Delft University of Technology
Faculty of Electrical Engineering, Mathematics and
Computer Science
Delft Institute of Applied Mathematics

Master of Science Thesis

Parameter Estimation in Reservoir Engineering Models via Data Assimilation Techniques

by

Mariya Victorovna Krymskaya

Delft, The Netherlands
July 2007

Members of the Committee

Chairperson of Graduate Committee:

Prof. Dr. Ir. A.W. Heemink

Graduate Committee:

Prof. Dr. Ir. A.W. Heemink	Faculty of Electrical Engineering, Mathematics and Computer Science, Department of Applied Mathematics, Section Mathematical Physics, Delft University of Technology, the Netherlands
Prof. Dr. Ir. J.D. Jansen	Faculty of Civil Engineering and Geosciences, Department of Geotechnology, Section Petroleum Engineering, Delft University of Technology, the Netherlands
Dr. D. Kurowicka	Faculty of Electrical Engineering, Mathematics and Computer Science, Department of Applied Mathematics, Section Operations Research and Risk Analysis, Delft University of Technology, the Netherlands
Dr. C.B.M Stroet	TNO Built Environment and Geosciences, Business Unit Geo Energy and Geo Information, TNO, the Netherlands
Dr. R.G. Hanea	TNO Built Environment and Geosciences, Business Unit Geo Energy and Geo Information, TNO, the Netherlands

Abstract

Title: **Parameter Estimation in Reservoir Engineering Models via Data Assimilation Techniques**

by

Mariya Victorovna Krymskaya

The mathematical modelling approach to the analysis of reservoir performance gains popularity throughout the years. However, the model can be used to forecast reservoir behavior only if it has been calibrated beforehand. The calibration stage, called 'history matching' in reservoir engineering context, aims at adjusting the parameters of reservoir simulation model in such a way that the computed values of observable variables at individual wells are consistent with available measurements of those quantities. As the models become more complicated and larger scaled, there increases a need of automatic history matching techniques.

The ensemble Kalman filter (EnKF) is normally applied nowadays to solving this problem. Meanwhile, EnKF has some shortcomings in reservoir engineering framework, such as heavy workload and sampling error. This fact is actually a reason to perform some further research in order to analyze specific modification of the above approach, namely iterative EnKF scheme (IEnKF).

The study has been focused on the analysis of the usage and applicability of ensemble Kalman filtering techniques to the history matching problem. There are found some practically valuable examples for which IEnKF algorithm demonstrates superior behavior. The algorithms are compared while estimating the permeability field on the basis of two-phase two-dimensional fluid flow model.

Key words: reservoir engineering, history matching, permeability, ensemble Kalman filter, iterative ensemble Kalman filter.

To my family

Contents

Acknowledgements

The two years spent as a master student within Risk and Environmental Modelling group of TU Delft provided me with invaluable professional and personal experience.

I would like to start repaying my dept of thankfulness by expressing the gratitude to Prof. R.M. Cooke and Dr. D. Kurowicka for giving me the opportunity to participate in the program.

I am grateful to Royal Dutch / Shell company which awarded me a scholarship, and a hypothetical possibility to follow a program became real. I hope I made a good use of it.

I tender many thanks to Prof. A.W. Heemink who has guided this research and follows an open door policy, such that any question can be directly answered.

Specially I must acknowledge Dr. R.G. Hanea for every day supervision, valuable advice and personal coaching.

I am grateful to Prof. J.D. Jansen for providing reservoir simulator `simsim` and ensemble of permeability fields for this study, his time explaining particular reservoir engineering matters and also sharing experience with Matlab Profile Toolbox.

Let me thank Dr. C.B.M Stroet for not only serving as a member of graduation committee and giving an opportunity to work on this topic, but also providing fruitful research conditions in Business Unit Geo Energy and Geo Information of TNO, Utrecht.

I would like to acknowledge the personnel of Delft Institute of Applied Mathematics and Business Unit Geo Energy and Geo Information of TNO, Utrecht, for their kind help and support.

I do think that the ground base level for my future was created during my study at the Ufa State Aviation Technical University (USATU), Russia.

This means that I also have to remember and acknowledge the professors and the staff of USATU in which I spent four years as a bachelor student.

Although I am really happy that I experienced these two years in TU Delft, I still remember my initial embarrassment while coming to another country and even simply moving out of my parents' house. At this point I have at least to mention the people who are responsible for the fact that I am still in a good mental condition. They are: my dear friends from Ufa — Olga, Andrey and Azamat — who still share with me all doings; my teammates, trainers and coaches from the PUNCH volleyball club, Delft, who created incredible team spirit; my colleagues within M.Sc. program, and, especially, Iwona, Lidia, Karolina and Lukasz who are always by my side keeping their PCs ready for a couple of more simulations if I need it.

Finally, I would like to express my gratitude to my parents for their unconditional and endless love. They know far too good what it means to let the child go his own way. I am sorry that my two years in Delft were not easy for you either. Thank you.

Let me just thank you all once again for never ever allowing me even to think of giving up.

List of Figures

List of Tables

Chapter 1

Introduction

Reservoir engineering is the science mostly focused on recovering the maximum amount of hydrocarbons from the field while minimizing expenses. To achieve such a goal one in particular needs an accurate prediction of the reservoir performance under different operation conditions which can be obtained by reservoir simulation.

1.1 Reservoir simulation

The use of reservoir simulation as a predictive tool gained popularity because of the possibility to take into account the factors substantially contributing to the economical risk which has to be minimized. These factors include rock properties, regional variation of fluid properties and relative permeability characteristics, and the most important physical processes taking place in reservoir system.

In turn reservoir simulation includes the following basic stages: objective definition, data analysis, deriving the model, model history matching, prediction cases running and reporting.

The scope of traditional reservoir simulation consists of analogical, experimental and mathematical approaches [7].

Analogical methods are based on predicting the performance of target reservoir involving the information on other (already developed) reservoirs with similar properties. Such approach can be used even at the stage preliminary to drilling when no measured data is available. However analogical methods provide reasonable results only if the development strategies for

target reservoir and reservoirs considered as the samples are similar. Moreover, it is impossible to investigate any scenario which differs from the scenarios known for the sample reservoir.

Experimental methods imply making direct measurements of flow properties. The experiments are run in laboratory environment and require scaling up the results to the entire reservoir. Although the behavior of experimental models reflects the performances of real reservoir, the weak point of experimental methods lies in scaling up the definite model features to the actual reservoir scale.

Nowadays, mathematical models are probably the most widely used techniques in reservoir simulation. These models describe the reservoir system via mathematical equations coupled with boundary and/or initial conditions. Since the field description is quite complex, it is not possible in general to solve the model analytically even using simplifying assumptions. This obstacle can be overcome by solving mathematical model via numerical methods and obtaining approximate results. The popularity of numerical methods is secured and supported by the progress of modern computer industry. Moreover this approach allows investigating any production scenario that crosses reservoir engineer's mind. In the current study we just concentrate our attention on the mathematical models which are solved numerically.

Whatever technique or model is applied to investigate future reservoir performance, the basic idea is to start by running the model within some time period for which an analyst already has the outcomes available. If the computed quantities match the actual data, the procedure can be accepted as correct and used further to make predictions. However, if these quantities do not match then some of the model parameters (e.g. permeability) have to be modified and the model has to be run again. The process of adapting the model parameters to match the computed reservoir outcome quantities and the real observations is called history matching [5]. The problem of history matching is of high magnitude and complexity.

1.2 History matching

History matching is an important part of any analysis related to reservoir performances. It addresses adjusting the parameters of reservoir simulation model (e.g. permeability) in such a way that the computed values of observable variables (e.g. rates, or pressures, or saturations) at individual

wells are consistent with available measurements of those quantities.

The following steps are usually performed during history matching procedure [7]:

1. set the objectives of the history matching process. The main objectives of history matching are improvement and validation of reservoir simulation model. Meanwhile if history matching is accomplished successfully, it is also possible to get some beneficial secondary objectives like better understanding of the processes occurring in reservoir;

2. determine the method to use in history match. The most commonly used history matching approaches are manual and automatic history matching;

3. settle the observable variables to be matched and the criteria to be used for determining a successful match. In relation to the given production schedule the observable variables are usually represented by (a) gas-oil ratios (GORs) and water-oil ratios (WORs); (b) average pressures (shut-in pressures) or pressures at observation wells; (c) flowing well pressures; (d) oil production rates [5];

4. determine the reservoir parameters that can be adjusted during the history match and the confidence range for these parameters. The typical parameter used for adjustment within history match is relative permeability;

5. run the simulation model with the best available input data;

6. compare the results of the simulation run with the observable variables chosen in step 3;

7. change the reservoir parameters selected in step 4 within the range of the confidence;

8. continue with steps 5–7 until the criteria established in step 3 are met.

There are distinguished manual and automatic approaches to history matching.

1.2.1 Manual history matching

Manual history matching can be considered in some sense as a kind of art. It includes running the simulation model for the historical period, comparing

the result to known field behavior and then manually adjusting the reservoir parameters to improve the match. The last step is performed by engineer and its success mostly depends on personal knowledge of the field under study and general experience.

Although there are no well-defined rules how to perform manual history matching in any case, the following tips are helpful [1]:

- change the parameters which have the largest uncertainty and also the largest influence on the solution;
- the match of average pressure is affected by fluid volumes in-place, size of the aquifer and the degree of communication between the reservoir and the aquifer. Moreover a poor match of GORs and WORs will result in poor average pressure matching;
- pressure draw-down depends on horizontal permeability and skin effects;
- GORs and WORs are primarily affected by draw-down, but also by the position of fluid contacts and the thickness of the transition zone (which in turn depends on capillary forces). The shapes of the GORs and WORs curves after breakthrough depend on the relative permeability curves but the breakthrough time depends mainly on the end points of the latter curves;
- matching breakthrough times is one of the most difficult tasks which is not often performed.

Manual history matching allows accumulating some knowledge during the history match procedure, being very time consuming though. This obstacle can be overcome by introducing an automatic approach.

1.2.2 Automatic history matching

Automatic history matching uses computer logics to adjust the reservoir parameters. This excludes human factor from the history matching process which may result in the loss of specific knowledge related to reservoir under study.

The main problem that has to be solved via automatic history matching is searching for the combination of reservoir parameters for which an error function (objective function) attains its minimum. This function represents

a difference between the observed reservoir performance and the results of simulation during the historical period [7]:

$$f_E = \sum_{i=1}^{n_{par}} \left[w_i \left(\mathbf{X}_{io} - \mathbf{X}_{is} \right)^2 \right], \tag{1.1}$$

where f_E denotes error function, n_{par} is the number of reservoir parameters, w_i - weighting coefficient, $\mathbf{X}_{io}$ and $\mathbf{X}_{is}$ correspondingly refer to observed and simulated data that have to be matched.

There are several approaches to automatic history matching which differ in the way they obtain parameter set that minimizes objective function. The choice of minimization technique is mainly based on the fact whether the error function has linear or nonlinear form.

1.2.3 Concluding remarks

Selection of history matching method depends on the study objectives, the put up company resources and the time constraints. However, no specific history matching method guarantees a successful history match. Moreover the term 'successful match' is not well-defined itself. Its meaning is up to company, particular project or individual performing the study. For instance, in general it is sufficiently realistic to consider pressures being matched if the difference between observed and computed quantities is within $\pm 10\%$ draw-down, while some specific study may require reducing the tolerance to $\pm 5\%$ or less, but not less than one percent [8].

The quality of the match is concerned with the amount of historical data available. With insufficient data the sets of reservoir parameters minimizing objective function significantly differ from those for the case of more representative data array. Even if successful match has been achieved with one bunch of historical data, there is no guarantee that further reservoir parameters adjustment is not needed as additional historical data become available.

History matching stage is considered to be extremely time consuming and usually contributes with a high rate to general study costs, being essential for practical needs of reservoir engineering though. The use of automatic history matching algorithms may resolve some difficulties related to the stage by reducing the time required for computations.

1.3 Automatic history matching strategies

The automatic history matching problem can be considered in a certain sense as a mathematical minimization problem. To solve it we need in general the algorithms suitable for minimization of objective functional (1.1).

1.3.1 Traditional automatic history matching

The history matching process has been investigated by joint efforts of reservoir engineering and mathematical community for around 40 years and is pioneered by [14]. The authors suggest dividing reservoir into a number of homogeneous blocks and computing so-called sensitivity coefficient for each resulting cell. These sensitivity coefficients are used to evaluate how the cost function changes with respect to small perturbations in the properties of particular block. The reservoir parameters are varied until a least-squares fit is obtained for observed and calculated pressures. The other early developed approach involves application of adjoint technique from the optimal control theory to the history matching problem and is presented in [4].

During the past decades a few automatic history matching techniques appeared, including e.g. representer method described in [2, 19]. The authors have assumed the model being perfect except for the errors in the model parameters. The measurements are considered as imperfect and containing some errors. All errors are assumed to be Gaussian. Then the minimizer of the cost function is obtained as a linear combination of the solution given by the model without measurements and so-called representers that depend on the measurements. For each measurement a representer is computed via the two step procedure: (i) solving backwards the adjoint equation forced by an 'impulse' correspondingly to the form of measurement, (ii) solving forward the above equation with involvement of the computed adjoint solution.

Among the recent techniques we should mention the streamline simulation approach [18]. The method exploits the fact that geometry of the streamlines represents the flow path which results from the field heterogeneities. This feature of streamlines is used by modifying reservoir parameters along the blocks mapped by streamlines.

The investigation of production problem requires the analysis of reservoirs filled in with the fluid consisting of up to three phases: oil, gas and water. Behavior of such a reservoir can be described by the multi-phase model discussed in Section 2.2.

Several researchers have analyzed history matching problem with respect to this model. For instance, one of the authors [20] computes sensitivity coefficients via adjoint method in such a way that the number of matrix problems which have to be solved is independent of the number of model parameters. Then minimization of the objective function can be performed by e.g. Gauss-Newton method exhibiting the property of approximate quadratic convergence in the neighborhood of the minimum.

However, most of these traditional history matching approaches are either limited to the small-scaled and simple reservoir models or inefficient in terms of computational costs. In general, these methods also perform the treatment of uncertainty via repeated history matching process for different initial models which results in even greater computational efforts. Moreover, traditional history matching also does not allow continuous model updating. Namely, as the new data become available for being included into the match, the whole history matching process has to be repeated using all observed data. At the same time, the amount of deployed sensors for permanent monitoring of pressure, temperature or flow rates increases. This fact yields the increase of data output frequency and rises up a problem of incorporating obtained data in the model as soon as it becomes available so that the model is always up-to-date.

1.3.2 Kalman filter in automatic history matching

The Kalman filtering techniques are known as the most popular methodology for assimilating the new measurements to continuously update the state of the system. Originally, the Kalman filter was developed for operating on the linear models, while non-linearity requires using some further modifications, e.g. the extended Kalman filter. However, when the model is highly non-linear or the scale of the space vector is too large, application of extended Kalman filter also meets difficulties. These difficulties can be overcome by applying the ensemble Kalman filtering (EnKF) algorithm based on Monte-Carlo approach.

The great majority of the problems in reservoir engineering are highly non-linear and characterized by a large number of variables, thus the idea to use EnKF in reservoir simulation seems to be natural. In particular it is presented in the publications [11, 25]. Other papers [10, 17] report the results of using EnKF approach in history matching process. They consider the application of EnKF to a PUNQ-S3 model. 'PUNQ' stands for Production forecasting with UNcertainty Quantification. The PUNQ-

S3 is a small-scaled synthetic tree-dimensional reservoir engineering model developed by a group of oil companies, research institutes and universities to compare methods for quantifying uncertainty in history matching. The model is based on a real field operated by Elf Exploration Production. Although these studies clearly show that EnKF is successful in assimilating production data to update initial reservoir model and its application allows reducing computational costs for history matching, there is still enough space for further investigation and improvement.

Specifically, the research described in [27] has shown that for some non-linear models the EnKF does not provide completely acceptable characterizations of the uncertainties. This leads to the idea of using improved EnKF modifications, namely various iterative EnKF schemes.

1.4 Research objectives

The carried project was focused on investigating ensemble Kalman filtering and in particular iterative ensemble Kalman filtering approaches to solving history matching problem for reservoir engineering. It was found out that the following intermediate objectives have to be gained in order to achieve the major goal:

- analyzing the features of the physical phenomena which influence reservoir performance;
- studying the physical origin of the mathematical model which reflects performance of reservoir with multi-phase flow;
- in-depth considering the algorithms which implement the Kalman filtering techniques in order to form a basis for their extended exploitation;
- proposing iterative modification of ensemble Kalman filtering;
- implementing the ensemble Kalman filtering techniques under investigation as a Matlab code and coupling them with forward simple reservoir simulator `simsim` provided by Prof. J.D. Jansen;
- setting up experimental framework and performing numerical experiments which allow comparing behavior of considered algorithms.

The thesis presents the results of the work devoted to the above tasks.

1.5 Thesis outline

The proposed objectives have naturally predetermined the structure of the thesis. It is organized as follows.

Chapter 2 provides the background necessary to analyze the main effects which influence reservoir performance and presents the mathematical model reflecting multi-phase fluid flow behavior.

In Chapter 3 the ensemble Kalman filtering technique which recently gained popularity as powerful tools for history matching is described, and the iterative approach to ensemble Kalman filtering is proposed in the framework of reservoir simulation.

Chapter 4 defines the experimental settings for the twin experiment, while the accomplishment of the case study is described in Chapter 5.

Finally, Chapter 6 presents relevant conclusions and recommendations.

Chapter 2

Reservoir properties and flow equations

Before performing any reservoir simulation an analyst has to specify the main effects influencing reservoir performance and develop an appropriate mathematical model of the system. The current chapter is devoted to consideration of such effects and in particular it describes the two-phase water-oil flow model. The subsequent theory is presented in a manner similar to [1, 7]. The form of the model equations essentially depends on the nature of reservoir, namely, the rock and fluids filling it.

A natural reservoir is usually heterogeneous meaning that its properties change with the space location. Such a fact has to be incorporated into the model. This causes appearance of additional problems with modelling.

In turn the nature of fluids filling the reservoir strongly depends on the stage of recovery. The following production stages can be distinguished [5]:

- *primary recovery* is the very early stage at which reservoir contains a single fluid (e.g. gas or oil) under the high pressure. The gas or oil is produced by natural decompression without any pumping effort at the wells. Usually around 70%–80% of hydrocarbons are left in the reservoir by the end of this stage;

- *secondary recovery (water flooding)* is oriented towards recovering the part of remaining oil (or gas). For that purpose a fluid (typically the water) is injected into so-called injection wells maintaining high reservoir pressure and producing oil through so-called production wells.

After this stage often 50% or more of hydrocarbons still remain in the reservoir. The water flooding although suffers from the water breakthrough event at which some injected water is produced back via production wells. The time for which at first the breakthrough occurs is known as the breakthrough time;

- *tertiary recovery (enhanced recovery)* is the stage at which complex chemical and thermal effects are involved to produce more hydrocarbons. These effects could be achieved by e.g. injecting into reservoir some materials which are not present there under normal conditions.

The water flooding stage is very common in practical recovery process and in particular an object of the current research. Consideration of oil and water features allows assuming fluids being immiscible, so there is no mass transform between phases. The phase which wets the porous medium stronger than the other one is said to be wetting phase and denoted by subscript w. The other phase is called non-wetting and labelled by subscript o. In the current case water and oil are wetting and non-wetting phases respectively.

The reservoir rock properties, the nature of fluids filling reservoir and the effects appearing within rock-fluid interaction strongly influence multiphase flow in porous medium. Let us present some considerable features of these phenomena.

2.1 Reservoir properties

The current section gives an insight into the main physical properties of the reservoir that have to be taken into account while developing an appropriate fluid flow model.

2.1.1 Rock properties

This part introduces basic reservoir-rock notions such as porosity and permeability which are assumed to be independent of fluid filling in the reservoir.

Porosity. The fluid in reservoir is contained in the pore spaces of the rock. Some of these pores are interconnected and some are not. The ratio of the pore space in the rock volume sample to the total rock volume sample

is called porosity. This quantity is usually denoted by ϕ. Depending on whether isolated pores are included or not into the computation of pore volume in the sample, one can distinguish total and effective porosity. Since only interconnected pores produce fluids we shall focus later on the effective porosity which for simplicity will be called 'porosity'.

Permeability is the capacity of a porous medium to transmit fluids through its interconnected points. We talk about absolute permeability if the porous medium is 100% saturated with a single phase. Permeability property depends on the point in space and even the fluid direction. In many practical processes it is reasonable to represent permeability as a tensor $\mathbf{k} = diag(k_x, k_y, k_z)$, where x, y and z are three principle directions. However, it is often possible to assume that horizontal permeability $k_H = k_x = k_y$, while being still different from vertical permeability $k_V = k_z$. If $k = k_x = k_y = k_z$ then porous medium shows isotropic property.

Rock compressibility is defined as [5]

$$c_R = \frac{1}{\phi}\frac{\partial \phi}{\partial p}, \tag{2.1}$$

where p denotes the pressure. After integration we obtain

$$\phi = \phi^0 e^{c_R(p-p^0)},$$

where ϕ^0 is the porosity at a reference pressure p^0. The use of approximation brings us to the expression

$$\phi \approx \phi^0(1 + c_R(p - p^0)).$$

Concluding remarks. Rock properties of the reservoir typically vary with location in space (e.g. porosity). If the property is constant and independent of location, then the porous medium is called homogeneous. In reality it is necessary to deal with heterogeneous reservoirs in which property varies in space.

Some other reservoir parameters (e.g. permeability) may suffer from directional dependency. If the property does not depend on the direction in which it is measured, the reservoir has isotropic property distribution. In the opposite case reservoir is said to be anisotropic in relation to that property.

2.1.2 Fluid properties

The current subsection addresses fluid properties that are of interest for reservoir modelers. The list of such properties includes fluid compressibility factors, fluid formation volume factors (FVF), fluid densities and fluid viscosities.

Fluid compressibility is defined as the relative volumetric change of a given mass to pressure change at constant temperature:

$$c_\alpha = -\frac{1}{V_\alpha}\frac{\partial V_\alpha}{\partial p_\alpha}\bigg|_{T^0},$$

where α is either w or o with respect to the phase, V represents the volume and T^0 is a reference temperature. Equivalently fluid compressibility may be expressed as

$$c_\alpha = \frac{1}{\rho_\alpha}\frac{\partial \rho_\alpha}{\partial p_\alpha}\bigg|_{T^0}, \tag{2.2}$$

where ρ_α is the density of the phase α. Note that under isothermal conditions

$$c_w = \frac{1}{\rho_w}\frac{\partial \rho_w}{\partial p_w}\bigg|_{T^0} \approx \frac{1}{\rho_w}\frac{\partial \rho_w}{\partial p_o}\bigg|_{T^0}. \tag{2.3}$$

FVF. A fixed mass of a fluid filling reservoir occupies different volumes at different reservoir pressures. The FVF is introduced to convert volumes at reservoir pressure and temperature to their equivalent volumes under standard conditions. The phase FVF is the ratio of the volume occupied by phase at reservoir pressure and temperature to the corresponding volume under standard conditions. For a single phase FVF may be written in terms of densities:

$$B_\alpha = \frac{\rho_{\alpha sc}}{\rho_\alpha},$$

where $\rho_{\alpha sc}$ denotes the density of phase α under standard conditions. In reservoir engineering standard conditions are usually defined as $p_{sc} = 100(kPa)$ and $T_{sc} = 15(^\circ C)$.

Fluid density is also pressure dependent in reservoirs. For a single-phase flow the density can be obtained via

$$\rho_\alpha = \frac{\rho_{\alpha sc}}{B_\alpha}.$$

Fluid viscosity measures how easy the fluid flows as pressure gradient is applied. Fluid viscosity is labelled by μ. Dilute fluid demonstrates low resistance to flow because fluid molecules are far from each other and perform random motions. On the other hand, fluid molecules in a dense fluid are close to each other and their random motion retards flow. Fluid viscosity is influenced by both pressure and temperature. The behavior of fluid viscosity and density are related, since the density represents the measure of the mean free path of molecules, hence, the measure of random molecular motions which affect viscosity.

2.1.3 Fluid-rock properties

This section presents conceptions of fluid saturations, capillary pressure, relative permeability and Corey-type two-phase relative permeability model.

