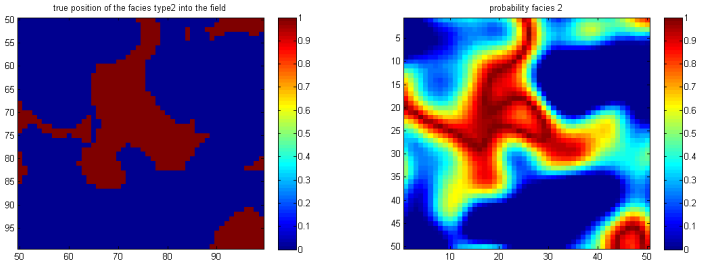


Figure 4.1: The binary field defined by a facies type (left) and the associated probabilities field (right)

The function used for the projection described above is $\varphi_m : \mathbf{R} \to [0,1]$

$$\varphi_m(t) = \begin{cases} \frac{-|t|}{m} + 1 & \text{if } t \in [-m, m] \\ 0 & \text{if } |t| > m \end{cases}, \tag{4.1}$$

where m is a parameter, which we have named truncation parameter (see Figure 4.2). Hence, a probabilities field α is defined as $\alpha = \varphi_m(y)$ (with the convention that $\alpha = \varphi_m(y)$ means $\alpha^{i,j} = \varphi_m\left(y^{i,j}\right)$ in every grid cell (i,j)) where y is a Gaussian field defined on the reservoir domain. In the grids where we have observations about the occurrence of a facies type, the associated Gaussian field will be generated with value 0, such that, his projection (the associated probabilities field) has value 1. This is consistent with the probability of the facies occurrence in observed grids. Using the function decreasing on the intervals [0,m] and [-m,0] the probabilities field will have in the surrounding grids of the observed grids, values that becomes smaller with the increasing

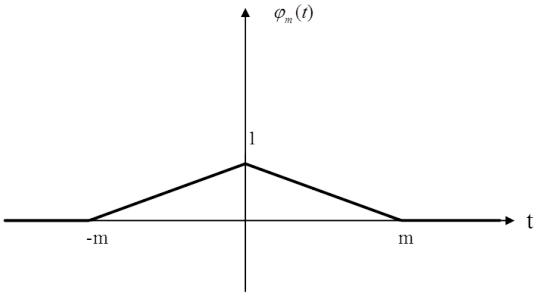


Figure 4.2: The projection function

distance to the observed grid. In addition, the function piece-wise monotony ensures the transfer of the correlation from the Gaussian fields to the probabilities fields. Knowing that, in some areas of the reservoir, a certain facies type does not occurs, we have used a threshold value, denoted m, over which the absolute values of the associated Gaussian field should pass, so that the associated probabilities field to have value 0.

4.2.2 Geological simulation model (Assigning the facies on the grid)

We start with two GRF's y_1 and y_2 defined on the entire reservoir domain and α_1 and α_2 the associated probabilities fields, corresponding to the facies types 1 and 2, where $\alpha_k = \varphi_{m_k}(y_k)_{k=1,2}$. The parameter m_k represent the truncation parameter for the random field y_k. In order to appoint the probabilities field for the third facies we can use the following rule in each grid cell:

$$\alpha_3^{i,j} = \begin{cases} 1 - \left(\alpha_1^{i,j} + \alpha_2^{i,j}\right) & \text{if } 0 \leq \alpha_1^{i,j} + \alpha_2^{i,j} \leq 1 \\ \\ 0 & \text{if } \alpha_1^{i,j} + \alpha_2^{i,j} > 1 \end{cases} \tag{4.2}$$

Thus, we have defined three probabilities fields corresponding to each facies type. At a certain location (i, j) we assign facies of type $k \in \{1, 2, 3\}$ if

$$\alpha_k^{i,j} = max\{\alpha_l^{i,j},\ l = \overline{1,3}\}, \tag{4.3}$$

with the convention that if $\alpha_1^{i,j} = \alpha_2^{i,j} \geq \alpha_3^{i,j}$ we assign facies type 1 and if $\alpha_2^{i,j} = \alpha_3^{i,j} > \alpha_1^{i,j}$ we assign facies type 2. The assignment of a facies type to a grid cells is for the facies type with the highest confidence level of occurrence in that grid. This maximization criterion has the role to draw in each probabilities field the location of the associated facies type, working for each probabilities field as a threshold. Also, it

generates in the bi-dimensional spaces of the probabilities fields and respectively of the Gaussian fields, two different maps that help to place this geological simulation model in the context of the general truncated plurigaussian simulation method. If we represent the point $\left(\alpha_1^{i,j}, \alpha_2^{i,j}\right)$ in the cartesian system $\alpha_1 0 \alpha_2$, presented in Figure 4.3 (left side), we can assign the facies type to the grid (i, j) based on the place where the point $\left(\alpha_1^{i,j}, \alpha_2^{i,j}\right)$ falls in that two dimensional space. The same maximization criterion generates in the space $(y_1\ y_2)$ a truncation map for the Gaussian random fields, which is presented at the right side of the Figure 4.3. This means that, we assign a facies type in a grid (i,j), depending on which zone of the map, the point $\left(y_1^{i,j}, y_2^{i,j}\right)$ belongs. These two maps are equivalent (but in different spaces). The truncation map of the Gaussian fields are uniquely defined by the parameters m_1 and m_2, each combination generating a different scheme with the same geometry and with the same assignment of the facies types in the map zones. The assignments of the facies types in areas of the truncation map is a result of the maximization criterion applied to the triplet of probabilities fields; either, if we refer to the truncation map of the probabilities or to the truncation map of the Gaussian fields. In the approach of Le Loc'h and Galli (1997) the truncation map

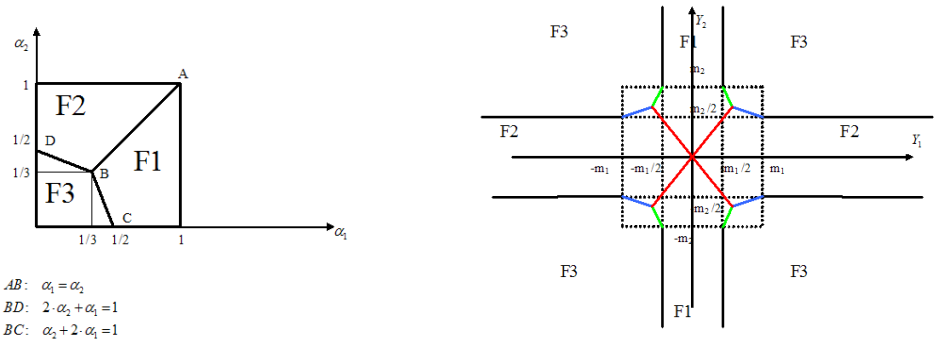


Figure 4.3: Truncation map for probabilities fields (left) and Truncation map for Gaussian random fields (right)

was formed with rectangular shapes (the lines used for intersecting have equations of type $y_1 = a$, $y_2 = b$) and Liu and Oliver (2005) truncation map was build with three intersecting lines having equation of type $y_2 = \tan(\theta - \frac{\pi}{2}) \cdot (y_1 - \frac{r}{\cos(\theta)})$, (the threshold line is perpendicular to the line passing through the origin with the slope θ and intersects the line at a distance r). The truncation map used in this study (Figure 4.3, right) is defined intersecting the curves with equations $\frac{|y_1|}{m_1} = \frac{|y_2|}{m_2}$ (red lines), $\frac{|y_1|}{m_1} + 2 \cdot \frac{|y_2|}{m_2} = 2$ (blue lines), $\frac{|y_2|}{m_2} + 2 \cdot \frac{|y_1|}{m_1} = 2$ (green lines), $|y_1| = \frac{m_1}{2}$, $|y_2| = \frac{m_2}{2}$ (black lines).

4.2.3 Truncation parameters

One of the challenges in the truncated plurigaussian method is the choice of the truncation map. In this study the truncation map is uniquely defined by two parameters (named truncation parameters). These parameters that initial occur in the definition of the projection function, have an important role in the simulation process due to their property to control the facies proportions. In the appendix of this chapter we emphasise the relation between the truncation parameters and the expected facies proportions. Also, the truncation parameters are variable that introduce uncertainty in the model, together with the geostatistical parameters of the Gaussian fields (the isotropy/anisotropy, the range correlations) and the Gaussian fields themselves. In order to quantify the influence of the truncation parameters in the estimation process, in one of the cases, we will introduce these parameters in the state vector and perform a comparison with the experiment where, we consider that the uncertainty in the model is introduced only by the Gaussian fields.

4.3 Ensemble Kalman filter implementation for facies update

In this section, we present the customized implementation of the EnKF in related with our case study.

4.3.1 State vector

The state vector for the j^{th} ensemble member at the k^{th} assimilation step is:

$$\mathbf{x}_j^k = \left[y_1^T \; y_2^T \; m_1^T \; m_2^T \; \mathbf{d}_{sim}^T\right]_j^T, \tag{4.4}$$

where $y_1$and y_2 represents the two Gaussian random fields, m_1, m_2 are the truncation parameters of the random fields, and $\mathbf{d}_{sim}$ are the simulated observations. The latter ones are a combination of simulated production data (oil and water rates, bottom hole pressure) and simulated facies observations $f_{sim} = [\alpha_1^w \; \alpha_2^w]$. So, the state vector from equation 4.4 becomes:

$$\mathbf{x}_j^k = \left[y_1^T \; y_2^T \; m_1^T \; m_2^T \; \mathbf{BHP}^T \; \mathbf{q}_w^T \; \mathbf{q}_o^T \; \alpha_1^{w,T} \; \alpha_2^{w,T}\right]_j^T \tag{4.5}$$

In this study, the EnKF is performed without dynamical variables in state vector. Therefore, after each assimilation time, we have chosen to rerun the simulator from time zero, even though this technique will increase the computational cost. For the small models

of the examples presented, this procedure is not a time consuming, but, for large scale models, the computational cost should be considered.

4.3.2 Measurements

The measurements available for the HM process are the observed production data and facies data at the well locations. Contrary to the numerical character of the production data, the facies observations are of categorial type. Therefore, to be able to handle them in the numerical scheme of the history matching algorithm, we need to associate them with numerical values. When at a certain grid we know that a facies type occurs then the probability of occurrence of this facies type is one and probability of occurrence of the other two types of facies is zero. Therefore we might use the following assumptions: If at the well location we observe

- facies 1 then
$$\begin{cases} \alpha_1^{i,j} = 1 \\ \alpha_2^{i,j} = 0 \end{cases}$$
- facies 2 then
$$\begin{cases} \alpha_1^{i,j} = 0 \\ \alpha_2^{i,j} = 1 \end{cases}$$
- facies 3 then
$$\begin{cases} \alpha_1^{i,j} = 0 \\ \alpha_2^{i,j} = 0 \end{cases}$$

4.3.3 The EnKF implementation

1. **Initialization**

   The uncertainty in the initial ensemble is given by the choice of the two Gaussian random fields and by the choice of the truncation parameters (if we include the truncation parameters in the state vector). Initially, we generate an ensemble of pairs of Gaussian fields y_1 and y_2 with sequential Gaussian simulation method, specifying the geostatistical properties (isotropy or anisotropy, principal directions and the range correlation). Throughout this study the number of the ensemble members set to 120. The Gaussian fields are generated with constraints given by the types of facies found in the grids where the wells are situated. If, in a grid with a well located, we have observation about the existence of facies type 1, then the value in this grid for y_1 is 0 (y_1 is generated with value 0 in this grid such that the probability of occurrence is 1) and, of course, if we have observation

about the existence of facies type 2, then we generate y_2 with value 0 in this grid. The truncation parameters are generated with normal distribution and with a mean determined such that we keep accounted for prior knowledge about facies proportions. After the generation of the random Gaussian fields ensemble we perform an iterative process to constrain the GRF pairs to the other facies observations (y_1 to observations of the facies type 2 and type 3 and y_2 to observation of the facies type 1 and type 3). This iterative process is made with the EnKF method, defining the state vector

$$\mathbf{x} = \left[y_1^T \; y_2^T \; \alpha_1^T \; \alpha_1^T \right]^T \tag{4.6}$$

and the measurement are facies observations at grids where wells are situated. This iterative process stops when α_1 and α_2 in the grid with observation of facies type 3 have reached the values such that $\alpha_1 + \alpha_2 < 0.1$ (resulting that $\alpha_3 > 0.9$ in that grid), α_1 in grid with observation of facies 2 has property $\alpha_1 < 0.05$ ($\alpha_3 = 0$ there, by definition) and α_2 in the grid with observation of facies 1 has property $\alpha_2 < 0.05$ (as well $\alpha_3 = 0$ there, by definition). The stopping criterion is chosen in this way because, using two Gaussian fields to model three probabilities fields, we need to ensure similar initial conditions for all facies types. Although, the facies observations are perfect observations (0 variance) we use the value 0.01 for standard deviations of facies observations measurement error. We do that, because we need that the term $\left(H \, C_{\mathbf{x}} \, H^\top + C_{obs}\right)^{-1}$, from the expression of Kalman gain to exist, and, the value $C_{obs} = 0.0001 \cdot I$ to not have great importance in the inverse (Agbalaka and Oliver 2008). However, it is possible that, for some initial y_1 and y_2, the iterative EnKF procedure described above to not lead to an initial ensemble with the prescribed constraints and actually the iterative process get stuck (either the value of $\alpha_1 + \alpha_2$ does not descend below 0.1 or α_i for $i = 1, 2$ does not descent below 0.05) and is not able to stop. Then, we try the procedure with other two random fields, until we find the initial ensemble with correct (prescribed) constraints of the facies observations. As a measure of the spread of the Gaussian fields, we calculate, for both fields, the mean variance and we will watch their evolutions in the history matching process.

$$\sigma_y^2 = \frac{1}{n_g \left(n_e - 1\right)} \sum_{j=1}^{n_e} \sum_{i=1}^{n_g} \left(y_i^j - \overline{y_i} \right)^2 \tag{4.7}$$

where n_g is the number of the grids, n_e is the number of the ensemble members and $\overline{y_i}$ is the mean value of the random field y in the grid i

$$\overline{\mathbf{y}}_i = \frac{1}{n_e} \sum_{j=1}^{n_e} y_i^j \tag{4.8}$$

The mean variance is an important indicator because the quality of the forecast error covariance matrix is directly connected to spreading of the random fields (Agbalaka and Oliver 2008). Also, the mean variance of the Gaussian fields is a measure of the variability existing in the ensemble at a certain moment. If one of the mean variances is very close to zero then, we can state that the ensemble has collapsed. It is well known that the ensemble collapse is one of the main issue of the EnKF, and, in order to overcome that, we might use either a localization scheme, or an inflation of the forecast covariance matrix. This is not the case here, where, at the end of the assimilation period, the mean variance of each Gaussian field is greater than 0.08, value that from our experience ensures enough variability in the final updated ensemble.

2. **Forecast step/Model update**

The forecasted state vector at time $k+1$:

$$\mathbf{x}_j^{k+1,f} = \left[y_1^{k+1,T}\ y_2^{k+1,T}\ m_1^{k+1,T}\ m_2^{k+1,T}\ \mathbf{BHP}^{k+1,T} \mathbf{q}_w^{k+1,T}\ \mathbf{q}_o^{k+1,T}\ (\alpha_1^w)^{k+1,T}\ (\alpha_2^w)^{k+1,T} \right]_j^{T,f}, \tag{4.9}$$

where

$$\begin{bmatrix} y_1 \\ y_2 \\ m_1 \\ m_2 \end{bmatrix}_j^{k+1,f} = \begin{bmatrix} y_1 \\ y_2 \\ m_1 \\ m_2 \end{bmatrix}_j^{k,a} \tag{4.10}$$

$$\begin{bmatrix} \mathbf{p} \\ \mathbf{s} \end{bmatrix}_j^{k+1,f} = \mathcal{M}\left(\begin{bmatrix} \mathbf{p} \\ \mathbf{s} \end{bmatrix}_0^{k}, \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_j^{k,a} \right) \tag{4.11}$$

$$\begin{bmatrix} \mathbf{BHP} \\ \mathbf{q}_w \\ \mathbf{q}_o \end{bmatrix}_j^{k+1,f} = g_{pred}\left(\begin{bmatrix} \mathbf{p} \\ \mathbf{s} \end{bmatrix}_j^{k+1,f}, \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_j^{k+1,f} \right) \tag{4.12}$$

$$\begin{bmatrix} \alpha_1^w \\ \alpha_2^w \end{bmatrix}_j^{k+1,f} = \begin{bmatrix} \alpha_2^w \\ \alpha_2^w \end{bmatrix}_j^{k,a}, \tag{4.13}$$

where g_{pred} is the prediction operator for the production data. Hence, there are no changes from time k to $k+1$ for the values of the Gaussian random fields (eq. 4.10). But, their values have an impact in the changes of the pressures and saturations at time $k+1$ which generates changes in the forecast in production data (eq. 4.12). The Gaussian random fields are changed in the update step of the EnKF when the production data is assimilated.

3. **Update step/Measurement update**

During the assimilation of production data the updated state vector is calculated as follows:

$$\begin{aligned}\mathbf{x}_j^{k+1,a} &= \mathbf{x}_j^{k+1,f} + \mathbf{C}^f_{\mathbf{x}^{k+1}}\mathbf{H}^\top_{k+1}\Big(\mathbf{H}_{k+1}\mathbf{C}^f_{\mathbf{x}^{k+1}}\mathbf{H}^\top_{k+1} + \mathbf{C}^{prod}_{\text{obs},k+1}\Big)^{-1} \\ &\quad \Big[\mathbf{d}^{prod}_{\text{obs}_{j,k+1}} - \mathbf{H}_{k+1}\mathbf{x}_j^{k+1,f}\Big], \end{aligned} \tag{4.14}$$

where $\mathbf{H}$ represents observation operator for the production data, $\mathbf{C}^{prod}_{\text{obs}}$ represents the error covariance matrix for the production data, $\mathbf{C}_{\mathbf{x}^f}$ represents the forecasted error covariance of the ensemble, and $\mathbf{d}^{prod}_{\text{obs}_j}$ are the observed production data. In order assimilate the facies data we separate part of the state vector as in:

$$\tilde{\mathbf{x}}_j^{k+1} = \Big[y_1^T\ y_2^T\ m_1^T\ m_2^T\ \alpha_1^{w,T}\ \alpha_2^{w,T}\Big]_j^{T,k+1}, \tag{4.15}$$

If, after the assimilation of the facies data, some of the ensemble members violate the position of the facies type 3 in the grids where we have observations about their existence, we perform an iterative enforcement on facies observations (because the random fields y_1 and y_2 are generated with value 0 in grids where facies type 1 and type 2 respectively occur, during the HM process the position of the facies type 1 and type 2 are not violated). The iterative process will stop when the value of α_3 in the grids where we have observation about existence of the facies type 3 is greater than α_1 and α_2 in those grids. In the iterative process explained above the uncertainty for the facies observations is represented by the error covariance matrix of $0.0001\mathbf{I}$.

4. **The estimated field**

One of the targets of this study is to find a field of facies distribution which is an estimator for the "truth" field (or reference field). In the previous studies using truncated plurigaussian method for facies simulations, an estimator of the truth field could not be directly presented, because the posterior mean of the random fields could not generate a plausible geological model. In this study, for each facies type, we calculate the mean of probabilities fields (named estimated probabilities field), that will be an estimator of the associated binary field (generated by the facies type). Using these three estimated probabilities fields we define the estimated field with the maximization criterion, meaning that the assignment of a facies type in a grid is made for the facies type with the highest value of probabil-

ity in that grid. For $k = 1$ or $k = 2$ we have first two estimated probabilities fields as

$$\overline{\alpha}_k^{i,j} = \frac{1}{n_e} \sum_{l=1}^{n_e} \varphi_{m_{k,l}} \left(y_{k,l}^{i,j} \right) \tag{4.16}$$

For the mean of the probability of occurrence of the facies type 3 we use the relation

$$\overline{\alpha}_3^{i,j} = \begin{cases} 1 - \left(\overline{\alpha}_1^{i,j} + \overline{\alpha}_2^{i,j} \right) & \text{if } 0 \leq \overline{\alpha}_1^{i,j} + \overline{\alpha}_2^{i,j} \leq 1 \\ 0 & \text{if } \overline{\alpha}_1^{i,j} + \overline{\alpha}_2^{i,j} > 1 \end{cases} \tag{4.17}$$

Consequently, in the estimated field we assign the facies type k in the grid (i, i) if

$$\overline{\alpha}_k^{i,j} = max\{\overline{\alpha}_l^{i,j},\ l = \overline{1,3}\}, \tag{4.18}$$

Also, we will use the same rule as in the main procedure, regarding the cases when the probabilities are equals in the grids.

4.3.4 Indicators for the quality of the estimations

The first comparison between updated fields and reference field or between different estimated fields of the same reference field is made by visualizing the facies maps. But this visualization is not sufficient, therefore, we define some quantitative indicators which can show the quality of the estimation.
(1) Root Mean Square Error (RMS) calculated for permeability:

$$\mathbf{RMS(k)} = \sqrt{\frac{1}{n_g} \sum_{i=1}^{n_g} (k_{ref_i} - k_{est_i})^2} \tag{4.19}$$

where n_g is the number of grid blocks $kref_i$ and $kest_i$ represent the values of permeability for reference field and estimated field for grid block i.
(2) Percentage of grid fit between the estimated field and the reference field:

$$\mathbf{Percentage} = \frac{1}{n_g} \sum_{i=1}^{n_g} m_i \times 100 \tag{4.20}$$

where $m_i = 1$ if in the grid i the estimated field and the reference field has the same facies type and $m_i = 0$ if we have a misfit between the facies type existing in estimated field and facies type existing in reference field.
(3) Root Mean Square Error (RMS) of the Water Rates (WR) prediction for the next 100 days (10 time steps), calculated for the estimated field and for the mean of the updated ensemble members and, respectively for Total Rates Production (TR) of the estimated

field and mean of the updated ensemble in the assimilation period . We denote with X the reference variable (respectively WR and TR).

$$\mathbf{RMS}(\mathbf{X_{meanupdate}}) = \sqrt{\frac{1}{4n_g}\sum_{j=t_0}^{t_1}\left(X_{ref,j} - \overline{X}_j\right)^2} \tag{4.21}$$

$$\mathbf{RMS}(\mathbf{X_{estimate}}) = \sqrt{\frac{1}{4n_g}\sum_{j=t_0}^{t_1}\left(X_{ref,j} - X_{est,j}\right)^2} \tag{4.22}$$

where $X_{ref,j}$ is the value obtained by the "truth" field at the time j, $X_{est,j}$ is the value obtained by the estimated field at the time j, $\overline{X}_j = \frac{1}{n_e}\sum_{i=1}^{n_e} X_i$ is the mean of the reference variable obtained by the updated ensemble at time j and n_e is the number of ensemble members. These quantities are calculated for every producer, and for this reason, we have a factor 4 at denominator.

4.4 Reservoir description

The simulation model is a 5-spot water flooding 2D-reservoir, black oil model with $50 \times 50 \times 1$ active grid blocks. The dimension of each grid block was set at $30 \times 30 \times 20$ ft and there is one injector situated at the center of the reservoir domain and there are 4 producers situated at the corners. The coordinates of the well positions are presented in Table 4.1. For all the examples presented we use the facies observations presented in Table 4.1. The values of the permeability and porosity corresponding to each facies type are presented in Table 4.2.

Water injection starts from the first day and continue thereafter a period of 201 days of

Table 4.1: The position of the wells in the reservoir domain and the facies observations

	Injector	Producer 1	Producer 2	Producer 3	Producer 4
x coordinate	25	5	5	45	45
y coordinate	25	5	45	5	45
Facies observation	Type 2	Type 1	Type 3	Type 1	Type 2

production. We assimilate data starting from first day until day 201 from 10 to 10 days which means that we have 21 assimilation steps. All the producers work under constant BHP value of 1500 pound per square inch (psi), the injector operates at 4500 stock tank barrels per day (STB/D) and constrained by a maximum 5000 psi for BHP. The measurement error for the production data is considered Gaussian with 0 mean and standard deviations respectively $\sqrt{5}$ STB/D for WR and OR at the producers and 3 psi for BHP at the injector. These values will be used for generation of the noise, which is added to the

Table 4.2: The petrophysical properties of each facies type

Facies type	Permeability	Porosity
Type 1	174 md	0.18
Type 2	732 md	0.25
Type 3	80 md	0.14

observations of the production data in the analyzed step of the EnKF process. The observations of the production data are obtained through forward integration of one model defined as reference (or "truth"), considered known in every example presented.

4.5 Cases presented

We test the model presented, using Gaussian random fields generated with different geostatistical properties (isotropic or anisotropic fields) and different levels for the uncertainty that we add in construction of the truncation parameter ensemble. In the first case we perform two experiments, one with uncertain truncation parameters and the other without uncertain truncation parameters. When we don't add uncertainty for the truncation parameters, actually we use for the truncation of the Gaussian fields the same truncation map for all the ensemble members. This case is made in order to reveal the influence of the truncation parameters in the estimates.
The second case contains one experiment, where the fields are generated with anisotropic characteristics and in the third case we introduce uncertainty in the range correlation with which the Gaussian fields are generated. In all the experiments we assume that, based on the prior information, we have an initial knowledge about the geostatistical properties of the Gaussian fields and the mean of the truncation parameters.
Case1. In this case the reference field was build using two Gaussian fields with isotropic characteristics having the range correlation 17 grid blocks. We project these two random fields using a truncation map constructed with values $\sqrt{2}$ for both thresholds generating the "truth" distribution of the facies type in the reservoir domain. The initial ensemble of Gaussian fields y_1 and y_2 was generated with the same isotropic characteristics and range correlation. This case contains two experiments. For the first experiment we constrain the ensemble of Gaussian fields to facies observations using value 1.5 for all truncation parameters, in the second experiment we generate an ensemble of thresholds with Gaussian distribution having mean 1.5 and standard deviation 0.2.
Case2. In the second case presented, the reference field was build using two Gaussian fields with anisotropic characteristics with long range correlation 16 grid blocks, short

range correlation 8 grid blocks and principal direction having angle 30° with X-axis and for the truncation we use value $\sqrt{2}$ for both thresholds. The initial ensemble of Gaussian fields y_1 and y_2 was generated with the same anisotropic characteristics, range correlations and directions. In this example we constrain the ensemble of Gaussian fields to facies measurements using an ensemble of truncation parameters having normal distribution with $\sqrt{2}$ mean and 0.2 standard deviation.

Case3. In the third case we introduce uncertainty in the range correlation with which we generate the Gaussian fields. As reference field we use the field presented in the first case. The initial ensemble of Gaussian fields is composed from three batches, each one having 40 members and generated isotropic with range correlation 16, 18 and respectively 20 grid blocks. This ensemble was constrained to facies observation with an ensemble of truncation parameters generated with mean 1.5 and standard deviation 0.1.

4.5.1 Case 1: Simulation model generated using isotropic random fields

The reference field used for this case is presented in Figure 4.4 where blue color represent facies type 1, green color represent facies type 2 and red color represent facies type 3. The black dots show the positions of the wells. The field has at the well location the facies observations presented in Table 4.1. Each facies type generates a binary field defined on the reservoir domain of which representation is shown in the first line of Figure 4.5.

