Continuous Updating of a Coupled Reservoir-Seismic Model Using an Ensemble Kalman Filter Technique

PhD Thesis

Jan-Arild Skjervheim

Department of Mathematics
University of Bergen
Norway

January 2007

ISBN

Abstract

This work presents the development of a method based on the ensemble Kalman filter (EnKF) for continuous reservoir model updating with respect to the combination of production data, 3D seismic data and time-lapse seismic data. The reservoir-seismic model system consists of a commercial reservoir simulator coupled to existing rock physics and seismic modelling software. The EnKF provides an ideal-setting for real time updating and prediction in reservoir simulation models, and has been applied to synthetic models and real field cases from the North Sea.

In the EnKF method, static parameters as the porosity and permeability, and dynamic variables, as fluid saturations and pressure, are updated in the reservoir model at each step data become available. In addition, we have updated a lithology parameter (clay ratio) which is linked to the rock physics model, and the fracture density in a synthetic fractured reservoir.

In the EnKF experiments we have assimilated various types of production and seismic data. Gas oil ratio (GOR), water cut (WCT) and bottom-hole pressure (BHP) are used in the data assimilation. Furthermore, inverted seismic data, such as Poisson's ratio and acoustic impedance, and seismic waveform data have been assimilated.

In reservoir applications seismic data may introduce a large amount of data in the assimilation schemes, and the computational time becomes expensive. In this project efficient EnKF schemes are used to handle such large datasets, where challenging aspects such as the inversion of a large covariance matrix and potential loss of rank are considered.

Time-lapse seismic data may be difficult to assimilate since they are time difference data, i.e. data which are related to the model variable at two or more time instances. Here we have presented a general sequential Bayesian formulation which incorporates time difference data, and we show that the posterior distribution includes both a filter and a smoother solution. Further, we show that when time difference data are used in the EnKF, a combination of the ensemble Kalman filter and the ensemble Kalman smoother has to be applied. However, the method is still completely recursive, with little additional cost compared to the

standard EnKF.

An underestimation of the uncertainty of the model variables may become a problem when few ensemble members and a large amount of data are used during the assimilation. Thus, to improve the reservoir model updating we have proposed a methodology based on a combination of a global and a local analysis scheme.

Preface

The dissertation is submitted as a partial fulfilment of the requirements for the degree of Doctor of Philosophy at the University of Bergen. The studies leading to the submission of the thesis started in February 2004, and have been done at the Department of Mathematics at the University of Bergen, and at the Centre for Integrated Petroleum Research.

Principal Researcher Professor II Sigurd Ivar Aanonsen has been the main adviser for the work, while Dr. Geir Evensen, Professor Tor Arne Johansen and Professor Magne Espedal have been co-advisers.

The thesis is organized in two parts. Part I focuses on the theoretical background and methodology used for updating of a coupled reservoir-seismic model. In Part II a set of research papers is presented. A brief outline of the thesis follows below:

Part I: Theoretical background

Chapter 1 gives an introduction to petroleum reservoir evaluation. Further, the main objectives of the research work are outlined, an overview of the papers and the results are presented.

In Chapter 2 the concept of reservoir characterization is presented. A general theory for inverse problems described in a Bayesian framework is outlined, and different sampling routines to explore the posterior distribution are discussed.

In Chapter 3 the inverse problem is introduced as a combined parameter and state estimation problem and formulated as a sequential data assimilation method using Bayesian statistics. Further, the Kalman filter and sampling of the Kalman filter distribution are discussed. Particularly the incorporation of time difference data are presented in detail.

Chapter 4 contains an overview of the ensemble Kalman filter methodology. A description of the ensemble Kalman filter, the ensemble Kalman smoother and the ensemble square root filter are presented, and a brief introduction to some other analysis schemes is given. Especially the challenging issue of dealing with large datasets involved in the assimilation are discussed, and a detailed explanation of

an efficient analysis scheme which handles such large amount of data is given.

Chapter 5 specifies different methods to generate Gaussian random fields which can be used in the sampling of ensemble realizations.

In Chapter 6 different challenges when incorporating seismic data in a data assimilation process are considered. It is pointed out that conditioning to seismic data requires a coupling method for combined modelling of reservoir fluid flow and seismic parameters. Seismic data may be available at different processing levels and may have different degrees of spatial resolution - this topic is discussed in detail. The procedure applied to incorporate time-lapse seismic data in the ensemble Kalman filter is explained thoroughly. Further, a brief description of the rock physics model and the seismic modelling approach is presented.

Part II: Papers and reports

Paper A: "Incorporating 4D Seismic Data in Reservoir Simulation Models Using Ensemble Kalman Filter". J. A. Skjervheim, G. Evensen, S. I. Aanonsen, B. O. Ruud and T. A. Johansen. Proceeding at the 2005 Annual Technical Conference and Exhibition held in Dallas, Texas, U.S.A, October 2005. Accepted for publication in the SPE Journal.

Paper B: "Ensemble Kalman filter with time difference data". J. A. Skjervheim, S. I. Aanonsen and G. Evensen. Submitted to Computational Geosciences, April 2006.

Paper C: "Combined inversion of 4D seismic waveform data and production data using ensemble Kalman filter". J. A. Skjervheim and B. O. Ruud. 76th Annual International Meeting, SEG, Expanded Abstracts, 1776-1780.

Paper D: "Using the Ensemble Kalman Filter with 4D Data to Estimate Properties and Lithology of Reservoir Rocks". J. A. Skjervheim, B. O. Ruud, S. I. Aanonsen, G. Evensen and T. A. Johansen. Proceeding at the 10^{th} European Conference on the Mathematics of Oil Recovery (ECMOR X), Amsterdam, The Netherlands, September 2006.

Paper E: "Estimating Lithology and other Reservoir Properties from 4D Seismic Data Using the Ensemble Kalman Filter". J. A. Skjervheim, S. I. Aanonsen, B.O. Ruud, G. Evensen and T. A. Johansen. To be submitted to the Journal of Petroleum Science and Engineering.

Paper F: "Characterization of fractured reservoirs by effective modelling and joint inversion of seismic and production data". M. Jakobsen J. A. Skjervheim

and S. I. Aanonsen. Presented at the 12th International Workshop on Seismic Anisotropy (12IWSA). Submitted to the Journal of Seismic Exploration.

Report 1: "Use of EnKF on a North Sea Field Case". J. A. Skjervheim, S. I. Aanonsen, M. Haverl and G. Evensen. Scientific/Technical Report.

Report 2: "Sequential Gaussian Simulation with Block Kriging". J. A. Skjervheim and S. I. Aanonsen. Scientific/Technical Report.

Acknowledgements

I want to express my sincere gratitude to all people I have worked with and who have assisted me during my study for the PhD degree.

I gratefully acknowledge my advisers Sigurd Ivar Aanonsen, Geir Evensen, Tor Arne Johansen and Magne Espedal. They have always taken their time to discuss the research challenges, and have shared their insight with me. Special thanks to Sigurd and Geir for having introduced the main principles of the inverse problem and the ensemble Kalman filter theory.

Further, I would like to thank Bent Ole Ruud and Morten Jakobsen for the guidance and contribution to this work.

A thanks to the people having assisted me with the proof reading of this thesis. Especially thanks to Geir Nævdal, Marco Haverl, David Moreno and Kristian Thulin for useful suggestions.

I will also thank my colleagues at Centre for Integrated Petroleum Research, at the Department of Mathematics at University of Bergen. They have been of great importance for making my PhD period into three good years.

I also owe many thanks to my ensemble Kalman filter colleagues at the International Research Institute of Stavanger, Hydro and Statoil, for valuable counsels and thorough feedback related to the work. My appreciation extend to the staff at Norsar for their contribution to the project.

Finally, I want to thank my family and friends, and especially my girlfriend Jofrid, for invaluable encouragement and support during my time as a PhD student.

Jan-Arild Skjervheim
January, 2007

Contents

Part I

General Background

Chapter 1

Introduction

In a petroleum reservoir evaluation the aim is to forecast the future production under different recovery strategies. The production forecast should be predicted with as little uncertainty as possible. Thus in a reservoir evaluation all the available information about the reservoir should be used in a consistent way.

Stochastic modelling is frequently used in reservoir characterization to quantify uncertainty and to integrate different types of information. The knowledge of the reservoir properties is initially very vague. Hence, a multivariate stochastic model should be used to describe the spatial variability of the reservoir properties such as porosity, permeability, lithology, initial reservoir pressure and fluid saturation. In addition, the stochastic model should represent the pressure and fluid saturation changes in accordance with the fluid flow model.

The fluid flow model is a forward function, defined from Darcy's law, and is represented by a numerical simulator, taking the reservoir characteristics and the recovery strategy as input. The fluid flow simulator describes the flow of water, oil, and gas in the reservoir and provides the behaviour of performance variables like bottom-hole pressure, gas-oil ratio, water-cut and cumulative oil production over time.

The information used in the stochastic model should include both the measurements from the reservoir and the geological knowledge. Reservoir measurements comprise production performance data, well log observations and seismic data achieved from the reservoir under production. The geological knowledge may consist of the knowledge from analog reservoir fields, well logs, core data, the evolution of the geological formations and the geological deposition. Further, important issues in a reservoir evaluation are the understanding of the procedures used in the data acquisition and the insight into the development of the physical properties and the geological model of the reservoir under study.

The geological model defines the reservoir in the geological context, and the geological model has several functions: It is utilized to design the geometry and

trap mechanisms and to unravel structure. It is used to delineate the stratigraphic architecture, lithology and lithofacies of the reservoir. Moreover, the interpretation of the reservoir properties is also based on the geological model. Core measurements and well log data are integrated to derive permeability and porosity. Based on production data, laboratory experiments and electrical logs, saturation profiles can be calculated. Production tests and well tests measure pressure and can be used to calculate permeability.

All this information is combined with different seismic attributes to design the fluid flow model. In this process geostatistical methods can be applied to capture the observed heterogeneity and complexity in a detailed description, and provide proper upscaled models to be used in the reservoir simulation model.

In a reservoir simulation model the different geological layers are characterized by their permeability and the porosity. The permeability is a measure of the rock's ability to transmit fluids through the porous medium. The permeability tensor is usually spatially dependent, since pores in the reservoir often show preferred alignment directions. The porosity, Φ, is the parameter which characterizes the porous medium, and is defined as the ratio of the pore volume to the total volume. The pore volume only includes the connected pores which contribute to fluid flow in a reservoir.

Figure 1.1 shows a reservoir simulation model of a North Sea field case. The reservoir model includes partly communicating fault blocks, and the heterogeneity of the permeability is shown in the oil zone.

The porous reservoir rock is sealed from above by an impermeable cap rock, typically a shale layer, preventing the hydrocarbons to escape from the reservoir. Below, the reservoir has an aquifer, which excites pressure support. Water, oil and gas are filling the porous rock and the individual fluid phases are separated by their densities, defining distinct fluid contacts: a water-oil (WOC) and a gas-oil (GOC) contact. The knowledge of the initial depth of both the water-oil and the gas-oil contact will be important in volumetric calculations when the production plan of the reservoir is designed. The main issue is to optimally place the production wells, and design water and gas injectors to effectively drain the reservoir. Pressure gradients towards the producers allow the oil start flowing into the wells. The purpose for the injectors will be to prevent pressure drop and to displace the oil towards the producers. The requirement for these reservoir simulation models should be that they are conditioned to reservoir specific observations (e.g production history, pressure, 3D seismic and time-lapse seismic data), and the aim is to adjust the model in such a way that the simulated performance of the model matches the observed history. This is referred to as the history matching problem of reservoir simulation models.

These conditioned models provide the basis for an uncertainty assessment through an ensemble of simulation studies, which, in turn, may be used to evalu-

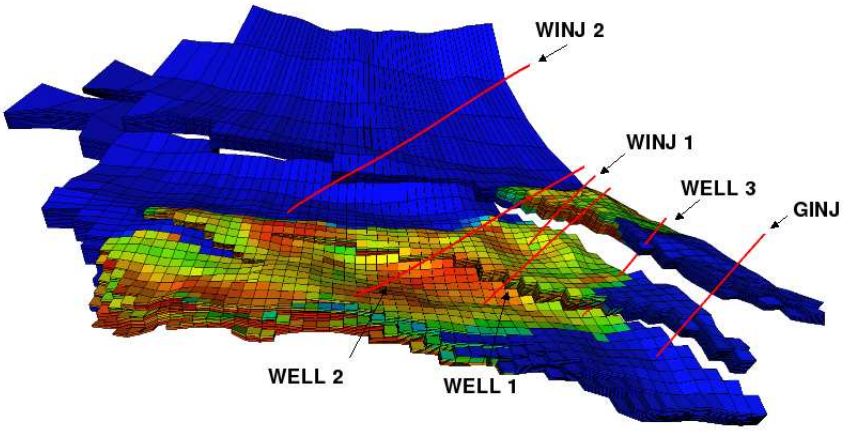


Figure 1.1: A reservoir simulation model of a real field case from the North Sea.

ate different recovery strategies. Such recovery strategies define the closing and opening criteria for production and injection wells, propose injection strategies of water, gas or chemicals and help to position infill wells. From an economical point of view, a reasonable description of the reservoir properties is a prerequisite for net-present value calculations. Optimal production techniques that will maximize the hydrocarbon production are assessed by simulating the fluid flow behaviour within the reservoir under various strategies.

The history matching process is known as an inverse problem, and a Bayesian framework is suitable for finding a solution [1]. A Bayesian framework for stochastic reservoir characterization includes a prior model and a likelihood function and both define the posterior model. The prior model typically involves the geological knowledge of the reservoir properties and should give a reasonable description of the prior uncertainty. The likelihood function normally includes both a forward modelling of the data acquisition procedure and a measurement error term which takes into account the uncertainty connected to the acquisition instruments.

In reservoir studies the posterior model will be of high dimension, including highly nonlinear forward models. Thus, an analytical derivation of the posterior model will not be feasible and sampling techniques are required to explore the uncertainty in the posterior model [2].

1.1 Main Objectives

Sampling from the posterior model requires that the reservoir models are conditioned to the dynamic data. The conditioning process is difficult and time consuming, and to date mainly done manually by trial and error approaches. Lately, computer-aided history matching processes have gained popularity and have been used in different reservoir applications. Traditionally, the reservoir models are then conditioned to dynamic data using an automatic history matching procedure in which the model is adjusted and the simulator is rerun in an interactive loop until a satisfactory match between the observed data and the simulated data has been achieved.

In this study we present a sequential data assimilation procedure, based on a Bayesian framework, where the model parameters and the state variables are updated continuously in time. When new data become available, the current model and the data are assimilated to improve the model. The idea is that the quality of the model will continuously improve with time. The focus lies on the predictions, and the model is not rerun after each update to check whether it matches the history. This methodology for continuous model updating is based on an ensemble Kalman filter (EnKF) technique [3], which is a Monte Carlo type sequential Bayesian inversion method.

The EnKF is related to the Kalman filter (KF) [4], which provides the optimal state estimate of a linear dynamic system, when the error statistics are assumed Gaussian. In Chapter 3 the Kalman filter is described and in Section 3.3.2 we present a correct sampling, called conditional simulation with kriging, of the Kalman filter distribution. Further, in Chapter 4 a complete description of the EnKF method is given, and in Section 4.1 we use the conditional simulation with kriging method to derive the EnKF sampling algorithm. In Section 4.2 we show that the same sampling strategy can be used when the data are dependent on more than one time instant.

In this work an assimilation of both production data and seismic data have been used. Production data are measurements of gas-oil ratio (GOR), water-cut (WCT) and flowing bottom-hole pressure (BHP). Regarding the seismic data, different ways of incorporating 3D seismic data and time lapse seismic data in the updating procedure are thoroughly investigated. Mainly, seismic data are measurements of acoustic impedance and Poisson's ratio. Grid cell pressure, grid cell saturation and seismic waveform data are also used as seismic data in the EnKF update. The sensitivities of both the size of the seismic measurement error and the correlation structure are studied.

The number of possible model variables to update in the EnKF process are usually very large. In this project we have mainly updated the static parameters, porosity and permeability, and the dynamic variables, saturations and pressures.

Moreover, we have updated lithology which is linked to a rock physics model, while the fracture density is updated in a synthetic fractured reservoir.

To incorporate both production data and seismic data in the data assimilation a coupled reservoir-seismic model is required. Hence, in this project a coupled reservoir-seismic model is developed for simulation of production data and repeated seismic surveys. Existing reservoir simulation software and a seismic modelling tool are applied in this task. In this work the coupled model is connected to rock physics models which are either based on the Gassmann relation [5] or effective medium theory (e.g Kuster-Toksöz method) [6]. A development of a coupling to a fast seismic modelling tool [7] is performed, where seismic waveform data are used as measurements in the assimilation. We have also developed a coupling tool for fractured reservoirs, where the hydraulic properties (i.e permeability) and the seismic attributes are derived at each simulation time step from the fracture density parameter by using a T-matrix approach [8].

