

**QUANTIFICATION OF
THE IMPACT OF DATA
IN RESERVOIR MODELING**

QUANTIFICATION OF THE IMPACT OF DATA IN RESERVOIR MODELING

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. ir. K.C.A.M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen
op woensdag 22 mei 2013 om 10:00 uur
door

Mariya Victorovna KRYMSKAYA

ingenieur in de toegepaste wiskunde
geboren te Ufa, Sovjet-Unie.

Dit proefschrift is goedgekeurd door de promotoren:

Prof. dr. ir. A.W. Heemink

Prof. dr. ir. J.D. Jansen

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Prof. dr. ir. A.W. Heemink,	Technische Universiteit Delft, promotor
Prof. dr. ir. J.D. Jansen,	Technische Universiteit Delft, promotor
Prof. B.A. Foss,	Norges Teknisk-Naturvitenskapelige Universitet, Norwegen
Prof. dr. ir. H.X. Lin,	Technische Universiteit Delft, Tilburg University
Prof. dr. ir. P.M.J. Van den Hof,	Technische Universiteit Eindhoven, Technische Universiteit Delft
Prof. dr. ir. C. Vuik,	Technische Universiteit Delft
Dr. R.G. Hanea,	Technische Universiteit Delft, TNO

Dr. R.G. Hanea heeft als begeleider in belangrijke mate aan de totstandkom-
ing van het proefschrift bijgedragen.

This research was carried out within the context of the "Integrated Systems Approach to Petroleum Production" (ISAPP) knowledge center. ISAPP is a long-term co-operation of the Dutch Organization for Applied Scientific Research (TNO), Shell International Exploration and Production (SIEP) and Delft University of Technology (TU Delft) to increase hydrocarbon recovery through the application of innovative reservoir development and management technologies.

ISBN: 978-90-8891-629-8

Copyright ©2013 by M.V. Krymskaya.

All rights reserved. No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without written permission of the author.

Cover design: Proefschriftmaken.nl || Uitgeverij BOXPress

Published by: Uitgeverij BOXPress, 's-Hertogenbosch

Contents

1	Introduction	1
1.1	Energy landscape	1
1.2	Closed-loop model-based reservoir management	3
1.2.1	History matching loop	5
1.2.2	Production optimization loop	7
1.3	Research objective and solution directions	7
1.3.1	Research objective	8
1.3.2	Solution directions	8
1.4	Thesis outline	10
2	Experimental environment	11
2.1	Introduction	11
2.2	Reservoir model	11
2.2.1	Governing equations	12
2.2.2	State space representation	13
2.2.3	Simple simulator <code>simsim</code>	13
2.2.4	Parameter estimation: augmented state vector approach	14
2.3	Twin experiment	15
2.3.1	'Truth' model	15
2.3.2	Synthetic noisy measurements	17
2.3.3	'Uncertain' model	17
3	Representer and accelerated representer algorithms	21
3.1	Introduction	21
3.2	The RM and the ARM	22
3.2.1	Generalized cost function	22
3.2.2	The RM: linear and non-linear cases	24
3.2.3	The ARM	26

3.2.4	Scaling	28
3.2.5	Assessing the quality of the algorithm performance . . .	28
3.3	Results and discussion	29
3.3.1	The RM as iterative procedure	29
3.3.2	Tuning the ARM	31
3.3.3	Comparison of computational performance	33
3.4	Conclusions	35
4	An ensemble Kalman filter and its modifications	39
4.1	Introduction	39
4.2	The ensemble Kalman filter	41
4.2.1	The confirming ensemble Kalman filter	43
4.2.2	The asynchronous ensemble Kalman filter	44
4.2.3	Scaling	46
4.3	Results and discussion	48
4.3.1	History matching with the EnKF	48
4.3.2	History matching with the AEnKF	49
4.4	Conclusion	52
5	Observation sensitivity matrix	53
5.1	Introduction	53
5.2	Concept of observation sensitivity	54
5.3	Structure of the observation sensitivity matrix	57
5.3.1	Symmetry of the observation sensitivity matrix	57
5.3.2	Scaled observation sensitivity matrix	58
5.3.3	Positive (semi)definiteness of the scaled observation sensitivity matrix	59
5.4	Matrix norm of the observation sensitivity matrix	60
5.5	Uncertainty in the analyzed model	63
5.6	Application of the observation sensitivity matrix to HM with the RM	65
5.6.1	Observation sensitivity matrix	65
5.6.2	SVD of the observation sensitivity matrix	68
5.6.3	Trace diagnostics of the observation sensitivity matrix	69
5.6.4	Role of measurement type and/or location	70
5.6.5	Role of prior fields	78
5.6.6	Role of 'true' fields	80
5.7	Application of the observation sensitivity matrix to HM with the AEnKF	83
5.7.1	Role of water breakthrough	84
5.7.2	Role of exciting input	88
5.8	Conclusion	97

6	Conclusions	99
6.1	Overview	99
6.2	Conclusions	100
6.3	Recommendations	102
	Bibliography	105
A	Fluid-rock properties	113
A.1	Capillary pressure	113
A.2	Relative permeability	113
A.3	Corey-type two-phase relative permeability model	114
B	Mathematical background	115
B.1	Condition number of a matrix	115
B.2	Vector and matrix norms	115
B.3	Singular value decomposition of a matrix	116
	Summary	119
	Samenvatting	123
	Acknowledgments	127
	Curriculum Vitae	129

Introduction

1.1 Energy landscape

Energy is essential for our lives. On a daily basis energy is used to power residential and commercial buildings, in the industrial sector and for transportation. As the world population grows and the living conditions improve, the energy demand increases.

The International Energy Agency expects the world energy consumption to grow by 53% in total in the period from 2008 to 2035 (Figure 1.1). In countries outside the Organization for Economic Cooperation and Development (non-OECD nations), where demand is driven by strong long-term economic growth, the energy consumption increases by 85%. Energy use in OECD economies increases by 18%*.

To satisfy this increasing demand, liquid, natural gas, coal, nuclear power and renewable fuel sources are extensively developed (Figure 1.2).

Although renewable energy is currently the world fastest growing form of energy, it is not yet able to provide sufficient and reliable energy supply. The earthquake and tsunami in Japan on March 11, 2011, resulted not only in extensive loss of life and infrastructure damage, but also made the long-term future of nuclear energy extremely uncertain due to the policies some countries have adopted with respect to continued operation of the existing nuclear plants. Fossil fuels (i.e. oil, natural gas and coal) will therefore most likely remain the largest source of energy for the world for many decades to come. Oil, in particular, is expected to provide about 30% of the world energy needs.

*As of September 1, 2010, the OECD member countries are the United States, Canada, Mexico, Austria, Belgium, Chile, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Luxembourg, the Netherlands, Norway, Poland, Portugal, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, the United Kingdom, Japan, South Korea, Australia, and New Zealand.

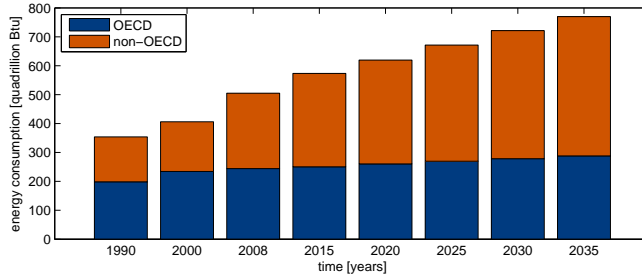


Figure 1.1: World energy consumption per region according to [1].

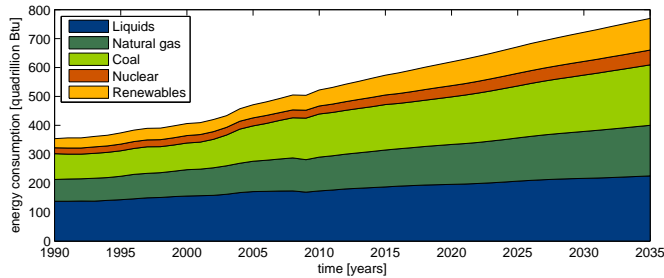


Figure 1.2: World energy consumption per fuel type according to [1].

Not only the global energy demand rises, so does the need to produce the energy resources responsibly, to minimize impact on air and water pollution, climate change and altogether to mitigate risks for environment and society. The landscape of operations for the energy companies thus becomes increasingly more complex.

To keep up with the demand and consumer expectations, the oil companies, in particular, need to continuously increase the rate of production from the existing fields and discover new fields to compensate for the produced amounts. To address the former, they need to develop new and enhance current technologies to reduce production costs, minimize the environmental effect and maximize the recovery.

The hydrocarbon recovery process can be divided into three phases [15]:

- *primary recovery* is the very early stage at which reservoir contains fluids (e.g. gas, oil) under the high pressure. The gas or oil is produced by natural decompression. 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 part of the remaining oil (or gas). For that purpose a fluid (typically the water) is injected into injection wells maintaining high reservoir pressure and

producing oil through production wells. After this stage often 50% or more of hydrocarbons still remain in the reservoir. Water flooding suffers from water breakthrough after which some injected water is produced back via production wells. The time at which the first 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 steam into reservoir to modify the viscosity of the oil and to make it flow easier.

Which mode of the operation will be the most successful for the given field, when secondary recovery has to be initiated to maximize recovery, what type of the enhanced oil recovery (EOR) is appropriate for the prospect, depends on the the reservoir rock properties, the nature of fluids filling reservoir and the effects appearing due to rock-fluid interaction.

Implementation of any of the recovery strategies requires vast amounts of investment, and therefore the energy companies rely on various tools to support their production and development decisions. Particularly, mathematical models are advantageous as they allow to describe and study the behavior of the system, evaluate associated risks and uncertainties, and forecast the production performance.

1.2 Closed-loop model-based reservoir management

Nowadays mathematical modeling is a widely used technique to describe the reservoir system and in-situ processes via mathematical equations coupled with boundary and/or initial conditions. The form of the model equations essentially depends on the nature of reservoir, namely, the rock and fluids filling it. The knowledge of the rock and fluid properties is generally very limited, though crucial for creating a representative model. Additional difficulties are caused by the fact that rock properties usually vary in space (i.e. are heterogeneous).

Since the field description is quite complex, it is practically impossible to solve the model analytically even using simplifying assumptions. This obstacle can be overcome by solving the mathematical model via numerical methods to obtain approximate results. The continuous development of the computer industry supports building and solving numerical models with an increasing number of details.

The mathematical models are employed in making initial predictions of reservoir performance and evaluating if development of a particular asset is

potentially profitable. Numerical models are also used to mimic the processes occurring in the reservoir during production, which allows investigation and comparison of various production scenarios, and determination of the most optimal production strategy based on predefined, usually economic, criteria.

Mathematical models form an important component of the integrated closed-loop model-based reservoir management (CLRM) framework [37, 38], and as such have to be considered in close connection to the other key elements of the CLRM process. Figure 1.3 displays reservoir management as a model-based controlled process.

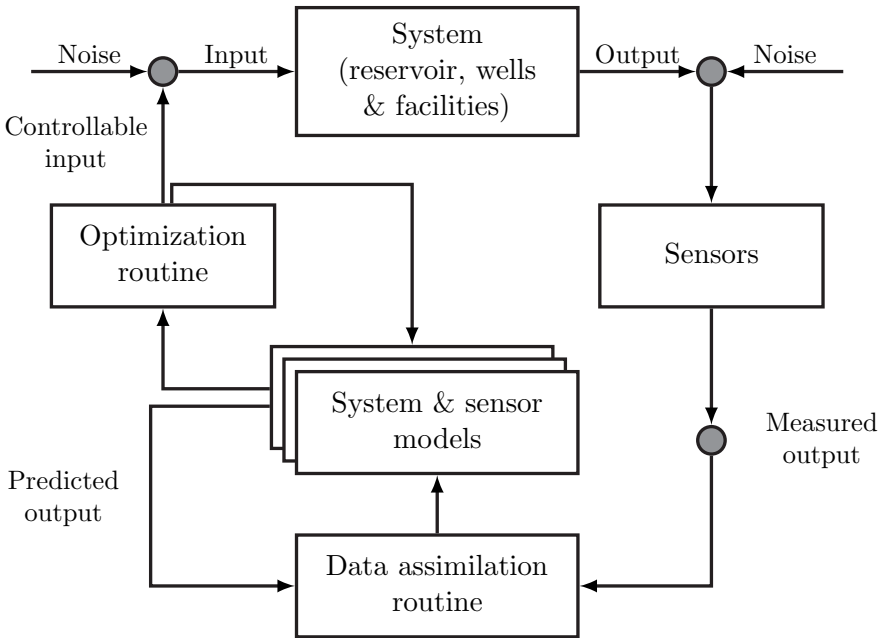


Figure 1.3: Closed-loop model-based reservoir management [38].

The system box on the top represents physical reality with one or more reservoirs, wells and facilities. The system is controlled by a number of input parameters, and system behavior is monitored with the help of sensors gathering various information about the system variables (e.g. pressures and flow rates in the wells). Both system inputs and outputs are imperfectly known and subject to noise.

In the most general form the system model is a mathematical relationship between the inputs and outputs. The system and sensor models may include static (geological), dynamic (reservoir flow) and well bore flow models set up to mimic the behavior of the real system. Here the multiple boxes represent the need for multiple geological models to quantify the range of uncertainty in the knowledge on subsurface structure for a given reservoir.

These system models are usually just a very crude approximation of reality, and the model parameters, such as permeabilities, are only known with large uncertainty. Therefore the predictive value of these models is limited and tends to deteriorate in time. Data assimilation (i.e. computer-assisted history matching) is used to calibrate such models and improve their predictive capability. Data assimilation employs the data that become available during the production life of the reservoir to update the values of poorly known model parameters in time, i.e. to adapt parameters such that simulated results are consistent with measured data.

After the history matching (HM) has been performed, the model can be used for developing optimal production strategy and decision making. This may involve e.g. determining the optimal number of wells, their locations and controls. The procedure is reiterated whenever the new data become available.

1.2.1 History matching loop

From the mathematical point of view history matching can be seen as a minimization problem, namely as searching for a combination of reservoir parameters that minimizes a least-squares objective function representing the difference between the observed reservoir performance and the results of simulation over a past period [56, 59]. History matching is usually an ill-posed problem, because there are more unknowns than constraints to resolve all undetermined quantities. Then the Gauss-Newton method applied to minimize cost function fails because the Hessian is ill-conditioned. Such a problem can be overcome by applying some regularization strategy (e.g. by use of prior geostatistical model [48]).

Formulated in this manner, the history matching problem can be addressed with data assimilation methods as used in meteorology and oceanography. In these fields, data assimilation is primarily used for state estimation, i.e. it is aimed at producing improved forecasts of model states (dynamic variables) by combining information available in measurements and prior models. The majority of data assimilation techniques can be easily adjusted to also produce improved forecasts of model parameters (static variables). An essential difference is that state estimation can be formulated as either a linear or a non-linear problem, whereas parameter estimation is always a non-linear problem. Most data assimilation methods were originally developed for linear state estimation problems. Application of those techniques to history matching problems in the reservoir engineering domain therefore requires adaptations to cope with the essential non-linearity of parameter estimation, which usually results in iterative procedures.

Traditionally, a distinction is made between two main approaches to data

assimilation: 1) variational techniques, where gradient-based minimization methods are used to minimize an objective function representing the mismatch between all modeled and measured observations over the entire assimilation period, and where exact gradients are used that are usually computed with the aid of an adjoint method (see e.g. [13, 14, 17, 61, 77] for early applications of adjoint-based variational algorithms in reservoir engineering and meteorology); 2) recursive techniques (also known as 'sequential techniques'), where the mismatch between modeled and measured observations is minimized at the moment that the measurements become available, and where stochastic approximations to the gradients are used (see e.g. [39] for Kalman filter methods, [23] for an overview of meteorological and oceanographic problems addressed with the ensemble Kalman filter (EnKF) and its modifications). In fact there is no reason to restrict the use of exact gradients to the assimilation of all observations over the entire period, or to restrict the use of stochastic gradients to sequential assimilation, and indeed both variational and the recursive methods may be utilized to sequentially assimilate groups of measurements taken over subintervals of the data assimilation period. Variational methods may be applied to solve deterministic or stochastic assimilation problems. Deterministic problems are those where uncertainty is assumed in neither the measurements nor the model (i.e. the states and/or the parameters). On the contrary, stochastic problems are those where we assume that errors may be present in the measurements and/or the model. Recursive methods are typically only used to solve stochastic assimilation problems, because computing a stochastic gradient presupposes the presence of uncertainty. It can be shown that for linear problems the stochastic versions of variational and recursive approaches lead to exactly the same results. For non-linear cases, however, different results can be expected depending on the particular linearization strategies and iterative procedures employed. During the past decades both approaches have been actively exploited for parameter estimation in reservoir engineering: Kalman filter algorithms for history matching are discussed in e.g. [2, 24, 29, 53] and examples of adjoint-based variational techniques as applied to history matching are given in e.g. [13, 14, 48, 59]. The performances of EnKF and variational algorithms in application to history matching have been compared in [32].

The history-matched model is further used to forecast field performance under specified development strategy. This allows investigation of various field production scenarios, their evaluation against predefined criteria and comparison with each other, and identification of potential risks during field development.

1.2.2 Production optimization loop

During production optimization, the goal mostly is to maximize a certain economic criterion. Frequently the net present value (i.e. the difference between the revenues and the costs) or the recovery factor is being maximized. The optimization parameters are the number, the type and the location of wells, and their control strategies.

Spatial heterogeneities of reservoir geological characteristics form additional obstacles for production optimization during water flooding, as they can cause preferential fluid flow paths outside of which significant quantities of hydrocarbons can be bypassed and remain not recovered. When water is injected through the injection wells in order to sweep the oil into production wells, the heterogeneities lead to irregularities in the shape of the moving water-oil front. Once this front has reached the production well, a preferential flow path is formed, through which most of the injected water will flow. Continuing water injection will result in production of water instead of oil. This can make production from a given well economically inviable, and the well will have to be shut-in. Had the wells been drilled in different locations or operated in a different manner, the water breakthrough could have been postponed.

Execution of the optimization loop in the CLRM framework can assist in identifying an optimal production strategy. For a given configuration of wells, the flooding process can be optimized over the producing life of the reservoir by appropriate selection of well rates and pressures (see e.g. [9, 73]). This problem has been studied both under assumption of perfect [43, 73] and imperfect [91] knowledge on model parameters. For early attempts on optimization of tertiary recovery processes such as polymer or carbon dioxide flooding see [66]. Optimizing production through sophisticated selection of well locations and well trajectories is discussed in e.g. [91] and [85] respectively.

1.3 Research objective and solution directions

Constituting the subject of closed-loop model-based reservoir management framework history matching (data assimilation) and optimization loops are very challenging.

The history matching phase mainly focuses on improving the model parameters such that the model will closely mimic the processes occurring in the reservoir during production. The solution of this problem is usually obtained by incorporating available measurements into the model, i.e. translating them into the corrections that have to be applied to the model parameters.

In a general case, the use of a larger measurement set potentially improves estimators of the model parameters and predictive capability of the model.

In that sense it is beneficial to involve all the measurements available into a history matching procedure. However, the information extracted from the measurements in the history matching phase is repeatedly found as not enough to provide a well-calibrated model with a high predictive value (see e.g. [63, 83]). Hence, consideration of additional data can be of particular help. To optimize the costs and effort associated with the collection of new data and computations, one needs to (i) identify which data or portions of them have the most effect on the model improvement and, hence, should be studied and gathered carefully; (ii) establish the value of new data in reducing uncertainty; (iii) estimate the type and amount of data required to be assimilated to obtain a model of the desired accuracy.

The selection of measurements becomes an issue also during the production optimization phase, since new wells to be drilled according to the field development plan will bring additional data on previously inaccessible areas of the reservoir. These new measurements are thereafter assimilated in the model. Naturally, production settings yielding the same value of optimization criterion do not necessarily provide observations having the same effect on model uncertainty when history-matched. Therefore a production strategy that results in gathering the most influential data can become preferential.

1.3.1 Research objective

The discussion on the use of observations raises an important question to be answered: can the information provided by the observations be quantified? A simple data count might be misleading as not all observations are equal in what they measure and in their accuracy. This leads to the following objective of this research:

To develop efficient tools for quantifying the impact of measured data on the outcome of history matching of reservoir models.

The tools are recognized as efficient if they provide meaningful quantification of the impact of observations, while requiring limited time and effort to be incorporated in the history matching phase of the closed-loop reservoir management framework.

1.3.2 Solution directions

Methods to assess the impact of observations (i.e. measurements of physical variables) have originally been developed for meteorological and oceanographic applications. In meteorology and oceanography the collected data are fed into the data assimilation process with the aim to improve the initial conditions of the model. The measurement equipment nowadays provides

enormous amount of data on a daily basis. Assimilating these data becomes too demanding for operational forecasting systems due to time constraints on obtaining the new forecast and limitations of computational resources. Hence efforts are made towards proper selection of the most influential measurements. The data assimilation process in meteorology and oceanography has many similarities to computer-assisted history matching in petroleum reservoir engineering. Therefore, this thesis will investigate the applicability of these methods to quantify the impact of observations to the history matching of reservoir models.

Three main approaches can be distinguished to determine the impact of observations:

- *Observation targeting* methodologies aim at identifying the regions where uncertainty in the initial conditions can result in a wide range of possible model predictions. The regions can be detected with the help of e.g. singular vectors [25, 60]. It is envisaged that assimilation of the measurements collected in these sensitive regions can reduce the uncertainty in the forecasts. The outcome of such an analysis can provide useful guidelines for surveillance campaigns by targeting locations for data gathering. However, there exists no direct way to translate these recommendations into the value of a particular observation;
- *An observation sensitivity matrix* can be constructed to quantify the effect of measured data on the model parameter update [11, 65]. The matrix can be used to evaluate the amount of information extracted from available data during the data assimilation phase and identify the observations that have contributed to the parameter update the most. Though no immediate conclusion can be made on the importance of these exceptional data points for the quality of model predictions, it is reasonable to expect that an improved model characterization will result in improved forecasts. This technique is easy to implement both for adjoint-based and EnKF types of data assimilation algorithms;
- *The forecast sensitivity to observations* concept considers the gradient of a forecast error cost function with respect to the vector of measurements. This gradient is used to diagnose the observations that, when assimilated, are likely to have a positive impact on subsequent model predictions [20, 26, 46]. The scalar forecast error cost function used in atmospheric applications usually includes a verifying analysis trajectory. In reservoir engineering applications, however, no verifying analysis is available. Moreover, implementation of the sensitivity calculations is cumbersome and relies on an adjoint, which is often not available. Though the authors of [49] attempted to overcome this drawback for

data assimilation systems involving EnKF, it has not yet resulted in the standardized procedure.

The above discussion on the available methods, their benefits and drawbacks, leads to selecting the observation sensitivity matrix approach as the most prospective candidate for tackling the stated research problem. Throughout the study we consider a field at the secondary recovery stage being operated through water flooding with multiple injectors and producers to test the influence of the data on the outcome of the history matching.

1.4 Thesis outline

The thesis is organized as follows.

Chapter 2 presents the mathematical model for multi-phase fluid flow behavior and provides the set-up of the case studies performed in the remainder of the thesis.

Chapter 3 investigates application of a particular variational method, the representer method, to history matching. Computational inefficiency of the representer method for assimilation of large data sets leads to testing the capabilities of its accelerated version.

Chapter 4 describes two EnKF techniques: the classical EnKF and its asynchronous modification. Both techniques are powerful tools for data assimilation, and are tested prior to being used as platforms to evaluate the impact of measured data. This chapter is partly based on [44].

Chapter 5 focuses on the details of the observation sensitivity matrix concept and its implementation for history matching with both variational- and EnKF-based algorithms, and discusses and interprets the results. In addition, possible theoretical extensions of the concept are considered. This chapter is partly based on [45].

Finally, Chapter 6 concludes the thesis by summarizing the findings and recommendations for further research.

Experimental environment

2.1 Introduction

To quantify the impact of data in reservoir modeling and history matching, first the history matching exercise itself has to be set up. We are considering a field at the secondary recovery stage being operated through water flooding with multiple injectors and producers, and test the influence of the data on the outcome of the history matching with the aid of a synthetic ('twin') experiment (Section 2.3). Setup of the history matching exercise includes:

- identifying the appropriate model to describe the processes in the reservoir at the water flooding stage of production;
- defining the known parameters of the fluid flow model;
- determining the uncertain parameters to be estimated and specifying the prior knowledge on these;
- generating synthetic data (or collecting real ones);
- implementing and testing the history matching algorithms.

This chapter defines the experimental settings and characteristics to be used in the case studies except for the history matching algorithms that are discussed in the next two chapters.

2.2 Reservoir model

We consider the equations for two-phase (oil-water) flow through porous media (see e.g. [4]). Moreover, we simplify the analysis by assuming a two-dimensional horizontal reservoir, isotropic permeability, small fluid and rock

compressibilities, and by disregarding gravity and capillary pressure. For the details on fluid-rock properties (e.g. capillary pressure) the reader is referred to e.g. [4, 21] and Appendix A of this thesis.

2.2.1 Governing equations

The mass balance equations can be expressed for each phase as

$$\nabla \cdot (\rho_i \mathbf{v}_i) + \frac{\partial (\rho_i \phi S_i)}{\partial t} - \rho_i q_i = 0, \quad (2.1)$$

where ρ is phase density at in-situ conditions, \mathbf{v} is the Darcy velocity, ϕ is porosity, S is saturation, q is flow rate per unit volume, t is time, and the subscript i indicates the phase (oil or water). The differential representation of Darcy's law for the simultaneous flow of more than one phase takes the form

$$\mathbf{v}_i = -\frac{k_{ri}}{\mu_i} k \nabla p_i, \quad (2.2)$$

where k is permeability, $k_{ri}(S_i)$ is relative permeability (see also Appendix A), μ is viscosity and p is pressure.

Substituting (2.2) in (2.1), using the closure equation

$$\sum_{i=1}^2 S_i = 1, \quad (2.3)$$

and working out the expressions yields

$$-\nabla \cdot \left(\rho_w \frac{k_{rw}}{\mu_w} k \nabla p \right) + \frac{\partial (\rho_w \phi S_w)}{\partial t} - \rho_w q_w = 0, \quad (2.4)$$

$$-\nabla \cdot \left(\rho_o \frac{k_{ro}}{\mu_o} k \nabla p \right) + \frac{\partial (\rho_o \phi (1 - S_w))}{\partial t} - \rho_o q_o = 0, \quad (2.5)$$

where the subscripts 'o' and 'w' for the pressure have been omitted, since the absence of capillary pressure implies that $p_o = p_w$.

With the aid of the definitions for small fluid and rock compressibilities

$$c_i(p) = \frac{1}{\rho_i} \frac{\partial \rho_i}{\partial p} \quad (2.6)$$

and

$$c_r(p) = \frac{1}{\phi} \frac{\partial \phi}{\partial p} \quad (2.7)$$

equations (2.4)–(2.5) can be rewritten as

$$-\frac{1}{\mu_w} \left[\frac{\partial}{\partial x} \left(k k_{rw} \frac{\partial p}{\partial x} \right) + \frac{\partial}{\partial y} \left(k k_{rw} \frac{\partial p}{\partial y} \right) \right] + \left[\phi S_w (c_w + c_r) \frac{\partial p}{\partial t} + \frac{\partial S_w}{\partial t} \right] - q_w = 0, \quad (2.8)$$

$$-\frac{1}{\mu_o} \left[\frac{\partial}{\partial x} \left(k k_{ro} \frac{\partial p}{\partial x} \right) + \frac{\partial}{\partial y} \left(k k_{ro} \frac{\partial p}{\partial y} \right) \right] + \left[\phi (1 - S_w) (c_o + c_r) \frac{\partial p}{\partial t} - \frac{\partial S_w}{\partial t} \right] - q_o = 0. \quad (2.9)$$

Initial conditions for equations (2.8)–(2.9) are set in the form of constant pressure and saturation fields. Boundary conditions are specified as no-flow conditions at external boundaries and prescribed bottom-hole pressures and/or total flow rates at the wells.

2.2.2 State space representation

Applying a central difference scheme with uniform grid to approximate the spatial differentials [4], the spatially discretized version of model (2.8)–(2.9) can be expressed as

$$\hat{\mathbf{E}}(\mathbf{x})\dot{\mathbf{x}} = \hat{\mathbf{A}}(\mathbf{x})\mathbf{x} + \hat{\mathbf{B}}(\mathbf{x})\mathbf{u}, \quad (2.10)$$

where $\hat{\mathbf{E}}$ is the accumulation matrix, $\hat{\mathbf{A}}$ is the transmissibility matrix, $\hat{\mathbf{B}}$ is the input matrix, \mathbf{u} is the input vector representing prescribed flow rates or bottom hole pressures in the wells, and $\mathbf{x}^T = [\mathbf{p}^T \ \mathbf{s}^T]$ is the state vector consisting of grid block pressures \mathbf{p} and water saturations \mathbf{s} . Expression $\dot{\mathbf{x}}$ designates the derivative of the state vector with respect to time.

2.2.3 Simple simulator `simsim`

The in-house simple simulator `simsim` used in the current investigation solves the system of equations (2.10). The reservoir is supposed to be operated under constraints on bottom hole pressures and/or total flow rates at the wells. The simulator uses implicit Euler integration with Newton iteration for solving (2.10). After initialization with grid block pressures and saturations, a given input strategy \mathbf{u} , and with parameters describing the rock and fluid properties (permeability, porosity, compressibility and viscosity), the user can obtain the state vector \mathbf{x} at each time point of interest t_n , i.e.

$$\mathbf{x}_n = \mathcal{M}(\mathbf{x}_{n-1}, \phi, \mathbf{k}),$$

where porosity ϕ and permeability \mathbf{k} are vectors containing values of porosity and permeability for each of the grid cells, and \mathcal{M} is an operator representing the action of the reservoir simulator.

The simulator also generates an output vector \mathbf{y}^{pr} of production data, consisting of the values of bottom hole pressures and water and oil flow rates at each of the wells, i.e.

$$\mathbf{y}^{pr} = \begin{bmatrix} \mathbf{p}_{bh}^{well} \\ \mathbf{q}_o^{well} \\ \mathbf{q}_w^{well} \end{bmatrix},$$

where \mathbf{p}_{bh}^{well} denotes the vector of bottom hole pressures, \mathbf{q}_o^{well} and \mathbf{q}_w^{well} stand, respectively, for the vector of oil flow rates and the vector of water flow rates. The output vector is related to the state vector according to

$$\mathbf{y}_n^{pr} = \mathcal{H}^{pr}(\mathbf{x}_n, \phi, \mathbf{k}),$$

where \mathcal{H}^{pr} is an operator providing output from the model.

2.2.4 Parameter estimation: augmented state vector approach

One of the approaches to perform parameter estimation involves incorporation of the parameters of interest into the state vector. Our study is focused on estimating porosity and permeability fields, hence, the relevant variables have to be included into the state vector.

The parameters values are often subject to a log-transform [59] in order to comply with the Gaussian assumptions classically used in the data assimilation methods (see e.g. [23, 39]). If porosity ϕ and permeability \mathbf{k} are transformed into $\ln(-\ln \phi)$ and $\ln \mathbf{k}$ respectively, then the transformed values belong to the interval $(-\infty, \infty)$. Log-transformation also resolves the problem of obtaining non-physical values of parameters after a linear data assimilation update.

Using these transformations, the augmented state space vector reads:

$$\mathbf{x}^* = \begin{bmatrix} \mathbf{x} \\ \ln(-\ln \phi) \\ \ln(\mathbf{k}) \end{bmatrix} = \begin{bmatrix} \mathbf{p} \\ \mathbf{s} \\ \ln(-\ln \phi) \\ \ln(\mathbf{k}) \end{bmatrix}. \quad (2.11)$$

Note that a linear assimilation update can potentially also result in non-physical values of the pressures \mathbf{p} and particularly of the saturations \mathbf{s} . Moreover, since the data assimilation update does not take into account the physical nature of the relationships between different variables, the updated values of pressures and saturations may be impossible for the system to reach from its initial condition given the newly estimated values of the parameters.

Appropriate transformations can be employed to ensure the updated values of the pressures and saturations are within physical bounds, e.g. the authors of [67] and [36] respectively suggest applying log-transform or inverse

error function transform to the saturations. This, however, does not resolve the possible inconsistency between the dynamic state of the system and its estimated parameters. Transforming the saturations is not considered in this thesis, as we are mostly interested in the estimated values of porosity ϕ and permeability \mathbf{k} , which in turn allow to consistently restore the dynamic state of the deterministic system given its initial state.

If the model parameters are considered as static, i.e. time invariant, then the model equations take the form

$$\mathbf{x}_n^* = \mathcal{M}^* (\mathbf{x}_{n-1}^*), \quad (2.12)$$

where

$$\mathcal{M}^* (\mathbf{x}_{n-1}^*) = \begin{bmatrix} \mathcal{M}(\mathbf{x}_{n-1}, \phi, \mathbf{k}) \\ \ln(-\ln \phi) \\ \ln(\mathbf{k}) \end{bmatrix}.$$

Expression (2.12) gives the state space representation of the model under study. To account for time invariance, operator \mathcal{M}^* acts as identity operator for the parts of the state vector \mathbf{x}^* corresponding to the transformed porosity and permeability values. In a simple case of linear transformation \mathcal{M} with a matrix \mathbf{M} , the transformation \mathcal{M}^* is also linear and given by a block diagonal matrix

$$\mathbf{M}^* = \begin{bmatrix} \mathbf{M} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_\phi & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_\mathbf{k} \end{bmatrix},$$

where \mathbf{I}_ϕ and $\mathbf{I}_\mathbf{k}$ are the identity matrices (operators) of appropriate sizes to be applied to the transformed vectors of porosity and permeability.

The remainder of the paper considers the model in the form (2.12) with the asterisk omitted to simplify the notation.

2.3 Twin experiment

We test the performance of the history matching algorithms with the aid of a twin experiment, i.e. we use a reservoir model representing the 'truth' to generate synthetic noisy measurements, and we assimilate these measurements in another model, different from the 'truth' but with similar characteristics.

2.3.1 'Truth' model

We use a simple two-dimensional two-phase (oil-water) simulation model to represent the 'truth' with a permeability field resembling the fields considered in earlier studies [64] and [70]. The model describes a square oil field with

a size of $700 \text{ m} \times 700 \text{ m}$ and a thickness of 2 m . The model has a uniform cartesian grid consisting of 21 grid cells in each direction.

The field is produced through waterflooding with five wells: a water injection well (INJ) at the center and four production wells, one each at the North-West (NW), South-West (SW), North-East (NE) and South-East (SE) corners. The 'true' permeability and porosity fields are depicted in Figure 2.1. They each represent a single member out of ensembles of permeability and porosity fields which will be described later.

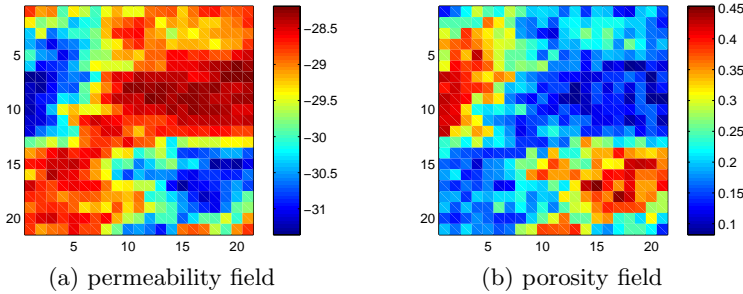


Figure 2.1: 'True' permeability and porosity fields. The scale represents transformed values of permeability [m^2] and porosity [-].

The relevant reservoir and fluid properties are listed in Table 2.1. The wells are controlled by prescribing an injection rate of $0.002 \text{ m}^3/\text{s}$ in the injection well and bottom hole pressures of $2.5 \cdot 10^7 \text{ Pa}$ in the producers.

Property	Value	Units
Oil viscosity	$0.5 \cdot 10^{-3}$	$\text{Pa} \cdot \text{s}$
Water viscosity	$1.0 \cdot 10^{-3}$	$\text{Pa} \cdot \text{s}$
Total compressibility	$1.0 \cdot 10^{-7}$	$1/\text{Pa}$
Initial reservoir pressure	$3.0 \cdot 10^7$	Pa
Initial water saturation	0.2	—
End point relative permeability, oil	0.9	—
End point relative permeability, water	0.6	—
Corey exponent, oil	2.0	—
Corey exponent, water	2.0	—
Residual oil saturation	0.2	—
Connate water saturation	0.2	—

Table 2.1: Reservoir and fluid properties.

2.3.2 Synthetic noisy measurements

The simulator is run from time $t_0 = 0$ [days] to $t_{\text{end}} = 900$ [days] providing synthetic measurements after every 30 days of production. We assume that it is possible to measure bottom hole pressure in the injector, and oil and water fluid flow rates in the producers. Therefore the output vector \mathbf{y}_n becomes

$$\mathbf{y}_n = \mathcal{H}(\mathbf{x}_n) = \begin{bmatrix} p_{bh}^{INJ} \\ q_o^{NW} \\ q_o^{NE} \\ q_o^{SW} \\ q_o^{SE} \\ q_w^{NW} \\ q_w^{NE} \\ q_w^{SW} \\ q_w^{SE} \end{bmatrix},$$

where \mathcal{H} is the non-linear observation operator that produces measurements from a given state vector \mathbf{x} predicted by the model.