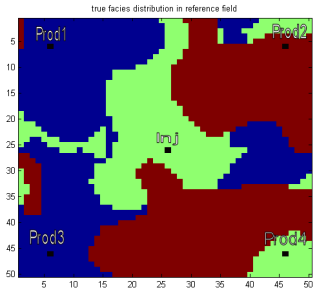


Figure 4.4: Reference field for example 1

This case contains two experiments. In the beginning, we generate two families of Gaussian fields which will be constrained to facies observations using different ensembles of truncation parameters. For the first experiment we use value 1.5 for all truncation parameters and for the second experiment we use an ensemble of truncation parameters generated with normal distribution, having mean 1.5 and standard deviation 0.2. This

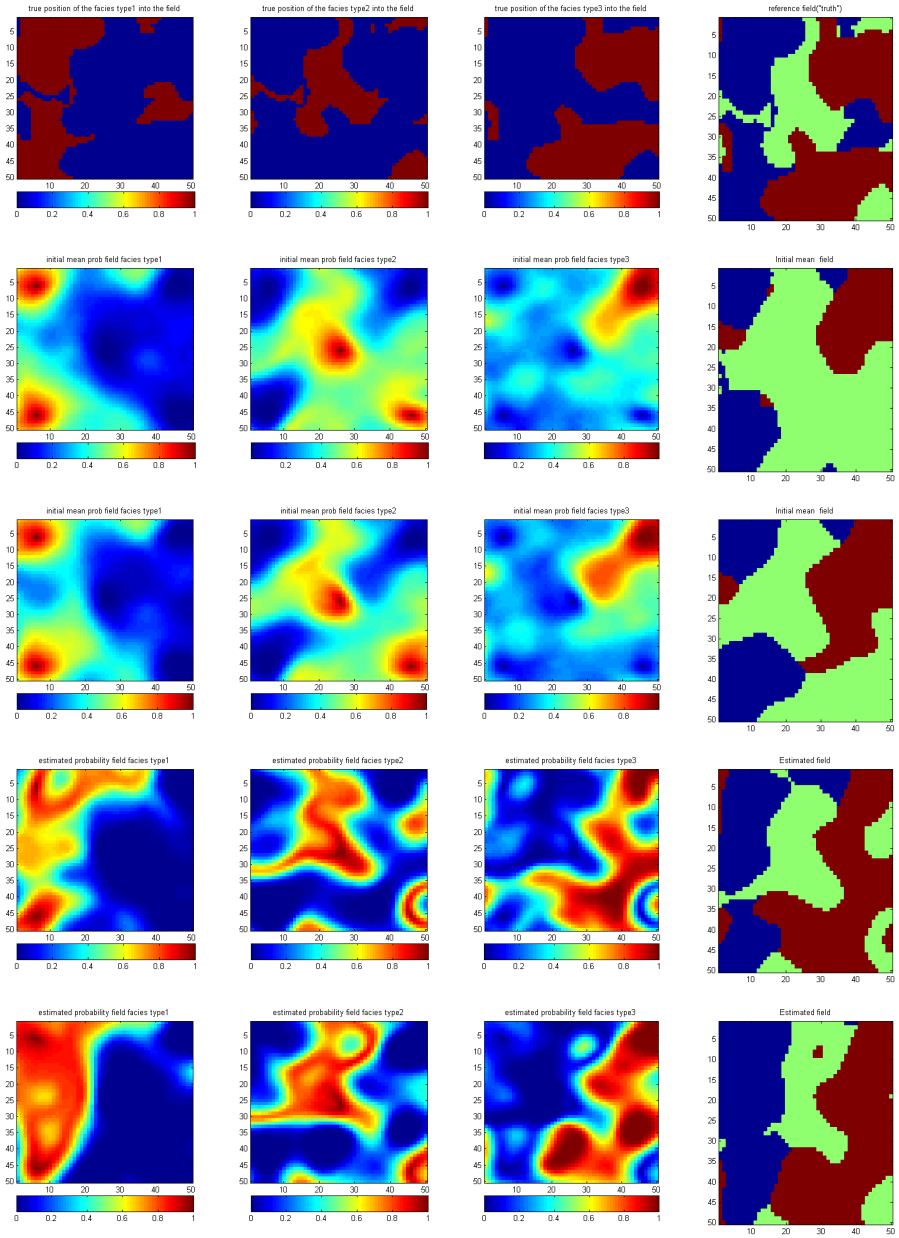


Figure 4.5: The true positions of the facies types in the reservoir domain (line 1), The initial fields with standard deviation for truncation parameters 0 (line 2), The initial fields with standard deviation for truncation parameters 0.2 (line 3), The estimated fields with standard deviation for truncation parameters 0 (line 4), The estimated fields with standard deviation for truncation parameters 0 (line 5)

means that in the first experiment we do not introduce uncertainty in the truncation parameters. The choice of 1.5 as the mean value for the truncation parameters is made taking into account the prior knowledge of the facies proportions of the reference field. We obtain two initial ensembles of Gaussian fields that are derived from the same ensemble of GRF which means that these two ensembles contain similar information (as we see from the initial mean probabilities fields presented in Figures 4.5 (lines 2 and 3)). In Figure 4.5, lines 2 and 3, we presents, for the experiments, the initial fields (which contains initial mean fields for the probabilities of occurrence of the facies type and the initial "mean" field) and in lines 4 and 5 the results obtained after the assisted HM performing. The results presented are the estimated probabilities fields for all the facies types and the estimated field. From these pictures we can observe that, for the initial probabilities fields of facies type 1 and type 2 the region with probabilities greater than 0 are diffusive distributed in the reservoir domain, whereas in the estimated probabilities fields are more clearly shaped. If in the initial fields the shapes are contoured with the values near to 0.4 in the estimated fields the shapes are contoured by the values near to 0.7. The variability in the initial ensemble of random fields can be observed in the diffusive distribution of the shapes of the facies type 1 and type 2 in the initial maps. Also, in the initial mean field (and also in the initial ensemble of facies maps) the facies type 3 have the smallest proportion because of two causes; first, only in the well 2 we have observation about occurrence of that facies type and second, because of the construction of the simulated model, where probability of occurrence of the facies type 3 is appointed based on fundamental rule of probabilities. After the assimilation period, the facies proportion in the estimated field is close to the true facies proportions ("truth" field has the facies proportions 0.36:0.24:0.40 and in the estimated field has the facies proportions 0.34:0.24:0.41). With a visual inspection of the estimated fields (Figure 4.5 lines 4 and 5 last pictures) we can conclude that these fields are good estimators for the reference field but for a better quantification of the estimation we will compare the indicators presented previously. In Table 4.3 we present the values for the indicators calculated for the estimated fields. We can easily observe that indicators of the estimated field resulting from HM process with uncertainty in the truncation parameters have better values than the indicators of the estimated field obtained without uncertainty in the truncation parameters. Knowing that the true facies proportions is 0.36:0.24:0.40, from the Table 4.3 we can observe that in the both cases the variability of the initial ensemble incorporates this true value. However, only the updated ensemble for the case with variability in the truncation parameters preserves this property. We have repeated the experiments for several initial ensembles and the conclusion is that, the updated fields of the experiment where we take into account the uncertainty in the truncation parameters, has a higher percentage of grid fitting and a better quantification of the facies proportions then the other updated fields. Concerning the RMS calculated for the rates, either for the ensembles, or for the

estimated fields, these values are very close to each other. This happens because, given two initial ensembles derived from the same ensemble of Gaussian fields, one generated using an ensemble of truncated parameters and one generated with constant value for truncation parameters, the reduction in variability, after the assimilation of production data and facies (hard) data, is similar, toward to the values of the truth. We can conclude that, for a better description of the facies types distribution in the field we should introduce uncertainty in the truncation parameters.

Table 4.3: Qualitative indicators calculated for the experiments

	$Sd_m = 0$	$Sd_m = 0.2$
Percentage of the grid fit	64.20%	72.16%
RMS(permeability)(md)	135.48	115.82
RMS(WR mean update)(STB/D)	0.52	1.19
RMS(WR estimate)(STB/D)	1.47	1.14
RMS(TR mean update)(STB/D)	4.36	3.46
RMS(TR estimate)(STB/D)	5.92	3.96
Facies proportions initial state	0.25:0.42:0.34	0.27:0.39:0.34
Standard deviation initial state (facies prop)	0.08:0.11:0.11	0.08:0.11:0.11
Max proportions initial state (facies prop)	0.47:0.68:0.70	0.45:0.64:0.64
Min proportions initial state (facies prop)	0.07:0.06:0.10	0.09:0.07:0.08
Facies proportions updated state	0.28:0.29:0.43	0.32:0.26:0.42
Standard deviation updated state (facies prop)	0.03:0.02:0.02	0.03:0.02:0.02
Max proportions updated state (facies prop)	0.34:0.32:0.48	0.40:0.32:0.49
Min proportions updated state (facies prop)	0.22:0.26:0.38	0.21:0.22:0.37

For the model with truncation parameters generated with standard deviation 0.2 we present the WR for all producers, for updated ensemble (Figure 4.6 line 2) and for initial ensemble (Figure 4.6 line 1) in the assimilation period. From these two pictures we can observe the initial variability (Figure 4.6 line 1) and the reduction in variability due the data assimilation procedure (Figure 4.6 line 2). In the line 3 of Figure 4.6 we present the evolution of WR for a period of 301 days, where the vertical blue line delimits the assimilation period (201 days) to the prediction period (100 days) for WR of the estimate field, reference field and mean of the ensemble. For the producers 1, 3 and 4 the prediction is very good, only for the producer 2 the water rate prediction of the estimated field and the mean of the ensemble exhibits some distance to the water rate measured at that producer in the reference field, in the interval 201-301 days. The reason for that the WR prediction at the producer 2 is not as good as in the rest producers is because the water cut at this well appears later (at assimilation time 14) whereas at the rest wells the water occurs earlier (around assimilation time 10).

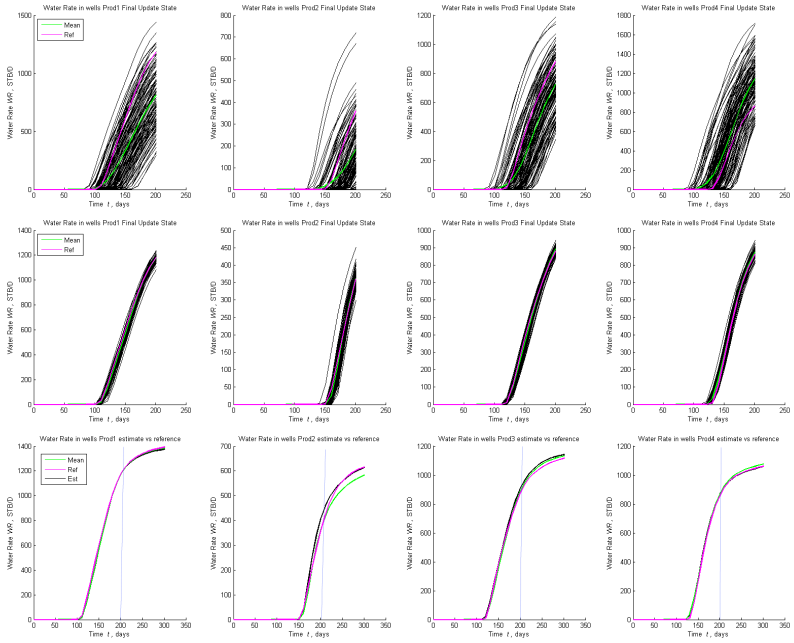


Figure 4.6: WR in initial state (line 1), WR in update state (line 2) and WR prediction for the next 100 days for estimated field and mean of the ensemble (line 3)

The evolution of the variability of random fields in the assimilation period is represented

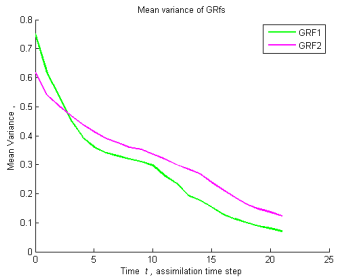


Figure 4.7: Mean variance of the random fields

by the mean variance calculated at every time step and, its distribution over the assimilation period (Figure 4.7). The reduction in the variability of both random fields is sharp at the initial assimilation times (when important data are assimilated) and when the water occurs at the producers (time step 10, when new relevant information is introduced in the process) After that, the decrease in variability is made with an approximately constant slope. From the updated ensemble of facies maps we can calculate based on a simple

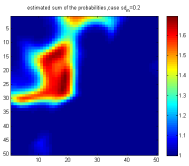


Figure 4.8: Sum of the estimated probabilities

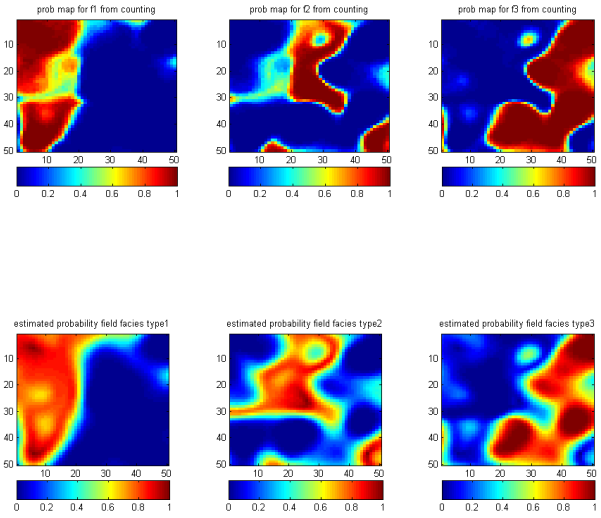


Figure 4.9: probabilities maps from estimation and from the facies maps ensemble

counting three probability maps associated with each facies type. The probabilities fields resulting from the estimated process are consistent with the probability maps calculated from the updated ensemble as can be seen from Figure 4.9. The differences between these maps are caused in principal because the probabilities fields derived from estimation (bottom of Figure 4.9) arise from a continuous medium, whereas the probability fields defined by the ensemble of facies maps (top of Figure 4.9) come from a discrete medium (is the result of a counting). The inverse modeling of the facies field with EnKF allows us the estimation of the state parameters, but the estimators are not perfect which means that, if between the true values of parameters exist some connectivity relations, these relations are not necessarily fulfilled by the estimators. In our case in every grid

cell the sum of the probabilities is 1 (actually two probabilities are 0 and one probability is 1) but it is possible that for the estimators of the probabilities, the sum it may be greater than 1 (Figure 4.8).

4.5.2 Case 2: Simulation model generated using anisotropic random fields

The directions of the facies type propagation in the reservoir domain, are related with the isotropy or anisotropy of the Gaussian fields. When the prior information suggests the propagation of the facies in certain direction, we model the facies maps (in the initial ensemble) with the aid of Gaussian fields generated anisotropically, with principal direction given by the propagation direction of the facies. The experiment performed in this case is addressed to the facies fields that exhibit a given orientation in the domain. We have used the same setup for the reservoir description as in the first case. The reference field is presented in the last picture of the first line of Figure 4.10, where facies type 1 has blue color, facies type 2 has green color and facies type 3 has red color. The petrophysical properties of every facies type are the same as in the previous experiments. The results for the estimated field and estimated probabilities fields after a period of 21 assimilation time steps are presented in the Figure 4.10 (line 3), whereas the true position of the facies types in the reference field are presented in Figure 4.10 (line 1). From these pictures, we can observe that the positions of the facies types over the reservoir domain are pretty good estimated, although the estimated probabilities fields for facies type 1 and 2 exhibits some features that are not present in reference field. The facies proportion in the initial mean field was 0.50:0.44:0.06 (this smaller proportion for facies type 3 can be seen from Figure 4.10 (line 2)) totally inconsistent with true facies proportions 0.305:0.271:0.417. However, the facies proportion calculated from the initial ensemble is 0.38:0.38:0.24 with a spread quantify in standard deviation of 0.1:0.1:0.12 which contains the true facies proportions. After the assimilation period the calculated facies proportions for the estimated field are 0.359:0.278:0.363 close to the true facies proportions and close to the facies proportions calculated from the updated ensemble, 0.32:0.29:0.4 (with a standard deviation of 0.045:0.025:0.035). The reduction in variability follows a similar scenario as in the experiments carried out with isotropic fields (as we can see with a visual inspection of the Figures 4.6 and 4.11). The variability existing in the initial ensemble can be inferred either from the initial maps presented in Figure 4.10 (line2), or from the water rates spread presented in Figure 4.11 (line 1). The facies fields that exhibit this type of propagation in the domain are very hard to be estimates because by modeling the Gaussian fields with anisotropic characteristics, having a very small short correlation range reported to the domain length, we introduce in the model many degrees of freedom.

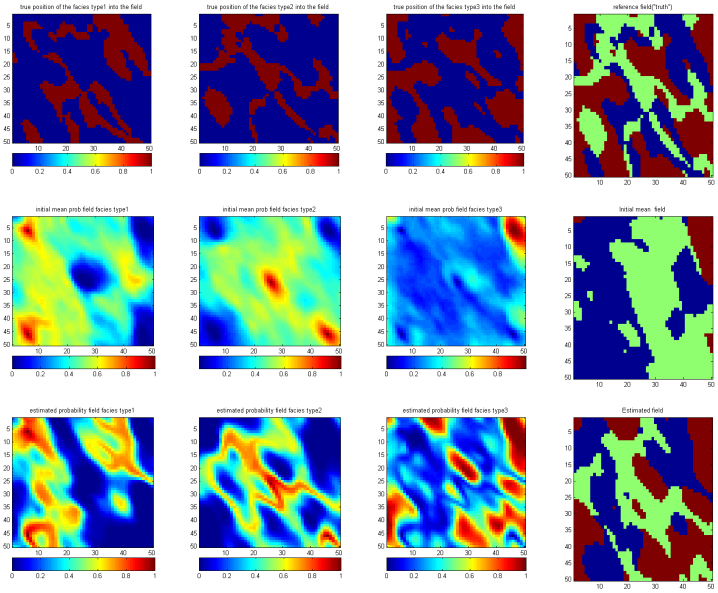


Figure 4.10: The binary fields defined by each facies type and the reference field (line1), The initial mean probabilities fields and the initial mean field (line 2), The estimated probabilities fields and the estimated field (line 3)

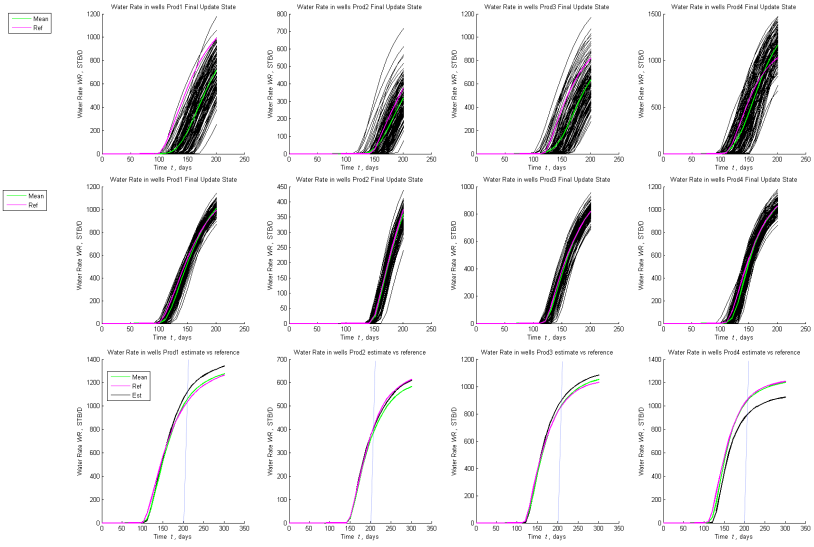


Figure 4.11: Water Rates in initial state (line 1), Water rates in updated state (line 2) add Water rates prediction in updated state for the reference field, mean of the ensemble and the estimated field

4.5.3 Case 3: Uncertainty in the geostatistical properties of the Gaussian fields

An important issue in the plurigaussian simulation is the choice of the geostatistical properties (variogram parameters) used for generating of the Gaussian fields (Jafarpour et al. (2011)). The shape of the facies continuity and the numbers of the bodies depend on the variogram parameters of the Gaussian fields but, it is not straightforward to establish a connection formula. The choice of the correlation range (or any other geostatistical property) shall be made taking into account the prior knowledge about geology description (facies proportions, the propagation of the facies areas in the reservoir domain, the numbers and the volume of bodies etc). This prior information about geology is often uncertain. Different geostastical properties of the Gaussian fields provide, after truncation, facies fields with similar characteristics. Therefore, in this experiment we want to explore the case in which we do not use a single value for the geostatistical properties of prior fields generation. For that, we generate three batches of isotropic Gaussian fields, each one having 40 members constructed with different levels for correlation range of 16, 18 and respectively 20 grid blocks. We have used these values because the maps resulting from truncation present similar topological characteristics.

Table 4.4: Qualitative indicators calculated for the experiments

	Case 1 ($Sd_m = 0.2$)	Case 2 (21 ts)	Case 3 (22 ts)
Percentage of the grid fit	72.16%	70.76%	71.12%
RMS(permeability)(md)	115.82	120.2	117.25
RMS(WR mean update)(STB/D)	1.19	1.87	1.36
RMS(WR estimate)(STB/D)	1.14	1.89	1.54
RMS(TR mean update)(STB/D)	3.46	2.2	2.20
RMS(TR estimate)(STB/D)	3.96	3.22	3.49

This ensemble is constrained to facies observations using an ensemble of truncation parameters generated with a normal distribution having the mean 1.5 and the standard deviation 0.1. The reference field used in this example is the field presented in first case study (Figure 4.12, line 1). The results are shown in Figure 4.12 line 3, where the estimated probabilities fields and the estimated field are presented. The water rates evolution for the initial ensemble and the updated ensemble, in the assimilation period has the same characteristics as the ones from the first case presented (Figure 4.13). Because the water occurs later at the producer 2 (Figure 4.13, time step 16 and more consistent 17) the prediction of the water rate in this producer is not the best. Therefore we introduce one more assimilation time step in order to reduce the WR variance in this producer.

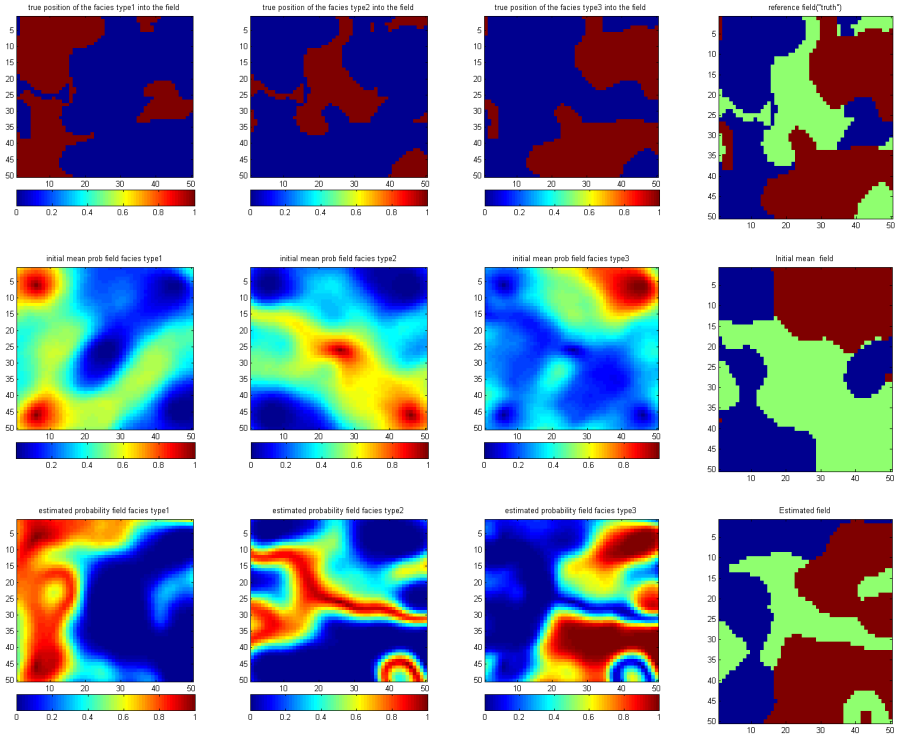


Figure 4.12: The binary fields defined by each facies type and the reference field (line 1), The initial mean probabilities fields and the initial mean field (line 2), The estimated probabilities fields and the estimated field (line 3)

As we can see from the Table 4.4 there is an improvement related with the percentage of grid fit (from 70.76 to 71.12) and for RMS calculated for permeability (from 120.2 md to 117.25 md). After several experiments with the same setup we have found that we could not see an improvement in the estimation, neither for the grid fitting, nor for the RMS. The percentage of grid fitting for the estimated field is situated around 68% for the both cases (the ensembles where all the Gaussian fields are generated with the same geostatistical characteristics and the ensembles generates as in the case 3).

4.6 Appendix: The facies proportion impact

The truncation parameters are variables that define the bounding lines in the truncation map and the volumes of the map zones. As a result, they have the ability to control the facies proportions in the reservoir domain.

In Figure 4.14 are five different domains generated with the same random fields but with different values for the truncation parameters. From this figure we can observe

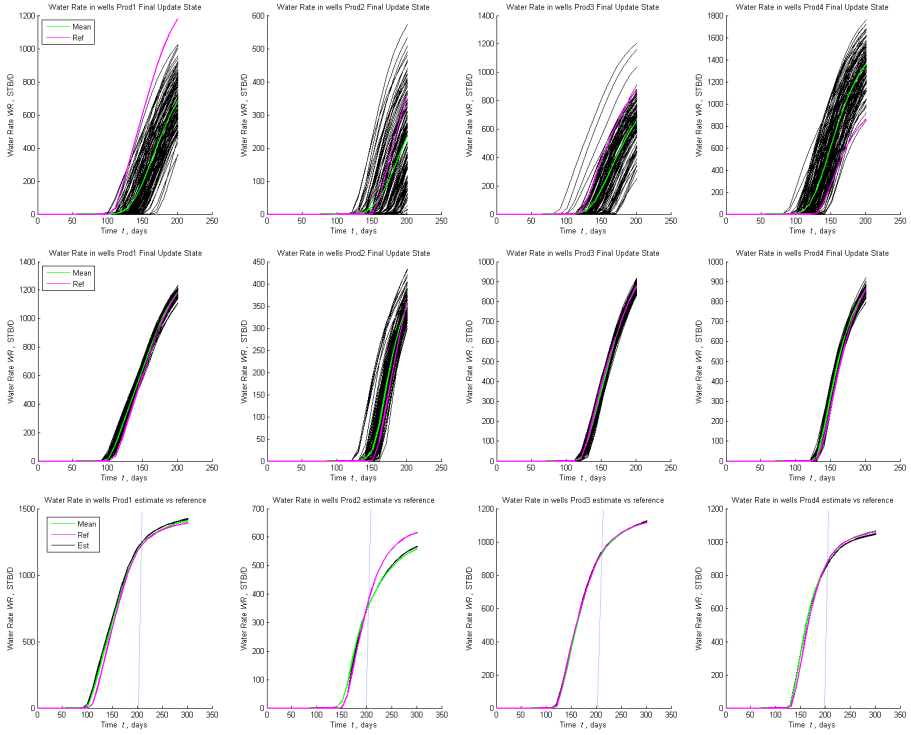


Figure 4.13: WR in initial state (line 1), WR in updated state (line 2) and WR prediction for the next 100 days (line 3)

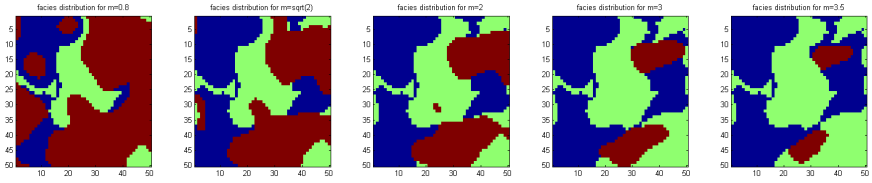


Figure 4.14: Facies maps generated with different truncation parameters

that, for both values of truncation parameters of 3.5 the proportion of the facies type 3 (red shade) is very small in comparison with the proportions of the facies type 1 (green) and type 2 (blue). Decreasing values of m_1 and m_2, facies type 1 and facies type 2 are eroded by the facies type 3. For example, if we assign for both parameters the value 0.8, the regions represented by the facies type 3 have the higher proportion. From the construction of the map, an increase for m_1 generates an increase of proportion of the facies type 1 together with a decrease in the proportion of the facies type 3 and type 2. The influence of the parameter m_1 in the proportion of the facies type 2 is smaller than the influence in the proportion of facies type 3. A similar approach is valid for

parameter m_2. For a given pair of Gaussian fields, we can define a pair of truncation parameters such that the associated facies map has established facies proportion. This pair of truncation parameters is defined based on an estimation procedure, using, as well, the ensemble Kalman filter method. For that, we define the state vector as:

$$\mathbf{x} = \left[m_1^T \; m_2^T \; prop_1^T \; prop_2^T \; prop_3^T \; \right]^T, \tag{4.23}$$

where, m_1 and m_2 signify the parameters and $prop_1$, $prop_2$, $prop_3$ are the simulated measurements which, in this case, represent the simulated facies proportions for each facies type. We start the procedure with an initial ensemble of truncation parameters, generated with normal distribution, having a prior mean that can be chosen such that the associated facies proportions are close to the established facies proportions (with a trial error procedure), and a standard deviation of 0.2. Also, we can define a general prior mean of 1.5 for both truncation parameters and a standard deviation of 0.4. These values ensure us enough variability in the initial ensemble, such that, the EnKF procedure to converge. In this case the observations are perfect (being represented by the given facies proportions). However, for the Kalman gain to exist, we set for the measurement errors a small value, of 0.01 for standard deviation. We apply the EnKF procedure iterative, with the stopping criterion set up such that the difference between the observations and measurements are in absolute value less than 0.01 for all the ensemble members. The ensemble mean for the both parameters are the required estimators. However, various pairs of Gaussian fields, having similar geostatistical properties, could generate, after truncation with the same scheme (a truncation map built with the same values for the parameters), facies maps with different facies proportions. Considering an ensemble of Gaussian fields pairs (in our case 120 realizations), we can determine (from the ensemble), the proportion mean for each facies type. These values constitute estimators for the expected facies proportions of each facies type for a given truncation map (in our case, for a given combination of the parameters m_1 and m_2) and for given geostatistical properties of the Gaussian fields. Depending on the generation of Gaussian fields ensemble, unconditioned or conditioned to facies observations, the associated expected facies proportions have different values. For example, considering a map defined by parameters, both having the value 1.5, and Gaussian fields generated isotropic with range correlation of 17 gb, we obtain the expected facies proportion 0.35:0.35:0.30 for unconditional generated GRF, 0.37:0.38:0.24 for conditional generated GRF (to observation of facies type 1 and 2) and 0.25:0.42:0.33 for conditional generated GRF (to all facies observations). Together with these mean values, the ensemble of facies maps, generates a spread of the facies proportions values quantified in standard deviation as, 0.11:0.11:0.11 for the unconditional generated GRF's and conditioned to observations of the first two facies types and respectively 0.08:0.11:0.11 for the GRF's constrained to all observations. Also, the extreme values for each facies proportions are 0.45:0.64:0.64 for the maximum values

and respectively 0.09:0.06:0.08 for the minimum values. In this study we do not use the facies proportions as measurements, even though we could have some prior knowledge about these (together with an uncertainty bandwidth). The choice for the prior values of truncation parameters is made, such that the spread of the facies proportions (calculated from the ensemble) to contains the expected facies proportions given by the geologist. This means that the choice of the prior values for the truncation parameters (and from here for the truncation map) could be made in various ways, as long we have enough variability in the initial ensemble which incorporates the prior knowledge.

Chapter 5

An adaptive plurigaussian simulation (APS) model for geological uncertainty quantification using the EnKF

5.1 Introduction

For the implementation of any assisted history matching (AHM) model in the area of reservoir engineering, one of the most important topics to address is the reservoir geology. When referring to geological description we have in mind facies distribution, faults position or top/base surfaces of the geological layers. Usually the initial knowledge about the geology is very uncertain, which is very attractive for stochastic AHM models. In this study we are interested in the facies distribution and try to reduce its initial uncertainty using additional information given by the production data when available.
In order to quantify the geological uncertainty, firstly, we have to define the geological simulation model. Using this model, we generate possible facies realizations (distributions) on the reservoir domain. These facies instances must be consistent with the prior knowledge about the reservoir geology (the number of the facies types, the possible transition between facies types, the direction of the facies, the geometry of the facies, etc.). Secondly, for the estimation purposes and for uncertainty propagation and hence, quantification, we use as history matching (HM) method the ensemble Kalman filter (EnKF).

This chapter is full reference of an article submitted to Computational Geosciences

The EnKF is part of the large family of Monte Carlo techniques. The representation of the uncertainty in the EnKF framework is carried out through an ensemble of possible realizations. All the statistical measures (mean, covariances) that occur in the Kalman filter are estimated from the ensemble. This ensemble of realizations is assumed to be samples from the distribution of the true state. Thereby, the updated ensemble quantifies the uncertainty of the reference parameters.

The geological simulation model used in this article is based on a plurigaussian simulation technique. The plurigaussian truncation scheme consists of a generation of the facies distribution using a projection from the multi-dimensional (usually bi-dimensional) space of Gaussian random fields (defined on the reservoir domain) to the discrete space (the facies fields space) through a truncation map. The truncation map is defined in the multi-dimensional space of the Gaussian random fields (GRF) and, is either user defined based on prior knowledge (Liu and Oliver 2005) or is defined through an internal construction of the model (Sebacher et al. 2013). The truncation plurigaussian simulation (TPS) method is a natural generalization of the Gaussian truncation method where a single Gaussian field was truncated using some thresholds defined on the real axis. The TPS method was first introduced by Galli et al. (1994) and further developed by Le Loc'h et al. (1994) and Le Loc'h and Galli (1997). In the paper of Le Loc'h et al. (1994), two Gaussian random fields were truncated using a truncation map, defined on the bi-dimensional real space, designed by intersection of some vertical and horizontal lines which give a rectangular geometry for the map. Each zone delimited by the truncation lines has assigned a facies type. Using different geostatistical properties of the Gaussian fields, different instances of facies distributions were generated, corresponding to different characteristics of the facies types. In the article of Le Loc'h and Galli (1997), the authors investigate the relation between the variogram of the Gaussian fields and the indicator variogram of the facies for the stationarity case, together with presenting the conditional simulation technique for the non-stationary case. Important prior information is related to the possible contacts among different facies types an issue characterized by Xu et al. through a binary dynamic contact relation matrix in the context of plurigaussian simulation model.