The assimilation of time-lapse seismic data may be difficult since time-lapse seismic data are time difference data related to a model state at different times. Thus, special care needs to be taken when this type of data are introduced in the EnKF scheme. In Section 3.2 we present a general sequential Bayesian formulation which incorporates time difference data, and we show that the posterior model includes both a filter and a smoother solution. Further, in Section 4.2 we utilize time difference data and a non-linear state space model to present the EnKF scheme. We show that for updating a reservoir simulation model a combination of the ensemble Kalman filter and ensemble Kalman smoother has to be applied.

The important issue of large amounts of data, introduced by 3D seismic and time-lapse seismic data, is governed by efficient EnKF schemes [9]. The challenging aspects such as the inversion of a large covariance matrix and potential loss of rank are considered. In Section 4.5 we present an efficient subspace EnKF algorithm where a proper sampling avoids the loss of rank. The motivation for this algorithm is based on the efficient square root algorithm presented by Evensen [9].

Problems related to an underestimation of the uncertainty in the low rank covariance matrix representation of the static and dynamic variables may occur when few ensemble members and a large amount of data are used during the assimilation [10]. Hence, to improve the results a methodology based on a combination of a global and a local analysis scheme is proposed. The global and local analysis are used to assimilate production data and seismic data, where the local scheme, see Section 4.4.1, assumes that only seismic data within a certain distance from a model variable will impact the analysis of this particular model variable.

The EnKF method has been tested on both synthetic and real field reservoir models. In complex reservoir systems the forward simulation requires high computational costs. Thus, in a EnKF methodology, with typical ensemble sizes of 100 members or more, a sequential forward simulation for each member is very

time demanding. Hence, in this project we have developed a parallel implementation of the EnKF algorithm, which ensures that all the members are evolved forward in time in parallel until the next time step where new data become available. At this time the simulated data and the measured data are assimilated before a new parallel forward simulation is started up.

Seismic data have a much higher areal resolution than the simulation model, while the opposite is the case for the vertical resolution. In addition, the geological model and well log data show different degrees of spatial resolution. When coupling the models, this effect has to be taken into account. This includes upscaling and downscaling of properties and uncertainties, as well as data correlations. In this project we have performed a simulation study of a specific downscaling algorithm called Sequential Gaussian Simulation with Block Kriging (SGSBK) [11].

1.2 Overview of Papers and Results

In Paper A we show how the ensemble Kalman filter method can be used to update a combined reservoir simulation/seismic model using production data and inverted 4D seismic data. The efficient subspace EnKF algorithm presented in Section 4.5 is used to handle the large amount of data resulting from 4D seismic. When the seismic data are given as a difference between two surveys, a combination of the ensemble Kalman filter and the ensemble Kalman smoother is used, see Section 4.2. The method is applied to a synthetic model and a real field case from the North Sea.

For the synthetic model, the EnKF model updating with production data only is not able to recover the trend in the true permeability field, although a good match to the data is obtained. Furthermore, it is shown that the introduction of seismic data gives a much better estimate of reservoir permeability. However using only the seismic difference measurement reduces the quality significantly.

For the field case, two cases of EnKF updating were run. The first case assimilated production data only, while the second case assimilated one seismic difference dataset in addition to the production data. The EnKF run with both production and seismic data gave a better match both to production data and seismic data than the base case, and the run with only production data. For the EnKF runs, the root mean square for the seismic data was reduced by 25% and 14% for the runs with and without seismic data, respectively. The introduction of seismic data resulted in considerable different porosity and permeability estimates than using only production data, while retaining the production match. Conditioning only to production data and seismic difference data are an obvious weakness of the procedure used here, as it was also seen in the synthetic example. In future real field reservoir studies, the base line seismic survey should be included in the

EnKF updating.

Paper B presents a more detailed overview of the ensemble Kalman filter technique for continuous model updating with respect to time difference data, i.e. data related to a model state at different times. It shows that the use of time difference data involves a combination of the ensemble Kalman filter and the ensemble Kalman smoother. In addition, Paper B focuses on handling a large amount of data in the assimilation schemes, and an efficient subspace EnKF algorithm is proposed. The method is applied to a 3D synthetic reservoir model, where acoustic impedance is used as seismic data.

The synthetic case shows that the combined assimilation of production data, 3D seismic data and time-lapse seismic data is performed in a statistical consistent manner and has a positive impact in the EnKF update. For the synthetic model, the EnKF method gives good history matching results, and it is shown that a better estimate of the porosity and the permeability is obtained by adding seismic data to the production data measurements. Especially an improvement of the porosity is obtained, and one reason might be that the acoustic impedance measurements do have a stronger correlation to the porosity than to the permeability in the reservoir.

Simulation runs have been performed to compare the traditional and the subspace EnKF scheme, where different combinations of a diagonal and a correlated seismic measurement error covariance matrix have been used. It is shown from the simulation runs that the results obtained from the subspace EnKF scheme, when a diagonal measurement error covariance matrix is used, are nearly identical to the traditional EnKF scheme. Further, the simulation runs show that the use of a correlated measurement error covariance matrix in the subspace EnKF scheme gives a lower reduction of the ensemble variance compared to the traditional scheme, but results in similar rms values and correlation values between the estimated and the true porosity and permeability fields.

The computation time of the subspace EnKF scheme, when a large amount of data are incorporated in the assimilation, is very small compared to the traditional EnKF algorithm. Furthermore, it is seen from the experiments that there are no loss of rank in the analyzed ensemble, when the subspace EnKF scheme has been used. The results are fairly consistent, although the subspace EnKF scheme introduces an approximate pseudo inverse, when a correlated measurement error covariance matrix is included.

In Paper C a combination of 4D seismic waveforms data with production data are used to improve a synthetic 2D reservoir model during production. The inversion is based on the ensemble Kalman filter and the forward method is a combination of a fast seismic modelling tool [7] and a reservoir simulator. The method is applied to synthetic 2D reservoir model, and it is shown that introduction of production data and seismic waveform data give a fairly good estimation of the porosity and permeability fields. The method demonstrate that the estimates of

both the porosity and permeability fields can be improved over time with repeated seismic surveys and assimilation of production data.

In Paper D an improvement of the reservoir characterization is performed by introducing an uncertainty connected to parameters in the rock physics model, and we show that when seismic data are available the lithology can also be estimated. In this work the lithology is quantified by the clay ratio. A large amount of seismic data in the assimilation step is a concern, and to improve the results we have proposed a methodology based on a combination of a global and a local analysis scheme. The technique is applied to synthetic 2D and 3D reservoir models. The global and the local analysis are used to assimilate the production data, 3D seismic and 4D seismic data, where the local scheme assumes that only seismic data within a certain distance from a state variable will impact the analysis of this particular state variable. The rock physics model used here to compute the effective properties of a porous, fluid filled medium is an inclusion based model.

The results show that the combined global/local scheme is able to recover the main features in the true lithology field, and it is demonstrated that taking into account uncertainty in lithology by updating the clay ratio may be important to obtain satisfactory porosity and permeability estimation results. A proper representation of the seismic data error model is also shown to be of importance during the assimilation. Large amount of data in a global scheme with few ensemble members may cause a too large reduction in the uncertainty of the model variables. Here, the combined global and local scheme gave promising results regarding the uncertainty estimation, even with few members applied.

Paper E is an extension of Paper D, where effects of using local versus global analysis schemes on inverted seismic data are investigated. The elastic parameters derived from the rock physics model had different sensitivity to the parameters in the model. Hence, to take into account the complementary information both acoustic impedance and the Poisson's ratio differences are used simultaneously at a given seismic assimilation step. Other evaluated factors are the effects of using different stencil sizes in the local scheme, and the sensitivity of the EnKF performance related to the seismic measurement error quantity.

The performance of the EnKF was clearly sensitive to the ensemble size and the error level of the seismic data. And in the local scheme an increase of the stencil size gave a gradually smoother ensemble mean and a decrease in the ensemble spread. All of these sensitivities affected the characterization of the reservoir and influenced the covariance estimation of the model variables, and thus the uncertainty in the predictions of the production data.

The ensemble spread obtained from the global/local scheme are all higher than the global scheme with the same ensemble size, and the uncertainty level for both the permeability and the oil production rate predictions are similar to the global scheme using a large number of ensemble members.

The results show that the lithology may have a strong impact on the seismic signal, and an estimation of the clay ratio will be of importance to achieve satisfactory predictions of the elastic parameters. The use of an improved sampling algorithm is also shown to have a positive influence on the quality of the EnKF results.

Paper F proposes a method for characterization of natural fractures in hydrocarbon reservoirs by quantitative integration of production data and seismic attributes. The method is based on a unified model for the effective elastic and hydraulic properties of fractured porous media and an ensemble Kalman filter method. The methodology is applied to a 2D synthetic reservoir model, and when data becomes available the fracture density in each reservoir grid block is updated.

An application to synthetic data suggested that one may obtain a significantly better estimate of the fracture density and permeability distributions within a fractured reservoir, by using time-lapse measurements of seismic attributes (acoustic impedance) in addition to reservoir production data (BHP, WCT, GOR) in the dynamic reservoir characterization process.

Report 1 presents a study of a North Sea field case, where real production data (BHP, WCT) and seismic time-lapse data (Poisson's ratio) have been assimilated using the ensemble Kalman filter methodology.

In this history matching study it has been difficult to obtain a good match of both production data and seismic data simultaneously. Matching the seismic data resulted in a too early water break-through in most of the wells, and the EnKF tried to compensate for this by performing large modifications of the permeability and the porosity. The difference in the absolute values between the simulated and measured 4D seismic data were very large, and the EnKF had problems adjusting the reservoir model properly to honor the data. An incorrect rock physics model may explain the discrepancy in the level, and in future work it would be desirable to introduce an uncertainty to parameters in the rock physics model which are sensitive to the seismic signal.

In Report 2 a simulation study of a specific downscaling algorithm, named Sequential Gaussian Simulation with Block Kriging (SGSBK), is performed.

The SGSBK results show that the method generates fine-scale geostatistical models constrained to coarse-scale average constraints in a proper way, that honor the prescribed histogram and the spatial covariance models.

Chapter 2

Reservoir Characterization in a Bayesian Setting

Following Tarantola [1] the study of a physical system can be divided into three major steps. The first step is called the *parametrization of the system*, where the main objective is to identify a minimal set of model parameters that are able to characterize the physical system completely.

The second step is named the *forward modelling*. A forward model is typically designed by a mathematical model, where a complete description of the physical system and the physical properties exists and a unique response prediction of the observable variables in the system is available. Thus, in reservoir characterization where also seismic attributes are calculated, the typical forward problem is represented by a combination of the fluid flow simulator and the seismic forward model.

The third step is the *inverse modelling*. The inverse modelling can be described as a determination of model parameters that characterize the system, given measurements of the observable variables of the system. In this work we focus on the inverse problem and we try to recover reservoir properties, which will be of importance to characterize the reservoir.

2.1 Bayesian Framework

The inverse problem is normally an ill-posed problem [12, 13], and the most general theory is achieved when using a probabilistic point of view. In a Bayesian framework one assumes that prior information on model parameters is given by a probability density function (pdf), and the available observations are linked to the parameters of interest through their respective likelihood models. Given prior information on the model parameters, the likelihood distribution of the measure-

ments, and an uncertain relation between the data and the model parameters, a posterior distribution can be established [14, 15]. The construction of the posterior pdf is often expressed as

$$\textbf{Posterior} \propto \textbf{Observed Likelihood} \times \textbf{Prior}, \tag{2.1}$$

and represents the solution of the inverse problem [16, 17]. To obtain a plausible solution, the model must be consistent with the physical constraints and the measured data. However for inverse problems in reservoir characterization infinitely many models may satisfy this criterion. In a Bayesian setting, such a model will typically be a sample from the posterior distribution [2, 18].

In a reservoir characterization approach of the inverse problem, also called history matching, the establishment of the prior distribution for the reservoir parameters (e.g porosity and permeability) are typically based on geological information. The prior is often assumed to be Gaussian distributed and the stochastic model is then fully explained by the mean and the covariance. The likelihood model for the well production data and the seismic data involve both the forward model and the measurement uncertainties, where the measurement error term is normally assumed Gaussian distributed. The characterization of the reservoir parameters given production and seismic data, are then found from the posterior pdf, where an estimate of the mean or the maximum a posteriori solution is a main objective. The forward model in reservoir cases is highly non-linear and a sampling strategy is needed to create a suite of realizations of reservoir parameters from the posterior pdf. Then, for each realization a reservoir prediction performance can be generated, and a determination of the uncertainty for each performance variable (e.g bottom-hole pressure, gas-oil ratio, water-cut, cumulative oil production) can be found from the predicted outcome statistics. It is important to notice that if the simulated reservoir properties do not represent a correct sampling of the posterior distribution, the realizations of the fields can not be expected to give a reliable characterization of the uncertainty in the predictions.

2.2 Inverse Problem

Let the model parameters be denoted by the vector $\boldsymbol{\alpha}$, the recovery strategy by η, the observable variables (data) by $\boldsymbol{d}_\eta$, and the forward model by the model operator $\boldsymbol{f}$. The theoretical relationship between the model parameters and the data can then be expressed as

$$\boldsymbol{d}_\eta = \boldsymbol{f}(\boldsymbol{\alpha}, \eta). \tag{2.2}$$

The forward model will normally not represent the physical system exactly, and the uncertain relation between the model parameters and the data can be described

as

$$\boldsymbol{d}_\eta = \boldsymbol{f}(\boldsymbol{\alpha}, \eta) + \boldsymbol{\epsilon}^m, \tag{2.3}$$

where $\boldsymbol{\epsilon}^m$ is a vector of random model errors with mean zero and covariance matrix $\boldsymbol{P}_{\epsilon^m}$, denoted on a short form as $\boldsymbol{\epsilon}^m \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{P}_{\epsilon^m})$. Normally it is assumed that $\boldsymbol{\epsilon}^m$ is Gaussian distributed, and that $\boldsymbol{\alpha}$ are independent of the model errors. Let from now the recovery strategy η be implicitly assumed in the forward model and not visible in the notation. The difference between the data, $\boldsymbol{d}_\eta$, and the model prediction, $\boldsymbol{f}(\boldsymbol{\alpha})$, is given by the following conditional probability density function

$$\begin{aligned} g(\boldsymbol{d}_\eta|\boldsymbol{\alpha}) &= g(\boldsymbol{\epsilon}^m = \boldsymbol{d}_\eta - \boldsymbol{f}(\boldsymbol{\alpha})) \\ &= const \cdot \exp\left(-\frac{1}{2}(\boldsymbol{d}_\eta - \boldsymbol{f}(\boldsymbol{\alpha}))^T \boldsymbol{P}_{\epsilon^m}^{-1}(\boldsymbol{d}_\eta - \boldsymbol{f}(\boldsymbol{\alpha}))\right), \end{aligned} \tag{2.4}$$

with *const* being the normalizing constant. In a similar way the measurement errors can be represented as a conditional probability density function for the observed data values, $\boldsymbol{d}$, given the data, $\boldsymbol{d}_\eta$, denoted $g(\boldsymbol{d}|\boldsymbol{d}_\eta)$. Assuming the measurement errors are Gaussian distributed with mean zero and covariance matrix $\boldsymbol{P}_{\epsilon^o}$, the probability density, $g(\boldsymbol{d}|\boldsymbol{d}_\eta)$, is given as

$$g(\boldsymbol{d}|\boldsymbol{d}_\eta) = const \cdot \exp\left(-\frac{1}{2}(\boldsymbol{d}_\eta - \boldsymbol{d})^T \boldsymbol{P}_{\epsilon^o}^{-1}(\boldsymbol{d}_\eta - \boldsymbol{d})\right). \tag{2.5}$$

The likelihood function, $g(\boldsymbol{d}|\boldsymbol{\alpha})$, defined as the conditional probability density for the observed data values given a model parameter value, is found from the following equation

$$g(\boldsymbol{d}|\boldsymbol{\alpha}) = \int g(\boldsymbol{d}|\boldsymbol{d}_\eta) g(\boldsymbol{d}_\eta|\boldsymbol{\alpha}) d\boldsymbol{d}_\eta, \tag{2.6}$$

where we assume that the measurement errors are independent of the model errors. If the model and the measurement errors are assumed Gaussian it can be shown that the likelihood function can be written as [1]

$$g(\boldsymbol{d}|\boldsymbol{\alpha}) = const \cdot \exp\left(-\frac{1}{2}(\boldsymbol{f}(\boldsymbol{\alpha}) - \boldsymbol{d})^T \boldsymbol{P}_d^{-1}(\boldsymbol{f}(\boldsymbol{\alpha}) - \boldsymbol{d})\right), \tag{2.7}$$

where the covariance matrix $\boldsymbol{P}_d = \boldsymbol{P}_{\epsilon^m} + \boldsymbol{P}_{\epsilon^o}$ combines the modelling and the measurement errors. The density function $g(\boldsymbol{d}|\boldsymbol{\alpha})$ is Gaussian with expectation $\boldsymbol{f}(\boldsymbol{\alpha})$ and covariance $\boldsymbol{P}_d$.