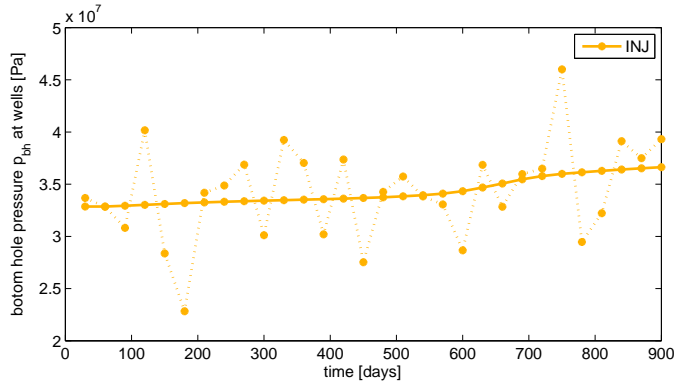
These 'perfect' measurements are transformed to noisy measurements by adding noise to the 'true' values of \mathbf{y} according to

$$\mathbf{y}_n = \mathcal{H}(\mathbf{x}_n) + \boldsymbol{\xi}_n,$$

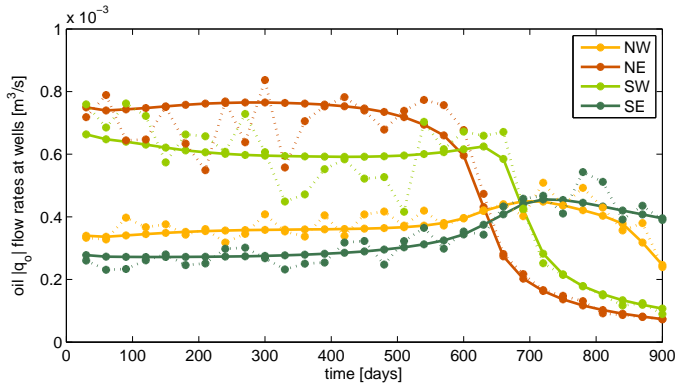
where $\boldsymbol{\xi}_n$ is Gaussian noise with mean zero and diagonal covariance matrix $\mathbf{R}_n \in \mathbb{R}^{9 \times 9}$, such that $\mathbf{R}_n^{i,i} = 0.1 \cdot |(\mathcal{H}(\mathbf{x}_n))^i|$, $i = 1, \dots, 9$, with superscript i denoting the relevant matrix or vector element. This strategy implies that the error in each observable variable is taken to be 10% of its actual value. The same covariance matrix is then used to represent the measurement noise in the data assimilation exercise. If the value of a generated synthetic measurement is infeasible, it is discarded. An example of 'perfect' and noisy measurements is given in Figure 2.2.

2.3.3 'Uncertain' model

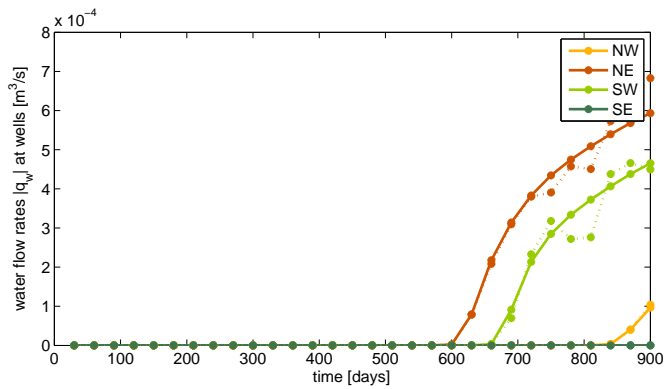
The data assimilation algorithm has to be initialized by generating an initial augmented state. We assume that the initial state (grid block pressures and water saturations) is equal to the initial condition of the 'true' model and does not contain any error. Moreover, we consider all inputs, reservoir dimensions, and all parameters describing rock and fluid properties to be known without uncertainty except for the grid block permeabilities \mathbf{k} and the grid block porosities ϕ . We however assume that $\ln \mathbf{k}$ and $\ln(-\ln \phi)$ are normally



(a) bottom hole pressure



(b) oil flow rate



(c) water flow rate

Figure 2.2: Example of synthetic measurements: solid line — 'perfect', dotted line — noisy measurements.

distributed. Finally, we avoid any model discretization error by using an 'uncertain' model for the data assimilation with the same number and geometry of grid blocks as the 'truth' model. These assumptions provide us with a very simple test case that allows to assess the performance of the algorithms under controlled conditions.

The mean and the spatial covariance matrices of the uncertain parameter fields \mathbf{k} and ϕ are obtained from an ensemble of 1000 reservoir models generated with an in-house developed geostatistical algorithm. The algorithm is based on ordinary principal component analysis of a training image (similar to [74]) and a matrix partitioning method [19] to constrain the parameter values at the well bore location to known values. The mean values of the permeability and porosity ensembles (depicted in Figure 2.3) are used as prior parameter fields for the data assimilation procedure. The two ensemble members that are selected as 'true' permeability \mathbf{k} and porosity ϕ fields are excluded from the mean and covariance computations throughout the assimilation procedure.

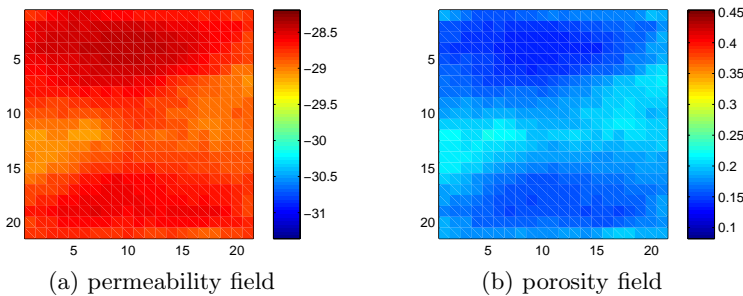


Figure 2.3: Prior permeability and porosity fields. The scale represents transformed values of permeability [m^2] and porosity [-].

The above twin experiment set-up is used as a base test case throughout the study, unless specified otherwise.

Representer and accelerated representer algorithms

3.1 Introduction

A particular variational method is the representer method (RM). The original RM was introduced by Bennett [8] for oceanography-related problems and designed to perform a single minimization of an objective function over a long assimilation time period using all available observations (typically a relatively small amount). The algorithm's key feature is the computation of a set of so-called representer describing the influence of a certain measurement on an estimation of the state and/or parameter. Traditionally, an important benefit of the RM method was claimed to be its capacity to account for model errors by imposing a model as a 'weak constraint'. However, e.g. [3, 80] have shown that the addition of model errors is possible also in other variational data assimilation schemes, although it may result in time consuming iterative procedures. For linear assimilation problems, the RM provides the answer to the weak constraint problem using a one-shot approach, but for non-linear problems this benefit is lost because iteration is required anyway. Computationally, the RM can be interpreted as a specific implementation of the Gauss-Newton method for minimization of an objective function. Moreover, the representer can be seen as sensitivity matrices that quantify the sensitivity of observations to changes in linear combinations of model states or parameters. For linear applications, it can be shown that these sensitivity matrices, and thus the representer, are cross-covariance matrices between the observations and the states or parameters [7]. The representer can also be interpreted as data-driven basis functions to reparameterize a high-dimensional system model in terms of the product of a small number of representer coefficients multiplying the same small number of representer [70].

Initially developed for the linear models, the algorithm was revised in [71] for non-linear models providing an iterative technique that uses the RM to solve a linearized problem at each iteration. Examples of the applications of the RM to parameter estimation in reservoir models are described in [5, 6, 35, 64, 70].

However, the RM encounters computational efficiency problems when used for systems with large amounts of measured data. This fact initiated a discussion on accelerating computations in [7]. Because meteorological measurements nowadays become available for assimilation near-continuously, thus tremendously increasing the computational workload, within meteorology Xu and co-authors [89] developed an accelerated version of the representer method (ARM) where direct computation of representer is avoided.

The features of the models and measurements used in reservoir engineering bear a similarity to those in ocean and atmospheric data assimilation. In particular, an increasing amount of measurements is becoming available, e.g. in the form of time-lapse seismic. That fact provided us with a motivation to test the ARM in reservoir engineering applications and compare its performance to that of the classical RM in terms of accuracy and computational speed. The methods are used for assimilating production data to estimate the permeability and porosity fields in a two-phase two-dimensional fluid flow model.

3.2 The RM and the ARM

The section addresses solving an inverse problem formulated in terms of minimizing an objective (cost) function of least-squares type. This approach employs the method of Lagrange multipliers that yields a coupled Euler-Lagrange system of equations, which are thereafter decoupled by means of the RM. The method was introduced in [8] within oceanography and has a continuous form in space and time. This section describes the algorithm in a manner similar to [87] following a meteorological convention, in which the relevant equations are first discretized in space and time.

3.2.1 Generalized cost function

Under the fundamental assumption that initial errors, model errors and observation errors are normally distributed, and considering model errors and measurement errors to be uncorrelated in time, the inverse problem can be viewed as a minimization of the following objective function:

$$J = J_0^p + J^q + J^r, \quad (3.1)$$

where

$$\begin{aligned} J_0^p &= \frac{1}{2} [\mathbf{x}_0^p - \mathbf{x}_0]^T \mathbf{P}_0^{-1} [\mathbf{x}_0^p - \mathbf{x}_0], \\ J^q &= \frac{1}{2} \sum_{n=1}^N [\mathbf{x}_n - \mathcal{M}(\mathbf{x}_{n-1})]^T \mathbf{Q}_n^{-1} [\mathbf{x}_n - \mathcal{M}(\mathbf{x}_{n-1})], \\ J^r &= \frac{1}{2} \sum_{n=1}^N [\mathbf{y}_n - \mathcal{H}(\mathbf{x}_n)]^T \mathbf{R}_n^{-1} [\mathbf{y}_n - \mathcal{H}(\mathbf{x}_n)] \end{aligned}$$

for the time period $t_0 \leq t_n \leq t_N$. J_0^p , J^q and J^r are the cost functions of the initial, model and observation errors respectively. Superscript p refers to the prior, subscripts 0 and n stand for the initial and the n th model time steps. The non-linear model \mathcal{M} is assumed to have a state vector of length I and N time steps with $1 \leq n \leq N$, \mathbf{x}_n is a state vector in the model space, and \mathbf{Q}_n is a rank I model error covariance matrix at time $t = t_n$. The prior \mathbf{x}_n^p is a forecast of length I at time $t = t_n$, \mathbf{P}_0 is a rank I covariance matrix of initial errors. A vector of observations of length k_n at a time $t = t_n$ is denoted by \mathbf{y}_n , where $\sum_{n=1}^N k_n = K$. Therefore K is the dimension of the full observation vector. \mathcal{H} is a non-linear observation operator that produces k_n measurements predicted by the model from a given state vector \mathbf{x}_n at time $t = t_n$ and \mathbf{R}_n is a rank k_n observation error covariance matrix at a time $t = t_n$.

The non-linear model \mathcal{M} is assumed to be imperfect:

$$\mathbf{x}_n - \mathcal{M}(\mathbf{x}_{n-1}) = \boldsymbol{\epsilon}_n, \quad (3.2)$$

where $\boldsymbol{\epsilon}_n$ represents model error forcing at time $t = t_n$. With the aid of the method of Lagrange multipliers, imposing (3.2) to the problem of minimizing cost function (3.1) leads to the following optimization problem:

$$\begin{aligned} J &= \frac{1}{2} [\mathbf{x}_0^p - \mathbf{x}_0]^T \mathbf{P}_0^{-1} [\mathbf{x}_0^p - \mathbf{x}_0] \\ &+ \frac{1}{2} \sum_{n=1}^N \boldsymbol{\epsilon}_n^T \mathbf{Q}_n^{-1} \boldsymbol{\epsilon}_n \\ &+ \frac{1}{2} \sum_{n=1}^N [\mathbf{y}_n - \mathcal{H}(\mathbf{x}_n)]^T \mathbf{R}_n^{-1} [\mathbf{y}_n - \mathcal{H}(\mathbf{x}_n)] \\ &+ \sum_{n=1}^N \boldsymbol{\lambda}_n^T [\mathbf{x}_n - \mathcal{M}(\mathbf{x}_{n-1}) - \boldsymbol{\epsilon}_n]. \end{aligned} \quad (3.3)$$

Taking the first variation of function (3.3) yields

$$\delta J = \frac{\partial J}{\partial \mathbf{x}_0} \delta \mathbf{x}_0 + \sum_{n=1}^N \frac{\partial J}{\partial \mathbf{x}_n} \delta \mathbf{x}_n + \sum_{n=1}^N \frac{\partial J}{\partial \boldsymbol{\epsilon}_n} \delta \boldsymbol{\epsilon}_n + \sum_{n=1}^N \frac{\partial J}{\partial \boldsymbol{\lambda}_n} \delta \boldsymbol{\lambda}_n,$$

hence,

$$\begin{aligned}
\delta J = & \{-(\mathbf{x}_0^p - \mathbf{x}_0)^T \mathbf{P}_0^{-1} - \boldsymbol{\lambda}_1^T \mathbf{M}_0\} \delta \mathbf{x}_0 \\
& + \sum_{n=1}^N \{-(\mathbf{y}_n - \mathcal{H}(\mathbf{x}_n))^T \mathbf{R}_n^{-1} \mathbf{H}_n\} \delta \mathbf{x}_n + \sum_{n=1}^N \{\boldsymbol{\lambda}_n^T - \boldsymbol{\lambda}_{n+1}^T \mathbf{M}_n\} \delta \mathbf{x}_n \\
& + \sum_{n=1}^N \{\boldsymbol{\epsilon}_n^T \mathbf{Q}_n^{-1} - \boldsymbol{\lambda}_n^T\} \delta \boldsymbol{\epsilon}_n \\
& + \sum_{n=1}^N \{\mathbf{x}_n - \mathcal{M}(\mathbf{x}_{n-1}) - \boldsymbol{\epsilon}_n\}^T \delta \boldsymbol{\lambda}_n,
\end{aligned} \tag{3.4}$$

where the matrix $\mathbf{M}_n = \frac{\partial \mathcal{M}(\mathbf{x}_n)}{\partial \mathbf{x}_n}$ is the $I \times I$ Jacobian matrix corresponding to the model operator $\mathcal{M}(\mathbf{x}_n)$ and the $k_n \times I$ matrix $\mathbf{H}_n = \frac{\partial \mathcal{H}(\mathbf{x}_n)}{\partial \mathbf{x}_n}$ is the Jacobian matrix for observation operator $\mathcal{H}(\mathbf{x}_n)$. For later convenience, a 'dummy' variable $\boldsymbol{\lambda}_{N+1}$ is chosen equal to $\mathbf{0}$.

If \mathbf{x}^a is the analysis (or the estimated) state, i.e. the value of \mathbf{x} that minimizes J , then $\delta J|_{\mathbf{x}=\mathbf{x}^a} = 0$. Under this requirement, it follows from (3.4) that

$$\boldsymbol{\lambda}_n - \mathbf{M}_n^T \boldsymbol{\lambda}_{n+1} = \mathbf{H}_n^T \mathbf{R}_n^{-1} (\mathbf{y}_n - \mathcal{H}(\mathbf{x}_n^a)), \quad n = 1, \dots, N-1, \tag{3.5}$$

subject to $\boldsymbol{\lambda}_N = \mathbf{H}_N^T \mathbf{R}_N^{-1} (\mathbf{y}_N - \mathcal{H}(\mathbf{x}_N^a))$ and

$$\mathbf{x}_n^a - \mathcal{M}(\mathbf{x}_{n-1}^a) = \mathbf{Q}_n \boldsymbol{\lambda}_n, \quad n = 1, \dots, N, \tag{3.6}$$

subject to $\mathbf{x}_0^a = \mathbf{x}_0^p + \mathbf{P}_0 \mathbf{M}_0^T \boldsymbol{\lambda}_1$. Note that $\mathbf{Q}_n \boldsymbol{\lambda}_n = \boldsymbol{\epsilon}_n$, $n = 1, \dots, N$.

Expression (3.5) is an equation for adjoint variable $\boldsymbol{\lambda}_n$, $n = 1, \dots, N$, and has to be integrated backward in time, whereas (3.6) integrates the analysis variable \mathbf{x}^a forward in time. Equations (3.5) and (3.6) form a coupled system that cannot be solved directly. The next subsection outlines a decoupling strategy for the system (3.5)–(3.6) via the RM.

3.2.2 The RM: linear and non-linear cases

Under the simplifications introduced in the previous section and assumptions of model and observation operator linearity (i.e. $\mathcal{M}(\mathbf{x}_n) = \mathbf{M}_n \mathbf{x}_n$ and $\mathcal{H}(\mathbf{x}_n) = \mathbf{H}_n \mathbf{x}_n$) the Euler-Lagrange equations (3.5)–(3.6) yield

$$\boldsymbol{\lambda}_n - \mathbf{M}_n^T \boldsymbol{\lambda}_{n+1} = \mathbf{H}_n^T \mathbf{R}_n^{-1} (\mathbf{y}_n - \mathbf{H}_n \mathbf{x}_n^a), \quad n = 1, \dots, N-1, \tag{3.7}$$

subject to $\boldsymbol{\lambda}_N = \mathbf{H}_N^T \mathbf{R}_N^{-1} (\mathbf{y}_N - \mathbf{H}_N \mathbf{x}_N^a)$ and

$$\mathbf{x}_n^a - \mathbf{M}_n \mathbf{x}_{n-1}^a = \mathbf{Q}_n \boldsymbol{\lambda}_n, \quad n = 1, \dots, N, \tag{3.8}$$

subject to $\mathbf{x}_0^a = \mathbf{x}_0^p + \mathbf{P}_0 \mathbf{M}_0^T \boldsymbol{\lambda}_1$.

The prior state estimate \mathbf{x}^p is obtained via model forecast

$$\mathbf{x}_n^p = \mathbf{M}_n \mathbf{x}_n + \boldsymbol{\epsilon}_n^p, \quad n = 1, \dots, N,$$

subject to the initial condition \mathbf{x}_0^p . Here $\boldsymbol{\epsilon}_n^p$ is a prior estimate of model noise.

Consider K representer functions γ_k , $1 \leq k \leq K$, with the adjoints $\boldsymbol{\alpha}_k$, $1 \leq k \leq K$, satisfying

$$\{\boldsymbol{\alpha}_k\}_n - \mathbf{M}_n^T \{\boldsymbol{\alpha}_k\}_{n+1} = \{\mathbf{H}_k^T\}_n, \quad n = 1, \dots, N, \quad (3.9)$$

and $\{\boldsymbol{\alpha}_k\}_{N+1} = \mathbf{0}$. Then the adjoint field $\boldsymbol{\alpha}_k$ for the k th observation can be obtained by integrating equation (3.9) backwards. Once $\boldsymbol{\alpha}_k$ is known, the representer functions can be found by solving

$$\{\gamma_k\}_n - \mathbf{M}_{n-1} \{\gamma_k\}_{n-1} = \mathbf{Q}_n \{\boldsymbol{\alpha}_k\}_n, \quad n = 1, \dots, N, \quad (3.10)$$

subject to initial condition $\{\gamma_k\}_0 = \mathbf{P}_0 \mathbf{M}_0^T \{\boldsymbol{\alpha}_k\}_1$.

The solution of (3.7)–(3.8) is sought as a sum of a prior and correction terms, whereas the correction is a linear combination of representer functions

$$\mathbf{x}^a = \mathbf{x}^p + \sum_{k=1}^K \gamma_k \beta_k = \mathbf{x}^p + \boldsymbol{\Gamma}^T \boldsymbol{\beta}, \quad (3.11)$$

where matrix $\boldsymbol{\Gamma}^T = (\gamma_1, \dots, \gamma_K)$ contains representers, $\boldsymbol{\beta} = (\beta_1 \dots \beta_k \dots \beta_K)^T$ is a vector of unknown scalar coefficients, and β_k is the representer coefficient for the k th representer function. Note that $\boldsymbol{\Gamma}$ is actually a prior error covariance between the observations and the model states. The representer coefficient vector $\boldsymbol{\beta}$ can be computed as

$$\boldsymbol{\beta} = [\mathbf{H} \boldsymbol{\Gamma}^T + \mathbf{R}]^{-1} [\mathbf{y} - \mathbf{H} \mathbf{x}^p], \quad (3.12)$$

where $\mathbf{y}^T = [\mathbf{y}_1^T \dots \mathbf{y}_n^T \dots \mathbf{y}_N^T]$ is a vector of all observations, \mathbf{R} is a block diagonal matrix with blocks $\mathbf{R}_1, \dots, \mathbf{R}_N$, \mathbf{H} is a block diagonal matrix with blocks $\mathbf{H}_1, \dots, \mathbf{H}_N$. For further details the reader is referred to [87].

Solving the coupled Euler-Lagrange equations (3.5) and (3.6) meets additional obstacles due to the fact that the operators \mathcal{M} and \mathcal{H} are non-linear. This problem can be overcome by producing an iterative sequence of coupled problems, each iteration step linearized about an analyzed basic state from the previous iteration [71]. The linearized problem is thereafter solved via the RM. The algorithm flowchart is presented below (Figure 3.1).

The procedure is expected to converge to an approximation of the non-linear solution if the non-linearities in the system are not too strong.

It turns out that for systems with a large amount of measurements direct computation of representer coefficients (3.12) is inefficient. This fact has led to discussion on accelerating computations in [7], and derivation of the so-called ARM introduced in [88] and further developed in [89].

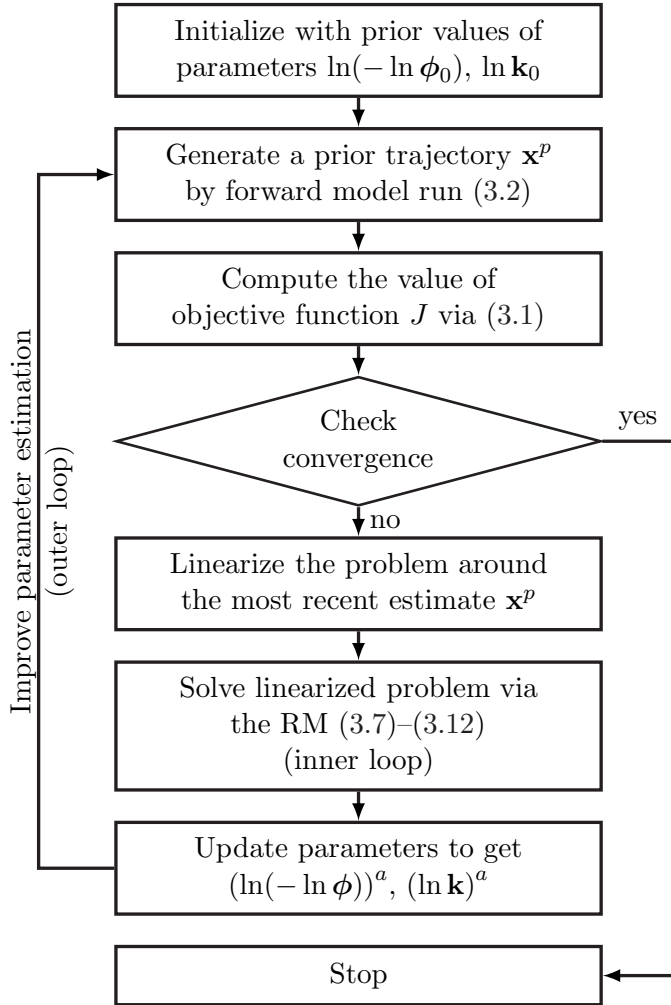


Figure 3.1: Iterative approach to the RM in the non-linear case.

3.2.3 The ARM

To reduce the computational costs while solving the Euler-Lagrange equations (3.7)–(3.8) via the RM, an accelerated procedure can be used. The accelerated algorithm, following [89], aims at improving the performance of the solver $[\mathbf{H}\mathbf{\Gamma}^T + \mathbf{R}]^{-1}[\mathbf{y} - \mathbf{H}\mathbf{x}^p]$ and the post-multiplication with $\mathbf{\Gamma}^T$ in equation (3.11).

The solver $[\mathbf{H}\mathbf{\Gamma}^T + \mathbf{R}]^{-1}[\mathbf{y} - \mathbf{H}\mathbf{x}^p]$

The problem can be described as solving a system of linear equations

$$[\mathbf{H}\mathbf{\Gamma}^T + \mathbf{R}]\mathbf{z} = [\mathbf{y} - \mathbf{H}\mathbf{x}^p] \quad (3.13)$$

with respect to \mathbf{z} , where $\mathbf{z}^T = [\mathbf{z}_1^T \dots \mathbf{z}_n^T \dots \mathbf{z}_N^T]$ and \mathbf{z}_n is a vector of length k_n corresponding to time t_n . Equation (3.13) is usually solved iteratively using descent methods such as the conjugate gradient (CG) algorithm. Within these algorithms a matrix-vector multiplication has to be performed at each iteration:

$$\mathbf{v} = [\mathbf{H}\mathbf{\Gamma}^T + \mathbf{R}]\mathbf{w},$$

where \mathbf{w} is a known vector of length K and \mathbf{v} is a vector of length K , or

$$\mathbf{v} = \mathbf{v}^* + \mathbf{v}^+ \quad (3.14)$$

with $\mathbf{v}^* = \mathbf{H}\mathbf{\Gamma}^T\mathbf{w}$ and $\mathbf{v}^+ = \mathbf{R}\mathbf{w}$.

Define $(\mathbf{v}^*)^T = [(\mathbf{v}_1^*)^T \dots (\mathbf{v}_n^*)^T \dots (\mathbf{v}_N^*)^T]$ and $\mathbf{w}^T = [\mathbf{w}_1^T \dots \mathbf{w}_n^T \dots \mathbf{w}_N^T]$, where $(\mathbf{v}_n^*)^T$ and \mathbf{w}_n^T are vectors of length k_n with \mathbf{w}_n^T assumed to be known for $n = 1, \dots, N$. Introduce two vectors \mathbf{f}_n and \mathbf{g}_n which are defined for each time t_n and have length I . Then the backward sweep reads as

$$\mathbf{f}_n = \mathbf{M}_n^T \mathbf{f}_{n+1} + \mathbf{H}_n^T \mathbf{w}_n, \quad n = 1, \dots, N-1,$$

subject to $\mathbf{f}_N = \mathbf{H}_N^T \mathbf{w}_N$.

The backward sweep is followed by the forward sweep

$$\mathbf{g}_n = \mathbf{M}_{n-1} \mathbf{g}_{n-1} + \mathbf{Q}_n \mathbf{f}_n, \quad n = 1, \dots, N,$$

subject to $\mathbf{g}_0 = \mathbf{P}_0 \mathbf{M}_0^T \mathbf{f}_1$.

Finally,

$$\mathbf{v}_n^* = \mathbf{H}_n \mathbf{g}_n, \quad n = 1, \dots, N,$$

and computation of \mathbf{v}_n via (3.14) is therefore straightforward.

The post-multiplication $\mathbf{\Gamma}^T$

The post-multiplication reduces to the problem of matrix-vector multiplication $\mathbf{\Gamma}^T \mathbf{z}$. Similarly to the previous section the algorithm consists of backward and forward sweeps, which are

$$\mathbf{f}_n = \mathbf{M}_n^T \mathbf{f}_{n+1} + \mathbf{H}_n^T \mathbf{z}_n, \quad n = 1, \dots, N-1,$$

subject to $\mathbf{f}_N = \mathbf{H}_N^T \mathbf{z}_N$ and

$$\mathbf{g}_n = \mathbf{M}_{n-1} \mathbf{g}_{n-1} + \mathbf{Q}_n \mathbf{f}_n, \quad n = 1, \dots, N,$$

subject to $\mathbf{g}_0 = \mathbf{P}_0 \mathbf{M}_0^T \mathbf{f}_1$.

Finally

$$\mathbf{x}_n^a = \mathbf{x}_n^p + \mathbf{g}_n, \quad n = 0, \dots, N.$$

Given the original forward simulation, a new state estimate \mathbf{x}^a for the ARM can be produced at the costs of two backward and two forward sweeps. The algorithm also avoids direct computation and, hence, storage of the representer matrix $\mathbf{\Gamma}$.

3.2.4 Scaling

Implementation of data assimilation methods and, in particular, representer-based algorithms often meets an additional obstacle, namely, different orders of magnitude of the entries in the state vector and the measurement vector (e.g. pressure values in the order of 10^7 Pa and production rates in the order of 10^{-3} m³/s). This fact becomes crucial when matrix-matrix and matrix-vector multiplications have to be performed. In these operations the small numbers tend to get lost because of finite machine precision, even though they are important for the process. To prevent such a loss of information and even divergence of the algorithm, we propose to rewrite the generalized cost function (3.1) as

$$\begin{aligned} J^* = & \frac{1}{2} [\mathbf{x}_0^p - \mathbf{x}_0]^T \mathbf{P}_0^{-1} [\mathbf{x}_0^p - \mathbf{x}_0] \\ & + \frac{1}{2} \sum_{n=1}^N [\mathbf{x}_n - \mathcal{M}(\mathbf{x}_{n-1})]^T \mathbf{Q}_n^{-1} [\mathbf{x}_n - \mathcal{M}(\mathbf{x}_{n-1})] \\ & + \frac{1}{2} \sum_{n=1}^N \left[\mathbf{R}_n^{-1/2} \mathbf{y}_n - \mathbf{R}_n^{-1/2} \mathcal{H}(\mathbf{x}_n) \right]^T \left[\mathbf{R}_n^{-1/2} \mathbf{y}_n - \mathbf{R}_n^{-1/2} \mathcal{H}(\mathbf{x}_n) \right]. \end{aligned} \quad (3.15)$$

If scaled versions of the measurement vector and the observation operator are considered, i.e. vector $\mathbf{y}_n^* = \mathbf{R}_n^{-1/2} \mathbf{y}_n$ and operator $\mathcal{H}^*(\mathbf{x}_n) = \mathbf{R}_n^{-1/2} \mathcal{H}(\mathbf{x}_n)$, then cost functions (3.15) and (3.1) are of the same form and representer methods can be applied to minimize (3.15). Note that the Jacobian matrix for observation operator $\mathcal{H}^*(\mathbf{x}_n)$ is $\mathbf{H}_n^* = \frac{\partial \mathcal{H}^*(\mathbf{x}_n)}{\partial \mathbf{x}_n} = \mathbf{R}_n^{-1/2} \frac{\partial \mathcal{H}(\mathbf{x}_n)}{\partial \mathbf{x}_n} = \mathbf{R}_n^{-1/2} \mathbf{H}_n$.

It is also possible and sometimes desired to work with the scaled version not only of the measurement vector/operator but also of the state space vector and model. In such a case, a state vector can be premultiplied with e.g. an inverse of a diagonal matrix that has absolute values of elements of the initial state vector on its diagonal [79, 82]. In this study, however, scaling the vector of observations was found to be sufficient to avoid numerical problems.

3.2.5 Assessing the quality of the algorithm performance

To evaluate the quality of the parameter estimate, we consider the value of the cost function at the parameter estimate. Denote this value of the cost function by J^{opt} . It can be shown that in case of linear parameter estimation problems the minimum of $2J$ is approximately distributed as a χ^2 with K degrees of freedom [78]. In practice this result remains valid also for non-linear parameter estimation problems [59]. Hence the expected value of cost function J at a minimum is approximately $K/2$ and the corresponding standard deviation is $\sqrt{K/2}$. Without further discussion, it is assumed in [59] that a reliable

realization of the distribution lies within five standard deviations from the mean and, therefore, the following inequality should hold for the obtained estimate

$$K - 5\sqrt{2K} \leq 2J^{\text{opt}} \leq K + 5\sqrt{2K}. \quad (3.16)$$

We can provide a rather rigorous justification of this fact with the aid of Chebyshev's inequality [31]

$$P(|X - \mu| \geq \alpha\sigma) \leq \frac{1}{\alpha^2}, \quad (3.17)$$

where X is a random variable, μ and σ correspond to its mean and standard deviation respectively, and α is a scalar. In our case $X = J^{\text{opt}}$, $\mu = K/2$, $\sigma = \sqrt{K/2}$ and α is suggested to be equal to 5. Then it follows from inequality (3.17) that a realization of the distribution lies outside five standard deviations from the mean with the probability of at most $\frac{1}{\alpha^2} = \frac{1}{5^2} = 0.04$, hence, in 96% cases satisfying (3.16). Thus if the estimated model parameters imply a value of J^{opt} that does not satisfy (3.16), their quality should be questioned.

In a synthetic case, moreover, there is an opportunity to compare estimated values of the parameters with the 'true' ones that were used to initialize the experiment. For that matter a root mean square (rms) error measure is defined as

$$E_{\text{rms}}(\mathbf{z}) = \sqrt{\frac{\|\mathbf{z} - \mathbf{z}_{\text{true}}\|_2^2}{I/4}}, \quad (3.18)$$

where \mathbf{z} stands for the vector of parameters of interest, e.g. permeability.

The performance of a data assimilation procedure can also be characterized by the quality of predictions obtained by using the history-matched model to forecast behavior of the physical process beyond the history-matched period. In particular, we will consider the prognosis of the water breakthrough time in the production wells.

3.3 Results and discussion

The study is divided into two parts. First a history matching problem is solved with the classical representer method and the accelerated representer method to ensure that under certain conditions both algorithms provide estimates of the same accuracy. Thereafter the computational performances of the RM and the ARM are compared.

3.3.1 The RM as iterative procedure

We start by performing a history matching experiment with the RM. The data assimilation is performed from time $t_0 = 0$ [days] till $t_{\text{end}} = 450$ [days],

which ensures that water breakthrough occurs in none of the production wells. The data is assimilated every 15 days for 20 times and since each measurement vector contains 9 observations, a total of 180 measurements is history-matched.

Table 3.1 displays the results of the experiment in terms of measures (3.16)–(3.18).

	Value of cost function	$E_{\text{rms}}(\phi)$	$E_{\text{rms}}(\mathbf{k})$
Initially	1948.881	0.12	1.16
1 st outer loop	117.840	0.07	0.60
2 nd outer loop	79.824	0.06	0.53
3 rd outer loop	79.312	0.06	0.53
4 th outer loop	79.305	0.06	0.53
5 th outer loop	79.304	0.06	0.53

Table 3.1: Results of history matching using the RM.

The largest drop in the value of the cost function occurs after the first iteration, while the subsequent iterations keep steering the cost function towards a local minimum. Overall, the cost function drops from 1948.881 to 79.304. Moreover, the computed minimum value of the cost function is close to the expected minimum value of 90 and satisfies criterion (3.16), which in this particular case has the form $42.567 \leq J^{\text{opt}} \leq 137.434$. Since we are running a synthetic experiment, we have an opportunity to compare the estimated permeability and porosity fields to the 'true' ones in terms of the rms error. The changes in rms error follow the same trend as the changes in the value of cost function. In particular, the biggest improvement is again obtained within the 1st outer loop of the algorithm. Overall, the rms error for permeability is reduced from 1.16 to 0.53 and for porosity from 0.12 to 0.06.

In line with these results, the updated permeability and porosity fields already display the main features of the 'true' fields after the first iteration (compare Figure 3.2a vs. 2.1a and Figure 3.3a vs. 2.1b). In particular, the high permeability (and porosity) streak between the SW and NE production wells can be recognized. Subsequent outer loops keep improving the details of the earlier determined structure (see Figures 3.2 and 3.3).

An important aspect of reservoir simulation is the accurate prediction of water breakthrough in the production wells. In our case, the prediction is obtained by running the model forward to time moment $t = 1500$ [days] with parameters as estimated at $t_{\text{end}} = 450$ [days]. Figure 3.4 illustrates that the prediction of water breakthrough is much more accurate when the calibrated model is used than when the initial model is applied.