The coupling of the truncated plurigaussian simulation method with a HM method was made first by Liu and Oliver (2005) when the facies boundaries were automatically adjusted when new measurements were available. The authors have used as HM algorithms the EnKF (as non-gradient based method) and the randomized maximum likelihood (as gradient based method) and they did a comparison between them; the conclusion was that the EnKF gave better results. For the facies simulation model, the truncation map used was built with three lines, whose intersections in the bi-dimensional real space generate seven regions, each having a facies type assigned. The measurements used in the HM were the production data and the facies observations at the well locations. The same

truncation map with EnKF as HM method was used by Agbalaka and Oliver(2008) in a 3D synthetic model, with 3 facies types (each two could be in contact), and using in addition, a distance-based localization scheme for the forecast covariance matrix. Because, after each production data assimilation, some of the facies fields generated with TPS, are not fulfilling the correct facies observation, an extra iterative procedure is performed to correct this misfit. For the facies observations operator a proxy function was used, where the image is defined by two values (0 if the facies type is correct observed and 1 if the facies type is not). Zhao et al. (2008) integrate the truncated plurigaussian simulation model, introduced by Liu and Oliver, into an iterative EnKF process, assimilating production data in addition to seismic data, for a 3D synthetic reservoir model with three vertically uncorrelated layers. A bi-dimensional truncation scheme defined by two ellipses was used for channels modeling in a plurigaussian context. Agbalaka and Oliver (2009) applied EnKF to facies fields that exhibit non-stationarity in facies proportions whilst simultaneously updating the petrophysical properties of the facies. They truncated a single Gaussian field using two thresholds calculated with respect to three predefined facies proportion maps, a model that generates facies fields where the transition between the facies types is unidirectional (two facies types are not in contact). A probabilistic operator for the facies observation has been used by Sebacher et al (2013), where the geological simulation model was introduced in the TPS context through the internal construction of the model. The authors defined a new object, named "the probabilities fields", with which the binary fields of each facies type were estimated. The probabilities fields were defined as projection of some Gaussian fields in [0,1] interval with a projection function, the function which plays the role of the observation operator of the facies occurrence. The truncation map that arises from the model is defined by a maximization criterion applied to the probabilities fields. The parameters that define the truncation map were introduced in the state vector and estimated. In addition, the probability maps (for each facies type) calculated from the updated ensemble are consistent with the mean of the probabilities fields.

This article starts from the assumption that the initial probability fields of each facies type can be a priori designed by a group of experts in the early phase of the reservoir exploration, using various data, such as core information, seismic data, well logs data, outcrops information, etc. Supplementary with the probability fields, we require extra information, e.g. the possible contact among the facies type, the expected facies proportions (global), the geometry of the facies, the relative position among the facies type. All this information is part of the expert knowledge of the reservoir and is used to construct a geological simulation model which is coupled with the EnKF, obtaining sequentially, updates for the probability fields. The AHM model offers an update of the probability maps (provided by an initial complex study of the reservoir) for the facies types every time when new information becomes available.

The main idea of our approach is to generate an ensemble of facies fields, geologically plausible, where the probability maps of the facies types are consistent with the probability maps offered by the experts. The generation of the initial ensemble is realized through a geological simulation model defined in the plurigaussian context, but not using of a truncation map in the multi-Gaussian space. Here we define adaptive simulation maps in a domains of measure 1 situated in multi-dimensional real spaces ([0,1] interval for one dimensional real space, $[0,1]^2$ for bi-dimensional real space, etc.). The simulation in this domain is carried out using objects defined by the projection of the Gaussian fields with the normal cumulative distribution function. The cumulative distribution function (cdf) links the Gaussian variables to the uniform variables (the probability integral transform Casella and Berger (2002)). The maps are adaptive in the sense that they are built in correlation with the probability maps provided by the experts using a layout that yields continuous modifications for neighbor grids. Because the initial probability maps incorporate the facies observations at the well locations (in probabilistic terms) the simulations always provide the correct facies types at the observation points. Therefore, throughout the assimilation period, the facies observations are kept for all ensemble members, without an extra procedure for repositioning (as in the previous studies, Liu and Oliver 2005, Agbalaka and Oliver 2008, Sebacher et al. 2013). When the data assimilation process starts, the initial ensemble contains facies distributions, with a large bandwidth for facies proportions. To ensure that the EnKF update will drive the ensemble towards a correct path, we introduce facies proportions as measurements, at each assimilation time step. The Gaussian fields are initially generated without any constraints (like in other plurigaussian approaches where the Gaussian fields were generated constrained to facies observations), being completely unconstrained and independents. The novelty consists on in how the expert knowledge is directly used as valuable prior information and is kept throughout assimilation period, using a simulation model that is able to create a link between the mathematical theory and real world.

The model is tested using a 2-D synthetic reservoir case. In the next sections we present the description of the geological simulation model, the EnKF implementation for the facies update and two examples where we test the model. The first example contains the case of reservoirs with three facies types each two could be in contact and the second example is for the case where four facies types generate a more complex geology because two of them are not in contact.

5.2 The geological simulation model.

5.2.1 The Motivation

Let us assume that we have a prior description of the subsurface geology in a certain reservoir, from different sources (seismic surveillance, core interpretations, outcrops, etc.). This prior knowledge includes for example, the number of the facies types that occurs and the possible contact (transition) between the facies types. This prior information is the result of reservoir exploration and is gathered by the experts (geologists, geophysicists, etc.) at the initial stage of deposit geology description. Several methods have been proposed to estimate probability maps for the facies occurrence by combining seismic data, well logs facies analysis, statistical rock physics (Avseth et. al 2001, Massonnat 1999, Sebastian Ng et. al 2008). However, these probability maps are not conditioned to the reservoir production outcome. Once the reservoir starts to produce and new information becomes available, an update of the probability maps is needed. To obtain an update of the probability maps, a solution is to couple a geological simulation model with an ensemble based AHM method. The geological simulation model should be able to generate an ensemble of facies maps conditioned to all prior information provided. In our case, the prior information refers to the prior probability maps, the possible contacts among facies types, the geometrical and the topological properties of the facies (the relative position among facies). In other words, we need to generate an ensemble of facies distribution, geologically plausible, of which probability maps are consistent with the probability maps provided by the experts. The challenge is to define the geological simulation model in a plurigaussian simulation context, that can conditions the topological prior knowledge to facies probability maps derived from expert knowledge. The benefit of the plurigaussian simulation model is the parameterization of the facies maps that is straightforward with the Gaussian fields. In addition, the geological acceptability is easily preserved in updates when proper AHM methods are used.

5.2.2 The adaptive plurigaussian simulation model (APS)

To be able to give a clear and precise description of the method, we present a case where three facies types occur in a reservoir domain. The approach will be applied in Section 5.4.2 for cases with three and four facies types.

The facies simulation map

In Figure 5.1, an example is shown with three possible probability maps, provided by the experts. In each map, we have value 1 at the well locations where the associated facies type occurs and value 0 at the well locations where the associated facies type

does not occurs (the facies observations at the wells location are incorporated in the maps). Another important piece of information is related with the transition (contact) between any two facies type (if there is contact or not). Suppose that any two facies types might intersect each other, in the domain. We denote by p_i the probability field

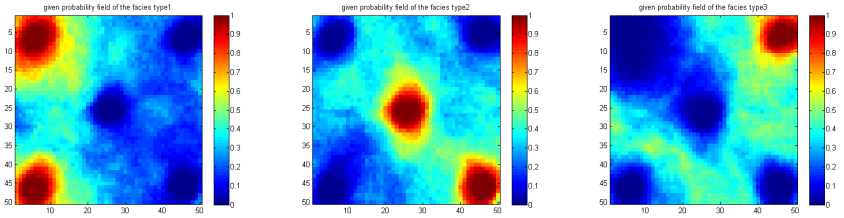


Figure 5.1: The probability map of each facies type provided by the experts

associated to facies type i (where $i = 1, 2, 3$), of which mapping is presented in Figure 5.1. For each grid cell j, we consider the discrete random variable, denoted "facies type" that express the distribution of the facies types at that location. The distribution of the random variable can be described with the equation

$$\textbf{facies type}^j \sim p_1^j \cdot \delta(\textbf{facies type}_1) + p_2^j \cdot \delta(\textbf{facies type}_2) + p_3^j \cdot \delta(\textbf{facies type}_3) \tag{5.1}$$

where δ is the Dirac function and p_1^j, p_2^j, p_3^j are the probabilities collected at the grid cell j from the probability maps provided by the experts. To sample from this discrete distribution we have to take a few steps:

1. Consider a domain A of measure 1, in a multidimensional real space X.
2. Split the domain in n subdomains of which measures are given by the probabilities, assigning a facies type to each subdomain.
3. Generate an element from a uniform distribution with support on A.
4. Assign the object depending on the subdivision of A where the generated element belongs to.

For instance, we can consider the domain A as $[0, 1]$ interval in the uni-dimensional real metric space or the $[0, 1]^2$ square in the real bi-dimensional metric space and, so on. The choice of the multidimensional real space should be done considering the topological properties of the facies types (the contact among facies type). In our particular case, knowing that each facies type can intersects any other facies type in the field we decide to work in the bi-dimensional real space, choosing the square $[0, 1]^2$ as the domain of measure (area) 1. Then, for each grid cell of the domain, we create a decomposition

of the square $[0,1]^2$ such that each subdomain is in contact with all other subdomains (Figure 5.2). In this quadratic domain, the total area is 1 and the area of each subdomain is equal to the associated probability found in the maps provided by the experts. We call this decomposition the facies simulation map. Consequently, each grid cell has his own simulation map because its form depends on the facies types probabilities (eq. 5.1). However, we need to design a parameterization of the simulation map layout, in order to obtain continuous modifications when simulate facies types for neighboring grids. The

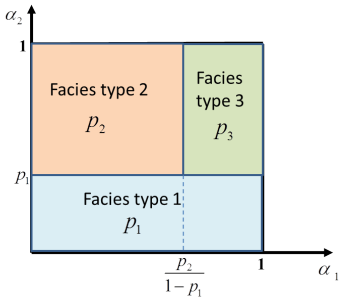


Figure 5.2: The facies simulation map layout

parameterization of the simulation map refers to the parameterization of the curves that delimitate the subdomains. In our specific case these are the parameters of the vertical and horizontal lines that separate the subdomains (Figure 5.2). As consequence of the parameterization, at the well locations, the map consists of a single square of surface 1, occupied by the facies type observed there. In addition, there might be some regions of the reservoir domain where only two facies types occur (in terms of probabilities, only two probabilities are greater than zero). In this case, the map contains only two rectangular regions, each with an area equal to the associated probability of the facies type that occupies that region.

The spatial correlated fields from [0,1] interval (the uniform random fields)

In each grid cell j, we need to simulate a facies type, conditioned to the probabilities given by the experts. Hence, we need to simulate from the distribution defined by eq.5.1. For that, having designed a simulation map, we generate two independent random numbers α_1^j, α_2^j from [0,1] interval; we represent the point (α_1^j, α_2^j) in the simulation map and, depending on which subdomain the point belongs to, we assign in the grid cell the corresponding facies type. To extend this approach to the reservoir domain and keep a well/realistic defined shape zones the fields α_1, α_2 must have spatial correlation. Consequently, we need to define spatial correlated fields, defined on the reservoir domain, with values in $[0,1]$ interval. These spatially correlated fields can be modeled by piece-

wise projection into [0,1] of any continuous distributed, spatial correlated random fields defined on the reservoir domain. The condition to be satisfied is that in each grid cell, at the initial ensemble simulation, the distribution of the variable **"facies type"** calculated from the ensemble, has to be consistent with the given distribution (eq.5.1). This condition is fulfilled if each component of the spatial correlated field (from [0,1] interval) is drawn from an uniform distribution with support on [0,1] interval. Consequently, we call these spatial correlated fields uniform random fields. In order to define these spatial correlated fields we use the probability integral transform (Casella and Berger (2002) [5]). The probability integral transform states that if X is a continuous random variable with the cumulative distribution function F_X, then the random variable $F_X(X)$ has an uniform distribution on $[0, 1]$. The cumulative distribution function links any continuous random variable with the uniformly distributed random variable on $[0, 1]$. Therefore, we can define the spatial correlated fields (having values in $[0, 1]$) as piece-wise projection of any spatial correlated fields (defined on the reservoir domain), of which distribution is continuous, through their cumulative distribution function. We choose to model these fields as piece-wise Gaussian fields projection because of two important reasons:

1. The generation of the Gaussian random fields can be done easily using any geostatistical tool.
2. The HM method used in this study (the EnKF) works well under the Gaussian assumption of the model parameters.

In this study we use two properties of the cumulative distribution function (cdf). The first property is to transfer the geostatistical properties of the Gaussian fields to the spatial correlated fields from the $[0, 1]$ interval; the second property is the consequence the probability integral transform. The projection with cdf ensures that, at the initial step of facies fields generation the probability map of each facies type is consistent with the originals.

The geological simulation model (the APS)

The geological simulation model defined in the plurigaussian context that samples from prior facies probability maps can be summarise as follows:

1. Design the layout and the parameterization of the facies simulation map.
2. Generate samples of two independent and unconditioned Gaussian random fields defined on the reservoir domain, denoted y_1 and y_2.
3. Calculate the spatial correlated fields (uniform random fields) $\alpha_1 = cdf(y_1)$ and $\alpha_2 = cdf(y_2)$ projecting component-wise the Gaussian fields in $[0, 1]$ using the cdf of the normal distribution used for the Gaussian fields generation.

4. Define a facies type in each grid cell j of the reservoir domain.

   - Design the facies simulation map associated to grid cell j using the parametrization of the simulation map.
   - Assign the facies type associated to the subdomain, in which the realization (α_1^j, α_2^j) falls in the simulation map.

We call this geological simulation model the adaptive plurigaussian simulation model (APS). The simulation scheme is adaptive in the sense that each grid cell has his own simulation map, depending on the pair of probabilities. At the well locations the simulation will provide each time the same facies type observed there, because the $[0,1]^2$ square is completely occupied by a single facies type. However, the main result of the simulation model is that, if we generate an ensemble of unconditioned and independent pairs of Gaussian fields, the facies probability maps calculated from the ensemble are consistent with the maps provided by the experts. This can be observed with a visual inspection of the Figure 5.3. At the top of this figure are presented the prior probability maps of the facies occurrence that comes from expert knowledge (the same maps with the ones presented in Figure 5.1) and at the bottom of the figure the probability maps calculated from an ensemble of 120 facies fields realizations generated with the APS model. This result is the consequence of the lemma presented in the Appendix of the

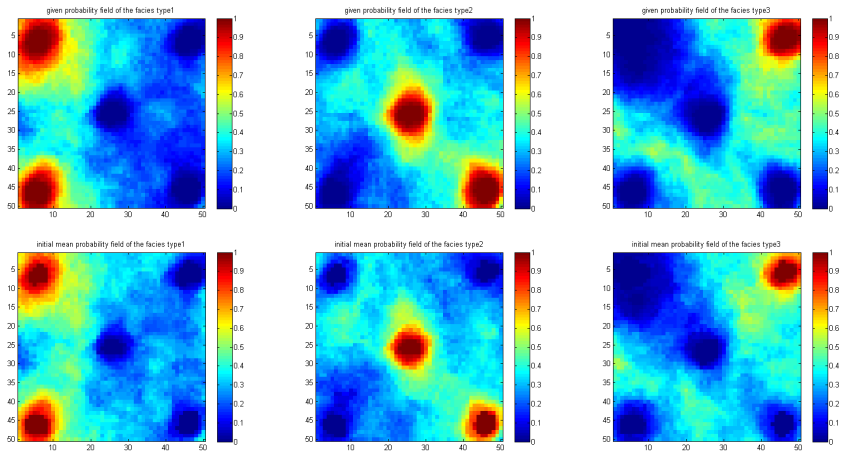


Figure 5.3: The probability map of each facies type provided by the experts (top) and the probability maps calculated from an ensemble of realizations (bottom)

paper coupled with the Monte Carlo approach of sampling from Gaussian distributions. The APS model can be generalized to as many facies type we want, to any types of

contact between facies types. For that, all we have to do is to define a reliable decomposition (parameterization) of the $[0,1]^2$ square. If the connectivity is very complex we may carry out the decomposition to higher dimension, using the domain $[0,1]^3$ and three Gaussian fields for facies maps simulation.

5.2.3 The impact of the facies simulation map

The geological acceptability of the simulated facies distributions is achieved by a reliable choice of the geostatistical properties of the Gaussian fields and of the simulation map layout. Usually, in the plurigaussian simulation model, the geostatistical properties of the Gaussian fields used for facies simulation are set based on the prior topological characteristics available for the facies types. We refer here at the geometry of the facies type, the number of the facies (rock bodies) that occurs in domain and the orientation of the facies.
In the following, we present an example where, small changes of the simulation map layout, yield different topological characteristics for the simulated facies maps, when preserved the geostatistical properties of the Gaussian fields unchanged. The contact among facies types can be handled by a reliable definition of the decomposition of the unitary domain. However, the relative position among facies types cannot be handled with any decomposition. Figure 5.4 shows four possible realizations of the facies distributions for reservoirs characterized by the existence of three different rock types, where any two could be in contact with each other. The green facies type is characterized by long anisotropy of which angle with the horizontal direction is obtuse. In contrast, the red facies type is characterized by a large amount of small bodies within the reservoir domain. The blue facies type is the geological environment in which the other two facies types are propagated. These geological realizations were generated using the same values of two Gaussian random fields, the first is anisotropic with long range correlation of 50 grid cells, short range correlation of 5 grid cells and with the principal direction of 120° with the horizontal. The second Gaussian random field was generated isotropic with a small range correlation of 5 grid cells. The parameterization and the layout of the facies simulation maps are presented at the bottom part of Figure 5.4. We have applied the adaptive plurigaussian simulation model, using uniform probability maps of 0.4 for the facies type 1 (blue) and type 2 (green) and respectively 0.2 for the facies type 3 (red). Even though there are small differences among the simulation map layouts, the facies realizations provided exhibit different topological properties. We refer here at the relative position of the red facies type with respect to the all facies types. In the first simulation, bodies of the red facies type can be inside both the green and blue facies types. In the second example, we may find bodies of the red facies type within the blue facies type, whereas the green facies type intersects the red facies type only on its border. In the third

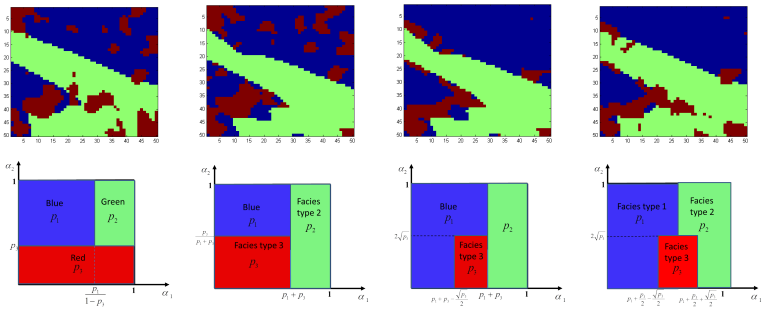


Figure 5.4: Example of four different simulation maps and theirs respective simulations

simulation, the deposits of red facies type are encountered mainly between the borders of the other two facies types, eroding the blue facies type. In the fourth simulation, the relative position of the red facies type with respect to the other facies types exhibits similar characteristics, remaining mainly on the edges of the borders. In this formation, the red facies type has eroded both other facies types. Responsible for these simulations is the layout of the simulation map, the way in which the subdomain assigned to the red facies type has been defined in the $[0, 1]$ square. Consequently, the design of the facies simulation map and the choice of the geostatistical properties of the Gaussian fields must be carried out by a group of experts, involving multidisciplinary disciplines (geology, mathematics, reservoir engineering, etc.) in such way that the simulation results reflects as close as possible reality.

5.3 Ensemble Kalman Filter (EnKF) implementation

Concerning the EnKF implementation, we define the state vector as:

$$\mathbf{x} = \left[y_1^\top \; y_2^\top \; \mathbf{d}_{sim}^\top \right]_j^\top , \tag{5.2}$$

where y_1 and y_2 are the Gaussian random fields used for the facies distribution simulation; $\mathbf{d}_{sim}$ are the simulated observations represented by the simulated production data e.g. the oil production rates (OR), the water production rates (WR) and the bottom hole pressures (BHP) together with the facies proportions (for some experiments). Consequently, the state vector for the ensemble member j at the k^{th} assimilation step can be defined as:

$$\mathbf{x}_j^k = \left[y_1^\top \; y_2^\top \; \text{BHP}^\top \; \text{WR}^\top \; \text{OR}^\top \; \text{facies\_prop}^\top \right]_j^{k,\top} \tag{5.3}$$

We define the statistical measure facies_prop for the ensemble member j, augmenting

the facies proportions calculated for each facies type i according to

$$\text{facies\_prop}^{\mathbf{i}}_{\mathbf{j}} = \frac{1}{n_g} \sum_{k=1}^{n_g} Ind_i(k) \tag{5.4}$$

the summation being performed for all grid cells (n_g grid cells), where the indicator function that express the belonging of the grid cell k to the facies type i is calculated as

$$Ind_i(k) = \begin{cases} 1 & \text{if } k \in \text{facies type i} \\ 0 & \text{if } k \notin \text{facies type i} \end{cases} \tag{5.5}$$

The state vector does not contain the petrophysical properties (permeability, porosity) of the facies type because we keep these values constant throughout assimilation period. We assume that the uncertainty in the system is due the poorly knowledge of the facies distribution in the field. The dynamical variables (the pressure and the saturation) are not updated together with the parameters because after each assimilation time we have chosen to rerun the simulator from time zero, even though this technique will increase the computational cost. This avoids the unphysical values of the dynamical parameters (e.g. negative pressures and saturations or saturations greater than 1) in the EnKF update and the inconsistency between the dynamical variables and the statical variables e.g. the permeability, porosity (the updated dynamical variables could not by an acceptable numerical solution of the partial differential equations describing the multi-phase flow, given the statical parameters). For small models used in our study this procedure is not very time consuming, but for large-scale models, the computational cost should be considered.

5.3.1 The Work Flow

- **Initialization**
  The uncertainty of the initial ensemble is given by the choice of the model parameters, which in our case are the Gaussian fields y_1 and y_2. Initially, we generate an ensemble of independent and unconditioned (stationary) pairs of Gaussian fields y_1 and y_2 (we have used the sequential Gaussian simulation technique implemented in S-GeMS [17]). In this study, the stationary Gaussian fields were generated with zero mean function, Gaussian covariance type and with each component having standard normal distribution. Then, using the adaptive plurigaussian simulation (APS) model, we generate the initial ensemble of facies maps. In this study, we have used 120 ensemble members in all the examples presented. The independence and the unconstraint properties of the Gaussian fields ensures generation of an ensemble of facies maps that represent accurately the prior uncertainty in the facies distribution (Figure 5.3).

- **Forecast step/Model update**
  We populate each facies map with proper petrophysical properties (permeability, porosity, etc.). In this study we have chosen to keep constant these values within each facies not taking into account the heterogeneity within each facies. For each ensemble member we apply the forward model from time zero until the next assimilation time. The forecasted state vector at time $k+1$ can be written as:

$$\mathbf{x}_j^{k+1,f} = \left[y_1^{f,\top}\ y_2^{f,\top}\ \text{BHP}^{f,\top} \text{WR}^{f,\top}\ \text{OR}^{f,\top}\ \text{facies\_prop}^{f,\top} \right]_j^{k+1,\top},$$

  where

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_j^{k+1,f} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_j^{k,a} \tag{5.6}$$

$$\begin{bmatrix} \mathbf{p} \\ \mathbf{s} \end{bmatrix}_j^{k+1,f} = \mathcal{M}\left(\begin{bmatrix} \mathbf{p} \\ \mathbf{s} \end{bmatrix}_0^{k}, \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_j^{k,a} \right) \tag{5.7}$$

$$\begin{bmatrix} \text{BHP} \\ \text{WR} \\ \text{OR} \end{bmatrix}_j^{k+1,f} = g_{pred}\left(\begin{bmatrix} \mathbf{p} \\ \mathbf{s} \end{bmatrix}_j^{k+1,f}, \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_j^{k+1,f} \right) \tag{5.8}$$

$$\begin{bmatrix} \text{facies\_prop} \end{bmatrix}_j^{k+1,f} = \begin{bmatrix} \text{facies\_prop} \end{bmatrix}_j^{k,a} \tag{5.9}$$

  where g_{pred} is the prediction operator for the production data. Hence, there are no changes from time k to $k+1$ for the values of the Gaussian random fields (eq. 5.6), because the forward model equations does not influence the Gaussian fields values. However, their values have an impact in the changes of the pressures and saturations at time $k+1$ based on the model (eq. 5.7) which generates changes in the forecast of the production data (eq. 5.8). This happens because the Gaussian fields values influence the spatial distribution of the permeability and porosity that generates modification of the dynamical variables (eq. 5.7) and of the production data (oil and water rates and pressures at the bottom of wells, eq. 5.8). The Gaussian fields values are adapted in the update step of the EnKF when the measurements are assimilated. The predicted values for the facies proportions will not changed, because the facies maps are not adapted during the forecast step (eq. 5.9).

- **Update step/Measurement update**
  At each assimilation time k+1, $k \in \{0, \ldots, n_i - 1\}$ (n_i is the number of the assimilation time steps) we assimilate the data (production data and facies proportions)

using the Kalman filter equations. Each ensemble member is modified according to:

$$\begin{aligned}\mathbf{x}_j^{k+1,a} &= \mathbf{x}_j^{k+1,f} + \mathbf{C}^f_{\mathbf{x}^{k+1}}\mathbf{H}^\top_{k+1}\left(\mathbf{H}_{k+1}\mathbf{C}^f_{\mathbf{x}^{k+1}}\mathbf{H}^\top_{k+1} + \mathbf{C}_{\text{obs},k+1}\right)^{-1} \\ &\quad \left[\mathbf{d}_{\text{obs}_j,k+1} - \mathbf{H}_{k+1}\mathbf{x}_j^{k+1,f}\right], \end{aligned} \tag{5.10}$$

where $\mathbf{H}$ represents observation operator for the observed data, $\mathbf{C}_{\text{obs}}$ represents the error covariance matrix for the observed data, $\mathbf{C}_{\mathbf{x}^f}$ represents the forecasted error covariance of the ensemble, and $\mathbf{d}_{\text{obs}_j}$ are the observed data (production data and facies proportions) of the ensemble member j

- **The updated facies maps**
  The update of the state vector using the measurements, implies modification of the Gaussian fields values (according to eq. 5.10). This update of the Gaussian fields values will change the facies maps. The updated facies maps are obtained by simulation from the probability maps given by the experts, but using the updated values of the Gaussian fields (conditioned to measurements). This means that we preserve the simulation maps designed at the beginning throughout assimilation process.
  Consequently, at this step, for each ensemble member i, at each grid cell j we assign the facies type depending on where the updated realization $(cdf((y_1^a)_i^j, cdf((y_2^a)_i^j)$ belongs in the simulation map designed for the grid cell j (cdf is the cumulative distribution function of the normal distribution used for generation of the Gaussian fields before any data assimilation is performed i.e. the cdf of the standard normal variables).

At the end of the assimilation period, we have obtained an ensemble of facies maps from which we calculate the updated probability map for each facies type. In addition, we have the uncertainty quantification for the geology description provided by the updated ensemble of the facies maps. The initial ensemble of the Gaussian fields pairs has initially a high variability (the Gaussian fields are generated unconditioned and independent) which decrease with the assimilation of the production data and facies data (in this case the facies proportions). The variability of the Gaussian field ensemble can be quantified using an indicator that measures its mean variance:

$$\sigma_Y^2 = \frac{1}{n_g\left(n_e - 1\right)}\sum_{j=1}^{n_e}\sum_{i=1}^{n_g}\left(Y_i^j - \overline{Y_i}\right)^2 \tag{5.11}$$

where, n_g is the number of the grids, n_e is the number of the ensemble members and $\overline{Y_i}$ is the mean value of the random field Y in the grid i

$$\overline{\mathbf{Y}}_i = \frac{1}{n_e} \sum_{j=1}^{n_e} Y_i^j \tag{5.12}$$

We calculate the mean variance, independently for both Gaussian fields y_1 and y_2, plotting theirs evolution during the assimilation period. The evolution of this indicator provides information about the convergence or collapse of the ensemble (Agbalaka and Oliver 2008, Sebacher et al 2013).

5.4 Case studies

In this section, we present two applications of the adaptive plurigaussian simulation method. In the first example we consider a reservoir where three facies types occur, any two of them could be in contact with the others. Schematically, we may represent these characteristics as $F_1 \Leftrightarrow F_2$, $F_2 \Leftrightarrow F_3$ and $F_3 \Leftrightarrow F_1$. In the second example, we present a case where in the reservoir four facies types occur, while two of the facies types do not have contact. For both cases presented, we consider the geostatistical properties of the Gaussian fields and the layout of the facies simulation map known. The reference fields and the prior probability maps of the facies occurrence are generated with the geological simulation model presented in Sebacher et al. 2013 .

5.4.1 The model with three facies types

We consider a model with three facies types with the distribution presented in the previous sections. We perform two experiments. In first experiment, we consider as measurements only the pressure at the injector and the oil and water rates at the producers. In the second experiment, we introduce as measurement the facies proportions. The experts have knowledge about these values, together with a uncertainty bandwidth which are translated into stochastic elements as measurement errors for facies proportions.
The simulation model is a 5-spot water flooding 2D-reservoir, black oil model with $50 \times 50 \times 1$ active grid cells. The dimension of each grid cell was set at $30 \times 30 \times 20$ ft and there is one injector situated at the center of the reservoir domain and there are four producers situated at the corners. The coordinates of the wells positions are presented in Table 5.1 and the facies observations are presented in Table 5.1. The values of the permeability and porosity corresponding to each facies type are shown in Table 5.2.