Fluid saturation is a fraction of the pore space that is occupied by the above fluid. If fluid saturations of water and oil are denoted by S_w and S_o respectively then the following equality holds for two-phase flow:

$$S_w + S_o = 1. \tag{2.4}$$

Capillary pressure occurs across an interface between any two immiscible flows and is a consequence of the interfacial tension. In a two-phase system capillary pressure is the difference between pressures of the non-wetting and wetting phases:

$$p_c = p_o - p_w = p_c(S_w). \tag{2.5}$$

Capillary pressure is a function of saturation for the given reservoir rock and fluids at a constant temperature.

Relative permeability. If two or more phases are saturating the porous medium the reservoir capacity to transmit any particular phase α is called the effective permeability to that phase and denoted by $\mathbf{k}_\alpha$. This definition yields that components of the effective permeability vector $\mathbf{k}_\alpha$ are not greater than those of the absolute permeability $\mathbf{k}$ of porous medium. Relative permeability $k_{r\alpha}$ to phase α indicates the tendency of phase α to wet the porous medium:

$$\mathbf{k}_\alpha = k_{r\alpha}\mathbf{k}.$$

Typical relative permeability curves for an oil-water system with water displacing oil are presented in Figure 2.1.

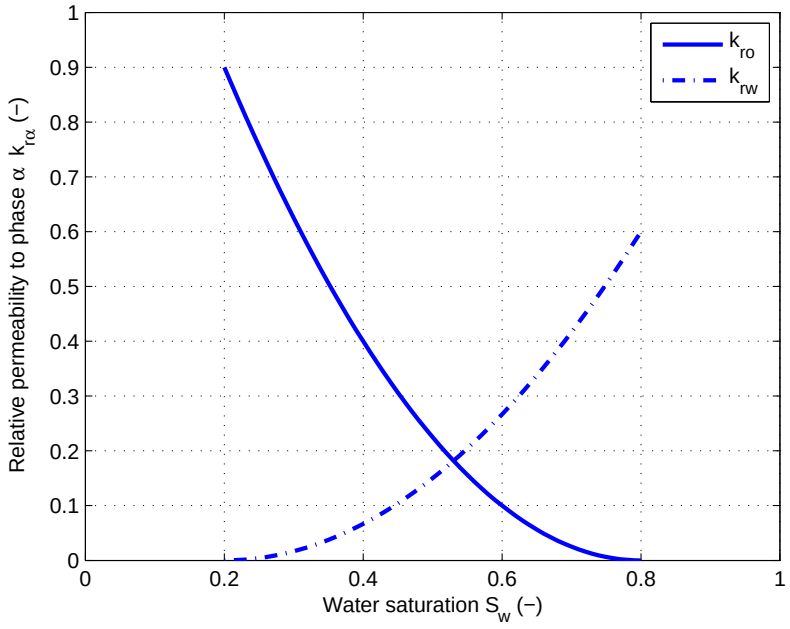


Figure 2.1: Typical relative permeability curves

The value of S_w at which water starts to flow is termed as connate water saturation S_{wc} and the value at which oil starts to flow — as residual oil saturation S_{or}. These notions are reversed for a drainage cycle.

Considering the relative permeability to phase α at critical saturation of the other phase we end up with the notion of the end point permeability $k^0_{r\alpha}$.

Corey-type two-phase relative permeability model. Although the values of relative permeability for each particular porous medium are subjects to the empirical or experimental determination, there exist mathematical models appropriate for describing a relationship between relative permeability and saturation of the wetting phase. We are going to discuss the Corey-type model.

Define the normalized saturation value as

$$S = \frac{S_w - S_{wc}}{1 - S_{wc} - S_{or}},$$

then Corey-type approximations of relative permeability for water and oil are

$$k_{rw} = k_{rw}^0 S^{n_w}, \; k_{ro} = k_{ro}^0 (1-S)^{n_o}, \tag{2.6}$$

where n_w and n_o are Corey exponents. In general these quantities can be obtained from measured data, but simplification $n_w = n_o = 2$ is found being appropriate.

The above discussion gives us basic ideas of some phenomena which influence reservoir performance and, hence, have to be incorporated into fluid flow model.

2.2 Two-phase fluid flow model

Consider a two-phase water-oil fluid under isothermal conditions. The current subsection is devoted to the derivation of governing partial differential equations (PDE's) for such a flow. This system of equations includes:

- mass balance equations;
- Darcy's law;
- capillary pressure equation (given by (2.5));
- relative permeability equations (given by Corey-type model (2.6));
- equations of state (given by (2.2) and (2.3)).

Mass balance equation. If the flow process is isothermal and no chemical reactions occur, then the basic idea of the equation representing conservation of mass per unit time for a unit control volume under standard conditions is illustrated by the following diagram:

mass in	-	mass out	-	mass accumulated	+	source term	=0

Figure 2.2: Mass balance per unit time for a unit control volume

In the case of two-phase flow the mass balance equations can be expressed for each phase as

$$\nabla \cdot \left(A \frac{\rho_\alpha}{B_\alpha} \mathbf{v}_\alpha \right) + A \frac{\partial \left(\frac{\rho_\alpha}{B_\alpha} \phi S_\alpha \right)}{\partial t} - A \rho_\alpha \bar{q}_\alpha = 0, \tag{2.7}$$

where A is geometrical factor, $\mathbf{v}_\alpha$ denotes flow rate of phase α per unit cross-sectional area perpendicular to flow direction (superficial velocity), $\bar{q}_w$ and $\bar{q}_o$ are the source terms for water and oil phase respectively. Source terms are currently expressed as flow rates per unit volume, and their positive values correspond to injection, while negative — to production.

Darcy's law. In addition to the mass conservation equation we need a relationship between the flow rate and pressure gradient for each phase. In the case of single-phase flow such a relationship was discovered by Darcy.

Differential form of Darcy's law for the simultaneous flow of more than one phase is

$$\mathbf{v}_\alpha = -\frac{k_{r\alpha}}{\mu_\alpha B_\alpha}\mathbf{k}\left(\nabla p_\alpha - \rho_\alpha g \nabla d\right), \tag{2.8}$$

where g is acceleration of gravity and $d = d(x, y, z)$ represents the depth. We can rewrite equation (2.8) in the form of

$$\mathbf{v}_\alpha = -\lambda_\alpha\left(\nabla p_\alpha - \rho_\alpha g \nabla d\right), \tag{2.9}$$

where $\lambda_\alpha = \dfrac{k_{r\alpha}}{\mu_\alpha B_\alpha}\mathbf{k}$ is phase mobility tensor.

To complete the mathematical model we have to determine appropriate boundary and initial conditions.

Initial conditions in reservoir models normally assign phase pressures and saturations for every grid block at the beginning of simulation.

Boundary conditions. The reservoir can be treated as an open system with external (i.e. defining reservoir limits) and internal (i.e. wells) boundaries. For such a system, flow in and out of it takes place only at the boundaries. The flow at the boundaries may be described by the following types of conditions [7]:

- *no-flow* — phase transmissibilities across the boundary are set to zero. This condition appears when the external reservoir boundary is sealed to flow;

- *constant-pressure* — there is no change of pressure while crossing the boundary. This condition arises if the rate of fluids withdrawn on one side of the boundary is equal to the rate of fluids being supplied or injected on the other side of the same boundary;

- *specified-efflux* — the boundary is replaced by a now-flow boundary condition and a fake production well. This condition becomes necessary when there exists a communication between target reservoir and some another one trough the boundary;

- *specified-influx* — the boundary is replaced by a now-flow boundary condition and a fake injection well. This condition pops up when there exists a communication between target reservoir and an aquifer supplying water influx through the boundary.

Finally, the presence of wells has to be incorporated into the model. For multi-phase flow this can be done by including bottom hole pressure, oil production/injection rate and/or water inhection/production rate.

Chapter 3

Kalman filtering

Kalman filtering is a powerful technique designed for solving data assimilation problems. Being named after Rudolph E. Kalman, who formulated this approach in his famous paper of 1960, it was taken further by a number of scientists and resulted nowadays in various modifications. This section of the report presents general idea of Kalman filtering in a manner similar to [21]. Let us restrict ourselves to the case of linear system that can be described by the following two equations:

$$\mathbf{X}_{k+1} = \mathbf{F}_k\mathbf{X}_k + \mathbf{B}_k\mathbf{U}_k + \mathbf{G}_k\mathbf{W}_k, \tag{3.1}$$

$$\mathbf{Z}_k = \mathbf{M}_k\mathbf{X}_k + \mathbf{V}_k, \tag{3.2}$$

where $\mathbf{F}_k, \mathbf{B}_k, \mathbf{G}_k, \mathbf{M}_k$ are matrices, k is the time index, $\mathbf{X}_k$ denotes the state of the system, $\mathbf{U}_k$ is a system input, $\mathbf{Z}_k$ is the vector of measurements, $\mathbf{W}_k$ is Gaussian white system noise process with zero mean and covariance matrix $\mathbf{Q}_k$, $\mathbf{V}_k$ is Gaussian white measurement noise process with zero mean and covariance matrix $\mathbf{R}_k$. Moreover processes $\mathbf{W}_k$ and $\mathbf{V}_k$ are assumed to be independent.

Vector $\mathbf{X}_k$ which contains information on the current system state cannot be directly observed. However it is possible to measure $\mathbf{Z}_k$ which is some function of $\mathbf{X}_k$ affected by noise process $\mathbf{V}_k$. The idea is to use the available measurements $\mathbf{Z}_k$ to estimate the state of the system $\mathbf{X}_k$. Naturally, we would like to have an accurate estimator of the true state. This desire leads to the requirements given below:

- expected value of the estimate has to be equal to the expected value

of the state;

- estimator should have the smallest possible error variance.

As it has been proven, Kalman filter is just the estimator satisfying the formulated criteria. We proceed by introducing the original Kalman filter algorithm.

3.1 The Kalman filter

To solve filtering problem (3.1)–(3.2) we have to determine the probability density of the state $\mathbf{X}_k$ conditioned on the history of available measurements $\mathbf{Z}_1, \ldots, \mathbf{Z}_l$. It turns out that this conditional density function is Gaussian, hence, it can be characterized by mean an covariance matrix. Moreover, for Gaussian distribution the mean is an optimal estimate of the state in a minimum variance (or least square) sense and also in maximum likelihood sense. Equations to obtain the mean $\mathbf{X}(k|l)$ and covariance matrix $\mathbf{P}(k|l)$ of probability density of state $\mathbf{X}_k$ at time k conditioned on the history of the measurements $\mathbf{Z}_1, \ldots, \mathbf{Z}_l$ can be formulated as follows [13]:

- Initial condition:

$$\mathbf{X}(0|0) = \mathbf{X}_0, \tag{3.3}$$

$$\mathbf{P}(0|0) = \mathbf{P}_0. \tag{3.4}$$

- Time update:

$$\mathbf{X}(k|k-1) = \mathbf{F}_k\mathbf{X}(k-1|k-1) + \mathbf{B}_k\mathbf{U}_k, \tag{3.5}$$

$$\mathbf{P}(k|k-1) = \mathbf{F}_k\mathbf{P}(k-1|k-1)\mathbf{F}_k^{\mathrm{T}} + \mathbf{G}_k\mathbf{Q}_k\mathbf{G}_k^{\mathrm{T}}. \tag{3.6}$$

- Measurement update:

$$\mathbf{X}(k|k) = \mathbf{X}(k|k-1) + \mathbf{K}(k)(\mathbf{Z}_k - \mathbf{M}_k\mathbf{X}(k|k-1)), \tag{3.7}$$

$$\mathbf{P}(k|k) = (\mathbf{I} - \mathbf{K}(k)\mathbf{M}_k)\mathbf{P}(k|k-1), \tag{3.8}$$

where $\mathbf{I}$ is identity matrix of the appropriate size, $\mathbf{K}(k)$ is the Kalman gain having a form of $\mathbf{K}(k) = \mathbf{P}(k|k-1)\mathbf{M}_k^{\mathrm{T}}(\mathbf{M}_k\mathbf{P}(k|k-1)\mathbf{M}_k^{\mathrm{T}} + \mathbf{R}_k)^{-1}$.

The Kalman filter is in fact a sequential filter method, because the model is integrated forward in time via expressions (3.5)–(3.6) and as soon as

measurements are available they are used to re-initialize the model through (3.7)–(3.8) before the next integration step. The algorithm flowchart is presented below.

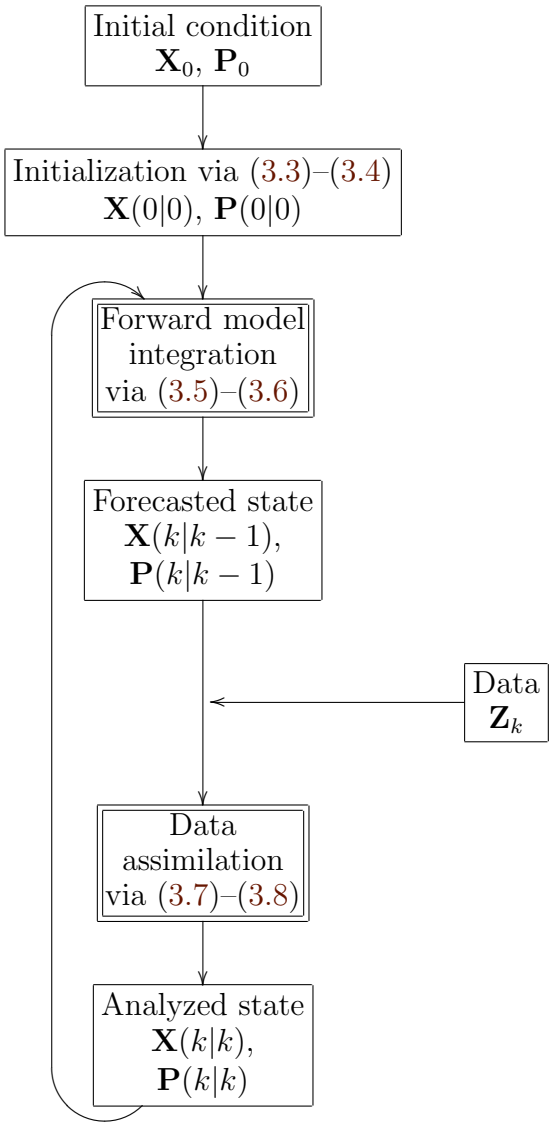


Figure 3.1: Kalman filter algorithm flowchart

Term $\mathbf{Z}_k - \mathbf{M}_k\mathbf{X}(k|k-1)$ on the right hand side of expression (3.7) is called innovation $\mathbf{I}_k$ of the filter and is used to determine whether or not the filter is performing well. If system equations (3.3)–(3.8) are correct (which happens only in ideal case of perfect stochastic description of the system)

then statistics of innovations are

$$E[\mathbf{I}_k] = 0,$$

$$E[\mathbf{I}_k \mathbf{I}_k^{\mathrm{T}}] = \mathbf{M}_k \mathbf{P}(k|k-1)\mathbf{M}_k^{\mathrm{T}} + \mathbf{R}_k.$$

Since these theoretical statistics are known they can be compared with the statistics of observed innovations. If the realizations of the innovations are consistent with theoretical statistics the filter is considered to be performing well, otherwise the filter is said to be divergent.

Let us note that

$$\lim_{\mathbf{R}_k \to 0} \mathbf{K}(k) = \mathbf{M}_k^{-1} \tag{3.9}$$

and

$$\lim_{\mathbf{P}(k|k-1) \to 0} \mathbf{K}(k) = 0. \tag{3.10}$$

It follows from the limit (3.9) that as the measurement covariance approaches zero (hence, measurements become more accurate) the Kalman gain puts more weight on the innovation within measurement update (3.7). Thus, actual measurements $\mathbf{Z}_k$ become more reliable and predicted measurements $\mathbf{M}_k \mathbf{X}(k|k-1)$ are trusted less and less. In turn, as covariance matrix $\mathbf{P}(k|k-1)$ approaches zero, the Kalman gain puts less weight on the innovation, thus predicted estimate $\mathbf{X}(k|k-1)$ is trusted more and more.

Unfortunately, applying the standard Kalman filter algorithm to large-scale systems implies a big computational effort needed. The most time consuming part is calculation of the term $\mathbf{F}_k \mathbf{P}(k-1|k-1)\mathbf{F}_k^{\mathrm{T}}$ in covariance matrix update expression (3.6). There are various methods which allow avoiding this computation, in particular the ensemble Kalman filter.

3.1.1 Parameter estimation with Kalman filter

Originally the Kalman filter was developed to update only dynamic state variables. However it turns out that parameter estimation via Kalman filtering algorithms is also possible. This can be done by constructing the following state vector:

$$\mathbf{X} = \begin{bmatrix} \mathbf{m} \\ \mathbf{Y} \end{bmatrix},$$

where $\mathbf{Y}$ consists of dynamic variables changing with time and $\mathbf{m}$ is a vector of static model parameters which are constant in time and have to be estimated.

Now the Kalman filter analysis is performed on augmented state vector. The forward step of the algorithm results in updating only the dynamic variables with time and conserving the values of static parameters. However, at the assimilation step the variables of both types are simultaneously updated providing corrected estimations of the state vector and, hence, model parameters.

3.2 The ensemble Kalman filter

The EnKF has been examined and applied in a number of studies since it was first introduced by Geir Evensen in 1994. This filtering approach is relatively easy to implement. It does not require any tangent linear operator in the case of non-linear model (in contrary to extended Kalman filter technique which was derived especially for non-linear cases) and has affordable computational cost.

The EnKF is based on a representation of the probability density of the state estimate by a finite number N of randomly generated system states $\xi_i,\ i = 1, \ldots, N$. The ensemble Kalman filter algorithm [13] is presented below:

- Initialization:

$$\xi_i(0|0) \sim \mathcal{N}(\mathbf{X}_0, \mathbf{P}_0),\ i = 1, \ldots, N. \tag{3.11}$$

- Time update:

$$\xi_i(k|k-1) = \mathbf{F}_k \xi_i(k-1|k-1) + \mathbf{B}_k \mathbf{U}_k + \mathbf{G}_k \mathbf{W}_k^i, \quad i = 1, \ldots, N, \tag{3.12}$$

$$\mathbf{X}(k|k-1) = \frac{1}{N} \sum_{i=1}^{N} \xi_i(k|k-1), \tag{3.13}$$

$$\mathbf{L}(k|k-1) = [\xi_1(k|k-1) - \mathbf{X}(k|k-1), \ldots, \xi_N(k|k-1) - \mathbf{X}(k|k-1)]^{\mathrm{T}}, \tag{3.14}$$

where $\mathbf{L}(k|k-1)$ defines an approximation of the covariance matrix

$\mathbf{P}(k|k-1)$ with rank N:

$$\mathbf{P}(k|k-1) = \frac{1}{N-1}\mathbf{L}(k|k-1)\mathbf{L}(k|k-1)^{\mathrm{T}}. \tag{3.15}$$

- Measurement update:

$$\mathbf{K}(k) = \frac{1}{N-1}\mathbf{L}(k|k-1)\mathbf{L}(k|k-1)^{\mathrm{T}}\mathbf{M}_k^{\mathrm{T}} * \left(\frac{1}{N-1}\mathbf{M}_k\mathbf{L}(k|k-1)\mathbf{L}(k|k-1)^{\mathrm{T}}\mathbf{M}_k^{\mathrm{T}} + \mathbf{R}_k\right)^{-1}, \tag{3.16}$$

$$\xi_i(k|k) = \xi_i(k|k-1) + \mathbf{K}(k)\left(\mathbf{Z}_k - \mathbf{M}_k\xi_i(k|k-1) + \mathbf{V}_k^i\right), \quad i = 1, \ldots, N, \tag{3.17}$$

$$\mathbf{X}(k|k) = \frac{1}{N}\sum_{i=1}^{N}\xi_i(k|k). \tag{3.18}$$

The algorithm (3.11)–(3.18) can be visualized by the flowchart presented in Figure 3.2, where the shadowed blocks represent the ensemble at different stages of the procedure.

Note that (3.17) involves generating additional noise $\mathbf{V}_k^i$ with the assumed statistics of the observation errors while constructing the measurement set corresponding to the ensemble. The perturbed measurements are necessary due to the fact that the absence of perturbation leads to the updated ensemble which has too low variance and causes the divergence of the algorithm [3].

Actually, the forward model integration step within reservoir engineering framework can be performed by making a forward run of the reservoir simulator. As only the outputs of the simulator are needed to proceed with EnKF algorithm, the simulator itself can be developed by another scientific group and used as a black box in EnKF analysis. This feature allows separating tasks between the researchers participating in the project, therefore optimizing the work organization.

The data assimilation procedure via EnKF technique is suitable for parallel programming, since time and measurement update can be performed separately on each ensemble member and run on different processors. The communication between the parallel processes is only needed while computing the state covariance matrix and Kalman gain.

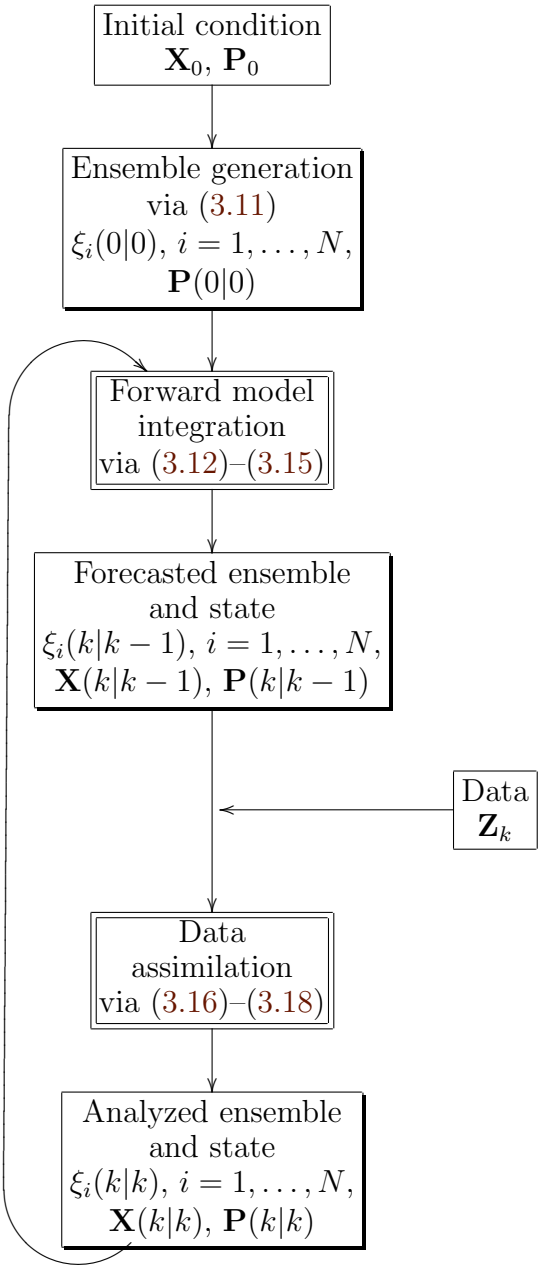


Figure 3.2: Ensemble Kalman filter algorithm flowchart

Despite EnKF has a lot of advantages which are already mentioned, it also faces an important practical problem, namely, standard deviation of the errors in the state estimate converges very slowly with the number of ensembles [13]. This makes the ensemble Kalman filter quite sensitive to the number of ensemble members used for simulation. The main workload is provided by performing time update for each ensemble member at each

time step. Hence, the speed of the EnKF algorithm is highly dependent on the model complexity and efficiency of reservoir simulator.

The model describing multi-phase fluid flow in reservoir is highly non-linear and the number of variables included into state space vector is very large, normally at least two per grid block. Although EnKF performs fairly good for this kind of problems, it sometimes fails to provide appropriate characterization of uncertainty. An example is given in [27] in relation to the case when the conditional pdf for reservoir model is multi-modal. Such phenomenon results from the fact that model non-linearity destroys the normality of a prior and a posterior distributions within Kalman filtering analysis.

We are going to consider the history matching via EnKF algorithm as the starting point for further investigations. We continue with some introduction into alternative EnKF techniques.

3.3 Iterative Kalman filtering

The current section presents the ideas of Kalman filtering algorithms that in our opinion can be alternatively applied to solving history matching problem.

Iterative forms of the Kalman filter are not completely new within the scope of reservoir engineering applications. These methods aim at obtaining any ensemble which provides improving the representation of the state distribution. There exist several approaches in petroleum engineering literature: the ad-hoc confirming EnKF method proposed by [25], iterative EnKF method recently analyzed by [26] from optimization point of view instead of Monte Carlo sampling methodology and ensemble randomized maximum likelihood filter developed by [12].