The prior probability density for the model parameters, $g(\boldsymbol{\alpha})$, is based on information which is found independently of the observations. Assuming that the a priori information on the model parameters is Gaussian

$$g(\boldsymbol{\alpha}) = const \cdot \exp\left(-\frac{1}{2}(\boldsymbol{\alpha}-\boldsymbol{\alpha}_{prior})^T \boldsymbol{P}_\alpha^{-1}(\boldsymbol{\alpha}-\boldsymbol{\alpha}_{prior})\right), \tag{2.8}$$

where $\boldsymbol{\alpha}_{prior}$ and $\boldsymbol{P}_\alpha$ are respectively the a priori mean and covariance.

In a Bayesian framework the solution of the inverse problem is given by the posterior pdf. The posterior pdf is the probability distribution for the model parameters, $\boldsymbol{\alpha}$, given the observations, $\boldsymbol{d}$, which takes into account both the likelihood function and the prior model [19],

$$g(\boldsymbol{\alpha}|\boldsymbol{d}) = \frac{g(\boldsymbol{d}|\boldsymbol{\alpha})g(\boldsymbol{\alpha})}{g(\boldsymbol{d})} = \frac{g(\boldsymbol{d}|\boldsymbol{\alpha})g(\boldsymbol{\alpha})}{\int g(\boldsymbol{d}|\boldsymbol{\alpha})g(\boldsymbol{\alpha})d\boldsymbol{\alpha}}. \tag{2.9}$$

The solution for Gaussian statistics can also be written as

$$g(\boldsymbol{\alpha}|\boldsymbol{d}) = const \cdot \exp(-\boldsymbol{J}(\boldsymbol{\alpha})), \tag{2.10}$$

where $\boldsymbol{J}(\boldsymbol{\alpha})$ is the objective function

$$\begin{aligned}\boldsymbol{J}(\boldsymbol{\alpha}) = &\frac{1}{2}\left((\boldsymbol{f}(\boldsymbol{\alpha})-\boldsymbol{d})^T \boldsymbol{P}_d^{-1}(\boldsymbol{f}(\boldsymbol{\alpha})-\boldsymbol{d})\right) \\ &+\frac{1}{2}\left((\boldsymbol{\alpha}-\boldsymbol{\alpha}_{prior})^T \boldsymbol{P}_\alpha^{-1}(\boldsymbol{\alpha}-\boldsymbol{\alpha}_{prior})\right).\end{aligned} \tag{2.11}$$

The traditional Bayesian maximum a posteriori (MAP) predictor for the model parameters is given as [20]

$$\begin{aligned}\boldsymbol{\alpha}_{map} &= argmax\{g(\boldsymbol{\alpha}|\boldsymbol{d})\} \\ &= argmin\{\boldsymbol{J}(\boldsymbol{\alpha})\},\end{aligned} \tag{2.12}$$

where *argmax* and *argmin* are the arguments which optimize the function. If the relation between the model parameters and the data is linear

$$\boldsymbol{d}_\eta = \boldsymbol{f}(\boldsymbol{\alpha}) = \boldsymbol{F}\boldsymbol{\alpha}, \tag{2.13}$$

the posterior pdf of interest, $g(\boldsymbol{\alpha}|\boldsymbol{d})$, will be Gaussian with mean [21]

$$\boldsymbol{\alpha}_c = \boldsymbol{\alpha}_{prior} + \boldsymbol{P}_\alpha \boldsymbol{F}^T(\boldsymbol{F}\boldsymbol{P}_\alpha\boldsymbol{F}^T + \boldsymbol{P}_d)^{-1}(\boldsymbol{d}-\boldsymbol{F}\boldsymbol{\alpha}_{prior}), \tag{2.14}$$

and covariance

$$\boldsymbol{P}_c = \boldsymbol{P}_\alpha - \boldsymbol{P}_\alpha \boldsymbol{F}^T(\boldsymbol{F}\boldsymbol{P}_\alpha\boldsymbol{F}^T + \boldsymbol{P}_d)^{-1}. \tag{2.15}$$

Notice that for the Gaussian linear model the mean in Equation (2.14) and the map solution in Equation (2.12) will give the same solution.

2.3 Sampling from the Posterior Distribution

To quantify the prediction uncertainty of the model parameters an assessment of the full posterior distribution, $g(\boldsymbol{\alpha}|\boldsymbol{d})$, is required. By using a non-linear forward model an analytical evaluation will be prohibited and the exploration of the posterior pdf can only be done by sampling. Rejection sampling and Markov chain Monte Carlo (MCMC) are two sampling routines which can be used to sample from the posterior pdf, and both of the algorithms can be shown to provide correct samples from the target pdf.

2.3.1 Rejection Sampling

In the rejection sampling method [22], we sample from a relatively simple proposal distribution $p(\boldsymbol{\alpha})$ rather than from the target probability density $g(\boldsymbol{\alpha}|\boldsymbol{d})$, and then apply a test to decide whether to accept it or not.

Specifically let c be a constant so that

$$\frac{g(\boldsymbol{\alpha}|\boldsymbol{d})}{p(\boldsymbol{\alpha})} \leq c \quad \forall \ \boldsymbol{\alpha}. \tag{2.16}$$

We then have the following technique for simulating samples from $g(\boldsymbol{\alpha}|d)$:

Algorithm 2.3.1

1. Generate a proposal sample $\boldsymbol{\alpha}^*$ from pdf $p(\boldsymbol{\alpha})$ and simulate a random number u from a uniform distribution U(0,1)

2. If $u \leq g(\boldsymbol{\alpha}^*|\boldsymbol{d})/cp(\boldsymbol{\alpha}^*)$ set $\boldsymbol{\alpha} = \boldsymbol{\alpha}^*$. Otherwise return to Step 1.

Here, all generated samples will be independent.

A high acceptance rate in the rejection sampling method is obtained by selecting a proposal distribution that is a close approximation to the posterior distribution. In reservoir applications, it can be very hard to find a simple proposal distribution, especially because the number of model parameters is large in such cases.

2.3.2 MCMC Algorithm

The aim behind the MCMC algorithm is to build a Markov chain that is easy to simulate and has a target distribution given by the distribution of interest. In a MCMC methodology this is done by an iterative procedure with each iteration containing a proposal and an accept/reject step. An advantage of the MCMC algorithm is that it provides samples from a probability density where the normalizing

constant is very difficult to compute since the unknown normalization constant cancels in the algorithm. The MCMC method is as follows [2, 23]:

Algorithm 2.3.2

1. For initiation $j = 1$, choose an initial state $\boldsymbol{\alpha}^{(1)}$ with $g(\boldsymbol{\alpha}^{(1)}|\boldsymbol{d}) > 0$
2. Generate a sample $\boldsymbol{\alpha}^*$ from the proposal pdf $p(\boldsymbol{\alpha}|\boldsymbol{\alpha}^{(j)})$
3. Calculate the acceptance probability
   $\beta(\boldsymbol{\alpha}^*, \boldsymbol{\alpha}^{(j)}) = \min\left(1, \frac{g(\boldsymbol{\alpha}^*|\boldsymbol{d})}{g(\boldsymbol{\alpha}^{(j)}|\boldsymbol{d})} \times \frac{p(\boldsymbol{\alpha}^{(j)}|\boldsymbol{\alpha}^*)}{p(\boldsymbol{\alpha}^*|\boldsymbol{\alpha}^{(j)})}\right)$
4. Assign $\boldsymbol{\alpha}^{(j+1)}$ by:
   $\boldsymbol{\alpha}^{(j+1)} = \boldsymbol{\alpha}^*$ with probability $\beta(\boldsymbol{\alpha}^*, \boldsymbol{\alpha}^{(j)})$
   $\boldsymbol{\alpha}^{(j+1)} = \boldsymbol{\alpha}^{(j)}$ else
5. $j \to j + 1$, return to Step 2.

For a sufficiently large number of iterations, this trajectory will eventually produce draws from $g(\boldsymbol{\alpha}|\boldsymbol{d})$. By constructing a suitable Markov chain, one is able to perform a Monte Carlo simulation of values from the equilibrium distribution $g(\boldsymbol{\alpha}|\boldsymbol{d})$, thus the name MCMC. There exists different possibilities for the acceptance criterion and in the algorithm above a Metropolis-Hastings criterion is used [24]. In the Metropolis-Hastings algorithm the procedure for assigning the initial state, $\boldsymbol{\alpha}^{(1)}$, and the choice of $p(\cdot|\cdot)$ will have an important influence of the chain mix and the convergence rate. Another scheme is provided by the Gibbs sampler algorithm [25], and a more detailed discussion can be found in [26].

No matter which scheme is used to establish the chain, a stream of values, $\boldsymbol{\alpha}^{(1)}$,$\boldsymbol{\alpha}^{(2)}$,..., is generated. Although consecutive values $\boldsymbol{\alpha}^{(j)}$ and $\boldsymbol{\alpha}^{(j+1)}$ are dependent, a random sample of size N from $g(\boldsymbol{\alpha}|\boldsymbol{d})$ can be formed by retaining N successive values after convergence has been ascertained.

Independent samples can be obtained by holding only the N samples lagged by κ units, for example $(\boldsymbol{\alpha}^{(n)}, \boldsymbol{\alpha}^{(n+\kappa)}, \ldots, \boldsymbol{\alpha}^{(n+(N-1)\kappa)})$, where κ is large enough to carry only residual correlation over the chain and n is large enough to ensure convergence has been achieved. A challenge in the applications of MCMC methods is to ensure convergence of the Markov chain. A common approach is to plot the averages of the components of the model parameters and assess by visual inspection whether the convergence has occurred. The degree of correlation may be quantified by an auto-correlation function.

The MCMC-algorithm has been applied to many history matching applications, see [23, 27, 28, 29]. However, the computer expenses to obtain an uncertainty assessment by the MCMC-algorithm are very large, and many papers use

approximate sampling algorithms, which include an optimization step, in order to reduce the computational cost. Randomized maximum likelihood (RML) is such an approximate algorithm.

2.3.3 Randomized Maximum Likelihood

Randomized maximum likelihood is known as an approximate sampling algorithm when the forward model is nonlinear. However, it can be shown that the algorithm samples correctly when the data relationship is linear and Gaussian statistics are assumed. To create samples with an RML method for large problems, will as for the MCMC-algorithm be a difficult task and require large computer processing resources. An important advantage by using the RML method is that the generated samples often obtain a satisfactory history match to the data and look reasonable compared to the input properties from the prior model. The algorithm was first introduced by Kitanidis [30] and Oliver et. al. [18], and they proposed that unconditional samples from a Gaussian distribution could be used to generate samples conditional to nonlinear data by an optimization process. In the algorithm, it is required that the prior covariance of the model parameters, $\boldsymbol{P}_\alpha$, and the data error covariance, $\boldsymbol{P}_d$, are known. Samples from the RML method can then be generated as follows:

Algorithm 2.3.3

1. Generate an unconditional sample from the prior distribution $g(\boldsymbol{\alpha})$:
   $\boldsymbol{\alpha}^* \sim \mathcal{N}(\boldsymbol{\alpha}_{prior}, \boldsymbol{P}_\alpha)$

2. Generate an unconditional sample of the data:
   $\boldsymbol{d}^* \sim \mathcal{N}(\boldsymbol{d}, \boldsymbol{P}_d)$

3. Compute the conditional model $\boldsymbol{\alpha}_c$ that minimizes
   $\boldsymbol{J}(\boldsymbol{\alpha}) = \frac{1}{2}\Big((\boldsymbol{f}(\boldsymbol{\alpha}) - \boldsymbol{d}^*)^T \boldsymbol{P}_d^{-1}(\boldsymbol{f}(\boldsymbol{\alpha}) - \boldsymbol{d}^*) + (\boldsymbol{\alpha} - \boldsymbol{\alpha}^*)^T \boldsymbol{P}_\alpha^{-1}(\boldsymbol{\alpha} - \boldsymbol{\alpha}^*)\Big)$

Step three in the algorithm performs a minimization comparable to the computation of the maximum a posteriori estimate in Equation (2.12). The difference is now that the regularization is with respect to unconditional samples of the data and the model instead of the observed data and the prior model.

Numerical Optimization Methods

The sampling of the posterior pdf performed by the RML algorithm and the maximum a posteriori predictor in Equation (2.12) shows that the inverse problem can be reformulated as an optimization problem. If the forward model is non-linear,

the inverse problem may have several local minima and in some cases no unique global minimum. This is typically the case when using a multi-phase fluid flow model, where the relationship between the physical properties and the output from the model is highly non-linear.

Inverse problems defined as a minimization problem can be solved by a gradient-based methodology, and different minimization algorithms can be used, such as steepest decent, quasi Newton, conjugate gradient or Levenberg-Marquart, and a variety of different methods for computing the gradient are available (e.g. sensitivity or adjoint equations).

One of the most efficient gradient-based methods solving the history matching problem, is the algorithm based on an adjoint method to calculate the gradient of the squared data mismatch, see for example [31], and the limited memory Broyden-Fletcher-Goldfarb-Shanno method (LBFGS) [32] to compute the direction of the change [33]. The adjoint method has gained popularity and is widely used in history matching applications [31, 34]. However, an implementation of the adjoint system can be very time consuming, and might also be difficult to perform since it requires a detailed knowledge of the numerical schemes used in the forward simulator. Such gradient-based methods are called deterministic optimization methods since they always produce the same result for a given starting condition.

An alternative to the deterministic optimization methods are the stochastic optimization methods. The stochastic optimization methods include randomness in the optimization process, and the randomness is commonly related to how the search process is performed. Some of the stochastic methods may obtain an ensemble of solutions, which can be used to represent the uncertainty of the solution. Typically stochastic methods are direct search methods such as simulated annealing, genetic algorithms and neighborhood approximation. The latter one was first applied in petroleum reservoir applications by Subbey et al. [35].

Chapter 3

Sequential Data Assimilation

In Chapter 2, the parameter estimation problem and different strategies for sampling the posterior probability density function were discussed. In this chapter we will present a statistically consistent formulation of the combined parameter and state estimation problem, and discuss the main concepts of the Kalman filter method.

As stated by Evensen [3], the joint parameter and state estimation problem for a dynamical model can in general form be formulated as: *"how to find the joint pdf of the parameters and model state, given a set of measurements and a dynamical model with known uncertainties"*. In this context, based on the Bayesian framework, it is possible to derive a statistical algorithm for the combined parameter and state estimation, which processes the information from measurements sequentially in time.

3.1 Parameter and State Estimation

The general smoother for non-linear dynamics can be formulated as a sequential method, that is, observations can be assimilated sequentially during a forward integration. The general data assimilation problem was introduced using a Bayesian formulation, see [36, 37], and following the probabilistic derivation is explained.

Let $\boldsymbol{\psi}(\boldsymbol{x},t)$ denote the unknown model variable

$$\begin{bmatrix} \boldsymbol{u}(\boldsymbol{x},t) \\ \boldsymbol{\alpha}(\boldsymbol{x}) \end{bmatrix}, \tag{3.1}$$

where $\boldsymbol{u}(\boldsymbol{x},t)$ represent an unknown model state in space and time, and $\boldsymbol{\alpha}(\boldsymbol{x})$ represent some poorly known model parameters, which are assumed to be constant in time. The unknown model variable can be described by a prior probability density $g(\boldsymbol{\psi})$, and the probability density for the observations, $\boldsymbol{d}$, is given by the likelihood

$$\begin{array}{ccccccccccc} \boldsymbol{\psi}_0 & \rightarrow & \boldsymbol{\psi}_1 & \rightarrow & \cdots & \rightarrow & \boldsymbol{\psi}_k & \rightarrow & \boldsymbol{\psi}_{k+1} & \rightarrow & \cdots \\ & & \downarrow & & & & \downarrow & & \downarrow & & \\ & & \boldsymbol{d}_1 & & & & \boldsymbol{d}_k & & \boldsymbol{d}_{k+1} & & \end{array}$$

Figure 3.1: Data assimilation scheme.

distribution $g(\boldsymbol{d}|\boldsymbol{\psi})$. From Bayesian theory the conditional probability density of a model variable $\boldsymbol{\psi}$ given a vector of observations, $\boldsymbol{d}$, can then be defined as

$$g(\boldsymbol{\psi}|\boldsymbol{d}) = \frac{g(\boldsymbol{d}|\boldsymbol{\psi})g(\boldsymbol{\psi})}{\int g(\boldsymbol{d}|\boldsymbol{\psi})g(\boldsymbol{\psi})d\boldsymbol{\psi}}. \tag{3.2}$$

Now assume that the observations are available at the same discrete set of times, $t_0, t_1, \ldots, t_k$, as where we want to estimate $\boldsymbol{\psi}$, and let $\boldsymbol{\psi}_k = \boldsymbol{\psi}(\boldsymbol{x}, t_k) \in \Re^q$ and $\boldsymbol{d}_k \in \Re^{m_k}$ denote the model variables and the observations at time t_k. Figure 3.1 illustrates a data assimilation scheme where $\boldsymbol{d}_k$ depend only on $\boldsymbol{\psi}_k$.