Water injection starts from the first day and continue thereafter a period of 201 days of production. We assimilate data starting from first day until day 201, with 10-day interval, having 21 assimilation steps. All the producers work under constant BHP value of 1500 pound per square inch (psi), the injector operates at 4500 stock tank barrels per

Table 5.1: The position of the wells in the reservoir domain and the facies observations

	Injector	Producer 1	Producer 2	Producer 3	Producer 4
x coordinate	25	5	5	45	45
y coordinate	25	5	45	5	45
Facies observation	Type 2	Type 1	Type 3	Type 1	Type 2

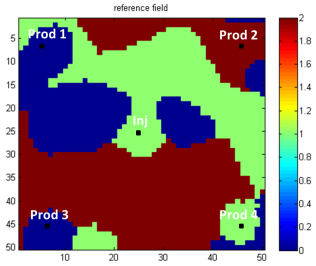


Figure 5.5: The reference field

Table 5.2: The petrophysical properties of each facies type

Facies type	Permeability	Porosity
Type 1	174 md	0.18
Type 2	372 md	0.25
Type 3	80 md	0.14

day (STB/D) and constrained by a maximum 5000 psi for BHP. The measurement error for the production data is considered Gaussian with 0 mean and standard deviations respectively 10 STB/D for WR and OR at the producers and 30 psi for BHP at the injector. These values will be used for the generation of the noise, which is added to the observations of the production data in the analyzed step of the EnKF process. The observations of the production data are obtained through forward integration of one model defined as reference field (or "truth field"). The reference field is presented in Figure 5.5, where, the blue colour is associated with facies type 1, the green colour is associated with facies type 2 and the red with the facies type 3. The probability maps provided by the experts are presented in the top of Figure 5.3 and the probability maps calculated from the ensemble in the bottom of Figure 5.3. For the facies maps simulation, we have used isotropic Gaussian fields with Gaussian variogram type with the correlation range of 17 grid cells (one third of domain length). The layout of the facies simulation map used is the one presented in Figure 5.2. When EnKF is applied, we expect that the reduction of

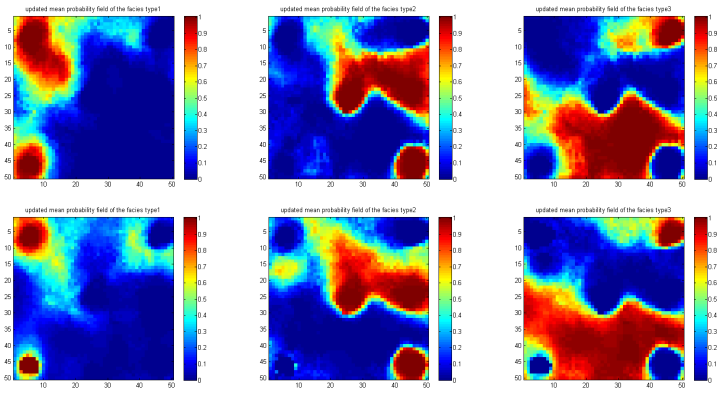


Figure 5.6: The probability maps for the case without facies proportions as measurements (bottom) and with facies proportions as measurements (top)

the variability due to the data assimilations results in an updated ensemble which spread to contains the "truth". In case of the facies distribution estimation two things are needed in the updated ensemble. One is that the updated fields reproduce the major features of the truth and the other is related to an improved predictability of the models. If the latter is easily fulfilled, the more difficult part is the reproduction of the facies geometry. Here an important role is played by the facies proportions specially when, the experts, based on preliminary studies, have some knowledge. Therefore, we perform an experiment where the facies proportions are used as measurement at each assimilation point, with an associated error, also given by the experts. In this study the facies proportions used as observations are set up as 0.25:0.30:0.45 with a measurement error following a normal distribution with mean 0 and 0.05 standard deviation. Although, we have used this information as measurement, we do not have observed a severe reduction in uncertainty comparing with the other case. This means that, this observation has the role of guiding the ensemble to some predefined values (of which the experts have knowledge). In Figure 5.7 is shown the evolution of the mean variance of the Gaussian fields in the assimilation period for both experiments (at the right side, for the experiment where facies proportions are not introduced in state vector as measurements). We observe that, the impact of facies proportions is not severe, the influence is more significant for the first Gaussian field, but the final values are comparable. The updated facies probability maps for the case where the facies proportions are not included as measurements are in Figure 5.6 (top line) and the updated facies probability maps with facies proportions used as measurements at the bottom line of the same figure. In this example, one may observe that, the averages fields capture the main features of the reference field the only difference consists on the facies proportions existing in the updated ensembles.

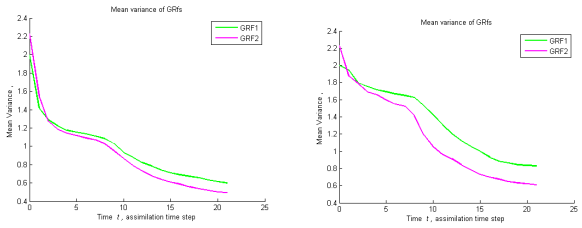


Figure 5.7: The evolution of mean variance of Gaussian felds in both experiments

Table 5.3: Facies proportions in the ensembles

	No proportions in state	Proportions in state
Reference field	0.26:0.30:0.44	0.26:0.30:0.44
Facies proportions initial state	0.365:0.344:0.291	0.365:0.344:0.291
Standard deviation initial state	0.17:0.16:0.18	0.17:0.16:0.16
Max proportions initial state	0.81:0.82:0.78	0.81:0.82:0.78
Min proportions initial state	0.02:0.03:0.02	0.02:0.03:0.02
Facies proportions updated state	0.195:0.34:0.465	0.246:0.306:0.447
Standard deviation updated state	0.08:0.04:0.06	0.022:0.024:0.02
Max proportions updated state	0.52:0.42:0.55	0.31:0.35:0.52
Min proportions updated state	0.08:0.22:0.22	0.18:0.26:0.38

Knowing that the true facies proportions are 0.254:0.30:0.446, in the Table 5.3 the expected facies proportions calculated from the ensemble together with theirs standard deviations, minimum values and maximum values are shown. We observe from this table, that the spread of the initial ensemble is high enough to contain the facies proportions values of the "truth" field. As expected, when EnKF is applied without facies proportions as measurement, the spread of the updated ensemble is higher than the spread of the other experiment. It can be seen from Table 5.3, that the spread contains the facies proportions of the "truth", but the expected facies proportions calculated from the updated ensemble is not that close to the truth. In the other experiment, where facies proportions are used as measurements, the expected facies proportions are much closer to the truth and the spread around the mean is within the bounds that the experts provide us. For the experiment with facies proportions used as measurements, the updated production rates are presented in Figure 5.8). From this figure it can be seen a good data match for both oil and water rates in all wells. At the top of Figure 5.9 we present the first four members of the facies maps in the initial ensemble and, at the bottom part, the same members in the updated ensemble (for the experiment with facies proportions used as observations). It can be seen by a visual comparison with the truth (Figure 5.5)

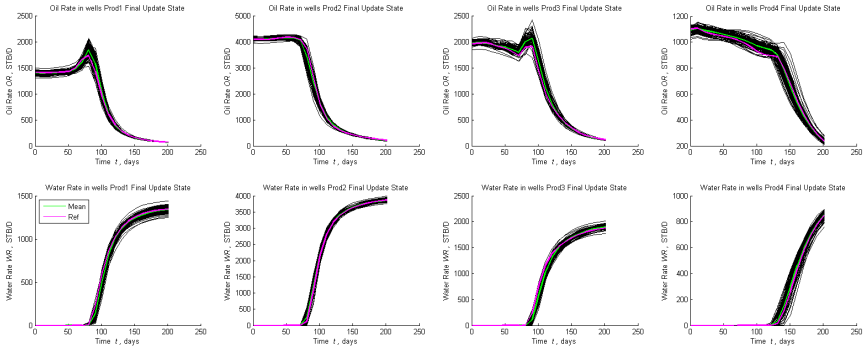


Figure 5.8: The updated rates at the producers (the case with facies proportions used as observations)

that the updated ensemble members capture its major features, preserving the topology. Even though we have used 21 assimilation time steps where, besides the production data the facies data have been assimilated, the ensemble does not collapsed. This can be seen by examining the evolution of the mean variance in the assimilation period where, is observed an asymptotic behavior to a value close to 0.5 (Figure 5.7).

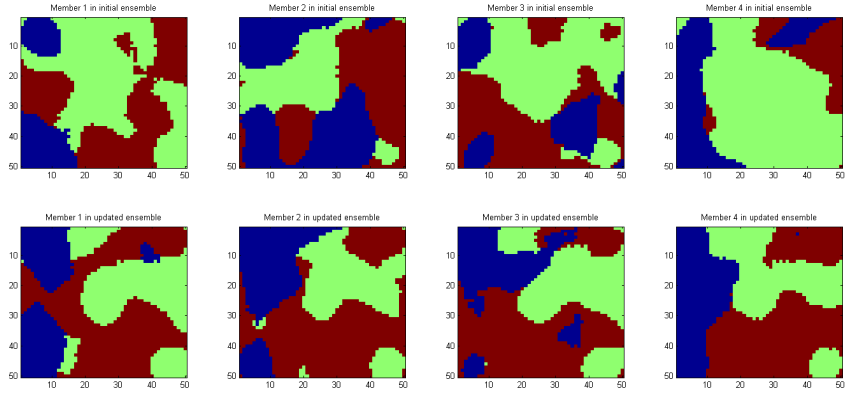


Figure 5.9: First four ensemble members in initial and updated ensemble (the case with facies proportions used as observations)

5.4.2 The model with four facies types

In this example, we consider a reservoir where four facies types occur, but of which two are not in direct contact with each other. This means that between them always is interposed another facies type. We denote the facies type F_1, F_2, F_3 and F_4 and the possible contacts are $F_1 \Leftrightarrow F_2$, $F_2 \Leftrightarrow F_3$, $F_1 \Leftrightarrow F_4$, $F_2 \Leftrightarrow F_4$, $F_3 \Leftrightarrow F_4$ (F_1 and F_3 are not in contact). The petrophysical properties of each facies type are presented in Table 5.4. The simulation model is an 8-spot water flooding 2D-reservoir, black oil

Table 5.4: The petrophysical properties of each facies type

Facies type	Permeability	Porosity
Type 1	2 md	0.1
Type 2	10 md	0.2
Type 3	50 md	0.2
Type 4	250 md	0.3

model with $100 \times 50 \times 1$ active grid cells. The dimension of each grid cell was set at $30 \times 30 \times 20$ ft. and there are two injectors situated at the center of the reservoir domain and there are six producers, four situated at the corners and two situated between the other four. The reference field used in this experiment is presented in Figure 5.10, where blue shade represents facies type 1, light blue shade represents facies type 2, yellow shade represents facies type 3 and red shade represents facies type 4. The light blue dots represent the positions of the producers and the light red dots represent the positions of the injectors. The coordinates of the wells positions and the facies observations are presented in Table 5.5.

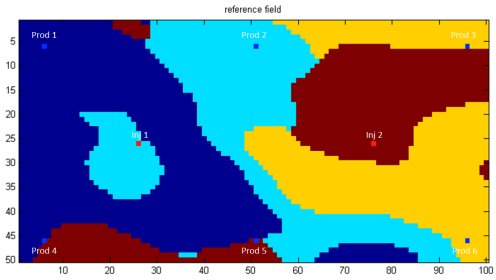


Figure 5.10: The reference field

Water injection starts from the first day and continues throughout 231 days of production. We assimilate data starting from first the day until day 231, with a 20-day inter-

Table 5.5: The position of the wells in the reservoir domain and the facies observations

	Inj 1	Inj 2	Prod 1	Prod 2	Prod 3	Prod 4	Prod 5	Prod 6
x coordinate	25	75	5	50	95	5	50	95
y coordinate	25	25	5	5	5	45	45	45
Facies observation	Type 2	Type 4	Type 1	Type 2	Type 3	Type 1	Type 4	Type 3

val and with 12 assimilation steps. All the producers work under constant BHP value of 2500 pound per square inch (psi), the injector operates at 4500 stock tank barrels per day (STB/D) and constrained by a maximum 100000 psi for BHP. The measurement error for the production data is considered Gaussian with 0 mean and standard deviations respectively 10 STB/D for WR and OR at the producers and 70 psi for BHP at the injector. From the exploration process the experts provide us four probability maps, associated with each facies type, maps that show the possible position of each facies type in the field (top of Figure 5.11). The expected facies proportions used as measurements were set at 0.35:0.25:0.25:0.25 whilst, the true facies proportions are 0.34:0.24:0.22:0.20.

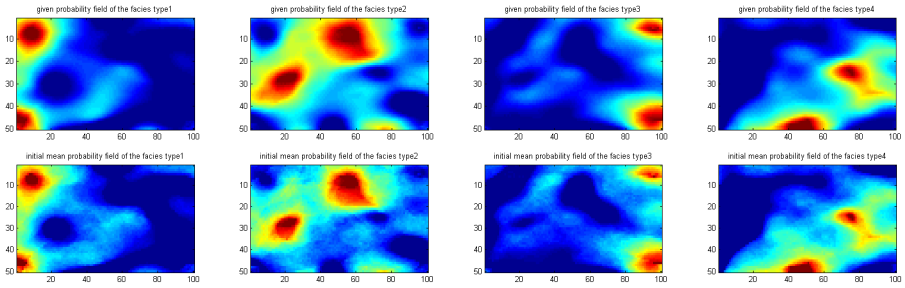


Figure 5.11: The probability maps provided by the experts (top) and The probability maps in initial ensemble (bottom)

In order to apply the adaptive simulation model, for each grid cell, we construct a decomposition of the $[0,1]^2$ square, like in Figure 5.12, where the area of each zone will is equals with the probability found in the maps provided by the experts. The Figure 5.12 shows also the parametrization of the simulation map, e.g. the parameters of the vertical and horizontal lines (two on $0\alpha_1$ axis and one on $O\alpha_2$ axis). This decomposition of the $[0,1]^2$ square ensures that the sub-domains associated to facies type 1 and facies type 3 are not in contact. We simulate the initial ensemble of facies fields (conditioned to the probability maps) generating an ensemble of 120 unconditioned and independent pairs of Gaussian fields. For this example we have used anisotropic Gaussian fields with Gaussian variogram type, having a long range correlation of 35 grid cells, a short range

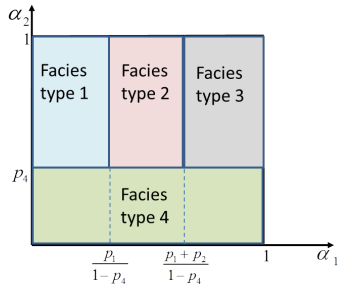


Figure 5.12: The layout and the parametrization of the simulation map for the case with four facies types

correlation of 17 grid cells and with the principal direction of $0°$. The probability maps of each facies type calculated from the initial ensemble are shown in the bottom part of Figure 5.11 and it can be seen that this is a good reproduction of the maps provided by the experts (the top part of Figure 5.11). After 12 assimilation time steps, the probability maps of each facies type calculated from updated ensemble (Figure 5.13, middle part), emphasize much better the real position of the facies types in reservoir domain (Figure 5.13, top part). The mean of the facies proportions calculated from the updated ensemble is 0.33:0.23:0.21:0.23, close to the real facies proportions 0.34:0.24:0.22:0.22. We have performed the experiment without assimilating the facies proportion as measurements and the results look worse. In the bottom part of Figure 5.13 the updated probability maps of the facies types in this experiment are shown. As we can see the facies types 2, 3, and 4 are not correctly estimated having the facies proportions mean in the updated ensemble of 0.41:0.17:0.11:0.31 far away from the true facies proportions.

In Figure 5.14 are presented the water rates forecast for 360 days in the initial ensemble (top) and in the updated ensemble for both experiments performed e.g. with facies proportions used as measurements (middle part) and without (bottom). The blue vertical line is the threshold that delimitates the assimilation period to the prediction period of 120 days. It can be seen the reduction in variability together with the improvement in prediction for all producers for both experiments. If we consider only the reduction in variability coupled with the production data match and prediction we cannot rank the quality among the experiments. Without conditioning to facies proportions, we obtain very good forecast with topologically plausible facies maps realizations but without realistic realizations regarding the statistical measure. If we look at the results of the experiment where the ensemble was conditioned to facies proportions we have obtained good data match, good predictions, plausible topologically realizations (Figure 5.15) of which facies maps are in concordance with the prior knowledge about the global facies proportions. Therefore, conditioning the ensemble to this statistical measure guides the

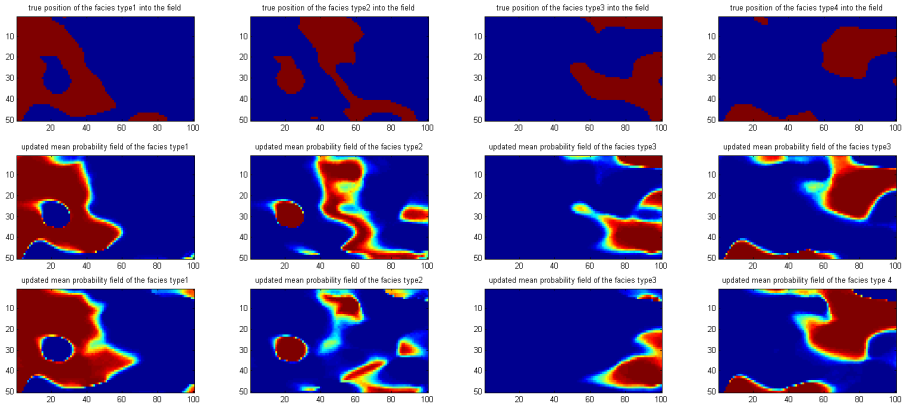


Figure 5.13: The real position of the facies types in field (top), the probability maps in updated ensemble with facies proportions used as measurements (middle) and the probability maps in updated ensemble without facies proportions used as measurements (bottom)

ensemble towards a correct path, providing in the same time good predictions and data match.
Regarding the Gaussian fields spread we can observe the same asymptotic behavior for the mean variance of the Gaussian fields as in the first case presented. In addition, the final variability of the Gaussian fields is comparable in both experiments performed (Figure 5.16), the ensemble collapse being avoided.

In the Figure 5.15, we show four members of the initial ensemble (top) and in the updated ensemble (bottom; the experiment with facies proportions used as measurements). As expected, comparing with a visual inspection this figure with the reference field (Figure 5.10), the important patterns of the reference field can be visualized in all updated members. In addition, the topological request (the lack of contact among yellow facies type and the dark blue facies type) is fulfilled in the updates.

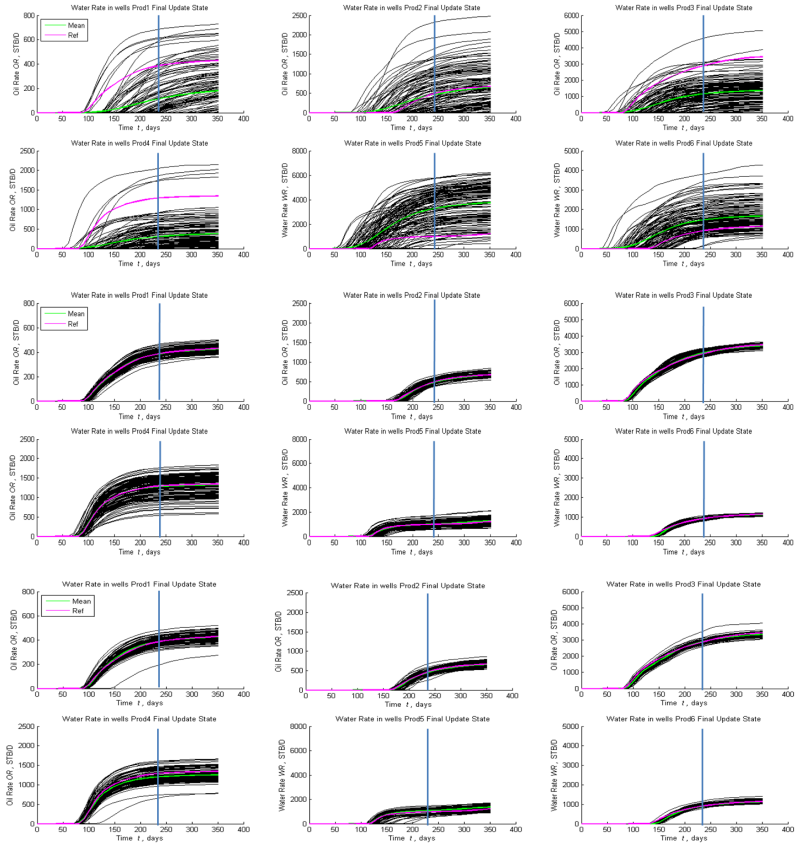


Figure 5.14: The Water rates forecast in initial ensemble (top), in updated ensemble conditioned to facies proportions (middle) and in updated ensemble unconditioned to facies proportions (bottom), for 360 days

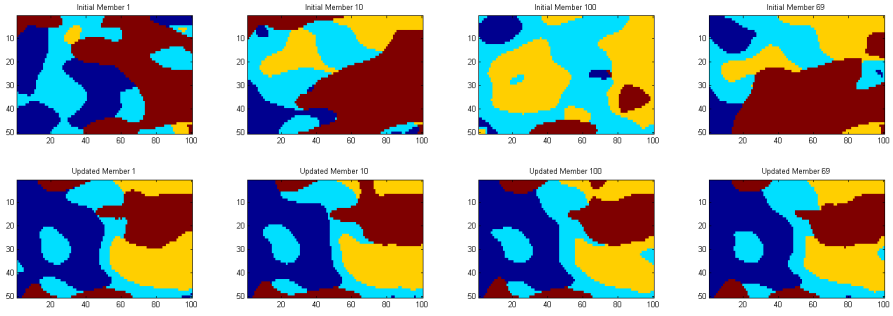


Figure 5.15: Four ensemble members in initial and updated ensemble

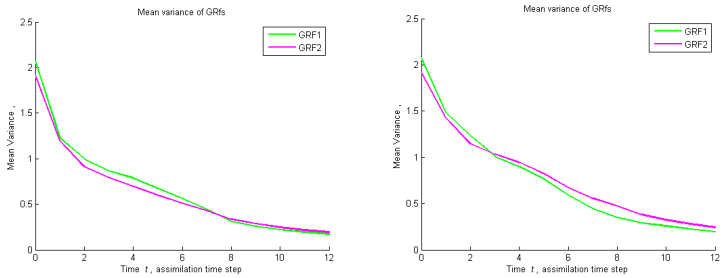


Figure 5.16: The evolution of mean variance of Gaussian fields in both experiments

5.5 Appendix: The probability integral transform

The probability integral transformation links any continuous distribution with the uniform distribution with support on $[0, 1]$ interval. The result of this transformation is given by the following theorem.

Theorem (Probability integral transform)

Let X be a random variable with continuous distribution of which cumulative distribution function is $F_X(X)$ and define the random variable Y, $Y = F_X(X)$. Then the random variable Y is uniformly distributed on $[0, 1]$.

The proof of this theorem can be found in Casella and Berger (2002). An application of this theorem is the following Lemma which helps in the implementation of the model presented before.

Lemma: Let D a sub-domain of the square $[0, 1]^2$, and two independent random variables $z_1 \sim N(0; 1), z_2 \sim N(0; 1)$. Then $P((cdf(z_1), cdf(z_2)) \in D) = area(D)$.

Proof: Let's consider an arbitrary sub-domain, denoted D, of the unitary square (Figure 5.17) and the function $\varphi : \mathbf{R}^2 \to (0, 1)^2$, $\varphi(z_1, z_2) = (cdf(z_1), cdf(z_2))$. This function is bijective, differentiable with partial derivatives continuous and with his inverse having the same properties (a C^1 difeo-morphism). We denote $D' = \varphi^{-1}(D)$ and we might write that $(\alpha_1, \alpha_2) \in D \Leftrightarrow \varphi^{-1}(\alpha_1, \alpha_2) = (cdf^{-1}(\alpha_1), cdf^{-1}(\alpha_2) \in D')$.

Then,

$$area(D) = \iint_D d\alpha_1 d\alpha_2 \tag{5.13}$$

In this double integral we perform a change of variables according to $\alpha_1 = cdf(z_1)$ and $\alpha_2 = cdf(z_2)$. We get

$$\iint_D d\alpha_1 d\alpha_2 = \iint_{D'} |Det(J_{(z_1,z_2)}(\alpha_1, \alpha_2))| dz_1 dz_2 \tag{5.14}$$

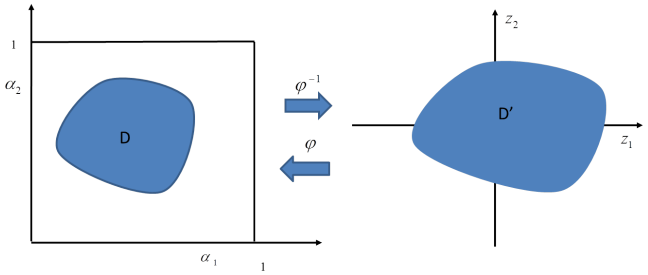


Figure 5.17: The relation between the domain in the unitary square and the domain from bi-dimensional real space

, where, J is the Jacobian for φ, the function of the variables change. On the other hand,

$$J_{(z_1,z_2)}(\alpha_1,\alpha_2) = \begin{pmatrix} \frac{\partial cdf(\alpha_1)}{\partial z_1} & \frac{\partial cdf(\alpha_1)}{\partial z_2} \\ \frac{\partial cdf(\alpha_2)}{\partial z_1} & \frac{\partial cdf(\alpha_2)}{\partial z_2} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2\pi}} e^{-\frac{z_1^2}{2}} & 0 \\ 0 & \frac{1}{\sqrt{2\pi}} e^{-\frac{z_2^2}{2}} \end{pmatrix} \tag{5.15}$$

Therefore,

$$\iint_D d\alpha_1 d\alpha_2 = \frac{1}{2\pi} \iint_{D'} e^{-\frac{z_1^2}{2} - \frac{z_2^2}{2}} dz_1 dz_2 = P((z_1, z_2) \in D^{'}) \tag{5.16}$$

The last equality stands because of the independence of the Gaussian variables z_1 and z_2. Thanks to the bijection of the function φ we have

$$P((z_1, z_2) \in D^{'}) = P((cdf(z_1), cdf(z_2)) \in D) \tag{5.17}$$

and consequently,

$$P((cdf(z_1), cdf(z_2)) = area(D). \tag{5.18}$$

Chapter 6

Bridging multiple point geostatistics and truncated Gaussian fields for improved estimation of channelized reservoirs with ensemble methods

6.1 Introduction

An important key in reservoirs engineering is the development of reliable reservoir models with high predictive capacity of production behavior in existing and potentially new drilled wells. The reservoir predictability is the starting point for guiding the production plan to an optimum. In order to achieve this, the geological description plays a crucial role. The initial description of the subsurface fields, such as permeability and porosity, could be carried out using geostatistical tools in specialized modules, defining thus, the simulation model and its uncertainty. When referred to a geological description of the reservoir, one may think of a description of the spatial distribution of the petrophysical properties (permeability and/or porosity) or more natural, a description of the spatial distribution of the geological deposits formations (facies). Prior information is related to the types of the facies that are collected in the initial stage of subsurface description. Based on this prior information, a reliable geological simulation model is chosen in order

This chapter is full reference of an article submitted to Computational Geosciences

to generate realistic reservoir instances. The geological simulation model can be defined either using object based simulation technique (Deutsch and Journel, 1998; Deutsch and Wang, 1996) or grid based simulation techniques such as the multi-point geostatistics (MPS)(Guardiano and Srivastava, 1993; Strebelle, 2002; Caers and Zhang, 2004) and also, the truncated pluri-Gaussian methods (TPS) (Galli et al., 1994) Irrespective of the geological simulation model used, the poor knowledge of geology leads to an initial reservoir model whose parameters are uncertain. This prior uncertainty is usually so large that the predictive capacity is limited. To improve the predictive capacity, the prior uncertainty is reduced in a process named in reservoir engineering as history matching (HM), either manual or assisted. Comprehensive description of some assisted history matching methods can be found in Oliver et al. (2008). This paper addresses the case where the subsurface is characterized by the presence of two facies types, of which one inhabits channelized geometry. This type of spatial distribution is hardly obtained using TPS simulation techniques, especially because of the TPS restriction of applicability to variogram models. More appropriate is the use of MPS simulation methods such as the single normal equation simulation method (SNESIM, Strebelle (2002)). These geostatistical tools enable generations of more complex and natural instances of facies distributions, relying on a general conceptual structural model (a training image) of which construction incorporates the prior knowledge of the subsurface geometry and topology. By combining ideas from permeability updates in an MPS setting (Jafarpour and Khodabakhshi, 2011) with a Gaussian parameterization of the marginal facies probabilities, we define a TPS method where the truncation map is estimated from samples using MPS and a training image. This allows for non-Gaussian dependence structures and we avoid pre-specification of a variogram/covariance function and truncation map. One of the most commonly used assisted HM methods is the ensemble Kalman filter (Evensen, 1994. 2003; Aanonsen et al., 2009) especially for its ease of implementation and its design to work with an ensemble of models instead of a single model. However, since the direct use of the facies types petrophysical properties (permeability and/or porosity) as model parameters does not conducts the EnKF process to geological realistic updates due to the Gaussian nature of the method, a (re)parameterization of the facies field is needed. In Sarma et al. (2008) a parameterization of the log-permeability fields (defined by the facies field), with kernel principal component analysis (PCA) methodology, carried out in a high-dimensional space was proposed. The link between the input space (the space where the K-L expansion were performed) and the feature space is realized through a nonlinear function where an inner product is defined by a kernel function. The authors used a polynomial kernel of order two or three. In Sarma et al. (2009) the EnKF is applied in the high-dimensional feature space, in which the log-permeability fields were carried with mapping functions, whose kernels are of polynomial type. For the updates, a back transform defined via a pre-image problem has to be solved. The results

showed improvement in preserving of the high order statistical moments of the updates; this could be seen in the random field continuity. However, because the parameterization involved the permeability fields the updates are not facies fields. Therefore, the posterior uncertainty refer to the log-permeability distribution uncertainty, not facies distribution uncertainty. Another unsolved issue with the kernel EnKF method is it's tendency for all the ensemble members to collapse onto a single model. A different approach in quantification of the permeability distribution uncertainty is presented in Jafarpour et al. (2007). The authors proposed a parameterization of the log-permeability, pressure and saturation fields using the Fourier discrete cosine transform (DCT) and coupled the reservoir model with EnKF for history matching. The results showed an improvement in the random field connectivity compared with the HM process applied directly to log-permeability fields. The facies distribution uncertainty quantification could be achieved with the parametrization of the facies fields using level set methods as in Lorentzen et al. (2009) were a level set function is used to assign a signed distance to the channel border, the positive sign being reserved for the grid cells inside of the channel, whereas, the other grid cells have negative distance values. Then, the EnKF is used for updating of the distance fields, generated after comparing it with the zero field. In the context of the MPS simulation models, Hu et al. (2013) proposed the parametrization of a uniform random field used for the simulation of a certain facies types at grid cells, based on the conditional probability distributions extracted from the training image. The Gaussian transformation of such uniform field is further updated with the EnKF, the results showed an improved uncertainty quantification of the facies distribution. Other facies parameterizations which involve the domain transformation have been proposed, such as, the discrete wavelet transform (Jafarpour, 2011) or the gradual deformation (Roggero and Hu, 1998). The choice of the MPS as geological simulation model is based on the information collected at the initial stage of reservoir description. However, if important data are revealed, such as core information (hard data) or information from seismic surveillance (soft data), this information can easily be incorporated in the MPS model using probabilistic models (Journel, 2002; Krishnan et al., 2005). The challenge, however, is how to incorporate dynamical production data. In Caers and Hoffman (2006) a model (the probability perturbation method, PPM) that links the dynamical (flow) measurements with the MPS geological simulation model is developed. The sampling from the prior model (represented by the training image) rely on a the "tau" probabilistic model (Journel, 2002), by perturbing a so called pre-posterior probability (the probability of occurrence of a facies type given the "dynamical" data). The pre-posterior perturbation is carried out with a convex relation that depends on a sub-unitary parameter estimation which involves an optimization procedure of the square misfit between the measurements and the response. The process is iterative because multiple random seeds should be used in order for the simulations to match the data. An ensemble based method that links