We would like to exploit the idea of iterating the filter globally [16]. Although it was originally suggested to iterate the extended Kalman filter, we modify the approach for the case of parameter estimation via EnKF technique.

The algorithm looks as follows. Incorporating all available data via EnKF starting with $\mathbf{X}(t_0|t_0)$ and $\mathbf{P}(t_0|t_0)$, we obtain the estimated values of $\mathbf{X}(t_{end}|t_{end})$ and $\mathbf{P}(t_{end}|t_{end})$, where t_0 and t_{end} denote respectively the starting and the end time point of data assimilation period. If the number of available measurements is sufficiently large we can expect that the estimated model parameter value $\mathbf{m}(t_{end}|t_{end})$ is closer to the 'true' one than

initial $\mathbf{m}(t_0|t_0)$. The estimated model parameter $\mathbf{m}(t_{end}|t_{end})$ replaces now $\mathbf{m}(t_0|t_0)$ and becomes a new initial guess for the next global iteration, which is done by rerunning the EnKF based on the same bunch of observations. Afterwards, this procedure can be repeated until no sufficient change in estimated model parameter is obtained. Note that when rerunning the filter we change only the mean estimator of initial guess about the model parameter and not the statistics $\mathbf{Y}(t_0|t_0)$ and $\mathbf{P}(t_0|t_0)$.

The flowchart of such an iterative EnKF (IEnKF) is presented in Figure 3.3, where the dashed blocks correspond to the steps which actually are the parts of EnKF algorithm.

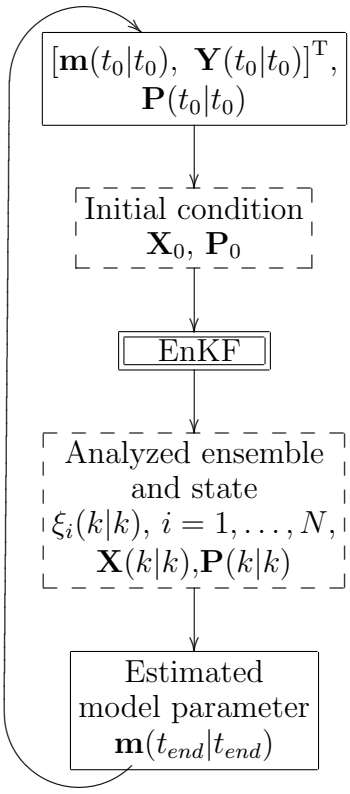


Figure 3.3: Iterative ensemble Kalman filter algorithm flowchart

Although the more educated choice of initial guess naturally should result in better estimation, there is no guarantee that iteration will converge. Thus, the plan is to investigate the features of the above IEnKF technique and to check whether it indeed allows improving the state vector estimations and the forecasts. The next chapter outlines the settings of the experiment used to test EnKF and IEnKF performances.

Chapter 4

Experimental environment

This chapter defines the experimental settings and characteristics to be used while approaching case study aimed at comparing the performance of algorithms considered earlier (Chapter 3).

The study is accomplished on the basis of two-dimensional two-phase fluid flow model described in detail in Appendix A. The model implementation is provided by the forward reservoir simulator '`simsim`' developed by Prof. J.D. Jansen. The above simulator is used as a black box to perform the time update in filtering algorithm.

4.1 Domain and process

The model is applied to a two-dimensional squared petroleum reservoir with a size of $700(m) \times 700(m)$ equipped with uniform cartesian grid consisting of 21 grid cells in each direction. The reservoir is taken to be $2(m)$ height, however, we assume that all quantities are vertically homogeneous, which allows considering fluid flow processes only in two dimensions.

We consider the water flooding stage of recovery process which is performed through the exploitation of the injection well located at the center of reservoir and four production wells established at the corners of the field (see Figure 4.1)

The injection well is constraint by prescribed injection rate of $0.002(m^3/s)$ and production wells — by bottom hole pressure of $2.5 \times 10^7(Pa)$. These parameters together with the others like connate water saturation, residual water saturation, etc. are taken the same as specified in the input file

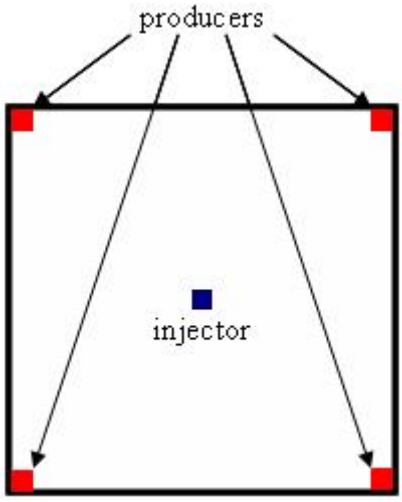


Figure 4.1: Wells locations

of `simsim` simulator corresponding to the case of heterogeneous reservoir operated under described recovery strategy. The list of input parameters required for `simsim` simulator is given in Appendix B.

4.2 State space representation

To analyze and compare filtering methods for history matching problem, we use the model (A.21):

$$\widehat{\mathbf{E}}(\mathbf{X})\dot{\mathbf{X}} - \widehat{\mathbf{A}}(\mathbf{X})\mathbf{X} - \widehat{\mathbf{B}}(\mathbf{X})\mathbf{U} = \mathbf{0}, \tag{4.1}$$

where $\widehat{\mathbf{E}}(\mathbf{X})$, $\widehat{\mathbf{A}}(\mathbf{X})$ and $\widehat{\mathbf{B}}(\mathbf{X})$ are respectively accumulation, system and input matrices in generalized state space form, $\mathbf{U}$ is an input vector containing constraints at the wells, $\mathbf{X}$ is a state vector formed as a vertical concatenation of vectors which consist of pressures $\mathbf{p}$ and water saturations $\mathbf{S}$ corresponding to each of the grid blocks:

$$\mathbf{X} = \begin{bmatrix} \mathbf{p} \\ \mathbf{S} \end{bmatrix}.$$

In order to obtain the state vector at each time moment t_k we have to find a solution of the above equation. For that purpose we use in-house reservoir simulator `simsim` which implements particular method of implicit

Euler integration with Newton iteration for solving (4.1).

Denote the output of the forward simulator run from time moment t_{k-1} up to time t_k by $[\mathbf{p}_k,\ \mathbf{S}_k]^{\mathrm{T}} = \mathbf{f}(\mathbf{p}_{k-1}, \mathbf{S}_{k-1})$.

While operating on a field, one can measure the following parameters at the wells: bottom hole pressures, oil and water flow rates, pressures and saturations. On the other hand, these quantities can be also estimated via mathematical modelling. As soon as the output of the simulator run is available, we may forecast the values of observable variables. The state vector $\mathbf{X}$ already contains estimations of well pressures and saturations. Let us consider the observable bottom hole pressures, oil and water flow rates at the wells as the components of the vector

$$\mathbf{Y} = \begin{bmatrix} \mathbf{p}_{well} \\ \mathbf{q}_{well,o} \\ \mathbf{q}_{well,w} \end{bmatrix}.$$

In fact predicted measurements $\mathbf{Y}_k$ can be seen as some function of the current field pressures and saturations $\mathbf{Y}_k = \mathbf{g}(\mathbf{p}_k, \mathbf{S}_k)$, hence, a function of the previous state $\mathbf{Y}_k = \mathbf{g}(\mathbf{f}(\mathbf{p}_{k-1}, \mathbf{S}_{k-1}))$. This kind of information can be also extracted from the forward simulator run which leads to the idea of augmenting state vector with the vector of observable variables. Note that we are interested only in bottom hole pressures and flow rates which are not a priori determined by the given constraints, hence, in bottom hole pressure at injection well and flow rates at production wells.

To perform parameter estimation we have to include the parameter of interest into the space vector. The study is focused on estimating permeability field. It turns out that the normal logarithm of permeability is normally distributed, hence, we would like to augment the state vector by the log-permeability.

Finally, the mega state space vector takes the following form:

$$\mathbf{X} = \begin{bmatrix} \log \mathbf{k} \\ \mathbf{p} \\ \mathbf{S} \\ \mathbf{p}_{well} \\ \mathbf{q}_{well,o} \\ \mathbf{q}_{well,w} \end{bmatrix},$$

where log-permeability $\log \mathbf{k}$ is a vector which consists of log-permeability values corresponding to each of the grid cells. The model parameter is

considered as static, i.e. time invariant. Meanwhile, the value of the static parameter is corrected within the data assimilation step.

The state space representation of the model under study can be now written in the following form

$$\mathbf{X}_k = \mathbf{F}(\mathbf{X}_{k-1}), \tag{4.2}$$

where $\mathbf{F}$ is the operator of reservoir simulator.

This equation has to be solved at the time update stage of EnKF algorithm. In the current project solving (4.2) is accomplished by `simsim` simulator. Note that the state vector consists of 441 permeability values, 441 pressure values, 441 water saturation values, 1 observed bottom hole pressure at the injection well, 4 observed oil flow rates and 4 observed water flow rates at the production wells or simply $\mathbf{X} \in \mathbf{R}^{1332}$.

It is also important to add a relation between the model variables and the measurements, which for our system reads as

$$\mathbf{Z}_k = \mathbf{M}_k \mathbf{X}_k, \tag{4.3}$$

where the measurement matrix $\mathbf{M}_k$ actually does not depend on time and is a block matrix $\mathbf{M} \in \mathbf{R}^{19\times 1332}$:

$$\mathbf{M} = \begin{bmatrix} \mathbf{0} & \mathbf{M}^1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{M}^1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{M}^2 \end{bmatrix}$$

with blocks $\mathbf{M}^1 \in \mathbf{R}^{5\times 441}$ and $\mathbf{M}^2 \in \mathbf{R}^{9\times 9}$ of the following form:

- elements of matrix $\mathbf{M}^1$ corresponding to the observations at the well grid blocks (i.e. elements indexed as $(1,1)$, $(2,21)$, $(3,221)$, $(4,421)$ and $(5,421)$ are set to one, the rest of the matrix if filled in with zeros);
- $\mathbf{M}^2$ is in fact an identity matrix.

Let us note that obtained system (4.2)–(4.3) is non-linear, although the measurement equation (4.3) has a linear form. However, the actual relation between the model state vector and observable variables is non-linear. The derived notation only shifts the source of non-linearity and does not vanish its effects.

We consider the model (4.2) as being perfect which might seem to be not very realistic. However, such an assumption specifies better environment for

investigating particular iterative EnKF method. We expect that in the case when the ensemble spread is not influenced by model noise, the iterative techniques have to demonstrate their specific features.

On the contrary, the values of observable variables are assumed to be imprecise. The inaccuracy of the measurements can be reflected by inserting into equation (4.3) a stochastic process $\mathbf{V}_k \in \mathbf{R}^{19}$ which is Gaussian white observation noise process with zero mean and time invariant covariance matrix $\mathbf{R}$. Hence relation (4.3) becomes

$$\mathbf{Z}_k = \mathbf{M}\mathbf{X}_k + \mathbf{V}_k. \tag{4.4}$$

Finally, equations (4.2) – (4.4) represent the model in a state space form.

4.3 Synthetic measurements generation

To test the performance of EnKF algorithms we are going to do a so-called 'twin experiment'. It requires that the 'true' values of observable variables are generated synthetically by a preliminary run of the model itself and the noisy observations are then created by permutating the 'true' values with the measurement error noise. This procedure ensures that the model is indeed able to match the data. Thereafter the synthetic data is used in the assimilation experiments.

The implementation of `simsim` simulator provides the 'true' permeability field which originates from the training image of meandering channels. The training image was created with the `snesim` algorithm based on the use of single normal equation [24] (see Figure 4.2).

Afterwards, the proper orthogonal decomposition was applied to the training image which resulted in multiple realizations of the permeability field with the identical statistics. Finally, one of the realizations was taken as a 'true' permeability field.

We were equipped with 1000 ensemble members and followed the given choice of the 'true' realization (see Figure 4.3).

Now it is possible to generate synthetic data initializing the simulator with 'true' permeability field, grid block pressures $p = 3 * 10^7(Pa)$ and water saturations $S_w = 0.2$. Figure 4.4 illustrates generated reference data consisting of pressures and water saturations at each grid cell, bottom hole pressure at injection well and flow rates at production wells. The legends establish a correspondence between the well which is located in (x, y) grid block of the field and the line coloring representation of a quantity at this

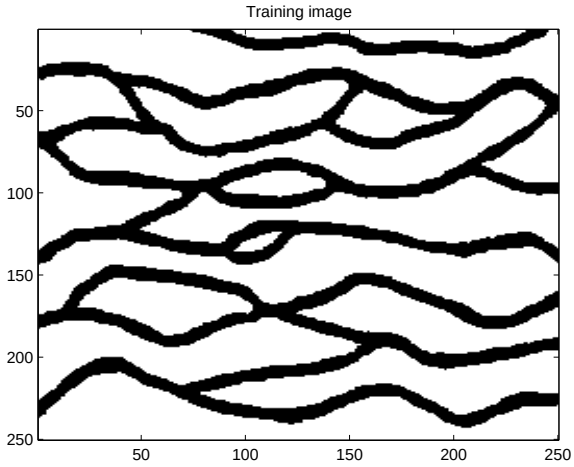


Figure 4.2: Training image of meandering channels

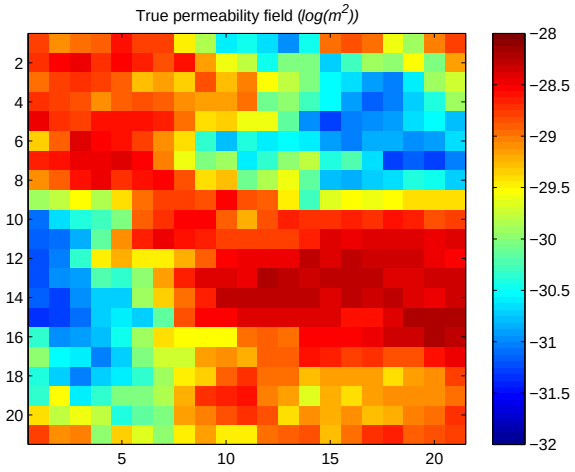


Figure 4.3: 'True' permeability field

well. The oil and the water flow rates plotted on the right-top subplot can be distinguished by solid and dashed lines respectively.

The measurements are now obtained from the state vector through relation (4.4) with the covariance matrix of observation noise taken in a block form

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}^1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}^2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R}^3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{R}^4 \end{bmatrix}, \tag{4.5}$$

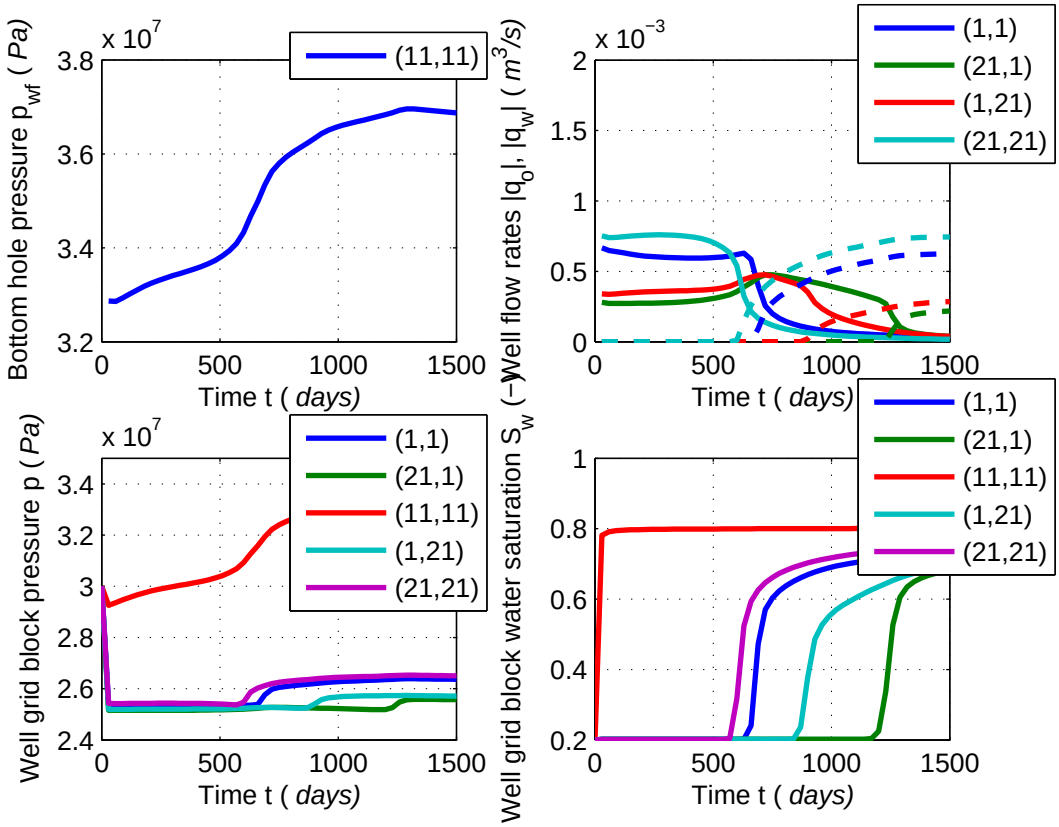


Figure 4.4: 'True' data

where $\mathbf{R}^1 = (0.05 * 10^7)^2 \cdot \mathbf{I} \in \mathbf{R}^{5\times5}$, $\mathbf{R}^2 = (0.05 * 10^{-1})^2 \cdot \mathbf{I} \in \mathbf{R}^{5\times5}$, $\mathbf{R}^3 = (0.05 * 10^7)^2 \cdot \mathbf{I} \in \mathbf{R}^{1\times1}$, $\mathbf{R}^4 = (0.05 * 10^{-3})^2 \cdot \mathbf{I} \in \mathbf{R}^{8\times8}$, i.e. the error in each observable variable is taken to be 5% of its actual scale. The same covariance matrix is then used to represent the measurements noise within data assimilation analysis. The simulator is run from time $t_0 = 0$ to $t_{end} = 1500(days)$ with a step of $30(days)$, providing the 'true' observations after each time step. If the value of generated observation is out of physically reasonable bounds for the process (e.g. $S < 0.2$ or $S > 0.8$), the simple truncation to the nearest bound is applied. Figure 4.5 has the same structure as Figure 4.4 and illustrates generated measurements.

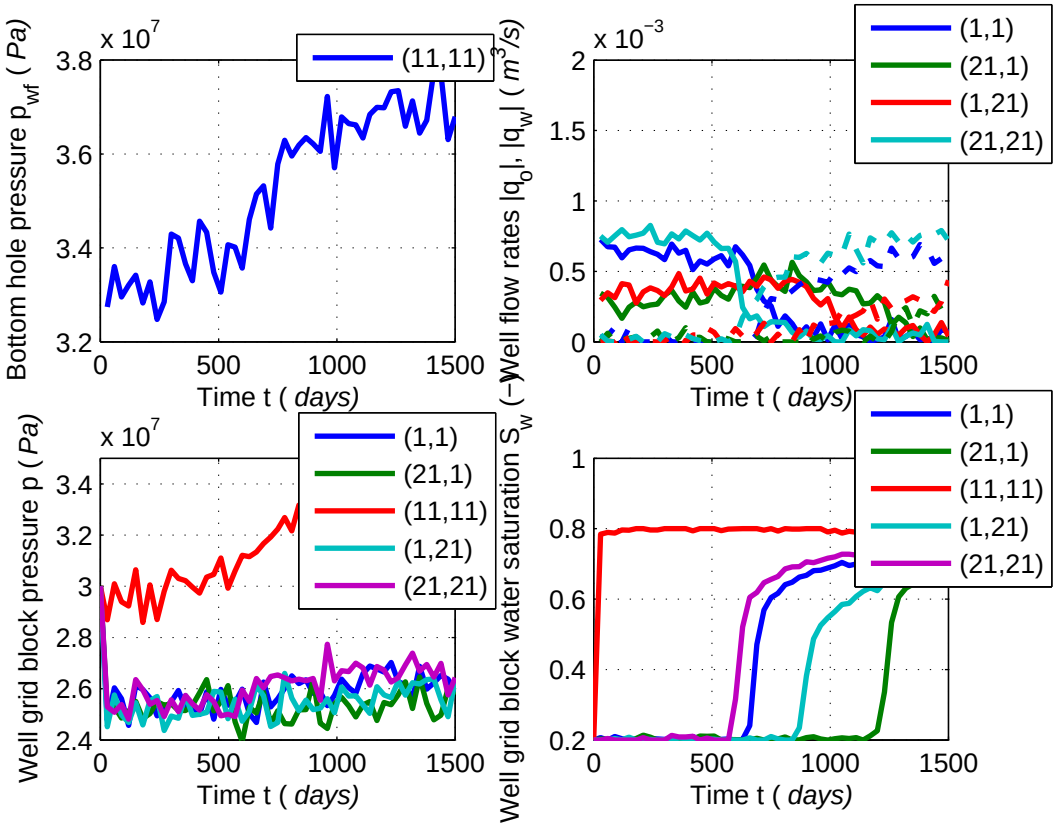


Figure 4.5: Generated measurements

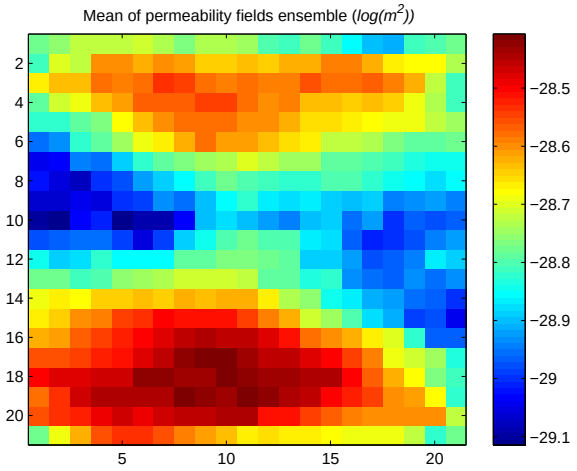
Although at the moment we have a model and a set of observations, EnKF analysis still is not possible, since it requires the decision on the initial state for the filtering procedure.

4.4 Filter initialization

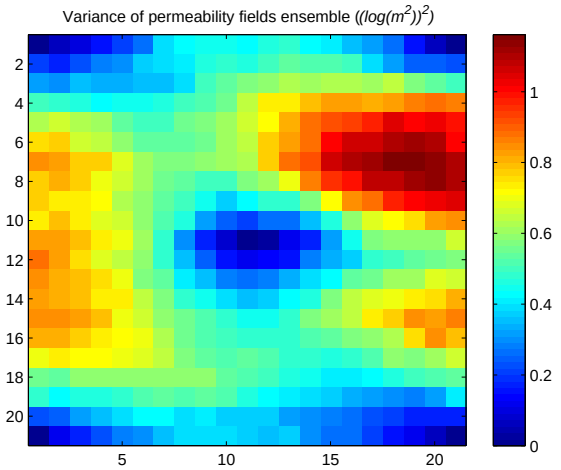
The filter has to be initialized by generating initial ensembles of only static and dynamic variables, because there is no production data available at the starting time.

Since the reservoir is typically in a state of equilibrium at the time when production starts, the initial dynamic variables (i.e., initial pressures and water saturations corresponding to each grid block) are assumed to be perfectly known (without uncertainty). Therefore they are the same for each ensemble member and equal to the initial condition of the 'true' model (i.e. $p = 3 * 10^7(Pa)$ and $S_w = 0.2$).

Thus, at the initial moment the only permutations contained in the initial ensemble are caused by initial ensemble of permeability models. Within the study we are going to use the initial permeability ensemble provided by Prof. J.D. Jansen (the ensemble consists of 999 members). The ensemble mean and variance are visualized in Figure 4.6, where the top picture corresponds to the ensemble mean and the bottom image — to the variance respectively.



(a)



(b)

Figure 4.6: Mean and variance of permeability fields ensemble

Now the state space representation (4.2)–(4.4) is derived, the set of synthetic observations is generated and the decision about initial ensemble

has been made, so we can start with the data assimilation procedure.

4.5 Measures of filter performance

While accomplishing EnKF analysis through different methods, one has to determine the ways of evaluating the filter performance. The measures of 'filter quality' that we are going to use for our purposes can be basically subdivided in two groups, namely into indices measuring the quality of either (i) history matching or (ii) forecast.

The measures in the first group aim at finding the difference between the 'true' observational data or model parameter and those which are estimated within the filter run.