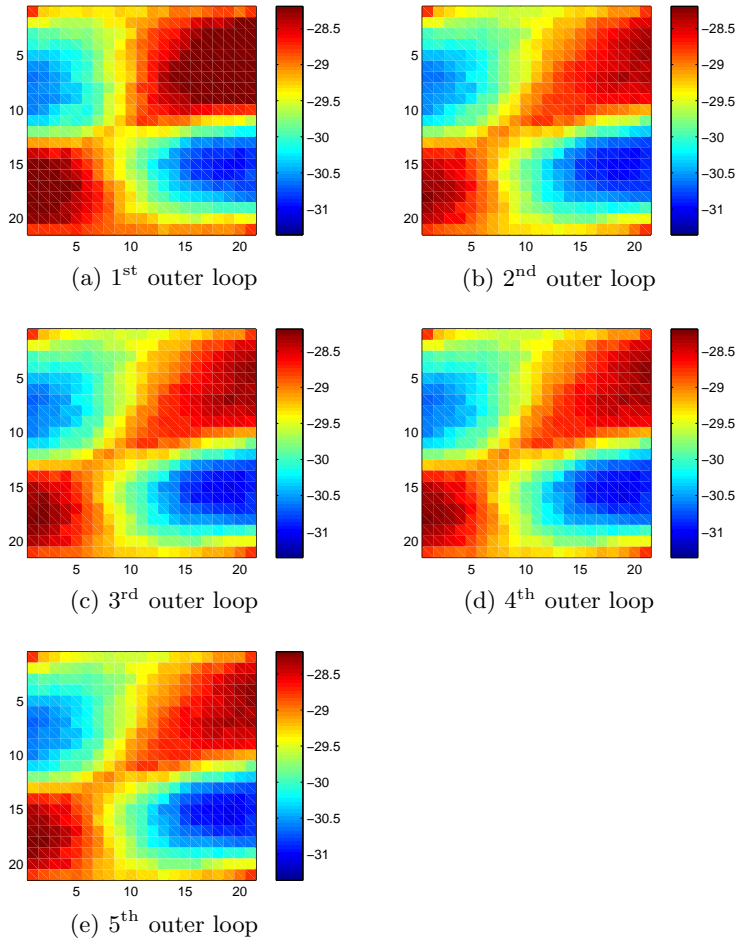


Figure 3.2: Estimated permeability field. The scale represents transformed values of permeability [m^2].

3.3.2 Tuning the ARM

To compare the computational performance of the RM and the ARM, it is first necessary to adjust the ARM in such a way that it produces results of the same accuracy and in the same number of outer loops as the RM. This can be done by selecting the appropriate termination criterion used to cease the conjugate gradient loop while solving equation (3.13).

For that purpose we ran one outer loop of the RM assimilating data at 20 times and one outer loop of the ARM with various termination criteria for the conjugate gradient solver using the same data. The experiment was repeated ten times and the averaged results are displayed in Figure 3.5.

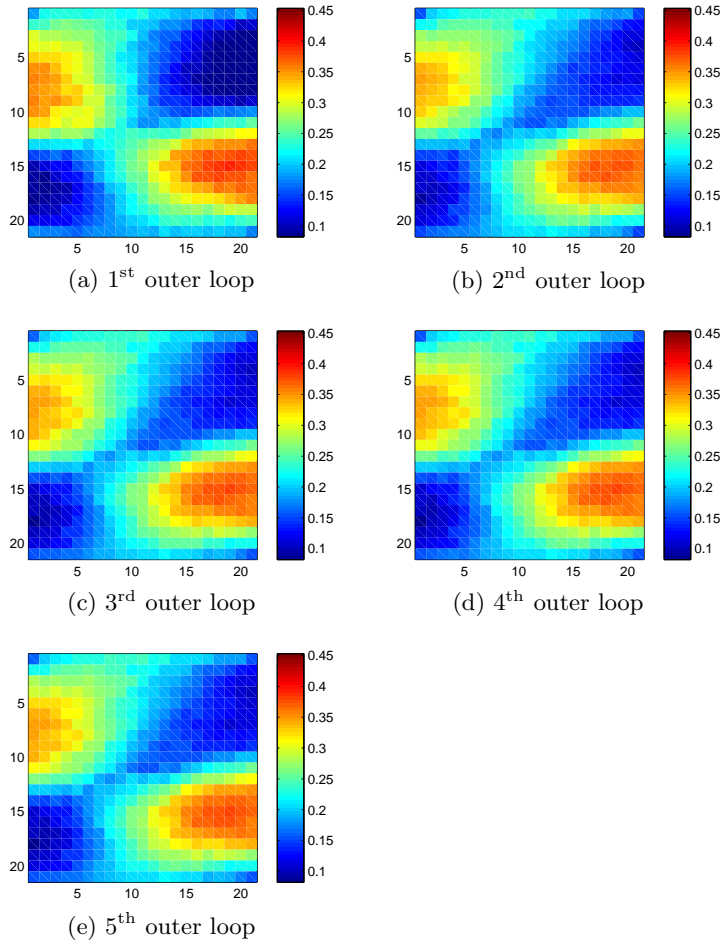


Figure 3.3: Estimated porosity field. The scale represents transformed values of porosity [-].

Naturally, a run of the ARM requires more time if a smaller value is used for the stopping criterion (Figure 3.5a). At the same time, the accuracy of the value of the cost function improves. It can be observed that if 10^{-2} is chosen as termination criterion, the ARM provides a cost function which value is reasonably close to the one obtained by the classical RM (Figure 3.5b). Moreover, Figures 3.5c and 3.5d illustrate that for this value of the termination criterion the parameters estimated by the two methods are of the same quality. Therefore, we used a termination criterion equal to 10^{-2} for the remainder of the study.

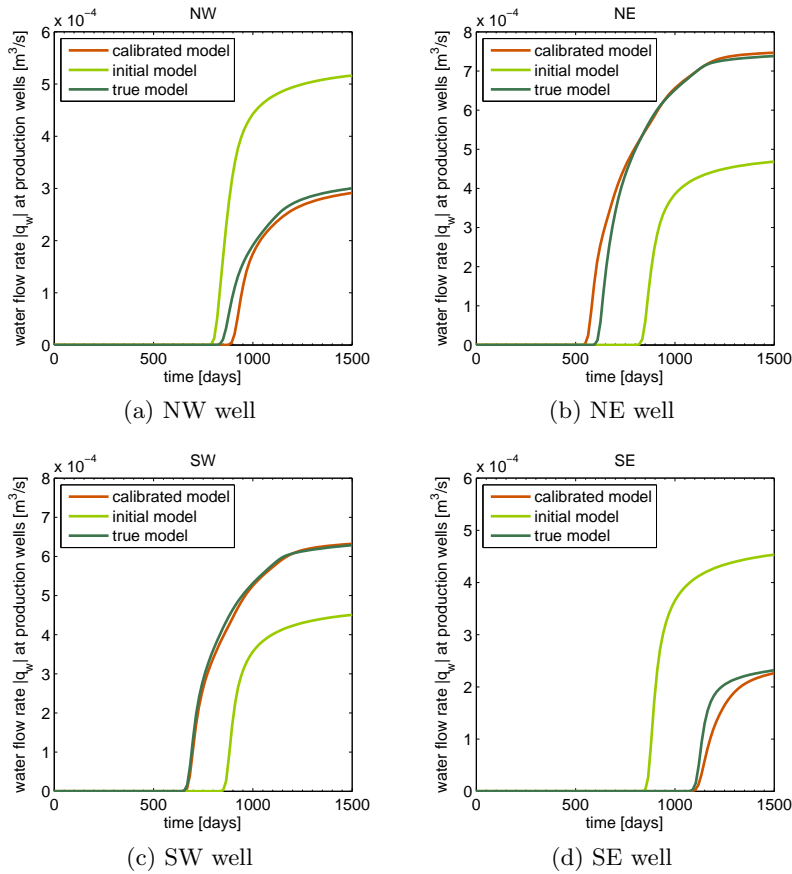


Figure 3.4: Forecast of water flow rate at production wells.

3.3.3 Comparison of computational performance

We first consider the approximate amount of elementary operations (summation and multiplication of two numbers) that have to be accomplished within one outer loop of the RM and the ARM. The computations are summarized in Tables 3.2 and 3.3. We discuss only the operations that correspond to the core of the methods and that are different for the algorithms.

Under the measurement strategy that we apply (namely, nine measurements are collected at a time, i.e. $K = 9N$ or $N = K/9$), the computational load of the RM is driven by terms of magnitude K^3 and K^4 , whereas the number of operations for the ARM grows as K^3 for large amounts of data. Using the estimates of the total number of operations performed by both methods (Tables 3.2 and 3.3) we can expect that for small amounts of measurements the original RM is faster than the ARM, but that for $K \geq 13$ the ARM

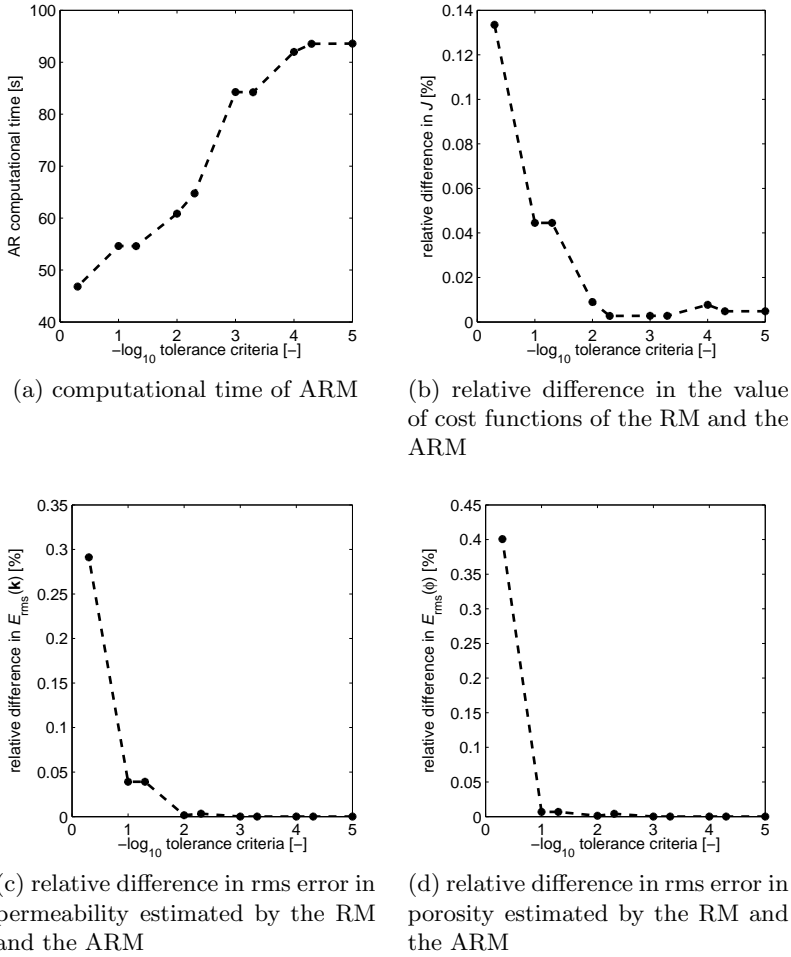


Figure 3.5: Different performance measures of the ARM vs. the stopping criterion for terminating the conjugate gradient loop.

demonstrates its efficiency. Indeed, the results of the computational experiment presented in Figure 3.6a confirm that the ARM starts outperforming the classical RM as the amount of assimilated data points increases. However, it can be observed that the ARM only outperforms the RM when more than 85 observations are used for the experiment, which is a considerably higher number than the 13 observations expected from the operations count.

The main reason for this delayed performance is that due to implementation matters, the ARM needs to accomplish some extra work to load the model Jacobians. In particular, there have to be loaded $2N$ and $2N + 2NK$ Jacobians within one outer loop of the RM and the ARM respectively. This

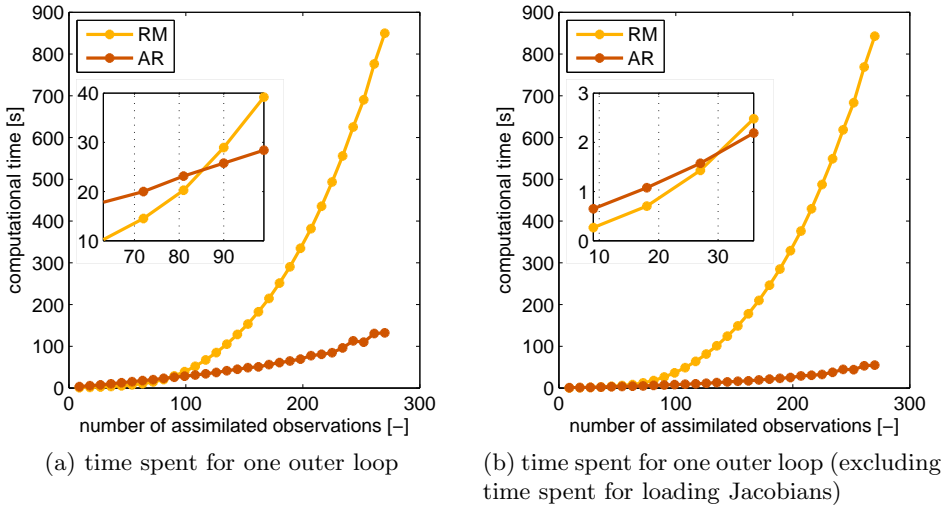


Figure 3.6: Computational time of the RM and the ARM vs. number of measurements.

effect is illustrated in Figure 3.6b, in which time spent by both methods for loading Jacobians has been excluded from the comparison.

3.4 Conclusions

We compared the (computational) performance of the classical RM and the ARM as applied to gradient-based history matching. In particular, we performed a twin experiment in which we estimated the uncertain permeability and porosity fields in a small two-phase two-dimensional subsurface fluid flow model from noisy production measurements in the wells. First the RM was tested and was found to provide a reasonable parameter estimate in such a case. Thereafter the possibilities to tune the ARM to produce results of the same accuracy as the original RM were investigated. It was found that the accuracy of the ARM can be controlled by the termination criterion for the conjugate gradient loop in the inner loop of the optimizer. When the value of this criterion was set to 10^{-2} (or smaller) both methods provided outcomes of the same accuracy at each iteration.

Finally, the computational performance of the two procedures was compared. The results indicate that the ARM outperforms the classical RM in terms of computational speed when the number of assimilated measurements increases. In our particular example the ARM became more efficient than the RM when more than 85 individual measurements had to be history-matched.

This conclusion makes ARM to be a promising tool for history matching applications, especially for cases where large amounts of data have to be assimilated.

Routine	Subroutine	Number of operations
Backward run (3.9)	$\{\alpha_k\}_n - \mathbf{M}_n^T \{\alpha_k\}_{n+1} = \{\mathbf{H}_k^T\}_n,$ $n = 1, \dots, N-1$	$2I^2(N-1)K$
Forward run (3.10)	$\{\gamma_k\}_n - \mathbf{M}_{n-1} \{\gamma_k\}_{n-1} = \mathbf{Q}_n \{\alpha_k\}_n,$ $n = 1, \dots, N$	$[2(2I-1) + 1]INK$
	$\{\gamma_k\}_0 = \mathbf{P}_0 \mathbf{M}_0^T \{\alpha_k\}_1$	$2(2I-1)IK$
Solver (3.12) via CG	$[\mathbf{H}\mathbf{\Gamma}^T + \mathbf{R}]\beta = \mathbf{y} - \mathbf{H}\mathbf{x}^p$	$\{2[2(2K-1) + 1] + 4K$ $+ (2K-1)K + I(N+1)(2K-1)K$ $+ [2I(N+1) - 1]K^2\}K$
Correction term (3.11)	$\mathbf{\Gamma}^T \beta$	$I(N+1)(2K-1)$
In total	$(1 + 4I + 4IN)K^3 + (11 - I - IN)K^2$ $+ (-2 + 2I^2 + IN + 6I^2N)K - IN - I$	

Table 3.2: Number of operations per outer loop of the RM.

Routine	Subroutine	Number of operations
Prerequisites (3.14), etc.		$\{2(2K - 1) + 1 + 4K + (2K - 1)K$ $+ [2I(N + 1) - 1]K$ $+ 2(2K - 1) + 1\}K$
The solver backward sweep	$\mathbf{f}_n = \mathbf{M}_n^T \mathbf{f}_{n+1} + \mathbf{H}_n^T \mathbf{w}_n,$ $n = 1, \dots, N - 1$	$\{I(N + 1)(2K - 1)$ $+ [(2I - 1) + 1]I(N - 1)\}K$
	$\mathbf{f}_N = \mathbf{H}_N^T \mathbf{w}_N$	$(2I - 1)IK$
The solver forward sweep	$\mathbf{g}_n = \mathbf{M}_{n-1} \mathbf{g}_{n-1} + \mathbf{Q}_n \mathbf{f}_n,$ $n = 1, \dots, N$	$[2(2I - 1) + 1]INK$
	$\mathbf{g}_0 = \mathbf{P}_0 \mathbf{M}_0^T \mathbf{f}_1$	$(2I - 1)IK$
Post-multiplication	$\mathbf{f}_n = \mathbf{M}_n^T \mathbf{f}_{n+1} + \mathbf{H}_n^T \mathbf{z}_n,$ $n = 1, \dots, N - 1$	$I(N + 1)(2K - 1) + [(2I - 1) + 1]I(N - 1)$
	$\mathbf{f}_N = \mathbf{H}_N^T \mathbf{z}_N$	$(2I - 1)I$
	$\mathbf{g}_n = \mathbf{M}_{n-1} \mathbf{g}_{n-1} + \mathbf{Q}_n \mathbf{f}_n,$ $n = 1, \dots, N$	$[2(2I - 1) + 1]IN$
	$\mathbf{g}_0 = \mathbf{P}_0 \mathbf{M}_0^T \mathbf{f}_1$	$(2I - 1)I$
In total	$2K^3 + (10 + 4I + 4IN)K^2$ $+ (-2 - I + 2I^2 + 6I^2N)K + 6I^2N - 2IN + 2I^2 - 3I$	

Table 3.3: Number of operations per outer loop of the ARM.

An ensemble Kalman filter and its modifications

4.1 Introduction

As follows from Chapter 3 representer-based history matching algorithms in particular (and variational data assimilation techniques in general) do not allow continuous model updating. Namely, as new data become available, 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 or flow rates increases. This fact yields an increase of data output frequency and creates a problem of incorporating the obtained data in the model as soon as it becomes available so that the model is always up-to-date.

Kalman filtering is known as the most popular methodology for assimilating new measurements to continuously update the state of the system. Originally, the Kalman filter was developed for operating on 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 size of the state vector is too large, application of the extended Kalman filter also meets difficulties. These difficulties can be overcome by applying the ensemble Kalman filtering (EnKF) algorithm based on a Monte-Carlo approach. The EnKF schemes do not require availability of the model adjoint, which makes them very attractive for data assimilation with large-scale complex non-linear models. This chapter focuses on examining the usage and the applicability of a number of ensemble Kalman filtering techniques to history matching.

As the great majority of the problems in reservoir engineering are highly non-linear and characterized by a large number of variables, the EnKF has been a natural choice for a wide range of history matching studies (see e.g. [29, 86]). Papers [28, 51] report the results of using the EnKF to history match

a PUNQ-S3 model. These studies not only demonstrate that the EnKF is successful in assimilating production data to update an initial reservoir model and that its application allows reducing computational costs for history matching, but also identify directions for further investigations and improvements.

Specifically, the research described in [90] has shown that for some non-linear models the EnKF does not provide completely acceptable characterizations of the uncertainties. The situation becomes more problematic if the a priori information on the reservoir structure is poor and the initial guess about the system state is far from the actual one. This leads to the idea of using EnKF modifications, namely iterative EnKF schemes, which aim at obtaining an ensemble that provides an improved representation of the state distribution. There exist several approaches in the petroleum engineering literature, e.g. the ad-hoc confirming EnKF method proposed by [86], the iterative EnKF method analyzed by [67] from an optimization point of view instead of a Monte Carlo sampling methodology, and the ensemble randomized maximum likelihood filter developed by [30]. Inspired by investigations on the iterative extended Kalman filtering approach [39], the authors of [44] suggest to iterate the filter globally. The proposed iterative technique allows to restart the procedure with a new initial guess that is closer to the actual solution and, hence, requires less improvement by the algorithm while providing better estimation of the parameters.

Continuous data assimilation is unfortunately not always feasible for large-scale forecasting applications because each assimilation interrupts ensemble integration, causes an update of the ensemble and a restart. These operations can become too expensive if performed frequently. The authors of [72] develop a modification of the EnKF — known as an asynchronous EnKF (AEnKF) — that allows assimilating collected observations in batches over a number of time windows and, hence, less often. The assimilation is then performed at times different to the times when measurements become available, therefore [72] takes measures to account for the evolution of the state and state error covariance over the length of a particular time window. The problem of asynchronous data assimilation also arises in reanalysis, when it is often desirable to improve the performed analysis by taking advantage of relationships between the observations in time, and usage of the variational data assimilation techniques is not possible due to unavailability of adjoint.

This chapter discusses the AEnKF and the classical EnKF algorithms applied to estimating the permeability and porosity fields in a two-phase two-dimensional fluid flow model.

4.2 The ensemble Kalman filter

Kalman filtering is a powerful technique designed for solving data assimilation problems. This section presents the general idea of Kalman filtering in a manner similar to [76] and of ensemble Kalman filtering as given in [23]. Let us restrict ourselves to the case of the following linear system:

$$\mathbf{x}_{n+1} = \mathbf{M}_n \mathbf{x}_n + \mathbf{B}_n \mathbf{u}_n + \mathbf{G}_n \mathbf{w}_n, \quad (4.1)$$

$$\mathbf{y}_n = \mathbf{H}_n \mathbf{x}_n + \mathbf{v}_n, \quad (4.2)$$

where $\mathbf{M}_n, \mathbf{B}_n, \mathbf{G}_n, \mathbf{H}_n$ are matrices, n is the time index, \mathbf{x}_n denotes the state of the system, \mathbf{u}_n is the system input, \mathbf{y}_n is the vector of measurements, \mathbf{w}_n is Gaussian white system noise process with zero mean and covariance matrix \mathbf{Q}_n , \mathbf{v}_n is Gaussian white measurement noise with zero mean and covariance matrix \mathbf{R}_n . The initial state is assumed to be Gaussian with mean \mathbf{x}_0 and covariance matrix \mathbf{P}_0 . Moreover, the variables \mathbf{x}_0 , \mathbf{w}_n and \mathbf{v}_n are assumed to be independent from each other.

Vector \mathbf{x}_n which contains information on the current system state cannot be directly observed. However it is possible to measure \mathbf{y}_n which is some function of \mathbf{x}_n affected by noise \mathbf{v}_n . The idea is to use the available measurements \mathbf{y}_n for estimating the state of the system \mathbf{x}_n .

To solve filtering problem (4.1)–(4.2) we have to determine the probability density of the state \mathbf{x}_n conditioned on the history of available measurements $\mathbf{y}_1, \dots, \mathbf{y}_{k_n}$. It turns out that this conditional density function is Gaussian, hence, it can be characterized by its mean and covariance matrix. However, for a non-linear model operator \mathbf{M}_n (as is the case for reservoir engineering applications) such a conditional density function can be represented by its first two moments only approximately.

The EnKF has been examined and applied in a number of studies since it was first introduced by Evensen in [22] and improved by Burgers in [10]. This filtering approach is relatively easy to implement and has affordable computational costs. The EnKF is based on a representation of the probability density of the state estimate at time n by a finite number N_{ens} of randomly generated system states $\mathbf{x}_{n,i}$, $i = 1, \dots, N_{ens}$. Equations to obtain mean \mathbf{x}_n and covariance matrix \mathbf{P}_n of the probability density of state \mathbf{x}_n at time n conditioned on the history of the measurements $\mathbf{y}_1, \dots, \mathbf{y}_{k_n}$ via the EnKF algorithm can be formulated as follows [23]:

Initialization step

$$\mathbf{x}_{0,i}^a \sim \mathcal{N}(\mathbf{x}_0, \mathbf{P}_0), \quad i = 1, \dots, N_{ens}. \quad (4.3)$$

Forward step

$$\mathbf{x}_{n,i}^p = \mathbf{M}_n \mathbf{x}_{n-1,i}^a + \mathbf{B}_n \mathbf{u}_n + \mathbf{G}_n \mathbf{w}_{n,i}, \quad i = 1, \dots, N_{ens}, \quad (4.4)$$

$$\mathbf{x}_n^p = \frac{1}{N_{ens}} \sum_{i=1}^{N_{ens}} \mathbf{x}_{n,i}^p, \quad (4.5)$$

$$\mathbf{L}_n^p = \left[\mathbf{x}_{n,1}^p - \mathbf{x}_n^p, \dots, \mathbf{x}_{n,N_{ens}}^p - \mathbf{x}_n^p \right]^T, \quad (4.6)$$

where \mathbf{L}_n^p defines an approximation of the covariance matrix \mathbf{P}_n^p with rank N_{ens} :

$$\mathbf{P}_n^p = \frac{1}{N_{ens} - 1} \mathbf{L}_n^p \mathbf{L}_n^{pT}. \quad (4.7)$$

Assimilation step

$$\mathbf{K}_n = \frac{1}{N_{ens} - 1} \mathbf{L}_n^p (\mathbf{H}_n \mathbf{L}_n^p)^T \left(\frac{1}{N_{ens} - 1} (\mathbf{H}_n \mathbf{L}_n^p) (\mathbf{H}_n \mathbf{L}_n^p)^T + \mathbf{R}_n \right)^{-1}, \quad (4.8)$$

where \mathbf{K}_n is the Kalman gain matrix determining the weights with which the measurements have to be incorporated into the model update outcome:

$$\mathbf{x}_{n,i}^a = \mathbf{x}_{n,i}^p + \mathbf{K}_n \left(\mathbf{y}_n - \mathbf{H}_n \mathbf{x}_{n,i}^p + \mathbf{v}_{n,i} \right), \quad i = 1, \dots, N_{ens}, \quad (4.9)$$

$$\mathbf{x}_n^a = \frac{1}{N_{ens}} \sum_{i=1}^{N_{ens}} \mathbf{x}_{n,i}^a. \quad (4.10)$$

Algorithm (4.3)–(4.10) can be visualized by the flowchart presented in Figure 4.1, where blocks with a shadow represent the ensemble at different stages of the procedure.

Note that (4.9) involves generating additional noise $\mathbf{v}_{n,i}$ while constructing the measurement set corresponding to the ensemble. This noise $\mathbf{v}_{n,i}$ has the same statistics as assumed for the observation errors. The perturbed measurements are necessary, because the absence of perturbation would lead to an updated ensemble that has a too low variance and causes divergence of the algorithm [10].

The forward step can be performed by making a forward run of the reservoir simulator (2.12), which is used as a black box in the EnKF analysis.

It turns out that parameter estimation via the EnKF is also possible. This can be done by constructing the augmented state vector (2.11) and performing the Kalman filter analysis on it. 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.

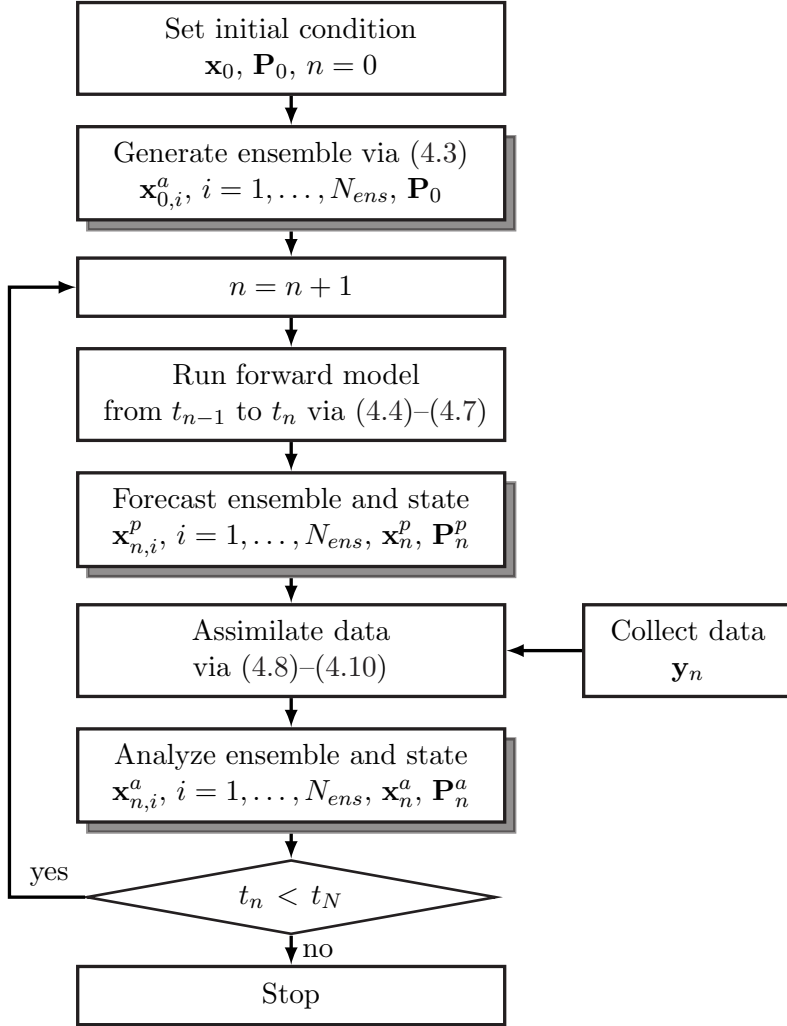


Figure 4.1: Ensemble Kalman filter algorithm.

4.2.1 The confirming ensemble Kalman filter

The use of the classical ensemble Kalman filter in reservoir engineering often leads to physically unreasonable values of the state variables. This problem originates from performing data assimilation on the extended state vector without any constraint coming from the physical nature of the parameters. Hence the updated dynamic variables may become unfeasible and inconsistent with the estimated static variables. The authors of [86] proposed to include one additional so-called confirmation step into the EnKF algorithm in order to ensure that the updated state is physically plausible and consistent with

the flow equations.

Within the confirmation step approach, we first perform a usual forecast update from time $n - 1$ up to time n and a data assimilation step. Afterwards taking only recently updated static model parameters, the flow simulator is run again from current time $n - 1$ to the next time moment n . The dynamic variables obtained replace those obtained after the measurement update and become the new initial guess for the next update step. This procedure guarantees that the updated state is consistent with the flow model. Figure 4.2 sketches integration of the confirmation step in the global workflow of the EnKF algorithm.

The inclusion of the confirmation step into the algorithm results in almost doubling the computational time due to the additional forward model run per ensemble member at each time step. In fact at the later times, when the production data carry less new and valuable information on reservoir heterogeneity, the differences between the EnKF with and without the confirmation step become smaller, and the EnKF can be switched back to the non-confirming mode to save some computation time [86].

We use the EnKF technique with confirmation step instead of classical ensemble Kalman filtering for our investigations. So from now on we imply the use of the confirmation step approach to EnKF when using the acronym 'EnKF'.

4.2.2 The asynchronous ensemble Kalman filter

The classical EnKF, particularly with confirmation step adjustment, is often computationally demanding for large-scale applications. The AEnKF technique is a modification of EnKF that offers a practical way to perform data assimilation in such cases by updating the system with batches of measurements collected at the times different to the time of the update.

The AEnKF requires only one forward run of the system to obtain and store data necessary for the analysis. Furthermore, it does not rely on an adjoint model, though it resembles the approach usually followed in variational methods.

The AEnKF method is based on the concept of propagation of corrections along the forecast trajectory. As long as the system can be considered linear within data assimilation window, the evolution of the corrections can be also treated in a linear way, which simplifies the analysis.

To perform data assimilation with the AEnKF algorithm, one needs to work with the actual measurements, the observation error covariance, and the ensemble of observations predicted by the model. The AEnKF solution can be obtained by formally replacing an ensemble of states $\mathbf{x}_{n,i}$, $i = 1, \dots, N_{ens}$ at time n by a joint ensemble $\tilde{\mathbf{x}}_{n,i}$, $i = 1, \dots, N_{ens}$, that contains variables to

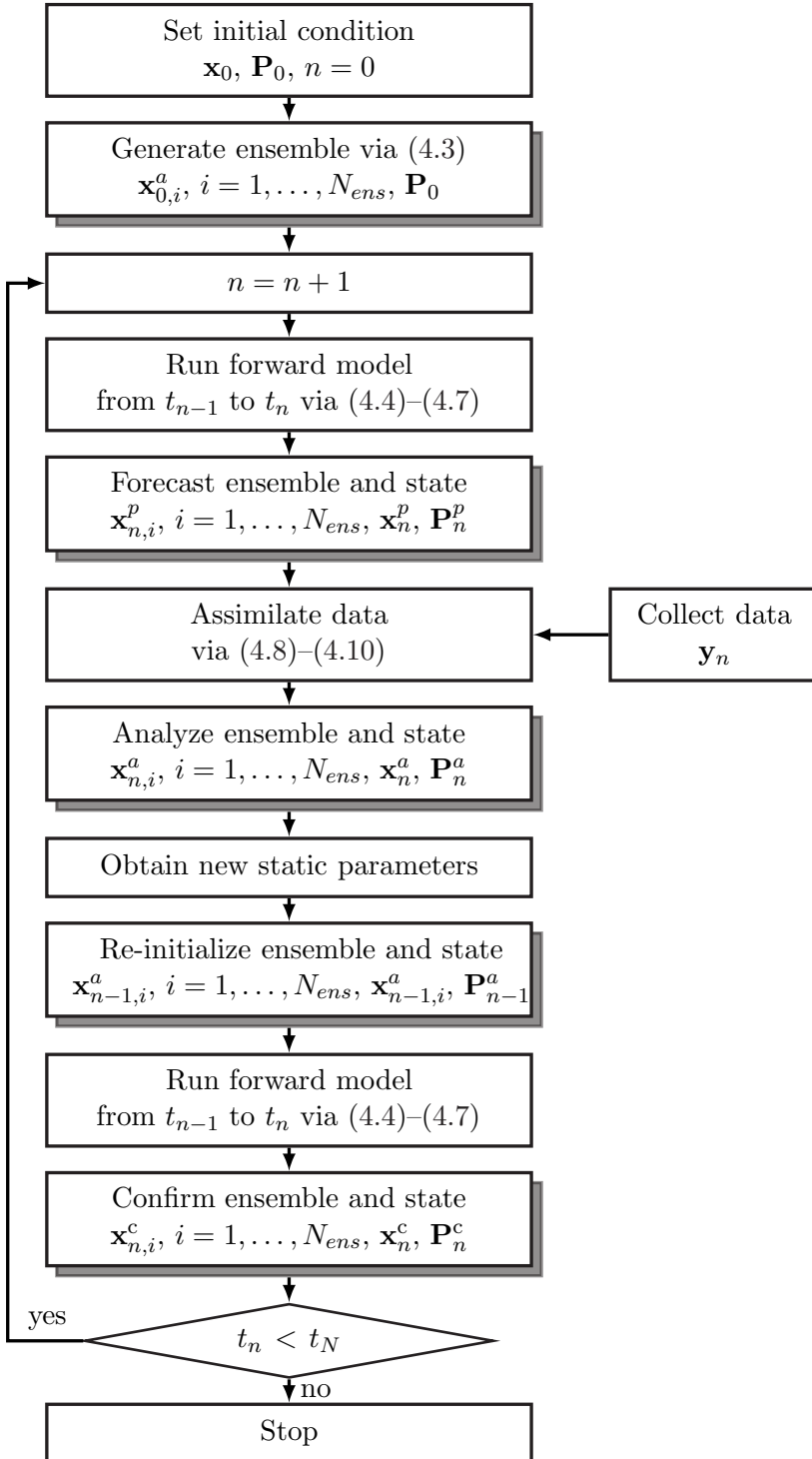


Figure 4.2: Ensemble Kalman filter with confirmation step.

be estimated (e.g. the state at the end of the data assimilation window) and predicted asynchronous data [72]. Hence, joint ensemble has the following structure:

$$\tilde{\mathbf{x}}_{n,i} = \begin{bmatrix} \mathbf{x}_{n,i} \\ \mathbf{H}_1 \mathbf{x}_{1,i} \\ \dots \\ \mathbf{H}_n \mathbf{x}_{n,i} \end{bmatrix}. \quad (4.11)$$

This joint ensemble (4.11) is thereafter updated within the regular EnKF framework. Figure 4.3 outlines the steps to be performed for data assimilation with the AEnKF algorithm.

It is important to note that although the AEnKF was originally introduced for an idealized framework that included assumptions of linearity for the model and observation operators, the algorithm can be used in practice as a suboptimal solution in situations when these assumptions are to certain extent violated [72].

4.2.3 Scaling

The use of different units and scales for the measurement variables may result in a problem to implement the filter algorithms, similar to the one discussed in Section 3.2.4. To avoid this computational problem and prevent divergence of the filter, we propose the following modification to the data assimilation step of the filtering procedure.