To simplify the notation we let $\boldsymbol{\psi}_{k:0}$ denote the sequence $\{\boldsymbol{u}_k, \ldots, \boldsymbol{u}_1, \boldsymbol{u}_0, \boldsymbol{\alpha}\}$. The general smoother in Equation (3.2) can be formulated as a sequential method [37], and the smoother for the time interval $t \in [t_0, t_k]$ can be expressed as

$$g(\boldsymbol{\psi}_{k:0}|\boldsymbol{d}_{k:1}) = const \cdot g(\boldsymbol{d}_k|\boldsymbol{\psi}_{k:0})g(\boldsymbol{\psi}_k|\boldsymbol{\psi}_{k-1:0})g(\boldsymbol{\psi}_{k-1:0}|\boldsymbol{d}_{k-1:1}). \tag{3.3}$$

We now assume that the model evolution is a first-order Markov process, and that the observations collected at different times are independent. Moreover, it is assumed that $\boldsymbol{d}_k$ will only depend on $\boldsymbol{\psi}_k$. We can then write the smoother as

$$g(\boldsymbol{\psi}_{k:0}|\boldsymbol{d}_{k:1}) = const \cdot g(\boldsymbol{d}_k|\boldsymbol{\psi}_k)g(\boldsymbol{\psi}_k|\boldsymbol{\psi}_{k-1})g(\boldsymbol{\psi}_{k-1:0}|\boldsymbol{d}_{k-1:1}). \tag{3.4}$$

In Equation (3.4) the density function $g(\boldsymbol{d}_k|\boldsymbol{\psi}_k)$ represent the likelihood term, $g(\boldsymbol{\psi}_k|\boldsymbol{\psi}_{k-1})$ corresponds to an integration of the solution forward in time from t_{k-1} to t_k, and the multiplication with the last two densities represent the prior distribution at time t_k.

In a similar probabilistic formalism the general filter solution can be derived from Equation (3.4) by integrating over the solutions $\boldsymbol{\psi}_{k-1:0}$, and the filter equation is given by

$$g(\boldsymbol{\psi}_k|\boldsymbol{d}_{k:1}) = const \cdot g(\boldsymbol{d}_k|\boldsymbol{\psi}_k)g(\boldsymbol{\psi}_k|\boldsymbol{d}_{k-1:1}), \tag{3.5}$$

where the prior density, $g(\boldsymbol{\psi}_k|\boldsymbol{d}_{k-1:1})$, is defined as

$$g(\boldsymbol{\psi}_k|\boldsymbol{d}_{k-1:1}) = \int g(\boldsymbol{\psi}_k|\boldsymbol{\psi}_{k-1})g(\boldsymbol{\psi}_{k-1}|\boldsymbol{d}_{k-1:1})d\boldsymbol{\psi}_{k-1}. \tag{3.6}$$

In the filter Equation (3.5) the information is carried only forward in time and the model variable $\boldsymbol{\psi}_k$ is dependent on all the previous data, $\boldsymbol{d}_1, \boldsymbol{d}_2, \ldots, \boldsymbol{d}_k$. Notice that the estimate at the final time, t_k, will be identical for the smoother and the filter.

3.2 Sequential Update Using Time Difference Data

When using time difference data the observations at a particular time are related to the model state at more than one time instant. Thus, in the Bayesian formulation of the sequential data assimilation the posterior distribution includes both a filter and a smoother solution. In this section we present how this can be performed in a statistical consistent manner, and a more detailed discussion is given in Paper A and B.

In the smoother (3.4) and filter (3.5) equations, we assumed that the data at a specific time were only related to the model variables at that time. In this section we will derive the update procedure for the smoother and filter solutions when data $\boldsymbol{d}_k = [\boldsymbol{d}_{kk}, \boldsymbol{d}_{kj}]^T$ at a particular time t_k are related to both $\boldsymbol{\psi}_k$ and $\boldsymbol{\psi}_j$, where $t_j < t_k$. The likelihood function is then represented by $g(\boldsymbol{d}_k|\boldsymbol{\psi}_k, \boldsymbol{\psi}_j)$. From now on $\boldsymbol{d}_{kk}$ denote an observation that depend only on $\boldsymbol{\psi}_k$, and $\boldsymbol{d}_{kj}$ denote an observation that depend both on $\boldsymbol{\psi}_k$ and $\boldsymbol{\psi}_j$. In our model we let $\boldsymbol{d}_{kj}$ be the time difference observation between $\boldsymbol{\psi}_k$ and $\boldsymbol{\psi}_j$. We still assume a Markov process and Figure 3.2 introduces the model, where the sequential updating procedure is as follows:

1. For the time interval $t_i \in (t_0, t_j]$ the data, $\boldsymbol{d}_i = [\boldsymbol{d}_{ii}]$, depend only on $\boldsymbol{\psi}_i$.

2. For the time interval $t_i \in (t_j, t_k)$ the update is based on the data, $\boldsymbol{d}_i = [\boldsymbol{d}_{ii}]$. Here $\boldsymbol{\psi}_j$ will be included in the update procedure as a smoother update, because at time t_k the observation $\boldsymbol{d}_k$ will include time difference data between $\boldsymbol{\psi}_k$ and $\boldsymbol{\psi}_j$.

3. When the time is equal t_k the update based on the data, $\boldsymbol{d}_k = [\boldsymbol{d}_{kk}, \boldsymbol{d}_{kj}]^T$, depend on $\boldsymbol{\psi}_k$ and $\boldsymbol{\psi}_j$.

4. When $t_i > t_k$ the update is based on the data, $\boldsymbol{d}_i = [\boldsymbol{d}_{ii}]$, and the observation depend only on $\boldsymbol{\psi}_i$.

From Equation (3.3) and Figure 3.2 the smoother at time t_k can be written as

$$g(\boldsymbol{\psi}_{k:0}|\boldsymbol{d}_{k:1}) \propto g(\boldsymbol{d}_k|\boldsymbol{\psi}_k, \boldsymbol{\psi}_j) g(\boldsymbol{\psi}_k|\boldsymbol{\psi}_{k-1}) g(\boldsymbol{\psi}_{k-1:0}|\boldsymbol{d}_{k-1:1}), \tag{3.7}$$

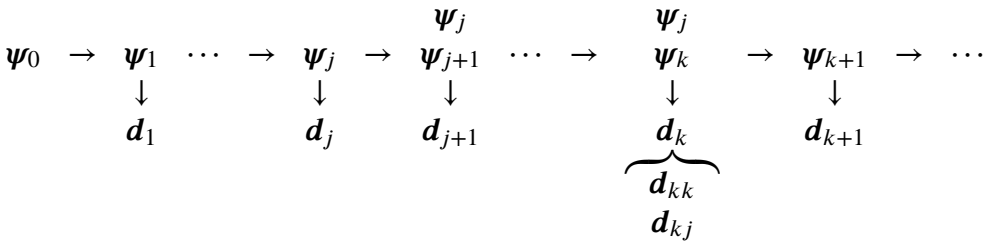


Figure 3.2: Data assimilation scheme where $\boldsymbol{d}_k$ depend both on $\boldsymbol{\psi}_k$ and $\boldsymbol{\psi}_j$.

and the filter solution for Equation (3.7) is given by

$$\begin{aligned} g(\boldsymbol{\psi}_k|\boldsymbol{d}_{k:1}) &\propto \int g(\boldsymbol{d}_k|\boldsymbol{\psi}_k,\boldsymbol{\psi}_j)g(\boldsymbol{\psi}_k|\boldsymbol{\psi}_{k-1})g(\boldsymbol{\psi}_{k-1:0}|\boldsymbol{d}_{k-1:1})d\boldsymbol{\psi}_{k-1:0} \\ &\propto \int\int g(\boldsymbol{d}_k|\boldsymbol{\psi}_k,\boldsymbol{\psi}_j)g(\boldsymbol{\psi}_k|\boldsymbol{\psi}_{k-1}) \\ &\cdot \left[\int\int g(\boldsymbol{\psi}_{k-1:0}|\boldsymbol{d}_{k-1:1})d\boldsymbol{\psi}_{k-2:j+1}d\boldsymbol{\psi}_{j-1:0}\right] d\boldsymbol{\psi}_{k-1}d\boldsymbol{\psi}_j \\ &\propto \int g(\boldsymbol{d}_k|\boldsymbol{\psi}_k,\boldsymbol{\psi}_j)\left[\int g(\boldsymbol{\psi}_k|\boldsymbol{\psi}_{k-1})g(\boldsymbol{\psi}_{k-1},\boldsymbol{\psi}_j|\boldsymbol{d}_{k-1:1})d\boldsymbol{\psi}_{k-1}\right] d\boldsymbol{\psi}_j \\ &\propto \int g(\boldsymbol{d}_k|\boldsymbol{\psi}_k,\boldsymbol{\psi}_j)g(\boldsymbol{\psi}_k,\boldsymbol{\psi}_j|\boldsymbol{d}_{k-1:1})d\boldsymbol{\psi}_j, \end{aligned} \tag{3.8}$$

where the forecast step defining the prior distribution is expressed by

$$g(\boldsymbol{\psi}_k,\boldsymbol{\psi}_j|\boldsymbol{d}_{k-1:1}) = \int g(\boldsymbol{\psi}_k|\boldsymbol{\psi}_{k-1})g(\boldsymbol{\psi}_{k-1},\boldsymbol{\psi}_j|\boldsymbol{d}_{k-1:1})d\boldsymbol{\psi}_{k-1}, \tag{3.9}$$

and the posterior distribution at time t_k is defined by

$$g(\boldsymbol{\psi}_k,\boldsymbol{\psi}_j|\boldsymbol{d}_{k:1}) \propto g(\boldsymbol{d}_k|\boldsymbol{\psi}_k,\boldsymbol{\psi}_j)g(\boldsymbol{\psi}_k,\boldsymbol{\psi}_j|\boldsymbol{d}_{k-1:1}). \tag{3.10}$$

Notice that the forecast step in Equation (3.9) requires the filter solution of $\boldsymbol{\psi}_{k-1}$ and the smoother solution of $\boldsymbol{\psi}_j$ obtained from $g(\boldsymbol{\psi}_{k-1},\boldsymbol{\psi}_j|\boldsymbol{d}_{k-1:1})$. The Bayesian formulation of the filter solutions in Equations (3.5) and (3.8) now gives the possibility to compute an update of the posterior distribution when we include time difference data. This sequential update procedure should also be possible to use for other types of data which depend on two or more time instances.

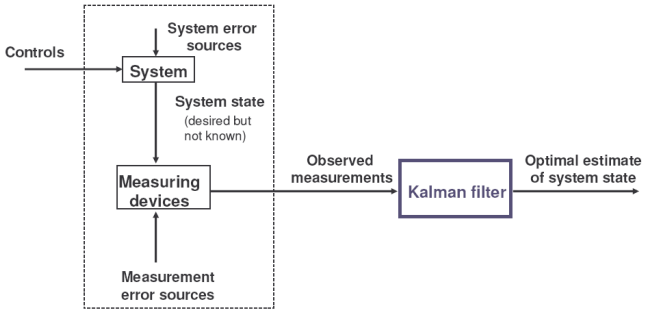


Figure 3.3: Typical Kalman filter application.

3.3 Kalman Filter

In 1960 Rudolph E. Kalman published his famous paper describing a recursive solution to the discrete-data linear filtering problem, and the method was named Kalman filter [4]. Figure (3.3), reproduced from [38], illustrates the application context in which the Kalman filter is used. The figure illustrates that a physical system (e.g. production of an oil-reservoir) is driven by a set of external controls or inputs, where the outputs of the system are typically evaluated by different measuring devices. The knowledge of the behavior of the system is then entirely given by the inputs and the measured outputs. The uncertainties and the errors in the process, such as the measuring device noise and the system errors, are then disclosed in the observations. Based on the control inputs and the observations, the aim of the filter is to find an estimate of the state of the system that optimizes a given criteria.

The Kalman filter can be viewed as an optimal recursive data processing algorithm, and it is a technique for assessing the posterior distribution, $g(\boldsymbol{\psi}_k|\boldsymbol{d}_{k:1})$, based on the likelihood pdf, $g(\boldsymbol{d}_k|\boldsymbol{\psi}_k)$, and the prior pdf of the model variables, $g(\boldsymbol{\psi}_k|\boldsymbol{d}_{k-1:1})$. Assuming linear dynamics and Gaussian statistics the Kalman filter is a statistical consistent method, and provides the best linear unbiased estimate (BLUE) of the posterior mean $\mathrm{E}(\boldsymbol{\psi}_k|\boldsymbol{d}_{k:1})$ and the posterior covariance $\mathrm{Cov}(\boldsymbol{\psi}_k|\boldsymbol{d}_{k:1})$. Since the likelihood and the prior distributions are assumed Gaussian, the posterior is also Gaussian and fully described by the mean and covariance, and the Kalman filter utilizes the dynamical equations to evolve the most probable model state and the error covariance matrix in time. The Kalman filter has been the subject of extensive research and application, and a more complete introductory discussion of the Kalman filter can be found in [38, 39] and at the web page: www.cs.unc.edu/∼whelch/kalman.

In this section the derivation of the Kalman filter, the Kalman smoother and

the extended Kalman filter are outlined, and we present a sampling algorithm from the Kalman filter distribution which is closely related to the sampling procedure used in the ensemble Kalman filter algorithm.

3.3.1 Derivation of the Kalman Filter

In the standard state space form, observations $\boldsymbol{d}_k$ are related to the model variables $\boldsymbol{\psi}_k$ by a linear observation equation

$$\boldsymbol{d}_k = \boldsymbol{H}_k \boldsymbol{\psi}_k + \boldsymbol{\epsilon}_k^o \qquad \boldsymbol{\epsilon}_k^o \sim \mathcal{N}_{m_k}(\boldsymbol{0}, \boldsymbol{P}_{\epsilon_k^o}), \tag{3.11}$$

where $\boldsymbol{H}_k$ is the observation operator, and $\boldsymbol{\epsilon}_k^o$ is a Gaussian observation noise process where $\boldsymbol{P}_{\epsilon_k^o}$ is the time-varying observation covariance matrix. The model state evolution is defined as

$$\boldsymbol{\psi}_k = \boldsymbol{F}_k \boldsymbol{\psi}_{k-1} + \boldsymbol{\epsilon}_k^m \qquad \boldsymbol{\epsilon}_k^m \sim \mathcal{N}_q(\boldsymbol{0}, \boldsymbol{P}_{\epsilon_k^m}), \tag{3.12}$$

where $\boldsymbol{F}_k$ is the linear model operator, and $\boldsymbol{\epsilon}_k^m$ is a Gaussian model noise sequence with error covariance matrix $\boldsymbol{P}_{\epsilon_k^m}$. The initial prior distribution has expectation $\mathrm{E}(\boldsymbol{\psi}_0) = \boldsymbol{\psi}_0^{init}$ and covariance $\mathrm{Cov}(\boldsymbol{\psi}_0) = \boldsymbol{P}_{\psi_0}^{init}$ and is assumed Gaussian

$$\boldsymbol{\psi}_0 \sim \mathcal{N}_q(\boldsymbol{\psi}_0^{init}, \boldsymbol{P}_{\psi_0}^{init}). \tag{3.13}$$

The mean and the covariance of the model variables are fully specified by assuming that $\boldsymbol{\epsilon}_k^o$, $\boldsymbol{\epsilon}_k^m$ and $\boldsymbol{\psi}_0$ are mutually independent. Here, we have also assumed that the noise processes have zero mean, although this need not to be the case in general.

Given the measurements $\boldsymbol{d}_{k:1}$, estimation of $\boldsymbol{\psi}_k$ is the primary objective, and following distributions will be of interest

- Filtering, i.e finding $g(\boldsymbol{\psi}_k | \boldsymbol{d}_{k:1})$
- Prediction, i.e finding $g(\boldsymbol{\psi}_k | \boldsymbol{d}_{j:1}), \;\; j < k$
- Smoothing, i.e finding $g(\boldsymbol{\psi}_k | \boldsymbol{d}_{l:1}), \;\; k < l$

This provides the key to perform statistical inference.