We determine the quality of estimating a 'true' permeability field by the following root mean square (RMS) error:

$$\mathrm{RMS}(\mathbf{k}) = \sqrt{\frac{1}{T \cdot \dim(\mathbf{k})} \sum_{k=1}^{T} \|\mathbf{k}_k - \mathbf{k}^{\mathrm{true}}\|^2}, \tag{4.6}$$

where T denotes the number of assimilation steps, $\dim(\mathbf{k})$ states for the size of vector of estimated parameters (i.e., $\dim(\mathbf{k}) = 441$ in our study), $\mathbf{k}_k$ is a permeability vector at time moment t_k and vector $\mathbf{k}^{\mathrm{true}}$ represents the 'true' permeability field.

Similarly we define a history match measure for other particular types of production data:

$$\mathrm{RMS}(\alpha) = \sqrt{\frac{1}{T * n(\alpha)} \sum_{k=1}^{T} \|\mathbf{X}_k(\alpha) - \mathbf{Z}_k^{\mathrm{true}}(\alpha)\|^2}, \tag{4.7}$$

where α states for the type of production data, namely, α corresponds to pressures, saturations, bottom hole pressures or water and oil production rates at the wells, $n(\alpha)$ denotes the number of measurements of appropriate type, $\mathbf{X}_k(\alpha)$ and $\mathbf{Z}_k^{\mathrm{true}}(\alpha)$ are respectively the vector of estimated data and 'true' measurements at time point t_k consisting of such elements of $\mathbf{X}_k$ and $\mathbf{Z}_k^{\mathrm{true}}$ that correspond to the data of type α.

On the other hand we may evaluate the filter performance on the basis of the quality of predicting production data. For that purpose we continue running the simulator forward in time starting the last data assimilation step. The ensemble of state vectors obtained at this step is used as the

initial condition for the forecast. Afterwards we determine the quality of the forecast with the help of the measure (4.7), where $\mathbf{X}$ denotes the forecasted state vector.

The next chapter examines performance of various EnKF techniques embedded into the given experimental setup on the basis of above measures.

Chapter 5

Case study

This chapter addresses the analysis of some experience associated with application of ensemble Kalman filter and iterative ensemble Kalman filter techniques to history matching problem.

The study of each particular algorithm can be divided into two parts. At first, the history matching problem is solved and the estimate of model static parameter (i.e. permeability) is obtained. This step involves the use of generated measurements and data assimilation procedure. The filter analysis is done from time $t_0 = 0(days)$ till $t_{stop} = 510(days)$, which ensures that the water breakthrough event occurs in none of the production wells (see Figure 4.4). At the second stage of investigations the state vector estimated at t_{stop} is used as initial condition for making a forecast of reservoir performance from time t_{stop} up to time moment $t_{end} = 1500(days)$ when breakthrough events at all production wells have been already observed.

The implementation of EnKF and, hence, IEnKF algorithms has disclosed some key problems that have to be solved before performing the analysis of the filtering technique.

5.1 State vector feasibility

The use of classical ensemble Kalman filter in reservoir engineering framework meets an important obstacle concerned with obtaining physically unreasonable values of the state variables. It originates from performing data assimilation on the state vector without any constraint coming from the

physical nature of the parameters. Hence the updated dynamic variables may become unfeasible and inconsistent with estimated static variables.

After some reviewing the literature and thinking over we have discovered the following approaches to overcome the infeasibility of state variables, most of which are oriented towards improving the saturation entries into the state vector:

simple truncation

the elements of the state vector, which become out of physically reasonable bounds after the time update step of filtering analysis, are simply truncated to the nearest bound. This method is used e.g. by [27], although in our opinion it leads to some losses in characterization of the a posterior distribution;

normal score transform

the state vector is composed of the normal score transform values [9] of water saturations instead of saturations themselves. However, it was shown in [11] that such a method provides reasonable values of saturations, but do not reduce their spatial oscillations;

Reynolds' transformation

the elements of state vectors which do not belong to the sets within physical bounds are scaled back to the interval of reasonable values through log transformation. The forward simulator run is then performed with such a vector as initial condition and the outcome of a run is scaled back with an appropriate inverse transform [26]. We are not going to use this method, since in our case the operator of transformation and the reservoir operator are not commutative;

saturation front location as a state variable

the state vector is composed of the variables describing the location of the shock front instead instead of water saturations [11]. Unfortunately this method is not very applicable for two- or three-dimensional problems;

iterated update

whenever the updated saturation values are detected to be out of bounds the simulator is rerun from the previous time step to recompute the dynamic variables using the updated static variables. Then

the state vector is updated again. The iteration is repeated until the corrected values of saturations fall into physically reasonable interval or the maximum number of iterations is exceeded [11];

confirming EnKF

the additional 'confirmation' step is involved into classical EnKF algorithm [25]. The details are given in the next subsection.

One may also think of modifying the approaches to constraint classical Kalman filtering [22, 23] and use of truncated Gaussian distribution in filtering analysis [6]. The last methods were nor tested on reservoir engineering problems neither they were developed for EnKF, although looking promising for further investigations.

We are going to use the confirming EnKF approach, which is known as producing reasonable results.

5.1.1 The confirming ensemble Kalman filter

The authors of [25] proposed to include one additional so-called 'confirming' step into EnKF algorithm in order to ensure that updated state is plausible and consistent with flow equations. The flowchart of confirmation step is presented in Figure 5.1.

The idea of confirmation step is following. Starting at time moment $k-1$ we, at first, perform a forecast update up to time k and then a data assimilation step. Then take only recently updated static model parameters and run again the flow simulator from current time k to the next time moment k. The dynamic variables obtained replace those got after measurement update stage of EnKF and become an initial guess for the next time update step. This procedure guarantees that the updated state is consistent with the flow model.

This confirmation step then occupies an appropriate place in the classical ensemble Kalman filter algorithm (see Figure 5.2).

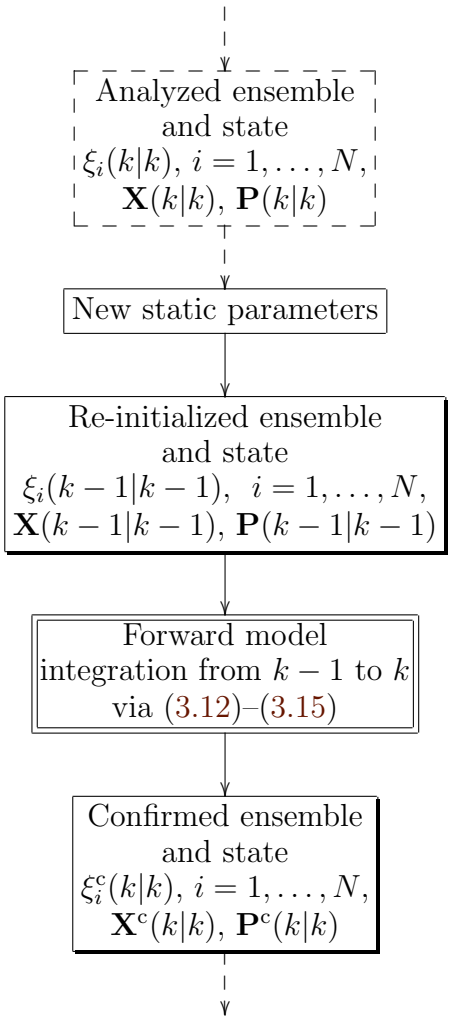


Figure 5.1: Confirmation step flowchart

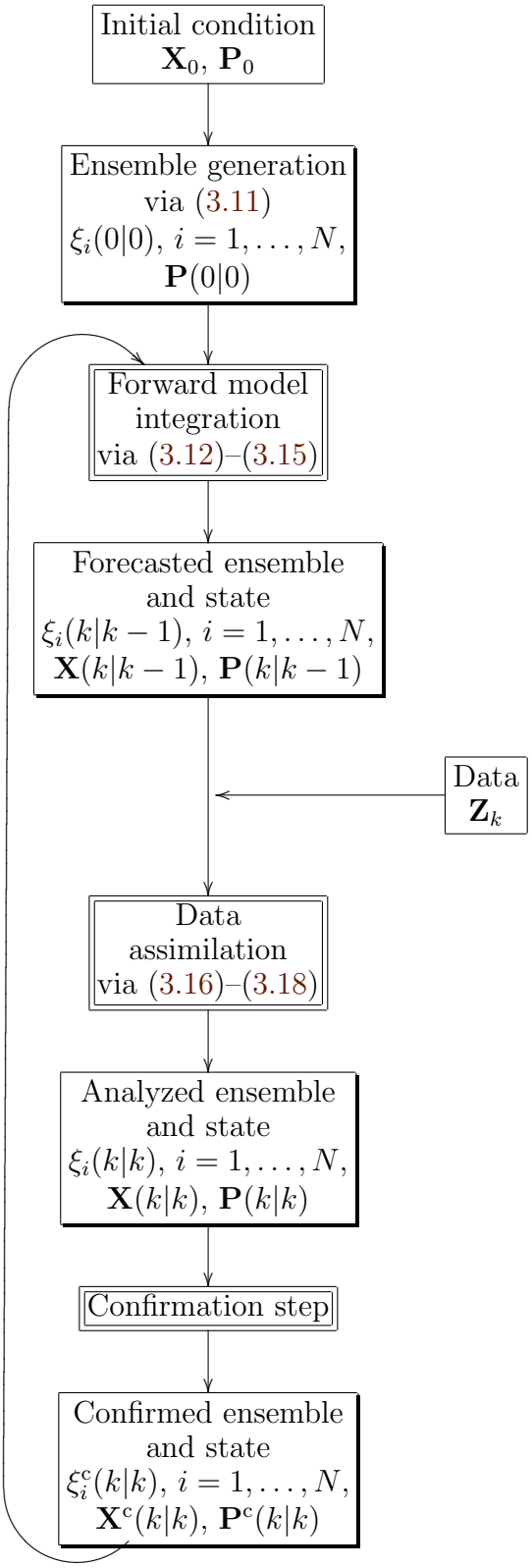


Figure 5.2: Confirming ensemble Kalman filter algorithm flowchart

The inclusion of the confirmation step into the algorithm results in almost doubling the computational time due to additional forward model run per ensemble member at each time step.

In fact we use the confirming EnKF technique instead of classical ensemble Kalman filtering for our investigations. So, from now and on we mean confirming EnKF technique under the abbreviation of EnKF.

5.2 Re-scaling state vector

The other problem related to implementation of the filter algorithms appears due to the different units and scales for measuring the variables, which leads to the block structure of the state vector where the two extremes are represented by pressure values of order 10^7 and production rates of order 10^{-3}, although both are equally important for the study. This fact becomes crucial when matrix-matrix and matrix-vector multiplications in measurement update step have to be done. Within these operations the small numbers tend to be lost because of computational error. Unfortunately, the computer is not able to handle properly a summation of two numbers if they essentially differ in the order of magnitude. To prevent such a loss of information and even divergence of the filter, we have proposed the following trick that is applied to data assimilation step of the filtering procedure.

Consider arbitrary time step k. Let us define the scaling matrices $\mathbf{A}_k \in \mathbf{R}^{19\times 1332}$ and $\mathbf{B}_k \in \mathbf{R}^{1332\times 1332}$ such that multiplications $\mathbf{A}_k\mathbf{M}_k\mathbf{L}(k|k-1)$ and $\mathbf{B}_k\mathbf{L}(k|k-1)$ result in matrices which elements do not sufficiently differ in order. We make the following choice for our study:

$$\mathbf{A}_k = \begin{bmatrix} \mathbf{A}_k^1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_k^2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{A}_k^3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{A}_k^4 \end{bmatrix}$$

and

$$\mathbf{B}_k = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_k^1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{B}_k^2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{B}_k^3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{B}_k^4 \end{bmatrix},$$

where $\mathbf{A}_k^1 = 10^{-7} \cdot \mathbf{I} \in \mathbf{R}^{5\times 5}$, $\mathbf{A}_k^2 = 10 \cdot \mathbf{I} \in \mathbf{R}^{5\times 5}$, $\mathbf{A}_k^3 = 10^{-7} \cdot \mathbf{I} \in \mathbf{R}^{1\times 1}$,

$\mathbf{A}_k^4 = 10^3 \cdot \mathbf{I} \in \mathbf{R}^{8\times 8}$ and $\mathbf{B}_k^1 = 10^{-7} \cdot \mathbf{I} \in \mathbf{R}^{441\times 441}$, $\mathbf{B}_k^2 = 10 \cdot \mathbf{I} \in \mathbf{R}^{441\times 441}$, $\mathbf{B}_k^3 = 10^{-7} \cdot \mathbf{I} \in \mathbf{R}^{1\times 1}$, $\mathbf{B}_k^4 = 10^3 \cdot \mathbf{I} \in \mathbf{R}^{8\times 8}$.

Denoting scaled versions of matrices $\mathbf{M}_k\mathbf{L}(k|k-1)$ and $\mathbf{L}(k|k-1)$ as $\mathbf{C}_k^1 = \mathbf{A}_k\mathbf{M}_k\mathbf{L}(k|k-1)$ and $\mathbf{C}_k^2 = \mathbf{B}_k\mathbf{L}(k|k-1)$ and accomplishing some transformations, we obtain the following expression to compute the Kalman gain (3.16):

$$
\begin{aligned}
&\mathbf{K}(k) = \frac{1}{N-1}\mathbf{L}(k|k-1)\mathbf{L}^{\mathrm{T}}(k|k-1)\mathbf{M}_k^{\mathrm{T}} \\
&* \left(\frac{1}{N-1}\mathbf{M}_k\mathbf{L}(k|k-1)\mathbf{L}^{\mathrm{T}}(k|k-1)\mathbf{M}_k^{\mathrm{T}} + \mathbf{R}_k\right)^{-1} \\
&= \frac{1}{N-1}\left(\mathbf{B}_k^{-1}\mathbf{B}_k\right)\mathbf{L}(k|k-1)\mathbf{L}^{\mathrm{T}}(k|k-1)\mathbf{M}_k^{\mathrm{T}}\left(\mathbf{A}_k\mathbf{A}_k^{-1}\right) \\
&* \left(\frac{1}{N-1}\mathbf{M}_k\mathbf{L}(k|k-1)\mathbf{L}^{\mathrm{T}}(k|k-1)\mathbf{M}_k^{\mathrm{T}} + \mathbf{R}_k\right)^{-1}\left(\mathbf{A}_k^{-1}\mathbf{A}_k\right) \\
&= \frac{1}{N-1}\mathbf{B}_k^{-1}\left(\mathbf{B}_k\mathbf{L}(k|k-1)\right)\left(\mathbf{A}_k\mathbf{M}_k\mathbf{L}(k|k-1)\right)^{\mathrm{T}} \\
&* \left(\frac{1}{N-1}\mathbf{A}_k\mathbf{M}_k\mathbf{L}(k|k-1)\left(\mathbf{A}_k\mathbf{M}_k\mathbf{L}(k|k-1)\right)^{\mathrm{T}} + \mathbf{A}_k\mathbf{R}_k\mathbf{A}_k\right)^{-1}\mathbf{A}_k \\
&= \frac{1}{N-1}\mathbf{B}_k^{-1}\mathbf{C}_k^2\left(\mathbf{C}_k^1\right)^{\mathrm{T}}\left(\frac{1}{N-1}\mathbf{C}_k^1\left(\mathbf{C}_k^1\right)^{\mathrm{T}} + \mathbf{A}_k\mathbf{R}_k\mathbf{A}_k\right)^{-1}\mathbf{A}_k
\end{aligned}
$$

or

$$\mathbf{K}(k) = \mathbf{B}_k^{-1}\mathbf{K}_k^1\mathbf{A}_k, \tag{5.1}$$

where $\mathbf{K}_k^1 = \frac{1}{N-1}\mathbf{C}_k^2\left(\mathbf{C}_k^1\right)^{\mathrm{T}}\left(\frac{1}{N-1}\mathbf{C}_k^1\left(\mathbf{C}_k^1\right)^{\mathrm{T}} + \mathbf{A}_k\mathbf{R}_k\mathbf{A}_k\right)^{-1}$.

Note that we have used the property of diagonal matrices $\mathbf{A}_k$ and $\mathbf{B}_k$ to be equal to their transposed versions $\mathbf{A}_k^{\mathrm{T}}$ and $\mathbf{B}_k^{\mathrm{T}}$ respectively. The Kalman gain factorization (5.1) involves matrix $\mathbf{K}_k^1$ which is computed via operations only on the scaled matrices. The ensemble update (3.17) is now performed as

$$
\begin{aligned}
&\xi_i(k|k) = \xi_i(k|k-1) + \mathbf{K}(k)\left(\mathbf{Z}_k - \mathbf{M}_k\xi_i(k|k-1) + \mathbf{V}_k^i\right) \\
&= \xi_i(k|k-1) + \mathbf{B}_k^{-1}\mathbf{K}_k^1\mathbf{A}_k\left(\mathbf{Z}_k - \mathbf{M}_k\xi_i(k|k-1) + \mathbf{V}_k^i\right) \\
&= \xi_i(k|k-1) + \mathbf{B}_k^{-1}\left(\mathbf{K}_k^1\left(\mathbf{A}_k\mathbf{Z}_k - \mathbf{A}_k\mathbf{M}_k\xi_i(k|k-1) + \mathbf{A}_k\mathbf{V}_k^i\right)\right).
\end{aligned} \tag{5.2}
$$

The expression (5.2) also involves operations on the scaled matrices except multiplication with $\mathbf{B}_k^{-1}$, which is not dramatic since $\mathbf{B}_k$ is a diagonal matrix.

We are now fully equipped to perform data assimilation analysis.

5.3 Results and discussion

A sequence of simulations has been accomplished to test the performance of EnKF and IEnKF assimilation algorithms in the framework of estimating the model parameters for two-phase two-dimensional fluid flow model. We proceed by describing particular instances and discussing the obtained results.

5.3.1 Sensitivity of EnKF to the number of ensemble members

It has been already notified that performance of EnKF method is quite sensitive to the size of the ensemble used. Thus, it seems reasonable to start by investigating how filter performance depends on a number of ensemble members and find some optimum ensemble size. Within the measurements update step the whole bunch of available data, namely, pressures and saturations at the well grid blocks, bottom hole pressures and water and oil flow rates, is assimilated every time period of $30(days)$. Data assimilation is performed from time $t_0 = 0(days)$ to time moment $t_{stop} = 510(days)$. We examine time averaged RMS difference (4.6) between the elements of ensemble mean and true state vectors, which correspond to the values of model parameter (i.e. permeability), for $10,\ 20,\ \ldots,\ 90$ ensemble members. The filter is initialized with the values of $\mathbf{X}_0$ and $\mathbf{P}_0$ fixed as in Section 4.4 and measurement covariance matrix $\mathbf{R}$ taken as (4.5). Despite this we also analyze CPU time needed to run EnKF for each ensemble.

Figure 5.3 confirms the fact stated in several publications e.g. [13]: as a number of members in ensemble increases RMS error slowly decreases. This leads to the natural idea of using for computations an ensemble of larger size. However, the CPU time required for performing simulation increases linearly with the number of ensemble members. Thus, it becomes necessary to decide which ensemble size to select keeping in mind two conditions: (i) we have to chose ensemble size which being increased further does not cause sufficient improvement of an accuracy of computations; (ii) simulation using ensemble with the chosen size should not require too much CPU time. Preliminary analysis shows that in our case it seems optimal to perform further EnKF runs on $N = 60$ ensemble members. Since iterative modification of EnKF actually has the same origin, we find it being appropriate to use the same ensemble size as being optimal also for IEnKF runs.

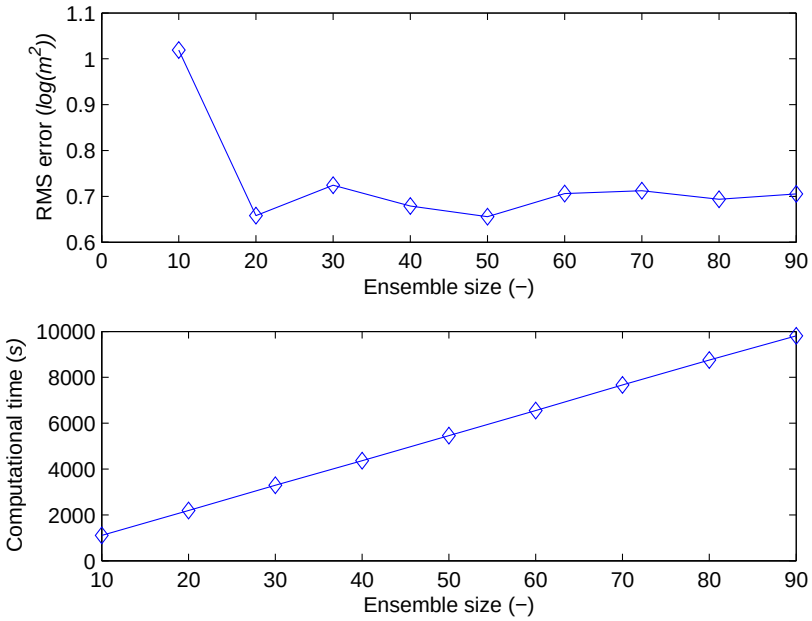


Figure 5.3: EnKF: RMS error and CPU time analysis subject to ensemble size

5.3.2 History matching via EnKF

Let us now present the outcome of data assimilation procedure accomplished via the EnKF algorithm with respect to the optimal number of ensemble members.

At first we consider the quality of estimating the model parameter. For that purpose space averaged RMS errors are plotted in time (see Figure 5.4). This quantities are related to the part of ensemble mean and ensemble members values corresponding to evaluated permeability. The graph demonstrates improvement of the parameter estimation in the first few data assimilation steps followed by stabilization of the error, and reduction of the uncertainty in estimated value (since the ensemble spread becomes narrower). This means that at the later times assimilated data carries less useful information on reservoir structure than at the early times. Indeed, we obtain a permeability field resembling the true one, although some underestimating the values in the upper right and overestimating the values in the bottom left corner area of the field (compare Figure 5.5(a) and left bottom chart on Figure 5.5(b)). The variance field is actually obtained as the diagonal terms of covariance matrix computed from the statistical

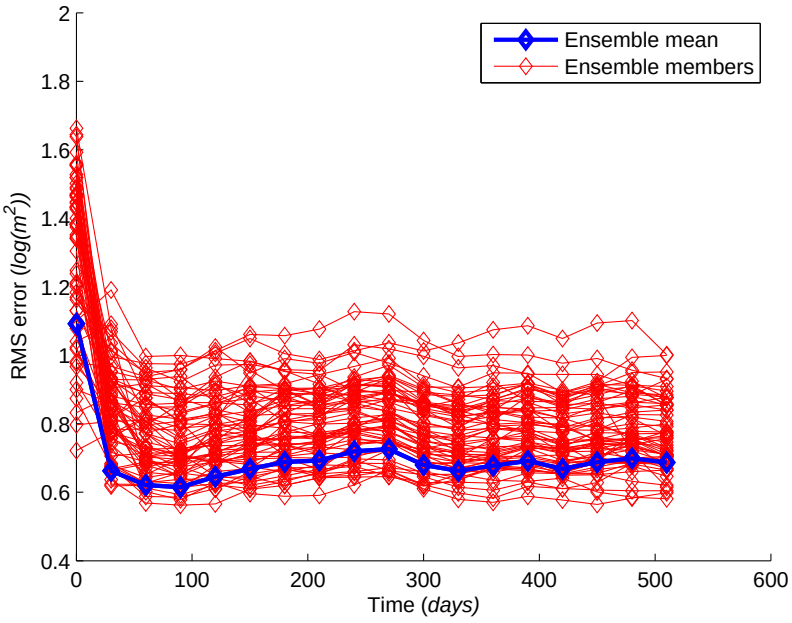
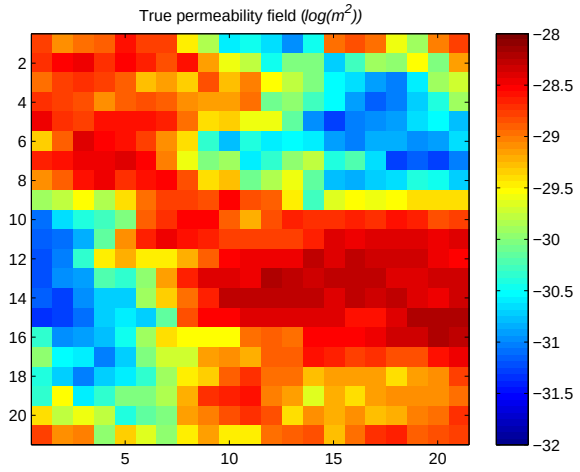


Figure 5.4: EnKF: RMS error in estimated permeability vs time

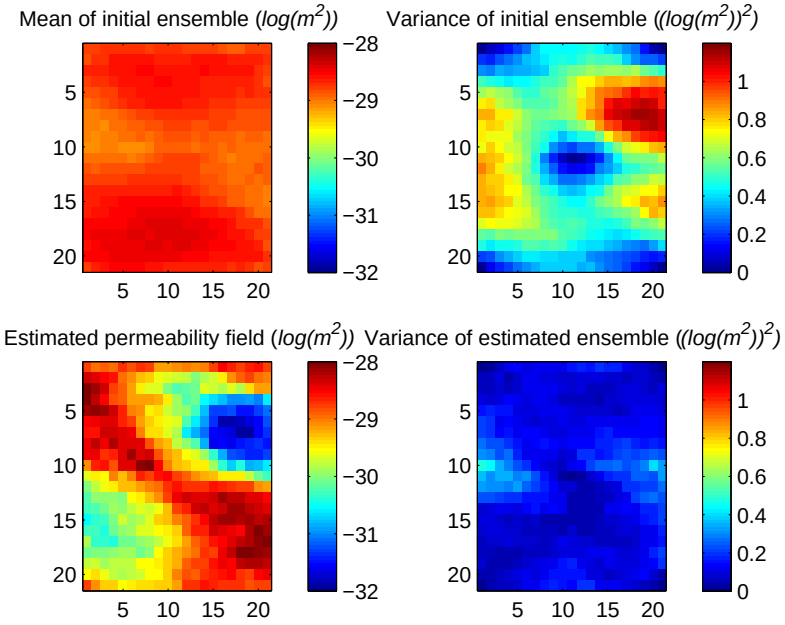
properties of the ensemble. The difference between the top right and the bottom right subplots in Figure 5.5(b) indicates reduction of the variance and therefore uncertainty in the estimation.