Consider an arbitrary time step n and rewrite the expression to compute the Kalman gain (4.8) as

$$\begin{aligned} \mathbf{K}_n &= \frac{1}{N_{ens} - 1} \mathbf{L}_n^p (\mathbf{H}_n \mathbf{L}_n^p)^T \left(\frac{1}{N_{ens} - 1} (\mathbf{H}_n \mathbf{L}_n^p) (\mathbf{H}_n \mathbf{L}_n^p)^T + \mathbf{R}_n \right)^{-1} \\ &= \frac{1}{N_{ens} - 1} \mathbf{L}_n^p (\mathbf{H}_n \mathbf{L}_n^p)^T \mathbf{R}_n^{-T/2} \\ &\quad * \mathbf{R}_n^{T/2} \left(\frac{1}{N_{ens} - 1} (\mathbf{H}_n \mathbf{L}_n^p) (\mathbf{H}_n \mathbf{L}_n^p)^T + \mathbf{R}_n \right)^{-1} \mathbf{R}_n^{1/2} \mathbf{R}_n^{-1/2} \\ &= \frac{1}{N_{ens} - 1} \mathbf{L}_n^p \left(\mathbf{R}_n^{-1/2} \mathbf{H}_n \mathbf{L}_n^p \right)^T \\ &\quad * \left(\frac{1}{N_{ens} - 1} \left(\mathbf{R}_n^{-1/2} \mathbf{H}_n \mathbf{L}_n^p \right) \left(\mathbf{R}_n^{-1/2} \mathbf{H}_n \mathbf{L}_n^p \right)^T + \mathbf{R}_n \right)^{-1} \mathbf{R}_n^{-1/2} \end{aligned}$$

or

$$\mathbf{K}_n = \mathbf{K}_n^* \mathbf{R}_n^{-1/2}, \quad (4.12)$$

$$\text{where } \mathbf{K}_n^* = \left(\frac{1}{N_{ens} - 1} \left(\mathbf{R}_n^{-1/2} \mathbf{H}_n \mathbf{L}_n^p \right) \left(\mathbf{R}_n^{-1/2} \mathbf{H}_n \mathbf{L}_n^p \right)^T + \mathbf{R}_n \right)^{-1}.$$

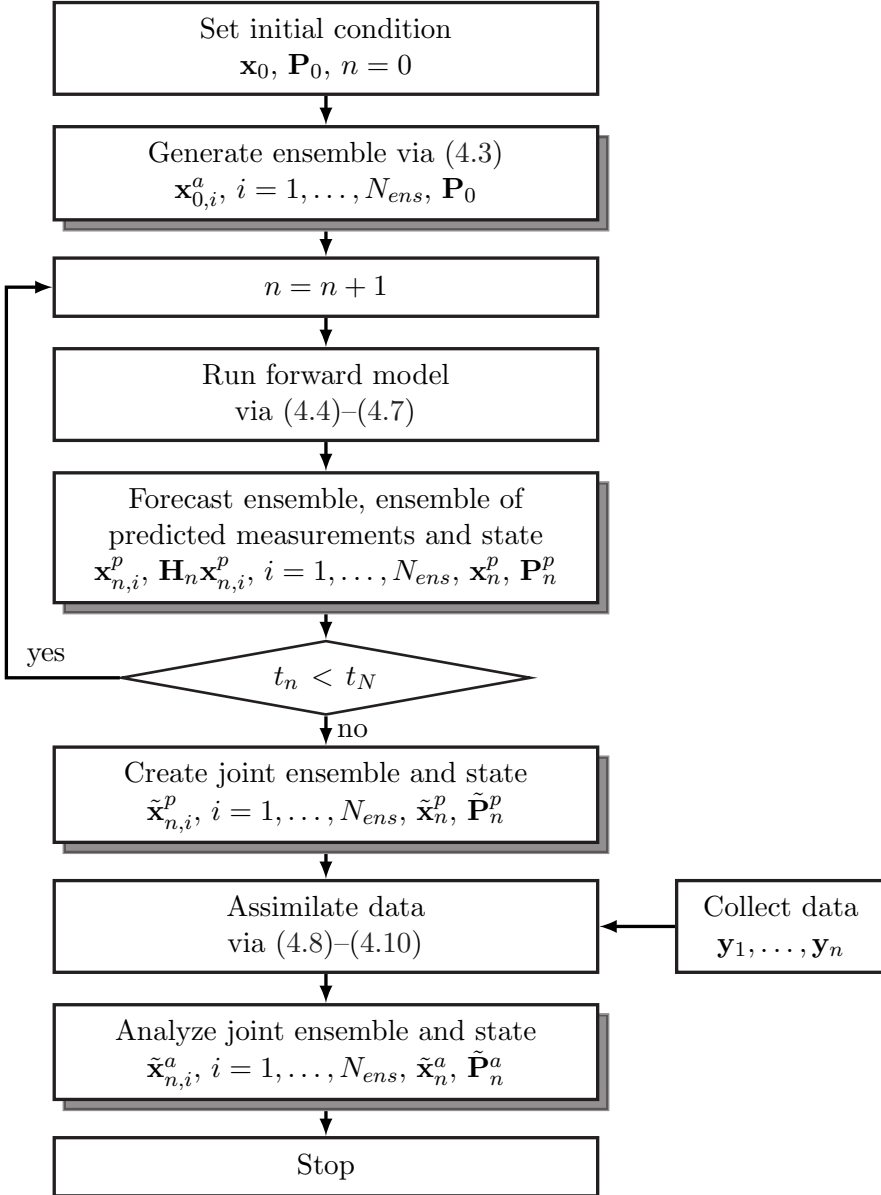


Figure 4.3: Asynchronous ensemble Kalman filter.

The Kalman gain factorization (4.12) involves matrix \mathbf{K}_k^* which is computed via operations only on the scaled matrices. The ensemble update (4.9)

is now performed as

$$\begin{aligned}
 \mathbf{x}_{n,i}^a &= \mathbf{x}_{n,i}^p + \mathbf{K}_n \left(\mathbf{y}_n - \mathbf{H}_n \mathbf{x}_{n,i}^p + \mathbf{v}_{n,i} \right) \\
 &= \mathbf{x}_{n,i}^p + \mathbf{K}_n^* \mathbf{R}_n^{-1/2} \left(\mathbf{y}_n - \mathbf{H}_n \mathbf{x}_{n,i}^p + \mathbf{v}_{n,i} \right) \\
 &= \mathbf{x}_{n,i}^p + \mathbf{K}_n^* \left(\mathbf{R}_n^{-1/2} \mathbf{y}_n - \mathbf{R}_n^{-1/2} \mathbf{H}_n \mathbf{x}_{n,i}^p + \mathbf{R}_n^{-1/2} \mathbf{v}_{n,i} \right).
 \end{aligned} \tag{4.13}$$

Expression (4.13) involves matrices and vectors scaled by the observation error covariance matrix.

4.3 Results and discussion

We tested the performance of the EnKF and AEnKF assimilation algorithms by estimating the model parameters of a two-phase two-dimensional fluid flow model. For each algorithm we solve the history matching problem and obtain estimates of permeability and porosity for the base case described in Section 2.3. The filter analysis is done from time $t_0 = 0$ [days] till $t_{\text{end}} = 450$ [days], which ensures that water breakthrough occurs in none of the production wells.

4.3.1 History matching with the EnKF

It turns out that EnKF faces an important practical problem, namely, the standard deviation of the errors in the state estimate converges very slowly with the number of ensembles. This makes the ensemble Kalman filter quite sensitive to the number of ensemble members used for simulation [24]. Preliminary analysis shows that in our case an ensemble consisting of $N_{\text{ens}} = 100$ members is sufficient for the EnKF runs. We use the same ensemble size also for the AEnKF runs.

We consider the quality of estimating the model parameters. For that purpose the rms error measure (3.18) is plotted in time both for permeability and porosity estimations (Figure 4.4). The quantities are related to the ensemble mean and ensemble members corresponding to evaluated permeability and porosity fields. The graph demonstrates an improvement of the parameter estimation in the first data assimilation step followed by stabilization of the error, and reduction of the uncertainty for estimated values (since the ensemble spread becomes narrower). This means that at the later times assimilated data carry less useful information on reservoir structure than at the early times.

We obtain permeability and porosity fields resembling the 'true' ones (compare Figures 2.1 and 4.5). The variance field is obtained through the diagonal terms of the covariance matrix computed from the statistical properties of the

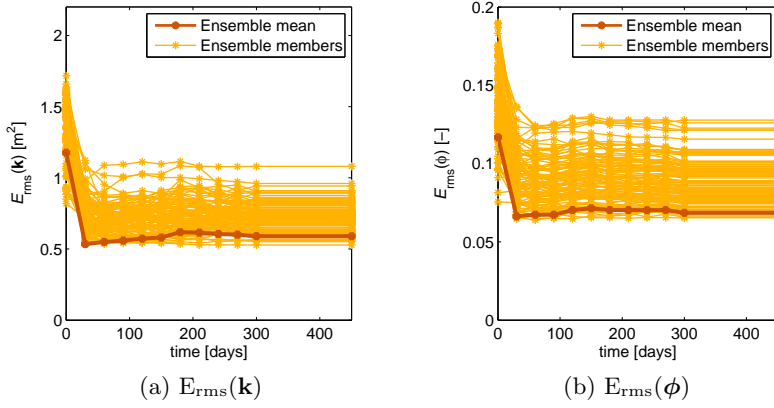


Figure 4.4: EnKF: rms error in estimated permeability and porosity vs time.

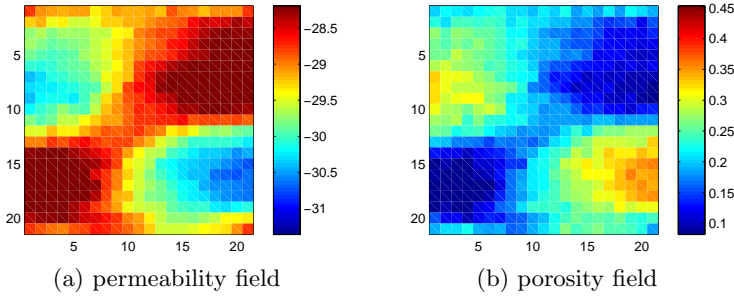


Figure 4.5: EnKF: estimated permeability and porosity fields. The scale represents transformed values of permeability [m^2] and porosity [-].

ensemble. The difference between the top and the bottom subplots in Figure 4.6 indicates a reduction of the variance and therefore of the uncertainty in the estimation of permeability and porosity.

Altogether the results presented in Figures 4.4–4.6 confirm the efficiency of the EnKF for history matching and estimation of model parameters in a small example.

4.3.2 History matching with the AEnKF

We proceed by performing history matching with the AEnKF in order to compare its performance to that of the EnKF. The joint ensemble is formed by running the model forward up to $t_{\text{end}} = 450$ [days] and collecting the model predictions on the measured variables every 30 [days] to augment the

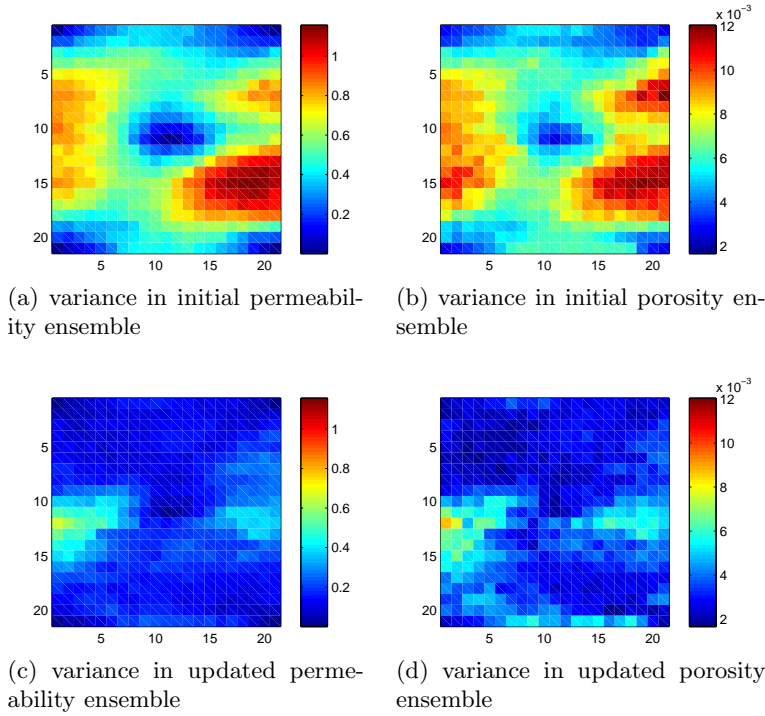


Figure 4.6: EnKF: ensemble variance. The scale represents variance for transformed values of permeability [m^2] and porosity [-].

state vector. The model update is thereafter performed only once at $t_{\text{end}} = 450$ [days].

The estimated porosity and permeability fields are presented in Figure 4.7. Visual comparison of the AEnKF results to the 'true' field and the field recovered by the EnKF confirms that the applied method leads to an acceptable estimation of the field features. The values of rms error measure support this judgment. Namely, the rms measure is 0.5744 for permeability and 0.0685 for porosity estimates, which is similar to the final rms values of 0.5897 and 0.0686 for the permeability and porosity fields as estimated by the EnKF respectively.

Alternatively, the quality of the history match can be assessed by analyzing the capability of the calibrated model to correctly forecast e.g. water breakthrough time in the production wells. Figure 4.8 illustrates that accurate prediction of water breakthrough is achieved with models obtained through data assimilation based on both the AEnKF and the EnKF techniques. These forecasts are also comparable to the predictions generated by the model that was history-matched via the RM.

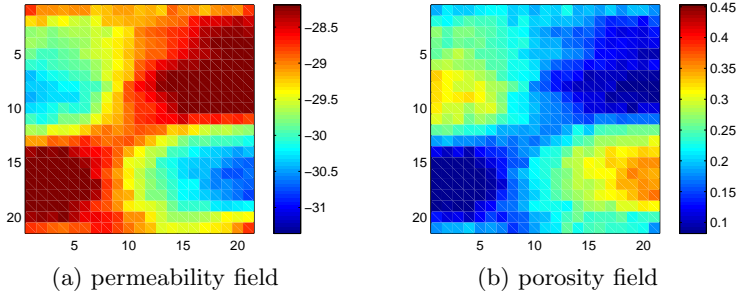


Figure 4.7: AEnKF: estimated permeability and porosity fields. The scale represents transformed values of permeability [m^2] and porosity [-].

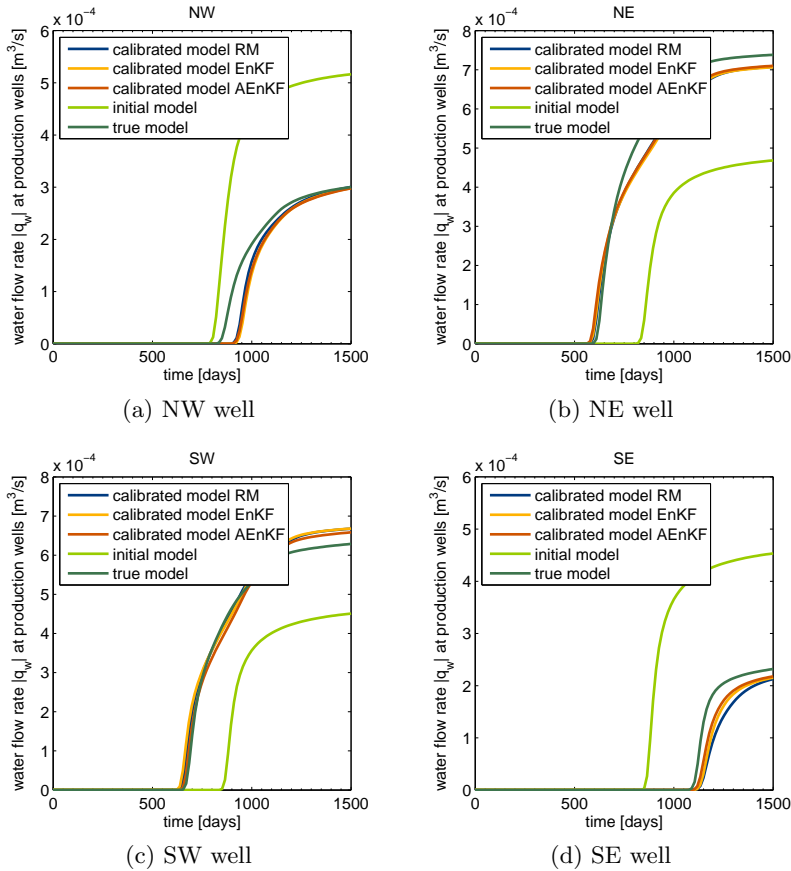


Figure 4.8: Forecast of water flow rate at production wells.

Thus the use of the AEnKF algorithm yields a reasonable estimate of the model parameters with estimated fields analogous to the ones obtained with the EnKF and the RM.

4.4 Conclusion

We have reviewed the use of ensemble Kalman filtering techniques as tools to solve the history matching problem. The EnKF method and its modifications belong to the family of recursive techniques as opposed to the variational algorithms (RM and ARM) discussed in Chapter 3.

First, the details of the EnKF algorithm have been presented alongside the considerations on occasional cost inefficiency of the EnKF for data assimilation in an operational environment. This discussion led to examination of the asynchronous modification of the EnKF technique and comparison of the performance of the two methods.

Results of history matching with the EnKF and the AEnKF are analogous and both demonstrate a considerable improvement of the model parameter estimates. In particular, the rms measure is 0.5744 for permeability and 0.0685 for porosity estimates obtained with the AEnKF, being similar to the final rms values of 0.5897 and 0.0686 for the estimates obtained by the EnKF. Also the capability of the model calibrated with the EnKF to predict the water breakthrough time is comparable to the one of the model history-matched with the AEnKF (see Figure 4.8). Since the AEnKF allows assimilation of all the data gathered throughout the observational period at once, it permits comparison of the effect of observations collected at different time instances. The equivalence of the AEnKF to variational techniques (e.g. RM) yields the possibility to evaluate if sequential and variational methods utilize the observations in a similar manner.

These features and the indifference of the AEnKF to the availability of an adjoint model make it a good candidate to be used further as a data assimilation method when quantifying the importance of the observations for history matching.

Observation sensitivity matrix

5.1 Introduction

The predecessor of the observation sensitivity matrix, the so-called influence matrix, originated in ordinary least-squares applications as a tool for monitoring statistical regression analysis (see e.g. [33]). Within regression analysis this matrix is called a hat matrix and provides a means to understand the influence that a data value will have on each fitted value and identify exceptional data points.

The concept was further modified and adapted within the inverse modeling domain under the name of 'model resolution matrix'. This matrix allows to investigate which parameters are unresolved or well-resolved by assimilation of the data. In [52] the model resolution matrix was discussed within a geophysical context; in [62] the model resolution matrix was applied to investigate the usefulness of chlorophyll data to estimate the parameters of an ocean model; [68] considered the model resolution matrix for inverse problems related to atmospheric sounding. The model resolution matrix approach has been also applied to the history matching problems in [18, 57, 58], where the possibility to resolve reservoir model parameters from the data has been studied.

The next step was taken by the authors of [65] and [11] who investigated the possibilities to choose an optimal subset of data for assimilation in models used for weather predictions via the so-called data resolution matrix and the related influence matrix (or observation sensitivity matrix) respectively. All of these matrices rely on information-theoretic concepts and determine the influence of individual observations on the quality of the 'analysis', i.e. the procedure to minimize the mismatch between measured data and model predictions of those data.

The issue of information content of observations was recently brought up in meteorological applications [65] and [11], because nowadays measurement equipment collects an enormous amount of atmospheric data on a daily basis. In meteorology the main use of observations for model updating is to improve the initial conditions of the model, a process referred to as data assimilation, which has many similarities to computer-assisted history matching in petroleum reservoir engineering. In general, increasing the amount of observations improves the results of the data assimilation procedure and the forecasting capability of the model. In that sense it is beneficial to involve all the measurements available in the assimilation. This, however, becomes too demanding for operational weather forecasting systems due to time constraints on providing the new forecast. Therefore a proper selection of the most influential data points is needed to reduce the computational effort while preserving the desired accuracy.

Contrary to the situation in meteorology, there are often too few observations available in reservoir engineering. The information extracted from these measurements in the history matching phase is frequently found to be not enough to provide a well-calibrated model with a high predictive value. This is mainly caused by the high cost of data gathering in the petroleum industry. Moreover, data collection requirements such as frequency and accuracy of production tests or in-well sensors have to be determined in advance, i.e. during the field development phase, and are difficult to adjust during the operational life of a field. Methods to assess the value of measurements for model parameter updating and subsequent model-based production forecasting are therefore needed, to allow up-front selection of the most influential measurements and their locations.

To assess the value of production measurements for history matching we apply a method originally developed for atmospheric problems. We consider assimilating production data to estimate the permeability field in a two-phase (oil-water) two-dimensional reservoir model. The most influential measurements are then found from an observation sensitivity matrix, which provides a fast and easy way to diagnose the influence of individual observations on the analysis.

5.2 Concept of observation sensitivity

The section outlines the key points of the observation sensitivity matrix diagnostics in a manner similar to [11] by considering history matching in a Bayesian framework (see e.g. [24, 59]). Then under the usual assumptions on Gaussianity of the model and measurement errors, the optimal solution to the analysis problem in a linear case can be written as a linear combination

of observed data and prior model data:

$$\mathbf{x}^a = \mathbf{x}^p + \mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}^p),$$

where \mathbf{x}^a is the analyzed (i.e. updated) vector of model state variables, \mathbf{K} denotes the Kalman gain matrix that balances the influence of the prior knowledge \mathbf{x}^p and the observations \mathbf{y} , \mathbf{H} stands for the observation operator establishing the link between the model state vector and the measurements. The weight factors in the Kalman gain depend on the covariances of the prior model errors and the measurement errors, namely $\mathbf{K} = \mathbf{P}_0\mathbf{H}^T (\mathbf{H}\mathbf{P}_0\mathbf{H}^T + \mathbf{R})^{-1}$, where \mathbf{P}_0 stands for the covariance of prior model errors and \mathbf{R} denotes covariance of measurement errors. Predicted observations $\hat{\mathbf{y}}$ of the analyzed model can then be expressed in the form of a weighted sum of predicted observations $\mathbf{H}\mathbf{x}^p$ of the prior model and the measured observations \mathbf{y} :

$$\hat{\mathbf{y}} = \mathbf{H}\mathbf{x}^a = \mathbf{H}\mathbf{x}^p + \mathbf{H}\mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}^p) = (\mathbf{I} - \mathbf{H}\mathbf{K})\mathbf{H}\mathbf{x}^p + \mathbf{H}\mathbf{K}\mathbf{y}, \quad (5.1)$$

where \mathbf{I} is an identity matrix of an appropriate size. Finally, the observation sensitivity matrix, i.e. the sensitivity matrix of the analyzed predicted observations with respect to the measured data, can be computed as

$$\mathbf{S} = \frac{\partial \hat{\mathbf{y}}}{\partial \mathbf{y}} = \mathbf{H}\mathbf{K}. \quad (5.2)$$

The observation sensitivity matrix \mathbf{S} , also referred to as the influence matrix, is a square matrix of dimension K , where K is the total amount of assimilated data. The diagonal elements S_{ii} are pure numbers and are called self-sensitivities. They represent the change in a predicted observation with respect to variations in the corresponding measured observation. If the covariance matrix of measurement errors is diagonal, the self-sensitivities can be proven to be in $[0, 1]$ bounds [11]. A zero self-sensitivity indicates that the i^{th} observation has no influence at all on the analysis, while $S_{ii} = 1$ shows that the entire information content of that observation has been devoted to updating those model parameters that influence the corresponding prediction.

The trace of sensitivity matrix $\text{Tr}(\mathbf{S})$ characterizes the amount of information extracted from the observations. Note that its maximum value is equal to the number of observations. The complementary trace $\text{Tr}(\mathbf{I} - \mathbf{S})$ describes the remainder of the information, which originates from the prior model. Starting from the observation sensitivity matrix, some further diagnostics can be performed. For instance, one can define the 'global average influence' (GAI) and 'partial averaged influence' (PAI) for any selected subset of data:

$$I_{\text{GA}} = \frac{\text{Tr}(\mathbf{S})}{K} \quad \text{and} \quad I_{\text{PA}} = \frac{\sum_{i \in \mathbb{K}} S_{ii}}{K_{\mathbb{K}}}, \quad (5.3)$$

where $K_{\mathbb{K}}$ is the amount of measurements in subset \mathbb{K} . The subset can, for example, represent a specific observation type or location. Note that measures (5.3) only quantify the impact of data on the history matching procedure, i.e. how much the outcome of the history matching would differ from the prior under the influence of measurements. Expressions (5.3) do not allow investigating if certain data have a positive or a negative influence on the quality of history-matched model.

To illustrate the concept we consider the two most extreme cases, when either a highly uncertain prior or highly uncertain observations are involved in the analysis.

Highly uncertain prior:

Recall that in such a case an algorithm exploits the measurements, since the prior model is unreliable as indicated by its high uncertainty. Indeed, the observation sensitivity matrix can be written in the form

$$\mathbf{S} = \mathbf{H}\mathbf{K} = \mathbf{H}\mathbf{P}_0\mathbf{H}^T (\mathbf{H}\mathbf{P}_0\mathbf{H}^T + \mathbf{R})^{-1}. \quad (5.4)$$

Since

$$\begin{aligned} \mathbf{S} &= \mathbf{H}\mathbf{P}_0\mathbf{H}^T (\mathbf{H}\mathbf{P}_0\mathbf{H}^T + \mathbf{R})^{-1} \\ &= \left((\mathbf{H}\mathbf{P}_0\mathbf{H}^T) (\mathbf{H}\mathbf{P}_0\mathbf{H}^T)^{-1} + \mathbf{R} (\mathbf{H}\mathbf{P}_0\mathbf{H}^T)^{-1} \right)^{-1} \\ &= \left(\mathbf{I} + \mathbf{R} (\mathbf{H}\mathbf{P}_0\mathbf{H}^T)^{-1} \right)^{-1} \\ &= \left(\mathbf{R}\mathbf{R}^{-1} + \mathbf{R} (\mathbf{H}\mathbf{P}_0\mathbf{H}^T)^{-1} \right)^{-1} \\ &= \left(\mathbf{R}^{-1} + (\mathbf{H}\mathbf{P}_0\mathbf{H}^T)^{-1} \right)^{-1} \mathbf{R}^{-1}, \end{aligned}$$

expression (5.4) is equivalent to

$$\mathbf{S} = \left(\mathbf{R}^{-1} + (\mathbf{H}\mathbf{P}_0\mathbf{H}^T)^{-1} \right)^{-1} \mathbf{R}^{-1}. \quad (5.5)$$

It is clear from (5.5) that if the problem is characterized by a highly uncertain prior, the observation sensitivity matrix degrades to an identity matrix of an appropriate size, i.e. $\mathbf{S} = \mathbf{I}$. Therefore $\text{Tr}(\mathbf{S}) = K$ and the GAI is equal to 1 yielding the conclusion that the entire information content of every observation has been used to update the model parameters.

Highly uncertain observations:

Highly uncertain observations force the model update to rely mostly on the prior, since the accuracy of the available measurements is poor. It

also follows from (5.4) that if the observations are very uncertain, the observation sensitivity matrix becomes a zero matrix of appropriate size, i.e. $\mathbf{S} = \mathbf{0}$. Hence, $\text{Tr}(\mathbf{S}) = 0$ and the GAI is equal to 0 indicating that the measurements have no influence at all on the update of the model.

The above discussion demonstrates that the trace of the sensitivity matrix can also be used as a tool to compare the quality of observations and prior.

In the following sections we investigate if other properties of the observation sensitivity matrix can provide additional insights into the influence of the measurements on the analysis.

5.3 Structure of the observation sensitivity matrix

The section addresses the structure and the features of the observation sensitivity matrix. We are particularly interested if (or under which conditions) the observation sensitivity matrix is symmetric positive semidefinite and all of its entries are dimensionless (i.e. pure real numbers). The properties to be derived will be used for the analysis presented in the remainder of this chapter.

5.3.1 Symmetry of the observation sensitivity matrix

To investigate the issue of symmetry of the square observation sensitivity matrix \mathbf{S} one needs to determine if and under which conditions

$$\mathbf{S} = \mathbf{S}^T$$

or equivalently

$$\mathbf{H}\mathbf{P}_0\mathbf{H}^T (\mathbf{H}\mathbf{P}_0\mathbf{H}^T + \mathbf{R})^{-1} = (\mathbf{H}\mathbf{P}_0\mathbf{H}^T + \mathbf{R})^{-1} \mathbf{H}\mathbf{P}_0\mathbf{H}^T. \quad (5.6)$$

Elaborating on (5.6) we obtain

$$(\mathbf{H}\mathbf{P}_0\mathbf{H}^T + \mathbf{R}) \mathbf{H}\mathbf{P}_0\mathbf{H}^T = \mathbf{H}\mathbf{P}_0\mathbf{H}^T (\mathbf{H}\mathbf{P}_0\mathbf{H}^T + \mathbf{R})$$

and finally

$$\mathbf{R}\mathbf{H}\mathbf{P}_0\mathbf{H}^T = \mathbf{H}\mathbf{P}_0\mathbf{H}^T\mathbf{R}. \quad (5.7)$$

It is now obvious that in a general case the observation sensitivity matrix \mathbf{S} is not symmetric unless condition (5.7) holds.

A particular example when symmetry is achieved can be given by a history matching case with independent measurements of the same type and accuracy. In this case the covariance matrix of measurement errors \mathbf{R} is given by $r^2\mathbf{I}$, where \mathbf{I} is identity matrix of appropriate size, and r is a scalar representing the

standard deviation of the measurement errors. Moreover, such data features ensure that both the diagonal and the off-diagonal entries of matrix \mathbf{S} are pure numbers.

Since the discussed case is very limiting, we proceed by introducing a version of matrix scaling that creates a dimensionless analogue of the observation sensitivity matrix with somewhat more attractive properties.

5.3.2 Scaled observation sensitivity matrix

To create a dimensionless version of predicted observations $\hat{\mathbf{y}}$ of the analyzed model we consider an invertible diagonal scaling matrix \mathbf{R}_s such that

$$\mathbf{R}_s^{-1/2}\hat{\mathbf{y}} = \mathbf{R}_s^{-1/2}\mathbf{H}\mathbf{x}^p + \mathbf{R}_s^{-1/2}\mathbf{H}\mathbf{K}\mathbf{R}_s^{1/2}(\mathbf{R}_s^{-1/2}\mathbf{y} - \mathbf{R}_s^{-1/2}\mathbf{H}\mathbf{x}^p). \quad (5.8)$$

and all the quantities constituting to (5.8) become dimensionless. There exist various choices of a scaling matrix that would lead to the desired outcome, however, the most natural choice is to prescribe \mathbf{R}_s to be the same as the covariance matrix of measurement errors \mathbf{R} , i.e. we consider $\mathbf{R}_s = \mathbf{R}$. Note that \mathbf{R}_s is then the same as the scaling matrix that appears throughout minimization of the generalized cost function in a history matching procedure to account for different orders of magnitude of the entries in the measurement vector.

We define the scaled version of the observation sensitivity matrix as

$$\tilde{\mathbf{S}} = \mathbf{R}_s^{-1/2}\mathbf{H}\mathbf{K}\mathbf{R}_s^{1/2} = \mathbf{R}_s^{-1/2}\mathbf{S}\mathbf{R}_s^{1/2}. \quad (5.9)$$

Elements of this scaled observation sensitivity matrix $\tilde{\mathbf{S}}$ are dimensionless. Importantly, the results of the trace diagnostic performed with matrix $\tilde{\mathbf{S}}$ are exactly the same as those obtained with original matrix \mathbf{S} due to the properties of matrix trace [47]

$$\text{Tr}(\tilde{\mathbf{S}}) = \text{Tr}(\mathbf{R}_s^{-1/2}\mathbf{S}\mathbf{R}_s^{1/2}) = \text{Tr}(\mathbf{S}\mathbf{R}_s^{1/2}\mathbf{R}_s^{-1/2}) = \text{Tr}(\mathbf{S}).$$

Hence, either matrix can be used to perform the original diagnostics of [11]. Moreover, note that formulation (5.9) exhibits a certain intuitively expected similarity to the dimensionless sensitivity matrix introduced in [92] to quantify the sensitivity of the predicted data to the model parameters.

The issue of symmetry of the scaled observation sensitivity matrix, however, requires further investigation. Similarly to Section 5.3.1 one needs to check conditions under which

$$\tilde{\mathbf{S}} = \tilde{\mathbf{S}}^T$$

or equivalently

$$\begin{aligned} \mathbf{R}_s^{-1/2}\mathbf{H}\mathbf{P}_0\mathbf{H}^T (\mathbf{H}\mathbf{P}_0\mathbf{H}^T + \mathbf{R})^{-1} \mathbf{R}_s^{1/2} \\ = \mathbf{R}_s^{1/2} (\mathbf{H}\mathbf{P}_0\mathbf{H}^T + \mathbf{R})^{-1} \mathbf{H}\mathbf{P}_0\mathbf{H}^T \mathbf{R}_s^{-1/2}. \end{aligned} \quad (5.10)$$

After several linear algebraic transformations on (5.10) we obtain

$$\mathbf{R}\mathbf{R}_s^{-1}\mathbf{H}\mathbf{P}_0\mathbf{H}^T = \mathbf{H}\mathbf{P}_0\mathbf{H}^T\mathbf{R}_s^{-1}\mathbf{R}. \quad (5.11)$$

Generally speaking, the scaled observation sensitivity matrix $\tilde{\mathbf{S}}$ is not necessarily symmetric. However, the matrix can be forced to become symmetric by choosing an invertible diagonal scaling matrix such that $\mathbf{R}_s = \mathbf{R}$.

5.3.3 Positive (semi)definiteness of the scaled observation sensitivity matrix

A symmetric matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is considered to be positive definite (PD) if $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$ for all nonzero $\mathbf{x} \in \mathbb{R}^n$ and positive semidefinite (PSD) if $\mathbf{x}^T \mathbf{A} \mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$ [47]. The following criteria can be used to determine if the matrix is P(S)D:

- $\mathbf{A} \in \text{PD}$ if the eigenvalues of \mathbf{A} are all positive,
- $\mathbf{A} \in \text{PSD}$ if the eigenvalues of \mathbf{A} are all nonnegative.

The covariance matrices are PSD due to their physical nature. Hence, in our case, matrices \mathbf{P}_0 and \mathbf{R} are PSD.

Since the scaled observation sensitivity matrix is symmetric when $\mathbf{R}_s = \mathbf{R}$, the notion of positive (semi)definiteness is applicable and can be further investigated. With the aid of (5.4) and (5.9) the scaled observation sensitivity matrix can be written as

$$\begin{aligned} \tilde{\mathbf{S}} &= \mathbf{R}^{-1/2} \mathbf{H} \mathbf{K} \mathbf{R}^{1/2} \\ &= \mathbf{R}^{-1/2} \mathbf{H} \mathbf{P}_0 \mathbf{H}^T (\mathbf{H} \mathbf{P}_0 \mathbf{H}^T + \mathbf{R})^{-1} \mathbf{R}^{1/2} \\ &= \left(\mathbf{R}^{-1/2} \mathbf{H} \mathbf{P}_0 \mathbf{H}^T \mathbf{R}^{-1/2} \right) \left(\mathbf{R}^{-1/2} \mathbf{H} \mathbf{P}_0 \mathbf{H}^T \mathbf{R}^{-1/2} + \mathbf{I} \right)^{-1}. \end{aligned} \quad (5.12)$$

To analyze (5.12) we use the following properties of positive (semi)definite matrices [34], where $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$, $\mathbf{C} \in \mathbb{R}^{m \times n}$:

- a PSD matrix is PD if and only if it is invertible,
- if $\mathbf{A} \in \text{PD}$, then $\mathbf{A}^{-1} \in \text{PD}$,
- if $\mathbf{A} \in \text{PSD}$ and k is any positive integer, then $\mathbf{A}^{1/k} \in \text{PSD}$,
- if $\mathbf{A} \in \text{PSD}$, then $\mathbf{C} \mathbf{A} \mathbf{C}^T \in \text{PSD}$,
- if $\mathbf{A} \in \text{PSD}$ and $\mathbf{B} \in \text{PD}$, then $\mathbf{A} + \mathbf{B} \in \text{PD}$,

and

- if \mathbf{A} , $\mathbf{B} \in \text{PSD}$ and $\mathbf{AB} = \mathbf{BA}$, then $\mathbf{AB} \in \text{PSD}$ [69].

Consider the terms constituting to (5.12) in detail:

- we consider \mathbf{R} to be invertible; then due to the first property of P(S)D matrices it is also PD,
- $\mathbf{R}^{-1/2}$ is PSD by the 2nd and the 3rd property,
- $\mathbf{R}^{-1/2}\mathbf{HP}_0\mathbf{H}^T\mathbf{R}^{-1/2}$ is PSD with the aid of the 4th property,
- \mathbf{I} is PD, since all of its eigenvalues are equal to 1 and, hence, are positive. Therefore $\mathbf{R}^{-1/2}\mathbf{HP}_0\mathbf{H}^T\mathbf{R}^{-1/2} + \mathbf{I}$ is PD by the 5th property and $(\mathbf{R}^{-1/2}\mathbf{HP}_0\mathbf{H}^T\mathbf{R}^{-1/2} + \mathbf{I})^{-1}$ is PD as well by the 2nd property,
- note that both $\mathbf{R}^{-1/2}\mathbf{HP}_0\mathbf{H}^T\mathbf{R}^{-1/2}$ and $\mathbf{R}^{-1/2}\mathbf{HP}_0\mathbf{H}^T\mathbf{R}^{-1/2} + \mathbf{I}$ are symmetric. Then

$$\begin{aligned} & \left(\mathbf{R}^{-1/2}\mathbf{HP}_0\mathbf{H}^T\mathbf{R}^{-1/2} \right) \left(\mathbf{R}^{-1/2}\mathbf{HP}_0\mathbf{H}^T\mathbf{R}^{-1/2} + \mathbf{I} \right)^{-1} \\ & \stackrel{(\tilde{\mathbf{S}}=\tilde{\mathbf{S}}^T)}{=} \left(\mathbf{R}^{-1/2}\mathbf{HP}_0\mathbf{H}^T\mathbf{R}^{-1/2} + \mathbf{I} \right)^{-T} \left(\mathbf{R}^{-1/2}\mathbf{HP}_0\mathbf{H}^T\mathbf{R}^{-1/2} \right)^T \\ & = \left(\mathbf{R}^{-1/2}\mathbf{HP}_0\mathbf{H}^T\mathbf{R}^{-1/2} + \mathbf{I} \right)^{-1} \left(\mathbf{R}^{-1/2}\mathbf{HP}_0\mathbf{H}^T\mathbf{R}^{-1/2} \right) \end{aligned}$$

and conditions of the 6th property hold leading to the final conclusion that scaled observation sensitivity matrix is PSD.