the dynamical data with the MPS simulation model was develop in Jafarpour and Khodabakhshi (2011). The authors propose a probability conditioning method (PCM) for facies simulation in the context of EnKF as HM method. The initial ensemble of facies realization are drawn from the training image conditioned on a uniform probability map and equal weights. At each assimilation point, the log-permeability fields are updated using the available measurements, with the EnKF, and, from the mean log-permeability field, updated probability maps for each facies type are inferred. Piece-wise linear functions have been used to project the mean log-permeability field to the probability field. Then, another initial ensemble is drawn from the prior model, using the tau model for conditioning the training image on the updated facies probability maps. This procedure will continue until the end of the assimilation period. The final ensemble has a realistic geometry structure (being drawn from training image) with a reduced data mismatch. However, the piece-wise linear assumption between the log-permeability and facies occurrence probability has an empirical component and the variability and the randomness introduced by the MPS model does not ensures a good data match in the updates. At the same time, the model is highly dependent on the permeability estimation from EnKF, so if there is not enough information in the data to detect facies from the updated log-permeability field , the procedure may fail. In other words it relies on the EnKF's ability to detect features in the continuous permeability field, which may not be the case if measurements in the reservoir domain is sparse, as is typical for real fields. This is in contrast to most of the proposed methodology discussed here, were the number of wells is large compared to the reservoir size. We present a new parametrization of the facies fields, in the context of MPS geological simulation model, which is further coupled with the iterative adaptive Gaussian mixture (Stordal and Lorentzen, 2014) for uncertainty quantification of the channelized reservoirs. The parameterization bridges ideas from MPS and TPS. The standard TPS parameterization ensures that the updated fields are facies fields, however, since the field is jointly Gaussian it is not trivial to characterize channel type facies structures via a covariance function (or variogram) in combination with a user defined truncation map. The MPS approach has the advantage of simulating channel type random fields, however, they are discrete-valued and it is not trivial to connect this with ensemble based inversion. We therefor suggest a combination of the two parameterizations. By estimating marginal facies probabilities from a sample from MPS in combination with a training image, we parameterize the marginal probability field with a (marginally) Gaussian random field where the truncation map is defined such that the integral for a facies type in a given grid cell in the Gaussian domain coincides with marginal grid cell facies probability estimated from the ensemble. This automatically gives us a truncation map in the Gaussian space that honors the channel structures since the dependence structure is inherited from the training image. Thus we can update the continuous-valued random field using any ensemble based history matching technique

and change the facies type in each grid cell according to the truncation map. This ensures us that the updated ensemble members are always facies fields. However, since all ensemble based methods are only approximating the Bayesian solution, the updated fields are not from the correct posterior distribution for these nonlinear problems. Typically, after an update, the non Gaussian prior is violated in the sense that the facies fields have discontinuous channels. We address this problem in two ways, first instead of using the EnKF we apply the recently published iterative adaptive Gaussian mixture filter. For non Gaussian models, this method is asymptotically better than any EnKF approach since the bias can be controlled. Secondly, it has a resampling step between iterations. The resampling allows us to combine this resampling step with the resampling using the updated probability fields estimated from the updated ensemble and the MPS as in Jafarpour and Khodabakhshi (2011), to better preserve channels. In addition this allows us to resample ensemble members using MPS after all data is assimilated in contrast to the EnKF approach where resampling is performed after each assimilation step. After resampling, a new iteration with AGM (Stordal et al., 2011) is then performed using the resampled ensemble as initial ensemble. After two or three iterations, the updated ensemble of facies maps remains geologically realistic, even without resampling, hence after the final iteration, we are left with ensemble members that have continuous channel structures and match the data. Placed in the context of the already existing parameterization we might also consider it as being a level set type, but the thresholds used for facies delimitation have a spatial distribution. This spatial distribution is given by the initial ensemble of realizations drawn from the training image. In Section 6.2 we introduced and described in details the new parameterization of the facies field. In Section 6.3 the history matching algorithm used (IAGM) is summarized and its advantages over more traditional techniques are presented. This is followed by simulation studies in Section 6.4 where we performed two experiments with synthetic 2D reservoirs of varying size. The first is a channelized reservoir model with 50*50 grid cells and the second has 100*100 grid cells. For the first experiment, a comparison with EnKF was performed in order to emphasize the benefits of the iterative method. Secondly, IAGM is also applied to update the permeability directly and then compute a probability map (Jafarpour and Khodabakhshi, 2011). This is to show that improvements are due to the combination of IAGM and the new parameterization and not IAGM alone. The second experiment with the reservoir model having 10000 grid cells, only the IAGM with the new parameterization is shown as the conclusion is the same as for the experiment.

6.2 Parameterization and truncation in the Gaussian space

A parametrization of the elements of a given set $\mathcal{Y}$ could be defined as a surjective function $\phi : \mathcal{X} \to \mathcal{Y}$, where $\mathcal{X}$ is a metric space. In our case $\mathcal{Y}$ is represented by the space of facies maps. Usually, if n is the dimension of the reservoir domain, $\mathcal{X} \subseteq \mathbb{R}^{n\times k}$, where k is a positive integer number. When a HM method is applied, the values in the $\mathcal{X}$-space are changed and, using the function ϕ, a new facies distribution is provided. The reason for using ϕ is to ensure that the updated ensemble members are facies realizations. When the facies fields are directly populated with petrophysical properties (permeability, porosity, etc.) the updated values does not easily back transform to facies realizations when ensemble methods such as EnKF or AGM are applied. Therefore a parametrization of the facies field is needed. One possibility is to truncate the permeability fields, however, this would requires a user specified truncation map. Further, when working directly with permeability fields, the updated fields are no longer facies fields hence there is an intrinsic need for truncation after the update which may violate the quality of the data match. Truncated (pluri) Gaussian fields also requires a user defined truncation map, however, given this map, the updated fields are guaranteed to be facies fields. The TPS method however is not capable of handling channel structures.

Here we introduce a truncated marginal Gaussian parameterization for channelized reservoirs with two facies types, although extension to more facies types is almost straight forward. The only problem with multiple facies is that the truncation map, estimated from the ensemble, is not unique. However, the facies proportions are directly estimated from an initial sample and the dependence structure is inherited from the training image and do not requires specification. In other words, the new parameterization consists of continuous valued random fields that are marginally Gaussian and where the joint distribution depends on the MPS tool used to generate initial facies models. The geological simulation model used is in this work is SNESIM, where a training image is used to generate initial ensembles of facies fields. In the following we present a parameterization of the facies field in a Gaussian space and the transformation function (truncation rule) ϕ that transforms elements in the Gaussian space to elements in the space of facies fields.

Let us consider an initial ensemble of facies fields (channelized fields) generated using the prior knowledge of the reservoir subsurface characterization. We denote the number of the ensemble members by n_e. From the ensemble, we estimate, in each grid cell, the marginal probability of the channel type facies. For grid cell i the estimate is given by the proportion of the ensemble that has channel in gird cell i. Then, if p_i is the estimated probability of the channel occurrence, the distribution of the random variable,

named "facies type", can be written as

$$faciestype_i \sim p_i \cdot \delta(channel) + (1 - p_i) \cdot \delta(nonchannel) \tag{6.1}$$

(where δ is the Dirac function). For this random variable we calculate the cumulative distribution function and define the threshold in the Gaussian space as $\alpha_i \in \mathbb{R}$, as $\alpha_i = \Phi^{-1}(p_i)$, where $\Phi : \mathbb{R} \to (0, 1)$, is the standard Gaussian cumulative distribution function, $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}}\, du$.

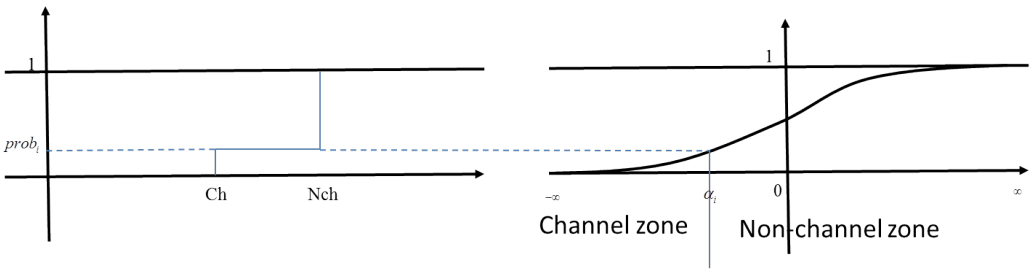


Figure 6.1: Gaussian representation of the discrete random variable "facies type"

This ensures that the marginal facies probabilities are the same in the Gaussian space as in the prior defined via the training image and MPS. That is, for a random variable $X \sim N(x|0, 1)$ we have the properties $P(X \leq \alpha_i) = p_i$ and $P(X > \alpha_i) = 1 - p_i$. This means that the real axis is split in two regions (intervals), $(-\infty, \alpha_i]$ and (α_i, ∞), corresponding to channel facies type and respectively to non-channel facies type (Figure 6.1).

The parametrization is then performed in the metric space $\mathbb{R}^n$, where n is the dimension of the reservoir domain. For each ensemble member (facies map) we define the variable θ on the reservoir domain such that in each grid cell i we have:

$$\theta_i = \begin{cases} \mathbf{E}(X|X \leq \alpha_i) & if \quad i \in channel \\ \mathbf{E}(X|X > \alpha_i) & if \quad i \in nonchannel \end{cases} \tag{6.2}$$

Given that grid cell i is inside the channel, we assign the conditional mean value of the standard Gaussian variable in grid cell i and vice versa. Alternatively, we could assign a value drawn from the conditional Gaussian distribution, however this is not straight forward since we do not explicitly know the dependence structure between grid cells. For extensions to three facies types we can split the plane in $\mathbb{R}^2$ into regions such that the area of each region corresponds to the marginal probability of each facies type. However, this truncation will not be unique.

Consequently, the function ϕ may be defined as
$\phi : R^n \to \{K_1, K_0\}^n$, where,

$$\phi(\theta) = \begin{cases} K_1 & if \quad \theta_i \leq \alpha_i \\ K_0 & if \quad \theta_i > \alpha_i \end{cases} \tag{6.3}$$

for any grid cell i. In this definition, K_1 and K_0 are represent the permeability value for each facies type in *FM*, the set of facies maps defined on the reservoir domain where the subindex 1 represents the channel and 0 the non-channel facies type. The function ϕ can, in the same manner, easily be extended to other petrophysical variables such as e.g. porosity. By construction we see that $\mathbf{E}(\theta) \equiv 0$ since, in grid cell i, we have

$$\begin{aligned} \mathbf{E}(\theta_i) &= p_i \mathbf{E}(X|X \leq \alpha_i) + (1-p_i)\mathbf{E}(X|X > \alpha_i) \\ &= p_i(\Phi(\alpha_i))^{-1} \int_{-\infty}^{\alpha_i} x \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right) dx \\ &= +(1-p_i)(1-\Phi(\alpha_i))^{-1} \int_{\alpha_i}^{-\infty} x \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right) dx \\ &= \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right) dx = 0, \end{aligned}$$

where we have used that $\Phi(\alpha_i) = p_i$.

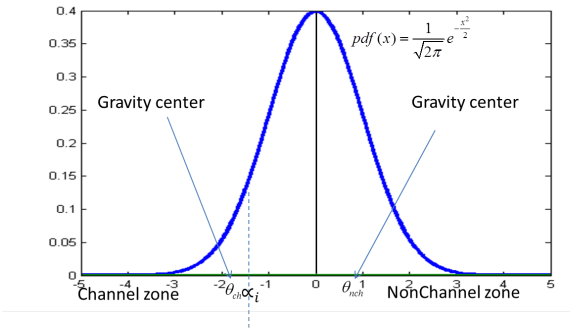


Figure 6.2: The pdf of the Gaussian variables

As a consequence the marginal prior mean of the parameter field with mean 0 in each grid cell (Figure 6.3, last picture). These conditional means have also a physical interpretation, as the gravity centers of the associated intervals of each facies type with respect to the density of the standard Gaussian variable (Figure 6.2). For an initial ensemble conditioned to hard data (facies observations at the well locations) for a given number of grid cells, these grid cells has θ values 0 with threshold values $\alpha = -\infty$ or $\alpha = \infty$ depending on the type of the observation (nonchannel and respectively channel). If, in the probability map, we have value 0.5 for the channel occurrence then $\alpha = 0$ and $\theta_{ch} = -\theta_{nch}$ (θ_{ch} and θ_{nch} are the gravity centers of the interval assigned to the channel

facies type and non-channel facies type respectively). Hence, this parameterization is a normalization of the discrete variable "facies type" and could be seen as a level set parametrization, but with the thresholds having a spatial distribution.

In Figure 6.3 are presented a probability occurrence map for the channels (left), cal-

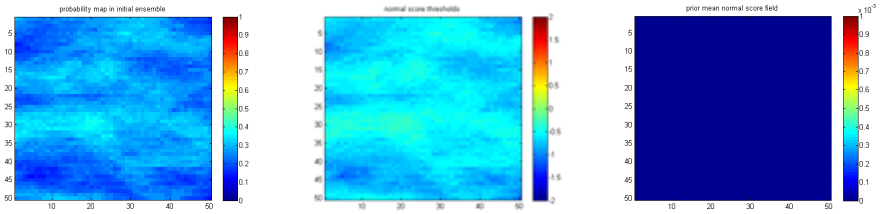


Figure 6.3: The probability map of the channel occurrence (left) and the thresholds map (middle) and the prior mean of the parameter field (right)

culated from an ensemble, and the associated thresholds field (middle) as consequence of the above methodology. Concerning the spatial distribution of the parameter field,

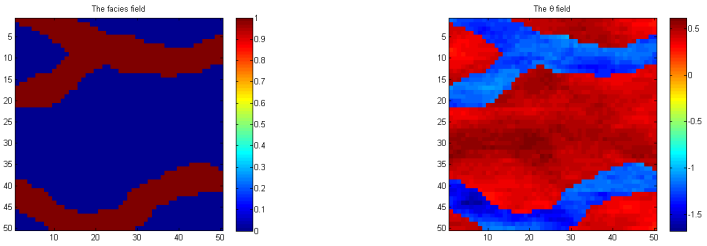


Figure 6.4: An arbitrary facies field (left) and the associated parameter field (right)

in Figure 6.4 we present an example of a channelized reservoir, which is part of an ensemble of facies fields (left) and the associated spatial distribution of the parameter field (right), calculated using the presented method.

6.3 Iterative Adaptive Gaussian Mixture Filter

An ideal way to estimate channels and in general facies in a reservoir model would be to apply an importance sampling approach. This method can work directly on the discrete facies field using samples from the training image and the likelihood function. One might also apply an MCMC approach but then the prior distribution has to be evaluated if it is not symmetric. Since we are using a training image, the prior has to be estimated from the samples using e.g. a transiogram which is cumbersome, suffers from

randomness and hence requires a lot of samples to achieve high precision. The importance sampling, although ideal for this discrete-continuous type of problems, cannot be applied to reservoirs as the number of parameters to estimate is simply too large. The number of samples required would simply be too large for practical purposes. From the authors experience 10.000 samples is not sufficient even for moderate sized reservoirs.

Since we have parameterized the field of discrete random variables with a field of continuous random variables (that is a random variable that has a density function and not a mass function) we can apply filters defined for continuous random variables. Several approaches to facies estimation that use the EnKF together with a parametrization and/or transformation of variables are available in the literature as discussed in the introduction. However, the nonlinear nature of this problem which causes severe bias with the EnKF, in addition to the fact that EnKF under estimates uncertainty in facies models we argue that the IAGM is a better suited method for this kind of problem. Alternatively, other iterative methods such as ES-MDA or EnRML might also be suitable. However we choose the IAGM since it has superior asymptotic properties, hence the results will continue to improve if we increase the sample size which is in contrast to e.g. EnKF which has a fixed asymptotic bias.

The IAGM was introduced as an iterative version of the adaptive Gaussian mixture filter (Stordal et al., 2011) to further improve the AGM on nonlinear models. The implementation of IAGM is very simple especially if EnKF or AGM is already implemented. We denote the ensemble of parameter fields by $\{\theta_i\}_{i=1}^N$. Further we denote by $\mathcal{G}_t = \mathcal{H}_t \circ \phi$, the composition of transformation of parameters to facies fields and the measurement operator. We augment our state vector with these variables. Hence

$$Z_t^i = [\theta_i^T \quad \mathcal{G}_t(\theta_i)^T]^T,\ i = 1, \ldots, N, \tag{6.4}$$

such that we can construct, for each assimilation time t, a binary matrix H_t with the relation $Y_t = H_t Z + \epsilon$. That is the measurement is a linear function of our augmented state vector with additive Gaussian white noise. As in the EnKF we denote by C_t the sample covariance matrix of $\{Z_t^i\}_{i=1}^N$. At each assimilation time, the augmented state vector is updated in the AGM (and IAGM) for each $i = 1, \ldots, N$ as

$$\hat{Z}_t^i = Z_t^i + C_t H_t^T (H_t C_t H_t^T + h^{-2} R)^{-1} \left(y_t - H_t Z_t^i + \epsilon_t^i\right), \tag{6.5}$$

where ϵ_t^i is a sample from the Gaussian measurement error distribution $N(0, R)$. Note that the only difference with a standard EnKF update is the scaling h^{-2} of the measurement error covariance matrix R. In other words the linear update is dampened where the dampening factor $h \in [0, 1]$ is the bandwidth of the Gaussian mixture (see Stordal et al. (2011)). In addition to a reduced linear update, importance weights are derived from the

Gaussian mixture and updated sequentially as

$$W_t^i \propto W_{t-1}^i N(y_t - H_t Z_t^i, h^{-2} HCH^T + R). \tag{6.6}$$

To avoid filter degeneracy that occurs in high dimension and complex systems a weight interpolation is introduced

$$W\widehat{W}_t^i = \alpha_t W_t^i + (1 - \alpha_t) N^{-1}, \quad \alpha_t \in [0, 1], \tag{6.7}$$

where

$$\alpha_t = N^{-1} (\sum_{i=1}^{N} (W_t^i)^2)^{-1}. \tag{6.8}$$

For details of the AGM we refer to Stordal et al. (2011).

A resampling and reweighting step in the algorithm before rerunning AGM is discussed in Stordal and Lorentzen (2014). Here we propose to resample from the training image with a new probability field obtained from the weighted mean of the ensemble members after the previous iteration. That is, after one assimilation cycle, a new probability map is constructed on a cell by cell basis as

$$p_i = \sum_{j=1}^{n_e} w_j 1(cell_i \in channel)$$

and used in SNESIM . This is equivalent to resampling marginal facies variables from the empirical distribution obtained with AGM (resampling) with the constraint that the dependence structure is given by the training image. To be completely rigorous we should then evaluate the ratio between the old and new prior for each ensemble member, however this would requires the estimation of the prior in addition to evaluation of weights in high dimensions. With the weight reduction leading to almost uniform weights due to the dimension, there would be more work than gain, and we instead make the assumption of uniform weights after resampling from the training image with the new probability field. The resampling and reweighting is a topic for future research. In most applications a few iterations are required, typically two or three. This is in agreement with the results we present below.

6.3.1 The State Vector

Concerning the HM methods implementation, the state vector for the i^{th} ensemble member at the t^{th} assimilation step will be:

$$\mathbf{X}_t^i = \left[\text{param}_i^T \mathcal{G}_t(\text{param}_i)^T\right]^T, \quad i = 1, \ldots, N \tag{6.9}$$

where $\mathcal{G}_t(\text{param}_i)$ are the simulated observations represented by the simulated produc-

tion data (oil and water rates, bottom hole pressures). The field param is the generic notation either for the permeability field or the parameter field θ presented above. The dynamical variables (the pressure and the saturation) are not updated together with the parameters because, after each assimilation time, we have chosen to rerun the simulator from time zero, even though this technique will increase the computational cost. For the small models of the examples presented, this procedure is not a time consuming, but, for large scale models, the computational cost should be considered. However it is possible to run the IAGM as a smoother, that is update all data at ones (Stordal et al. 2012).

6.3.2 The Resampling

After an iteration of IAGM we are able to calculate, from the weighted ensemble of facies fields, a probability map of the channel facies type occurrence. In order to reconstruct the channels geometry of the facies fields, we perform a resampling from the training image using the same SNESIM algorithm. The SNESIM has incorporated the tau model (Journel 2002 [15], Krishnan et al. 2005 [11]) which enables integration of the probabilities coming from soft data and training image. Usually, the probability maps used as soft data are coming from seismic interpretations, but here, we have used as soft data the probability maps calculated after an iteration of IAGM. If we consider, at a particular location i, $A(i)$ being event which consist of the occurrence of a certain facies type and $P(A(i)|B(i))$ and $P(A(i)|C(i))$ the conditional probabilities coming from the training image and weighted probability map respectively, then, the conditional probability $P(A(i) \mid B(i), C(i))$ is calculated as:

$$P(A(i) \mid B(i), C(i)) = \frac{1}{1 + x(i)}, \tag{6.10}$$

where "the distance" $x(i)$ is calculated as

$$\frac{x(i)}{x_0(i)} = \left(\frac{x_1(i)}{x_0(i)}\right)^{\tau_1} \left(\frac{x_2(i)}{x_0(i)}\right)^{\tau_2}, \text{where } \tau_1, \tau_2 \in \mathbf{R}, \tag{6.11}$$

and the "distances" $x_0(i)$, $x_1(i)$, $x_2(i)$ are calculated as follows

$$x_0(i) = \frac{1 - P(A(i))}{P(A(i))}, x_1(i) = \frac{1 - P(A(i)|B(i))}{P(A(i)|B(i))}, x_2(i) = \frac{1 - P(A(i)|C(i))}{P(A(i)|C(i))} \tag{6.12}$$

P(A(i)) is the probability with which the facies type occurs at the location i and is the target marginal probability of that facies type. In this study the weights τ_1 and τ_2 used are both equal to 1, which correspond to the equal importance assigned to the training image and weighted probability map respectively.

6.4 Applications

In this section we present two applications of the presented methodology; one which involves a small reservoir with 50*50 grid cells, having 3 injectors and 3 producers, and the other a reservoir with 100*100 grid cells with 4 injectors and 9 producers. The reference fields and the initial ensembles used are sampled from the same training image (Strebelle 2002) with 250*250 grid cells (Figure 6.5) using the SNESIM algorithm (Strebelle 2002) implemented in SGeMS.

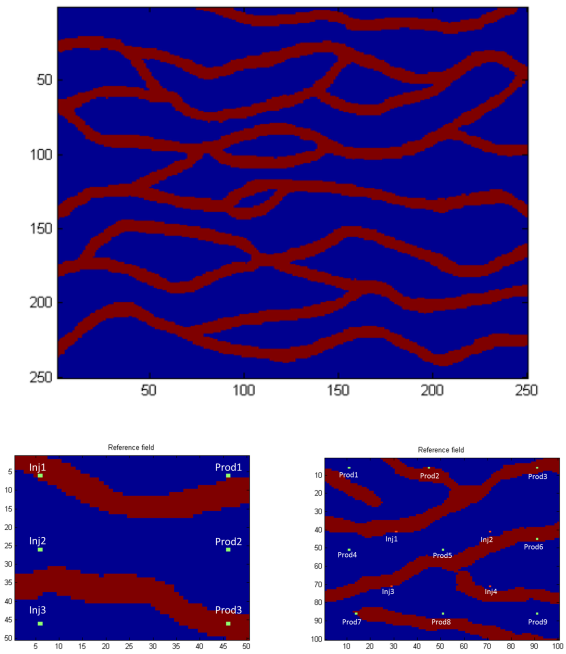


Figure 6.5: The Training Image (top) and the reference fields used for the experiments (bottom)

Throughout this study the red shades represents the channel facies type (typically sand, highly permeable) and the blue shade represents the non-channel facies type (typically shale with low permeability). In all the simulations, the petrophysical properties of each facies type are considered known and kept constant, not taking into account the heterogeneity within each facies. Hence, the uncertainty in the model is generated by the channels distribution in the reservoir domain.
The initial ensembles of facies fields were generated without conditioning the training image to facies observations (hard data). Only for the resampling, in the IAGM steps,

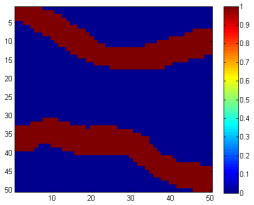
the initial ensembles were generated conditioning the training image to the previous estimated probability map. The number of iterations in the experiments was established taking into account the properties of the updated ensemble. This should fulfill two requirements; good data match and predictions, and the updated ensemble members should not be in conflict with our prior believes, that is, the ensemble should have continuous channel structures. With that in mind, we terminate the assimilation process when the ensemble members are matching the data satisfactory at the same time as they have continuous channels. The latter can be achieved by resampling from the training image using a probability map projected from the mean permeability field (Jafarpour and Khodabakhshi, 2011), however the quality of the data match is then decreased. Since our parameterization is guarantees that our updated ensemble members are facies fields, we do not need to perform resampling after the final iteration. This gives our method an advantage over previous work.

6.4.1 Channelized reservoir with 2500 grid cells

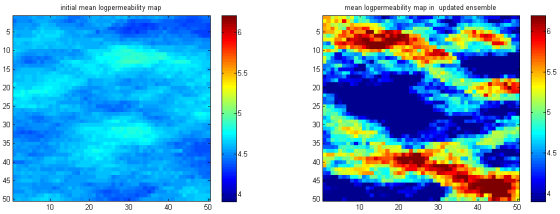
In this example we present the case of a small channelized 2D reservoir, with 50 grid cells in each direction of which dimension was set at $30 \times 30 \times 20$ ft. We design the reservoir with a 6-spot water flooding black oil model, having three injection wells at the left side and three production wells at the right side. The reservoir is initially filled with oil at a constant uniform saturation of 0.8 (which generates a connate water saturation of 0.2) and with a uniform pressure of 5000 psi in every grid cell. The producers works under constant bottom hole pressure (BHP) value of 1500 psi and the injectors operate at 2500 stock tank barrels per day (STB/D) constrained by a maximum 100000 psi for BHP. The measurement errors of the production data are assumed Gaussian with 0 mean and standard deviations of 30 STB/D for water rates (WR) and oil rates (OR) at the producers, and 80 psi for BHP at the injectors, respectively. These values are used when generating the reference measurements, and also when adding noise to the production data in the analyzed step of the HM processes. Water injection starts from the first day and continues for a period of 191 days. We assimilate data starting from the first day until day 191, with a 20-day interval, resulting in 10 assimilation steps. For each experiment the measurements were obtained through forward simulation of the synthetic model presented as the "reference" in which the channels geometry is defined by two non-intersected narrow bands which propagate longitudinally (Figure 6.5, bottom left). The permeability values were set at 500 mD and 50 mD for the channel facies type and for the non-channel facies type, respectively, while the porosity of both facies types is considered 0.2.
In this subsection we perform the following experiments:

- EnKF with log-permeability estimation.
- EnKF with the parameter field.
- IAGM with h=0.25 two iterations and with the parameter field.

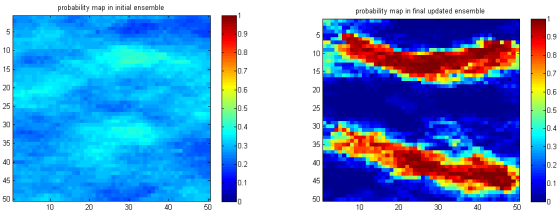
In the top of Figure 6.6 the true position of the channels in the field is presented and in the following pictures we present the prior mean and posterior mean of all the results of the experiments. In Figure 6.6(b) the prior mean and posterior mean refer to the mean log-permeability, whereas for all the rest we refer to the mean probability map of the channels occurrence.



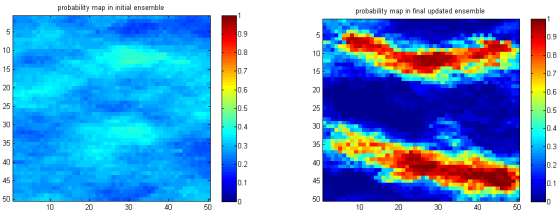
(a) The reference



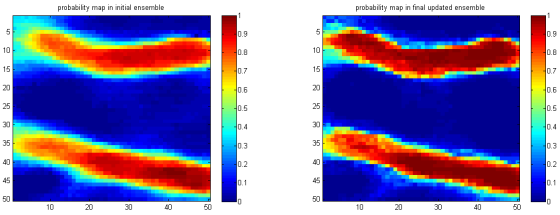
(b) EnKF with log-permeability



(c) EnKF with parametrization



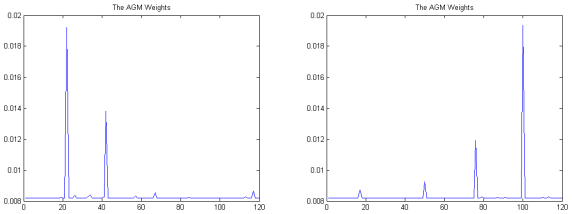
(d) IAGM h=0.25 with parametrization iteration 1



(e) IAGM h=0.25 with parametrization iteration 2

Figure 6.6: The probability maps of the channel occurrence in all the experiments

In Figure 6.7 are the weights of the ensemble members after both iterations.

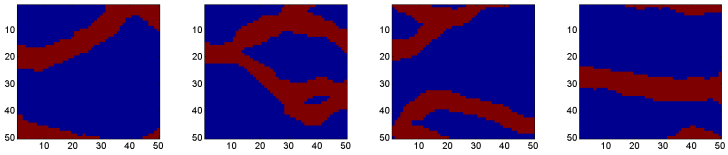


(a) Iteration 1 (b) Iteration 2

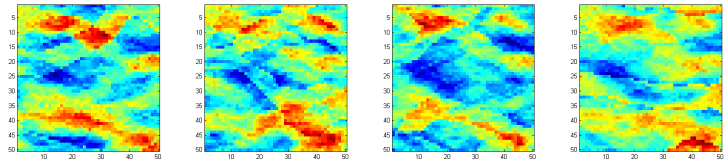
Figure 6.7: The AGM weights of the iterations

The initial ensemble used in the second iteration of the IAGM (Figure 6.6(e) left) is calculated with SNESIM, conditioning the training image to the facies probabilities map calculated with the weights presented in Figure 6.7(a). In Figure 6.8(a) we present the first four ensemble members in the initial ensemble of the experiments and in Figure 6.8(b, c ,d) the same members in the updated ensembles.