The next point of evaluating the EnKF algorithm performance is concerned with the quality of matching the actual data which comes from the true model. The outcome of data assimilation together with the reference model are visualized by Figures 5.6–5.9.

It can be seen that in general the results obtained from EnKF method match the reference model. The only serious problem appears while matching water saturation data at $(21, 21)$ grid block, which is connected with injection well grid block through a high permeability layer. In this case the time of water breakthrough event is underestimated. The same kind of problem occurs when matching water saturation data at $(1, 1)$ grid cell, since some of the ensemble members are already indicating the water breakthrough. Note that this upper left corner of reservoir is also connected to the injector via the high permeability area.



(a)



(b)

Figure 5.5: EnKF: True, initial and estimated permeability fields and corresponding variances

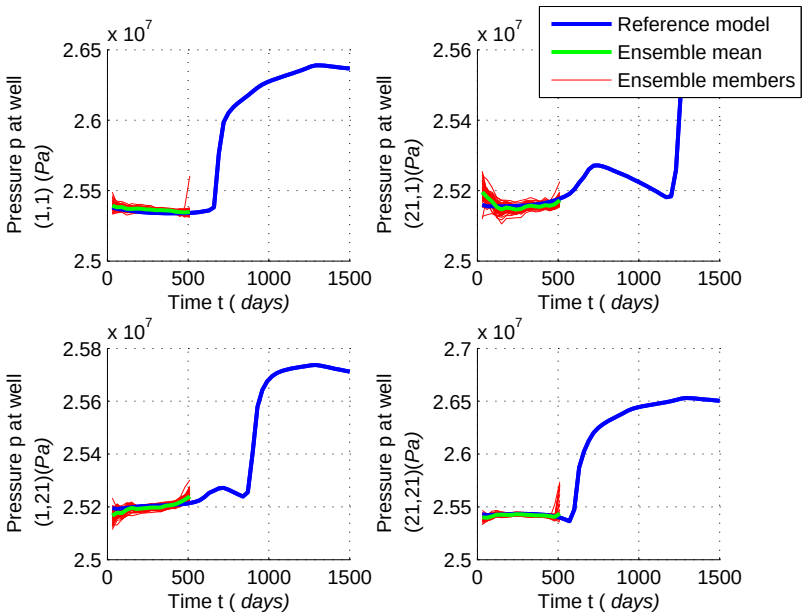


Figure 5.6: Historical match of pressure data at production well grid blocks

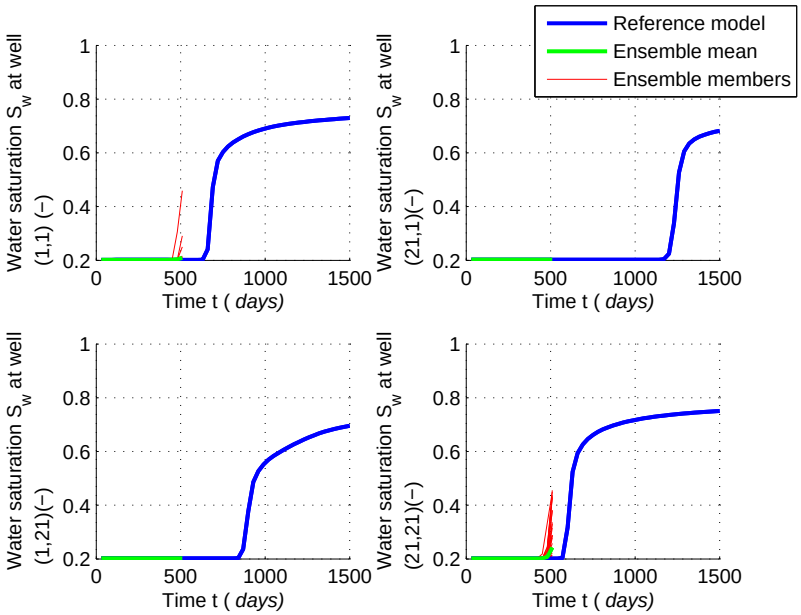


Figure 5.7: Historical match of water saturation data at production well grid blocks

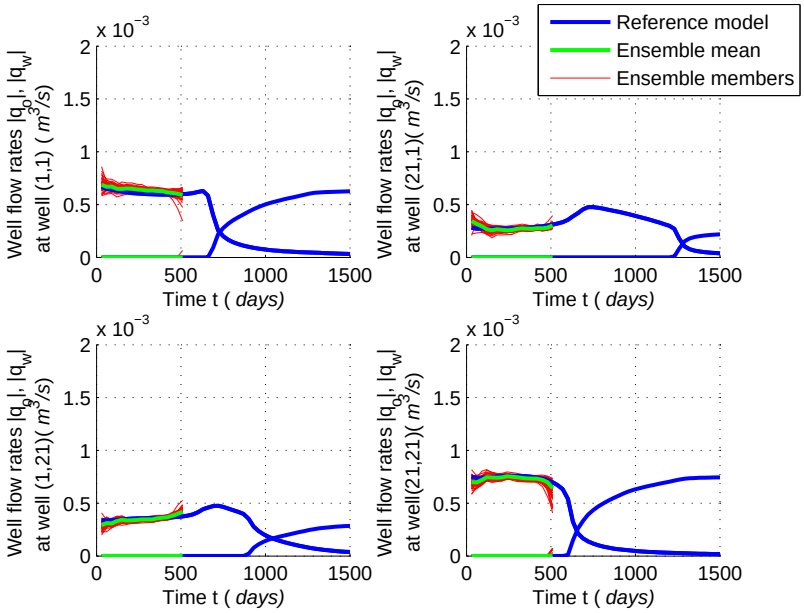


Figure 5.8: Historical match of water and oil flow rates data at production well grid blocks

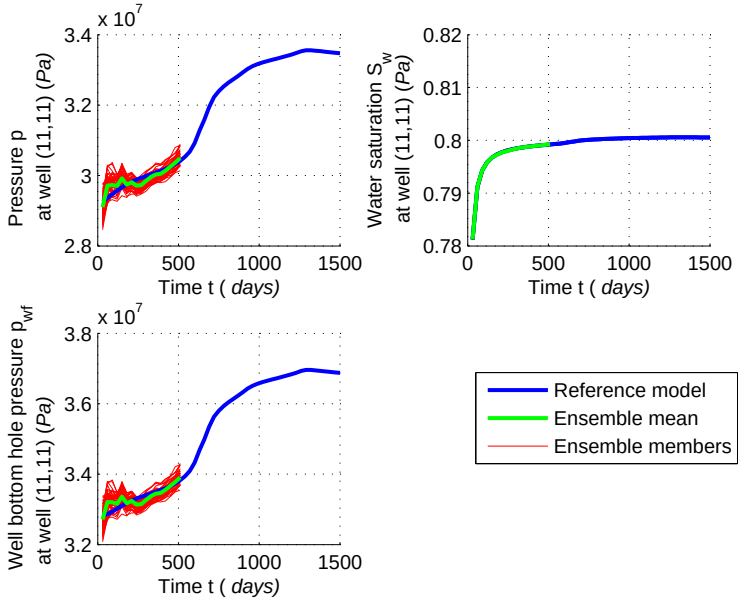


Figure 5.9: Historical match of pressure, water saturation and bottom hole pressure data at injection well grid block

Well		RMS error				
grid block	type	$p\ (Pa)$	$S\ (-)$	$q_o\ (m^3/s)$	$q_w\ (m^3/s)$	$p_{wf}\ (Pa)$
$(1,1)$	production	16501	0.001846	0.000029	0	-
$(21,1)$	production	13163	0.000005	0.000023	0	-
$(11,11)$	injection	158024	0.000044	-	-	157713
$(1,21)$	production	12548	0.000005	0.000022	0	-
$(21,21)$	production	14713	0.009995	0.000025	0.000002	-

Table 5.1: EnKF: History match measure for various types of assimilated data

The results presented in Figures 5.6–5.9 are supported by computing history match error (4.7) for each type of assimilated data. The values of the error are summarized in the Table 5.1.

It follows from the Table 5.1 that value corresponding to time averaged RMS error in water saturation at well grid block $(21,21)$ and the scale of saturation variable have nearly the same order of magnitude. This indicates the problems in matching water saturation data at the bottom right corner of the reservoir.

However, the complete analysis of the numerical information presented in the Table 5.1 confirms that the algorithm performs reasonably well. Thus, we may consider a history matching stage as being passed successfully and perform a forecast based on the estimated permeability field.

5.3.3 Forecasting reservoir performance

Consider the predictive stage of reservoir simulation. With the estimated model parameter in hand and the ensemble used for data assimilation we perform the following forecasting experiments:

- forecast from time t_0 up to time moment $t_{end} = 1500(days)$ with initial ensemble taken to be the same as the initial ensemble for data assimilation;
- forecast from time t_{stop} up to time moment $t_{end} = 1500(days)$ where initial condition at t_{stop} is taken to be the ensemble of state vectors estimated at $t_{stop} = 510(days)$;
- forecast from time t_0 up to time moment $t_{end} = 1500(days)$ where ensemble of initial model parameters is taken to be the same as ensemble of estimated parameters at $t_{stop} = 510(days)$, and initial pressures and saturations have the typical values of $3 * 10^7(Pa)$ and 0.2 respectively.

The numerical results of the experiments are presented in Tables 5.2–5.4 and visualized in Figures 5.10–5.13; the curves drawn there actually correspond to the mean of the ensemble of state vectors at each moment of time. Note that the results are plotted starting time $t_1 = 30(days)$ and that for the second experiment the parts of the curves corresponding to times t_1 up to t_{stop} are appropriate outcomes of data assimilation procedure.

Well		RMS error				
grid block	type	$p\ (Pa)$	$S\ (-)$	$q_o\ (m^3/s)$	$q_w\ (m^3/s)$	$p_{wf}\ (Pa)$
$(1,1)$	production	261284	0.093006	0.000115	0.000124	-
$(21,1)$	production	404078	0.191811	0.000154	0.000190	-
$(11,11)$	injection	1447839	0.000358	-	-	1447107
$(1,21)$	production	313572	0.100296	0.000123	0.000171	-
$(21,21)$	production	392179	0.124873	0.000175	0.000208	-

Table 5.2: Forecast measure for various types of data obtained within forecast from $t_0 = 0(days)$ to $t_{end} = 1500(days)$ (without parameter estimation)

Well		RMS error				
grid block	type	$p\ (Pa)$	$S\ (-)$	$q_o\ (m^3/s)$	$q_w\ (m^3/s)$	$p_{wf}\ (Pa)$
$(1,1)$	production	139145	0.093061	0.000104	0.000069	-
$(21,1)$	production	57054	0.062875	0.000049	0.000037	-
$(11,11)$	injection	634638	0.000416	-	-	634401
$(1,21)$	production	104269	0.072727	0.000065	0.000027	-
$(21,21)$	production	135055	0.070946	0.000075	0.000057	-

Table 5.3: Forecast measure for various types of data obtained within forecast from $t_0 = 510(days)$ to $t_{end} = 1500(days)$ (with parameter estimation)

Well		RMS error				
grid block	type	$p\ (Pa)$	$S\ (-)$	$q_o\ (m^3/s)$	$q_w\ (m^3/s)$	$p_{wf}\ (Pa)$
$(1,1)$	production	121140	0.081353	0.000092	0.000060	-
$(21,1)$	production	40444	0.049285	0.000042	0.000031	-
$(11,11)$	injection	518001	0.000339	-	-	517795
$(1,21)$	production	79003	0.053132	0.000048	0.000019	-
$(21,21)$	production	105763	0.053564	0.000056	0.000044	-

Table 5.4: Forecast measure for various types of data obtained within forecast from $t_0 = 0(days)$ to $t_{end} = 1500(days)$ (with parameter estimation)

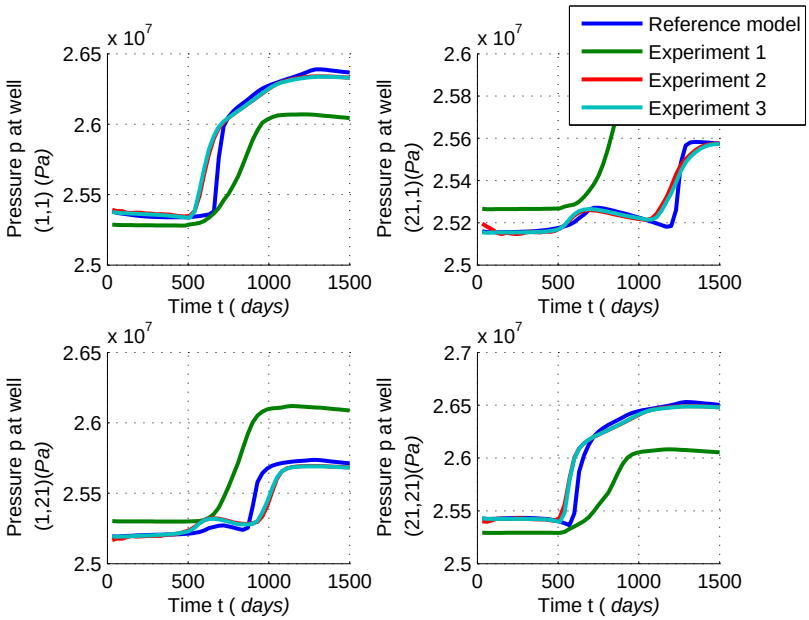


Figure 5.10: Various forecasts of pressure data at production well grid blocks

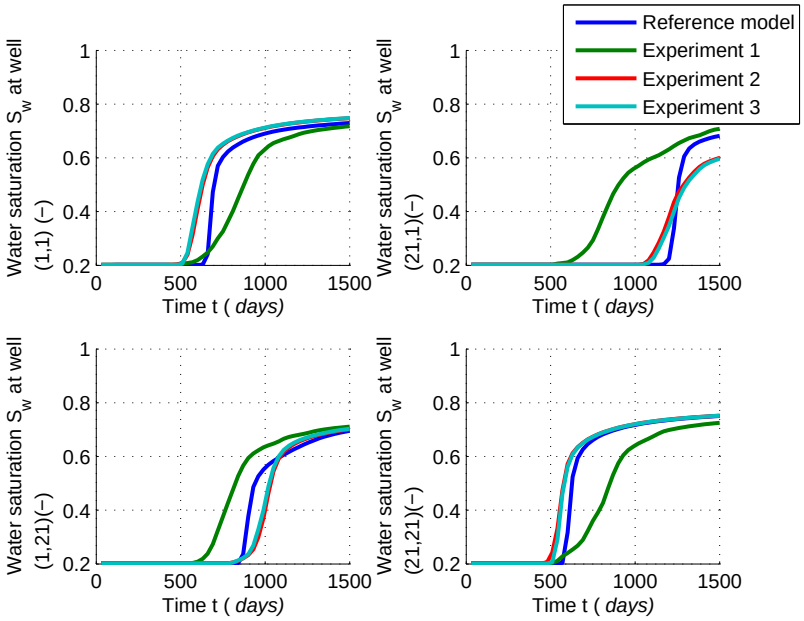


Figure 5.11: Various forecasts of water saturation data at production well grid blocks

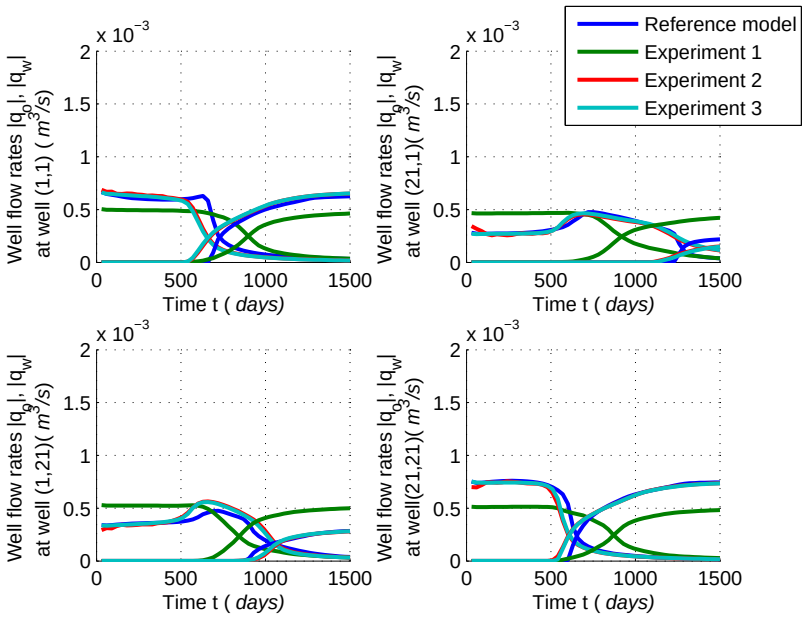


Figure 5.12: Various forecasts of water and oil flow rates data at production well grid blocks

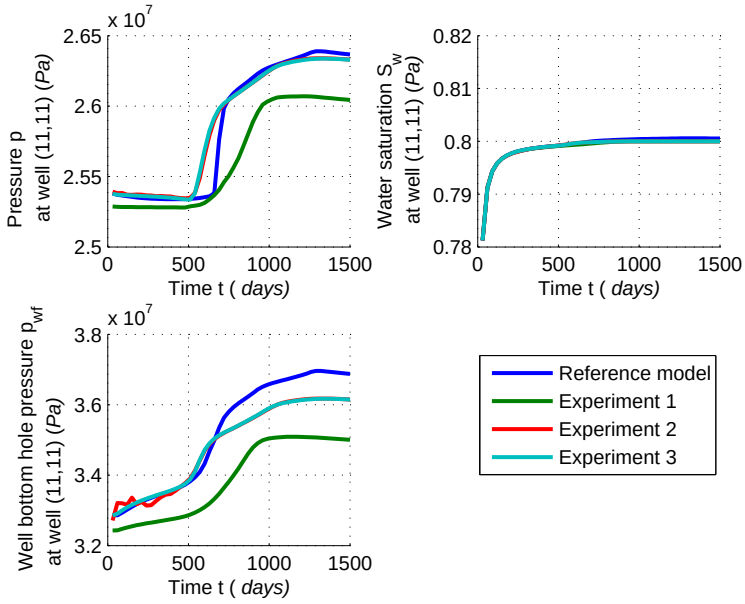


Figure 5.13: Various forecasts of pressure, water saturation and bottom hole pressure data at injection well grid block

Studying the results of the first experiment we find out especially serious problems with forecasting proper values of water saturations at grid blocks hosting production wells, pressure at injector grid block and bottom hole pressure at the injection well. Note that in general at wells located at the top right and bottom left corners of the reservoir the pressures and oil and water production rates are overestimated, whereas the time of water breakthrough event is underestimated. This phenomenon can be explained by the fact that reservoir simulation is performed based on the ensemble of permeability fields which mean is much higher than the true values in the top right and bottom left areas of the field. In turn, the values of the mean permeability field at the layer connecting top left, bottom right production wells and injector, which are lower than the true ones, cause underestimation of the pressures, water and oil production rates, and bottom hole pressure at injection well.

Since from the practical point of view it is very important to provide an accurate prediction of the water breakthrough event at production wells, the results of undertaken experiment clearly indicate the necessity of history matching stage of simulation and, hence, the need of parameter estimation to improve the subsequent forecast of reservoir behavior.

The second and third tests that we have run give the results shown in Table 5.3 and Table 5.4 respectively. Note that although do not indicating any critical obstacles within the forecast, the entries of the Table 5.3 are not directly comparable with corresponding cells from the Table 5.2 and Table 5.4, since the averaged RMS error was computed for a different time intervals.

It turns out that simulations based on calibrated model (i.e. involving estimated values of permeability) represent the matter better, which meets our expectations. The prediction curves obtained for various type of data within experiments 2 and 3 are similar. This can be explained by the following reasoning: (i) starting time t_{stop} we actually use the same values of model parameter for forecast in both cases; (ii) at time t_{stop} the elements of state vectors that do not represent permeability values are not equal. However, the vector used in second experiment comes from data assimilation which ensures a consistency with the true state vector coming from the reference model. In turn the state vector at time t_{stop} in the third experiment is obtained as an outcome of model run based on the estimated value of model parameter, which was computed via history matching procedure in such a way that it makes simulated values of observable variables to be consistent with true values of those quantities; (iii) the model used in ex-

periments is deterministic. So, one can save some computational time by performing a forecast only running the model from the end of data assimilation period and further based on the state vector and model parameter finally estimated within data assimilation.

Although history matching on the basis of EnKF technique has demonstrated its efficiency for proper estimating model parameter and further predictions, there is still space for improvement. We may aim at obtaining better representation of reservoir heterogeneous structure, which in turn will result in increasing the quality of forecasts.

5.3.4 Motivation for parameter estimation via IEnKF: global carbon-dioxide model

Let us provide some heuristical justification of IEnKF technique applicability to parameter estimation problems. To investigate the performance of IEnKF algorithm we would like to perform a trial experiment with relatively 'transparent' small-scaled model. Consider for that purpose the global carbon-dioxide (CO2) model.

The global CO2 model describes concentration of carbon-dioxide at several ocean layers and in simple case has the following approximate discrete representation [13]:

$$\mathbf{X}_{k+1} = \exp(\mathbf{A}\Delta t)\mathbf{X}_k + \mathbf{B}\Delta t, \tag{5.3}$$

where

$$\mathbf{A} = \begin{bmatrix} -k_{12} - k_{13} & k_{21} & k_{13} & 0 \\ k_{12} & -k_{21} & 0 & 0 \\ k_{13} & 0 & -k_{31} - k_{34} & k_{43} \\ 0 & 0 & k_{34} & -k_{43} \end{bmatrix}, \ \mathbf{B}(t) = [0.06(t - 1890)\ 0\ 0\ 0]^{\mathrm{T}},$$

$k_{12} = 1/33$, $k_{21} = 1/40$, $k_{13} = 1/5$, $k_{31} = 1/6$, $k_{34} = 1/6.2$ and $k_{43} = 1/300$ are model parameters, Δt is $10(years)$. We assume the measurements of CO2 in the top layer of the ocean being available and present measurement relation (3.2) in a form of

$$\mathbf{Z}_k = \mathbf{M}\mathbf{X}_k + \mathbf{V}_k, \tag{5.4}$$

where $\mathbf{M} = [0\ 0\ 1\ 0]$ is the measurement matrix and $\mathbf{V}_k$ is the measurement noise with zero mean an covariance $\mathbf{R} = 0.01$. Then the data set is generated

according to the above model with the initial condition at $t_0 = 1890(years)$ being taken as $\mathbf{X}_0 = [51\ 62.2\ 61.7\ 2986.4]^{\mathrm{T}}$.