Thus, the developed scaling technique allows consideration of a dimensionless (and for many cases symmetric positive definite) analogue of the observation sensitivity matrix that can be analyzed by means of trace-based quantities (5.3), matrix norms (Sections 5.4 and 5.6.1) and singular values (Section 5.6.2). From now on we deal only with the scaled versions of the observation sensitivity matrix and the (predicted) observations and we refer to them simply as the 'observation sensitivity matrix' and the '(predicted) observations'.

5.4 Matrix norm of the observation sensitivity matrix

The observation sensitivity matrix \mathbf{S} is defined by (5.2) to describe how predicted analyzed observations would get changed due to changes introduced into actually measured data. The expression (5.2) can be reformulated as

$$\mathbf{S} \cdot \Delta \mathbf{y} = \Delta \hat{\mathbf{y}}, \quad (5.13)$$

where $\Delta \mathbf{y}$ and $\Delta \hat{\mathbf{y}}$ denote the changes occurring in actual data and predicted measurements respectively. The norm of the observation sensitivity matrix can be therefore used to investigate if a change in the quality of the measurements will trigger a sufficient subsequent change in the analyzed model. This knowledge is useful for taking decisions on e.g. the need for collecting more accurate data. The derivations in this section are inspired by considering the condition number of a matrix (Appendix B.1).

Suppose we are in possession of two sets of measured data, \mathbf{y}_1 and \mathbf{y}_2 , obtained with different accuracy and, hence, described by measurement errors $\boldsymbol{\epsilon}_1$ and $\boldsymbol{\epsilon}_2$. The observation model links measured erroneous data to 'perfect' data through

$$\mathbf{y}_i = \mathbf{y} + \boldsymbol{\epsilon}_i, \quad i = 1, 2,$$

where \mathbf{y} stands for 'perfect' (i.e. noise free) measurements, $\boldsymbol{\epsilon}_i$ represents a Gaussian process with zero mean and a diagonal covariance matrix $\mathbf{R}_i \in \mathbb{R}^{K \times K}$. The covariance matrix \mathbf{R}_i can be written as

$$\mathbf{R}_i = (0.01)^2 \cdot \text{diag} \left((\alpha_{i1}|y_1|)^2, \dots, (\alpha_{ij}|y_j|)^2, \dots, (\alpha_{iK}|y_K|)^2 \right),$$

$$i = 1, 2, \quad j = 1, \dots, K,$$

where $|y_j|$, $j = 1, \dots, K$, is an absolute value of the coordinate j of vector \mathbf{y} , and α_{ij} is a scalar greater or equal than 0. This strategy implies that the error in each observed variable y_{ij} , $i = 1, 2$, $j = 1, \dots, K$, is taken to be $\alpha_{ij}\%$ of the actual value y_j .

If the simulator is initialized with the 'true' reservoir parameters and is not subject to simulation errors, the outcome of the forward run complies with the 'perfect' measurements \mathbf{y} , and history matching has no effect: $\hat{\mathbf{y}} = \mathbf{y}$.

History matching with the first and second sets will result in predicted analyzed observations $\hat{\mathbf{y}}_1$ and $\hat{\mathbf{y}}_2$, where $\hat{\mathbf{y}}_i = \hat{\mathbf{y}} + \hat{\boldsymbol{\epsilon}}_i$, $i = 1, 2$, and $\hat{\boldsymbol{\epsilon}}_i$ represents a difference between the analyzed model and the 'true' model. The difference in analyzed predicted data $\Delta \hat{\mathbf{y}} = \hat{\mathbf{y}}_2 - \hat{\mathbf{y}}_1 = \hat{\boldsymbol{\epsilon}}_2 - \hat{\boldsymbol{\epsilon}}_1$ is connected to the difference in measured data $\Delta \mathbf{y} = \mathbf{y}_2 - \mathbf{y}_1 = \boldsymbol{\epsilon}_2 - \boldsymbol{\epsilon}_1$ by the sensitivity matrix \mathbf{S} . In fact, $\Delta \mathbf{y}$ can be interpreted as the change in measurement error of the observed data due to e.g. use of more accurate measurement equipment. Then $\Delta \hat{\mathbf{y}}$ corresponds to the subsequent change in the predicted analyzed observations and, hence, the analyzed model.

The relative change in the error in \mathbf{y} is given by $\frac{\|\Delta \mathbf{y}\|_2}{\|\mathbf{y}\|_2}$. Similarly, the relative change in the error in predicted analyzed observations is $\frac{\|\Delta \hat{\mathbf{y}}\|_2}{\|\hat{\mathbf{y}}\|_2}$. Then with the aid of (5.13)

$$\frac{\|\Delta \hat{\mathbf{y}}\|_2}{\|\hat{\mathbf{y}}\|_2} \leq \frac{\|\mathbf{S}\|_2 \cdot \|\Delta \mathbf{y}\|_2}{\|\hat{\mathbf{y}}\|_2} = \|\mathbf{S}\|_2 \cdot \frac{\|\Delta \mathbf{y}\|_2}{\|\mathbf{y}\|_2}. \quad (5.14)$$

Hence, the norm of the observation sensitivity matrix \mathbf{S} provides a bound on how large the relative change of the error in $\hat{\mathbf{y}}$ can be. For details on matrix and vector norms the reader is referred to Appendix B.2 and [27, 34, 69].

The independent measurement errors ϵ_{ij} , $i = 1, 2$, $j = 1, \dots, K$, have a Gaussian distribution with mean 0 and standard deviation $0.01 \cdot \alpha_{ij} |y_j|$. Then the difference $\epsilon_{2j} - \epsilon_{1j}$, $j = 1, \dots, K$, has a Gaussian distribution with mean 0 and standard deviation $0.01 \cdot \sqrt{\alpha_{1j}^2 + \alpha_{2j}^2} |y_j|$, for which p of the drawn values lie within δ standard deviations from the mean.

Suppose $\alpha_{ij} = \alpha_i$, for all $j = 1, \dots, K$. Then with probability p the change in the measurement error $\epsilon_{2j} - \epsilon_{1j}$ falls within bounds $\left[-\delta \cdot 0.01 \cdot \sqrt{\alpha_1^2 + \alpha_2^2} |y_j|; \delta \cdot 0.01 \cdot \sqrt{\alpha_1^2 + \alpha_2^2} |y_j| \right]$, $j = 1, \dots, K$. Therefore, the maximum change one may expect to achieve in the model by e.g. improving the accuracy of observations is given by

$$\begin{aligned} \left| \frac{\|\hat{\epsilon}_2\|_2}{\|\hat{\mathbf{y}}\|_2} - \frac{\|\hat{\epsilon}_1\|_2}{\|\hat{\mathbf{y}}\|_2} \right| &\leq \frac{\|\hat{\epsilon}_2 - \hat{\epsilon}_1\|_2}{\|\hat{\mathbf{y}}\|_2} = \frac{\|\Delta \hat{\mathbf{y}}\|_2}{\|\hat{\mathbf{y}}\|_2} \\ &\stackrel{(5.14)}{\leq} \|\mathbf{S}\|_2 \cdot \frac{\|\Delta \mathbf{y}\|_2}{\|\mathbf{y}\|_2} = \|\mathbf{S}\|_2 \cdot \frac{\|\epsilon_2 - \epsilon_1\|_2}{\|\mathbf{y}\|_2} \\ &\leq 0.01 \cdot \delta \cdot \|\mathbf{S}\|_2 \cdot \sqrt{\alpha_1^2 + \alpha_2^2}. \end{aligned} \quad (5.15)$$

Thus, the upper bound in relation (5.15) is determined by the norm of observation sensitivity matrix and the accuracy of the measured data. Depending on the observation sensitivity matrix, introduction of more accurate measurements might not always lead to sufficient changes in the analyzed model.

The underlying assumption for the derivation of (5.15) exploits probabilistic properties of the measurement error quantity $\epsilon_2 - \epsilon_1$. Hence the final inequality (5.15) holds with a certain probability, and there is a chance that introduction of more accurate measurements will lead to greater changes in the analyzed model than expected. The probability with which (5.15) holds can be evaluated as follows:

$$\begin{aligned} &\mathbf{P} \left(\left| \frac{\|\hat{\epsilon}_2\|_2}{\|\hat{\mathbf{y}}\|_2} - \frac{\|\hat{\epsilon}_1\|_2}{\|\hat{\mathbf{y}}\|_2} \right| \leq 0.01 \cdot \delta \cdot \|\mathbf{S}\|_2 \cdot \sqrt{\alpha_1^2 + \alpha_2^2} \right) \\ &= \mathbf{P} \left(\|\mathbf{S}\|_2 \cdot \frac{\|\epsilon_2 - \epsilon_1\|_2}{\|\mathbf{y}\|_2} \leq 0.01 \cdot \delta \cdot \|\mathbf{S}\|_2 \cdot \sqrt{\alpha_1^2 + \alpha_2^2} \right) \\ &= \mathbf{P} \left(\|\epsilon_2 - \epsilon_1\|_2^2 \leq 0.01^2 \cdot \delta^2 \cdot (\alpha_1^2 + \alpha_2^2) \|\mathbf{y}\|_2^2 \right) \\ &= \mathbf{P} \left(\sum_{j=1}^K |\epsilon_{2j} - \epsilon_{1j}|^2 \leq 0.01^2 \cdot \delta^2 \cdot (\alpha_1^2 + \alpha_2^2) \sum_{j=1}^K |y_j|^2 \right), \end{aligned}$$

hence,

$$\begin{aligned}
& \mathbb{P} \left(\left| \frac{\|\hat{\epsilon}_2\|_2}{\|\hat{\mathbf{y}}\|_2} - \frac{\|\hat{\epsilon}_1\|_2}{\|\hat{\mathbf{y}}\|_2} \right| \leq 0.01 \cdot \delta \cdot \|\mathbf{S}\|_2 \cdot \sqrt{\alpha_1^2 + \alpha_2^2} \right) \\
& \geq \prod_{j=1}^K \mathbb{P} (|\epsilon_{2j} - \epsilon_{1j}|^2 \leq 0.01^2 \cdot \delta^2 \cdot (\alpha_1^2 + \alpha_2^2) |\mathbf{y}_j|^2) \\
& = \prod_{j=1}^K \mathbb{P} \left(|\epsilon_{2j} - \epsilon_{1j}| \leq 0.01 \cdot \delta \cdot \sqrt{\alpha_1^2 + \alpha_2^2} |\mathbf{y}_j| \right) \quad (5.16) \\
& = \prod_{j=1}^K p \\
& = p^K.
\end{aligned}$$

Thus one can manipulate the value of probability p involved into (5.16) to obtain a sharper bound in (5.15) and vice versa. When the δ and p are chosen and fixed, the inequality (5.15) holds with a probability greater than p^K , where the probability clearly depends on the amount of measurements used for history matching. If e.g. $\delta = 4$, then $p = 0.9999$ and the probability (5.16) remains to be high even for large data sets. This allows the inequality (5.15) to be used as part of a decision making procedure regarding the benefit of collecting more accurate measurements vs. the associated costs and efforts.

5.5 Uncertainty in the analyzed model

The ultimate goal of this section is to assess how the change in the quality of the observations used for history matching influences the quality of the resulting analyzed model. The quality of the analyzed model is studied based on the uncertainty in the analyzed predicted observations.

Recall that history matching is performed with the aid of measurement vector \mathbf{y} , where the covariance matrix of the corresponding measurement error is $\mathbf{R} \in \mathbb{R}^{K \times K}$. The analyzed model is then normally distributed with mean vector \mathbf{x}^a and covariance matrix $\mathbf{P} = (\mathbf{I} - \mathbf{KH})\mathbf{P}_0$. Therefore the predicted analyzed observations obtained from this model are normally distributed with mean vector $\mathbf{H}\mathbf{x}^a$ and covariance matrix $\mathbf{P}^H = \mathbf{H}(\mathbf{I} - \mathbf{KH})\mathbf{P}_0\mathbf{H}^T$. Reformu-

lating in terms of observation sensitivity matrix we obtain

$$\begin{aligned}
\mathbf{P}^H &= \mathbf{H}(\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}_0\mathbf{H}^T \\
&= \mathbf{H} \left(\mathbf{I} - \mathbf{P}_0\mathbf{H}^T (\mathbf{H}\mathbf{P}_0\mathbf{H}^T + \mathbf{R})^{-1} \mathbf{H} \right) \mathbf{P}_0\mathbf{H}^T \\
&= \mathbf{H}\mathbf{P}_0\mathbf{H}^T - \mathbf{H}\mathbf{P}_0\mathbf{H}^T (\mathbf{H}\mathbf{P}_0\mathbf{H}^T + \mathbf{R})^{-1} \mathbf{H}\mathbf{P}_0\mathbf{H}^T \\
&= \mathbf{H}\mathbf{P}_0\mathbf{H}^T (\mathbf{H}\mathbf{P}_0\mathbf{H}^T + \mathbf{R})^{-1} (\mathbf{H}\mathbf{P}_0\mathbf{H}^T + \mathbf{R} - \mathbf{H}\mathbf{P}_0\mathbf{H}^T) \\
&\stackrel{(5.4)}{=} \mathbf{S}\mathbf{R}.
\end{aligned} \tag{5.17}$$

Note that covariance matrices \mathbf{P}^H and \mathbf{R} are symmetric, hence, $\mathbf{S}\mathbf{R} = \mathbf{R}\mathbf{S}^T$.

Suppose we are in possession of two sets of measured data \mathbf{y}_1 and \mathbf{y}_2 obtained with different accuracy. The covariance matrices of the corresponding measurement errors are $\mathbf{R}_1 \in \mathbb{R}^{K \times K}$ and $\mathbf{R}_2 \in \mathbb{R}^{K \times K}$. History matching with the first and second sets will result in two analyzed models. These models yield predicted analyzed observations that are normally distributed with mean vectors $\mathbf{H}\mathbf{x}_1^a$ and $\mathbf{H}\mathbf{x}_2^a$ and covariance matrices $\mathbf{P}_1^H = \mathbf{S}_1\mathbf{R}_1$ and $\mathbf{P}_2^H = \mathbf{S}_2\mathbf{R}_2$. After some linear algebraic transformations we obtain the following relation between covariance matrices \mathbf{P}_2^H and \mathbf{P}_1^H :

$$\begin{aligned}
\mathbf{P}_2^H &= \mathbf{S}_2\mathbf{R}_2 \\
&= \mathbf{S}_1\mathbf{R}_1 + \mathbf{S}_2\mathbf{R}_2 - \mathbf{S}_1\mathbf{R}_1 \\
&= \mathbf{P}_1^H + \mathbf{R}_2\mathbf{S}_2^T - \mathbf{S}_1\mathbf{R}_1 \\
&= \mathbf{P}_1^H + \mathbf{S}_1 \left(\mathbf{S}_1^{-1}\mathbf{R}_2 - \mathbf{R}_1\mathbf{S}_2^{-T} \right) \mathbf{S}_2^T \\
&\stackrel{(5.4)}{=} \mathbf{P}_1^H \\
&\quad + \mathbf{S}_1 \left((\mathbf{P}_0^H + \mathbf{R}_1) (\mathbf{P}_0^H)^{-1} \mathbf{R}_2 - \mathbf{R}_1 (\mathbf{P}_0^H)^{-1} (\mathbf{P}_0^H + \mathbf{R}_2) \right) \mathbf{S}_2^T \\
&= \mathbf{P}_1^H + \mathbf{S}_1 (\mathbf{R}_2 - \mathbf{R}_1) \mathbf{S}_2^T,
\end{aligned} \tag{5.18}$$

where $\mathbf{P}_0^H = \mathbf{H}\mathbf{P}_0\mathbf{H}^T$.

Note that for a 2-norm (Appendix B.2) $\|\mathbf{S}_2^T\|_2 = \|\mathbf{S}_2\|_2$. Then the relative change in the uncertainty in the analyzed model due to a change in the accuracy of the observations is bounded by

$$\begin{aligned}
\frac{\|\mathbf{P}_2^H - \mathbf{P}_1^H\|_2}{\|\mathbf{P}_1^H\|_2} &\stackrel{(5.18)}{=} \frac{\|\mathbf{S}_1 (\mathbf{R}_2 - \mathbf{R}_1) \mathbf{S}_2^T\|_2}{\|\mathbf{P}_1^H\|_2} \\
&\leq \frac{\|\mathbf{S}_1\|_2 \cdot \|\mathbf{R}_2 - \mathbf{R}_1\|_2 \cdot \|\mathbf{S}_2\|_2}{\|\mathbf{P}_1^H\|_2}.
\end{aligned} \tag{5.19}$$

Expression (5.19) demonstrates that the upper bound on relative change in the uncertainty is proportional to the norms of observation sensitivity matrices

corresponding to the history matching systems with two different observation sets and the norm of the difference between the covariance matrices of the measurement errors. Moreover, (5.19) can be executed prior to a history matching experiment with a second data set allowing to assess its potential contribution to the model update.

5.6 Application of the observation sensitivity matrix to history matching with the RM

A typical strategy to investigate sensitivity of a history matching algorithm to the measurements by means of the observation sensitivity matrix is demonstrated on the basis of the exercise described in Section 2.3. History matching is performed from 0 [days] till 450 [days] with the data available every 30 [days] from 30 [days] till 300 [days], using the representer algorithm. The choice of the time interval ensures that water breakthrough has not yet occurred.

5.6.1 Observation sensitivity matrix

A first insight into the sensitivity of the history match to observations is provided by the entries of the observation sensitivity matrix. In case of history matching with the representer method, the observation sensitivity assessment needs to be performed at each outer iteration separately, because the problem is linear only within an outer iteration. The best possible linear approximation to the history matching problem is however obtained at the last iteration of the representer method, since then the model is linearized in the close neighborhood of a (local) minimum. Therefore, we focus on the observation sensitivity matrix computed at the last iteration of the representer algorithm.

It is instructive to plot the elements of the observation sensitivity matrix row-wise as a 2D field (see Figure 5.1, where Figure 5.1b displays a detailed version of the top left block of Figure 5.1a).

The following remarks can be made based on careful visual examination of the observation sensitivity matrix:

- the diagonal entries often are of high magnitude. Indeed, a particular predicted measurement is naturally the most sensitive to the actual data of the same type collected at the same time moment at exactly the same location;
- consider e.g. high magnitude values of observation sensitivity coefficients S_{2j} for $j = 9k + 2$, $k = 0, 1, \dots, 9$. It is now easy to notice that a particular predicted measurement is highly influenced by the actual data

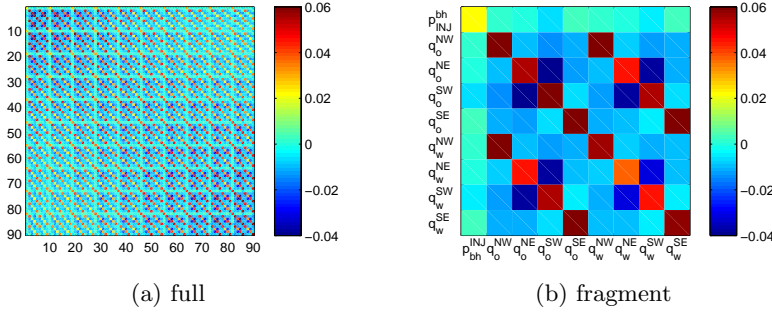


Figure 5.1: Observation sensitivity matrix (obtained with the RM).

of the same type collected at exactly the same location at different time instances. The measurements collected just a few time instances away from the current time have a higher impact than those collected further away in time;

- consider e.g. high magnitude values of observation sensitivity coefficients S_{2j} for $j = 9k + 6$, $k = 0, 1, \dots, 9$. These coefficients illustrate that a particular predicted oil flow rate measurement is highly influenced by the water flow rate measurements collected at exactly the same location. The high magnitudes of e.g. S_{6j} for $j = 9k + 2$, $k = 0, 1, \dots, 9$, demonstrate the impact of the oil flow rate data on the predicted water flow rates;
- observation sensitivity coefficients $S_{ij} \approx 0$ for every $i = 9k + l$, $k = 0, 1, \dots, 9$, $l = 2, 3, \dots, 9$, and $j = 9k + 1$, $k = 0, 1, \dots, 9$. This indicates that bottom hole pressure measurements p_{bh} at the injection well have almost no influence on the analysis;
- by the same token, observation sensitivity coefficients S_{ij} are approximately zero for $i = 9k + 1$, $k = 0, 1, \dots, 9$, and every $j = 9k + l$, $k = 0, 1, \dots, 9$, $l = 2, 3, \dots, 9$. Hence, the predicted values of the bottom hole pressure measurements are also hardly influenced by assimilation of any other available data in the current measurement network set-up;
- observation sensitivity coefficients S_{ij} for $i = 9k + 3$ and $j = 9k + 4$, $k = 0, 1, \dots, 9$, are negative with large magnitudes. These coefficients describe the sensitivity of the predicted q_o measurement at the NE well to the q_o data from the SW well. Note that both wells are connected by a high-permeable area and are somewhat 'competing' for the flow. Hence, an increase of the actually measured oil flow rate in the SW well

immediately results in adjusting the model to decrease the predicted oil flow rate at the NE well, and vice versa. The same reasoning explains the values of the observation sensitivity coefficients S_{ij} for $i = 9k + 4$ and $j = 9k + 3$, $k = 0, 1, \dots, 9$. Analogous arguments hold for sensitivity coefficients S_{ij} for $i = 9k + 7$ and $j = 9k + 8$, $k = 0, 1, \dots, 9$, and S_{ij} for $i = 9k + 8$ and $j = 9k + 7$, $k = 0, 1, \dots, 9$. In general, similar phenomena can be observed for all sensitivity coefficients related to the predicted and measured phase flow rates.

Suppose one has now available for history matching the same production data as in Section 2.3, but twice as accurate, i.e. with 5% error. Then according to inequality (5.15) the maximum change one may expect to achieve in the model due to such an improvement in the accuracy of the observations is given by

$$\left| \frac{\|\hat{\epsilon}_2\|_2}{\|\hat{\mathbf{y}}\|_2} - \frac{\|\hat{\epsilon}_1\|_2}{\|\hat{\mathbf{y}}\|_2} \right| \stackrel{(5.15)}{\leq} 0.04 \cdot \|\mathbf{S}\|_2 \cdot \sqrt{\alpha_1^2 + \alpha_2^2} = 0.04 \cdot 0.997 \cdot \sqrt{10^2 + 5^2} = 0.45.$$

Improving the accuracy of the data further, i.e. reducing the value of α_2 to zero in the above inequality, we obtain the following assessment on the upper boundary for the achievable change in the model accuracy:

$$\left| \frac{\|\hat{\epsilon}_2\|_2}{\|\hat{\mathbf{y}}\|_2} - \frac{\|\hat{\epsilon}_1\|_2}{\|\hat{\mathbf{y}}\|_2} \right| \leq \lim_{\alpha_2 \rightarrow 0} 0.04 \cdot \|\mathbf{S}\|_2 \cdot \sqrt{\alpha_1^2 + \alpha_2^2} = 0.04 \cdot 0.997 \cdot \sqrt{10^2} = 0.40.$$

This result yields that the accuracy of the calibrated model can be somewhat influenced by choosing more precise measurements for the history matching, but only to a limited extent. Hence, if the measurements comprising the original data set did not have enough influence on the data assimilation to produce a reasonable history matched model, it might be not possible to obtain a better model by improving only the accuracy of these measurements. One might need to reconstruct the measurement network to include different data.

However, if we consider the change in the uncertainty of the estimate through (5.19), we obtain

$$\frac{\|\mathbf{P}_2^H - \mathbf{P}_1^H\|_2}{\|\mathbf{P}_1^H\|_2} \leq \frac{\|\mathbf{S}_1\|_2 \cdot \|\mathbf{R}_2 - \mathbf{R}_1\|_2 \cdot \|\mathbf{S}_2\|_2}{\|\mathbf{P}_1^H\|_2} = 0.75.$$

Such a high boundary implies that by improving the accuracy of the observations from 10% to 5% (i.e. twice) one may expect to change the uncertainty associated with the history matched model up to 75%. If the calibrated model is accurate enough, narrowing the uncertainty range down can be beneficial. To the contrary, if the calibrated model is far from the true one, a significantly

reduced uncertainty range may not span the true model which is undesirable. The analyst, therefore, needs to carefully manage both the accuracy of the history-matched model and the associated uncertainty.

We proceed by studying the SVD of the observation sensitivity matrix.

5.6.2 SVD of the observation sensitivity matrix

Consider the outcome of a SVD performed on the observation sensitivity matrix (for the general theory on the SVD of a matrix see Appendix B.3). The singular values of matrix \mathbf{S} are plotted on logarithmic scale in Figure 5.2a and the corresponding singular vectors are given in Figure 5.2b. Then right singular vectors of \mathbf{S} determine linear combinations of data that are satisfied by the adjustments due to parameter estimation. Moreover, in our case the right and left singular vectors coincide as we actually have an eigenvalue decomposition.

The shape of the curve presented in Figure 5.2a provides additional insights into the features of the measurement network. If the curve is flat, all the linear combinations of the measurements can be concluded to have influenced the history matching process. However, sharp bending of the curve down to zero illustrates that some linear combinations of the measurements are not important for the analysis and can be omitted, simplifying the data assimilation. Further investigation is needed to formulate the criteria for selection of the redundant data based on the shape of the curve.

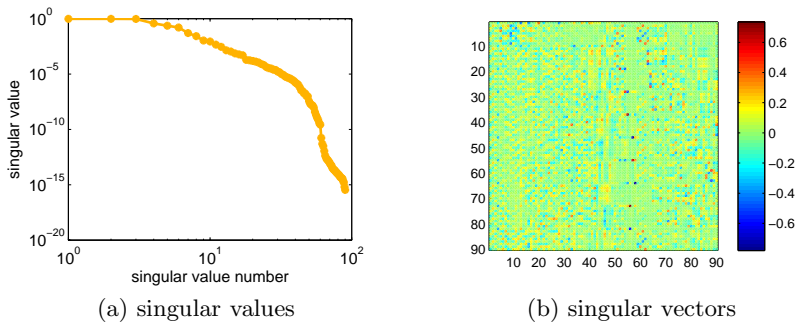


Figure 5.2: Singular values and singular vectors of the observation sensitivity matrix.

A closer look at the singular vectors corresponding to the three largest singular values (Figure 5.3) provides some insight in measurement combinations that have the largest influence on the history matching process. One can notice a clear repetition of the most influential patterns of measurements in time in the first three singular vectors, e.g. it follows from the first singular

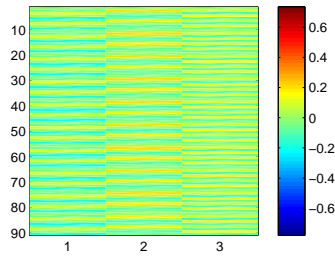


Figure 5.3: First three singular vectors of the observation sensitivity matrix.

vector that the combination of phase flow rates at the NE and SW wells is important for the analysis.

5.6.3 Trace diagnostics of the observation sensitivity matrix

The original observation sensitivity assessment via (5.3) can now be performed at each iteration separately. For our test case the GAI of the observations per iteration and the corresponding GAI for the prior I_P are given in Table 5.1.

Iteration	Value of cost function J	I_{GA}	I_P
initial	1001.48	—	—
1 st	55.87	0.0414	0.9586
2 nd	47.35	0.0411	0.9589
3 rd	47.23	0.0428	0.9572
4 th	47.23	0.0428	0.9572
5 th	47.23	0.0428	0.9572

Table 5.1: GAI and PAI per iteration of representer method.

However, one is naturally interested in more general measures of observational impact. One of the questions to be answered is how an observation influences the final outcome of the estimation procedure. For this matter one might focus only on the observation influence calculated for the last iteration of the algorithm, since then the linearization is performed in the neighborhood of the (local) minimum and, hence, provides the best possible linear approximation to the problem.

Alternatively, one could also look at the importance of the observations from a point of view of algorithm convergence. It was already observed that the main improvement of the model occurs in the first iteration of the history matching algorithm, and even if the algorithm has not yet converged, the

model has already a reasonable quality. Therefore the observations providing such an improvement need to receive a higher score. Then the overall observation influence can be defined as a linear combination of observation influences with weights related to the overall drop of the cost function and the drop per iteration:

$$I_{GA} = \sum_{i=1}^{N_{iter}} \frac{|J_i - J_{i-1}|}{|J_{N_{iter}} - J_0|} I_{GA}^i,$$

where N_{iter} stands for the total number of executed outer loops of the representer algorithm.

In the current exercise the fluctuations in the influence of observations between different iterations are negligible and lead to equivalence of these two approaches. In a general case, however, it is probably best to work with the sensitivity matrix obtained at the last iteration of the RM.

Finally, the GAI for the current exercise is only 4%. The GAI of the prior is therefore 96%. Thus, in our example the majority of the information in the history matching exercise results from the prior. This result is in line with the practical experience that the outcome of computer-assisted history matching with production data is strongly influenced by the prior model.

Additional exercises are required to draw general conclusions about the dependency between the location / type of the measurement and its importance.

5.6.4 Role of measurement type and/or location

To obtain a better insight into the trace diagnostics defined in Section 5.2 we perform a number of history matching exercises based on the set-up described in Section 2.3. This set-up is slightly adjusted to investigate the dependency between the measurement location / type and its importance to history matching. The test cases involve various combinations of well control strategies and, hence, different combinations of measurements, whereas the rest of the setting remains as in Section 2.3. The summary of the set-ups is given in Table 5.2. Note that Exercise 5.6.4.1 was discussed in detail in Sections 5.6.1–5.6.3.

The estimates of the permeability and porosity fields obtained in Exercises 5.6.4.1–5.6.4.6 are given in Figures 5.4 and 5.5 respectively illustrating that a reasonable reconstruction of the field features is achieved in all cases.

There are, however, some visual differences in the quality of estimated permeability and porosity fields resulting from Exercises 5.6.4.1–5.6.4.6. To investigate the link between the quality of the estimate, measurement strategy and information content of the data we summarize the average information content of the observations and the PAI for the measurements of different types

	prescribed		measured	
	p_{bh}	q_t	p_{bh}	q_o, q_w
Exercise 5.6.4.1	NW, NE, SW, SE	INJ	INJ	NW, NE, SW, SE
Exercise 5.6.4.2	—	NW, NE, SW, SE, INJ	NW, NE, SW, SE, INJ	—
Exercise 5.6.4.3	NW, SE, INJ	NE, SW	NE, SW	NW, SE, INJ
Exercise 5.6.4.4	NE, SW, INJ	NW, SE	NW, SE	NE, SW, INJ
Exercise 5.6.4.5	NW, NE, INJ	SW, SE	SW, SE	NW, NE, INJ
Exercise 5.6.4.6	SW, SE, INJ	NW, NE	NW, NE	SW, SE, INJ

Table 5.2: Settings of the exercises with various measurement strategies.

computed via (5.3), and the rms error in the obtained parameter estimates in Table 5.3.

	K	I_{GA}	I_{PA}			E_{rms}	
			p_{bh}	q_o	q_w	\mathbf{k}	ϕ
Exercise 5.6.4.1	90	0.0428	0.0360	0.0492	0.0381	0.5541	0.0665
Exercise 5.6.4.2	50	0.0200	0.0200	—	—	0.7500	0.0830
Exercise 5.6.4.3	70	0.0649	0.0377	0.1013	0.0588	0.6626	0.0758
Exercise 5.6.4.4	70	0.0759	0.0837	0.0994	0.0550	0.5012	0.0617
Exercise 5.6.4.5	70	0.0651	0.0661	0.0883	0.0490	0.5955	0.0708
Exercise 5.6.4.6	70	0.0663	0.0658	0.0898	0.0510	0.5503	0.0660

Table 5.3: GAI, PAI and rms error in the estimated parameters for exercises with various measurement strategies.

We start by comparing the outcomes of the Exercises 5.6.4.3 – 5.6.4.6, since the measurement scenarios simulated in these exercises are somewhat similar. Namely, we collect bottom hole pressure data at two production wells and at the injector, and phase flow rates at the other two producers. The choice of the production wells for p_{bh} measurements varies per exercise. The following remarks can be made based on investigating the four bottom rows of the Table 5.3:

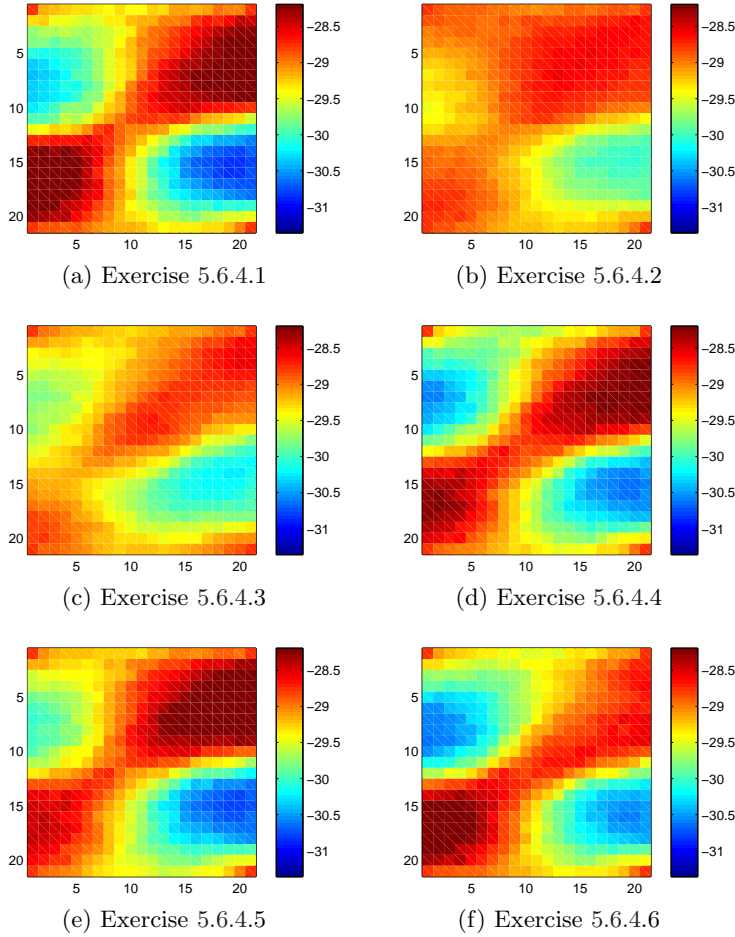


Figure 5.4: Estimated permeability field in $[m^2]$ on natural log scale.

- in Exercise 5.6.4.4 the measurements have the largest impact on the history match ($I_{GA} = 0.0759$), and the best estimates of the model parameters ($E_{rms}(\mathbf{k}) = 0.5012$, $E_{rms}(\phi) = 0.0617$) are achieved. In this exercise p_{bh} is measured at the NW and SE wells located in low permeability zones;
- in Exercise 5.6.4.3 the measurements have the smallest impact on the history match ($I_{GA} = 0.0649$), and the worst estimates of the model parameters ($E_{rms}(\mathbf{k}) = 0.6626$, $E_{rms}(\phi) = 0.0758$) are obtained. In this exercise p_{bh} measurements are collected at the NE and SW wells placed in the high permeability streak;

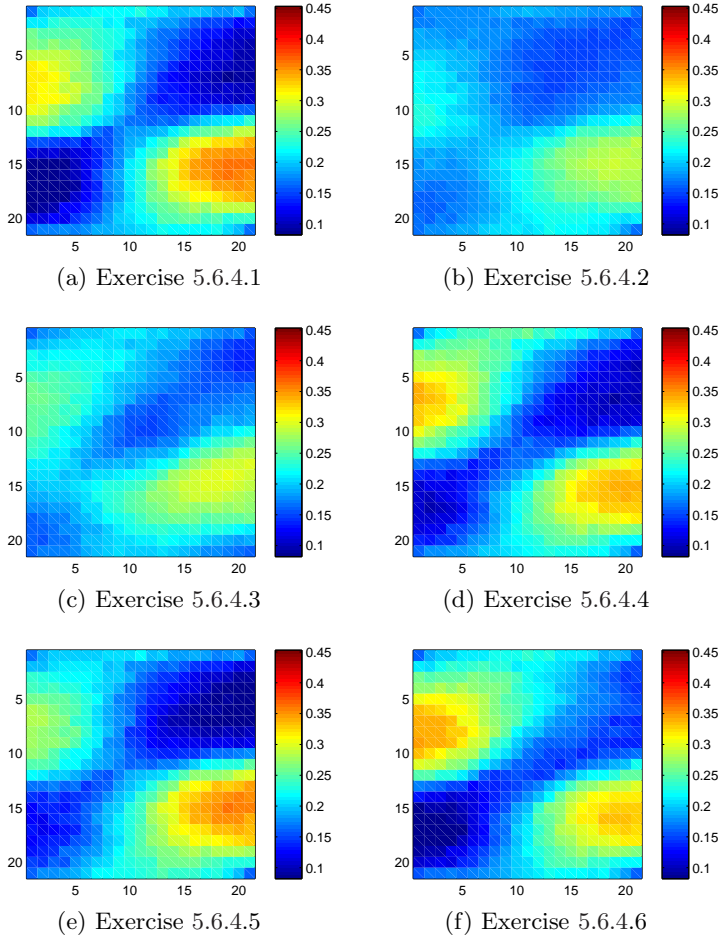


Figure 5.5: Estimated porosity field in $[-]$ on natural $\log(-\log)$ scale.