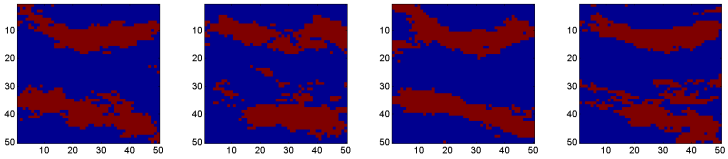
In Figure 6.8(e) are the first four members in the initial ensemble obtained after the resampling and on the position Figure 6.8(f) are the members in the updated ensemble after the second iteration is performed. When EnKF is performed with log-permeability in the state vector, not only ist structure not preserved (Figure 6.8(b)), but the estimate of the channel positions are also poor (Figure 6.6(a)). The main reason is due to the fact that EnKF has a tendency to increase the log-permeability values around the wells drilled in the channels and decrease the log-permeability values in the wells drilled outside of the channels, especially when a small numbers of wells are used and the position of the wells in the field is are not very informative (lack of connectivity). When, EnKF and the first iteration of IAGM is applied with the parametrization proposed, the position of the channels is better estimated (Figure 6.6(c) and (d)), but the structure in the updates is unfortunately not preserved (Figure 6.8(c) and (e)). This is a direct result of the updated limitation of the HM methods to the first two statistical moments. When the resampling is carried out using the training image and the updated probability map, the channel structure is regained, but the price of that is the loss of the data match (Figure 6.9(c)). To obtain data match and preserve channel structure in posterior ensemble, we performed an iteration with this new initial ensemble using the small update offered by AGM for a small tuning parameter (h=0.25). The result is an updated ensemble with a good channel structure (Figure 6.8(f)), with a good estimation of the position of the channels (Figure 6.6(e)) and with a good data match and prediction (Figure 6.9 (d)).



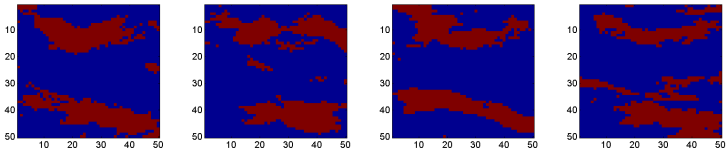
(a) Initial ensemble



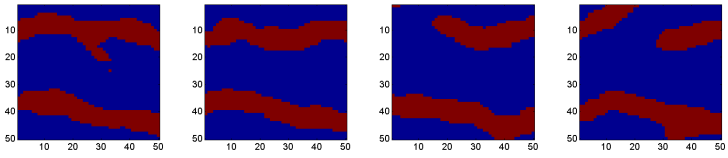
(b) Updated ensemble EnKF with log-permeability



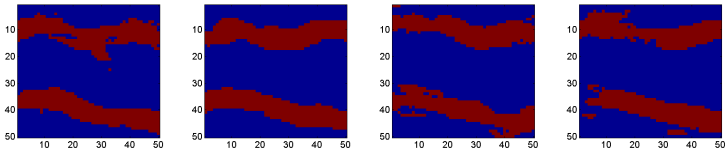
(c) Updated ensemble EnKF with parametrization



(d) Updated ensemble IAGM with parametrization iteration 1

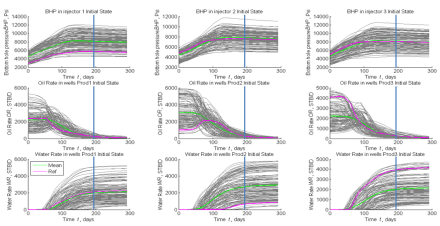


(e) Initial ensemble IAGM with parametrization iteration 2

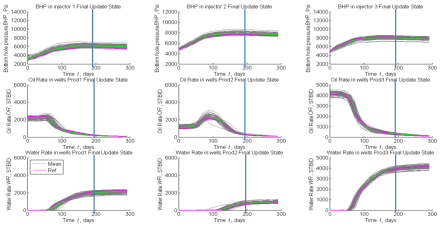


(f) Updated ensemble IAGM with parametrization iteration 2

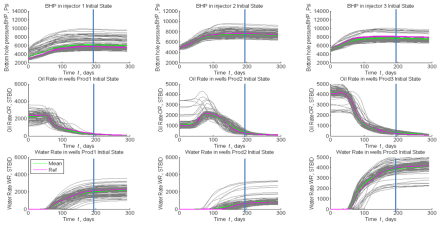
Figure 6.8: The first four ensemble members in all the experiments



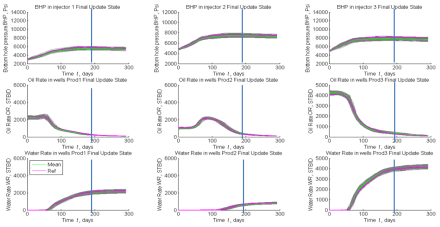
(a) Pressures and rates in initial ensemble iteration 1



(b) Pressures and rates in updated ensemble iteration 1



(c) Pressures and rates in initial ensemble iteration 2



(d) Pressures and rates in updated ensemble iteration 2

Figure 6.9: The production data in the AGM experiments

In Figure 6.9 we present the prediction for the next 100 days for pressures and production rates for all the wells in initial and updated ensemble of the both iterations of IAGM (the data assimilation period is delimited by the prediction period through a blue line). When the initial ensemble is generated, the production response has a huge uncertainty (Figure 6.9(a)) which decrease together with the production data assimilation at the first iteration of IAGM (Figure 6.9(b)). The predictions are satisfactory; however, since the channel structure is lost we perform the resampling. This introduces additional variability in the ensemble (Figure 6.9(c)), and after the second iteration is performed this is again reduced (Figure 6.9(d)) but preserving the channel structure (Figure 6.8(f)). These two goals, data match with good prediction and channel structure in updates is achieved in this case only after two iterations.

Updating log-perm with projection maps

Since the IAGM is applied to facies estimation for the first time here, we also want to apply it to the permeability directly and construct probability maps by projections in order to compare this method directly with our new parameterization. We have rerun the above experiment with the same initial ensemble. After updating the log-permeability field we use a projection function, see Fig 6.10, for each facies type with truncation at 50 and 500 mD together with a piecewise linear function to construct the probability map.

The probability map (for the channel type facies) is first truncated at 0.1 and 0.9 and then a second run is performed where the probability map goes from 0 to 1 (projection function borrowed from Jafarpour and Khodabakhshi, 2011). If $p_{\max}$ and $p_{\min}$ denote the upper and lower bound of the probability projection (for the channel type facies), then, after updating the log permeability fields we set for any grid cell $p = p_{\min}$ if the mean log permeability in the grid cell is less than 50 mD, $p = p_{\max}$ if the mean log permeability is greater than 500mD. If the mean log permeability is in between, we use linear functions so that the probability sums to one (Figure 6.10).

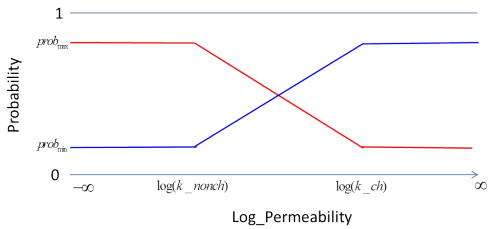
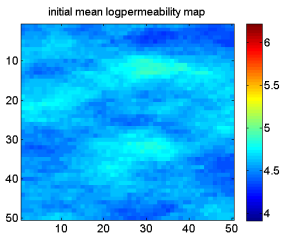
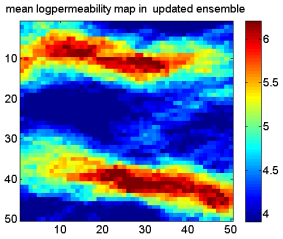


Figure 6.10: The two projection functions after Jafarpour and Khodabakhshi, 2011

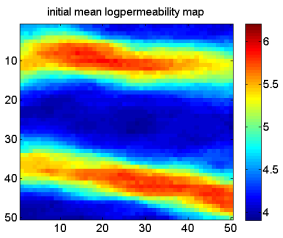
To demonstrate the dilemma of this methodology, as discussed in the introduction, we show the results after the second iteration both with and without resampling in order to enhance the geology vs data match issue that occur. In Figure 6.11-6.13 we have summarized the results of the first run in terms of the probability maps, the four first ensemble members and the production response. There are at least two important findings here. First, the method do not produce as good data match with two iterations as the above parameterization. Secondly, as stated previously, we see from Figure 6.12 and Figure 6.13 that we either obtain facies fields or a reasonable data match. We see some improvement in the data match when the probability field is not truncated (to [0.1,0.9]), however it is not of the same quality as the results of the previous section. In Figure 6.15-6.16 we show the results of the second run where the probability map is not truncated. The conclusions are the same for both experiments.



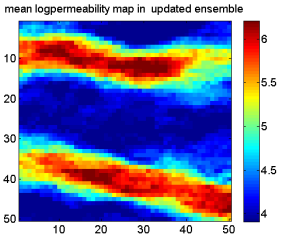
(a) Initial mean log-perm



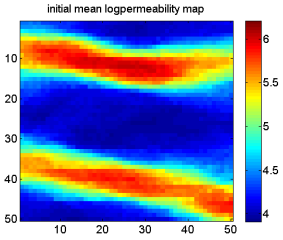
(b) Mean log-perm after first iteration



(c) Mean log-perm after first resampling

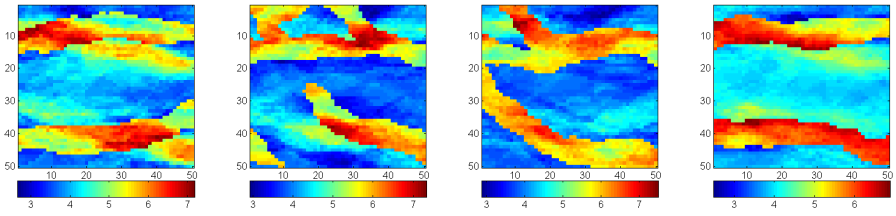


(d) Mean log-perm after second iteration

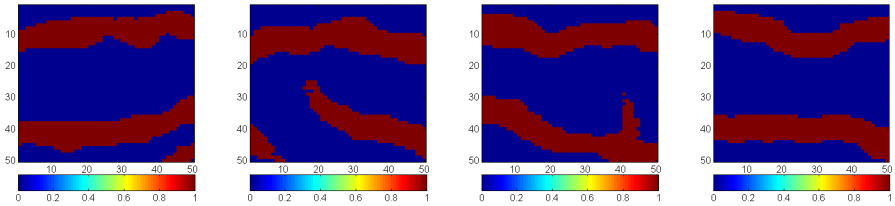


(e) Mean log-perm second resampling

Figure 6.11: Permeability fields, experiment 1

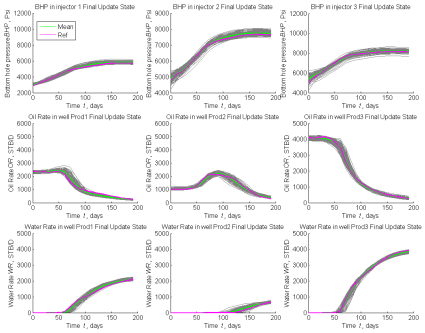


(a) After second update

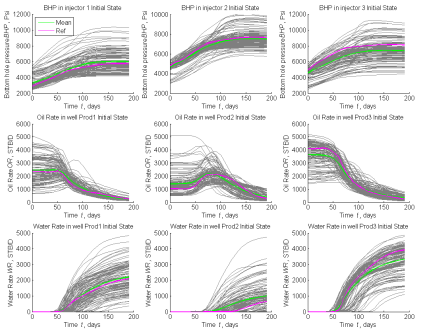


(b) After second resampling

Figure 6.12: First four ensemble members, experiment 1

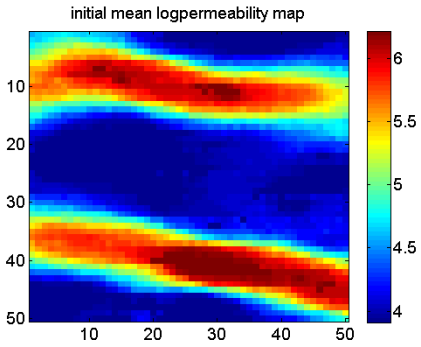


(a) Production pressures and rates after second iteration

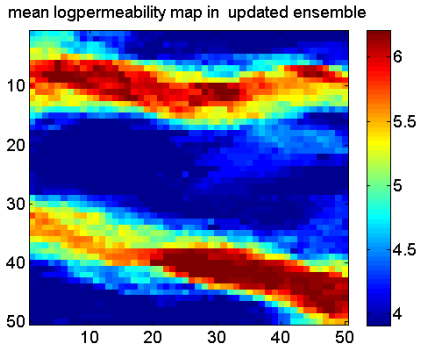


(b) Production pressures and rates after second resampling

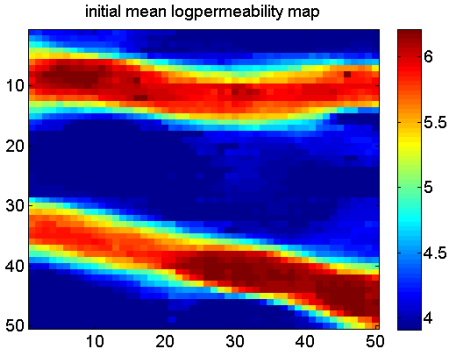
Figure 6.13: Data match after second resampling, experiment 1



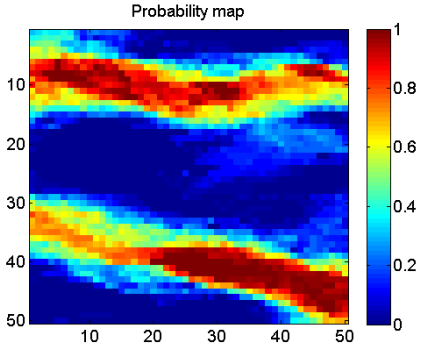
(a) Mean log-perm after first resampling



(b) Mean log-perm after second update

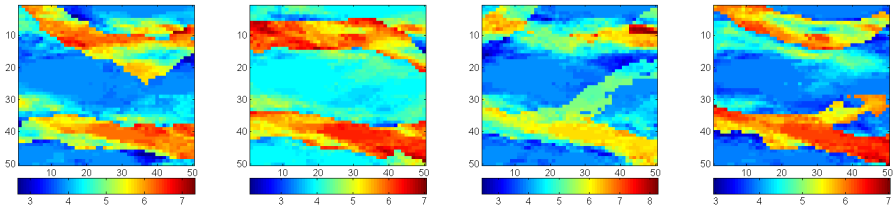


(c) Mean log-perm after second resampling

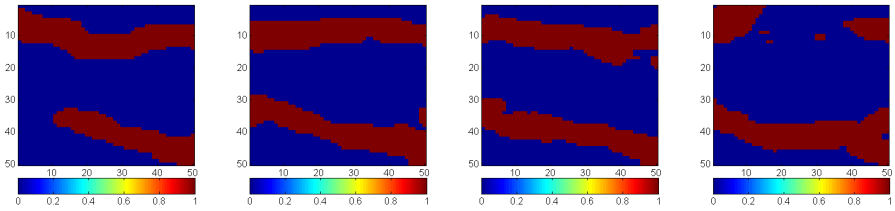


(d) Probability map after second update

Figure 6.14: Permeability fields, experiment 2

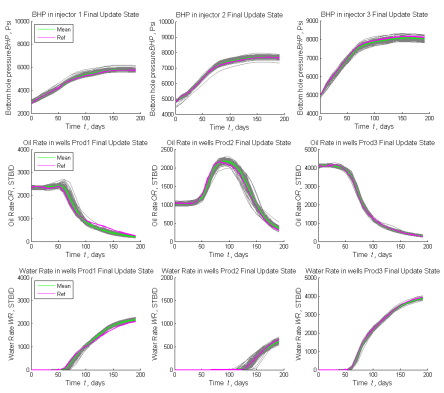


(a) After second update

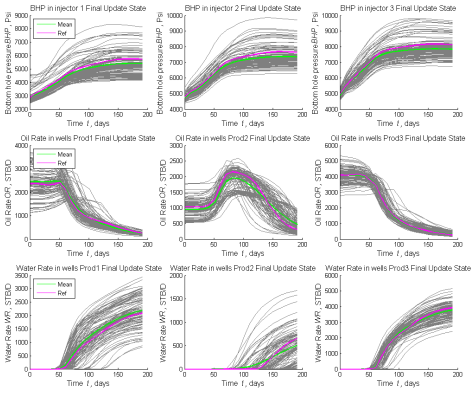


(b) After second resampling

Figure 6.15: First four ensemble members, experiment 2



(a) Production pressures and rates after second iteration



(b) Production pressures and rates after second resampling

Figure 6.16: Data match after second resampling, experiment 2

6.4.2 Channelized reservoir with 10000 grid cells

In the second example we present a channelized 2D synthetic reservoir with 100 grid cells in the X and Y direction. The dimension of each grid cell is $30 \times 30 \times 20$ ft. The reservoir is a 13-spot water flooding black oil model, having four injection wells at the center of reservoir domain and nine production wells surrounding the injection wells (Figure 6.5 bottom, last picture). The reservoir is initially filled with oil at a constant uniform saturation of 0.8 (which again generates a connate water saturation of 0.2) and with a uniform pressure of 5000 psi in every grid cell. The producers works under constant bottom hole pressure (BHP) with a value of 3000 psi and the injectors operate at 3500 STB/D constrained by a maximum BHP of 100000 psi.

We assimilate data starting from first day until day 351, with 60-day interval, having 6 assimilation steps. For each experiment the measurements were obtained through forward simulation of the synthetic model presented as the "reference" in which the channels geometry is presented in the top of Figure 6.17. The measurement errors of the production data are assumed Gaussian with 0 mean and standard deviations of 70 STB/D for water rates (WR) and oil rates (OR) at the producers, and 200 psi for BHP at the injectors. These values are used to generate noisy observations from the reference model in addition predicted observations of the production data used in the analysis step of the HM processes. Water injection starts from the first day and continue thereafter a period of 351 days of production. We assimilate data at a 60-day interval resulting in a total of 6 assimilation steps. The permeability values were set at 9 mD and 1 mD for the channel facies type and for the non-channel facies type, respectively, while the porosity of both facies types is set to 0.2 and considered as known.

We only show the results from IAGM with the proposed parameterization for this model as the findings are the same as in the previous example. The IAGM is run with $h = 0.25$ for three iterations with the proposed Gaussian parameterization and resampling between each iteration. In Figure 6.17 we present the reference field and the probability maps of channel occurrence for the initial (resampled) and updated ensemble in each iteration. Comparing the reference field with the probability map of the initial ensemble (before any data assimilation) we infer that there is very little information about the channels in the initial ensemble, that is, the variance of the initial ensemble is large. In Figure 6.17(c) we see the probability map after the first update which gives us an idea of what geological information is contained in the production data. We can already see that there are a few regions of the reservoir where our approximative Bayesian solution is certain that there is a channel type facies. After three iterations we are able to identify the two main channels however, the top left and bottom channel seems difficult to estimate. This may be due to the fact that there is not enough information in the production data to constrain the model. To be critical, the low channel probability in the bottom

where there is a channel tells us that we have either found only a local region of the posterior or that the bias from ensemble based methods is larger than expected. This motivates us to explore the full posterior using more expensive techniques in the future. Since the IAGM has a theoretical bias that decrease with n_e, it would be interesting to see how much the results improve if we increase the number of ensemble members. That is, however, beyond the scope of this study.

Figure 6.18 and Figure 6.19 tell the same story as in the previous example. Assimilation without resampling breaks up the channels in the first iteration whereas resampling reconnects them but increase the data mismatch. After three iteration (in contrast to two in the previous example) we are able to maintain the channel structure without resampling and hence preserve the good data match and prediction from the assimilation.

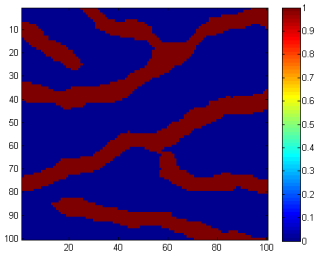
We, also, performed the fourth iteration with the only purpose to observe the ensemble behavior. The small differences between the root mean squared error (RMSE) values of the updated ensembles defined as

$$RMSE^2 = (n_d)^{-1} \sum_{i=1}^{n_e} \sum_{j=1}^{n_d} w_i (Y_j - \mathcal{G}_j(x_i))^2,$$

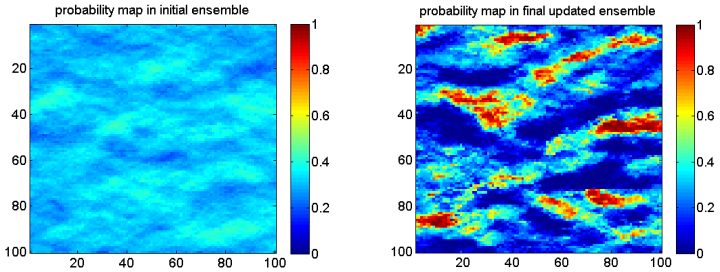
where n_d is the number of measurements, x_i denotes ensemble member number i and w_i the corresponding likelihood weight. Table 6.1 indicates that convergence is achieved.

Table 6.1: RMSE for total production rates

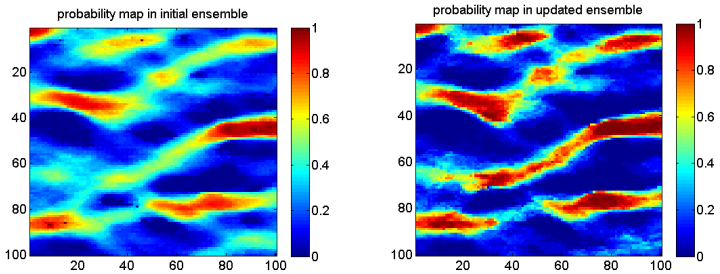
	Iteration 1	Iteration 2	Iteration 3	Iteration 4
RMSE in initial ensemble (STB/D)	827.58	183.58	150.49	147.97
RMSE in updated ensemble (STB/D)	158.75	104.05	101.67	98.94



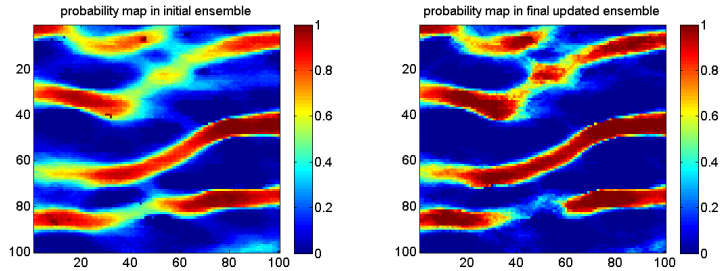
(a) The reference



(b) IAGM h=0.25 with parametrization iteration 1

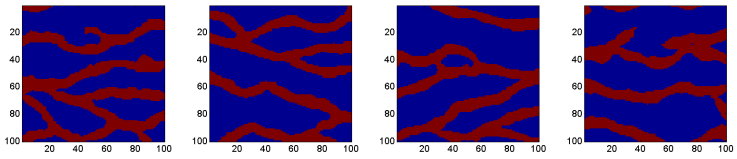


(c) IAGM h=0.25 with parametrization iteration 2

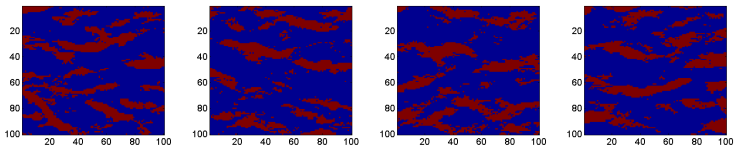


(d) IAGM h=0.25 with parametrization iteration 3

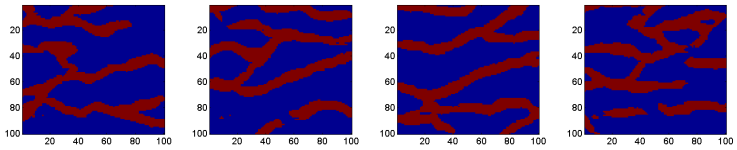
Figure 6.17: The probability maps of the channel occurrence



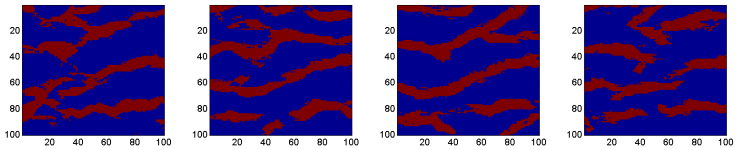
(a) Initial ensemble iteration 1



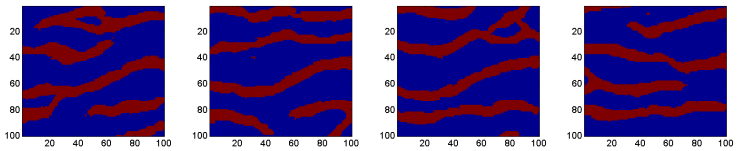
(b) Updated ensemble iteration 1



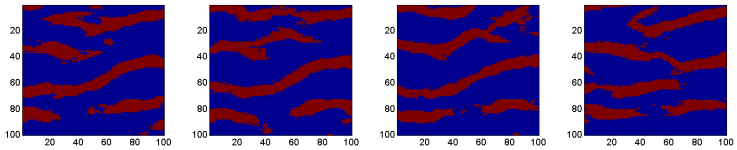
(c) Initial ensemble iteration 2



(d) Updated ensemble iteration 2

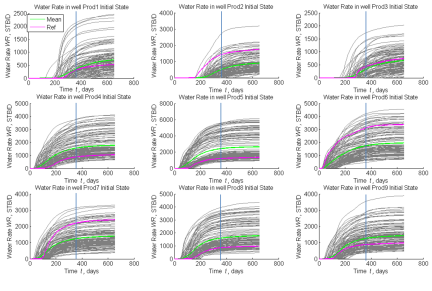


(e) Initial ensemble iteration 2

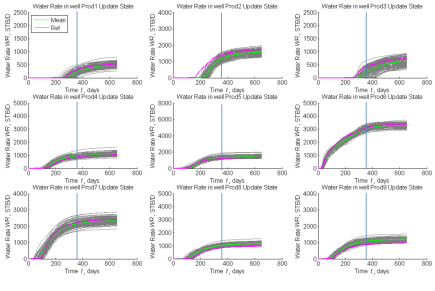


(f) Updated ensemble iteration 3

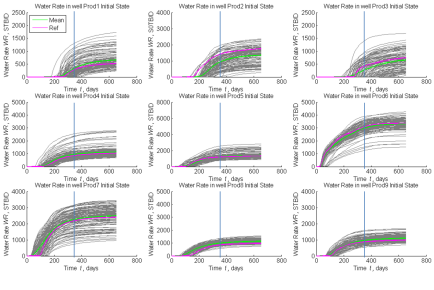
Figure 6.18: The first four ensemble members in all iterations



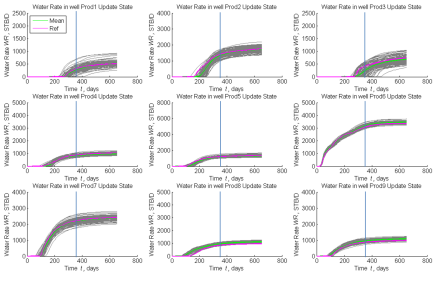
(a) Production water rates in initial ensemble iteration 1



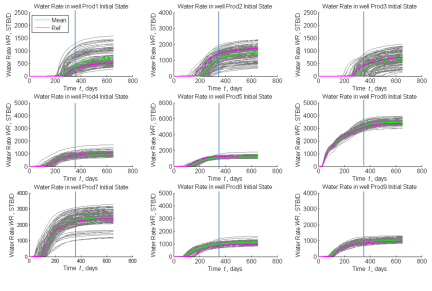
(b) Production water rates in updated ensemble iteration 1



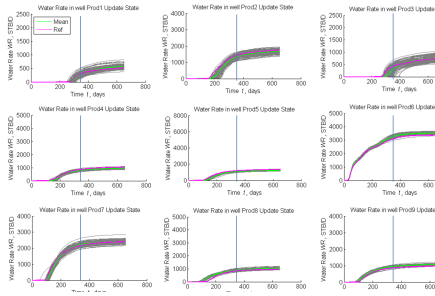
(c) Production water rates in initial ensemble iteration 2



(d) Production water rates in updated ensemble iteration 2



(e) Production water rates in initial ensemble iteration 3



(f) Production water rates in updated ensemble iteration 3

Figure 6.19: The water production rates in the AGM experiments

Chapter 7

Realistic facies simulation with the APS model for a North Sea field case

7.1 Introduction

In reservoir characterization, modern reservoir modeling and history matching aim at delivering integrated models with quantified uncertainty, constrained on all available data. The "holy grail" of model updating/conditioning in the framework of Assisted History Matching(AHM) is to obtain models which: match the past data within the model and measurements uncertainties; incorporate all types of information (different measurements types) and preserve geological realism (consistent with the initial geo-model).
The Fast-Model-Update workflow (FMU) resolves all of the above mentioned requirements by handling reservoir models in the big-loop model updating framework and it provides: a) a model chain, from depth conversion to simulation model, that is repeatable and updatable; b) an ensemble framework for representation and prediction of uncertainty: c) ensemble methods for conditioning the model chain on dynamic data. The major principle of the FMU workflow is an automation and integration of all the modeling steps such that the whole model chain from depth conversion to simulation can be run in batch.
In the implementation of any AHM process, one of the most important things to address is the reservoir geological structure. When referring to geological structure we think of

This chapter is based on an article presented at 76th EAGE Conference and Exhibition, Amsterdam June 16-19, 2014 where the coauthors are R.G. Hanea, T. Ek, J. Saetrom and D.B. Sollien

faults position and top/base surfaces of the geological layers. The solution of integrating structural uncertainties in the big loop approach was given by Seiler et al., 2010. The natural transition from structural geological uncertainties towards the rock and fluid uncertainties is represented by the facies distribution.

In this study we apply a facies modeling process that can be used in an integrated modeling workflow (FMU). Therefore, we should ensure the flexibility, efficiency and geologically realism of the process. In our opinion from a range of possible options (e.g multipoint geostastatitics, level set, object and process based methods) the "low hanging fruit" methodology that fits our requirements is the plurigaussian concept (Sebacher et al., 2013). By using this concept, we generate possible facies realizations (distributions) in the reservoir domain.

These facies instances must be consistent with the prior knowledge about the reservoir geology (the number of the facies types, the possible transition between facies types, the direction of the facies, the geometry of the facies, etc.) and honor the hard data (well log data) and the sand probability cubes (derived from seismic data). The solution applied here is the adaptive plurigaussian simulation (APS) approach. The APS model consists of generating facies distributions conditioned to prior facies probability occurrence maps. It is still part of the boarder truncated plurigaussian simulation (TPS) technique, using truncation maps, not in the multi-dimensional (usually bi-dimensional) space of GRF, but in a multi-dimensional probability space.

In a broader view we are interested in quantifying and reducing the uncertainties in the facies distributions in a multiple realizations world. Therefore, the proposed process should be updateable/conditioned based on all the available data by an ensemble based method/tool.