Suppose now that model parameter k_{31} is unknown and has to be estimated. This can be done by considering a parameter as additional system state variable. Let us perform parameter estimation via ensemble Kalman filtering starting with a wrong guess $k_{31} = 0.5$ and error covariance $\mathbf{R} = 0.01$. The initial vector $\mathbf{X}_0$ is assumed to be perfect and the covariance of the initial model parameter guess is taken as 0.09. The estimated parameter value at the end of the run is equal to 0.1666. The change of parameter estimation value in time is shown in the Figure 5.14.

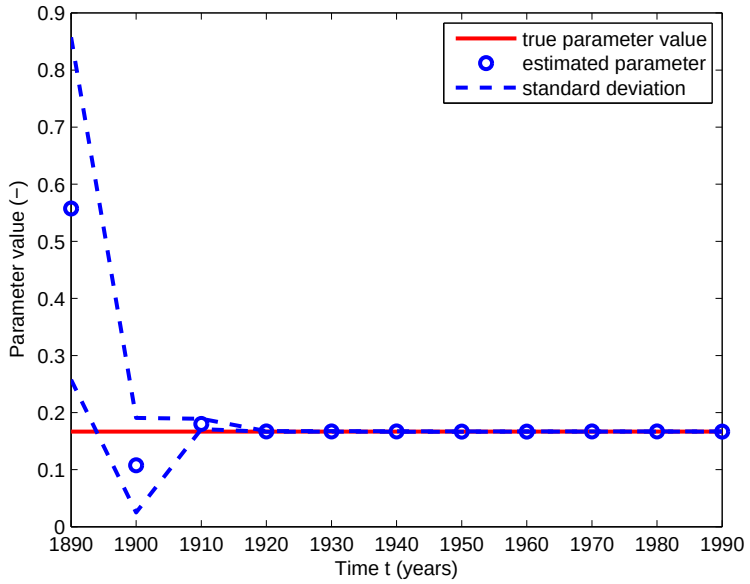


Figure 5.14: CO2 model: EnKF parameter estimation with initial guess normally distributed with parameters $(0.5, 0.09)$ and covariance of measurement noise $\mathbf{R} = 0.01$

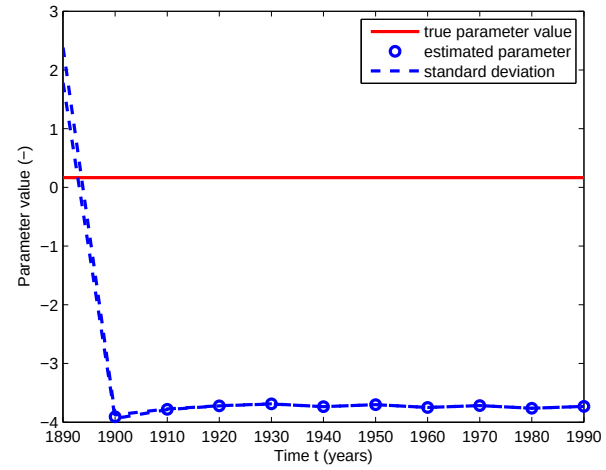
In this example the sufficient accuracy of the estimation was achieved by one global EnKF run. However if we start with initial guess for the parameter being much far from the true one, e.g. 1.5 and repeat the experiment, then after all available data is assimilated we end up with the estimation being equal to -3.7305. Hence, the idea of IEnKF described in Section 3.3 can be applied. Note that at the second iteration of EnKF algorithm we assume the model parameter coming from $\mathcal{N}(-3.7305, 0.09)$, i.e. the mean is equal to the obtained estimation and the variance stays the same as in

the very beginning of the experiment. The second global EnKF iteration provides us with an estimated value of the parameter equal to -3.4180, which is closer to the true one, but still needs to be improved. It turns out that 14 global iterations have to be performed in such a case to obtain reasonable estimation of value 0.1669. The first few global EnKF iterations are presented in Figure 5.15.

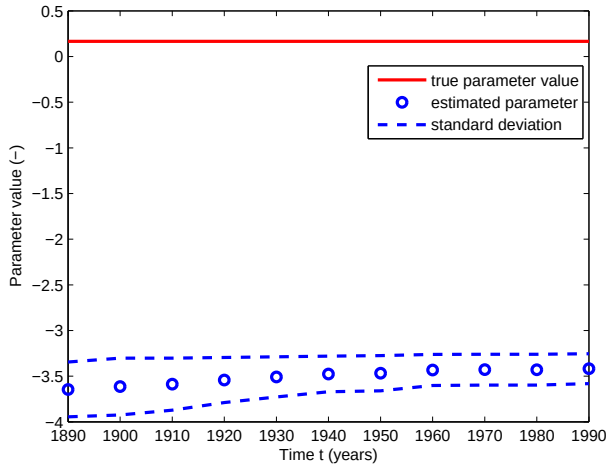
Although the IEnKF algorithm has demonstrated its efficiency in such a test, we can get faster convergence by taking larger value of measurement noise covariance, e.g. $\mathbf{R} = 10.0$. the increase of measurement error covariance implies keeping the spread of the ensemble sufficient as more data is assimilated. Then only two global iterations are needed to obtain an accurate parameter estimation which equals 0.1680 (see Figure 5.16).

The other issue is concerned with evaluating our believes in the quality of initial guess related to the model parameter. We consider initial guess of the model parameter to be $\mathcal{N}(2.0, 0.01)$ or $\mathcal{N}(2.0, 4)$, and use typical measurement noise of$\mathbf{R} = 0.01$. Therefore in the first case we are quite certain about the guessed parameter value, whereas in the second one our uncertainty is much higher. It turns out that within the first experimental framework an IEnKF algorithm provides an accurate estimation of the model parameter by means of 16 iterations. In the second test filter demonstrates divergent behavior (see Figure 5.17 for the first few global iterations). The divergence of the filter in such a case should not be surprising, since initially linear problem (5.3)–(5.4) becomes non-linear due to involvement of the model parameter into state vector. In general, for the case of non-linear filtering problem, there is no guarantee that the filter performance is satisfactory and the filter does not diverge.

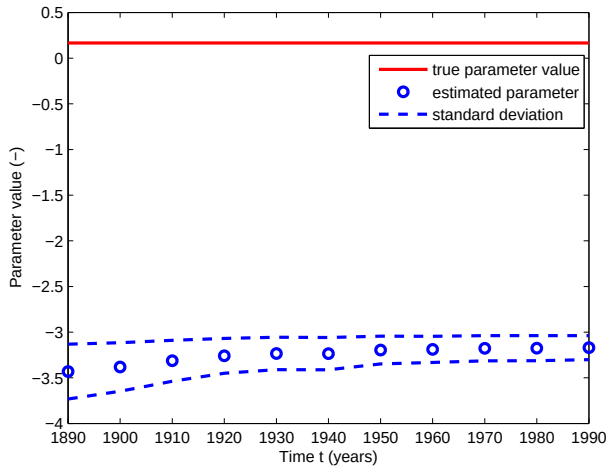
Thus, IEnKF technique is applicable to solving parameter estimation problems and does demonstrate some features superior to classical EnKF algorithm in the case of various experiments with the global carbon-dioxide model.



(a)

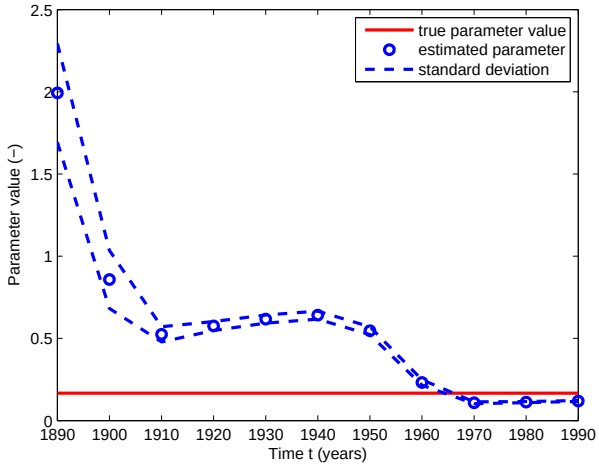


(b)

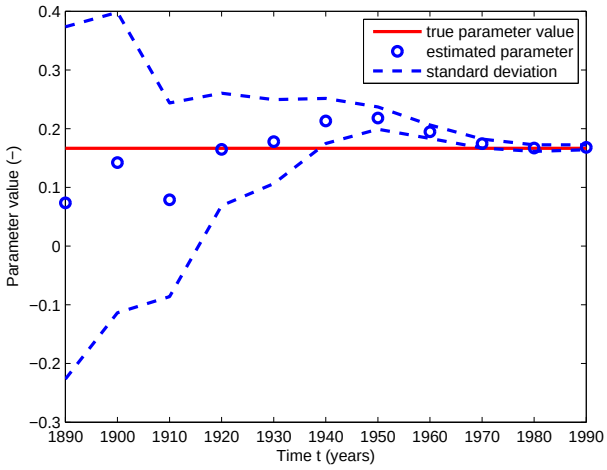


(c)

Figure 5.15: CO2 model: First iterations of IEnKF parameter estimation with initial guess normally distributed with parameters $(2.0, 0.09)$ and covariance of measurement noise $\mathbf{R} = 0.01$

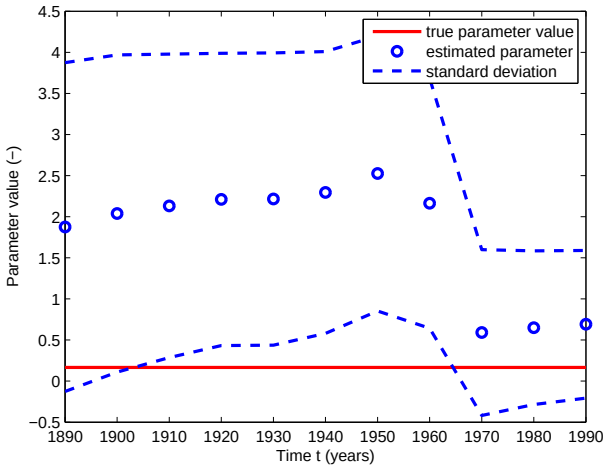


(a)

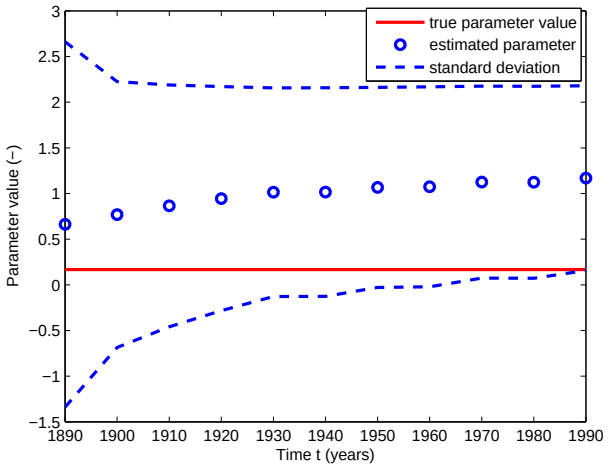


(b)

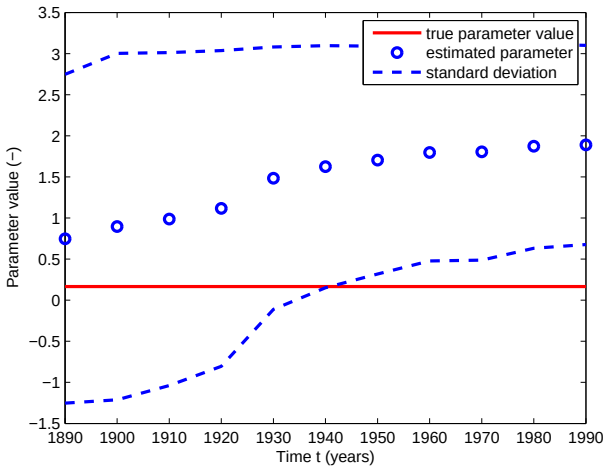
Figure 5.16: CO2 model: IEnKF parameter estimation with initial guess normally distributed with parameters $(2.0, 0.09)$ and covariance of measurement noise $\mathbf{R} = 10.0$



(a)



(b)



(c)

Figure 5.17: CO2 model: First iterations of IEnKF parameter estimation with initial guess normally distributed with parameters $(2.0, 4.0)$ and covariance of measurement noise $\mathbf{R} = 0.01$

5.3.5 History matching via IEnKF

Let us come back to estimating parameters of the two-phase two-dimensional fluid model. We proceed by running IEnKF algorithm for the trial example which is discussed throughout Section 5. In fact we accomplish the second global iteration of EnKF method. Space averaged RMS errors are plotted in time (see Figure 5.18) to evaluate the quality of estimating the model parameter. The graph demonstrates improvement for neither parameter es-

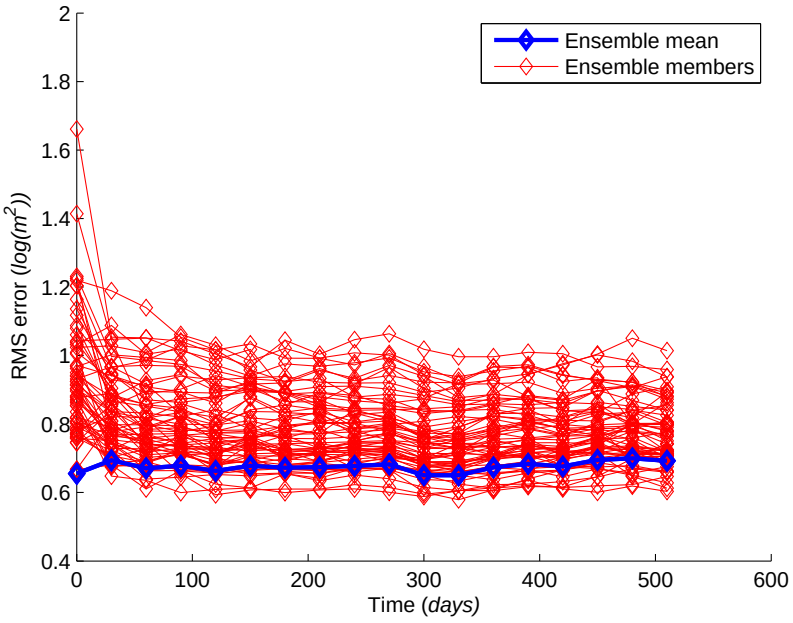
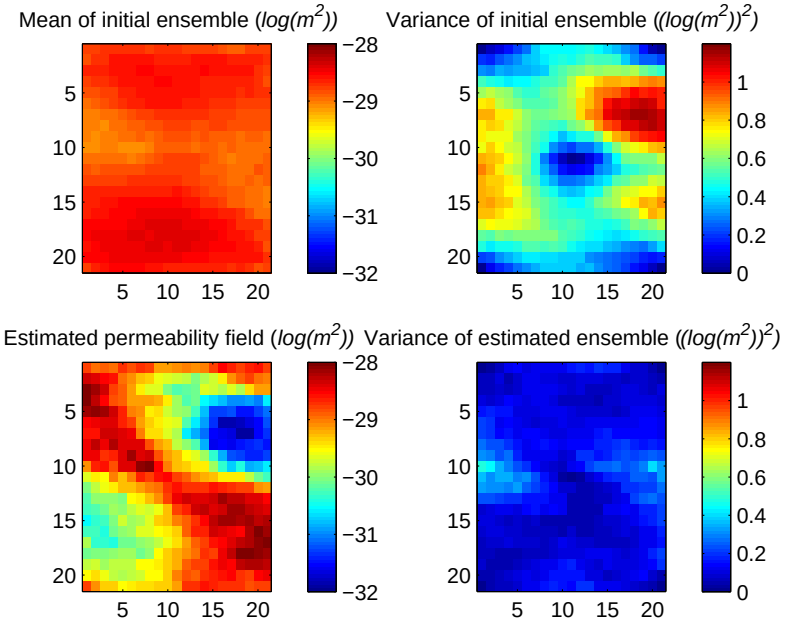
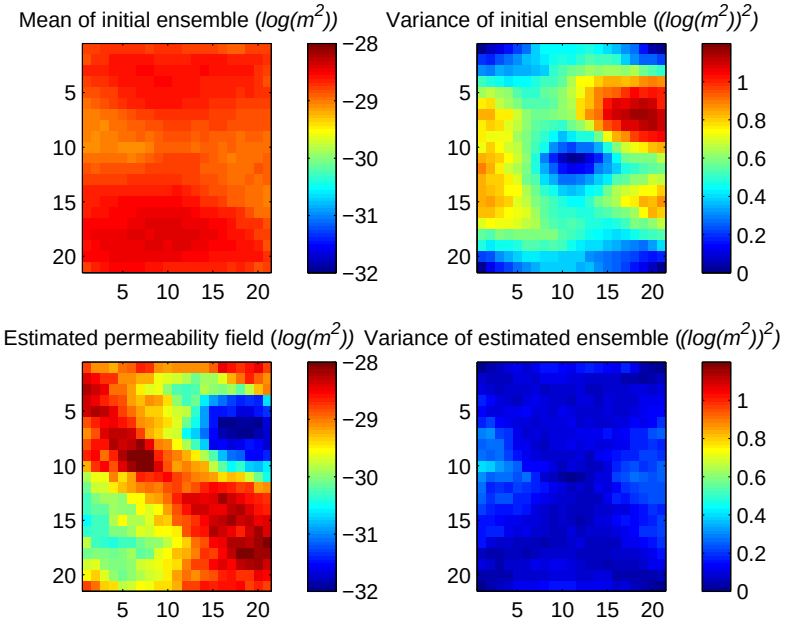


Figure 5.18: IEnKF: RMS error in estimated permeability vs time

timation nor uncertainty characterization, which can be expected since the first EnKF iteration does not provide reducing the parameter estimation error in later times and actually gives us relatively accurate estimate. Indeed, there is almost no visual difference between permeability fields obtained with EnKF and IEnKF algorithms (compare Figures 5.19(a) and 5.19(b)).



(a) EnKF algorithm



(b) IEnKF algorithm

Figure 5.19: IEnKF: RMS error for estimated permeability vs time

Consider now situation when a priori information on the values of model parameters is far from real. For that purpose we take the initial ensemble of log-permeability fields and shift each member of it by adding a vector $5 * \mathbf{I}_{shift}$, where shifting vector $\mathbf{I}_{shift}$ consists of ones and $\mathbf{I}_{shift} \in \mathbf{R}^{1\times 441}$. Note that such a shift does not affect the variance statistics, hence, the structure of initial ensemble is kept. The data assimilation is performed from time $t_0 = 0(days)$ up to time moment $t_{stop} = 300(days)$ for these initial parameters within usual experimental environment. It turns out that EnKF method allows obtaining some improved, although not yet enough accurate, estimation of the parameters, whereas IEnKF algorithm diverges.

At this point we try to benefit from the analysis of the trial CO2 model and check whether a choice of larger measurement error can provide better estimations. We now upscale the measurement noise covariance matrix $\mathbf{R}$ with a factor 10^4. Such parameters for data assimilation indeed allow the IEnKF algorithm to demonstrate its features by some reducing the error in estimation (see Figure 5.20).

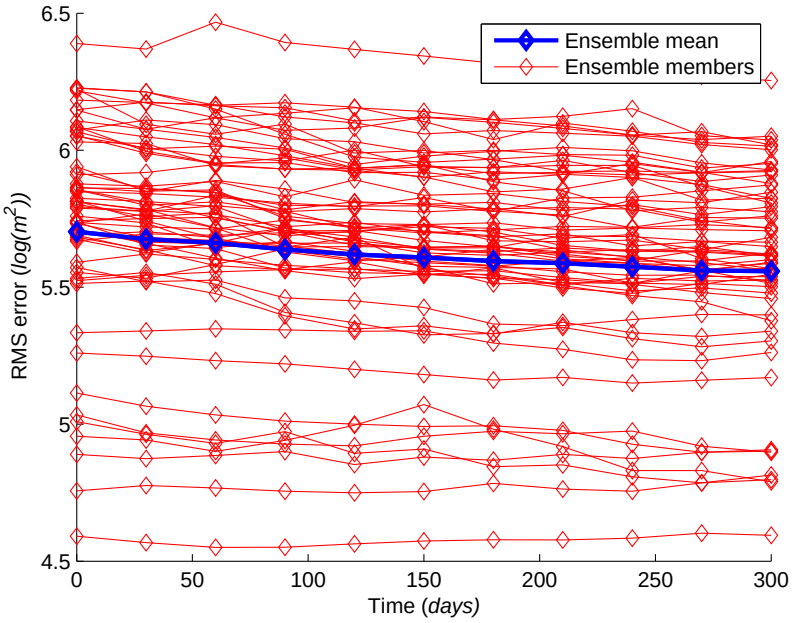
The same observation can be made based on visual comparison of Figures 5.21(a) and 5.21(b). Note that although regularly providing overestimated values, the filter tends to capture the structure of true permeability field. This happens because the ensemble of permeability fields, used for the current test, is only shifted version of the one previously used for investigations. Such an ensemble contains some information on the field structure which simple shifting does not affect, since a shift changes the ensemble mean and not the covariance. The given initial statistics cannot be changed, because it comes from the statistics of ensemble population. Thus the possibilities of improving parameter estimation by varying statistics of initially guessed values of model parameter are in certain sense restricted.

We proceed by the more representative example of IEnKF usage within reservoir engineering framework. Consider now the initial ensemble of log-permeability fields being shifted by vector $0.5 * \mathbf{I}_{shift}$. The data assimilation is performed from time $t_0 = 0(days)$ up to time moment $t_{stop} = 510(days)$ and the covariance of measurement error is scaled by the factor of 10^2 to prevent filter divergence. Such parameters for data assimilation allow some reducing the error in estimation of permeability values performed via (see Figure 5.22). Indeed, we obtain a permeability field with a structure resembling the true one, although some overestimating the values corresponding to low permeability areas of the field (see Figure 5.23). The parameter values corresponding to these areas are in particular improved after global iteration (compare Figures 5.23(a) and 5.23(b)). The difference between

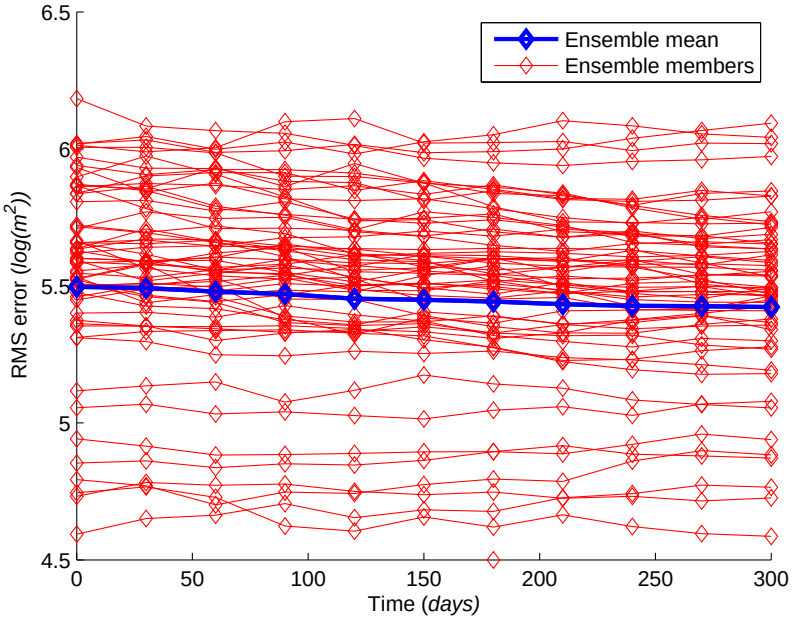
the bottom right charts in Figures 5.23(a) and 5.23(b) indicates reduction of the variance and therefore uncertainty in the estimation.

Although demonstrating a usage of IEnKF approach to estimating permeability values, the performed tests rise up additional problems to be solved. The list of such problems includes finding criteria to evaluate whether global filter iteration is needed in real case, when no true permeability values are available. We suppose that one may consider the RMS differences between the parameter values obtained at two sequential data assimilation steps. Another issue is concerned with determining appropriate error statistics that can have a great impact on improvement of the estimations and the number of global EnKF iterations needed for that purpose.

Summarizing, we may assert that history matching on the basis of IEnKF technique has demonstrated its efficiency for improving model parameter estimation.