- the change in GAI for the various exercises is mostly caused by respective changes in the information content of the p_{bh} observations;
- in Exercise 5.6.4.5 and 5.6.4.6 the measurement influence is similar, and also the PAI's for various data types are approximately the same. Note that in these exercises the p_{bh} measurements are collected in a well located within the low permeability area and a well placed outside this zone;
- p_{bh} data have less impact on the history matching procedure than q_o data;
- q_o data are more influential than q_w data.

The last two conclusions can be also drawn from Figure 5.6 demonstrating the influence of the measurements of different types, i.e. the PAI of bottom hole pressure observations and phase flow rates.

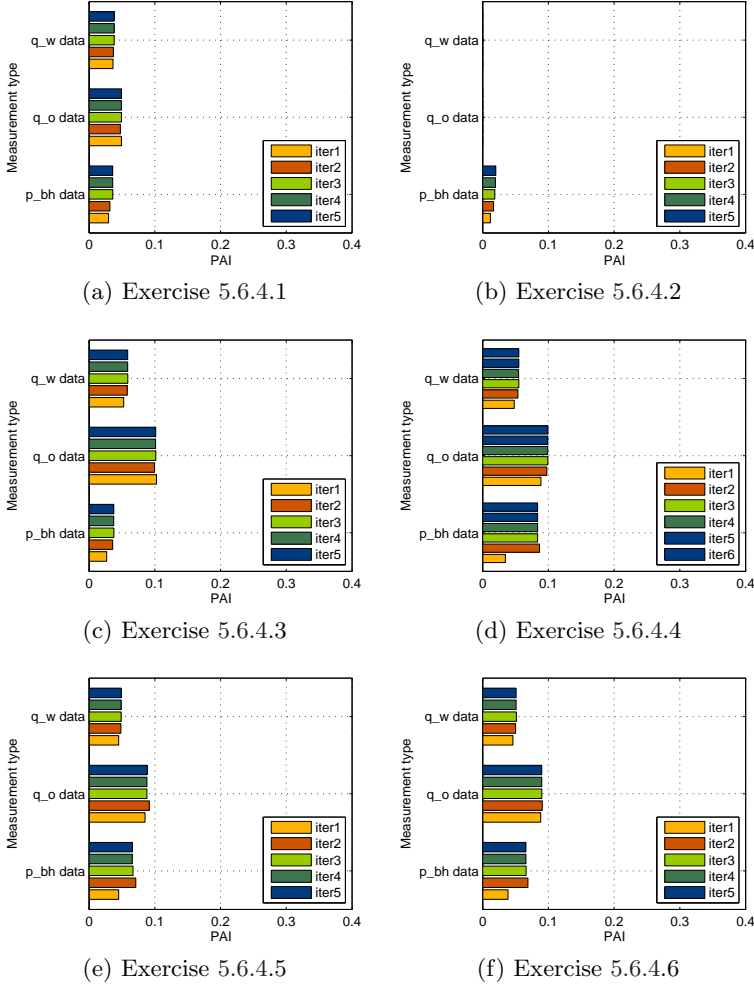
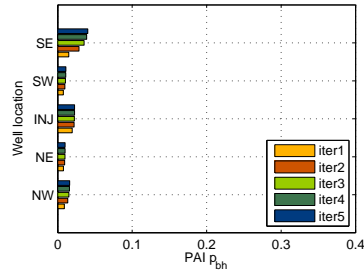


Figure 5.6: PAI for sets of p_{bh} , q_o and q_w measurements for the exercises with various measurement strategies.

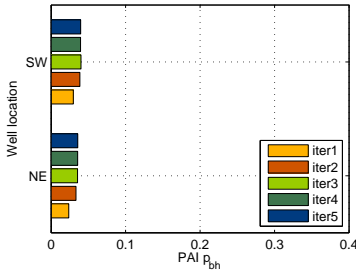
In Exercise 5.6.4.1 p_{bh} data also have less impact on the history matching procedure than q_o data, and q_o data are more influential than q_w data. However, there is notably no connection between the measurement impact and the quality of the estimate achieved in the various exercises. The measurements used in Exercise 5.6.4.1 are clearly less influential than those in Exercise 5.6.4.3 ($I_{GA} = 0.0428$ versus $I_{GA} = 0.0649$), however, the estimated

model parameters are better in the former case (e.g. $E_{\text{rms}}(\mathbf{k}) = 0.5541$ versus $E_{\text{rms}}(\mathbf{k}) = 0.6626$). This fact illustrates that the observation sensitivity matrix for history matching is indeed unable to distinguish between positive and negative influences of the data on the quality of the estimated parameters (see Section 5.2).

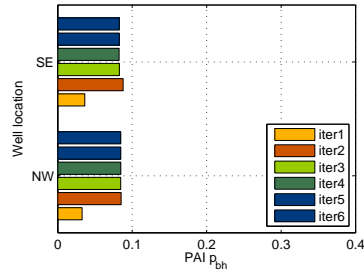
The relation between the location of p_{bh} measurements and their importance for the history matching procedure is illustrated in Figure 5.7, whereas Figures 5.8 and 5.9 give the PAI for q_o and q_w measurements collected at various locations.



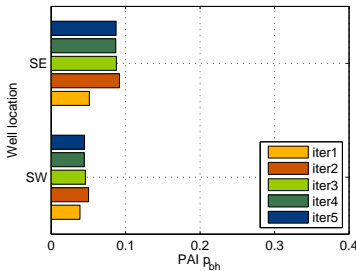
(a) Exercise 5.6.4.2



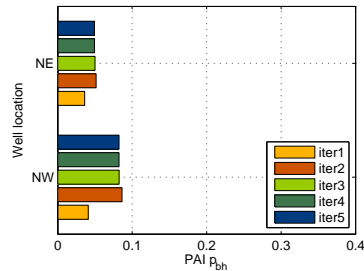
(b) Exercise 5.6.4.3



(c) Exercise 5.6.4.4

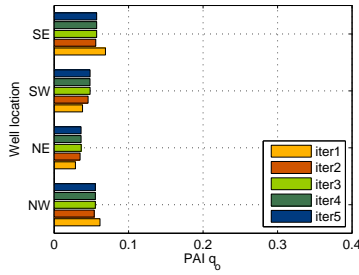


(d) Exercise 5.6.4.5

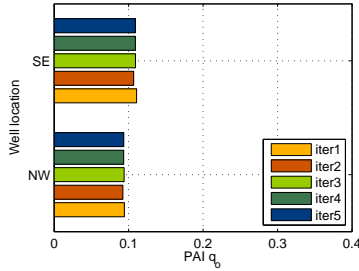


(e) Exercise 5.6.4.6

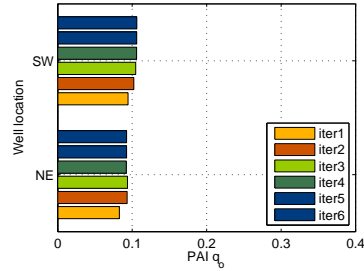
Figure 5.7: PAI for p_{bh} measurements at different locations for the exercises with various measurement strategies.



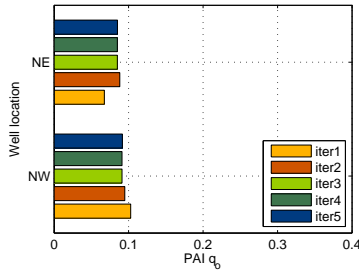
(a) Exercise 5.6.4.1



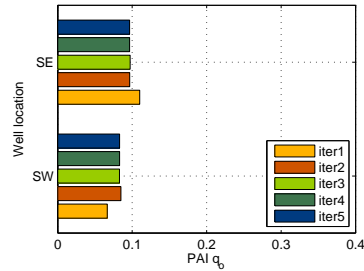
(b) Exercise 5.6.4.3



(c) Exercise 5.6.4.4



(d) Exercise 5.6.4.5

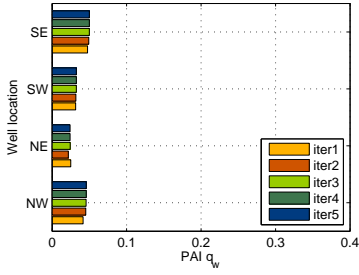


(e) Exercise 5.6.4.6

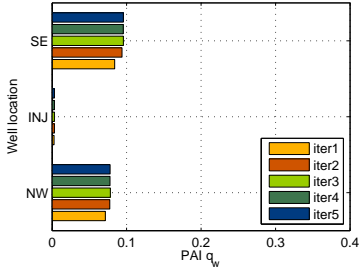
Figure 5.8: PAI for q_o measurements at different locations for the exercises with various measurement strategies.

Careful examination of Figures 5.6–5.9 yields the following remarks:

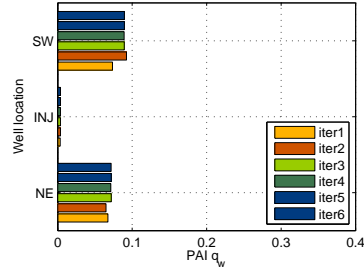
- q_w data collected at injection well are not influential compared to the observations from the other wells;
- p_{bh} (respectively q_o and q_w) data collected at the wells located in the low permeable zone have a higher influence than if collected at the wells located in the high permeability areas. This is in agreement with the discussion of Table 5.3.



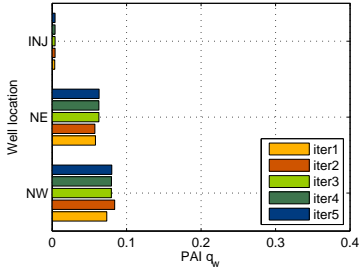
(a) Exercise 5.6.4.1



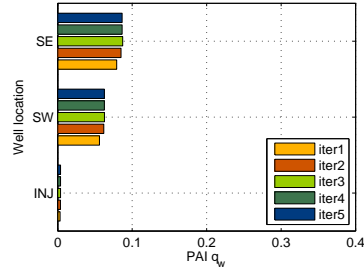
(b) Exercise 5.6.4.3



(c) Exercise 5.6.4.4



(d) Exercise 5.6.4.5



(e) Exercise 5.6.4.6

Figure 5.9: PAI for q_w measurements at different locations for the exercises with various measurement strategies.

Note that in the discussed exercises the low permeability areas are not present in the prior, therefore, it is reasonable to expect that data collected at these locations will contribute the most to the parameter update. To elaborate on that hypothesis we proceed with a study case initialized with a different prior model.

5.6.5 Role of prior fields

To investigate how the observation influence depends on the prior model, we consider the base case described in Section 2.3 and a modified version with different permeability and porosity prior fields (see Figure 5.10). Note that the prior fields given in Figure 5.10 are mirrored versions of the fields in Figure 2.3.

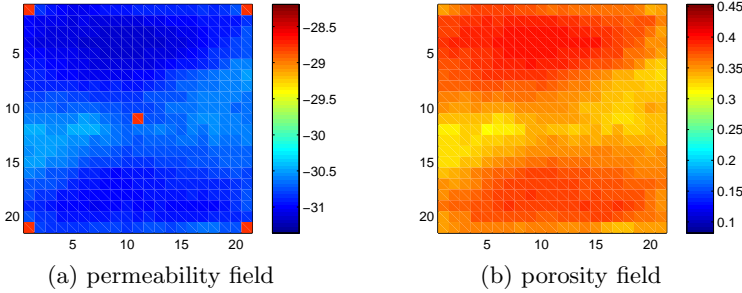


Figure 5.10: Prior permeability and porosity fields. The scale represents transformed values of permeability [m^2] and porosity [-].

History matching with prior fields as in Figure 5.10 shows good results. The value of the cost function drops from 1086.307 to 71.239, while the computed minimum value of the cost function satisfies criterion (3.16), which in this particular case has the form $42.567 \leq J^{\text{opt}} \leq 137.434$. The rms errors in the estimated permeability and porosity fields are reduced from 1.65 to 0.77 and from 0.16 to 0.11 respectively. Though the estimated permeability and porosity fields plotted in Figure 5.11 might be not fully resembling the 'true' fields in Figure 2.1, they contain the main features of the 'true' fields and the prediction of water breakthrough clearly gets improved if the history-matched model is used.

Figure 5.12 illustrates that the predictions of water breakthrough with calibrated models resulting from both regular prior 1 (Figure 2.3) and modified prior 2 (Figure 5.10) have comparable quality.

The results of the sensitivity diagnostics via (5.3) for exercises with both prior fields are given in Tables 5.4 – 5.6.

As the prior permeability field in Exercise 5.6.5.2 does not contain a high permeability pass from the SW to the NE well, it is reasonable to expect that the history matching procedure in this case should rely more on the available production data to reconstruct the truth. This fact is also illustrated by the computed GAI for Exercises 5.6.5.2 and 5.6.5.1 (0.0454 versus 0.0428). The slight increase in the measurement influence is mostly due to the increased importance of the p_{bh} measurements.

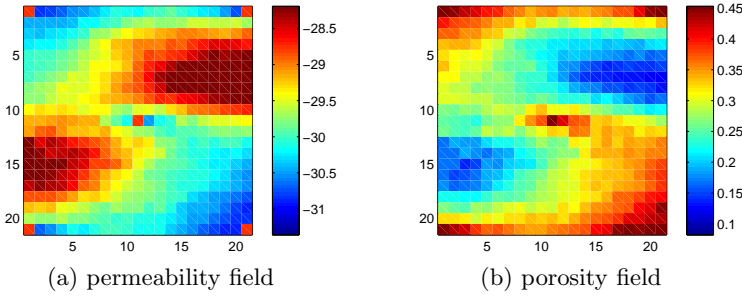


Figure 5.11: Estimated permeability and porosity fields. The scale represents transformed values of permeability [m^2] and porosity [-].

	Prior	I_{GA}	I_{PA}		
			p_{bh}	q_o	q_w
Exercise 5.6.5.1	Figure 2.3	0.0428	0.0360	0.0492	0.0381
Exercise 5.6.5.2	Figure 5.10	0.0454	0.0663	0.0491	0.0364

Table 5.4: GAI and PAI for sets of p_{bh} , q_o and q_w measurements for cases with various priors.

	Prior	I_{PA}			
		NW	NE	SW	SE
Exercise 5.6.5.1	Figure 2.3	0.0554	0.0362	0.0483	0.0571
Exercise 5.6.5.2	Figure 5.10	0.0542	0.0375	0.0495	0.0553

Table 5.5: PAI for q_o measurements collected at different locations for cases with various priors.

	Prior	I_{PA}			
		NW	NE	SW	SE
Exercise 5.6.5.1	Figure 2.3	0.0458	0.0241	0.0325	0.0500
Exercise 5.6.5.2	Figure 5.10	0.0435	0.0238	0.0313	0.0471

Table 5.6: PAI for q_w measurements collected at different locations for cases with various priors.

In both cases measurements collected outside high permeable area are more influential than measurements of the same type collected at other locations, i.e. inside the high permeability zone. This allows reformulating the hypothesis

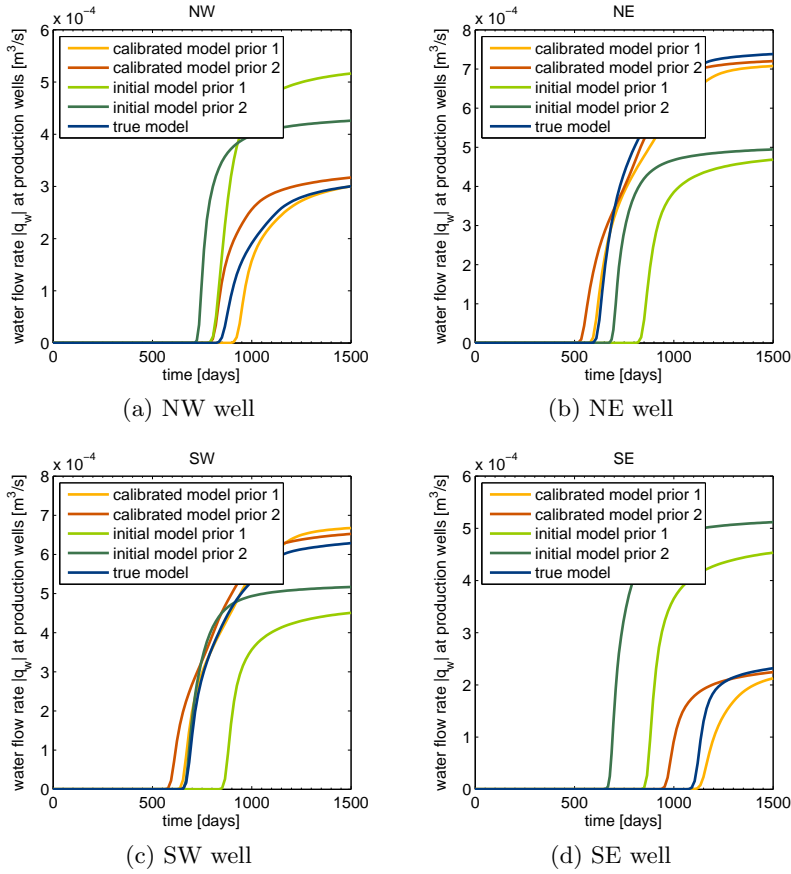


Figure 5.12: Forecast of water flow rate at production wells for the case of various priors.

from the Section 5.6.4: even if the information on low permeability areas is present in the prior, the measurements collected in these regions remain more important than the data of the same type collected in high permeability zones.

5.6.6 Role of 'true' fields

The structure of the 'true' permeability and porosity fields so far used in the synthetic exercises is favorable for history matching (given the utilized 5-spot configuration of the wells). The high permeability streak connecting the SW and NW production wells and the injector is easy to be estimated from the available prior and measurements. Therefore the simulations based on the 'true' fields given in Figure 2.1 might be not enough to draw general conclusions about the features of the history matching process [40, 63]. This

led us to consider somewhat more complex 'true' fields given in Figures 5.13a and 5.14a. These are the fields used in Exercise 5.6.6.1. We enlarge this study by also running synthetic Exercise 5.6.6.2 based on the 'true' fields given in Figures 5.13b and 5.14b. The rest of the set-up remains as described in Section 2.3.

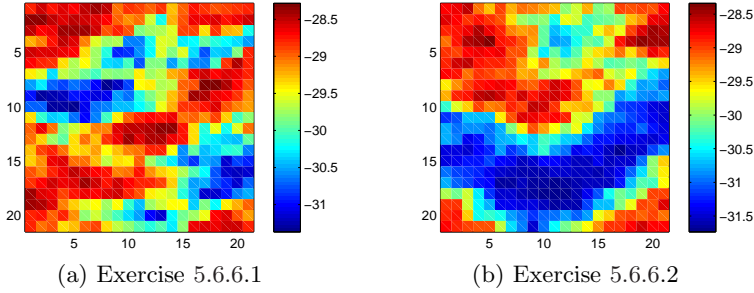


Figure 5.13: Various 'true' permeability fields. The scale represents transformed values of permeability [m^2].

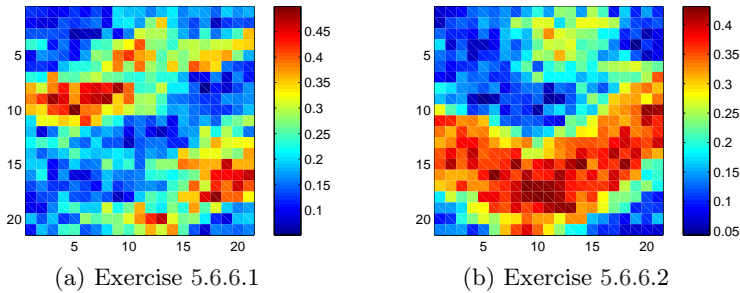


Figure 5.14: Various 'true' porosity fields. The scale represents transformed values of porosity [-].

The estimated permeability and porosity fields are given in Figures 5.15 and 5.16, demonstrating that the main features of the 'true' fields were recovered during the history matching process.

Tables 5.7–5.9 contain the results of the sensitivity diagnostics for Exercises 5.6.6.1–5.6.6.2 and the reference case with the regular truth given in Figure 2.1. Recall that the reference case was discussed in detail throughout Sections 5.6.1–5.6.3.

The results demonstrate similar low values of GAI of the observations for all the performed study cases. Considering the importance of the measurements of the same type collected at different locations (Tables 5.8–5.9), one

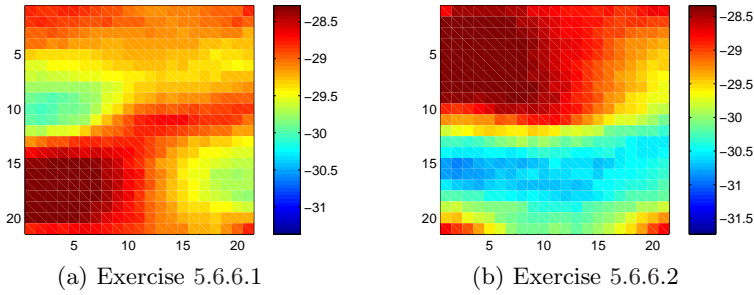


Figure 5.15: Estimated permeability fields for the case of various truth. The scale represents transformed values of permeability [m^2].

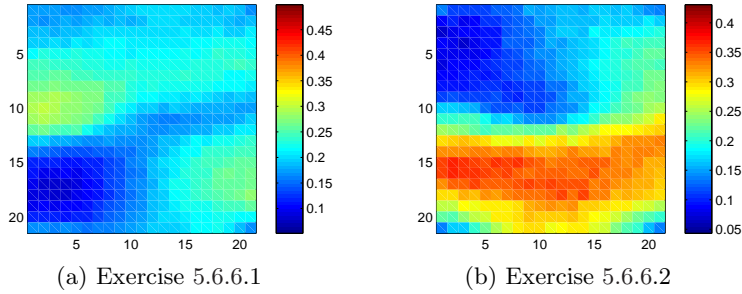


Figure 5.16: Estimated porosity fields for the case of various truth. The scale represents transformed values of porosity [-].

	Truth	I_{GA}	I_{PA}		
			p_{bh}	q_o	q_w
Exercise 5.6.6.0	Figure 2.1	0.0428	0.0360	0.0492	0.0381
Exercise 5.6.6.1	Figures 5.15a, 5.16a	0.0427	0.0372	0.0491	0.0376
Exercise 5.6.6.2	Figures 5.15b, 5.16b	0.0462	0.0431	0.0519	0.0414

Table 5.7: GAI and PAI for sets of p_{bh} , q_o and q_w measurements for cases with various truth.

can finally conclude that the data gathered at production wells that do not have a direct high permeability connection to the injector are more important for the history matching procedure than the data of the same type collected at producers linked to the injector by a high permeability streak. A possible explanation could be that if the producer is separated from the injector by a low permeability zone, the flow that reaches this producer contains more

	Truth	I_{PA}			
		NW	NE	SW	SE
Exercise 5.6.6.0	Figure 2.1	0.0554	0.0362	0.0483	0.0571
Exercise 5.6.6.1	Figures 5.15a, 5.16a	0.0519	0.0510	0.0411	0.0524
Exercise 5.6.6.2	Figures 5.15b, 5.16b	0.0287	0.0471	0.0671	0.0646

Table 5.8: PAI for q_o measurements collected at different locations for cases with various truth.

	Truth	I_{PA}			
		NW	NE	SW	SE
Exercise 5.6.6.0	Figure 2.1	0.0458	0.0241	0.0325	0.0500
Exercise 5.6.6.1	Figures 5.15a, 5.16a	0.0408	0.0400	0.0273	0.0422
Exercise 5.6.6.2	Figures 5.15b, 5.16b	0.0168	0.0345	0.0581	0.0561

Table 5.9: PAI for q_w measurements collected at different locations for cases with various truth.

information about the interior structure of the reservoir as it has to find ways to bypass the low permeability region. In turn, if injection and production wells are directly connected by a high permeability streak, the flow immediately shoots through to the producer, thus limiting the amount of information about the reservoir structure gathered on the way.

5.7 Application of the observation sensitivity matrix to history matching with the AEnKF

Until now we were solving the history matching problem by means of the RM which is a variational technique. It allowed for a straightforward calculation of the sensitivity matrix, but needed the model adjoint for the experiment. This requirement is often not feasible, so application of the sensitivity matrix concept to an adjoint-free history matching system is of practical interest. In the current application, the history matching problem is solved via the AEnKF, which has several favorable features for our purposes:

- The AEnKF does not require an adjoint model (like none of the methods from the family of ensemble Kalman filtering techniques);
- The AEnKF allows for asynchronous data assimilation, i.e. history matching observations made at a time different from the time of the

update. Hence, all observations collected during a certain time-window can be assimilated at once at the end of observational period. This allows for comparison of the influence of the observations collected at different times.

For this exercise we employ the usual set-up described in Section 2.3. The history matching is performed from 0 [days] till 450 [days] with data available every 30 [days] from 30 [days] till 300 [days] with the AEnKF algorithm. The choice of the time interval ensures that water breakthrough has not yet occurred and we can compare the results to the ones obtained with the RM.

The elements of the observation sensitivity matrix are plotted row-wise as a 2D field (see Figure 5.17, where Figure 5.17b displays a detailed version of the top left block of Figure 5.17a).

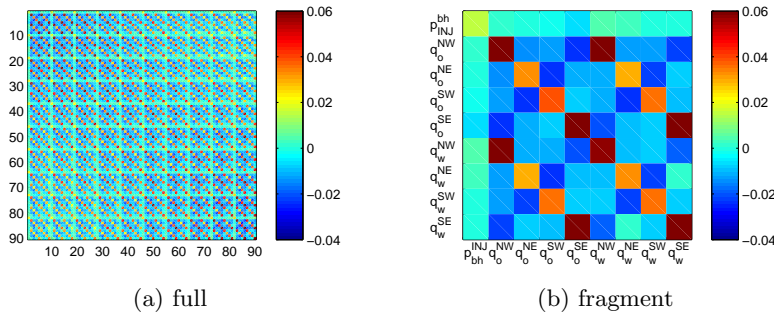


Figure 5.17: Observation sensitivity matrix (obtained with the AEnKF).

Comparison of the GAI calculated for the exercises with the AEnKF algorithm (0.0428) and the RM (0.0428), and Figures 5.1 and 5.17 allows to conclude that both the RM and the AEnKF algorithm have utilized the available measurements in a similar manner. We can therefore proceed with investigating the importance of the measurements for the history matching exercises in general by considering either the AEnKF or the RM history matching technique.

5.7.1 Role of water breakthrough

We have so far observed (see Section 5.6.4) that the amount of information brought to the history-matched system by measurements changes if the flow behavior in reservoir is altered. In that respect water breakthrough is known to substantially influence the system, and it is therefore of interest to investigate the impact of the measurements collected after water breakthrough in at least one well. As our current implementation of the RM does not permit to perform

history matching for the time interval including water breakthrough time, we use the AEnKF algorithm instead.

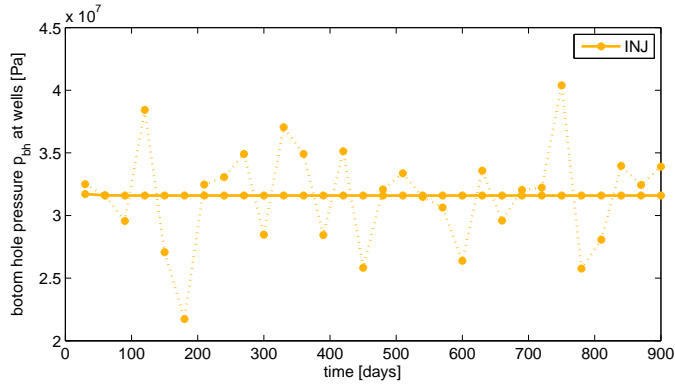
For this exercise we employ the usual set-up described in Section 2.3 (Exercise 5.7.1.1). The history matching is performed from 0 [days] till 900 [days] with the data available every 30 [days] from 30 [days] till 900 [days] with the AEnKF algorithm. The choice of the time interval ensures that water breakthrough has occurred, and we can evaluate the importance of the measurements collected within various time intervals containing or not containing water breakthrough. For comparison we also consider an exercise with a unit mobility set-up (Exercise 5.7.1.2) that utilizes a linear model for relative permeabilities. The description of the differences in both exercises is given in the Table 5.10.

	Viscosity	End point relative permeability	Corey exponents
Exercise 5.7.1.1	$\mu_o = 5 \cdot 10^{-4}$ [Pa · s]	$k_{ro}^0 = 0.9$ [–]	$n_o = 2.0$ [–]
	$\mu_w = 1 \cdot 10^{-3}$ [Pa · s]	$k_{rw}^0 = 0.6$ [–]	$n_w = 2.0$ [–]
Exercise 5.7.1.2	$\mu_o = 1 \cdot 10^{-3}$ [Pa · s]	$k_{ro}^0 = 1.0$ [–]	$n_o = 1.0$ [–]
	$\mu_w = 1 \cdot 10^{-3}$ [Pa · s]	$k_{rw}^0 = 1.0$ [–]	$n_w = 1.0$ [–]

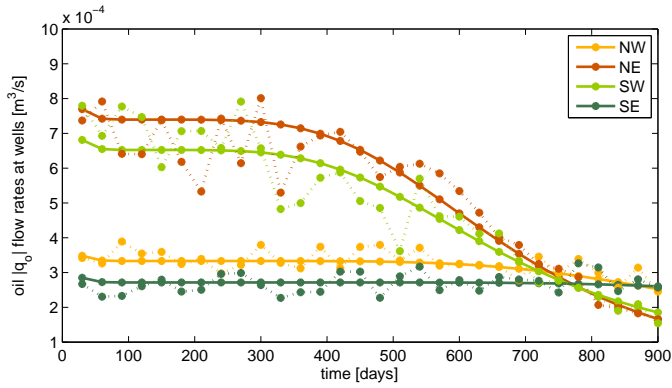
Table 5.10: Settings of the exercises with various relative permeability models.

Synthetic measurements generated for Exercise 5.7.1.2 are plotted in Figure 5.18.

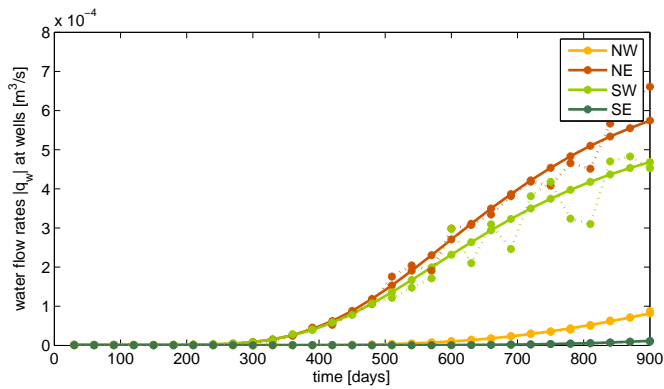
The general observation sensitivity assessment via (5.3) is now performed for each study case and the results are given in Table 5.11. This table also contains the PAIs of the measurements taken within two time intervals: 0 [days] – 300 [days] and 300 [days] – 900 [days], i.e. before any water breakthrough and thereafter. In addition Table 5.11 lists the PAIs of the p_{bh} , q_o and q_w groups of measurements computed for different time windows. The resulting $I_{GA} = 0.1668$ for Exercise 5.7.1.1 indicates a higher influence of the data on the history matching process when history matching is performed for a longer time period with the inclusion of the water-breakthrough-related data. The increase in the global averaged influence is mostly due to the more extensive utilization of the q_w data for time intervals containing water breakthrough. The PAI of the q_w measurements collected in the time window 0 [days] – 300 [days] (no water breakthrough has yet occurred) is 0.0192. The PAI of the same type of measurements gathered during the time interval 300 [days] – 900 [days] (water breakthrough in SW, NE and NW producers) equals 0.3926. Similar phenomena can be observed during the PAI assess-



(a) bottom hole pressure



(b) oil flow rate



(c) water flow rate

Figure 5.18: Synthetic measurements (Exercise 5.7.1.2): solid line — 'perfect', dotted line — noisy measurements.

	I_{GA}	I_{PA}	Time window in [days]	
			0 – 300	300 – 900
Exercise 5.7.1.1	0.1668		0.0201	0.2401
		p_{bh}	0.0087	0.0118
		q_o	0.0240	0.1446
		q_w	0.0192	0.3926
Exercise 5.7.1.2	0.0666		0.0511	0.0743
		p_{bh}	0.0060	0.0060
		q_o	0.0176	0.0382
		q_w	0.0958	0.1275

Table 5.11: GAI and PAI for groups of measurements collected at different time intervals for cases with various relative permeability models.

ment in Exercise 5.7.1.2, though the increase of observation influence is less pronounced.

Note that more influential measurements are collected at times when substantial alterations to the behavior of the fluid flow are observed, e.g there is no change in the p_{bh} in Exercise 5.7.1.2 (see Figure 5.18a), hence, no difference in the importance of p_{bh} measurements gathered at various times is observed; on the contrary, the value of the p_{bh} in Exercise 5.7.1.1 is slightly changing in time (see Figure 2.2a) yielding a small increase in the corresponding PAI.

Figures 5.19 and 5.20 summarize for each well the PAIs of the q_o and q_w measurements collected at different times, and water flow rates predicted by the initial, 'true' and calibrated models. Analysis of these figures provides further details on the importance of q_o and q_w data collected at various time instances / locations:

- the importance of q_w measurements sharply increases just before water breakthrough and sharply drops at the time of water breakthrough in case of the non-linear relative permeability model (see Figure 5.19);
- in case of the non-linear relative permeability model the importance of q_o measurements increases from the time of water breakthrough and on, or at the end of the data assimilation interval if no water breakthrough was observed for the duration of the history matching;
- in the unit mobility (linear permeability model) exercise, the importance of q_w measurements both sharply increases and drops before the water breakthrough event (see Figure 5.20);

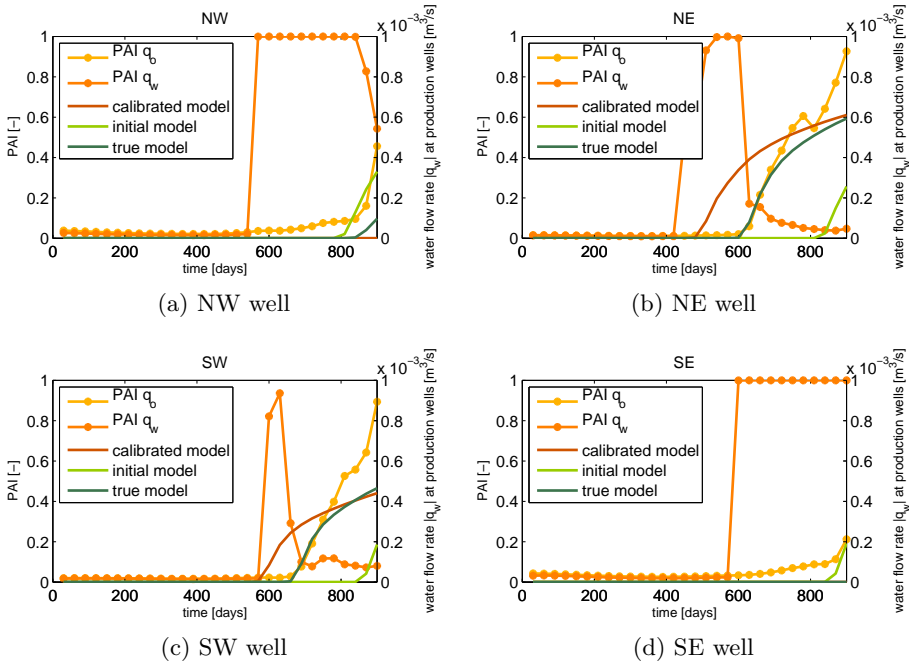


Figure 5.19: PAI for q_o and q_w measurements collected at different locations/time instances. Non-linear relative permeability model (Exercise 5.7.1.1).