If until now the APS technology was applied with great success for synthetic cases ranging different depositional environments (See Sebacher et al, 2013, 2014), this is the first attempt when APS is implemented and applied for a real field case in the North Continental Shelf (NCS). Due to confidentiality issues we can not use the real name of the field, so for simplicity we will name it "Loki" field. The novelty of this work consists of realistic facies simulations for a real field case, consistent with the hard and soft data as well as with the geological concept. This simulation process is repeatable and automated/assisted by being integrated in the FMU framework of the big loop conditioning philosophy. We are proposing an integrated workflow for facies generation and estimation using a simulation model that is able to create a link between the mathematical theory and real world.

7.2 The geological setup

We apply the pluri-gaussian facies modelling approach to the specific geological settings of the Loki field, located in the North Sea, offshore mid-Norway. The field consists of multiple westerly tilted blocks, bounded by faults of Devonian age generated during the opening of the North Viking Graben (Fossen et al., 2000), and later reactivated during Triassic and essentially Jurassic period. In term of sedimentology, the Loki field consists of Triassic deposits of the typical "redbeds" fluvial sandstones of the Upper Lunde and Statfjord Formations. The Statfjord Formation is strictly speaking of Rhaetian and Early Jurassic age, but commonly attached to Triassic deposits as it consists of genetically similar sediments, deposited in the same environment, with no major unconformity in the stratigraphic sequence. During the Triassic period, the climate evolved from arid to semi-arid continental conditions (Dore, 1992; Van Der Zwan and Spaak, 1992). The sedimentary structures consisted on terminal sections of alluvial fans (Tunbridge, 1984) with ephemeral braided streams terminating in sheetfloods, muddy flood plain or ephemeral lake. Episodic water flooding and collapsing of channel banks resulted in the deposition of crevasse splays along the margins of the channel systems. As a result of the arid conditions, most of the sediments were trapped within the continental environment (Collinson, 1986), hence the high accumulation of continental sediments observed in the Triassic sediment record of the North Sea. In the Loki field, the two reservoir formations consist of sandy fluvial deposits interbedded with floodplain mudstones, for a total thickness of 1000 m. A vertical progression in the style of fluvial bodies has been early recognized and extensively studied for integration in the geological model in an attempt to match production data. Depending on the relative sea level variations, channel style varies from proximal high energy braided channels located at the foothill of alluvial fans, to more distal sinuous fluvial channels in a floodplain environment. Fluvial channels are migrating laterally, the extent of the lateral migration being dependent on the slope of the depositional paleotopography and the velocity of the current: at the foothill of alluvial fans, braided channels are confined and the channel belts they form are limited in term of width (approximately 1 to 2 km wide); in the floodplain, channel belts become wider (approximately 5 to 10 km wide) as the system is closer to the seafront, with channels becoming more sinuous and meandering. The variations in the style of fluvial bodies are accompanied by variations of facies and net-to-gross, grain sizes and porosity, grain sorting and permeability. Subsequently, the reservoir formations have been subdivided in a number of reservoir units displaying similar characteristics in terms of channel geometry to fit to the subsurface observations: orientation, sinuosity, dimensions (width, depth), and proportions over floodplain mudstones. Channel belts are typically modelled with stochastic modelling techniques, most commonly with object-based modelling which insures that the continuity and geometry of channels are preserved in the

reservoir model. In association with channels, crevasse splays are modelled stochastically along the banks of the channel belts as elliptical patches in the different reservoir units, with user defined variograms constraining the dimensions, orientation and connectivity of the patches. One of the key issues is that the facies modelling process is not constrained by seismic data. As a result, the position of channels in any layer is not constrained, and only submitted to the modeler's appreciation. Through history matching, a more adequate position of the channels can be deduced after numerous realizations are tested but many of the scenarios can match the production data.

7.3 Integrated workflow

7.3.1 Fast modelling update (FMU) concept

For the facies modeling and estimation we use an integrated model workflow. The modeling implementation is based on the Fast Model Update (FMU) work as presented by Zachariassen et al. (2011) and Skjervheim et al. (2012). In the model workflow all the model processes, from depth conversion to saturation modeling and model upscaling, are organized in a consistent, repeatable and automatic workflow. Great care is taken to ensure that changes in the data input, will result in a final reservoir model that is consistent with the input. The data input used and the individual model building steps, are the same as in a traditional model workflow. The real step change is that the model workflow can be run in batch mode, without any human interactions. Some additional scripting steps are included to make sure that the model workflow can be run in batch.

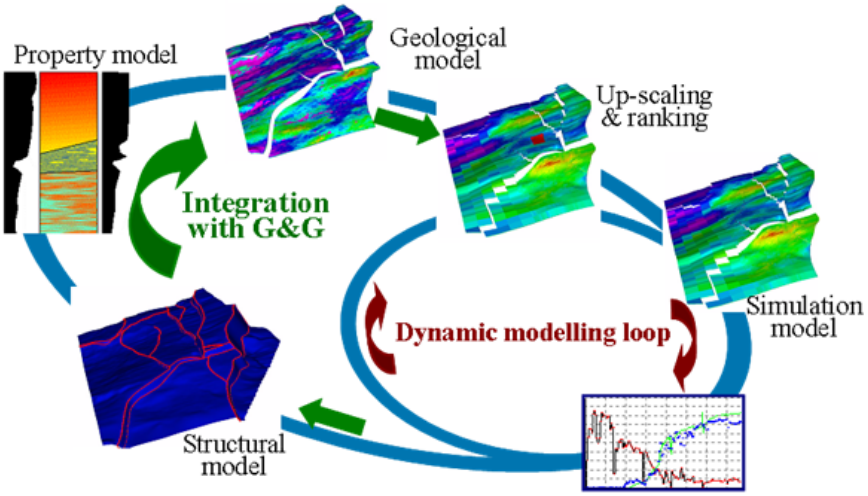


Figure 7.1: Integrated workflow

7.3.2 Reservoir model building workflow

One of the main ingredients in a big-loop conditioning workflow is the implemented model building workflow. This model building workflow consists of the following steps:

- Data preparation
- Structural modeling (horizons and faults)
- Gridding
- Well blocking
- Facies modeling
- Petrophysical modeling
- Upscaling
- Well modeling
- Export functions

These are all common modeling steps. But as mentioned above, each step has been made able to handle and propagate uncertainty, from uncertainty in the structural framework up till how this affects the well completion properties. The work presented here is done using the Roxar RMS software. This tool allows for a flexible and robust model building workflow. All the traditional steps are performed, but for some of them we use some extra software tools and smart scripts to allow for representation of the uncertainty in the input data. The facies modeling part of this workflow is of particular interest for us, as it is here we implement the suggested adaptive plurigaussian simulation (APS) method. The input to the facies modeling is the grid, the blocked well elements, information from seismic data, and our prior knowledge of the geological concept.

7.4 Set-up of the field case and results

As mentioned in the introduction, the paper introduces a facies modeling process that can be used in an integrated modeling workflow presented in the left side Figure 7.2.
We have a prior description of the subsurface geology for the reservoir, from different sources (seismic surveillance, core interpretations, outcrops, etc.). This prior knowledge is the result of reservoir exploration and is given by the experts (geologists, geophysicists, etc.) in the initial stage of deposit geology description. This includes the number of the facies types that occurs and the possible contact (transition) between the facies

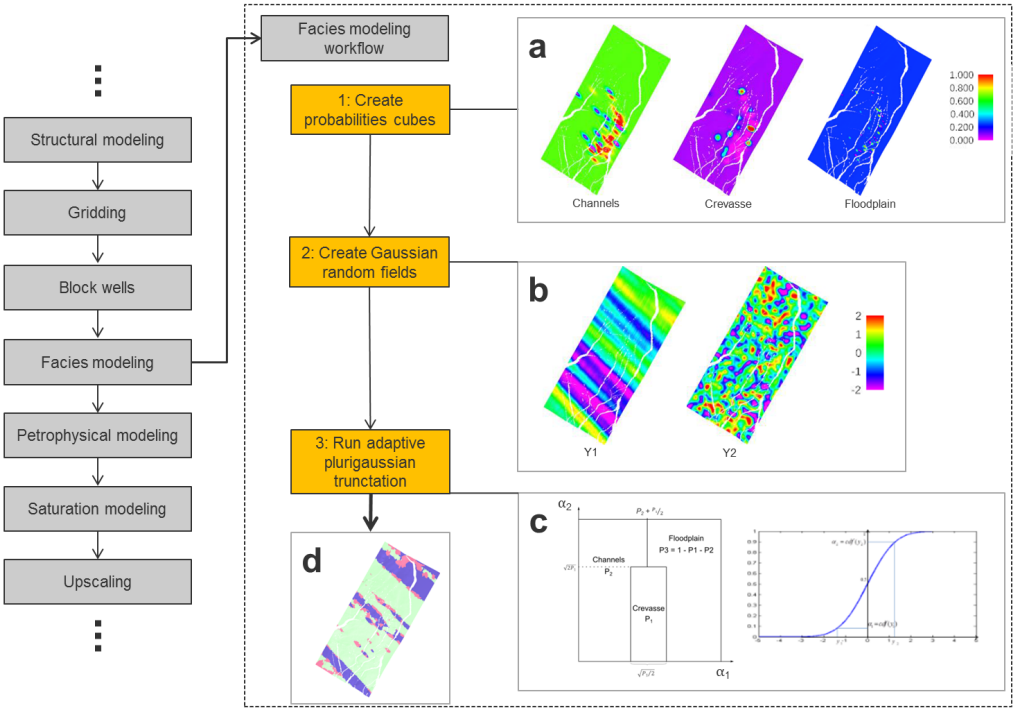


Figure 7.2: A visual representation of our integrated reservoir modeling workflow, with a detailed presentation on the facies modeling workflow. (a) shows the facies probability maps, (b) shows the random gaussian fields, (c) presents the truncation map and the normal cumulative distribution function and (d) gives an example of a facies realization.

types. Based on all prior information collected (e.g. seismic, core analysis), an occurrence probability map can be created for each facies type (see Figure 7.2a). The probabilities maps are constrained in each well on the observation of the facies type situated in subsurface (hard data). In each map we have value "1" at the well locations where the associated facies type occurs and value "0" at the well locations where the associated facies type does not occur. These maps are representing the initial input for the facies modeling process and they will not be changed throughout the whole process.

Moreover, these initial probabilities maps are honored and their geological properties are kept throughout the generation and updating/conditioning process of the facies realizations. This is ensured by the geological concept represented by the choice of two GRF's (Y_1 and Y_2 see Figure 7.2b). The two GRF's are the control variables of our optimization process and will be updated by the available observations/measurements.

For each grid cell of the domain, we create a bi-dimensional map, like in left side of Figure 7.2c. We define this map as the simulation map. In this quadratic domain, the total area is 1 and the area of the regions occupied by each facies type is equal with the associated probabilities found in the maps provided by the experts. The distribution of the regions in the square allows modeling the contact between any two facies type and the topological characteristics of the facies types. At the well locations, the map consists on a single square of surface 1, occupied by the facies type observed there. Consequently, there might be some parts of the reservoir domain where only two facies types occur (in terms of probabilities, only two probabilities are greater than zero). In this case, the map contains only two rectangular regions, each with area equals with the associated probability of the facies type that occupies the region.

Now, in each grid cell i, we need to simulate a facies type, conditioned to the probabilities given by the experts. For that, we apply the APS model. We generate samples for the Gaussian fields Y_1 and Y_2 and assign the facies type depending on where the point $(cdf(Y_1^i), cdf(Y_2^i))$ falls in the simulation map built for the grid cell i. We generate an ensemble of 100 realizations conditioned to the probability maps provided by the experts and the probability maps calculated from the ensemble is consistent with the originals (Figure 7.3). In this manner we have defined an ensemble of facies maps realizations, conditioned to the probability maps provided by the geologists. These realizations depend only on the Gaussian fields used for simulations and are consistent with the facies observations at the well locations.

After ensuring the above mentioned properties of the facies realization, another key issues in a multiple realizations world is the variability of the prior ensemble ensemble, which ensures that we are spanning the full uncertainty space (see Figure 7.4). Finally, the applicability and the success of the implementation of APS methodology for a real field case is shown by the fact that the facies realization obtained were more consistent and closer with the geological concept and the hard and soft data than the ones obtain previously for the same field using different methodologies (indicator kriging). In the Figures 7.5, 7.6 and 7.7 realizations of facies maps obtained with the indicator kriging (bottom left) and with the APS (bottom right) are shown. At the top of the figures we present the realizations of the Gaussian fields that provide the facies maps with the APS methodology. A valuable prior information is related to the relative position of the crevasses (green) with respect to the floodplain (yellow) and the channel belt (red). The crevasses are encountered mainly on the border of the channel belt. With a visual inspection, one can observe that this restriction cannot be kept with the indicator kriging methodology, whereas the APS methodology was able to provide facies maps with plausible topology.

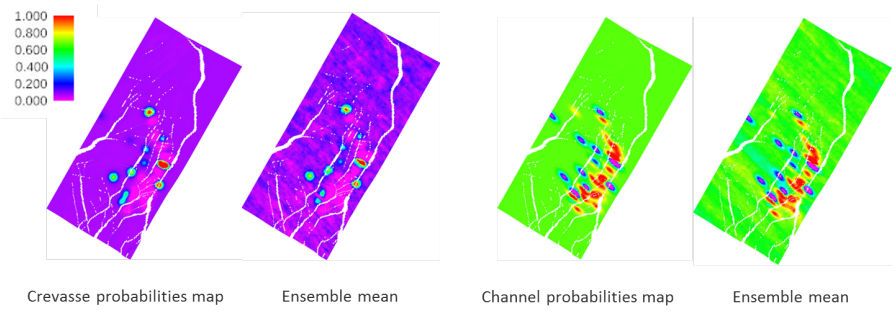


Figure 7.3: The probability map of each facies type provided by the experts and the probability maps calculated from the ensemble of 100 realizations

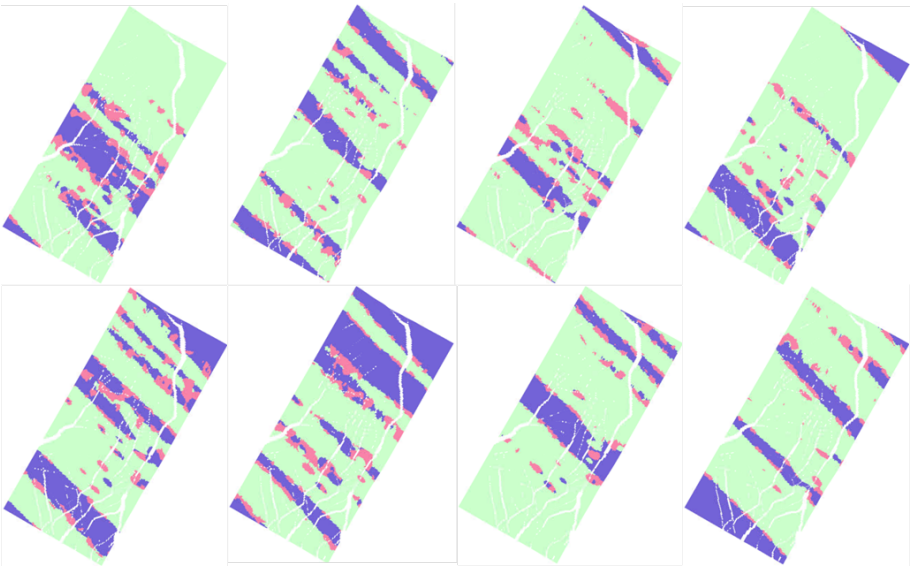
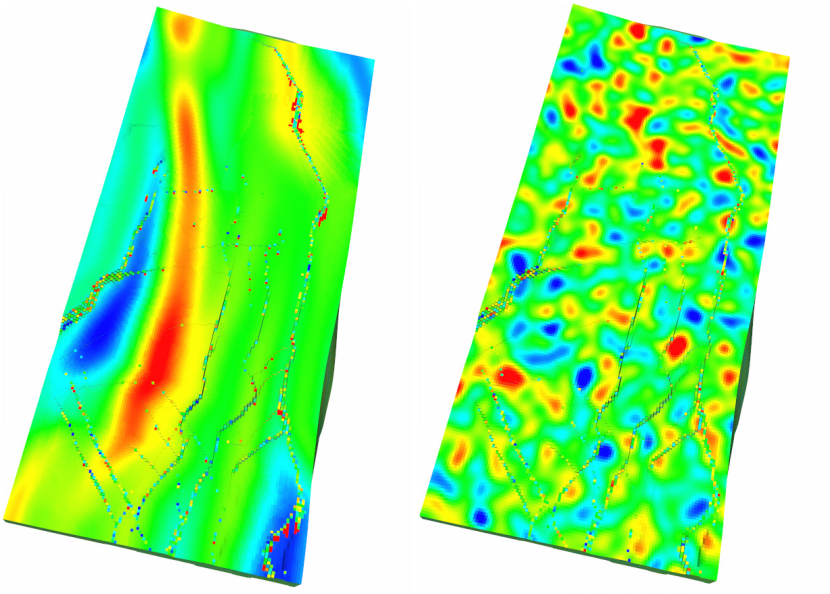
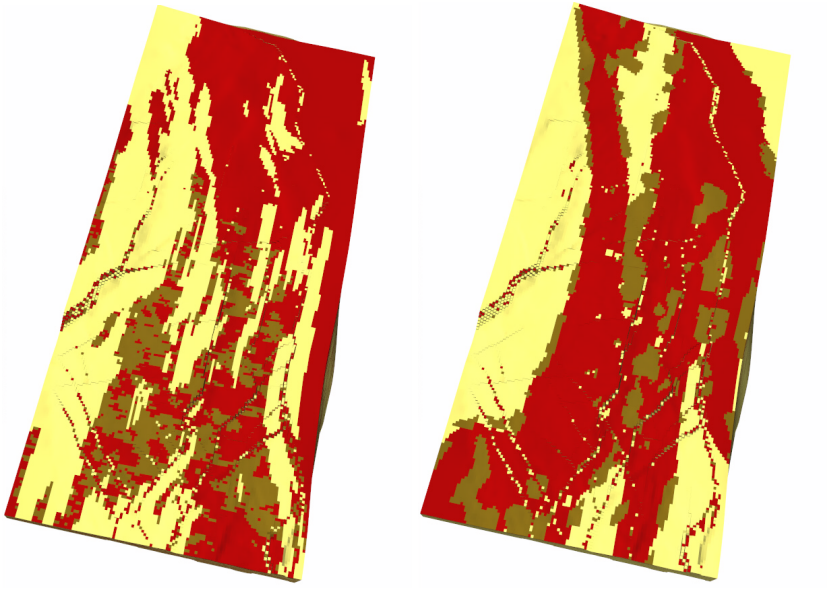


Figure 7.4: An example of 8 different members of the prior ensemble

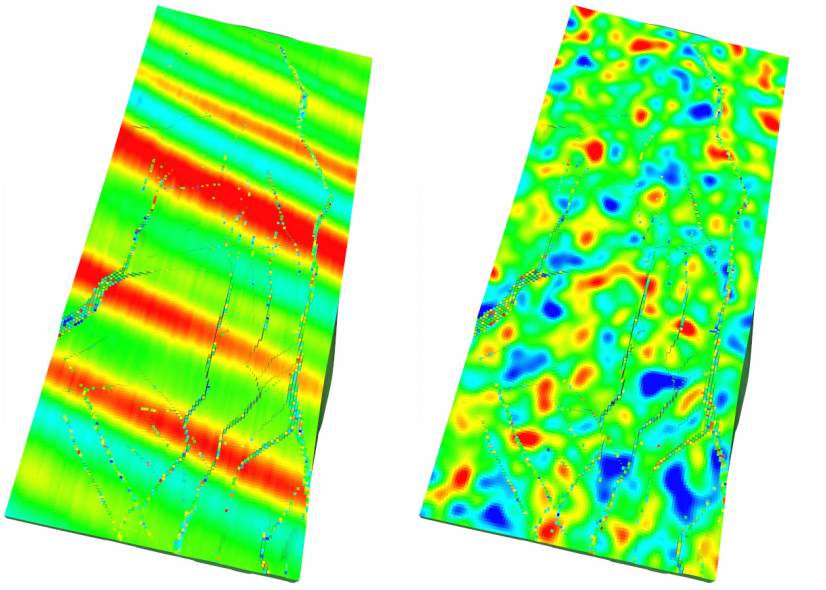


(a) 1a (b) 1b

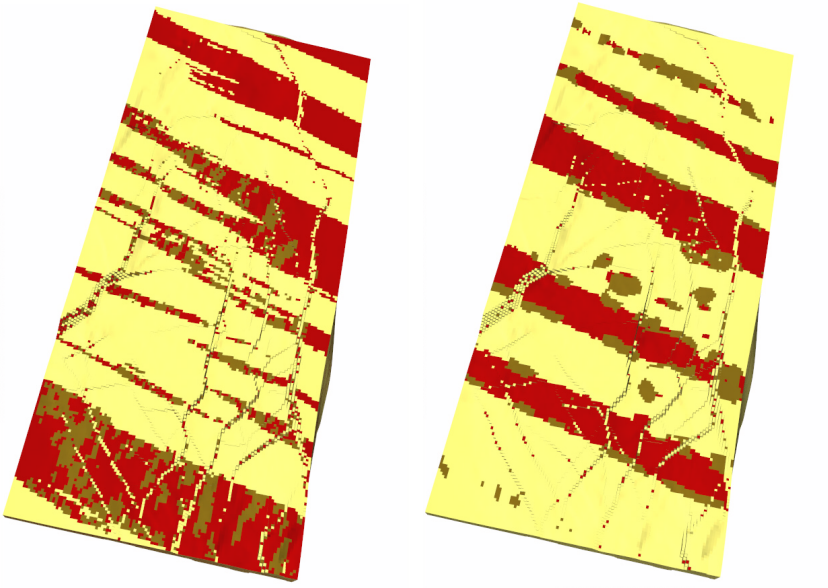


(c) 1c (d) 1d

Figure 7.5: Gaussian fields for channel belt and the crevasse for layer 37 (1a and 1b); Generated facies realizations with indicator kriging 1c and the APS methodology 1d

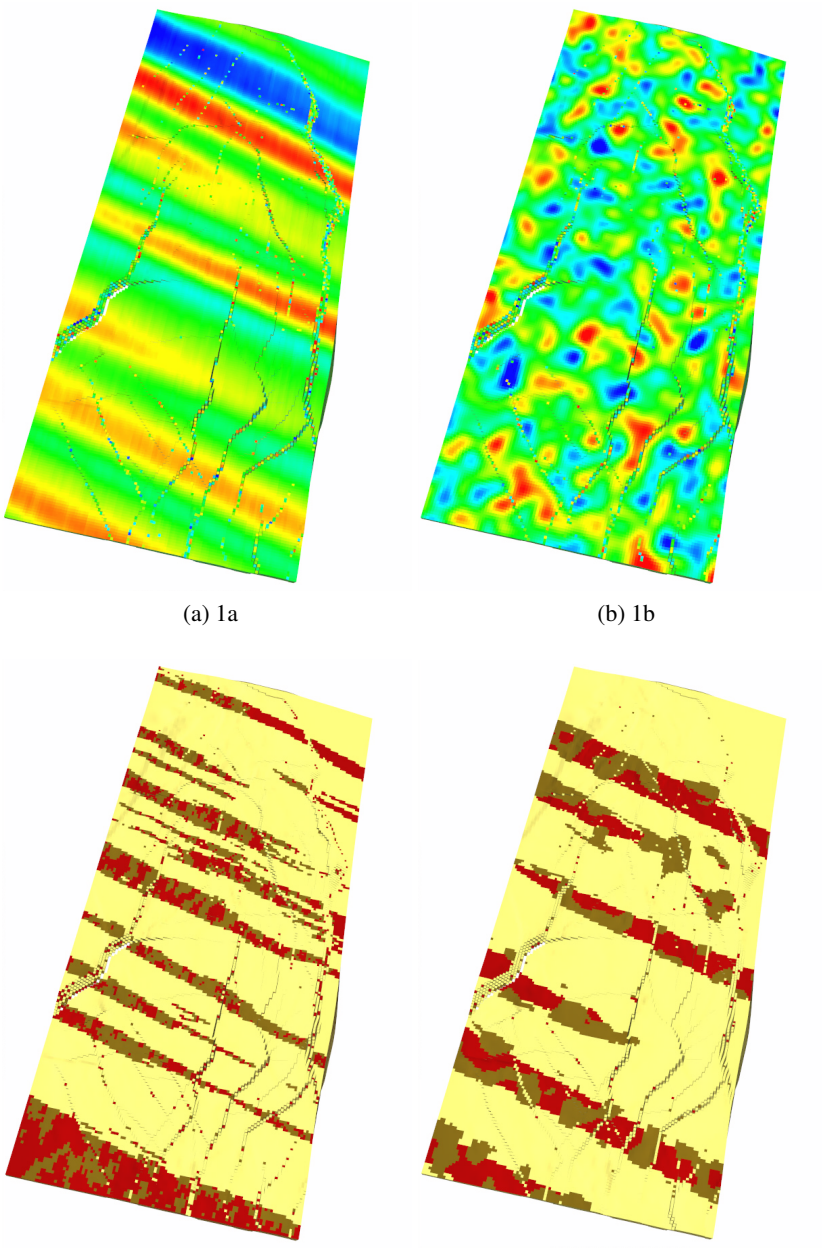


(a) 1a (b) 1b



(c) 1c (d) 1d

Figure 7.6: Gaussian fields for channel belt and the crevasse for layer 85 (1a and 1b); Generated facies realizations with indicator kriging 1c and the APS methodology 1d



(a) 1a (b) 1b

(c) 1c (d) 1d

Figure 7.7: Gaussian fields for channel belt and the crevasse for layer 182 (1a and 1b); Generated facies realizations with indicator kriging 1c and the APS methodology 1d

Chapter 8

Conclusions

This research focussed on new parameterizations of facies fields, which are coupled with reliable AHM methods for geology estimation and uncertainty quantification. Two directions have been used for facies fields generation: the plurigaussian simulation model (TPS, Chapter 4, 5 and 7) and the multiple point geostatistical model (MPS, Chapter 6). For the TPS model, the parameterization is straightforward with the Gaussian fields themselves, for the MPS model a parameterization is not easy to define. The parameterization proposed is carried out in a multidimensional Gaussian space by bridging the TPS methodology with the MPS.

8.1 Conclusions

- In Chapter 4 we presented a new approach of the facies estimation problem with EnKF, applied in the context of the truncated plurigaussian method. The truncation map that is used for generation of the facies maps arises from a probabilistic approach of the facies type fields. Actually, each probability can be viewed as a "distance" (to 0 or 1) that allows us to see how far we are to each facies type. This "distance" does not expresses a remoteness from a border (such as in the level set method) but shows the remoteness from two instances ("to be" which is 1 and not "to be" which is 0). Starting with the assumption that the highest "distance" to 0 gives us the facies type in a certain location, we have built a geological simulation model, which is further introduced in the EnKF context. The model is tested for the case with three facies types occurring, with the property that we could have transition between each two. The model is tested for different geostatistical properties of the Gaussian fields and different levels of uncertainty

introduced in generation of the model parameters. The novelty of the method has three directions.

- The first one is related to the truncation map. In the previous studies, where the truncation of two Gaussian fields has been made with a bidimensional map, it was constructed in an empirical manner and including the assignment of each facies types in the demarcated areas generated by the intersecting lines of the map. The truncation map that appears in our study arises from the internal construction of the model and also the assignment of the facies type in each zone of the map. The map depends only on two parameters which, if changed, perturb neither the geometry of the map, nor the assignment of the facies types in the map zones. In addition, the parameters have the ability to control the facies proportions. Consequently, the map parameters can be introduced in the state vector for a better description of the uncertainty, improving in the estimation process.
- The second is related to the observation operator of the facies type at the well locations. In the previous studies the facies observation operator is a proxy function which measures the affiliation of the pair of Gaussian fields to a specific area, in terms of yes or no. In our study we use an operator that measures the occurrence probability of each facies type. In this way we have a multiple condition for the Gaussian fields, in the places where a facies type occurs (the probability is 1) and in places where a facies type does not occurs (the probability is 0). Consequently, our Gaussian field ensemble is not conditioned to the map, but to the observations. From here arises the third main difference.
- With the introduction of "the probabilities fields" in the context of the truncation plurigaussian approach, we have tried to explain how the Gaussian fields can be interpreted in the truncation simulation. In the past work the Gaussian fields did not have a meaning, they were only useful for the facies maps generation. With this new approach, we gave a meaning to the Gaussian fields in the sense that their projection could be seen as probabilities. As a consequence of the model (and hence, of the map), the probability maps for each facies type calculated from the facies field ensemble are consistent with the ensemble averages of each probabilities fields (Figure 4.9)

At the end of the assimilation period, as result of the parameters estimation, we are able to offer a field, which we have named "estimated field", that is an estimator of the truth field.

- In the exploration phase a geological description of the subsurface is developed by the geologists and geophysicists. One important feature is represented by the facies probability maps obtained based on all available information (well logs, seismic data etc.). In this study, we have introduced a methodology that relates the work developed in the exploration phase with the data assimilation process, as valuable prior information.

  In Chapter 5, we have designed a geological simulation model in a plurigaussian simulation context, which is able to sample facies maps from the prior probability maps of the facies types. We named the geological simulation model the adaptive plurigaussian simulation (APS) model. The main differences with other approaches involving Gaussian fields truncation/simulation resides in the way in which the facies simulation/truncation maps are introduced and by the fact that not only a single map is used. In our approach, we do not construct a truncation map in the Gaussian space, but for each grid cell, we construct a map in a domain of measure 1 of a metric space. We called this map the facies simulation map. The map is based on the decomposition of the unitary domain in some zones assigned to the facies types of which areas must be equal to the probabilities found in the maps provided by the experts. The decomposition should be made in such a way that the contact between the facies types and the topology of the facies field reflect the reality. Therefore, a layout and a parametrization of the simulation map is needed. Consequently, the core information collected at the wells, being incorporated in the prior facies probability maps, will drive the simulation to always provide the correct facies type. This means that, throughout all the assimilation period, the facies types at the well location are not changed and satisfy the observations and is not necessary to include extra procedures to reposition the facies types at the well locations, as in previous studies. To sample from the discrete variable named "facies type", of which distribution is known, we connect the uniformly distributed variables with the Gaussian distributed variables through the normal cumulative distribution function (cdf) using the probability integral transform. Based on the cdf projection we have created spatially correlated fields with values in the [0,1] interval which at the starting point have uniform distribution. For that, at initial ensemble generation the Gaussian fields are generated uncorrelated and unconstrained (stationary Gaussian random fields). This will ensure that at each grid cell the facies type simulation will have the correct statistics at the beginning of the data assimilation. At each grid cell, the facies simulation map remains unchanged after each update being the object that relates the prior information provided by the experts with the plurigaussian simulation methodology. Basically, at the initial moment we simulate from the map with random variables from uniform distribution on [0,1] and once the available measurements condition

the model parameters, we sample from the same map with random variables with values in [0,1] but not having the same uniform distribution. Keeping the simulation map (for each grid cell) throughout assimilation process means two important things related with the expert knowledge.