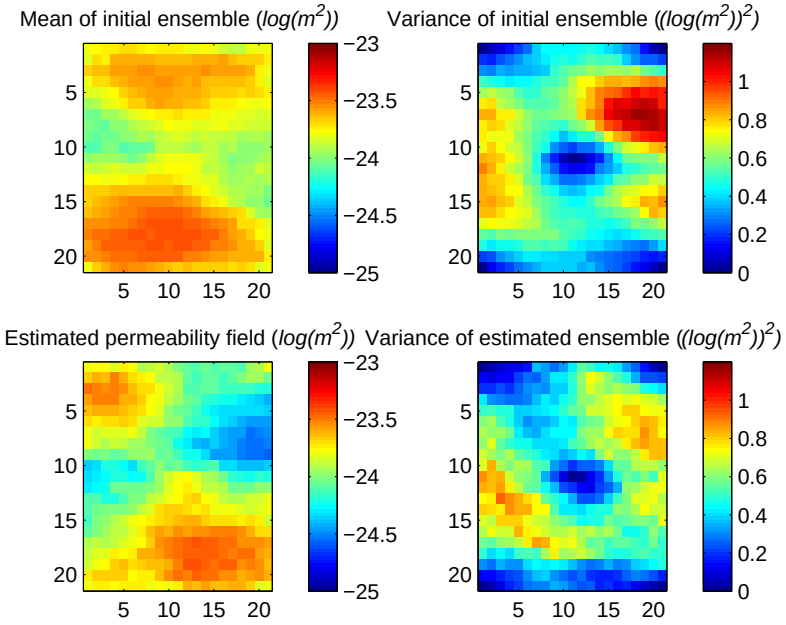


(a) First iteration

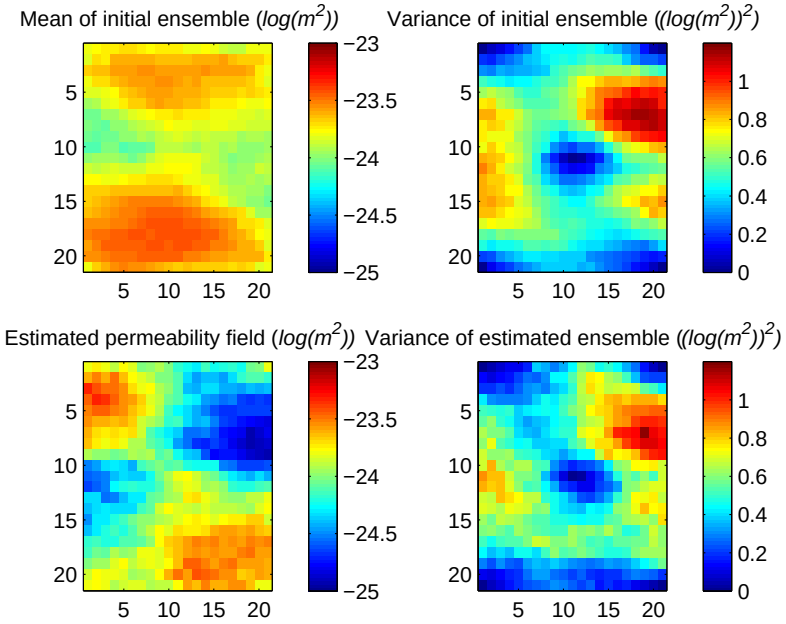


(b) Second iteration

Figure 5.20: IEnKF: RMS error for estimated permeability vs time (shifted initial ensemble with $5 * \mathbf{I}_{shift}$ and measurement error covariance matrix $10^4 * \mathbf{R}$ are used in experiment)

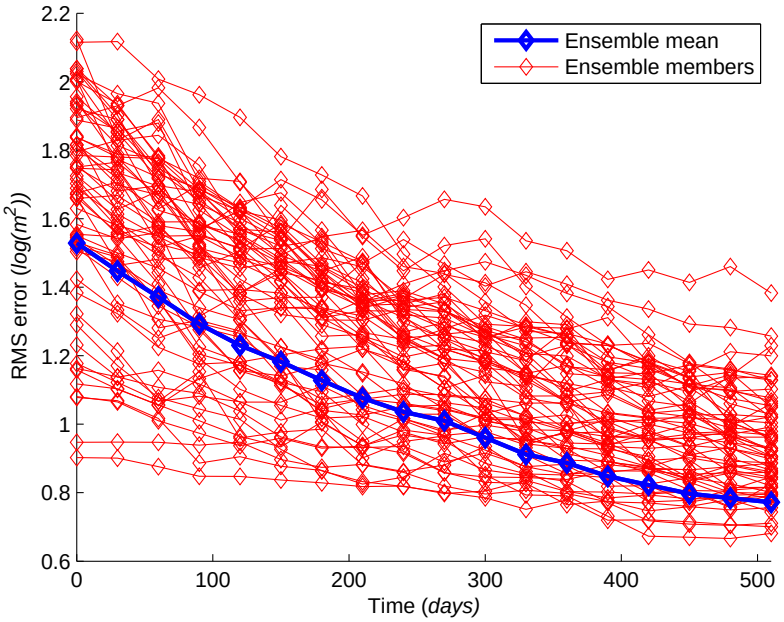


(a) First iteration

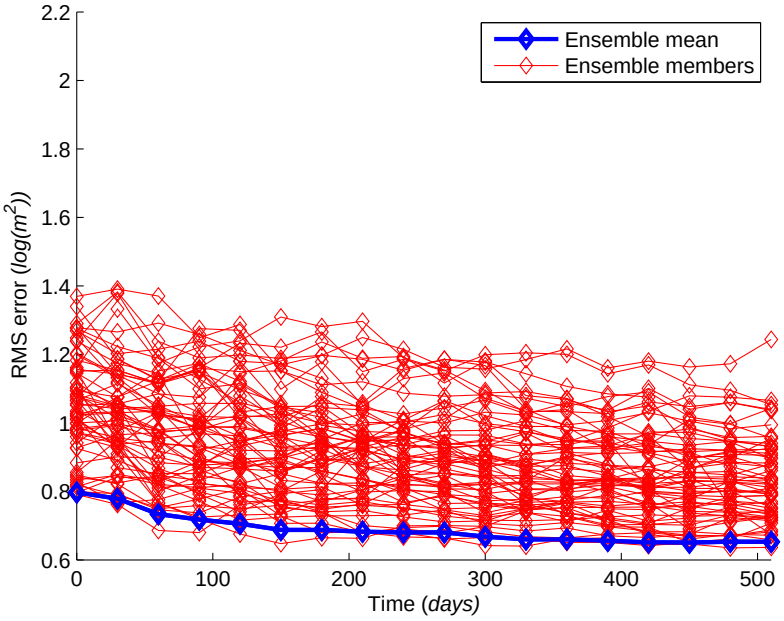


(b) Second iteration

Figure 5.21: IEnKF: Initial and estimated permeability fields and corresponding variances (shifted initial ensemble with $5 * \mathbf{I}_{shift}$ and measurement error covariance matrix $10^4 * \mathbf{R}$ are used in experiment)

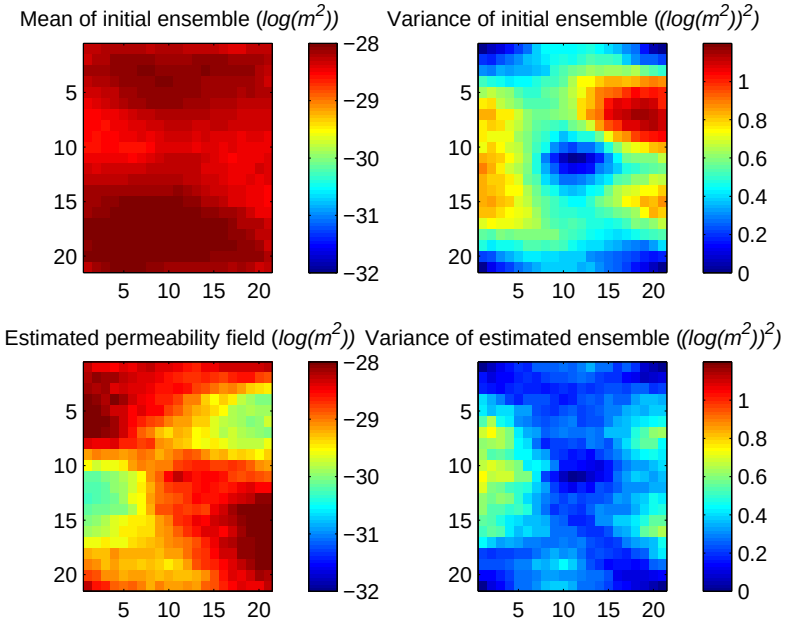


(a) First iteration

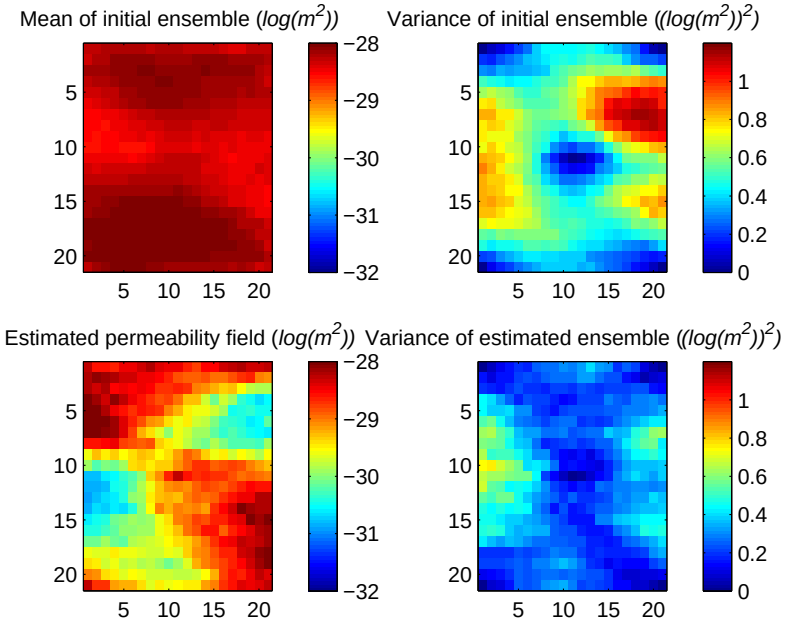


(b) Second iteration

Figure 5.22: IEnKF: RMS error for estimated permeability vs time (shifted initial ensemble with $0.5 * \mathbf{I}_{shift}$ and measurement error covariance matrix $10^2 * \mathbf{R}$ are used in experiment)



(a) First iteration



(b) Second iteration

Figure 5.23: IEnKF: Initial and estimated permeability fields and corresponding variances (shifted initial ensemble with $5 * \mathbf{I}_{shift}$ and measurement error covariance matrix $10^2 * \mathbf{R}$ are used in experiment)

Chapter 6

Conclusion

The study has been focused on the analysis of the usage and applicability of ensemble Kalman filtering techniques with respect to history matching stage of reservoir simulation process. In turn history matching is known to be very important for the whole reservoir investigation as it results in the calibrated model which can be later used to predict reservoir performance.

Summarizing the results obtained within the current research we can formulate some general conclusions.

First, the success in creating a basis for reservoir simulation methodology depends on deriving an appropriate model which reflects the knowledge of reservoir and fluid physical properties. Then the above model has to be discretized. With the model in hand one can perform a history matching process, which aims at adapting the model parameters to match the computed reservoir outcome quantities and the real observations.

There are two basic approaches to history matching: manual and automatic. In the latest years great research effort has been devoted to automatic history match, which demonstrates a potential to decrease time expenses and provide more accurate estimates of the model parameters. However, traditional automatic history matching approaches are either limited to the small-scaled and simple reservoir models or inefficient in terms of computational costs. Moreover, there rises up a problem of continuous real-time model updating.

The mentioned obstacles are tried to be overcome by the use of Kalman filtering techniques and especially ensemble Kalman filter method which is easy for implementation and computationally efficient. Meanwhile it turns out that application of EnKF to the problem of estimating parameters re-

lated to reservoir engineering model is not straightforward.

As the model describes a real physical process, the state vector obtained at each time step via data assimilation procedure has to be feasible. Nothing in Kalman filtering methodology guarantees that such a requirement is satisfied. Hence, one needs to make sure that he/she gets physically reasonable estimates. The confirming EnKF algorithm used in the current study just allows taking this circumstance into account by introducing additional 'confirmation' step into the body of each data assimilation routine [25]. However, this is not very efficient, since inclusion of additional step results in doubling computational costs.

Another issue is concerned with performing the data assimilation step, since the results of operations on matrices which elements significantly differ in scale are very sensitive to the accuracy of initial data. This feature may cause the critical errors in the estimation of the state vector and final divergence of the filter. We make use of specially introduced 'scaling' matrices to solve the problem.

Following the idea presented in [16], there is proposed an iterative modification of ensemble Kalman filter. Tested on trial global carbon-dioxide model, such an algorithm demonstrates superior features comparing to classical EnKF approach for some particular instances, when a priori knowledge of the possible parameter values is far from reality.

To investigate the perspectives opened by these ensemble Kalman filtering algorithms we have performed the investigations based on the use of in-house reservoir simulator, which provides a forward integration of two-phase (water-oil) two-dimensional fluid flow model. The accomplished case study has confirmed the usefulness of EnKF technique for solving the history matching problem and estimating reservoir model parameter. The experiments clearly indicate the necessity to find a proper model parameter value for performing further forecast of reservoir behavior. There might occur problems at which EnKF algorithm does not provide results of sufficient accuracy. An appropriate use of IEnKF method in such a case can improve the estimations.

Finally we conclude that EnKF methodology and its special modification (iterative EnKF algorithm) have a promising future as the powerful tools for solving the important problems related to reservoir engineering. The progress in this area can be definitely expected.

Bibliography

[1] K. Aziz and A. Settari. *Petroleum Reservoir Simulation*. Applied Science Publishers LTD, London, 1979. [cited at p. 4, 11, 86]

[2] J. Baird and C. Dawson. The representer method for data assimilation in single-phase darcy flow in porous media. *Computational Geosciences*, 9:247–271, 2005. [cited at p. 6]

[3] G. Burgers, P. Leeuwen, and G. Evensen. Analysis scheme in the ensemble kalman filter. *Monthly Weather Review*, 126:1719–1724, 1998. [cited at p. 26]

[4] G. Chavent, M. Dupuy, and P. Lemonnier. History matching by use of optimal control theory. *SPE Journal*, 15:74–86, 1975. [cited at p. 6]

[5] Z. Chen, G. Huan, and Y. Ma. *Computational Methods for Multiphase Flows in Porous Media*. Society for Industrial and Applied Mathematics, Philadelphia, 2006. [cited at p. 2, 3, 11, 13]

[6] M. Dueker. Kalman filtering with truncated normal state variables for bayesian estimation of macroeconomic models. Unpublished, March 2006. [cited at p. 45]

[7] T. Ertekin, J. H. Abou-Kassen, and G. R. King. *Basic Applied Reservoir Simulation*. Society of Petroleum Engineers, Richardson, 2001. [cited at p. 1, 3, 5, 11, 18, 90]

[8] J. R. Fanchi. *Principles of Applied Reservoir Simulation*. Gulf Publishing Company, Houston, 1997. [cited at p. 5]

[9] P. Goovaerts. *Geostatistics for Natural Resources Evaluation*. Oxford University Press, New York, 1997. [cited at p. 44]

[10] Y. Gu and D. S. Oliver. History matching of the punq-s3 reservoir model using the ensemble kalman filter. SPE 89942. SPE Annual Technical Conference and Exhibition, 2004. [cited at p. 7]

[11] Y. Gu and D. S. Oliver. The ensemble kalman filter for continuous updating of reservoir simulation models. *Journal of Energy Resources Technology*, 128:79–87, 2006. [cited at p. 7, 44, 45]

[12] Y. Gu and D. S. Oliver. An iterative ensemble kalman filter for multiphase fluid flow data assimilation. submitted to SPE Journal, 2007. [cited at p. 28]

[13] A. W. Heemink. Ch.5: Data assimilation methods. Unpublished. [cited at p. 22, 25, 27, 50, 61]

[14] P. Jacquard and C. Jain. Permeability distribution from field pressure data. *SPE Journal*, 5:281–294, 1965. [cited at p. 6]

[15] J. D. Jansen. Systems theory for reservoir management. Unpublished, February 2007. [cited at p. 83, 85, 86, 90]

[16] A. H. Jazwinski. *Stochastic Processes and Filtering Theory*. Academic Press, New York and London, 1970. [cited at p. 28, 76]

[17] R. J. Lorentzen, G. Nævdal, B. Vallès, and A. M. Berg. Analysis of the ensemble kalman filter for estimation of permeability and porosity in reservoir models. SPE 96375. SPE Annual Technical Conference and Exhibition, 2005. [cited at p. 7]

[18] C. Maschio and D. J. Schiozer. Assisted history matching using streamline simulation. *Petroleum Science and Technology*, 23:761–774, 2005. [cited at p. 6]

[19] J. K. Przybysz-Jarnut, R. G. Hanea, J. D. Jansen, and A. W. Heemink. Application of the representer method for parameter estimation in numerical reservoir models. *Computational Geosciences*, pages 73–85, 2007. [cited at p. 6]

[20] L. Ruijian, A. C. Reynolds, and D. S. Oliver. History matching of three-phase flow production data. *SPE Journal*, December 2003. [cited at p. 7]

[21] D. Simon. Kalman filtering. *Embedded Systems Programming*, 2001. [cited at p. 21]

[22] D. Simon and T. L. Chia. Kalman filtering with state equality constraints. Unpublished. [cited at p. 45]

[23] D. Simon and D.L. Simon. Kalman filtering with inequality constraints for turbofan engine health estimation. Tm2003-212111, NASA, 2003. [cited at p. 45]

[24] S. Strebelle. Conditional simulation of complex geological structures using multiple-point statistics. *Mathematical Geology*, 34:1–21. [cited at p. 35]

[25] X.-H. Wen and W. C. Chen. Real-time reservoir model updating using ensemble kalman filter. SPE 92991. SPE Reservoir Simulation Symposium, 2005. [cited at p. 7, 28, 45, 76]

[26] M. Zafari, G. Li, and A. C. Reynolds. Iterative forms of the ensemble kalman filter. 10th European Conference of the Mathematics of Oil Recovery, 2006. [cited at p. 28, 44]

[27] M. Zafari and A. C. Reynolds. Assessing the uncertainty in reservoir description and performance predicitions with the ensemble kalman filter. SPE 95750. SPE Annual Technical Conference and Exhibition, 2005. [cited at p. 8, 28, 44]

Appendices

Appendix A

Two-Phase Two-Dimensional Fluid Flow Model

This section addresses deriving the governing PDEs for two-phase (water-oil) two-dimensional fluid flow model and presenting the above equations in a discrete form. Such a goal is achieved with the help of the the guidelines given in [15].

A.1 Governing equations

Consider a two-phase water-oil fluid under isothermal conditions. This means that FVFs are not required for deriving the governing PDEs. We formulate the equations in terms of in-situ volumes.

In the case of two-phase flow the mass balance equations (2.7) can be expressed for each phase as

$$\nabla \cdot (A\rho_w \mathbf{v}_w) + A\frac{\partial (\rho_w \phi S_w)}{\partial t} - A\rho_w \bar{q}_w = 0, \tag{A.1}$$

$$\nabla \cdot (A\rho_o \mathbf{v}_o) + A\frac{\partial (\rho_o \phi S_o)}{\partial t} - A\rho_o \bar{q}_o = 0 \tag{A.2}$$

and differential representation of Darcy's law for the simultaneous flow of

more than one phase is (2.8), which in our situation takes the form

$$\mathbf{v}_w = -\frac{k_{rw}}{\mu_w}\mathbf{k}\left(\nabla p_w - \rho_w g \nabla d\right), \tag{A.3}$$

$$\mathbf{v}_o = -\frac{k_{rw}}{\mu_o}\mathbf{k}\left(\nabla p_o - \rho_o g \nabla d\right). \tag{A.4}$$

Substituting expressions (2.4), (2.5), (A.3) and (A.4) into (A.1) and (A.2) we obtain

$$-\nabla\cdot\left(A\rho_w\frac{k_{rw}}{\mu_w}\mathbf{k}\left[\left(\nabla p_o - \frac{\partial p_c}{\partial S_w}\nabla S_w\right) - \rho_w g\nabla d\right]\right) + A\frac{\partial\left(\rho_w\phi S_w\right)}{\partial t} - A\rho_w\bar{q}_w = 0, \tag{A.5}$$

$$-\nabla\cdot\left(A\rho_o\frac{k_{ro}}{\mu_o}\mathbf{k}\left[\nabla p_o - \rho_o g\nabla d\right]\right) + A\frac{\partial\left(\rho_o\phi(1-S_w)\right)}{\partial t} - A\rho_o\bar{q}_o = 0. \tag{A.6}$$

The term $\nabla\cdot\left(A\rho_w\frac{k_{rw}}{\mu_w}\mathbf{k}\frac{\partial p_c}{\partial S_w}\nabla S_w\right)$ in equation (A.5) reflects the nonlinear diffusion effect caused by the capillary pressure.

Let us investigate the accumulation term in equations (A.5) and (A.6) on the basis of expressions (2.1), (2.2) and (2.3):

$$\begin{aligned}\frac{\partial\left(\rho_w\phi S_w\right)}{\partial t} &= \frac{\partial\rho_w}{\partial t}\phi S_w + \rho_w\frac{\partial\phi}{\partial t}S_w + \rho_w\phi\frac{\partial S_w}{\partial t}\\ &= \frac{\partial\rho_w}{\partial p_o}\frac{\partial p_o}{\partial t}\phi S_w + \rho_w\frac{\partial\phi}{\partial p_o}\frac{\partial p_o}{\partial t}S_w + \rho_w\phi\frac{\partial S_w}{\partial t}\\ &= \rho_w\phi\left(S_w(c_w+c_R)\frac{\partial p_o}{\partial t} + \frac{\partial S_w}{\partial t}\right),\end{aligned} \tag{A.7}$$

$$\begin{aligned}\frac{\partial\left(\rho_o\phi(1-S_w)\right)}{\partial t} &= \frac{\partial\rho_o}{\partial t}\phi\left(1-S_w\right) + \rho_o\frac{\partial\phi}{\partial t}\left(1-S_w\right) - \rho_o\phi\frac{\partial S_w}{\partial t}\\ &= \frac{\partial\rho_o}{\partial p_o}\frac{\partial p_o}{\partial t}\phi\left(1-S_w\right) + \rho_o\frac{\partial\phi}{\partial p_o}\frac{\partial p_o}{\partial t}\left(1-S_w\right) - \rho_o\phi\frac{\partial S_w}{\partial t}\\ &= \rho_o\phi\left(\left(1-S_w\right)(c_o+c_R)\frac{\partial p_o}{\partial t} - \frac{\partial S_w}{\partial t}\right).\end{aligned} \tag{A.8}$$

During water flooding on reservoir scale the dispersion caused by geological heterogeneities is usually much stronger than the diffusion caused by capillary pressures. Moreover, solving the discretized equations numerically often results in the presence of numerical dispersion. This numerical dispersion is of the same order or even larger than the dispersion and diffusion caused by physical phenomena. With respect to the above discussion and following the idea described in [15] we neglect capillary forces and dispersion.

This assumption and the use of expressions (A.7) and (A.8) allow simplifying the equations (A.5) and (A.6) to the form

$$-\nabla\cdot\left(A\rho_w\frac{k_{rw}}{\mu_w}\mathbf{k}\left[\nabla p-\rho_w g\nabla d\right]\right) + A\rho_w\phi\left[S_w(c_w+c_R)\frac{\partial p}{\partial t}+\frac{\partial S_w}{\partial t}\right]-A\rho_w\bar{q}_w=0, \quad \text{(A.9)}$$

$$-\nabla\cdot\left(A\rho_o\frac{k_{ro}}{\mu_o}\mathbf{k}\left[\nabla p-\rho_o g\nabla d\right]\right) + A\rho_o\phi\left[(1-S_w)(c_o+c_R)\frac{\partial p}{\partial t}-\frac{\partial S_w}{\partial t}\right]-A\rho_o\bar{q}_o=0, \quad \text{(A.10)}$$

where the subscript 'o' for the pressure is omitted, since the absence of capillary pressure and (2.5) imply $p_o = p_w$.