- in both exercises the importance of q_o measurements remains relatively low until the water breakthrough time suggesting that these measurements provide redundant information to the data assimilation system.

These findings suggest including both self-sensitivities of the q_w measurements and the predictions obtained from the initial and calibrated models into considerations on the times of water breakthrough. The self-sensitivities of the q_w observations can be also obtained and used directly after history matching to provide the first rough estimates of the water breakthrough times, i.e. already before the simulation with the calibrated model is completed. These estimates are somewhat conservative, which however allows enough time for taking decisions on the necessary adjustments in the production strategy and their implementation.

5.7.2 Role of exciting input

The results so far indicate that production data generally have a low influence on the history match, and the estimation mostly relies on prior information.

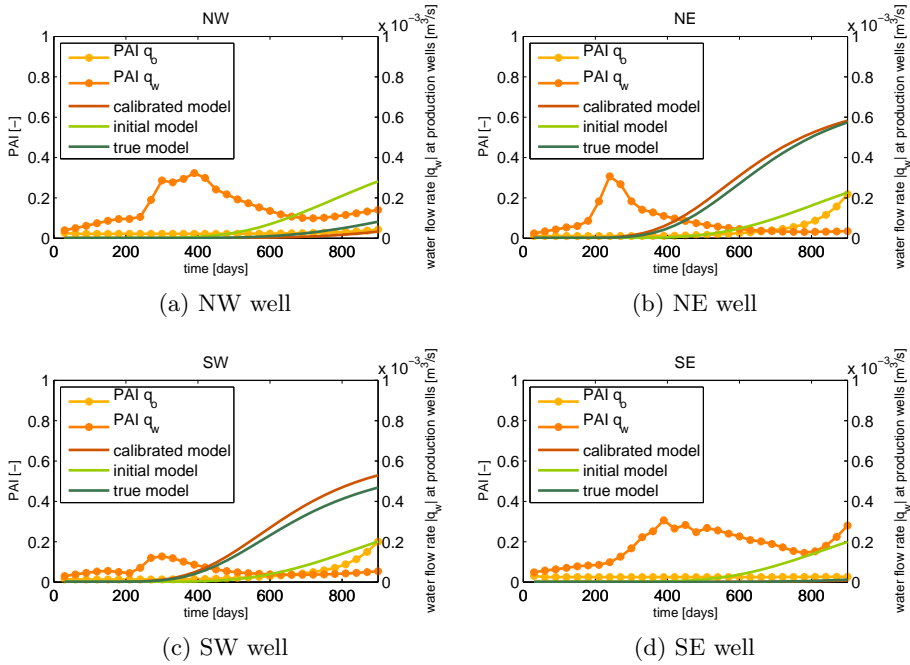


Figure 5.20: PAI for q_o and q_w measurements collected at different locations/time instances. Linear relative permeability model (Exercise 5.7.1.2).

These facts are in agreement with practical experience and results of previous studies (see e.g. [82, 83]). Section 5.7.1 nevertheless demonstrates that the importance of the data is highly increased if water breakthrough related measurements are used for history matching. However, in practice we would like to obtain a calibrated model before this kind of data becomes available, such that the model is able to actually predict water breakthrough. Hence, we need to look for other ways to obtain more influential data.

It is known within the field of system identification, that some parameters of a dynamic system may not be identifiable from the measurements (i.e. system output signals) unless the system input is varied [54]. The data, collected under such a changing input, allows to restore more of the unknown model parameters. Therefore we have found a number of concepts from system identification theory and experiment design to be particularly useful. The idea is to supply the system with such an exciting input signal that the obtained measurements can be utilized by history matching algorithms more effectively. In this respect one needs to decide on: (i) which system input can be used as an exciting signal, (ii) what shape does this input signal need to have, and (iii) what is the maximum permitted change in the value of the input.

In the framework of our base case exercise described in Section 2.3 the injection rate is the most natural choice of the input signal to be altered. To keep the method practically feasible, we propose the injection rate to be changed at discrete time moments that coincide with the time instances when measurement data are collected.

Concerning the shape of the exciting signal, we follow [50] and [81] and consider a number of typical exciting signals, i.e. a step function, a random binary signal (RBS) and a harmonic signal. In all exercises the amplitude of the signal is taken to be 15% of its initial value in order to stay consistent, since a larger amplitude frequently led to convergence problems in the software. The rest of the history matching set-up remains as in Section 2.3.

Step function

We use a step function to describe the water injection strategy and change the system behavior. The exact shape of the input signal is given in Figure 5.21, where the change in the value of the injection rate occurs at 210 [days]. The response of the system to such an input is given in Figure 5.22. Note

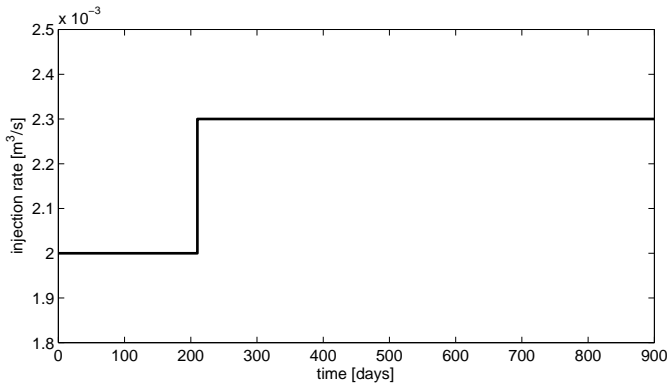
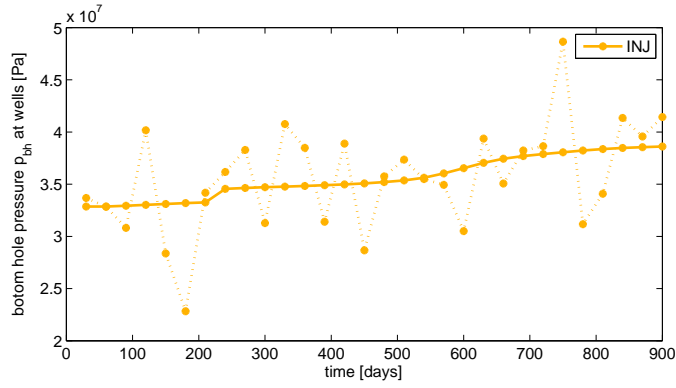
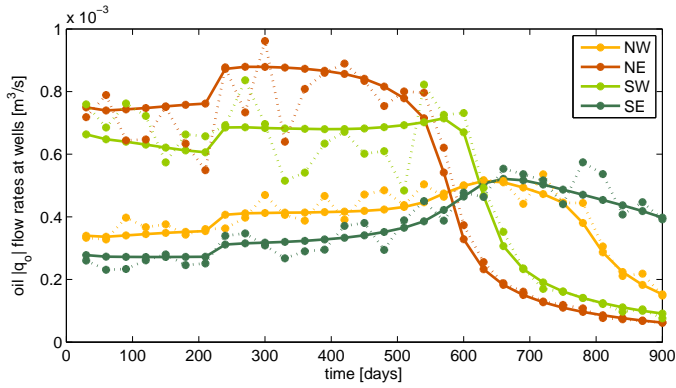


Figure 5.21: Input signal determined by a step function.

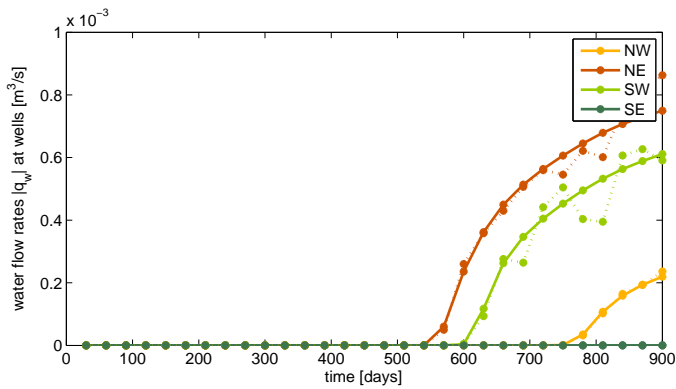
that a stepwise increase in the injection rate results in higher production rates and, hence, in a somewhat faster water breakthrough (compare Figure 5.22 to the base case Figure 2.2). Although the system is exposed to a sudden change in the input, the system response turns out to be unable to provide the history matching algorithm with extra knowledge on the model parameters ($I_{GA} = 0.0434$ vs. $I_{GA} = 0.0428$ in the regular exercise).



(a) bottom hole pressure



(b) oil flow rate



(c) water flow rate

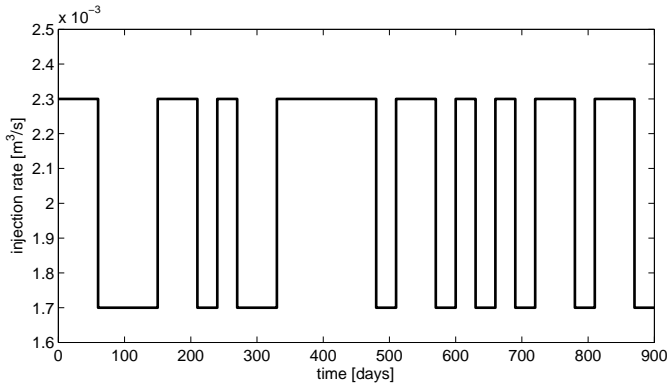
Figure 5.22: Synthetic measurements for stepwise input signal: solid line — 'perfect', dotted line — noisy measurements.

Random binary signal

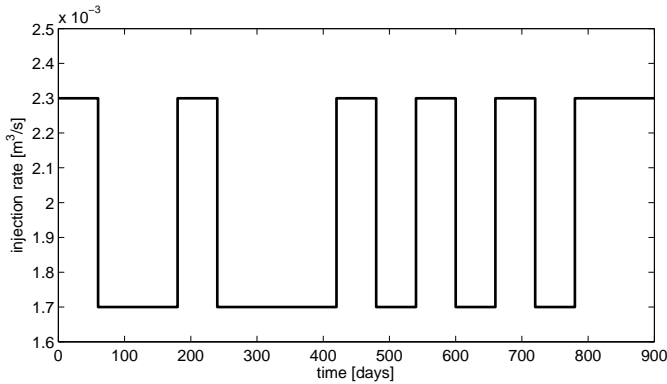
Let $\{w(t)\}$ be a stochastic white noise process, then

$$u(t) = c \cdot \text{sign}[w(t)] \quad (5.20)$$

is a binary signal that assumes values $\pm c$, i.e. c equals 15% of the initial value of the injection rate in our case. The reasons to use such a signal in a general case are discussed in [81]. In the current study, however, the main advantage is that an injection rate excited according to (5.20) can be straightforwardly obtained in the existing physical configuration by e.g. opening or closing a valve at discrete times for discrete amounts. We allow the signal to switch sign every time the new data are collected. The generated RBS signals with clock period chosen as the time step of data gathering ($T_c = 30$ [days]) and with an increased clock period ($T_c = 60$ [days]) are given in Figure 5.23. The



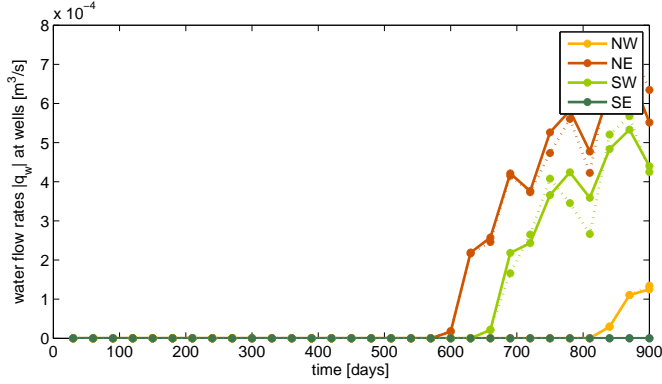
(a) $T_c = 30$ [days]



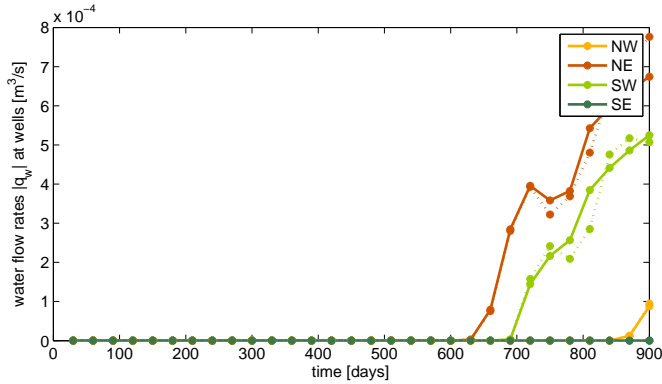
(b) $T_c = 60$ [days]

Figure 5.23: Input signal determined by RBS.

RBS input signal of the same amplitude as the considered above step function introduces only a minor shift in the water breakthrough time (see Figure 5.24).



(a) $T_c = 30$ [days]



(b) $T_c = 60$ [days]

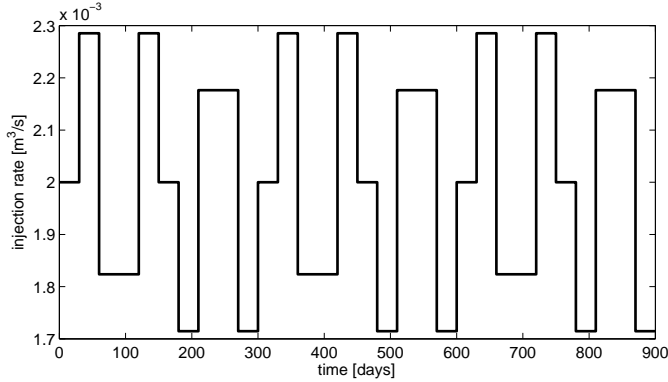
Figure 5.24: Synthetic measurements for RBS input signal: solid line — 'perfect', dotted line — noisy measurements.

The history matching exercises initialized with injection rate varied according to each of the curves in Figure 5.23 however demonstrate no increased sensitivity to the measurements ($I_{GA} = 0.0429$ and $I_{GA} = 0.0424$ vs. $I_{GA} = 0.0428$ in the regular exercise).

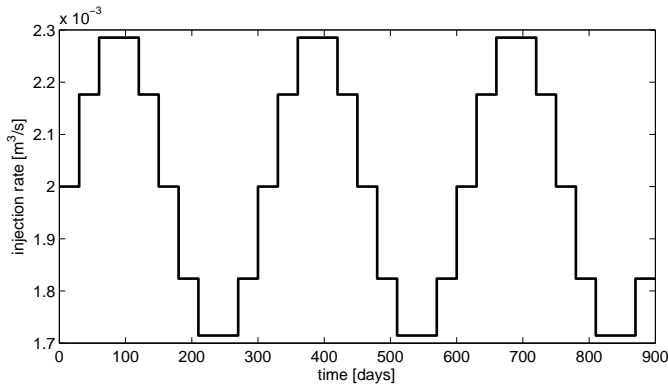
Harmonic signal

Another typical input signal that is used to obtain extra knowledge about a system is described by harmonic functions: sine and cosine. As we intend to alter the injection rate only at discrete time instances, these continuous

functions need to be approximated by piecewise constant functions (see e.g. Figure 5.25 that gives a piecewise constant approximation of sine functions with clock period T_c of 100 [days] and 300 [days]). Note that input in the form



(a) $T_c = 100$ [days]



(b) $T_c = 300$ [days]

Figure 5.25: Input signal determined by sine.

of the approximated harmonic signal does not affect the water breakthrough time (see Figure 5.26).

One needs to decide on the range of possible frequencies to be investigated. The minimum length of the sampling interval in our case is chosen to be 30 [days], i.e. the same as the length of the time intervals used for collecting data. This according to the Nyquist-Shannon theorem [55, 75] allows approximation of the sine functions with frequencies higher than $\frac{1}{2 \cdot 30}$ [days⁻¹]. We therefore consider only sine functions with the period $T_c > 60$ [days].

We perform history matching for 300 [days] and naturally would like to observe within this time at least one full period of the sine wave to cover the range of possible responses of the system. This translates into considering

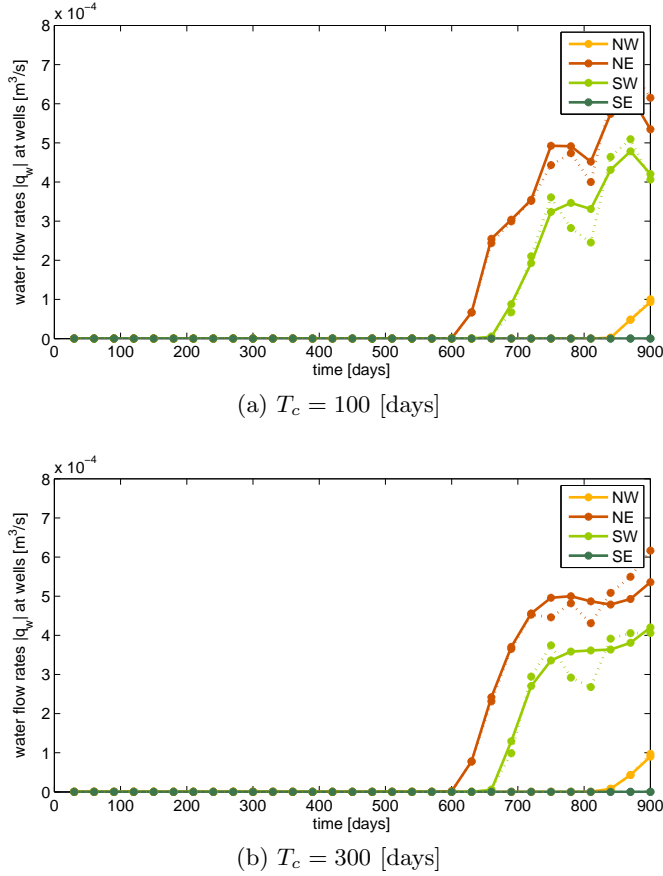


Figure 5.26: Synthetic measurements for approximated sine input signal: solid line — 'perfect', dotted line — noisy measurements.

sine functions with a period $T_c \leq 300$ [days] (i.e. with a frequency lower than $\frac{1}{300}$ [days $^{-1}$]).

We perform simulations for two extreme cases of high and low frequencies: $T_c = 100$ [days] and $T_c = 300$ [days] with the largest amplitude permitted by the software, i.e. 15% of the initial value of the injection rate. Nevertheless, there is no actual difference in the importance of the measurements comparing to the base case ($I_{GA} = 0.0431$ and $I_{GA} = 0.0429$ vs. $I_{GA} = 0.0428$).

Modified staircase signal

The simulations discussed above demonstrated that typical variations in the shape of the input signal do not provide the history matching algorithm with more influential production data. Only the stepwise input signal has shown

some potential in that respect. Unfortunately the magnitude of the step cannot be increased due to software limitations. Therefore we propose an approach that creates an input signal of a staircase type allowing to increase the values of injection rate in total up to 90% of its initial value (see Figure 5.27). This, in turn, results in an earlier water breakthrough times (see Fig-

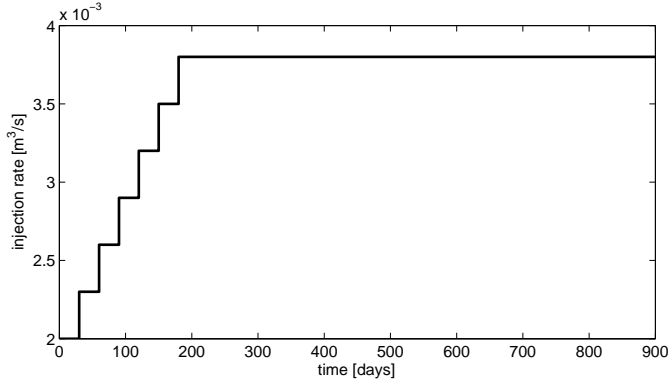


Figure 5.27: Input signal determined by stepwise function.

ure 5.28) and increased importance of the measurements ($I_{GA} = 0.0718$ vs. $I_{GA} = 0.0428$).

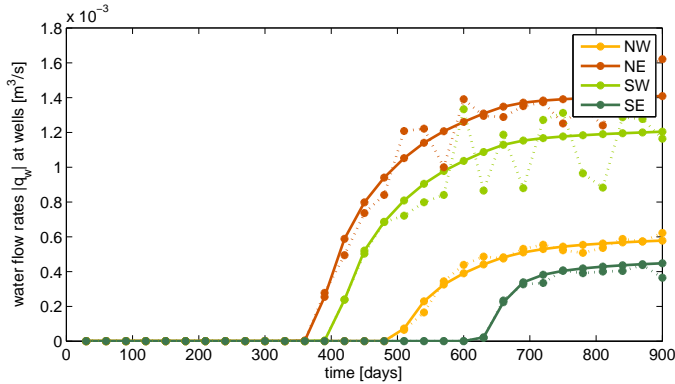


Figure 5.28: Synthetic measurements for modified staircase type input signal: solid line — 'perfect', dotted line — noisy measurements.

The increase in the GAI of the measurements can be explained by the fact that the higher water injection rate causes earlier water breakthrough. Hence, history matching within the regular time window 0 [days] – 300 [days] includes assimilating measurements collected in the vicinity of the water breakthrough. As has been observed in Section 5.7.1, the measurements collected in the neigh-

borhood of the water breakthrough time have a larger influence throughout the history matching procedure. Note that in practice accelerating the occurrence of the water breakthrough is most often not desired, though it can provide more influential measurements.

5.8 Conclusion

We applied a measure developed in meteorology, the observation sensitivity matrix, to a history matching problem in reservoir engineering. In particular, we used the diagonal elements of the matrix, known as self-sensitivities, as a quantitative measure of the influence of observed measurements on predicted measurements. Based on the examples investigated, the main findings are:

- the observation sensitivity matrix can be used as a tool to analyze the importance of the collected measurements for the history matching;
- the norm of the sensitivity matrix can be applied to assessing the magnitude of possible change in the accuracy of the updated model due to the respective change in the accuracy of collected observations (see (5.15) and (5.19));
- RM and AEnKF utilize the data with comparable effectiveness;
- for a simple test case the global averaged influence of the observed measurements is only 4%. This is a rather low value compared to the 96% global averaged influence of the prior;
- bottom hole pressure measurements are in general less influential than fluid flow rates measurements;
- data gathered at production wells that do not have a direct high permeability connection to the injector are more important for the history matching procedure than data of the same type collected at producers linked to the injector by a high permeability streak;
- the fluid flow rate measurements collected at times close to water breakthrough are highly important for history matching;
- together with predictions provided by the initial and calibrated models, the self-sensitivities of the q_w measurements can be used for obtaining the estimations of the water breakthrough times;
- regular modifications in the shape of the function describing the water injection strategy (step function, RBS and harmonic function) do not

result in providing the history matching algorithm with more influential production data, unless also the magnitude of the injection rate is increased yielding earlier water breakthrough times.

Further work is required to assess which influence the assimilated observations have on forecast beyond the data assimilation period.

Conclusions

6.1 Overview

Energy companies worldwide are currently facing a challenging operational landscape. To sustain the company competitiveness on the global energy market, it becomes increasingly important to ensure that appropriate production and development decisions are taken. Reservoir modeling is nowadays a standard tool used in the decision making process allowing analysis and prediction of the reservoir flow behavior, identification of beneficial production strategies and evaluation of the associated risks.

The models used for reservoir simulation contain a large number of imperfectly known parameters characterizing the reservoir flow, e.g. permeability and porosity of the reservoir rock. Therefore the predictive value of such models is limited and tends to deteriorate in time. History matching is employed to update the values of poorly known model parameters in time with the help of the production data which become available during the production life of the reservoir, i.e. to adapt parameters such that simulated results are consistent with measured production data. Such an approach generally improves estimators of the model parameters and predictive capability of the model.

Remarkably, the information extracted from the measurements in the history matching phase is repeatedly found as not enough to provide well-calibrated model with a high predictive value. Hence, consideration of additional data can be of particular help. To optimize the costs and effort associated with collection of new data and computations, up-front selection of the most influential measurements and their locations is desirable. Methods to assess the impact of measurements on model parameter updating are therefore needed.

The research described in this thesis has been focused on quantifying the impact of measured data on the outcome of history matching of reservoir

models, and leads to the following conclusions and recommendations for future studies.

6.2 Conclusions

This research addressed history matching of a two-dimensional two-phase five-well reservoir model with production data (bottom hole pressure at the injection well and oil and water flow rates at the production wells). The study was twofold. It first concerned the applicability and implementation of a number of data assimilation algorithms. The selected algorithms were thereafter used for data assimilation when quantifying the measurements impact on history matching. The following conclusions can be drawn based on the synthetic test cases considered in the thesis.

- A particular variational method, the representer method (RM), has been tested and was found to provide a reasonable parameter estimate. This iterative algorithm is designed to perform the minimization of an objective function over the whole time window using a relatively small amount of observations. The method requires availability of the adjoint model. If an increasing amount of measurements is becoming available the RM turns out to become computationally inefficient. In such a case an accelerated version of the representer method (ARM) can be used. Comparison of the performance of the classical RM and the ARM has demonstrated that the ARM can be controlled to provide outcomes of the same accuracy as those of the RM. Furthermore, the computational load of the RM per iteration has been found to be driven by terms of magnitude K^3 and K^4 , where K is the total number of measurements, whereas the number of operations for the ARM grows as K^3 for large sets of data. This causes the ARM to outperform the classical RM in terms of computational speed when the number of assimilated measurements increases. In a particular example the ARM became more efficient than the RM when more than 85 individual measurements had to be history-matched.
- Variational methods do not allow continuous model updating. Instead of assimilating new data on the fly, i.e. as soon as they become available, the whole history matching process will have to be repeated. The ensemble Kalman filter (EnKF) based techniques are the methods commonly used to deal with the problem. Furthermore, these methods do not require the adjoint model to be available. The classical EnKF has been tested and was found to give improvements in parameter estimates similar to the RM and ARM methods. Unfortunately, EnKF can be

occasionally cost inefficient in the operational environment, when the forward model run has to be interrupted too frequently to assimilate the most recent data. It was therefore decided to also examine the asynchronous modification of the EnKF (AEnKF) technique, which allows assimilating data in batches at times different from when the data became available. Results of history matching with the EnKF and the AEnKF have been demonstrated to be close.

- The RM and the AEnKF methods are theoretically related, although they belong to different families of data assimilation techniques. In our examples, the two methods also yielded analogous estimates of the model parameters. Moreover, since both methods allow assimilation of all the data gathered throughout the observational period at once, it is possible to study the effect of observations collected at different time instances. The RM and AEnKF have been therefore chosen as a backbone for evaluating the data impact on the history matching. Based on the performed tests, the RM and AEnKF were found to utilize the data with similar effectiveness.
- The concept of an observation sensitivity matrix has been found applicable to quantify the impact of measured data on the outcome of the history matching process. This matrix describes the changes in measurements predicted by the model due to assimilation of the actual measurements. The indicators derived from the observation sensitivity matrix can be used to characterize the influence of the entire data set, each particular observation or any selected subset of data on the model update. Consequently, analysis of the impact of measurements of different types or originating from different locations becomes possible. In addition it has been proposed how to elicit the magnitude of a possible change in the accuracy of the updated model due to a change in the accuracy of collected observations from the norm of the sensitivity matrix.
- In the considered test cases the global averaged influence of the measurements was found to be only 4%. The global averaged influence of the prior to the analysis is therefore 96%. Hence the outcome of the history matching is mostly influenced by the prior model. This result is in line with practical experience. The findings indicate that assimilating production data only has almost no impact on the estimates of model parameters, and other types of data, e.g. seismic or gravity data, need to be included in the history matching exercises. Additionally, extra focus is needed on obtaining accurate prior models and ensuring proper uncertainty specification.

6.3 Recommendations

The proposed sensitivity matrix based approach to quantifying impact of measurements on the history matching procedure requires additional testing before potential deployment to industrial applications. The following directions for further investigations are recommended:

- In this study the method was implemented and tested on a reasonably simple and small-scale two-dimensional two-phase model. The feasibility of the computations carried out to obtain the sensitivity matrix needs to be verified for large-scale three-dimensional models. Ensemble Kalman filter based techniques are frequently used for history matching such models, as an adjoint is seldom available making variational data assimilation difficult. The users also often take advantage of the fact that the EnKF algorithm and its modifications are 'embarrassingly parallelizable', which yields considerable speed-up of the history matching process. To ensure that the cost of the observation sensitivity matrix computation does not diminish the positive effect of the parallelization, an approach to parallelizing the computation of the sensitivity matrix needs to be developed.
- This thesis considers matching the production history (bottom hole pressures and phase rates at the wells) directly, i.e. a so-called 'amplitude matching'. Such an approach is well known to result in a highly nonlinear inverse problem, often leading to an inadequate history match. The authors of [16] suggest lining up observed data and model predictions at some reference time, e.g. water breakthrough time. This approach, called 'generalized time inversion', has demonstrated superior behavior for the case studies in [16]. The generalized time inversion seems to offer a conceptually different treatment of the field data. The sensitivity matrix can be used as a tool to rigorously evaluate if this method leads to more efficient utilization of the available measurements and, hence, can become particularly beneficial for history matching scarce measurements.
- The global averaged influence (GAI) indicator can be used to compare the effectiveness of various measurement strategies. Both the GAI and the other indicators deduced from the observation sensitivity matrix can be fed to an optimization procedure to determine the optimal configuration of the measurement network. The network configuration is considered optimal in a sense that it provides the data assimilation algorithm with the most influential measurements, i.e. the data that can result in the largest contribution to the updated values of model param-

eters. It cannot be directly concluded to which degree this sophisticated choice of data contributes to the development of a better production strategy. The way forward could include identifying how to incorporate the relevant indicators in an objective function, such that the network configuration can be determined as a part of the optimal field production strategy within a production optimization loop. This approach is also expected to allow accounting for the costs related to the execution of the derived field surveillance program.

Bibliography

- [1] International energy outlook. Technical report, U.S. Energy Information Administration, 2011.
- [2] S.I. Aanonsen, G.Nævdal, D.S. Oliver, A.C. Reynolds, and B. Vallès. The ensemble Kalman filter in reservoir engineering — a review. *SPE J.*, 14(3):393–412, 2009. SPE-117274-PA.
- [3] S. Akella and I.M. Navon. Different approaches to model error formulation in 4D-Var: a study with high-resolution advection schemes. *Tellus A*, 61(1):112–128, 2009.
- [4] K. Aziz and A. Settari. *Petroleum Reservoir Simulation*. Applied Science Publishers LTD, London, 1979.
- [5] J. Baird and C. Dawson. The representer method for data assimilation in single-phase Darcy flow in porous media. *Computat. Geosci.*, 9(4):247–271, 2005.
- [6] J. Baird and C. Dawson. The representer method for two-phase flow in porous media. *Computat. Geosci.*, 11(3):235–248, 2007.
- [7] A.F. Bennett. *Inverse Modeling of the Ocean and Atmosphere*.
- [8] A.F. Bennett and P.C. McIntosh. Open ocean modeling as an inverse problem: Tidal theory. *J. Phys. Oceanogr.*, 12(10):1004–1018, 1982.
- [9] D.R. Brouwer and J.D. Jansen. Dynamic optimization of water flooding with smart wells using optimal control theory. *SPE J.*, 9(4):391–402, 2004.
- [10] G. Burgers, P. Leeuwen, and G. Evensen. Analysis scheme in the ensemble Kalman filter. *Mon. Weather Rev.*, 126:1719–1724, 1998.

- [11] C. Cardinali, S. Pezzulli, and E. Andersson. Influence-matrix diagnostic of a data assimilation system. *Q. J. R. Meteorol. Soc.*, 130:2767–2786, 2004.
- [12] M.A. Cardoso and L.J. Durlofsky. Linearized reduced-order models for subsurface flow simulation. *J. Comput. Phys.*, 229(3):681–700, 2010.
- [13] G. Chavent, M. Dupuy, and P. Lemonnier. History matching by use of optimal control theory. *SPE J.*, 15(1):74–86, 1975.
- [14] W.H. Chen, G.R. Gavalas, and M.L. Wasserman. A new algorithm for automatic history matching. *SPE J.*, 14(6):593–608, 1974.
- [15] Z. Chen, G. Huan, and Y. Ma. *Computational Methods for Multiphase Flows in Porous Media*. Society for Industrial and Applied Mathematics, Philadelphia, 2006.
- [16] H. Cheng, A. Datta-Gupta, and Z. He. A comparison of travel-time and amplitude matching for field-scale production-data integration: Sensitivity, nonlinearity, and practical implications. *SPE J.*, 10(1):75–90, 2005. SPE-84570-PA.
- [17] P. Courtier and O. Talagrand. Variational assimilation of meteorological observations with the adjoint vorticity equation. II: Numerical results. *Q. J. R. Meteorol. Soc.*, 113:1329–1347, 1987.
- [18] A. Datta-Gupta, D.W. Vasco, and J.C.S. Jong. On the sensitivity and spatial resolution of transient pressure and tracer data for heterogeneity characterization. *SPE Formation Eval.*, 12(2):137–144, 1997.
- [19] M.W. Davis. Production of conditional simulations via the LU triangular decomposition of the covariance matrix. *Math. Geol.*, 19:91–98, 1987.
- [20] R.M. Errico. Interpretations of an adjoint-derived observational impact measure. *Tellus A*, 59(2):273–276, 2007.
- [21] T. Ertekin, J.H. Abou-Kassen, and G.R. King. *Basic Applied Reservoir Simulation*. Society of Petroleum Engineers, Richardson, 2001.
- [22] G. Evensen. Sequential data assimilation with a non-linear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, 99(C5):10143–10162, 1994.
- [23] G. Evensen. The ensemble Kalman filter: theoretical formulation and practical implementation. *Ocean Dynam.*, 53(4):343–367, 2003.

- [24] G. Evensen. *Data Assimilation: The Ensemble Kalman Filter*. Springer, Berlin, 2nd edition, 2009.
- [25] R. Gelaro, R.H. Langland, G.D. Rohaly, and T.E. Rosmond. An assessment of the singular-vector approach to targeted observing using the FASTEX dataset. *Q. J. R. Meteorol. Soc.*, 125(561):3299–3327, 1999.
- [26] R. Gelaro, Y. Zhu, and R.M. Errico. Examination of various-order adjoint-based approximations of observation impact. *Meteorologische Zeitschrift*, 16(6):685–692, 2007.
- [27] G.H. Golub and C.F. van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore and London, 1996.
- [28] Y. Gu and D.S. Oliver. History matching of the PUNQ-S3 reservoir model using the ensemble Kalman filter. *SPE J.*, 10(2):217–224, 2005.
- [29] Y. Gu and D.S. Oliver. The ensemble Kalman filter for continuous updating of reservoir simulation models. *J. Energ. Resour. — ASME*, 128(1):79–87, 2006.
- [30] Y. Gu and D.S. Oliver. An iterative ensemble Kalman filter for multiphase fluid flow data assimilation. *SPE J.*, 12(4):438–446, 2007.
- [31] A. Gut. *Probability: A Graduate Course*. Springer, New York, 2005.
- [32] R.G. Hanea, J.K. Przybysz-Jarnut, M.V. Krymskaya, A.W. Heemink, and J.D. Jansen. The choice of a 'best' assisted history matching algorithm. In *SPE/EUROPEC EAGE Annual Conference and Exhibition*, SPE 131088, Barcelona, Spain, 2010.
- [33] D.C. Hoaglin and R.E. Welsch. The hat matrix in regression and ANOVA. *Am. Stat.*, 32(1):17–22, 1978.
- [34] L. Hogben. *Handbook of Linear Algebra*. Chapman and Hall / CRC, 2007.
- [35] M.A. Iglesias and C. Dawson. An iterative representer-based scheme for data inversion in reservoir modeling. *Inverse Probl.*, 25(3), 2009.
- [36] B. Jafarpour and D. B. McLaughlin. History matching with an ensemble Kalman filter and discrete cosine parametrization. In *SPE Annual Technical Conference and Exhibition*, SPE 108761, Anaheim, California, USA, 2007.