The derivation of the Kalman filter is as follows. Assume that the posterior distribution at time t_{k-1}, $g(\boldsymbol{\psi}_{k-1} | \boldsymbol{d}_{k-1:1})$, is known,

$$\boldsymbol{\psi}_{k-1} | \boldsymbol{d}_{k-1:1} \sim \mathcal{N}_q(\boldsymbol{\psi}_{k-1}^a, \boldsymbol{P}_{\psi_{k-1}}^a), \tag{3.14}$$

where the conditional expectation and the error covariance matrix is given as

$$\boldsymbol{\psi}_{k-1}^{a} = \mathrm{E}(\boldsymbol{\psi}_{k-1}|\boldsymbol{d}_{k-1:1}), \tag{3.15}$$

$$\boldsymbol{P}_{\psi_{k-1}}^{a} = \mathrm{E}((\boldsymbol{\psi}_{k-1} - \boldsymbol{\psi}_{k-1}^{a})(\boldsymbol{\psi}_{k-1} - \boldsymbol{\psi}_{k-1}^{a})^{T}|\boldsymbol{d}_{k-1:1}). \tag{3.16}$$

Combining Equation (3.14) and the system equation (3.12) yields the prior distribution at time t_k, $g(\boldsymbol{\psi}_k|\boldsymbol{d}_{k-1:1})$,

$$\boldsymbol{\psi}_k|\boldsymbol{d}_{k-1:1} \sim \mathcal{N}_q(\boldsymbol{\psi}_k^f, \boldsymbol{P}_{\psi_k}^f), \tag{3.17}$$

where the prior mean, $\boldsymbol{\psi}_k^f$, and the prior covariance, $\boldsymbol{P}_{\psi_k}^f$, are defined by

$$\boldsymbol{\psi}_k^f = \mathrm{E}(\boldsymbol{\psi}_k|\boldsymbol{d}_{k-1:1}), \tag{3.18}$$

$$\boldsymbol{P}_{\psi_k}^f = \mathrm{E}((\boldsymbol{\psi}_k - \boldsymbol{\psi}_k^f)(\boldsymbol{\psi}_k - \boldsymbol{\psi}_k^f)^{T}|\boldsymbol{d}_{k-1:1}). \tag{3.19}$$

For the linear state space model in Equation (3.12) the prior moments are given as

$$\boldsymbol{\psi}_k^f = \boldsymbol{F}_k \boldsymbol{\psi}_{k-1}^a, \tag{3.20}$$

$$\boldsymbol{P}_{\psi_k}^f = \boldsymbol{F}_k \boldsymbol{P}_{\psi_{k-1}}^a \boldsymbol{F}_k^T + \boldsymbol{P}_{\epsilon_k^m}. \tag{3.21}$$

Based on the prior distribution in Equation (3.17), the one-step forecast distribution of $\boldsymbol{d}_k$ given $\boldsymbol{d}_{k-1:1}$ is derived using the observation equation (3.11)

$$\boldsymbol{d}_k|\boldsymbol{d}_{k-1:1} \sim \mathcal{N}_{m_k}(\boldsymbol{d}_k^f, \boldsymbol{P}_{d_k}), \tag{3.22}$$

where the one-step forecast mean and the one-step forecast covariance are

$$\boldsymbol{d}_k^f = \boldsymbol{H}_k \boldsymbol{\psi}_k^f, \tag{3.23}$$

$$\boldsymbol{P}_{d_k} = \boldsymbol{H}_k \boldsymbol{P}_{\psi_k}^f \boldsymbol{H}_k^T + \boldsymbol{P}_{\epsilon_k^o}. \tag{3.24}$$

The derivation of the posterior distribution at time t_k, $g(\boldsymbol{\psi}_k|\boldsymbol{d}_{k:1})$, is found from the joint distribution of $g(\boldsymbol{\psi}_k, \boldsymbol{d}_k|\boldsymbol{d}_{k-1:1})$. The joint distribution is established by observing from Equation (3.11) that

$$\mathrm{Cov}(\boldsymbol{\psi}_k, \boldsymbol{d}_k|\boldsymbol{d}_{k-1:1}) = \mathrm{Cov}(\boldsymbol{\psi}_k, \boldsymbol{H}_k\boldsymbol{\psi}_k + \boldsymbol{\epsilon}_k^o|\boldsymbol{d}_{k-1:1}) = \boldsymbol{P}_{\psi_k}^f \boldsymbol{H}_k^T, \tag{3.25}$$

and hence the joint distribution can be expressed as

$$\begin{pmatrix} \boldsymbol{\psi}_k \\ \boldsymbol{d}_k \end{pmatrix} \Bigg| \boldsymbol{d}_{k-1:1} \sim \mathcal{N}_{q+m_k}\left(\begin{pmatrix} \boldsymbol{\psi}_k^f \\ \boldsymbol{d}_k^f \end{pmatrix}, \begin{bmatrix} \boldsymbol{P}_{\psi_k}^f & \boldsymbol{P}_{\psi_k}^f \boldsymbol{H}_k^T \\ \boldsymbol{H}_k \boldsymbol{P}_{\psi_k}^f & \boldsymbol{P}_{d_k} \end{bmatrix} \right). \tag{3.26}$$

Recall that when we have a jointly Gaussian distribution $g(\boldsymbol{z}_1, \boldsymbol{z}_2)$:

$$\begin{pmatrix} \boldsymbol{z}_1 \\ \boldsymbol{z}_2 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \begin{bmatrix} \boldsymbol{P}_1 & \boldsymbol{P}_{12} \\ \boldsymbol{P}_{21} & \boldsymbol{P}_2 \end{bmatrix} \right), \tag{3.27}$$

then the conditional probability distribution $g(\boldsymbol{z}_1|\boldsymbol{z}_2)$ is also Gaussian with mean $\boldsymbol{\mu}_{1|2}$ and covariance $\boldsymbol{P}_{1|2}$, where [21]

$$\boldsymbol{\mu}_{1|2} = \boldsymbol{\mu}_1 + \boldsymbol{P}_{12}(\boldsymbol{P}_2)^{-1}(\boldsymbol{z}_2 - \boldsymbol{\mu}_2), \tag{3.28}$$

$$\boldsymbol{P}_{1|2} = \boldsymbol{P}_1 - \boldsymbol{P}_{12}(\boldsymbol{P}_2)^{-1}\boldsymbol{P}_{21}. \tag{3.29}$$

By applying these formulas to the joint distribution in Equation (3.26) the posterior distribution, $g(\boldsymbol{\psi}_k|\boldsymbol{d}_{k:1})$, can be expressed as

$$(\boldsymbol{\psi}_k|\boldsymbol{d}_{k:1}) \sim \mathcal{N}_q(\boldsymbol{\psi}_k^a, \boldsymbol{P}_{\psi_k}^a). \tag{3.30}$$

Here the analysis mean and covariance are given as

$$\boldsymbol{\psi}_k^a = \boldsymbol{\psi}_k^f + \boldsymbol{K}_k(\boldsymbol{d}_k - \boldsymbol{H}_k\boldsymbol{\psi}_k^f), \tag{3.31}$$

$$\boldsymbol{P}_{\psi_k}^a = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)\boldsymbol{P}_{\psi_k}^f, \tag{3.32}$$

where the Kalman gain matrix, $\boldsymbol{K}_k$, is given by

$$\boldsymbol{K}_k = \boldsymbol{P}_{\psi_k}^f \boldsymbol{H}_k^T(\boldsymbol{H}_k\boldsymbol{P}_{\psi_k}^f\boldsymbol{H}_k^T + \boldsymbol{P}_{\epsilon_k^o})^{-1}. \tag{3.33}$$

Notice that for the case with a linear dynamical model and Gaussian distributions the linear variance minimizing analysis or map estimate also can be found from the minimum of

$$\boldsymbol{\psi}_k^a = argmin\{\boldsymbol{J}(\boldsymbol{\psi}_k)\}, \tag{3.34}$$

where the objective function, $\boldsymbol{J}(\boldsymbol{\psi}_k)$, is defined as

$$\begin{aligned} \boldsymbol{J}(\boldsymbol{\psi}_k) = &\frac{1}{2}\left((\boldsymbol{d}_k - \boldsymbol{H}\boldsymbol{\psi}_k)^T(\boldsymbol{P}_{\epsilon_k^o})^{-1}(\boldsymbol{d}_k - \boldsymbol{H}\boldsymbol{\psi}_k)\right) \\ &+ \frac{1}{2}\left((\boldsymbol{\psi}_k - \boldsymbol{\psi}_k^f)^T(\boldsymbol{P}_{\psi_k}^f)^{-1}(\boldsymbol{\psi}_k - \boldsymbol{\psi}_k^f)\right). \end{aligned} \tag{3.35}$$

Taking the derivative of $\boldsymbol{J}(\boldsymbol{\psi}_k)$ with respect to $\boldsymbol{\psi}_k$, and let it equal to zero, the best estimate of $\boldsymbol{\psi}_k$ is obtained and equal to the result in Equation (3.31).

Based on the results above the forecast distribution, $g(\boldsymbol{\psi}_k|\boldsymbol{d}_{j:1})$, where $j < k$, can be derived in a similar manner, and it can be shown that $g(\boldsymbol{\psi}_k|\boldsymbol{d}_{j:1})$ also will be Gaussian distributed [40].

3.3.2 Sampling from the Kalman Filter Distribution

From the Kalman filter theory the state space models are assumed linear with Gaussian statistics and the posterior distribution, $g(\boldsymbol{\psi}_k|\boldsymbol{d}_{k:1})$, is then Gaussian and subject to analytical evaluation. The objective is to sample from the posterior pdf $g(\boldsymbol{\psi}_k|\boldsymbol{d}_{k:1})$, see Equation (3.30), and to generate a sample from this distribution, different sampling strategies can be applied. Here, we present a standard algorithm based on conditional simulation with kriging, and the algorithm is as follows (see e.g. [41]):

Algorithm 3.3.2

1. Generate a sample from the posterior distribution $g(\boldsymbol{\psi}_{k-1}|\boldsymbol{d}_{k-1:1})$:
$$\boldsymbol{\psi}_{k-1}^{a,*} \sim \mathcal{N}_q(\boldsymbol{\psi}_{k-1}^{a}, \boldsymbol{P}_{\psi_{k-1}}^{a})$$
2. Forecast the sample $\boldsymbol{\psi}_{k-1}^{a,*}$ using Equation (3.12):
$$\boldsymbol{\psi}_{k}^{f,*} = \boldsymbol{F}_k\boldsymbol{\psi}_{k-1}^{a,*} + \boldsymbol{\epsilon}_k^{m,*}, \quad \boldsymbol{\epsilon}_k^{m,*} \sim \mathcal{N}_q(0, \boldsymbol{P}_{\epsilon_k^m})$$
   - $\boldsymbol{\psi}_k^{f,*}$ represent a sample from the prior distribution $g(\boldsymbol{\psi}_k|\boldsymbol{d}_{k-1:1})$.
3. Generate a sample from the measurement distribution:
$$\boldsymbol{d}_k^* \sim \mathcal{N}_{m_k}(\boldsymbol{d}_k, \boldsymbol{P}_{\epsilon_k^o})$$
4. Use the kriging formula to generate a sample from the posterior distribution $g(\boldsymbol{\psi_k}|\boldsymbol{d}_{k:1})$:
$$\boldsymbol{\psi}_k^{a,*} = \boldsymbol{\psi}_k^{f,*} + \boldsymbol{P}_{\psi_k}^{f}\boldsymbol{H}_k^T(\boldsymbol{H}_k\boldsymbol{P}_{\psi_k}^{f}\boldsymbol{H}_k^T + \boldsymbol{P}_{\epsilon_k^o})^{-1}(\boldsymbol{d}_k^* - \boldsymbol{H}_k\boldsymbol{\psi}_k^{f,*})$$

To prove that the algorithm provides samples from the correct posterior pdf, the expectation and covariance need to be identical with the moments given in Equations (3.31)-(3.32). Let now $(\boldsymbol{\psi}_{k-1}^{a,s}, \boldsymbol{\epsilon}_k^{m,s}, \boldsymbol{\psi}_k^{f,s}, \boldsymbol{d}_k^s)$ represent the stochastic interpretation of $(\boldsymbol{\psi}_{k-1}^{a,*}, \boldsymbol{\psi}_k^{f,*}, \boldsymbol{\epsilon}_k^{m,*}, \boldsymbol{d}_k^*)$.

To show that $\boldsymbol{\psi}_k^{f,*}$ represents a sample from the prior distribution $g(\boldsymbol{\psi}_k|\boldsymbol{d}_{k-1:1})$ given in Equation (3.17), the first and second order moments have to be equal to the moments in Equations (3.20)-(3.21). The expectation and the covariance are given as

$$\begin{aligned}
\mathrm{E}(\boldsymbol{\psi}_k^{f,s}) &= \boldsymbol{F}_k\mathrm{E}(\boldsymbol{\psi}_{k-1}^{a,s}) + \mathrm{E}(\boldsymbol{\epsilon}_k^{m,s}) \\
&= \boldsymbol{F}_k\boldsymbol{\psi}_{k-1}^{a} \\
\mathrm{Cov}(\boldsymbol{\psi}_k^{f,s}) &= \boldsymbol{F}_k\mathrm{Cov}(\boldsymbol{\psi}_{k-1}^{a,s})\boldsymbol{F}_k^T + \mathrm{Cov}(\boldsymbol{\epsilon}_k^{m,s}) \\
&= \boldsymbol{F}_k\boldsymbol{P}_{\psi_{k-1}}^{a}\boldsymbol{F}_k^T + \boldsymbol{P}_{\epsilon_k^m},
\end{aligned}$$

where $(\boldsymbol{\psi}_{k-1}^{a,s}, \boldsymbol{\epsilon}_k^{m,s})$ are assumed independent. Hence, it is shown that $\boldsymbol{\psi}_k^{f,*}$ will be a sample from the correct Gaussian prior distribution.

The analysis step can now be written as

$$\boldsymbol{\psi}_k^{a,s} = \boldsymbol{\psi}_k^{f,s} + \boldsymbol{P}_{\psi_k}^f \boldsymbol{H}_k^T (\boldsymbol{H}_k \boldsymbol{P}_{\psi_k}^f \boldsymbol{H}_k^T + \boldsymbol{P}_{\epsilon_k^o})^{-1} (\boldsymbol{d}_k^s - \boldsymbol{H}_k \boldsymbol{\psi}_k^{f,s}),$$

where $(\boldsymbol{\psi}_k^{f,s}, \boldsymbol{d}_k^s)$ are independent with distributions as specified in the algorithm. Due to $(\boldsymbol{\psi}_k^{f,s}, \boldsymbol{d}_k^s)$ being stochastic, $\boldsymbol{\psi}_k^{a,s}$ will be stochastic, and the expectation and the covariance are given as

$$\begin{aligned}
\mathrm{E}(\boldsymbol{\psi}_k^{a,s}) &= \mathrm{E}(\boldsymbol{\psi}_k^{f,s}) + \boldsymbol{P}_{\psi_k}^f \boldsymbol{H}_k^T (\boldsymbol{H}_k \boldsymbol{P}_{\psi_k}^f \boldsymbol{H}_k^T + \boldsymbol{P}_{\epsilon_k^o})^{-1} (\mathrm{E}(\boldsymbol{d}_k^s) - \boldsymbol{H}_k \mathrm{E}(\boldsymbol{\psi}_k^{f,s})) \\
&= \boldsymbol{\psi}_k^f + \boldsymbol{P}_{\psi_k}^f \boldsymbol{H}_k^T (\boldsymbol{H}_k \boldsymbol{P}_{\psi_k}^f \boldsymbol{H}_k^T + \boldsymbol{P}_{\epsilon_k^o})^{-1} (\boldsymbol{d}_k - \boldsymbol{H}_k \boldsymbol{\psi}_k^f) \\
\mathrm{Cov}(\boldsymbol{\psi}_k^{a,s}) &= \mathrm{Cov}(\boldsymbol{\psi}_k^{f,s}) + \boldsymbol{P}_{\psi_k}^f \boldsymbol{H}_k^T (\boldsymbol{H}_k \boldsymbol{P}_{\psi_k}^f \boldsymbol{H}_k^T + \boldsymbol{P}_{\epsilon_k^o})^{-1} \\
&\quad \cdot (\mathrm{Cov}(\boldsymbol{d}_k^s) + \boldsymbol{H}_k \mathrm{Cov}(\boldsymbol{\psi}_k^{f,s}) \boldsymbol{H}_k^T) (\boldsymbol{H}_k \boldsymbol{P}_{\psi_k}^f \boldsymbol{H}_k^T + \boldsymbol{P}_{\epsilon_k^o})^{-1} \boldsymbol{H}_k \boldsymbol{P}_{\psi_k}^f \\
&\quad - 2 \boldsymbol{P}_{\psi_k}^f \boldsymbol{H}_k^T (\boldsymbol{H}_k \boldsymbol{P}_{\psi_k}^f \boldsymbol{H}_k^T + \boldsymbol{P}_{\epsilon_k^o})^{-1} \boldsymbol{H}_k \mathrm{Cov}(\boldsymbol{\psi}_k^{f,s}) \\
&= \boldsymbol{P}_{\psi_k}^f - \boldsymbol{P}_{\psi_k}^f \boldsymbol{H}_k^T (\boldsymbol{H}_k \boldsymbol{P}_{\psi_k}^f \boldsymbol{H}_k^T + \boldsymbol{P}_{\epsilon_k^o})^{-1} \boldsymbol{H}_k \boldsymbol{P}_{\psi_k}^f.
\end{aligned}$$

Hence, it is proven that the algorithm samples correctly.

3.3.3 Kalman Smoothing

In Kalman smoothing [40] the main objective is to find the conditional distribution of $\boldsymbol{\psi}_k$ given observations $\boldsymbol{d}_{l:1}$, where time $t_k \leq t_l$. Using a similar approach as for the Kalman filter, the posterior distribution, $g(\boldsymbol{\psi}_k | \boldsymbol{d}_{l:1})$, can be shown to be Gaussian

$$(\boldsymbol{\psi}_k | \boldsymbol{d}_{l:1}) \sim \mathcal{N}_q(\boldsymbol{\psi}_{k|l}^a, \boldsymbol{P}_{\psi_{k|l}}^a), \tag{3.36}$$

where $\boldsymbol{\psi}_{k|l}^a$ and $\boldsymbol{P}_{\psi_{k|l}}^a$ are respectively the smoother mean and the smoother covariance.