Assuming isotropic permeability, pressure independence of the parameters and absence of gravitational forces, we can rewrite equations (A.9) and (A.10) in scalar two-dimensional form

$$-\frac{A}{\mu_w}\left[\frac{\partial}{\partial x}\left(kk_{rw}\frac{\partial p}{\partial x}\right)+\frac{\partial}{\partial y}\left(kk_{rw}\frac{\partial p}{\partial y}\right)\right] + A\left[\phi S_w(c_w+c_R)\frac{\partial p}{\partial t}+\frac{\partial S_w}{\partial t}\right]-A\bar{q}_w=0, \quad \text{(A.11)}$$

$$-\frac{A}{\mu_o}\left[\frac{\partial}{\partial x}\left(kk_{ro}\frac{\partial p}{\partial x}\right)+\frac{\partial}{\partial y}\left(kk_{ro}\frac{\partial p}{\partial y}\right)\right] + A\left[\phi\,(1-S_w)\,(c_o+c_R)\frac{\partial p}{\partial t}-\frac{\partial S_w}{\partial t}\right]-A\bar{q}_o=0. \quad \text{(A.12)}$$

The model (A.11)–(A.12) describes the two-phase water-oil fluid flow. Now

the numerical solution of the model can be obtained by the finite difference approach.

A.2 Model discretization

In the current study we follow the most natural approach to dealing with (A.11)–(A.12) by solving equations simultaneously. This is the so-called simultaneous solution method as described originally in [1]. However, we mainly use the guidelines sketched in [15].

Let us start by discretizing the first term in equation (A.11) as

$$\frac{A}{\mu_w}\frac{\partial}{\partial x}\left(kk_{rw}\frac{\partial p}{\partial x}\right) \approx \frac{A}{\mu_w}\frac{(kk_{rw})_{i+\frac{1}{2},j}\,(p_{i+1,j}-p_{i,j})-(kk_{rw})_{i-\frac{1}{2},j}\,(p_{i,j}-p_{i-1,j})}{(\Delta x)^2}, \quad \text{(A.13)}$$

where absolute permeability k is computed through harmonic averages

$$k_{i-\frac{1}{2},j}=\frac{2}{\frac{1}{k_{i-1,j}}+\frac{1}{k_{i,j}}}$$

and relative permeability k_{rw} is obtained with the help of upstream weighting [1]:

$$(k_{rw})_{i+\frac{1}{2},j}=\begin{cases}(k_{rw})_{i,j}\,, & \text{if } p_{i,j}\geq p_{i+1,j};\\ (k_{rw})_{i+1,j}\,, & \text{if } p_{i,j}<p_{i+1,j}.\end{cases}$$

The second term in equation (A.11) can be rewritten in a similar manner as

$$\frac{A}{\mu_w}\frac{\partial}{\partial y}\left(kk_{rw}\frac{\partial p}{\partial y}\right) \approx \frac{A}{\mu_w}\frac{(kk_{rw})_{i,j+\frac{1}{2}}\,(p_{i,j+1}-p_{i,j})-(kk_{rw})_{i,j-\frac{1}{2}}\,(p_{i,j}-p_{i,j-1})}{(\Delta y)^2}. \quad \text{(A.14)}$$

Combining terms (A.13) and (A.14) results in the following discretization

of (A.11):

$$V\left[\phi S_w(c_w+c_R)\frac{\partial p}{\partial t}+\frac{\partial S_w}{\partial t}\right]_{i,j}-(T_w)_{i-\frac{1}{2},j}\,p_{i-1,j}-(T_w)_{i,j-\frac{1}{2}}\,p_{i,j-1}$$
$$+\left[(T_w)_{i-\frac{1}{2},j}+(T_w)_{i,j-\frac{1}{2}}+(T_w)_{i,j+\frac{1}{2}}+(T_w)_{i+\frac{1}{2},j}\right]p_{i,j}$$
$$-(T_w)_{i,j+\frac{1}{2}}\,p_{i,j+1}-(T_w)_{i+\frac{1}{2},j}\,p_{i+1,j}=V\,(\bar{q}_w)_{i,j}\,, \quad \text{(A.15)}$$

where transmissibility $(T_w)_{i-\frac{1}{2},j}$ denotes the following term

$$(T_w)_{i-\frac{1}{2},j}=\frac{\Delta y}{\Delta x}\frac{A}{\mu_w}\,(kk_{rw})_{i-\frac{1}{2},j}$$

and $(T_w)_{i,j-\frac{1}{2}}$ stays for

$$(T_w)_{i,j-\frac{1}{2}}=\frac{\Delta x}{\Delta y}\frac{A}{\mu_w}\,(kk_{rw})_{i,j-\frac{1}{2}}\,.$$

Analogously to the (A.15) we obtain a discretized version of equation (A.12):

$$V\left[\phi(1-S_w)(c_o+c_R)\frac{\partial p}{\partial t}-\frac{\partial S_w}{\partial t}\right]_{i,j}-(T_o)_{i-\frac{1}{2},j}\,p_{i-1,j}-(T_o)_{i,j-\frac{1}{2}}\,p_{i,j-1}$$
$$+\left[(T_o)_{i-\frac{1}{2},j}+(T_o)_{i,j-\frac{1}{2}}+(T_o)_{i,j+\frac{1}{2}}+(T_o)_{i+\frac{1}{2},j}\right]p_{i,j}$$
$$-(T_o)_{i,j+\frac{1}{2}}\,p_{i,j+1}-(T_o)_{i+\frac{1}{2},j}\,p_{i+1,j}=V\,(\bar{q}_o)_{i,j}\,. \quad \text{(A.16)}$$

Now let us combine equations (A.15) and (A.16) in a matrix form:

$$\begin{bmatrix}\mathbf{V}_{wp} & \mathbf{V}_{ws}\\ \mathbf{V}_{op} & \mathbf{V}_{os}\end{bmatrix}\begin{bmatrix}\dot{\mathbf{p}}\\ \dot{\mathbf{S}}\end{bmatrix}+\begin{bmatrix}\mathbf{T}_w & \mathbf{0}\\ \mathbf{T}_o & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{p}\\ \mathbf{S}\end{bmatrix}=\begin{bmatrix}\mathbf{q}_w\\ \mathbf{q}_o\end{bmatrix}, \quad \text{(A.17)}$$

where vector $\mathbf{p}$ contains pressures

$$\mathbf{p}^T=[p_{i-1,j}\ldots p_{i,j-1}\ p_{i,j}\ p_{i,j+1}\ldots p_{i+1,j}],$$

vector $\mathbf{S}$ includes water saturations

$$\mathbf{S}^T=[(S_w)_{i-1,j}\ldots(S_w)_{i,j-1}\ \ (S_w)_{i,j}\ \ (S_w)_{i,j+1}\ldots(S_w)_{i+1,j}],$$

sub-matrices $\mathbf{V}_{wp},\ \mathbf{V}_{ws},\ \mathbf{V}_{op},\ \mathbf{V}_{os}$ consist of accumulation terms

$$\mathbf{V}_{wp} = V\phi(c_w + c_R)[0\dots0\ \ (S_w)_{i,j}\ \ 0\dots0],\ \mathbf{V}_{ws} = V\phi[0\dots0\ 1\ 0\dots0],$$

$$\mathbf{V}_{op} = V\phi(c_o + c_R)[0\dots0\ \ (1-S_w)_{i,j}\ \ 0\dots0],\ \mathbf{V}_{os} = -V\phi[0\dots0\ 1\ 0\dots0],$$

sub-matrices $\mathbf{T}_w$ and $\mathbf{T}_o$ are made of transmissibility terms

$$\mathbf{T}_w = \Big[-(T_w)_{i-\frac{1}{2},j}\cdots-(T_w)_{i,j-\frac{1}{2}}\ \ \Big((T_w)_{i-\frac{1}{2},j}+(T_w)_{i,j-\frac{1}{2}}+(T_w)_{i,j+\frac{1}{2}}+(T_w)_{i+\frac{1}{2},j}\Big) - (T_w)_{i,j+\frac{1}{2}}\cdots-(T_w)_{i+\frac{1}{2},j}\Big],$$

$$\mathbf{T}_o = \Big[-(T_o)_{i-\frac{1}{2},j}\cdots-(T_o)_{i,j-\frac{1}{2}}\ \ \Big((T_o)_{i-\frac{1}{2},j}+(T_o)_{i,j-\frac{1}{2}}+(T_o)_{i,j+\frac{1}{2}}+(T_o)_{i+\frac{1}{2},j}\Big) - (T_o)_{i,j+\frac{1}{2}}\cdots-(T_o)_{i+\frac{1}{2},j}\Big],$$

and vectors $\mathbf{q}_w$ and $\mathbf{q}_o$ contain the flow rates (source terms)

$$\mathbf{q}_w^T = V[(\bar{q}_w)_{i-1,j}\dots(\bar{q}_w)_{i,j-1}\ \ (\bar{q}_w)_{i,j}\ \ (\bar{q}_w)_{i,j+1}\dots(\bar{q}_w)_{i+1,j}],$$

$$\mathbf{q}_o^T = V[(\bar{q}_o)_{i-1,j}\dots(\bar{q}_o)_{i,j-1}\ \ (\bar{q}_o)_{i,j}\ \ (\bar{q}_o)_{i,j+1}\dots(\bar{q}_o)_{i+1,j}].$$

Actually the terms $q_o = V\bar{q}_o$ and $q_w = V\bar{q}_w$ cannot be always prescribed directly. Normally it is possible to determine a total flow rate q_t and obtain oil and water rates through the fractional flows $f_o = \dfrac{\lambda_o}{\lambda_o+\lambda_w}$ and $f_w = \dfrac{\lambda_w}{\lambda_o+\lambda_w}$ as

$$q_o = f_o q_t,\ q_w = f_w q_t.$$

Hence the vector of source terms has the following form

$$\begin{bmatrix}\mathbf{q}_w\\ \mathbf{q}_o\end{bmatrix} = \begin{bmatrix}\mathbf{F}_w\\ \mathbf{F}_o\end{bmatrix}\mathbf{q}_t,$$

where $\mathbf{q}_t$ is the vector of total rates

$$\mathbf{q}_t^T = [(q_t)_{i-1,j}\dots(q_t)_{i,j-1}\ \ (q_t)_{i,j}\ \ (q_t)_{i,j+1}\dots(q_t)_{i+1,j}],$$

sub-matrices $\mathbf{F}_w$ and $\mathbf{F}_o$ are diagonal with non-zero entries containing frac-

tional flows, namely,

$$\mathbf{F}_w = [0 \dots 0 \;\; (f_w)_{i,j} \;\; 0 \dots 0],$$

$$\mathbf{F}_o = [0 \dots 0 \;\; (f_o)_{i,j} \;\; 0 \dots 0].$$

Now (A.17) can be written as

$$\begin{bmatrix} \mathbf{V}_{wp} & \mathbf{V}_{ws} \\ \mathbf{V}_{op} & \mathbf{V}_{os} \end{bmatrix} \begin{bmatrix} \dot{\mathbf{p}} \\ \dot{\mathbf{S}} \end{bmatrix} + \begin{bmatrix} \mathbf{T}_w & \mathbf{0} \\ \mathbf{T}_o & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{p} \\ \mathbf{S} \end{bmatrix} = \begin{bmatrix} \mathbf{F}_w \\ \mathbf{F}_o \end{bmatrix} \mathbf{q}_t$$

or

$$\widehat{\mathbf{E}}(\mathbf{X})\dot{\mathbf{X}} - \widehat{\mathbf{A}}(\mathbf{X})\mathbf{X} - \widehat{\mathbf{B}}(\mathbf{X})\mathbf{U} = \mathbf{0}, \tag{A.18}$$

where

$$\widehat{\mathbf{E}} = \begin{bmatrix} \mathbf{V}_{wp} & \mathbf{V}_{ws} \\ \mathbf{V}_{op} & \mathbf{V}_{os} \end{bmatrix}, \; \widehat{\mathbf{A}} = -\begin{bmatrix} \mathbf{T}_w & \mathbf{0} \\ \mathbf{T}_o & \mathbf{0} \end{bmatrix}, \; \widehat{\mathbf{B}} = \begin{bmatrix} \mathbf{F}_w \\ \mathbf{F}_o \end{bmatrix} \mathbf{L}_{qu},$$
$$\mathbf{X} = \begin{bmatrix} \mathbf{p} \\ \mathbf{S} \end{bmatrix}, \; \mathbf{U} = \mathbf{L}_{uq}\mathbf{q}_t,$$

vector $\mathbf{U}$ represents non-zero elements of the total flow rate vector $\mathbf{q}_t$, then $\mathbf{L}_{uq}$ is a location matrix consisting of zeros and ones at appropriate places and $\mathbf{L}_{qu} = \mathbf{L}_{uq}^{\mathrm{T}}$.

The system of equations (A.18) is the discretized version of model (A.11)–(A.12) which can be taken as the subject to future simulation.

A.3 Well model

In general reservoir simulation aims at providing an accurate forecast for the well production data and pressure and saturation distributions. For that purpose a model has to reflect the presence of wells in the field. The well treatment is considered to be a separate task with specific theory behind.

To describe a well performance one has to know the average grid block pressure p, the so-called flowing sandface pressure p_{wf} and the total production rate q_t. Since the well grid block has an additional unknown variable (either flowing sandface pressure or production rate), it is necessary to relate it to the known quantities. The basic assumption requires consideration of a flow of incompressible fluid as being steady-state cylindrical radial towards a well in the center of a grid block. Under such conditions the following

pressure distribution is obtained [7]:

$$p = p_{wf} - \frac{\mu q}{2\pi k_H h} \log\left(\frac{r_e}{r_{well}}\right), \tag{A.19}$$

where h is the grid block hight, r_e denotes external radius at which the analytical solution for pressure and numerical solution on a fine grid are equal, r_{well} states for the well-bore radius. Following the idea in [15] the external radius can be expressed in the form of

$$r_e = 0.14\sqrt{\Delta x^2 + \Delta y^2}. \tag{A.20}$$

Combining (A.19) and (A.20) we end up with

$$p = p_{wf} - \frac{\mu q}{2\pi k_H h} \log\left(\frac{0.14\sqrt{\Delta x^2 + \Delta y^2}}{r_{well}}\right)$$

or

$$p = p_{wf} - \frac{q}{J_{well}},$$

where $J_p = \dfrac{\mu}{2\pi k_H h} \log\left(\dfrac{0.14\sqrt{\Delta x^2 + \Delta y^2}}{r_{well}}\right)$ is a well index.

Now this concept has to be incorporated into equation (A.18).

A.4 Simple simulator `simsim`

The in-house simple simulator `simsim` used in the project solves the system of equations describing the two-phase two-dimensional fluid flow model in the reservoir with five-spot injection-production configuration. The system has generalized state space form and reads as follows [15]:

$$\widehat{\mathbf{E}}(\mathbf{X})\dot{\mathbf{X}} - \widehat{\mathbf{A}}(\mathbf{X})\mathbf{X} - \widehat{\mathbf{B}}(\mathbf{X})\mathbf{U} = \mathbf{0}, \tag{A.21}$$

where

$$\widehat{\mathbf{E}} = \begin{bmatrix} \mathbf{V}_{wp} & \mathbf{V}_{ws} \\ \mathbf{V}_{op} & \mathbf{V}_{os} \end{bmatrix}, \quad \widehat{\mathbf{A}} = -\begin{bmatrix} \mathbf{T}_w + \mathbf{F}_w\mathbf{J}_p & \mathbf{0} \\ \mathbf{T}_o + \mathbf{F}_o\mathbf{J}_p & \mathbf{0} \end{bmatrix}, \quad \widehat{\mathbf{B}} = \begin{bmatrix} \mathbf{F}_w \\ \mathbf{F}_o \end{bmatrix} [\mathbf{I}_q + \mathbf{J}_p]\,\mathbf{L}_{qu},$$

$$\mathbf{X} = \begin{bmatrix} \mathbf{p} \\ \mathbf{S} \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} \breve{\mathbf{p}}_{well} \\ \breve{\mathbf{q}}_{well} \end{bmatrix},$$

$\breve{\mathbf{p}}_{well}$ denotes prescribed bottom hole pressures and $\breve{\mathbf{q}}_{well}$ — prescribed total flow rates at the wells respectively, $\mathbf{J}_p$ is a diagonal matrix with non-zero well-indices J_p on the diagonal at the rows corresponding to the grid blocks with prescribed bottom hole pressures, $\mathbf{I}_q$ states for a diagonal matrix with ones on the diagonal at the rows corresponding to the grid blocks with prescribed total flow rates and other elements being zeros.

Reservoir simulator simsim implements particular method of implicit Euler integration with Newton iteration for solving (A.21). After initializing the simulator, the user can obtain the state vector $\mathbf{X}$ at each time point of interest.

Appendix B

Simsim input parameters

This chapter lists input data needed to initialize simulator simsim. The required parameters are:

- number of grid blocks in each direction: 21;
- field length and width: 700(m) each;
- grid block height: $h = 2(m)$;
- rock compressibility: $c_R = 1.0 \times 10^{-8}\,(Pa^{-1})$;
- oil compressibility: $c_o = 1.0 \times 10^{-8}\,(Pa^{-1})$;
- water compressibility: $c_w = 1.0 \times 10^{-8}\,(Pa^{-1})$;
- oil viscosity: $\mu_o = 5.0 \times 10^{-4}(Pa \cdot s)$;
- water viscosity: $\mu_w = 1.0 \times 10^{-3}(Pa \cdot s)$;
- porosity: $\phi = 0.3$;
- end point relative permeability for oil: $k^0_{ro} = 1.0$;
- end point relative permeability for water: $k^0_{rw} = 0.5$;
- Corey exponent for oil: $n_o = 2$;
- Corey exponent for water: $n_w = 2$;
- residual oil saturation: $S_{or} = 0.2$;

- connate water saturation: $S_{wc} = 0.2$;
- grid block number of injection well: 221;
- prescribed total flow rate for injector: $q_t = 0.002\,(m^3/s)$;
- grid block numbers of production wells: 1, 21, 421 and 441;
- prescribed bottom hole (flowing sandface) pressures for producers: $p_{wf} = 2.5 \times 10^7 (Pa)$;
- well-bore radius: $r_{well} = 0.1143(m)$.

The parameters given above correspond to a square five-spot injection-production situation which is considered in the project.

One needs to complete the input data with the time interval at which integration has to be performed, and also with appropriate initial pressures and water saturations at each grid cell.

List of Symbols and Abbreviations

Abbreviation	Description
EnKF	ensemble Kalman filter
FVF	formation volume factor
GOR	gas-oil ratio
PDE	partial differential equation
pdf	probability density function
PUNQ-S3	Production forecasting with UNcertainty Quantification (model)
RMS	root mean square (error)
simsim	simple simulator (our in-house reservoir simulator)
WOR	water-oil ratio

Symbol	Description	SI units
A	geometrical factor	$-,\ m,\ m^2$
$\mathbf{A}$	scaling matrix	-
$\widehat{\mathbf{A}}$	system matrix in generalized state space form	-
B_α	FVF of phase α	-
$\mathbf{B}$	matrix controlling model input	-
$\mathbf{B}$	scaling matrix	-
$\widehat{\mathbf{B}}$	matrix controlling model input in generalized state space form	-
c_R	rock compressibility	$1/Pa$
c_α	compressibility of phase α	$1/Pa$

Symbol	Description	SI units
$\mathbf{C}$	matrix	-
d	depth	m
$\widehat{\mathbf{E}}$	accumulation matrix in generalized state space form	-
f_E	error function	-
f_α	fractional flow for phase α	-
$\mathbf{f}$	non-linear vector-function	-
$\mathbf{F}$	system matrix	-
$\mathbf{F}$	operator of reservoir simulator	-
$\mathbf{F}_\alpha$	fractional flow matrix for phase α	-
g	acceleration of gravity	m/s^2
$\mathbf{g}$	non-linear vector-function	-
$\mathbf{G}$	matrix controlling model noise	-
h	grid block height	m
$\mathbf{I}$	identity matrix	-
$\mathbf{I}$	innovation of the filter	-
J_p	well index	$m^3/Pa \cdot s$
$\mathbf{J}_p$	well index matrix	$m^3/Pa \cdot s$
k	discrete time index	
k	permeability	m^2
$\mathbf{k}$	permeability tensor	m^2
$\mathbf{k}_\alpha$	effective permeability to phase α	m^2
$k_{r\alpha}$	relative permeability to phase	-
${k_{r\alpha}}^0$	end point permeability to phase α	-
$\mathbf{K}$	Kalman gain matrix	-
l	discrete time index	
$\mathbf{L}$	approximation of covariance matrix $\mathbf{P}$	-
$\mathbf{L}_{uq}$	location matrix	-
$\mathbf{L}_{qu}$	inverse of location matrix $\mathbf{L}_{uq}$	-
$\mathbf{m}$	vector of static model parameters	-
$\mathbf{M}$	linear measurement operator	-
n_{par}	the number of reservoir parameters	-
n_α	Corey exponent for phase α	-
$n(\alpha)$	number of measurements corresponding to the production data of type α	-
N	ensemble size	-
p	pressure	Pa
p^0	reference pressure	Pa
p_c	capillary pressure	Pa

Symbol	Description	SI units
p_{wf}	bottom hole (sandface) pressure	Pa
$\mathbf{p}$	vector of grid block pressures	Pa
$\breve{\mathbf{p}}_{well}$	vector of prescribed bottom hole pressures at the wells	Pa
$\mathbf{p}_{well}$	vector of non-prescribed bottom hole pressures at the wells	Pa
$\mathbf{P}$	system state covariance matrix	-
$\bar{q}_\alpha$	source term for phase α	s^{-1}
q_t	total flow rate	m^3/s
$\mathbf{q}_\alpha$	vector of flow rates for phase α	m^3/s
$\mathbf{q}_t$	vector of total flow rates	m^3/s
$\breve{\mathbf{q}}_{well}$	vector of prescribed total flow rates at the wells	m^3/s
$\mathbf{q}_{well,\alpha}$	vector of non-prescribed flow rates of phase α at the wells	m^3/s
$\mathbf{Q}$	model noise covariance matrix	-
r	normalized RMS ratio	-
r_e	external radius	m
r_{well}	well bore radius	m
$\mathbf{R}$	observational noise covariance matrix	-
Ra	RMS ratio	-
S	fluid saturation	-
S	normalized saturation	-
S_{or}	residual oil saturation	-
S_{wc}	critical water saturation	-
$\mathbf{S}$	vector of grid block water saturations	-
t	time	s
T	temperature	$^\circ C$
T	number of data assimilation steps	-
T^0	reference temperature	$^\circ C$
T_α	transmissibility of phase α	$m^3/Pa \cdot s$
$\mathbf{T}_\alpha$	matrix of transmissibility terms corresponding to phase α	$m^3/Pa \cdot s$
$\mathbf{U}$	model input vector	-
$\mathbf{v}_\alpha$	superficial velocity for phase α	m/s
V	volume	m^3
$\mathbf{V}$	Gaussian white measurement noise vector process	-
$\mathbf{V}_{wp}$	sub-matrix of matrix of accumulation terms	-
$\mathbf{V}_{ws}$	sub-matrix of matrix of accumulation terms	-

Symbol	Description	SI units
$\mathbf{V}_{op}$	sub-matrix of matrix of accumulation terms	-
$\mathbf{V}_{os}$	sub-matrix of matrix of accumulation terms	-
w	weighting coefficient	-
$\mathbf{W}$	Gaussian white model noise vector process	-
x	x-direction in Cartesian coordinate system	m
$\mathbf{X}$	state space vector	-
$\mathbf{X}_{io}$	vector of observed data	-
$\mathbf{X}_{is}$	vector of simulated data	-
y	y-direction in Cartesian coordinate system	m
$\mathbf{Y}$	vector of dynamic variables	-
$\mathbf{Y}$	vector of bottom hole pressures, oil and water flow rates at the wells	-
z	z-direction in Cartesian coordinate system	m
$\mathbf{Z}$	measurement vector	-
α	type of production data	-
Δx	step for spatial grid discretization in x-direction	m
Δy	step for spatial grid discretization in y-direction	m
λ_α	mobility tensor for phase α	$m^2/Pa \cdot s$
μ_α	fluid viscosity of phase α	$Pa \cdot s$
ξ	system state vector (member of ensemble)	-
ρ_α	density of phase α	kg/m^3
ϕ	porosity	-
ϕ^0	porosity at the reference pressure p^0	-

Subscript	Description
0	initial
α	phase
H	horizontal
i	discrete counter
j	discrete counter
k	discrete time index
l	discrete time index
o	oil (non-wetting phase)
sc	standard conditions
V	vertical
w	water (wetting phase)
x	x-direction in Cartesian coordinate system

Subscript	Description
y	y-direction in Cartesian coordinate system
z	z-direction in Cartesian coordinate system

Superscript	Description
c	confirmed
i	discrete counter
true	true value
T	transpose

Operator	Description
∇	gradient
∂	partial derivative
$\|\cdot\|$	2-norm
dim	dimension
E	expectation
$\mathcal{N}$	normal distribution