- [37] J.D. Jansen, O.H. Bosgra, and P.M.J. Van den Hof. Model-based control of multiphase flow in subsurface oil reservoirs. *J. Process Contr.*, 18: 846–855, 2008.
- [38] J.D. Jansen, S.D. Douma, D.R. Brouwer, P.M.J. Van den Hof, O.H. Bosgra, and A.W. Heemink. Closed-loop reservoir management. In *SPE Reservoir Simulation Symposium*, SPE 119098, The Woodlands, Texas, USA, 2009.
- [39] A.H. Jazwinski. *Stochastic Processes and Filtering Theory*. Academic Press, New York and London, 1970.
- [40] M.P. Kaleta. *Model-Reduced Gradient-Based History Matching*. PhD thesis, Delft University of Technology, 2011.
- [41] M.P. Kaleta, R.G. Hanea, A.W. Heemink, and J.D. Jansen. Model-reduced gradient-based history matching. *Computat. Geosci.*, 15(1):135–153, 2011.
- [42] G. Kerschen, J.-C. Golinval, A.F. Vakakis, and L.A. Bergman. The method of proper orthogonal decomposition for dynamical characterization and order reduction of mechanical systems: an overview. *Nonlinear Dynam.*, 41:147–169, 2005.
- [43] J.F.B.M. Kraaijevanger, P.J.P. Egberts, J.R. Valstar, and H.W. Buurman. Optimal waterflood design using the adjoint method. In *SPE Reservoir Simulation Symposium*, SPE 105764, Houston, Texas, USA, 2005.
- [44] M. Krymskaya, R. Hanea, and M. Verlaan. An iterative ensemble Kalman filter for reservoir engineering applications. *Computat. Geosci.*, 13, 2009.
- [45] M.V. Krymskaya, R.G. Hanea, J.D. Jansen, and A.W. Heemink. Observation sensitivity in computer-assisted history matching. In *SPE/EUROPEC EAGE Annual Conference and Exhibition*, Barcelona, Spain, 2010.
- [46] R.H. Langland and N.L. Baker. Estimation of observation impact using the NRL atmospheric variational data assimilation adjoint system. *Tellus A*, 56(3):189–201, 2004.
- [47] D.C. Lay. *Linear Algebra and Its Applications*. Pearson Education, 3rd edition, 2006.
- [48] R. Li, A.C. Reynolds, and D.S. Oliver. History matching of three-phase flow production data. *SPE J.*, 8(4):328–340, 2003.

- [49] J. Liu and E. Kalnay. Estimating observation impact without adjoint model in an ensemble Kalman filter. *Q. J. R. Meteorol. Soc.*, 134:1327–1335, 2008.
- [50] L. Ljung. *System Identification: Theory for the User*. Prentice Hall PTR, Upper Saddle River, New Jersey, 2nd edition, 1987.
- [51] 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. In *SPE Annual Technical Conference and Exhibition*, SPE 96375, Dallas, Texas, USA, 2005.
- [52] W. Menke. *Geophysical Data Analysis: Discrete Inverse Theory*. Academic Press, Inc., Orlando, Florida, 1984.
- [53] G. Nævdal, L.M. Johnsen, S.I. Aanonsen, and E.H. Vefring. Reservoir monitoring and continuous model updating using ensemble Kalman filter. *SPE J.*, 10(1):66–74, 2005. SPE-84372-PA.
- [54] G. De Nicolao. System identification: Problems and perspectives, tutorial plenary lecture. In *Proc. 11th Int. Workshop on Qualitative Reasoning*, pages 379–386, Cortona, Italy, 1997.
- [55] H. Nyquist. Certain topics in telegraph transmission theory. *P. IEEE*, 90(2):280–305, 2002.
- [56] D. S. Oliver and Y. Chen. Recent progress on reservoir history matching: a review. *Computat. Geosci.*, 15(1):185–221, 2011.
- [57] D.S. Oliver. Incorporation of transient pressure data into reservoir characterization. *In Situ*, 18(3):243–275, 1994.
- [58] D.S. Oliver, A.C. Reynolds, Z. Bi, and Y. Abacioglu. Integration of production data into reservoir models. *Pet. Geosci.*, 7(Sp. Iss. SI):S65–S73, 2001.
- [59] D.S. Oliver, A.C. Reynolds, and N. Liu. *Inverse theory for petroleum reservoir characterization and history matching*. Cambridge University Press, New York, 2008.
- [60] T.N. Palmer, R. Gelaro, J. Barkmeijer, and R. Buizza. Singular vectors, metrics, and adaptive observations. *J. Atmos. Sci.*, 55:633–653, 1998.
- [61] D. F. Parrish and J. C. Derber. The national meteorological center’s spectral statistical-interpolation analysis system. *Mon. Weather Rev.*, 120(8):1747–1763, 1992.

- [62] P. Prunet, J.-F. Minster, D. Ruiz-Pino, and I. Dadou. Assimilation of surface data in a one-dimensional physical-biogeochemical model of the surface ocean. 1. Method and preliminary results. *Global Biogeochem Cy.*, 10(1):111–138, 1996.
- [63] J.K. Przybysz-Jarnut. *Hydrocarbon Reservoir Parameter Estimation Using Production Data and Time-Lapse Seismic*. PhD thesis, Delft University of Technology, 2010.
- [64] 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. *Computat. Geosci.*, 11(1):73–85, 2007.
- [65] F. Rabier, N. Fourrié, D. Chafaï, and P. Prunet. Channel selection methods for infrared atmospheric sounding interferometer radiances. *Q. J. R. Meteorol. Soc.*, 128:1011–1027, 2002.
- [66] W.F. Ramirez. *Applications of Optimal Control Theory to Enhanced Oil Recovery*. Elsevier, Amsterdam, 1987.
- [67] A.C. Reynolds, M. Zafari, and G. Li. Iterative forms of the ensemble Kalman filter. In *Proc. of the 10th European Conference on the Mathematics of Oil Recovery*, Amsterdam, The Netherlands, 2006.
- [68] C.D. Rodgers. *Inverse Methods for Atmospheric Sounding: Theory and Practice*. World Scientific Publishing, Singapore, 2000.
- [69] S. Roman. *Advanced Linear Algebra*. Springer, Berlin, 3rd edition, 2007.
- [70] J.R. Rommelse, J.D. Jansen, and A.W. Heemink. An efficient weak-constraint gradient-based parameter estimation algorithm using representer expansions. *SPE J.*, 15(1):18–30, 2010.
- [71] T. Rosmond and L. Xu. Development of NAVDAS-AR: non-linear formulation and outer loop tests. *Tellus A*, 58(1):45–58, 2006.
- [72] P. Sakov, G. Evensen, and L. Bertino. Asynchronous data assimilation with the EnKF. *Tellus A*, 62(1):24–29, 2010.
- [73] P. Sarma, K. Aziz, and L. Durlofsky. Implementation of adjoint solution for optimal control of smart wells. In *SPE Reservoir Simulation Symposium*, SPE 92864, Houston, Texas, USA, 2005.
- [74] P. Sarma, L. Durlofsky, and K. Aziz. Kernel principal component analysis for efficient differentiable parameterization of multipoint geostatistics. *Math. Geosci.*, 40(1):3–32, 2008.

- [75] C.E. Shannon. Communication in the presence of noise. *P. IEEE*, 86(2): 447–457, 1998.
- [76] D. Simon. Kalman filtering. *Embedded Systems Programming*, 14(6): 72–79, 2001.
- [77] O. Talagrand and P. Courtier. Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory. *Q. J. R. Meteorol. Soc.*, 113:1311–1328, 1987.
- [78] A. Tarantola. *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation*. Elsevier, Amsterdam, 1987.
- [79] R. Tavakoli and A. C. Reynolds. History matching with parameterization based on the SVD of a dimensionless sensitivity matrix. *SPE J.*, 15(2): 495–508, 2011.
- [80] Y. Trémolet. Accounting for an imperfect model in 4D-Var. *Q. J. Roy. Meteor. Soc.*, 132(621):2483–2504, 2006.
- [81] P.M.J. van den Hof. System identification. TU Delft, 2001. Lecture notes for course TN3111.
- [82] J. F. M. Van Doren, S. G. Douma, P. M. J. Van den Hof, J. D. Jansen, and O. H. Bosgra. Identifiability: from qualitative analysis to model structure approximation. In *Prepr. 15th IFAC Symposium on System Identification*, pages 664–669, St. Malo, France, 2009.
- [83] J.F.M. van Doren. *Model Structure Analysis for Model-based Operation of Petroleum Reservoirs*. PhD thesis, Delft University of Technology, 2010.
- [84] J.F.M. van Doren, R. Markovinović, and J.D. Jansen. Reduced-order optimal control of water flooding using proper orthogonal decomposition. *Computat. Geosci.*, 10(1):137–158, 2006.
- [85] S. Vlemmix, G.J.P. Joosten, D.R. Brouwer, and J.D. Jansen. Adjoint-based well-trajectory optimization in a thin oil rim. In *SPE/EUROPEC EAGE Annual Conference and Exhibition*, SPE 121891, Amsterdam, The Netherlands, 2009.
- [86] X.-H. Wen and W. C. Chen. Real-time reservoir model updating using ensemble Kalman filter with confirming option. *SPE J.*, 11(4):431–442, 2006.
- [87] L. Xu and R. Daley. Towards a true 4-dimensional data assimilation algorithm: application of a cycling representer algorithm to a simple transport problem. *Tellus A*, 52(2):109–128, 2000.

-
- [88] L. Xu and R. Daley. Data assimilation with a barotropically unstable shallow water system using representer algorithms. *Tellus A*, 54(2):125–137, 2002.
 - [89] L. Xu, T. Rosmond, and R. Daley. Development of NAVDAS-AR: formulation and initial tests of the linear problem. *Tellus A*, 57(4):546–559, 2005.
 - [90] M. Zafari and A. C. Reynolds. Assessing the uncertainty in reservoir description and performance predictions with the ensemble Kalman filter. In *SPE Annual Technical Conference and Exhibition*, SPE 95750, Dallas, Texas, USA, 2005.
 - [91] M.J. Zandvliet, M. Handels, G.M. van Essen, D.R. Brouwer, and J.D. Jansen. Adjoint-based well-placement optimization under production constraints. *SPE J.*, 13(4):392–399, 2008. SPE-105797-PA.
 - [92] F. Zhang, A.C. Reynolds, and D.S. Oliver. Evaluation of the reduction in uncertainty obtained by conditioning a 3D stochastic channel to multiwell pressure data. *Math. Geol.*, 34(6):715–742, 2002.

Fluid-rock properties

A.1 Capillary pressure

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:

$$p_c = p_o - p_w = p_c(S_w).$$

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

A.2 Relative permeability

If two or more phases are saturating the porous medium the reservoir capacity to transmit any particular phase i is called the effective permeability to that phase and denoted by k_i . This definition yields that the effective permeability k_i is not greater than the absolute permeability k of the porous medium. The relative permeability k_{ri} to phase i indicates the tendency of the other phases to block the flow of phase i :

$$k_i = k_{ri}k.$$

Typical relative permeability curves for an oil-water system with water displacing oil are presented in Figure A.1.

The value of S_w at which water starts to flow is termed the connate water saturation S_{wc} and the value at which oil starts to flow — the residual oil saturation S_{or} . Considering the relative permeability to phase i at its critical saturation of the other phase (i.e. the saturation at which it starts to flow) we end up with the notion of the end point permeability k_{ri}^0 .

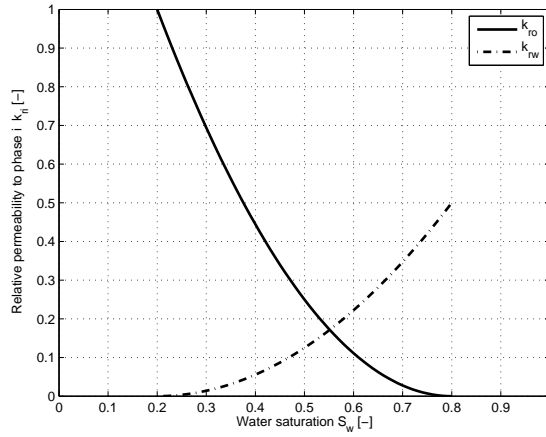


Figure A.1: Typical relative permeability curves

A.3 Corey-type two-phase relative permeability model

Although the values of relative permeability for each particular porous medium are subject to experimental determination, there exist mathematical models to describe a relationship between relative permeability and saturation of the wetting phase. In this study the Corey-type model is used.

Define the normalized saturation value as

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

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

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

where n_w and n_o are Corey exponents. In general these quantities can be obtained from measured data, but in this thesis the simplification $n_w = n_o = 2$ is used.

Mathematical background

B.1 Condition number of a matrix

Consider a system of linear equations $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{A} \in \mathbb{R}^{n \times n}$ is a system matrix, $\mathbf{x} \in \mathbb{R}^{n \times 1}$ stands for a vector of unknowns and $\mathbf{b} \in \mathbb{R}^{n \times 1}$ is a known vector at the right-hand side. A slight change in the right-hand side of the equation to e.g. $\mathbf{b} + \Delta\mathbf{b}$ for some vector $\Delta\mathbf{b}$ triggers a change of the solution to $\mathbf{x} + \Delta\mathbf{x}$, where $\Delta\mathbf{x}$ satisfies $\mathbf{A}(\Delta\mathbf{x}) = \Delta\mathbf{b}$.

The relative change in \mathbf{b} (or the relative error in \mathbf{b} if $\Delta\mathbf{b}$ represents the possible error in the entries of \mathbf{b}) is given by $\frac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|}$. Similarly, the relative change in the solution is $\frac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}\|}$. If \mathbf{A} is invertible, the condition number of \mathbf{A} , written as $\text{cond}(\mathbf{A})$, provides a bound on how large the relative change in \mathbf{x} can be [27]

$$\frac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}\|} \leq \text{cond}(\mathbf{A}) \cdot \frac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|}, \quad (\text{B.1})$$

where

$$\text{cond}(\mathbf{A}) = \|\mathbf{A}\| \cdot \|\mathbf{A}^{-1}\|. \quad (\text{B.2})$$

Hence, if the condition number is large, even a small error in vector \mathbf{b} may cause a large error in solution \mathbf{x} . To the contrary, if the condition number is relatively small, the error in the solution \mathbf{x} will be comparable to the error introduced into the vector on the right-hand side.

B.2 Vector and matrix norms

A vector norm on \mathbb{R}^n is a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ that satisfies the following properties [27]:

- $f(\mathbf{x}) \geq 0$, $\mathbf{x} \in \mathbb{R}^n$, ($f(\mathbf{x}) = 0$ if and only if $\mathbf{x} = \mathbf{0}$),
- $f(\mathbf{x} + \mathbf{y}) \geq f(\mathbf{x}) + f(\mathbf{y})$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,
- $f(\alpha\mathbf{x}) = |\alpha|f(\mathbf{x})$, $\alpha \in \mathbb{R}$, $\mathbf{x} \in \mathbb{R}^n$.

Such a function is usually denoted with a double bar: $f(\mathbf{x}) = \|\mathbf{x}\|$. One of the most commonly used vector norms is the so-called Euclidean norm or 2-norm:

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^n x_i^2}, \text{ where } x_i \text{ is the } i\text{th coordinate of the vector } \mathbf{x}.$$

The definition of a matrix norm is similar to the definition of a vector norm. In particular, $f : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}$ is a matrix norm if the following properties hold [27]:

- $f(\mathbf{A}) \geq 0$, $\mathbf{A} \in \mathbb{R}^{m \times n}$, ($f(\mathbf{A}) = 0$ if and only if $\mathbf{A} = \mathbf{0}$),
- $f(\mathbf{A} + \mathbf{B}) \geq f(\mathbf{A}) + f(\mathbf{B})$, $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}$,
- $f(\alpha\mathbf{A}) = |\alpha|f(\mathbf{A})$, $\alpha \in \mathbb{R}$, $\mathbf{A} \in \mathbb{R}^{m \times n}$.

As with vector norms, a double bar notation is used to designate matrix norms, i.e. $\|\mathbf{A}\| = f(\mathbf{A})$.

If $\|\cdot\|$ is a family of vector norms on \mathbb{R}^n , then the matrix norm on $\mathbb{R}^{m \times n}$ can be considered in connection to $\|\cdot\|$ [34]:

$$\|\mathbf{A}\| = \max_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|}.$$

One of the most frequently used matrix norms is a spectral norm or 2-norm: $\|\mathbf{A}\| = \sqrt{\lambda(\mathbf{A}^T \mathbf{A})}$, where $\lambda(\mathbf{A}^T \mathbf{A})$ is the largest eigenvalue of $\mathbf{A}^T \mathbf{A}$. This norm has the following extra properties:

- $\|\mathbf{A}\mathbf{x}\|_2 \leq \|\mathbf{A}\|_2 \cdot \|\mathbf{x}\|_2$ for $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{x} \in \mathbb{R}^n$,
- $\|\mathbf{A}\mathbf{B}\|_2 \leq \|\mathbf{A}\|_2 \cdot \|\mathbf{B}\|_2$ for $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$.

B.3 Singular value decomposition of a matrix

A singular value decomposition (SVD) can be performed for any real matrix $\mathbf{A} \in \mathbb{R}^{n \times m}$. The SVD factorizes the matrix such that [27]

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T,$$

where $\mathbf{U} \in \mathbb{R}^{m \times m}$ is an orthogonal matrix containing the left singular vectors, $\mathbf{\Sigma} \in \mathbb{R}^{m \times n}$ is a pseudo-diagonal and positive semidefinite matrix with

diagonal entries containing the singular values σ_i , $i = 1, \dots, \min(m, n)$ and $\mathbf{V} \in \mathbb{R}^{n \times n}$ is an orthonormal matrix containing the right singular vectors. A common convention is to order the singular values σ_i , $i = 1, \dots, \min(m, n)$, in descending order.

If matrix \mathbf{A} is viewed as a standard matrix of a linear operator T that maps vectors from \mathbb{R}^m to \mathbb{R}^n , then the geometrical interpretation of the SVD is as follows. One can find orthonormal bases of \mathbb{R}^m and \mathbb{R}^n such that T maps the basis vectors of \mathbb{R}^m into basis vectors of \mathbb{R}^n . These bases contain left and right singular vectors of \mathbf{A} . With respect to the new bases, transformation T is represented by a diagonal matrix with non-negative real diagonal entries.

One can also interpret the outcome of SVD in terms of oriented 'energy' [42]. Note that the term 'energy' is used in a somewhat loose sense. In this framework the square of singular value σ_i , $i = 1, \dots, \min(m, n)$, describes the oriented 'energy' in the direction of the i^{th} left singular vector. This explains the extensive use of SVD in applications that involve model reduction algorithms (see e.g. [12, 41, 84] for illustrations of model reduction performed for different types of reservoir engineering problems).

When the matrix \mathbf{A} is square, i.e. when it corresponds to a transformation T from one vector space onto itself, the eigenvalues and eigenvectors lead to a more straightforward interpretation. The transformation T can then be viewed as a composition of three actions: rotation, scaling and another rotation. However for such a square matrix \mathbf{A} an eigenvalue decomposition does not necessarily exist, although it does exist if \mathbf{A} is symmetric. Moreover if the matrix \mathbf{A} is positive semidefinite, the eigenvalues of $\mathbf{A}\mathbf{A}^T$ and of $\mathbf{A}^T\mathbf{A}$ coincide with the singular values of \mathbf{A} squared. In such a case the eigenvectors (singular vectors) of the matrix are the principal axes of the associated quadratic form which is an n -dimensional ellipsoid centered at the origin of Euclidean space. In other words, the eigenvectors (singular vectors) indicate directions of the shrinking or expansion of the unit sphere ordered by the magnitude of the effect provided by the transformation [47].

Summary

Global energy use is increasing. As societies advance, they will continue to need energy to power residential and commercial buildings, in the industrial sector, for transportation and other vital services. To satisfy this rising demand, liquid, natural gas, coal, nuclear power and renewable fuel sources are extensively developed. Particularly fossil fuels (i.e. oil, natural gas and coal) remain the largest source of energy for the world. Petroleum exploration and production companies continuously develop new and enhance current production technologies to increase recovery from the existing fields. These companies rely on various tools to support their production and development decisions. Reservoir modeling is a standard tool used in the decision making process allowing analysis and prediction of the reservoir flow behavior, identification of beneficial production strategies and evaluation of the associated risks.

The models used for reservoir simulation contain a large number of imperfectly known parameters characterizing the reservoir flow, e.g. permeability and porosity of the reservoir rock. Therefore the predictive value of such models is limited and tends to deteriorate in time. History matching is employed to update the values of poorly known model parameters in time with the help of the production data which become available during the production life of the reservoir, i.e. to adapt parameters such that simulated results are consistent with measured production data. Such an approach generally improves estimates of the model parameters and the predictive capability of the model.

Remarkably, the information extracted from the measurements in the history matching phase is repeatedly found as not enough to provide well-calibrated model with a high predictive value. Hence, consideration of additional data can be of particular help. To optimize the costs and effort associated with collection of new data and computations, up-front selection of the most influential measurements and their locations is desirable. Meth-

ods to assess the impact of measurements on model parameter updating are therefore needed. The research objective of this thesis was to develop efficient tools for quantifying the impact of measured data on the outcome of history matching of reservoir models, i.e. tools that provide a meaningful quantification of the impact of observations, while requiring limited time and effort to be incorporated in the history matching algorithms.

This research addressed history matching two-dimensional two-phase reservoir model representing water flood with production data (bottom hole pressure at injection well and oil and water flow rates at production wells).

First, the applicability and implementation of a number of history matching algorithms were investigated. The representer method (RM) has been considered as an example of variational techniques. The algorithm's key feature is the computation of a set of so-called representer describing the influence of a certain measurement on an estimation of the state and/or parameter. The RM was found to provide a reasonable parameter estimate, although it is computationally inefficient for dealing with large data sets. This fact yielded testing of the accelerated representer method (ARM), where direct computation of representer is avoided. The results indicate that the accuracy of the ARM can be controlled to provide an outcome of the same accuracy as the RM, and that the ARM outperforms the classical RM in terms of computational speed when the number of assimilated measurements increases. In this thesis we developed a strategy to evaluate the number of operations performed by the methods to assess the amount of data for which the ARM becomes beneficial to use.

The RM and the ARM require the model adjoint and are not intended for continuous (sequential) history matching, namely for incorporating obtained data in the model on the fly. Instead they perform history matching over a rather long time window using all available observations.

The ensemble Kalman filter (EnKF) has been discussed as it is the algorithm for continuous history matching. The EnKF schemes do not require the model adjoint, which makes them very attractive for data assimilation with complex non-linear models. The use of the EnKF in reservoir engineering however is prone to producing physically unreasonable values of the state variables. The problem can be overcome by including a so-called confirmation step in the algorithm. The EnKF, particularly with a confirmation step, is often computationally demanding for large-scale applications. The asynchronous EnKF (AEnKF) is a modification of the EnKF which offers a practical way to perform history matching in such cases by updating the system with batches of measurements collected at the times different to the time of the update. Hence, all observations collected during a certain time-window can be history-matched at once at the end of observational period. This allows for comparison of the influence of the observations collected at different

times. Furthermore, it does not rely on an adjoint model, though it resembles the approach usually followed in variational methods. Both the EnKF and the AEnKF demonstrated considerable improvement of the model parameter estimates compared to the prior and gave acceptable history matches.

Since the AEnKF allows for history matching all the data gathered throughout the observational period at once, it permits comparison of the effect of observations collected at different time instances. The equivalence of the AEnKF to variational techniques (e.g. the RM) yields the possibility to evaluate if ensemble Kalman filtering and variational methods utilize the observations in a similar manner. The representer method and the AEnKF were selected to be used as platforms for quantification of the measurements impact on history matching.

Secondly, in this thesis we developed a tool to quantify the impact of measured data on the outcome of history matching. The method has been inspired by the recent advancements in meteorology and oceanography, and is based on a so-called sensitivity matrix. This matrix can be used to evaluate the amount of information extracted from available data during the data assimilation phase and identify the observations that have contributed to the parameter update the most. In particular, we used the diagonal elements of the matrix, known as self-sensitivities, as a quantitative measure of the influence of observed measurements on predicted measurements. Additionally, we have proposed a way to use the norm of the sensitivity matrix for assessing the magnitude of possible change in the accuracy of the model due to the respective change in the accuracy of collected observations. The observation sensitivity matrix is fast and easy to compute both for adjoint-based and EnKF types of history matching algorithms. The analysis performed with the aid of the observation sensitivity matrix has confirmed that the RM and the AEnKF utilize the data with comparable effectiveness. Remarkably, for a simple test case the global averaged influence of the observed measurements is only 4%. This is a rather low value compared to the 96% global averaged influence of the prior. The observation sensitivity matrix can be also used to investigate the dependency between the measurement location / type and its importance to history matching.

Samenvatting

Het wereldwijde energieverbruik neemt gestaag toe. Als maatschappijen zich blijven ontwikkelen, zal de behoefte aan energie groeien om woon- en bedrijfsgebouwen van stroom te kunnen voorzien, voor in de industrie, voor in het vervoer en voor in andere onmisbare voorzieningen. Om aan deze toenemende eisen te kunnen voldoen, worden de bronnen van aardgas, steenkool, kern-energie, en vernieuwbare brandstoffen op grote schaal geproduceerd. Daarbij blijven de fossiele brandstoffen voorlopig de grootste energiebron ter wereld. Aardolie exploratie- en productiemaatschappijen zijn voortdurend bezig met het ontwikkelen van de nieuwe en het verbeteren van de huidige productie-technnologiën om olie- en gaswinning uit bestaande velden te verhogen. Deze bedrijven gebruiken daarbij diverse middelen voor het ondersteunen van hun productie- en veldontwikkelingsplannen. Het modelleren van reservoirs is een veelgebruikt middel in de besluitvorming, omdat het onderzoek en voorspelling van het gedrag van de stroming in reservoir, identificatie van een voordelig productiebeleid, en bepaling van verbonden risico's mogelijk maakt.

Modellen, die voor reservoirsimulatie worden gebruikt, bevatten een groot aantal onbekende parameters (b.v. de permeabiliteit en de porositeit van het reservoirgesteente) voor het karakteriseren van de stroming in het reservoir. Daarom hebben deze modellen een beperkte voorspellende waarde en zijn ze geneigd op den duur in kwaliteit achteruit te gaan. "History matching" is een techniek die wordt gebruikt om de waarden van de onvoldoende bekende parameters aan te passen met behulp van productiedata, die in de loop van de reservoirproductie zijn verkregen. Daarbij worden de parameters zodanig aangepast dat het resultaat van de simulatie in de vorm van voorspelde productiedata overeenkomt met de gemeten productiedata. In het algemeen verbetert een dergelijke aanpak de schatting van de modelparameters en het voorspellende vermogen van het model.

Dikwijls is de informatie die uit de metingen in de "history matching" fase

is gehaald, niet voldoende om tot een goed aangepast model met een hoge voorspellende waarde te leiden. In dat geval kan het meenemen van aanvullende data van bijzonder belang zijn. Om de kosten en de moeite verbonden aan de verzameling van de nieuwe data en de berekening daarvan te optimaliseren, wenst men de meest invloedrijke metingen en hun ligging vooraf te kunnen bepalen. Er zijn dus methoden nodig voor het schatten van de invloed van de nieuwe metingen op de aanpassingen van de modelparameters. Het doel van dit proefschrift was om efficiënte methodes te ontwikkelen voor de kwantificatie van de invloed van metingen op het resultaat van de "history matching" van reservoirmodellen. In andere woorden, methodes die een zinvol getal opleveren terwijl ze in een beperkte tijd en met beperkte moeite kunnen worden berekend.

Dit onderzoek richt zich op "history matching" van twee-dimensionale twee-fase (olie-water) reservoirmodellen met behulp van productiedata (d.w.z. drukmetingen in de injectieputten en debietmetingen in de productieputten).

Ten eerste werden de bruikbaarheid en de implementatie van aantal bestaande "history matching" algoritmen onderzocht. De "representer method" (RM) is als een voorbeeld van een variationele techniek in overweging genomen. Het belangrijkste aspect van het algoritme is berekening van een aantal zogenaamde "representers" die de invloed van bepaalde metingen op de schatting van de modeltoestand en/of modelparameter beschrijven. De RM is in staat om een redelijk parameterschatting te kunnen maken, maar is niet efficiënt voor grote hoeveelheden metingen. Dat feit leidde tot het testen van "accelerated representer method" (ARM), waarbij direct berekening van de representers wordt vermeden. Het onderzoek toont aan dat de vereiste nauwkeurigheid van de ARM zodanig kan worden gekozen dat de resultaten van dezelfde nauwkeurigheid zijn als die van de RM, en dat de ARM de klassieke RM op gebied van rekensnelheid overtreft als de hoeveelheid van de geassimileerde metingen toeneemt. In dit proefschrift ontwikkelden wij een strategie om het aantal van de voor elke methode benodigde rekenoperaties te bepalen, zodat de datahoeveelheid waarvoor het gebruik van de ARM voordelig wordt vooraf kan worden geschat.

De RM en de ARM hebben het geadjungeerde model van het reservoir-model nodig en zijn niet bedoeld voor "continuous history matching", d.w.z. voor het onmiddellijk verwerken van data in het model zodra ze beschikbaar komen. Integendeel, ze voeren "history matching" doorgaans uit over een lang tijdsinterval met behulp van alle beschikbare metingen.

Het ensemble Kalman filter (EnKF) is beschouwd, omdat het een algoritme voor "continuous history matching" is. De EnKF methodes hebben geen geadjungeerd model nodig, en zijn om die reden aantrekkelijk voor dataassimilatie met gecompliceerde niet-lineaire modellen. Het gebruik van het EnKF in de reservoirtechniek is echter geneigd om te leiden tot fysisch onrealistisch waar-

den van de modeltoestand. Dat probleem kan worden opgelost door toevoeging van een zogenaamde "confirmatie stap" aan het algoritme. Het EnKF, vooral me de confirmatie stap, is vaak rekentechnisch veeleisend voor grootschalige toepassingen. Het asynchrone EnKF (AEnKF) is een variant van het EnKF, die in zulke gevallen tot een praktische methode voor "history matching" leidt door het systeem aan te passen met behulp van metingen verzameld op de tijdstippen verschillend van het tijdstip waarop de aanpassingen worden gemaakt. Daartoe worden alle metingen, verzameld binnen een bepaald tijdsinterval, eenmalig aan het einde van het meetinterval gebruikt voor de "history match". Dat maakt het mogelijk om het invloed van de op verschillende tijdstippen verzamelde metingen te vergelijken. Bovendien heeft het AEnKF geen geadjungeerd model nodig, ondanks dat het de reguliere aanpak van de variationele methoden lijkt te volgen. Zowel het EnKF als het AEnKF resulteren, in vergelijking met het beginmodel, in aanzienlijke verbeteringen van de schatting van de modelparameters en een redelijk "history match".

Aangezien het AEnKF "history matching" van de alle gedurende de meetperiode verzamelde data in één keer mogelijk maakt, laat het een vergelijking toe van het effect van de op verschillende tijdstippen verzamelde metingen. Gelijkwaardigheid van het AEnKF en de variationele technieken (b.v. de RM) levert de mogelijkheid op om te bepalen of het EnKF en de variationele methoden hetzelfde gebruik van de metingen maken. De RM en het AEnKF zijn daarom als methodes gekozen voor het kwantificeren van de invloed van metingen op de "history match".

Ten tweede ontwikkelden wij in dit proefschrift een middel om de invloed van de gemeten data op het resultaat van de "history match" te bepalen. Deze methode is geïnspireerd door recente ontwikkelingen in de meteorologie en de oceanografie, en is gebaseerd op de zogenaamde "gevoeligheidsmatrix". Deze matrix kan worden gebruikt om de hoeveelheid van de informatie te bepalen die uit beschikbare data verkregen is, en om metingen te identificeren, die in de grootste mate aan parameteraanpassingen bijdroegen. Wij gebruikten met name diagonaalelementen van de matrix, bekend onder de naam "zelfgevoeligheden", als een kwantitatieve maat voor de invloed van de verzamelde metingen op de door het aangepaste model voorspelde metingen. Bovendien hebben wij een manier voorgesteld om de norm van de gevoeligheidsmatrix te gebruiken voor het inschatten van de grootte van mogelijke verandering in de nauwkeurigheid van het aangepaste model door de respectieve verandering in de nauwkeurigheid van de verzamelde metingen. De gevoeligheidsmatrix kan snel en eenvoudig berekend worden voor zowel geadjungeerd-gebaseerd als EnKF-achtige "history matching" algoritmen. Het onderzoek gedaan met behulp van de gevoeligheidsmatrix had bevestigd dat de RM en het EnKF met dezelfde effectiviteit gebruik maken van de data. Het is opmerkelijk dat, voor een eenvoudige test-situatie, de globale gemiddelde invloed van de verzamelde

metingen maar 4% is. Dat is een relatief kleine waarde vergeleken met de 96% globale gemiddelde invloed van het uitgenagsmodel. De gevoeligheidsmatrix kan ook worden gebruikt voor onderzoek naar het verband tussen de plaats en/of het type van een meting en zijn waarde voor de "history match".

Acknowledgments

It has been a long journey to this thesis, often hard or simply impossible without active support, contribution and encouragement of many people. I am grateful to all of you.

I would like to acknowledge invaluable input, guidance and high standards of professional excellence of my supervisors Arnold Heemink and Jan Dirk Jansen. I truly enjoyed working with Arnold, who oversees the whole research process and beyond, and allows enough research freedom for the student. I appreciate constructive criticism of Jan Dirk, his efforts to improve my command of English, and help with translating propositions and summary to Dutch.

Many heartfelt thanks go to my daily supervisor Remus Hanea who makes sure to take care not only of the day-to-day research activities, but also of the general well-being of a student. I am honored I can call Remus and his extended family my personal friends.

I am indebted to Arnold, Jan Dirk and Remus for particularly their patience, understanding and setting aside (often private) time for me in the extended last phase of the thesis preparation.

I would like to express my gratitude to Bjarne Foss for finding time to travel all the way from Norway to take part in the committee. I thank Paul van den Hof for introducing me to systems and control. Kees Vuik and Hai Xiang Lin have equipped me with fundamentals of applied mathematics during MSc courses they taught at Delft University of Technology (TU Delft). I am grateful to the committee members for providing a thorough feedback.

It has been a great pleasure to share this research experience with fellow colleagues at TU Delft. My officemates Wiktoria, Svetlana and Cristi always created enjoyable atmosphere in the office. I appreciate the exchange of ideas over many liters of coffee consumed with Alex Chernetsky, Alex Kirichek, Alex Pischansky, Umer, Julius, Eka, Niels, Tina, Birgit and Marcin. We had a good time at the workshops together with Ivan, Gosia, Justyna, Asia and Maryam.

I thank Christiaan for endless lunch discussions and many practical tips.

Completing a PhD also involves a lot of administration and seeming technicalities, which could have been overwhelming without the help of Evelyn, Dorothee, Kees Lemmens, Anke and Marlijn.

Doing this research has become possible thanks to all the prior education I obtained. I acknowledge the professors and the staff of Ufa State Aviation Technical University, Russia, in which I spent four years as a bachelor student, and these of TU Delft. I am indebted to Prof. R.M. Cooke for giving me an opportunity to follow a master program at TU Delft.

Last phase of the research had to be combined with a full-time job, which introduces its challenges. I am grateful to all colleagues at Shell for making me feel warmly welcome to the team, sharing professional expertise, passion to modeling and life experiences.

Volleyball has been a serious hobby for me for more than a decade by now and it was great to share the team spirit, wins and losses with the teams of PUNCH and Sovicos clubs. My teammates, trainers and coaches have often gone beyond just volleyball helping with various matters and making sure I am well-integrated into society. I thank Nicolien for her support through the most of my PhD time. I am particularly grateful to Viannette, Yvonne and Willemien for friendship and introducing me to the concept of 'spelletjesavond'.

We regularly have great fun and time together with my master study friends Karolina and Kasia. Hope we keep it this way.

My dear friends Iwona, Lidia and Lukasz and their partners have been always there for me in good and bad times. I am thrilled that despite time and some distances nothing has changed since I have been putting the same three names in the acknowledgements of my MSc thesis. When I started working and this thesis has been taking far too long, Iwona has invented a 'PhD-thesis-Sunday' concept, to make sure I spend at least every Sunday a week to finalize the thesis. Iwona and Lukasz, I am honored to have you as paranympths.

I am very lucky we keep our close friendship with my almost quarter-of-a-century friend Olga despite time and considerable distances. This friendship has become a friendship between families by now. I am thankful to Nikos for bringing a sense of adventure to my life, taking me to the great places, introducing me to the Greek cuisine and, last but not least, providing professional support to my PC.

I am grateful to the members of my extended family for their warmth, support and advice.

No words are good enough to express my gratitude to my parents. I am indebted to them for all the support, care, patience, endless love and belief in me. Thank you!

Curriculum Vitae

Mariya Victorovna Krymskaya was born on January 2, 1985, in Ufa, USSR (Russia). She completed her secondary education between 1991 and 2001 in Gymnasium No. 3 in Ufa, Russia. She has been awarded a Golden Medal of the Russian Ministry of Education for Excellence in Study.

From 2001 to 2005 she studied Mathematics and Informatics at the Ufa State Aviation Technical University in Ufa, Russia, from which she received a Bachelor of Science degree (with honors) in 2005. In the same year she started following a Master of Science program in Applied Mathematics with specialization in Probability, Risks and Statistics at Delft University of Technology, and obtained her Master of Science degree (with honors) in 2007. She has been awarded a UfD-Stud prize as the best graduate 2007 of the Faculty of Electrical Engineering, Mathematics and Computer Science at Delft University of Technology.

From September 2007 to August 2011 she continued her academic education in the group of Mathematical Physics, Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, as a PhD student under ISAPP.

In September 2011 she joined Shell Global Solutions International B.V. as a reservoir engineer.