The Kalman smoother consists of backwards recursions for time $t_l, \ldots, t_1$, and the smoother moments can be derived as

$$\boldsymbol{\psi}_{k|l}^a = \mathrm{E}(\boldsymbol{\psi}_k | \boldsymbol{d}_{l:1}) = \boldsymbol{\psi}_k^a + \boldsymbol{B}_k (\boldsymbol{\psi}_{k+1|l}^a - \boldsymbol{\psi}_{k+1}^f), \tag{3.37}$$

and

$$\boldsymbol{P}_{\psi_{k|l}}^a = \mathrm{Cov}(\boldsymbol{\psi}_k | \boldsymbol{d}_{l:1}) = \boldsymbol{P}_{\psi_k}^a + \boldsymbol{B}_k (\boldsymbol{P}_{\psi_{k+1|l}}^a - \boldsymbol{P}_{\psi_{k+1}}^f) \boldsymbol{B}_k^T, \tag{3.38}$$

with

$$\boldsymbol{B}_k = \boldsymbol{P}^a_{\psi_k} \boldsymbol{F}^T_{k+1} (\boldsymbol{P}^f_{\psi_{k+1}})^{-1}. \tag{3.39}$$

Here at each update step, the smoothing estimate $\boldsymbol{\psi}^a_{k|l}$ is achieved from the Kalman filter estimate $\boldsymbol{\psi}^a_k$ by adding the weighted difference between the smoothing estimate $\boldsymbol{\psi}^a_{k+1|l}$ of the previous step and the prediction estimate $\boldsymbol{\psi}^f_{k+1}$. From Equations (3.37) and (3.38) it is clear that the implementation of the Kalman smoother moments requires that the analysis and forecast moments $\boldsymbol{\psi}^a_k$, $\boldsymbol{P}^a_{\psi_k}$, $\boldsymbol{\psi}^f_{k+1}$ and $\boldsymbol{P}^f_{\psi_{k+1}}$ are computed from the Kalman filter.

3.3.4 Extended Kalman Filter

Consider the non-linear state space model

$$\boldsymbol{\psi}_k = \boldsymbol{f}_k(\boldsymbol{\psi}_{k-1}) + \boldsymbol{\epsilon}^m_k \qquad \boldsymbol{\epsilon}^m_k \sim \mathcal{N}_q(\boldsymbol{0}, \boldsymbol{P}_{\epsilon^m_k}), \tag{3.40}$$

$$\boldsymbol{d}_k = \boldsymbol{h}_k(\boldsymbol{\psi}_k) + \boldsymbol{\epsilon}^o_k \qquad \boldsymbol{\epsilon}^o_k \sim \mathcal{N}_q(\boldsymbol{0}, \boldsymbol{P}_{\epsilon^o_k}), \tag{3.41}$$

where $\boldsymbol{f}_k : \Re^q \to \Re^q$ is a non-linear model operator with model errors ϵ^m_k, and $\boldsymbol{h}_k : \Re^q \to \Re^{m_k}$ is the non-linear measurement operator, relating the model variable $\boldsymbol{\psi}_k$ to the observations $\boldsymbol{d}_k$ allowing for measurement errors $\boldsymbol{\epsilon}^o_k$.

In a non-linear state space model the model variables undergoing nonlinear transformations, Equations (3.40) and (3.41), and as a result the posteriori distribution $g(\boldsymbol{\psi}_k|\boldsymbol{d}_{k:1})$ can no longer be explicitly described by a Gaussian distribution. To solve the estimation problem a non optimal approach, in the frame of linear filters, can be used, and it is referred to as an extended Kalman filter (EKF) [38, 42, 43]. The EKF gives an approximation of the optimal minimum mean-square estimate by linearization, and at each step the non-linear dynamics are linearized around the last consecutive predicted and filtered state estimates. Based on the linearized dynamics the EKF applies the Kalman filter to obtain estimates of first and second order moments of the posterior distribution $g(\boldsymbol{\psi}_k|\boldsymbol{d}_k)$. To achieve a valid solution, this linearization should be a good approximation of the non-linear state space model in all the uncertainty domain related to the model variable estimate.

Assume that the mean, $\boldsymbol{\psi}^a_{k-1}$, and the covariance, $\boldsymbol{P}^a_{\psi_{k-1}}$, of the posterior distribution $g(\boldsymbol{\psi}_{k-1}|\boldsymbol{d}_{k-1:1})$ are known. One iteration of the EKF is composed by the

following steps:

$$
\begin{aligned}
\boldsymbol{F}_k &= \left[\frac{\partial \boldsymbol{f}_k(\boldsymbol{\psi}_{k-1})}{\partial \boldsymbol{\psi}_{k-1}}\right]_{\boldsymbol{\psi}_{k-1}=\boldsymbol{\psi}^a_{k-1}} \\
\boldsymbol{\psi}^f_k &= \boldsymbol{f}_k(\boldsymbol{\psi}^a_{k-1}) \\
\boldsymbol{P}^f_{\psi_k} &= \boldsymbol{F}_k \boldsymbol{P}^a_{k-1} \boldsymbol{F}^T_k + \boldsymbol{P}_{\epsilon^m_k} \\
\boldsymbol{H}_k &= \left[\frac{\partial \boldsymbol{h}_k(\boldsymbol{\psi}_k)}{\partial \boldsymbol{\psi}_k}\right]_{\boldsymbol{\psi}_k=\boldsymbol{\psi}^f_k} \\
\boldsymbol{K}_k &= \boldsymbol{P}^f_{\psi_k} \boldsymbol{H}^T_k (\boldsymbol{H}_k \boldsymbol{P}^f_{\psi_k} \boldsymbol{H}^T_k + \boldsymbol{P}_{\epsilon^o_k})^{-1} \\
\boldsymbol{\psi}^a_k &= \boldsymbol{\psi}^f_k + \boldsymbol{K}_k(\boldsymbol{d}_k - \boldsymbol{h}_k(\boldsymbol{\psi}^f_k)) \\
\boldsymbol{P}^a_{\psi_k} &= (\boldsymbol{I} - \boldsymbol{K}_k \boldsymbol{H}_k) \boldsymbol{P}^f_{\psi_k}.
\end{aligned}
$$

Notice that the EKF is not an optimal filter, but rather it is constructed on the basis of a set of approximations. Thus the matrix $\boldsymbol{P}^f_{\psi_k}$ and $\boldsymbol{P}^a_{\psi_k}$ do not represent the true covariance of the estimate of the model variables.

Chapter 4

Ensemble Kalman Filter (EnKF)

The Ensemble Kalman filter (EnKF) is a Monte Carlo type sequential Bayesian inversion method, and was first introduced by Geir Evensen in 1994 for data assimilation of non-linear ocean models [44]. Since its introduction the EnKF has been applied and examined in a number of studies related to atmospheric and oceanographic systems [45, 46]. An example is the data assimilation of Topaz, (see e.g. [3] and the web page: topaz.nersec.no), which is a high dimensional operational ocean prediction system for the Atlantic and Arctic oceans, where the state vector consist of almost 80 million variables.

Recently, several investigations have also shown promising results of the EnKF method for continuous updating of reservoir simulation models, as an alternative to traditional history-matching. This was first proposed by Nævdal et al. [47] who estimated the reservoir permeability on a simplified reservoir model by using the EnKF methodology. Other publications and studies discussing the estimation problem within petroleum reservoir applications using EnKF can be found in [3, 34, 48, 49, 50, 51].

EnKF has a promising potential in reservoir and seismic characterization, since it has a simple conceptual formulation and can be implemented relatively easy for reservoir topics. The method requires no integrations backward in time, and avoids computation of the gradient operator or adjoint equations. An important advantage is that any reservoir simulator and seismic forward model can be used in the EnKF history matching process without additional amount of work.

Lately, the growth in deployment of permanent sea bed sensors for seismic monitoring has increased rapidly, and as a result seismic datasets become available with a higher time frequency. Thus, in a reservoir problem, both production data and seismic data may be available at different time steps during the oil production history. The EnKF is a method for sequentially updating estimates of model variables and for assimilating different types of data and, therefore, well suited to solve such problems.

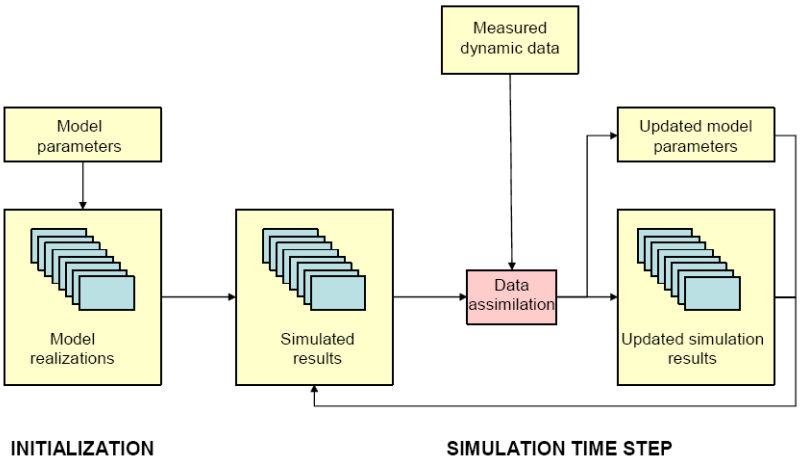


Figure 4.1: Illustration of ensemble Kalman filter methodology

The EnKF was introduced as a Monte Carlo alternative to the traditional EKF. Evensen [44] has shown that a Monte Carlo approach can be utilized to solve an equation for the time evolution of the pdf of the model variable, as a stochastic alternative to using the approximate error covariance equation in the EKF. The general evolution of the model variable can be written as stochastic differential equation

$$d\boldsymbol{\psi} = \boldsymbol{f}(\boldsymbol{\psi})dt + \boldsymbol{w}(\boldsymbol{\psi})d\boldsymbol{\theta}. \tag{4.1}$$

This equation implies that a model variable increment $d\boldsymbol{\psi}$ can be described by the forward model over the time increment considered, $\boldsymbol{f}(\boldsymbol{\psi})dt$, which in addition is influenced by the term $\boldsymbol{w}(\boldsymbol{\psi})d\boldsymbol{\theta}$, called the stochastic forcing and representing the model errors. If the dynamical model is written as in Equation (4.1), one can derive the Fokker-Plank or Kolmogorov's equation for the time evolution of the pdf, which includes all the information about the prediction error statistics. A detailed description is given in Jazwinski [52].

The EnKF applies a Monte Carlo method to solve the Fokker-Plank equation and is based on the use of an ensemble representation for the pdfs of interest. The EnKF is a sequential data assimilation method, and provides an approximate solution to the combined parameter and state estimation problem [3]. The result is an ensemble of solutions approximating the posterior pdf. In the EnKF each ensemble member is updated according to the traditional scheme of the Kalman filter, but the main difference is that the model variable error covariance matrix in the EnKF is estimated from an ensemble of typical short-range, non-linear forecasts. At the

analysis times, the members of the ensemble are updated so that the perturbations of the ensemble approximate a random sample from the analysis error distribution. The mean and the spread of the ensemble of the model variables at a given time step will then reflect both an estimate of the true state and an uncertainty of that estimate. Each member of the ensemble will be evolved independently forward in time, until new measurements from the system are available. At that point the mean and the covariance from the ensemble are used to generate the new state estimate and the reduced uncertainty introduced by the measurements. Here, the EnKF provides a linear update to a non-linear forecast ensemble and can thus be viewed as an intermediate step between the Kalman filter and particle filters, see for example Bertino et al. [53]. An illustration of the EnKF method is presented in Figure 4.1.

Evensen [3, 54] provides both a comprehensive review of the progress on the EnKF since its introduction, and a detailed interpretation and description of the EnKF as a data assimilation method.

4.1 Derivation of the EnKF Algorithm

Let the non-linear state space model be represented by the Equations (3.40) and (3.41). In the EnKF setting the nonlinear state space model is reformulated, so that the observation equation can be treated as linear [54]. To obtain a linear observation equation it is possible to augment the model variable with a diagnostic variable which is the model prediction of the measurement. The linear observation equation can then be written as

$$\boldsymbol{d}_k = \tilde{\boldsymbol{H}}_k \tilde{\boldsymbol{\psi}}_k + \boldsymbol{\epsilon}_k^o, \tag{4.2}$$

where $\tilde{\boldsymbol{\psi}}_k = (\boldsymbol{\psi}_k, \boldsymbol{h}_k(\boldsymbol{\psi}_k))^T$ and $\tilde{\boldsymbol{H}}_k$ being the linear measurement operator that picks data from $\tilde{\boldsymbol{\psi}}_k$. For simplicity we omit the augmented notation and let $\boldsymbol{\psi}$ from now include the parameters, the states and the model prediction of the measurements. The derivation of the EnKF algorithm is as follows.

Define the matrix holding the ensemble members, $\boldsymbol{\psi}_k \in \Re^q$, at time t_k, as

$$\boldsymbol{\Psi}_k = \boldsymbol{\Psi}(\boldsymbol{x}, t_k) = (\boldsymbol{\psi}_k^1, \boldsymbol{\psi}_k^2, \cdots, \boldsymbol{\psi}_k^N) \in \Re^{q \times N}, \tag{4.3}$$

where N is the ensemble size. The ensemble mean can be defined as

$$\bar{\boldsymbol{\Psi}}_k = \boldsymbol{\Psi}_k \mathbf{1}_N, \tag{4.4}$$

where $\mathbf{1}_N \in \Re^{N \times N}$ is the matrix where each element is equal to $1/N$. Based on Equations (4.3) and (4.4) the ensemble perturbation matrix,$\boldsymbol{\Psi}'_k$, can be expressed

as

$$\boldsymbol{\Psi}'_k = \boldsymbol{\Psi}_k - \bar{\boldsymbol{\Psi}}_k = \boldsymbol{\Psi}_k(\boldsymbol{I} - \mathbf{1}_N). \tag{4.5}$$

The ensemble covariances matrix, $\boldsymbol{P}^e_{\psi_k} \in \Re^{q\times q}$, can then be defined as

$$\boldsymbol{P}^e_{\psi_k} = \frac{\boldsymbol{\Psi}'_k(\boldsymbol{\Psi}'_k)^T}{N-1}. \tag{4.6}$$

Given a vector of observations, $\boldsymbol{d}_k \in \Re^{m_k}$, the N vectors of perturbed observations can be expressed as

$$\boldsymbol{d}^i_k = \boldsymbol{d}_k + \boldsymbol{\epsilon}^{o,i}_k, \quad i = 1, 2\cdots, N, \tag{4.7}$$

where the perturbed observations can be stored in the columns of a matrix

$$\boldsymbol{D}_k = (\boldsymbol{d}^1_k, \boldsymbol{d}^2_k, \cdots, \boldsymbol{d}^N_k) \in \Re^{m_k\times N}. \tag{4.8}$$

The ensemble of observation perturbations can be represented in the columns of the matrix

$$\boldsymbol{E}_k = (\boldsymbol{\epsilon}^{o,1}_k, \boldsymbol{\epsilon}^{o,2}_k, \cdots, \boldsymbol{\epsilon}^{o,N}_k), \tag{4.9}$$

and the ensemble representation of the error covariance matrix of the observations becomes

$$\boldsymbol{P}^e_{\epsilon^o_k} = \frac{\boldsymbol{E}_k\boldsymbol{E}^T_k}{N-1}. \tag{4.10}$$

The EnKF sampling can be viewed as a conditional simulation with kriging, and the algorithm is as follows:

Algorithm 4.1

0. Generate N independent samples, $\boldsymbol{\psi}^{a,i}_{k-1}$, from $g(\boldsymbol{\psi}_{k-1}|\boldsymbol{d}_{k-1:1})$
1. Forecast each of the samples, $\boldsymbol{\psi}^{a,i}_{k-1}$ using Equation (3.40):

$$\boldsymbol{\psi}^{f,i}_k = \boldsymbol{f}_k(\boldsymbol{\psi}^{a,i}_{k-1}) + \boldsymbol{\epsilon}^{m,i}_k, \quad \boldsymbol{\epsilon}^{m,i}_k \sim \mathcal{N}_q(\mathbf{0}, \boldsymbol{P}_{\epsilon^m_k})$$

   - $\boldsymbol{\psi}^{f,i}_k$ represent a sample from the prior distribution $g(\boldsymbol{\psi}_k|\boldsymbol{d}_{k-1:1})$.

2. Use Monte Carlo integration to estimate the first and second order moments from the prior pdf using Equations (4.4) and (4.6)

3. Generate N samples from the measurement distribution:

$$\boldsymbol{d}_k^i \sim \mathcal{N}_{m_k}(\boldsymbol{d}_k, \boldsymbol{P}_{\epsilon_k^o}) \;\; \forall \;\; i = 1:N$$

4. Use the kriging formula to generate a sample from the posterior distribution $g(\boldsymbol{\psi}_k|\boldsymbol{d}_{k:1})$:

$$\boldsymbol{\psi}_k^{a,i} = \boldsymbol{\psi}_k^{f,i} + \boldsymbol{P}_{\psi_k}^{e,f}\boldsymbol{H}_k^T(\boldsymbol{H}_k\boldsymbol{P}_{\psi_k}^{e,f}\boldsymbol{H}_k^T + \boldsymbol{P}_{\epsilon_k^o})^{-1}(\boldsymbol{d}_k^i - \boldsymbol{H}_k\boldsymbol{\psi}_k^{f,i})$$

- where $\boldsymbol{\psi}_k^{a,i}$ represent a sample from the filter distribution $g(\boldsymbol{\psi}_k|\boldsymbol{d}_{k:1})$.