- Sampling from the same statistics at each moment but with variables (parameters) at the beginning unconditioned and then conditioned to measurements (due the Bayesian inversion scheme)
- Keeping the topological request of the facies distribution.

An important novelty is the guiding direction that we constrain the EnKF to follow, introducing the facies observations as measurements at each assimilation time step. This extra constraint does not reduce the variability in the updated ensemble drastically and only ensures that the ensemble goes in the right direction.

- In Chapter 6 we have presented a new parameterization of facies fields generated using multi point statistics (MPS). Since the truncated (pluri) Gaussian method is well suited for facies updating with ensemble methods, but struggles with channel structures (since the variogram has to be specified) we suggested estimating a marginal facies probability field from an ensemble generated using MPS, which is well suited for channel type reservoirs, and parameterize this probability field with a truncated marginally Gaussian field preserving the facies probabilities. The advantage is that the dependence structure is inherited from the training image and do not have to be specified. Again we stress that prescribing an analytical covariance function that generates channels in combination with a truncation map is (probably) an impossible task. Since the result is a truncated Gaussian field it is well suited for ensemble based methods and we are guaranteed that the updated permeability fields are facies fields without specifying a truncation. In addition, we combined the parameterization with the iterative adaptive Gaussian mixture (IAGM) filter for ensemble based Bayesian inversion. There were two main reasons for this. First, the IAGM is more suited for highly nonlinear problems then the EnKF, secondly, the IAGM has a resampling step which allowed us to use the already existing technique of resampling from the training image using updated probability maps. However, since the IAGM can run assimilate all the data several times, there is no need for resampling after each update step and we only have to perform resampling after each iteration (typically 2-4). We performed experiments on two channel type synthetic reservoirs and the conclusion drawn from the experiments were the same. The combination of the parameterization and the IAGM resulted in an advantage over existing methods in the sense that after a

few iterations, the updated fields was able to keep the channel structure without performing another resampling. Since the parameterization guarantees that the updated fields are facies fields, the output of the algorithm was an ensemble of channelized reservoir models that had low data mismatch and good predictive capabilities. If we had worked directly with log-permeability fields with we have to make a compromise.

- In Chapter 7 we have presented a proof of concept for the APS methodology by successfully implemented it and got very reliable results for a real field case in the North Continental Shelf (NCS). We have obtained realistic facies simulations for a real field case, consistent with the hard and soft data as well as with the geological concept. This simulation process is repeatable and automated/assisted by being integrated in the Fast Modeling Update (FMU) framework of the big loop conditioning philosophy. We are proposing an integrated workflow for facies generation and estimation using a simulation model that is able to create a link between the mathematical theory and real world.

Bibliography

1. Aanonsen S. I., Naevdal G., Oliver D. S., Reynolds A. C., and B. Vall'es, The Ensemble Kalman filter in reservoir engineering a review, SPE Journal, pp. 393-412, 2009.

2. Abrahamsen P., A review of Gaussian Random Fiedls and Correlation Functions, 1997.

3. Agbalaka C. C., Oliver D. S., Application of the EnKF and localization to automatic history matching of facies distribution and production data, Mathematical Geosciences, 40, 353-374, 2008.

4. Agbalaka C. C., Oliver D. S., Automatic history matching of production and facies data with non-stationary proportions using EnKF, in SPE Reservoir Simulation Symposium, The Woodlands, Texas, USA, 2009.

5. Armstrong M., Galli A., Beucher H., Le Loc'h G., Renard D., Doligez B., Eschard R. and Geffroy F., Plurigaussian Simulations in Geosciences. Springer Berlin Heidelberg, 2nd revised edition edition, 2011.

6. Arpat G., Caers J., A multi-scale pattern based approach to sequential simulation. Geostatistics Banff, Vol 1, pp 255-264, Springer, Dordrecht, 2004.

7. Avseth P., Mukerji T., Jorstad A., Mavko G., Veggeland T.: Seismic reservoir mapping from 3-D AVO in a North Sea turbidite system. Geophysics vol. 66. no. 4, pp. 1157-1176, 2001.

8. Bengtsson T., Snyder C., Nychka D., Toward a nonlinear ensemble filter for high-dimensional systems. J. Geophys. Res., 108:35-45, 2003.

9. Bengtsson T., Bickel P., Li B., Curse-of-dimensionality revisited: Collapse of particle filter in very large scale systems. Probability and Statistic, 2:316-334, 2008.

10. Burgers G, van Leeuwen P. J., Evensen G., Analysis scheme in the ensemble Kalman Filter. Mon Weather Rev 126, 1719-1724, 1998.

11. Caers J. and Zhang T., Multiple-point geostatistics: A quantitative vehicle for integrating geologic analogs into multiple reservoir models. In: Grammer GM (ed) AAPG memoir,integration of outcrop and modern analogs in reservoir modeling: AAPG Memoir 80, p. 383-394, 2004.

12. Caers J. and Hoffman T., The probability perturbation method:a new look at Bayesian inverse modeling. Math. Geol. 38, 81-100, 2006.

13. Casella G., Berger R. L., Statistical Inference, Duxbury, Thomson Learning, second edition, 2002.

14. Chang H., Zhang D. and Lu Z., History matching of facies distribution with the EnKF and level set parametrization, Journal of Computational Physics, 229(20), 8011-8030, 2010.

15. Chen W. H., Gavalas G. R., Seinfeld J. H., and Wasserman M. L., A new algorithm for automatic history matching. SPE Journal 14(4), 593-608, 1974.

16. Christoakos G., Random field models in earth sciences, Academic Press, New York, 1992.

17. Christoakos G., On the problem of permissible covariance and variogram models, J Water Resources Res 20, No 2, pp 251, Feb 1984.

18. Collinson J. D., Alluvial sediments. 20-61 in Sedimentary environments and facies (second edition). Reading, H G (editor). (Oxford: Blackwell Scientific Publications.) 1986.

19. Deutsch C. V., Journel A. G., GSLIB Geostatistical Software Library and User's Guide,Applied Geostatistics Series, second ed., Oxford University Press, 1998.

20. Deutsch C. V. , Wang L., Hierarchical object-based modeling of fluvial reservoirs. Mathematical Geology 28, 857-880, 1996.

21. Dore A. G., Synoptic palaeogeography of the Northeast Atlantic Seaway: late Permian to Cretaceous. 421-446 in Basins on the Atlantic Seaboard: Petroleum Geology Sedimentology and Basin Evolution. Parnell, J (editor). SPeicla Publication of the Geological Society of London, n. 62, 1992.

22. Doucet A., de Freitas N., Gordon N. (eds.): Sequential Monte-Carlo methods in practice. Springer, New York, 2001.

23. Dovera L. and Della Rossa E., Multimodal ensemble Kalman filtering using Gaussian mixture models, Computational Geosciences, Volume 15, Issue 2, pp 307-323, 2011.

24. Evensen G., Sequential data assimilation with a non-linear quasigeostrophic model using Monte-Carlo methods to forecast error statistics. J. Geophys. Res. 99, pp 143-162, 1994.

25. Evensen G., The Ensemble Kalman Filter: theoretical formulation and practical implementation. Ocean. Dyn. 53, pp 343-367, 2003.

26. Evensen G., Data Assimilation: The Ensemble Kalman Filter, Springer, 2006.

27. Fossen H., Odinsen T., Farseth R. B., and Gabrielsen T. H., Detachments and low-angle faults in the northern North Sea rift system. 105-131 in Dynamics of the Norwegian Margin. Nottvedt, A (editor). Special Publication of the Geological Society of London, n. 167, 2000.

28. Galli A., Beucher H., Le Loc'h G., Doligez B., Heresim Group, The pros and cons of the truncated Gaussian method. In: Geostatistical simulations. Kluwer Academic, Dordrecht, pp 217-233, 1994.

29. Gao G., Gaoming L., Reynolds A. C., A Stochastic Optimization Algorithm for Automatic History Matching, SPE Journal, Vol 12, Issue 2, pp 196-208, 2007.

30. Gaspari G. , Cohn, S. E., Construction of correlation functions in two and three dimensions. Q J R Meteor Soc 125,723-757, 1999.

31. Guardiano F., Srivastava M., Multivariate geostatistics:beyond bivariate moments. In: Geostatistics-Troia, vol1, pp. 133-144. Kluwer Academic, Dordrecht, Holland, 1993.

32. Hanea R. G., Ek T., Sebacher B., Saetrom J., Sollien D.B., Geologically Realistic Facies Updates for a North Sea Field, Extended abstract In Proceedings of the 76th EAGE Conference and Exhibition, Amsterdam June 16-19, 2014

33. Honarkhah M., Caers J., Direct Pattern-Based Simulation of Non-stationary Geostatistical Models, Math Geosci Volume 44, pp 651-672, 2012.

34. Hoteit I., Pham D. T., Korres G., Triantafyllou G., Particle Kalman filtering for data assimilation in meteorology and oceanography. In 3rd WCRP International Conference on Reanalysis, pp.6, Tokyo, Japan, 2008a.

35. Hoteit I., Pham D.T., Triantafyllou G., Korres G., A New approximative solution of the optimal nonlinear filter for Data Assimilation in Meteorology and Oceanography. Mon. Wea. Rev., 136:317-334, 2008b.

36. Hu L. Y., Zhao Y., Liu Y., Scheepens C., and Bouchard A., Updating multipoint simulations using the ensemble Kalman filter. Computers and Geosciences, 51(0):7-15, 2013.

37. Hu L.Y., Le Ravalec-Dupin M., An Improved Gradual Deformation Method for Reconciling Random and Gradient Searches in Stochastic Optimizations, Mathematical Geology, Vol. 36, No. 6, 2004.

38. Hu L.Y., Liu Y., Scheepens C., Shultz A. W., Thompson R. D., Multiple-Point Simulation with an Existing Reservoir Model as Training Image, Math Geosci, Volume 46, Issue 2, pp 227-240, 2014.

39. Kelkar M. and Perez G., Applied Geostatistics for Reservoir Characterization: Society of Petroleum Engineers, 264 p., Richardson Texas, 2002.

40. Krishnan S., Boucher A. and Journel A. G., Evaluating information redundancy through the tau model, in O. Leuangthong and C.V. Deutsch (eds.), Geostatistics Banff Vol2, Springer, pp. 1037-1046, 2005.

41. Jafarpour B., Wavelet reconstruction of geologic facies from nonlinear dynamic flow measurements. IEEE Trans Geosci Remote Sens. doi:10.1109/TGRS.2010.2089464, 2010.

42. Jafarpour B., Khodabakhshi M., A probability conditioning method (PCM) for nonlinear flow data integration into multipoint statistical facies simulation. Math. Geosci. 43, 133-164, 2011.

43. Jafarpour B., McLaughlin D., Efficient permeability parametrization with discrete cosine transform. In: The Proceedings of the SPE Reservoir Simulation Symposium, Houston, TX, Paper SPE 106453, 2007.

44. Jafarpour B., Tarrahi M., Assessing the performance of the ensemble Kalman filter for subsurface flow data integration under variogram uncertainty, Water Resources Research, 47, W05537, 26 PP., 2011.

45. Jazwinski A., Stochastic processes and filtering theory. Academic Press, San Diego, 1970.

46. Journel A. G., Combining knowledge from diverse data sources: An alternative to traditional data independence hypothesis: Math. Geol., v. 34, p. 573-596, 2002.

47. Journel A. G., Multiple-point Geostatistics: A state of the art. In Report 16, Stanford Center for Reservoir Forecasting, Stanford CA, 1993.

48. Lantuejoul C., Geostatistical simulation: models and algorithms. Spinger, Berlin, 2002.

49. Lake L. W., Holstein E.D., Reservoir Engineering and Petrophysics-Petroleum Engineering Handbook, Vol V, Society of Petroleum Engineers, 2007

50. Le Loc'h G., Beucher H., Galli A., Doligez B., Heresim Group, Improvement in the truncated Gaussian method: combining several Gaussian functions. In: Proceedings of ECMOR IV: fourth European conference on the mathematics of oil recovery, 1994.

51. Le Loc'h G., Galli A., Truncated plurigaussian method: theoretical and practical points of view. In: Baafi EY, Schofield NA (eds) Geostatistics wollongong '96, vol 1. Kluwer Academic, Dordrecht, pp 211-222, 1997.

52. Liu N. and Oliver D. S., Critical evaluation of the ensemble Kalman filter on history matching of geologic facies, SPE Reservoir Evaluation and Engineering, 8(6), 470-477, 2005a.

53. Liu N. and Oliver D. S., Ensemble Kalman filter for automatic history matching of geologic facies, Journal of Petroleum Science and Engineering, 47(3-4), 147-161, 2005b.

54. Lorentzen R. J., Flornes K. M., and Naevdal G., History matching channelized reservoirs using the ensemble Kalman filter, SPE Journal, Vol 17, Issue 1, pp 137-151, 2012.

55. Lorentzen R. J., Naevdal G.,Shafieirad A., Estimating facies fields using the ensemble Kalman filter and distance functions, applied to shallow-marine environments, SPE Journal, Vol 3, Issue 1, pp 146-158, 2012.

56. Mandel J., Beezley J. D., An ensemble Kalman-particle predictor-corrector filter for non-Gaussian data assimilation. In ICCS 2009: Proceedings of the 9th International Conference on Computational Science, pages 470-478, Berlin, Heidelberg, Springer-Verlag, 2009.

57. Massonnat G. J., Breaking of a Paradigm: Geology Can Provide 3D Complex Probability Fields for Stochastic Facies Modelling. SPE Annual Technical Conference and Exhibition, Houston, Texas, 3-6 October, 1999.

58. Matheron G., Beucher H., de Fouquet C., Galli A., Gue'rillot D., Ravenne C., Conditional simulation of the geometry of fluvio-deltaic reservoirs. SPE 1987 Annual technical conference and exhibition, Dallas, Texas, pp 591-599. SPE 16753, 1987.

59. Moreno D., Coupling Level Set Methods with the Ensemble Kalman Filter for Conditioning Geological Facies Models to Well and Production Data, Ph.D. thesis, University of Bergen, Norway, 2009.

60. Oliver D. S., Moving averages for Gaussian simulation in two and three dimensions, Mathematical Geology, 27(8), 939-960, 1995.

61. Oliver D. S. and Chen Y., Recent progress on reservoir history matching: a review. Computational Geosciences 15(1), 185-221, 2011.

62. Oliver D. S., Reynolds A. J., Liu N., Inverse theory for petroleum reservoir characterization and history matching. Cambridge University Press, Cambridge, 2008.

63. Osher S., Fedkiw R., Level Set Methods and Dynamic Implicit Surfaces, Springer, 2003.

64. Remy N., Boucher A., Wu J., Applied geostatistics with SGeMS a user's guide, Cambridge University Press, 2009.

65. Roggero F., Hu L., Gradual deformation of continuous geostatistical models for history matching. SPE 49004, presented at the SPE annual technical conference and exhibition, New Orleans, LA, 27-30 September, 1998.

66. Sarma P., Durlofsky L., Aziz K., Kernel principal component analysis for efficient, differentiable parameterization of multipoint geostatistics. Math Geosci 40(1),3-32, 2008.

67. Sarma P., Chen W., Generalization of the Ensemble Kalman Filter using kernels for non-Gaussian Random Fields. In SPE Reservoir Simulation Symposium, The Woodlands, Texax, February 2-4, 2009.

68. Sebacher B., Hanea R., Heemink A., A probabilistic parametrization for geological uncertainty estimation using the ensemble Kalman filter (EnKF), Computational Geosciences, 2013, Volume 17, Issue 5, pp 813-832, 2013.

69. Sebastian Ng., Dahle Pg., Hauge R., Estimation of facies probabilities on the Snorre field using geostatistical AVA inversion. SEG Technical Program Expanded Abstracts 01/2008, 27(1), DOI:10.1190/1.3059282, 2008.

70. SGeMS The Stanford Geostatistical Modeling Software. http://sgems. sourceforge. net/. Accessed 1 March 2012.

71. Stordal A., Karlsen H., Naevdal G., Skaug H. and Valles B. Bridging the ensemble Kalman filter and particle filters. Computational Geosciences, 15(2), 293-305, 2011.

72. Stordal A., Lorentzen R. J., An iterative version of the adaptive Gaussian mixture filter. Computational Geosciences DOI 10.1007/s10596-014-9402-6, 2014.

73. Strebelle S., Conditional simulation of complex geological structures using multiple-point statistics. Mathematical Geology, vol 34, pp.1-22,2002.

74. Strebelle S., Zhang T., Non-Stationary multiple-point geostatistical models, In O. Leuangthong and C. V. Deutsch (eds.), Geostatistics Banff, pp. 235-244, 2005.

75. Skjervheim J. A., Evensen G., An Ensemble Smoother for Assisted History Matching, SPE Reservoir Simulation Symposium, 21-23 February, The Woodlands, Texas, USA, 2011.

76. Talagrand O., Courtier P., Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory. Quart. J. Roy. Meteor. Soc., 113, 1311-1328, 1987.

77. Tunbridge I. P., Facies model for a sandy ephemeral stream and clay playa complex; the Middle Devonian Trentishoe Formation of North Devon, UK. Sedimentology, Vol. 31, 697-715, 1984.

78. Van den Zwan C. J., and Spaak P., Lower to Middle Triassic sequence stratigraphy and climatology of the Netherlands, a model. Palaeogeography, Palaeoclimatology, Palaeoecology, Vol. 91, 277-290, 1992.

79. Vasco D. W. and Datta-Gupta A., Integrating field production history in stochastic reservoir characterization. SPE Formation Evaluation 12(3), 149-156, (36567-PA), 1997.

80. Vasco D. W., S. Soon and A. Datta-Gupta, Integrating dynamic data into high-resolution reservoir models using streamline-based analytic sensitivity coefficients, SPE Journal 4(4), 389-399, 1999.

81. Xu C., Dowd P. A., Mardia K. V., Fowell R. J., A flexible true plurigaussian code for spatial facies simulations. Computational Geosciences 32, 16291645, 2006.

82. Zachariassen E., Skjervheim J.A., Vabo J. G., Lunt I., Hove J. and Evensen G., Integrated Work Flow for Model Update Using Geophysical Monitoring Data, 73rd EAGE Conference & Exhibition, Vienna, 2011.

83. Zafari M., Reynolds A.C., Assessing the Uncertainty in Reservoir Description and Performance using the Ensemble Kalman filter. SPE Journal 12(3), 381:391, 2007.

84. Zhang F. and Reynolds A. C., Optimization algorithms for automatic history matching of production data. In Proceeding of the 8th European Conference on the Mathematics of Oil Recovery, Freiberg, Germany, 2002.

85. Zhang T., Journel A.G., Switzer P., Filter-based classification of training image patterns for spatial simulation, Mathematical Geology, Vol. 38(1), pp 63-80, 2006.

86. Zhao Y., Reynolds A. C., Li G., Generating facies maps by assimilating production data with EnKF paper spe113990. In:Proceedings of SPE/DOE symposium on improvement oil recovery, Tulsa 20-23 April, 2008.

87. Zhou H., Li L., Gomez-Hernandez J. J., Hendricks Franssen H.J., Pattern Recognition in a Bimodal Aquifer Using the Normal-Score Ensemble Kalman Filter. Mathematical Geosciences, Vol. 44, 169:185, 2012.

Summary

An important key in reservoirs engineering is the development of reliable reservoir models with high predictive capacity of production behavior in the existing and new drilled wells. The reservoir predictability is the starting point for guiding the production plan to an optimum. In order to achieve this, the internal geological description of the reservoir plays a crucial role. When referring to the internal geological description of the reservoir, one may think of a description of the spatial distribution of the petrophysical properties (permeability and/or porosity) or more natural, on a description of the spatial distribution of the geological deposit formations (facies). The prior modeling of the facies distribution is carried out using geostatistical tools in specialized modules, defining thus, the geological simulation model and its uncertainty. The choice of the geological simulation model for facies field simulation takes into account the geological particularities of the reservoir. Irrespective of the geological simulation model used, the poorly knowledge of geology leads to an initial reservoir model whose parameters are uncertain. This prior uncertainty is usually so large that the predictive capacity is limited. To improve the predictive capacity, the prior uncertainty is reduced in an Assisted History Matching (AHM) process. To be able to couple the geological simulation model with an AHM method, a parameterization of the facies fields is needed. The research objective of this thesis was to elaborate reliable parameterizations of the facies fields in a such manner that, introduced in an AHM process, the updates to be able to: match the past data within the model and measurement uncertainties; incorporate all types of information (different measurement types) and preserve the geological realism (consistency with the initial geomodel). The research was oriented in two directions, depending on the geological simulation model used.

In the context of truncated plurigaussian simulation model (TPS) two methodologies were elaborated. The first methodology is tailored for cases where probability maps of the facies occurrence are not available, but only geometrical and topological informa-

tion. In this case, we parameterized the binary field defined by each facies type with a new object denoted "the probabilities field". The probabilities fields are defined by projection of Gaussian fields in [0,1] interval with a piece-wise linear projection function. We have created a geological simulation model, placed in the broader TPS methodology, with a truncation map not user defined but arising from the internal construction of the model. The AHM method used in this case is the Ensemble Kalman filter, conditioning the model parameters to production data (bottom hole pressures at the injectors and oil and water rates at the producers) and hard data (facies observations at the well locations). The conditioning on hard data was performed within an iterative process using a probabilistic operator of the facies occurrence. In addition, is the first time when the parameters that define the truncation map are introduced in the estimation process, showing improvement in the uncertainty quantification of the facies distribution.

The second methodology is reserved for the case where probability maps of the facies occurrence are available from a prior investigation carried out in the exploration phase of geology description. We introduced a geological simulation model for simulation of the facies fields so that the probability maps of the facies occurrence calculated from an ensemble to be consistent with the given probability maps. The simulation is realized at the grid cell level using uniform distributed variables, from a geometrical map designed in an unitary domain. The map layout is chosen so that the simulated facies fields obey to geometrical and topological requirements. The uniform variables used for facies simulation are obtained by projection of Gaussian fields with the cumulative distribution function. We called the geological simulation model the adaptive plurigaussian simulation model (APS). We coupled the APS model with the EnKF obtaining updates for the facies probability maps together with an uncertainty quantification of the geology. The model parameters (the Gaussian fields) are constrained to production data and global facies proportions for a better quantification of uncertainty. The model was tested on two synthetic models with different complexity and one real field, reporting promising results.

The second research direction refers to channelized reservoirs, modeled with a MPS geological simulation models (SNESIM). We present a new parametrization of the facies fields, which is further coupled with the iterative adaptive Gaussian mixture (IAGM) for history matching. The parameterization bridges ideas from MPS and TPS. The standard TPS parameterization ensures that the updated fields are facies fields, however, since the field is jointly Gaussian it is not trivial to characterize channel type facies structures via a covariance function (or variogram) in combination with a user defined truncation map. The MPS approach has the advantage of simulating channel type random fields, however, they are discrete-valued and it is not trivial to connect this with ensemble based

inversion. We therefore suggest a combination of the two parameterizations. By estimating marginal facies probabilities from a sample from MPS in combination with a training image, we parameterize the marginal probability field with a (marginally) Gaussian random field where the truncation map is defined such that the integral for a facies type in a given grid cell in the Gaussian domain coincides with marginal grid cell facies probability estimated from the ensemble. In addition, we combined the parameterization the iterative adaptive Gaussian mixture filter for ensemble based Bayesian inversion. There are two main reasons for this. First, the IAGM is more suited for highly nonlinear problems then the EnKF, secondly, the IAGM has a resampling step which allowed us to use the already existing technique of resampling from the training image using updated probability maps.

Samenvatting

En belangrijke component in reservoir technologie is de ontwikkeling van betrouwbare reservoir modellen met een hoog voorspellend vermogen van het productie gedrag in bestaande en nieuw geboorde putten. De voorspelbaarheid van het reservoir is het uitgangspunt voor een optimaal productie plan. De geologische beschrijving van het reservoir speelt een cruciale rol om dit te bereiken. Bij een geologische beschrijving kan men denken aan een omschrijving van de ruimtelijke verdeling van de petrofysische eigenschappen (permeabiliteit en/of porositeit) of meer gangbaar, een omschrijving van de ruimtelijke verdeling van de geologische afzettingen (facies). Het voorafgaande modelleren van de facies verdeling wordt gedaan in gespecialiseerde modellen met behulp van geostatistische technieken, waardoor de onzekerheid van het geologisch simulatie model zodoende hier wordt gedefinieerd. In de keuze van het geologisch simulatie model voor simulatie van facies velden wordt rekening gehouden met de geologische bijzonderheden. Een slechte kennis van de geologie leidt, ongeacht welk geologisch model wordt gebruikt, tot een reservoir model waarvan de parameters onzeker zijn. Deze voorafgaande onzekerheid is normaal gesproken zo groot dat het voorspellend vermogen beperkt is. Om het voorspellend vermogen te verbeteren wordt de voorafgaande onzekerheid verminderd in een Assisted History Matching (AHM) proces. Daarvoor is een parametrisering van facies velden noodzakelijk zodat het geologisch simulatie model verbonden kan worden met een AHM methode. Het doel van dit proefschrift is om een betrouwbare parametrisering van facies velden te ontwikkelen zodat de updates, geintroduceerd in het AHM proces, in staat zijn: de bestaande gegevens in het model en de meet onzekerheden te matchen; alle soorten van informatie the verenigen (verschillende type metingen) en geologisch realisme te behouden (in overeenstemming met het oorspronkelijke geologisch model). Het onderzoek was georienteerd in twee richtingen, afhankelijk van het gebruikte geologisch simulatie model.

Twee methoden werden ontwikkelend in de context van afgeknotte meervoudige gaussische simulatie modellen (TPS). De eerste methodologie is gericht op gevallen

waar geen waarschijnlijkheidskaarten van facies voorkomen beschikbaar zijn, maar alleen geometrische en topologische informatie. In dit geval wordt het binaire veld geparametriseerd en gedefinieerd door elk facies type met een nieuw object, aangeduid als "het waarschijnlijkheidsveld". Deze waarschijnlijkheidsvelden werden gedefinieerd door een projectie van de gaussiche velden in een [0,1] interval met stuksgewijze lineaire projectie functie. Een geologisch simulatie model is gecreeerd, geplaatst in breed TPS model, met een afgeknotte kaart niet gedefinieerd door de gebruiker, maar voortgekomen uit de interne constructie van het model. De AHM methode gebruikt in dit geval is de Ensemble Kalman filter (EnKF), waarin de parameters van het model worden geconditioneerd aan productie data (onderste put druk van de injectoren en de olie en water snelheden van de producenten) en harde gegevens (facies waarnemingen in de put). De conditionering op harde data werd uitgevoerd door een iteratief proces met behulp van een waarschijnlijkheidsoperator van facies voorkomen. Het was ook de eerste keer dat de parameters die de afgeknotte kaart definieren werden geintroduceerd in het schattingsproces en verbetering vertoonden in de onzekerheid kwantificering van de facies distributie.

De tweede methode is gericht op gevallen waar de waarschijnlijkheidskaarten van facies voorkomen, verkregen uit een voorafgaand geologisch onderzoek, beschikbaar zijn. In deze methodiek werd een geologisch simulatie model geintroduceerd voor facies simulatie velden waar de waarschijnlijkheidskaarten van facies voorkomen berekent van een ensemble in overeenstemming zijn met de voorafgaande waarschijnlijkheidskaarten. De simulatie wordt gerealiseerd op grid cel niveau van een geometrische kaart ontworpen in een unitair domein met behulp van uniform verdeelde variabelen. De lay-out van de kaart is gekozen zodat de gesimuleerde facies velden voldoen aan de geometrische en topologische voorwaarden. De uniforme variabelen gebruikt voor de facies simulatie werden verkregen door een projectie van gaussische velden met een cumulatieve distributie functie. Wij noemen het geologische simulatie model het adaptieve meervoudige gaussische simulatie model (APS). Het APS-model wordt gekoppeld aan EnKF verkregen actualisering voor de waarschijnlijkheidskaarten van facies voorkomen samen met een onzekerheid kwantificering van de geologie. De model parameters (de gaussische velden) werden geconditioneerd aan productie gegevens en globale facies verhouding voor een betere kwantificering van de onzekerheid. Het model werd getest op twee synthetische modellen met verschillende complexiteit en een daadwerkelijk veld, de uitkomsten toonde veelbelovende resultaten.

De tweede onderzoek richting concentreerde zich op gekanaliseerde reservoirs, gemodelleerd met MPS geologische simulatie modellen. Wij presenteren een nieuwe parametrisatie van facies velden die worden gekoppeld aan de iteratieve adaptieve gaussische mix (IAGM) voor geschiedenis conditionering. De parametrisering overbrugt ideeen van MPS en TPS. In de standaard TPS parametrisering zijn de facies velden de

geactualiseerde velden, het veld is echter gedeeltelijk gaussisch waardoor het niet triviaal is om gekanaliseerde facies structuren te karakteriseren door een covariante functie (of variogram) in combinatie met een door de gebruiker gedefinieerde afgeknotte kaart. De MPS benadering heeft het voordeel van willekeurige veld simulatie van een kanaal type, ze hebben een discrete waarde en het is niet triviaal om deze te verbinden met een ensemble gebaseerde inversie methode. Wij stellen daarom een combinatie van twee parametrisaties voor. Door het schatten van de waarschijnlijkheid van marginale facies, van een monster van MPS in combinatie met een trainingsafbeelding, parametriseren wij de marginale waarschijnlijkheidsveld met een (marginale) willekeurig gaussisch veld waar de afgeknotte kaart wordt gedefinieerd door de integraal van een facies type in een bepaalde grid cel in het gaussische domein samenvalt met waarschijnlijkheid van de marginale grid cel facies die geschat wordt op basis van de ensemble. Daarnaast hebben wij de parametrisering van de iteratieve adaptieve gaussische mix filter gecombineerd voor een ensemble gebaseerde Bayesian inversie. Hiervoor zijn twee belangrijke redenen. Allereerst, de IAGM is meer geschikt dan EnKF voor zeer niet lineaire problemen, ten tweede, de IAGM heeft een herbemonsteringsstap die het toelaat om een trainingsafbeelding te hermonsteren met behulp van geactualiseerde waarschijnlijkheidskaarten.

About the Author

Marius Bogdan Sebacher was born on 3th of August, 1975 in Bucharest, Romania. He received the secondary education between 1989 and 1993 at Liceum "Alexandru Ioan Cuza" in Bucharest, Romania. From 1993 to 1997 he studied Mathematics at University of Bucharest, Faculty of Mathematics and Computer Science. In 1997 he started a Master of Science Program in Algebra at the same University and obtained the M.Sc. degree in noncommutative Algebra in 1999. From 1999 he started to teach Mathematics as assistant professor at Technical University of Civil Engineering in Bucharest, Romania. From 2010 until 2014 he continued the academic education in the group of Mathematical Physics, Faculty of Electrical Engineering, Mathematics and Computer Science as a Ph.D candidate, based on a research contract financed by the TNO, The Netherlands.