The ensemble representation of the analysis equation can be expressed as

$$\boldsymbol{\Psi}_k^a = \boldsymbol{\Psi}_k^f + \boldsymbol{P}_{\psi_k}^{e,f}\boldsymbol{H}_k^T(\boldsymbol{H}_k\boldsymbol{P}_{\psi_k}^{e,f}\boldsymbol{H}_k^T + \boldsymbol{P}_{\epsilon_k^o})^{-1}(\boldsymbol{D}_k - \boldsymbol{H}_k\boldsymbol{\Psi}_k^f). \tag{4.11}$$

At time t_k we define the ensemble of innovation vectors, $\boldsymbol{D}_k' \in \Re^{m_k \times N}$, as

$$\boldsymbol{D}_k' = \boldsymbol{D}_k - \boldsymbol{H}_k\boldsymbol{\Psi}_k^f, \tag{4.12}$$

the measurement of the ensemble perturbations, $\boldsymbol{S}_k \in \Re^{m_k \times N}$, as

$$\boldsymbol{S}_k = \boldsymbol{H}_k\boldsymbol{\Psi}_k^{f'}, \tag{4.13}$$

and the matrix $\boldsymbol{C}_k \in \Re^{m_k \times m_k}$ as

$$\boldsymbol{C}_k = \boldsymbol{S}_k\boldsymbol{S}_k^T + (N-1)\boldsymbol{P}_{\epsilon_k^o}. \tag{4.14}$$

By using the Equations (4.12)-(4.14) and the ensemble error covariance matrix definition in Equation (4.6), the analysis in Equation (4.11) can be written as

$$\begin{aligned}
\boldsymbol{\Psi}_k^a &= \boldsymbol{\Psi}_k^f + \boldsymbol{\Psi}_k^{f'}\boldsymbol{\Psi}_k^{f'T}\boldsymbol{H}_k^T(\boldsymbol{H}_k\boldsymbol{\Psi}_k^{f'}\boldsymbol{\Psi}_k^{f'T}\boldsymbol{H}_k^T + (N-1)\boldsymbol{P}_{\epsilon_k^o})^{-1}\boldsymbol{D}_k' \\
&= \boldsymbol{\Psi}_k^f + \boldsymbol{\Psi}_k^f(\boldsymbol{I} - \boldsymbol{I}_N)\boldsymbol{S}_k^T\boldsymbol{C}_k^{-1}\boldsymbol{D}_k' \\
&= \boldsymbol{\Psi}_k^f(\boldsymbol{I} + (\boldsymbol{I} - \boldsymbol{I}_N)\boldsymbol{S}_k^T\boldsymbol{C}_k^{-1}\boldsymbol{D}_k') \\
&= \boldsymbol{\Psi}_k^f(\boldsymbol{I} + \boldsymbol{S}_k^T\boldsymbol{C}_k^{-1}\boldsymbol{D}_k') \\
&= \boldsymbol{\Psi}_k^f\boldsymbol{X}_k,
\end{aligned} \tag{4.15}$$

where we have used the ensemble perturbation matrix in Equation (4.5) and $\boldsymbol{1}_N\boldsymbol{S}_k^T \equiv 0$. Hence, the analyzed ensemble can be considered as a combination of the forecast ensemble. Assuming a linear model and Gaussian statistics Equation (4.15) converges towards the correct posterior density solution of the Bayesian problem with increasing ensemble size [3].

A detailed description of how the EnKF analysis can be implemented and computed efficiently for practical applications can be found in Evensen [54] and on the web-page: enkf.nersc.no.

4.1.1 Ensemble Kalman Smoother (EnKS)

As described in Evensen and van Leeuwen [37], the general smoother for non-linear dynamics, see Equation (3.4), can be formulated as a sequential method, so that measurements can be assimilated sequentially during a forward integration. The general filter, see Equation (3.5), can be found from the smoother, and at the final time the general filter and smoother will have identical solutions. The EnKS is closely related to the EnKF algorithm. The difference is that every time new observations become available, an update of the ensemble is performed for all previous times up to this time. Hence, the EnKF solution will be the first guess for the smoother, and the estimate of the smoother provides an improvement of this. The analysis utilizes the space-time covariance between the model state at a prior time and the model forecast at measurement locations. Thus, similar to the ensemble update in Equation (4.15) the EnKS update for a prior time t_j, based on a new set of measurements at time $t_k > t_j$, is given by

$$\begin{aligned}\boldsymbol{\Psi}_j^a &= \boldsymbol{\Psi}_j + \boldsymbol{\Psi}_j' \boldsymbol{\Psi}_k'^T \boldsymbol{H}_k^T (\boldsymbol{H}_k \boldsymbol{\Psi}_k' \boldsymbol{\Psi}_k'^T \boldsymbol{H}_k^T + (N-1)\boldsymbol{P}_{\epsilon_k^o})^{-1} \boldsymbol{D}_k' \\ &= \boldsymbol{\Psi}_j \boldsymbol{X}_k. \end{aligned} \tag{4.16}$$

The matrix of coefficients $\boldsymbol{X}_k$ is used on the analysis ensemble at prior times t_j to update the smoother estimate at time t_j, and the equation is updated at each time new observations are introduced at future times. An estimate of the smoother at a prior time t_j where $t_j < t_{j+1} \leq t_k$, using measurements from future times, i.e., $(t_{j+1}, t_{j+2}, \ldots, t_k)$, can then be written as

$$\boldsymbol{\Psi}_{\mathrm{EnKS},j}^a = \boldsymbol{\Psi}_{\mathrm{EnKF},j} \prod_{i=j+1}^{k} \boldsymbol{X}_i. \tag{4.17}$$

The multiplication sequence of $\boldsymbol{X}_i$ in Equation (4.17) will for each repetitive update result in a slight change of the ensemble mean and a slight reduction of the ensemble spread.

4.2 EnKF and Time Difference Data

In Section 3.2 we derived a sequential Bayesian formulation which incorporated time difference data, where the update included both a filter and a smoother solution. In this Section we consider the same data assimilation scheme, given in Figure 3.2 where the observations at time t_k depend both on $\boldsymbol{\psi}_k$ and $\boldsymbol{\psi}_j$, and we derive the corresponding EnKF scheme and show that a combination of the ensemble Kalman filter and the ensemble Kalman smoother has to be applied.

Let the non-linear state space model for times $t_i \in (t_j, t_k]$ be defined as

$$\tilde{\boldsymbol{\psi}}_i = \begin{bmatrix} \boldsymbol{f}_i(\boldsymbol{\psi}_{i-1}) \\ \boldsymbol{\psi}_j \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon}_i^m \\ 0 \end{bmatrix} \qquad \boldsymbol{\epsilon}_i^m \sim \mathcal{N}_q(0, \boldsymbol{P}_{\epsilon_i^m}) \tag{4.18}$$

$$\boldsymbol{d}_i = \tilde{\boldsymbol{H}}_i \tilde{\boldsymbol{\psi}}_i + \boldsymbol{\epsilon}_i^o \qquad \boldsymbol{\epsilon}_i^o \sim \mathcal{N}_{m_i}(0, \boldsymbol{P}_{\epsilon_i^o}) \tag{4.19}$$

where $\tilde{\boldsymbol{\psi}}_i = [\boldsymbol{\psi}_i, \boldsymbol{\psi}_j]^T \in \Re^{\tilde{q}}$, $\boldsymbol{f}_i$ is a non-linear model operator operating only on $\boldsymbol{\psi}_{i-1}$ with model errors $\boldsymbol{\epsilon}_i^m$. The model errors are only added to the terms where $\boldsymbol{f}_i$ is operating on $\boldsymbol{\psi}_{i-1}$ and are zero elsewhere. The linear measurement operator, $\tilde{\boldsymbol{H}}_i$, is relating the model state $\tilde{\boldsymbol{\psi}}_i$ to the observations $\boldsymbol{d}_i$ in accordance with the data assimilation scheme in Figure 3.2 (i.e. $\boldsymbol{d}_k$ include time difference observations between $\boldsymbol{\psi}_k$ and $\boldsymbol{\psi}_j$) allowing for measurement errors $\boldsymbol{\epsilon}_i^o$.

Consider now a joint ensemble matrix including the ensemble matrices for the different time instants, $\boldsymbol{\Psi}_i = \boldsymbol{\Psi}(\mathbf{x}, t_i), \boldsymbol{\Psi}_j = \boldsymbol{\Psi}(\mathbf{x}, t_j)$

$$\tilde{\boldsymbol{\Psi}}_i = \begin{bmatrix} \boldsymbol{\Psi}_i \\ \boldsymbol{\Psi}_j \end{bmatrix} \in \Re^{\tilde{q} \times N}. \tag{4.20}$$

The ensemble mean and the ensemble perturbation matrix at time t_i can be written as

$$\bar{\tilde{\boldsymbol{\Psi}}}_i = \tilde{\boldsymbol{\Psi}}_i \mathbf{1}_{\boldsymbol{N}} \tag{4.21}$$

$$\tilde{\boldsymbol{\Psi}}'_i = \tilde{\boldsymbol{\Psi}}_i - \bar{\tilde{\boldsymbol{\Psi}}}_i = \tilde{\boldsymbol{\Psi}}_i (\boldsymbol{I} - \mathbf{1}_{\boldsymbol{N}}), \tag{4.22}$$

where $\mathbf{1}_N \in \Re^{N \times N}$ is the matrix where each element is equal to $1/N$. The ensemble covariances, $\boldsymbol{P}^e_{\tilde{\psi}_i} \in \Re^{\tilde{q} \times \tilde{q}}$, can then be estimated as

$$\boldsymbol{P}^e_{\tilde{\psi}_i} = \frac{\tilde{\boldsymbol{\Psi}}'_i (\tilde{\boldsymbol{\Psi}}'_i)^T}{N-1}. \tag{4.23}$$

The sampling algorithm for a case with time difference data will then be as follows:

Algorithm 4.2

0. Generate N independent samples, $\boldsymbol{\psi}_j^{a,l}$ from $g(\boldsymbol{\psi}_j | \boldsymbol{d}_{j:1})$.

For $i = j+1, j+2, \ldots, k$

1. Forecast each of the samples, $\boldsymbol{\psi}_{i-1}^{a,l}$, using Equation (4.18):

$$\tilde{\boldsymbol{\psi}}_i^{f,l} = \begin{bmatrix} \boldsymbol{f}_i(\boldsymbol{\psi}_{i-1}^{a,l}) \\ \boldsymbol{\psi}_{j,i-1}^{a,l} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon}_i^{m,l} \\ 0 \end{bmatrix} \quad \boldsymbol{\epsilon}_i^{m,l} \sim \mathcal{N}_q(\mathbf{0}, \boldsymbol{P}_{\epsilon_i^m}) \ \forall \ l = 1:N$$

- $\boldsymbol{\psi}_{j,i-1}^{a,l}$ is the smoother update of $\boldsymbol{\psi}_j^{a,l}$ at the previous time step.
- $\tilde{\boldsymbol{\psi}}_i^{f,l}$ represents a sample from the prior distribution $g(\tilde{\boldsymbol{\psi}}_i|\boldsymbol{d}_{i-1:1})$.

2. Use Monte Carlo integration to estimate the mean and the covariance from the prior density, given in Equations (4.21) and (4.23).

3. Generate N samples from the measurement distribution:

$$\boldsymbol{d}_i^l \sim \mathcal{N}_{m_i}(\boldsymbol{d}_i, \boldsymbol{P}_{\epsilon_i^o}) \ \forall \ l = 1:N$$

4. Use the kriging formula to generate a sample from the posterior distribution $g(\tilde{\boldsymbol{\psi}}_i|\boldsymbol{d}_{i:1})$:

$$\tilde{\boldsymbol{\psi}}_i^{a,l} = \tilde{\boldsymbol{\psi}}_i^{f,l} + \boldsymbol{P}_{\tilde{\psi}_i}^{e,f} \tilde{\boldsymbol{H}}_i^T (\tilde{\boldsymbol{H}}_i \boldsymbol{P}_{\tilde{\psi}_i}^{e,f} \tilde{\boldsymbol{H}}_i^T + \boldsymbol{P}_{\epsilon_i^o})^{-1} (\boldsymbol{d}_i^l - \tilde{\boldsymbol{H}}_i \tilde{\boldsymbol{\psi}}_i^{f,l})$$

- where $\tilde{\boldsymbol{\psi}}_i^{a,l} = [\boldsymbol{\psi}_i^{a,l}, \boldsymbol{\psi}_{j,i}^{a,l}]^T$ and $\boldsymbol{\psi}_i^{a,l}$ represent a sample from the filter distribution $g(\boldsymbol{\psi}_i|\boldsymbol{d}_{i:1})$.

If we now at time t_i define the measurement of the ensemble perturbations, $\boldsymbol{S}_i \in \Re^{m_i \times N}$, as

$$\boldsymbol{S}_i = \tilde{\boldsymbol{H}}_i \tilde{\boldsymbol{\Psi}}_i^{f'}, \tag{4.24}$$

the ensemble of innovation vectors as

$$\boldsymbol{D}_i' = \boldsymbol{D}_i - \tilde{\boldsymbol{H}}_i \tilde{\boldsymbol{\Psi}}_i^f, \tag{4.25}$$

and the matrix $\boldsymbol{C}_i \in \Re^{m_i \times m_i}$ as

$$\boldsymbol{C}_i = \boldsymbol{S}_i \boldsymbol{S}_i^T + (N-1)\boldsymbol{P}_{\epsilon_i^o}, \tag{4.26}$$

we have from Equation (4.15) that the ensemble representation of the analysis step is

$$\tilde{\boldsymbol{\Psi}}_i^a = \tilde{\boldsymbol{\Psi}}_i^f \boldsymbol{X}_i. \tag{4.27}$$

The EnKF update in Equation (4.27) can be written as

$$\begin{bmatrix} \boldsymbol{\Psi}_i^a \\ \boldsymbol{\Psi}_{j,i}^a \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Psi}_i^f \boldsymbol{X}_i \\ \boldsymbol{\Psi}_{j,i-1}^a \boldsymbol{X}_i \end{bmatrix}, \tag{4.28}$$

where $\boldsymbol{\Psi}_{j,i}^a$ is the smoother solution at time t_j when the observations at time t_i have been assimilated.

This updating process continues until time t_k where the time difference data are to be assimilated, and we have the augmented ensemble as

$$\begin{bmatrix} \boldsymbol{\Psi}_k^f \\ \boldsymbol{\Psi}_{j,k-1}^a \end{bmatrix}. \tag{4.29}$$

At time t_k we use the time difference measurement operator which relates the measurements to both $\boldsymbol{\Psi}_k^f$ and $\boldsymbol{\Psi}_{j,k-1}^a$ and compute a standard EnKF analysis. More discussion of this procedure can be found in Paper A and B.

This method is particularly useful when assimilating the difference of the seismic data, since such data depend on the model state at two different time instances. The method has been applied to all EnKF runs throughout this work, where different types of seismic time-lapse data have been used. The overall results show that adding seismic time-lapse data to the production data measurements give an improvement of the characterization of the reservoir. However, special challenges, such as inversion of a large covariance matrix and potential loss of rank, are involved in the assimilation of the large amount of data comming from 4D seismic data. Hence, the use of efficient EnKF algorithms handling the problems properly are required.

4.3 Ensemble Square Root Filter

In the EnKF the traditional analysis scheme uses a randomization or a stochastic perturbation of the observations to ensure that the updated ensemble attain the correct variance. The so called deterministic analysis is an alternative algorithm which avoids the perturbation of observations, and the scheme is based on a square-root formulation. Here, the updates of the ensemble mean and the ensemble perturbations are performed separately [55].

Several variations of implementing the square root filter are introduced, caused by the non-uniqueness of analysis perturbations that can be used to represent the covariance matrix of the analysis error [56]. Existing publications of the square root filter approach are the ensemble adjustment filter [10], ensemble transform Kalman filter [57], local ensemble Kalman filter [58], and the ensemble square root filter (EnSRF) [56, 59, 9]. In this section the EnSRF is based on the formulation presented by Evensen [9]. Here we have dropped the time subscript and the forecast notation f to simplify the notation.

The scheme is used to update the ensemble perturbations, and the derivation of the algorithm is based on the traditional analysis equation for the covariance update of the Kalman filter in Equation (3.32). The ensemble representation for the error covariance matrix, $\boldsymbol{P}_{\psi}^e$, is defined in Equation (4.6), and by using this

and the definitions for $\boldsymbol{S}$ and $\boldsymbol{C}$ from Equations (4.13) and (4.14), respectively, the EnSRF equation can be expressed as

$$\boldsymbol{\Psi}^{a'}\boldsymbol{\Psi}^{a'T} = \boldsymbol{\Psi}'\boldsymbol{\Psi}'^{T} - \boldsymbol{\Psi}'\boldsymbol{\Psi}'^{T}\boldsymbol{H}^{T}(\boldsymbol{H}\boldsymbol{\Psi}'\boldsymbol{\Psi}'^{T}\boldsymbol{H}^{T} + (N-1)\boldsymbol{P}_{\epsilon^o})^{-1}\boldsymbol{H}\boldsymbol{\Psi}'\boldsymbol{\Psi}'^{T} \tag{4.30}$$

$$= \boldsymbol{\Psi}'(\boldsymbol{I} - \boldsymbol{S}^{T}\boldsymbol{C}^{-1}\boldsymbol{S})\boldsymbol{\Psi}'^{T} \tag{4.31}$$

$$= \boldsymbol{\Psi}'(\boldsymbol{I} - \boldsymbol{Z}\boldsymbol{\Lambda}\boldsymbol{Z}^{T})\boldsymbol{\Psi}'^{T} \tag{4.32}$$

$$= \boldsymbol{\Psi}'\boldsymbol{Z}(\boldsymbol{I} - \boldsymbol{\Lambda})\boldsymbol{Z}^{T}\boldsymbol{\Psi}'^{T} \tag{4.33}$$

$$= \left(\boldsymbol{\Psi}'\boldsymbol{Z}\sqrt{\boldsymbol{I} - \boldsymbol{\Lambda}}\right)\left(\boldsymbol{\Psi}'\boldsymbol{Z}\sqrt{\boldsymbol{I} - \boldsymbol{\Lambda}}\right)^{T}, \tag{4.34}$$

where $\boldsymbol{Z}$ and $\boldsymbol{\Lambda}$ represent respectively the eigenvectors and the eigenvalues found from an eigenvalue decomposition of the second term within the brackets of Equation (4.30). Hence a given solution of the ensemble perturbations is:

$$\boldsymbol{\Psi}'^{a} = \boldsymbol{\Psi}'\boldsymbol{Z}\sqrt{\boldsymbol{I} - \boldsymbol{\Lambda}}. \tag{4.35}$$

The updated mean is computed by a similar equation to the standard Kalman filter analysis (see Equation (3.31)) and is given as

$$\overline{\boldsymbol{\psi}^{a}} = \overline{\boldsymbol{\psi}^{f}} + \boldsymbol{\Psi}'\boldsymbol{S}^{T}\boldsymbol{C}^{-1}(\boldsymbol{d} - \boldsymbol{H}\overline{\boldsymbol{\psi}^{f}}), \tag{4.36}$$

where $\boldsymbol{d}$ are the unperturbed observations. As long as the measurement perturbations average to zero Equation (4.36) will give exactly the same updated ensemble mean as the EnKF. The above equations are similar to the equations solved in the EnSRF algorithms discussed by Tippett et al. [56].

Simulation of non-linear problems have shown that an ensemble collapse may occur when deterministic filters are applied. An insertion of a random orthogonal matrix product $\boldsymbol{I} = \boldsymbol{\Theta}^{T}\boldsymbol{\Theta}$ can avoid this problem, as shown by Leewuenburgh et. al [55] and Evensen [9], i.e

$$\boldsymbol{\Psi}^{a'}\boldsymbol{\Psi}^{a'T} = \left(\boldsymbol{\Psi}'\boldsymbol{Z}\sqrt{\boldsymbol{I} - \boldsymbol{\Lambda}}\right)\boldsymbol{\Theta}^{T}\boldsymbol{\Theta}\left(\boldsymbol{\Psi}'\boldsymbol{Z}\sqrt{\boldsymbol{I} - \boldsymbol{\Lambda}}\right)^{T}, \tag{4.37}$$

and the randomized EnSRF update equation can be written as

$$\boldsymbol{\Psi}'^{a} = \boldsymbol{\Psi}'\boldsymbol{Z}\sqrt{\boldsymbol{I} - \boldsymbol{\Lambda}}\boldsymbol{\Theta}^{T}. \tag{4.38}$$

This solution is still a square root of the updated covariance. A random rotation of the eigenvectors in $\boldsymbol{Z}$ is obtained by the multiplication with $\boldsymbol{\Theta}^{T}$, and the effect is a randomly redistribution of the variance among all the ensemble members.

The formulation of the square root scheme presented above can also be written in the same simple form as the EnKF equation (4.15). To obtain this form

the analysis is written as the updated ensemble mean plus the updated ensemble perturbations, i.e

$$\boldsymbol{\Psi}^a = \overline{\boldsymbol{\Psi}^a} + \boldsymbol{\Psi}^{a'}. \tag{4.39}$$

The analyzed mean using Equation (4.11), can be expressed as

$$\overline{\boldsymbol{\Psi}^a} = \boldsymbol{\Psi}\mathbf{1}_N + \boldsymbol{\Psi}(\boldsymbol{I} - \mathbf{1}_N)\boldsymbol{S}^T\boldsymbol{C}^{-1}(\boldsymbol{D} - \boldsymbol{H}\boldsymbol{\Psi})\mathbf{1}_N, \tag{4.40}$$

and the perturbations are, from Equation (4.38),

$$\boldsymbol{\Psi}^{a'} = \boldsymbol{\Psi}(\boldsymbol{I} - \mathbf{1}_N)\boldsymbol{Z}\sqrt{\boldsymbol{I} - \boldsymbol{\Lambda}}\boldsymbol{\Theta}^T \tag{4.41}$$

Combining the previous equations we get $\boldsymbol{\Psi}^a = \boldsymbol{\Psi}\boldsymbol{X}$ with $\boldsymbol{X}$ defined as

$$\boldsymbol{X} = \mathbf{1}_N + (\boldsymbol{I} - \mathbf{1}_N)\boldsymbol{S}^T\boldsymbol{C}^{-1}(\boldsymbol{D} - \boldsymbol{H}\boldsymbol{\Psi})\mathbf{1}_N + (\boldsymbol{I} - \mathbf{1}_N)\boldsymbol{Z}\sqrt{\boldsymbol{I} - \boldsymbol{\Lambda}}\boldsymbol{\Theta}^T. \tag{4.42}$$

Hence, we search for the EnSRF solution as a combination of ensemble members as discussed in Evensen [54]. As pointed out by Evensen [9] the forming of $\boldsymbol{X}$ and then computing the matrix multiplication $\boldsymbol{\Psi}^a = \boldsymbol{\Psi}\boldsymbol{X}$ is the most efficient algorithm for computing the analysis when many measurements are used.

The EnSRF scheme has been used in Report 1 to assimilate real production data and seismic time-lapse data, and in the North Sea field study the EnSRF method was able to improve the match during the assimilation. An efficient subspace EnSRF implementation was used, when seismic data were assimilated (see [9] and Paper B). In this example we observed the necessity of using a random orthogonal matrix $\boldsymbol{\Theta}^T$ in the EnSRF scheme to avoid loss of rank.

4.4 Other EnKF Analysis Schemes

4.4.1 Local Analysis

Local analysis is applied in many data assimilation application to reduce the impact of a finite ensemble size in the EnKF, especially when the model system is high dimensional and involves a large amount of observations [54, 58, 60, 61]. The local scheme assumes that only observations within a given influence region from a grid cell will impact the update in this specific grid cell. The local analysis updates each grid cell independently, and accounts for the observations only in the influence region surrounding the grid cell.

The local analysis is an approximate scheme and does not solve the original problem posed. However, the approximation is dependent of the size of the influence region, and when the region includes all the observations for all the grid cells, the local scheme solution is equal to the solution from the global scheme.

An improvement by using a local analysis is dependent on the range size, and the input of the range parameter will become a tuning problem. Thus, when using the local analysis it will be important that the size of the region is large enough to include all observations, which will significantly impact the analysis of the grid cell, but small enough to suppress spurious correlations between uncorrelated variables.

The local scheme is given as follows. First construct the matrices involved in the global scheme, i.e. the innovations $\boldsymbol{D}'$, the measured ensemble perturbations $\boldsymbol{S}$, and the measurement covariance matrix $\boldsymbol{P}_{\epsilon^o}$. For each grid cell the rows from these ensemble matrices, corresponding to observations that will be used in the local update, are copied to local matrices and used in the standard EnKF to calculate the local update $\boldsymbol{\Psi}_l^a$ for grid cell l. The EnKF update at grid cell l, is given by

$$\boldsymbol{\Psi}_l^a = \boldsymbol{\Psi}_l \boldsymbol{X}_l, \tag{4.43}$$

where $\boldsymbol{X}_l$ is the solution for a local analysis corresponding to grid cell l where only the observations inside the influence region are used in the analysis.

A combination of a global and a local analysis scheme is investigated in Papers D and E. The local scheme is used to assimilate seismic data, and in each grid cell update, the scheme accounts for the seismic observations only in the local region surrounding the cell. In reservoir simulation models the dimension of the model variable and the corresponding covariance matrix is often very large compared to the space spanned by the ensemble members, and in practical applications an ensemble size of typically 100 is used. Seismic data introduce a large number of data points in the assimilation step, and the results show that problems related to an underestimation of the uncertainty in the low rank covariance matrix representation of the model variables may occur when few ensemble members are used during the assimilation.

The advantage of using a local scheme is that the ratio between the dimension of the model variable and the ensemble size is smaller than for a standard global EnKF scheme, and this allows for a larger flexibility to obtain different model solutions, since the scheme will use a different combination of ensemble members for each grid point. This is also seen in the results in Papers D and E, and the combined global/local scheme shows an improvement in the characterization of the reservoir and the uncertainty predictions compared to the global scheme with the same ensemble size.

4.4.2 An Iterative EnKF Method

An iterative form of the EnKF algorithm is proposed for the case when the posterior pdf of interest may be non-Gaussian and the relation between model variables

and the predicted data may be highly nonlinear. Lately, different publications have shown that the standard EnKF may not provide an appropriate characterization of uncertainty in such problems [62].

The iterative algorithm incorporates some of the main features of EnKF, but the implementation of iterative EnKF requires far more computationally effort than the standard EnKF. This is due to the fact that the iterative EnKF algorithm requires gradients of the objective function, which in most reservoir applications can only be obtained by the adjoint method.

The motivation of the iterative algorithm was that the EnKF update equations can be derived as an approximation to the Gauss-Newton method, with an "average" sensitivity matrix. The iterative scheme combining RML, described in Section 2.3.3, and the EnKF and can be referred to as an RML/EnKF algorithm. In Gao and Reynolds [34] a comparison study between the RML/EnKF and the standard EnKF scheme was performed on a reservoir model, and the methods obtained similarly results.

4.4.3 Ensemble Particle Filters

An alternative filtering for ensemble data assimilation, called ensemble particle filter, has been investigated in several articles [63, 64, 65]. The ensemble particle filter also relies on the assumption that a relatively small ensemble size represents the uncertainty in a high dimensional model system. The particle filter, more generally known as a sequential Monte-Carlo method [66], is a fully nonlinear filter, where the first and second order moments of the error distribution are not directly used in the calculation. Hence the particle filter has potential to explore non-Gaussian probability distributions, where linearization of the observational and dynamical operators are omitted.

When the ensemble particle filter is used in data assimilation problems , where the non-linear dynamical models contain multimodal properties, this advantage can be exploited. For more information see the Sequential Monte Carlo Methods Particle Filtering web-page: www-sigproc.eng.cam.ac.uk/smc/

4.5 EnKF and Large Amount of Data

In simulation of reservoir models, spatial data such as 3D seismic and time-lapse seismic data may introduce large datasets in the assimilation and we have $m \gg N$. Thus, in reservoir cases, it will be important to have the possibility to use efficient EnKF schemes which handle such large amount of data. When working with efficient EnKF schemes, special challenges such as inversion of the $m \times m$ matrix $\boldsymbol{C}$ in Equation (4.14), and potential loss of rank are involved.

It has been pointed out by Evensen [9] that the EnKF scheme may have problems in cases where the number of observations is larger than the number of ensemble members or when the matrix $\boldsymbol{C}$, defined in Equation (4.14), has poor conditioning.

Large datasets may introduce numerical singularity in $\boldsymbol{C}$ even when a full rank error covariance matrix $\boldsymbol{P}_{\epsilon^o}$ is used. When $\boldsymbol{C}$ becomes singular it is possible to use the pseudo inverse $\boldsymbol{C}^+$ of $\boldsymbol{C}$. The pseudo inverse of the quadratic matrix $\boldsymbol{C}$ with eigenvalue factorization

$$\boldsymbol{C} = \boldsymbol{Z}\boldsymbol{\Lambda}\boldsymbol{Z}^T \tag{4.44}$$

is given as

$$\boldsymbol{C}^+ = \boldsymbol{Z}\boldsymbol{\Lambda}^+\boldsymbol{Z}^T. \tag{4.45}$$

The matrix $\boldsymbol{\Lambda}^+$ is diagonal and with $p = \text{rank}(\boldsymbol{C})$ it is defined as

$$\text{diag}(\boldsymbol{\Lambda}^+) = (\lambda_1^{-1}, \ldots, \lambda_p^{-1}, 0, \ldots, 0) \tag{4.46}$$

with eigenvalues $\lambda_i \geq \lambda_{i+1}$.

In cases with many observations, the computational effort increases significantly since an order of $\mathcal{O}(Nm^2)$ operations are required to form the matrix $\boldsymbol{C}$ and the eigenvalue decomposition requires $\mathcal{O}(m^3)$ operations. Hence, an ensemble representation of the error covariance matrix, defined in Equation (4.10), can be used to reduce the computational cost [54]. As stated by Kepert [67] the use of an ensemble representation, $\boldsymbol{P}^e_{\epsilon^o}$ for $\boldsymbol{P}_{\epsilon^o}$ in some cases may lead to loss of rank in the ensemble when $m > N$. The rank problem may occur both using the standard EnKF analysis scheme with perturbations of observations and using the square root algorithm.

Evensen [9] showed that the rank problem can be avoided if the observations are projected onto a subspace $\mathcal{L}_S$, spanned by the column vectors of S, defined in Equation (4.13). Based on this result, an alternative inversion algorithm which reduces the factorization of the $m \times m$ matrix to a factorization of an $N \times N$ matrix was proposed [9]. The algorithm computes the inverse in the N-dimensional ensemble space rather than the m-dimensional measurement space.

In the following we will first present an alternative EnKF expression and then we use this interpretation to derive an efficient subspace EnKF algorithm.

4.5.1 An Alternative EnKF Expression

The alternative EnKF expression is based on the singular value decomposition (SVD) of the measurement of the ensemble perturbations $\boldsymbol{S}$, and is

$$\boldsymbol{S} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T, \tag{4.47}$$

with $\boldsymbol{U} \in \mathbb{R}^{m\times m}, \Sigma \in \mathbb{R}^{m\times N}$ and $\boldsymbol{V} \in \mathbb{R}^{N\times N}$. The matrix $\boldsymbol{S}\boldsymbol{S}^T$ can then be written as

$$\boldsymbol{S}\boldsymbol{S}^T = \boldsymbol{U}\Sigma\boldsymbol{V}\boldsymbol{V}^T\Sigma^T\boldsymbol{U}^T = \boldsymbol{U}\Sigma\Sigma^T\boldsymbol{U}^T, \tag{4.48}$$

since the eigenvectors in $\boldsymbol{V}$ are orthogonal and, thus, $\boldsymbol{V}\boldsymbol{V}^T = \boldsymbol{I}$. Furthermore, the columns of $\boldsymbol{U}$ are orthogonal, thus, $\boldsymbol{U}\boldsymbol{U}^T = \boldsymbol{I}$.

We define the rotation of the measurements, $\hat{\boldsymbol{d}} \in \mathbb{R}^m$, as

$$\hat{\boldsymbol{d}} = \boldsymbol{U}^T\boldsymbol{d}, \tag{4.49}$$

and the rotated measurement operator $\hat{\boldsymbol{H}} \in \mathbb{R}^{m\times q}$ as

$$\hat{\boldsymbol{H}} = \boldsymbol{U}^T\boldsymbol{H}. \tag{4.50}$$

The new observation equation will then be given by

$$\hat{\boldsymbol{d}} = \hat{\boldsymbol{H}}\boldsymbol{\psi} + \hat{\boldsymbol{\epsilon}}^o \qquad\qquad \hat{\boldsymbol{\epsilon}}^o \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{P}_{\hat{\epsilon}^o}). \tag{4.51}$$

Here the rotated measurement errors are defined by $\hat{\boldsymbol{\epsilon}}^o = \boldsymbol{U}^T\boldsymbol{\epsilon}^o$, and the error covariance matrix is given by $\boldsymbol{P}_{\hat{\epsilon}^o} = \boldsymbol{U}^T\boldsymbol{P}_{\epsilon^o}\boldsymbol{U}$. If we now define the rotated measurements of the ensemble perturbations, $\hat{\boldsymbol{S}} \in \mathbb{R}^{m\times N}$ as

$$\hat{\boldsymbol{S}} = \hat{\boldsymbol{H}}\boldsymbol{\Psi}' = \boldsymbol{U}^T\boldsymbol{H}\boldsymbol{\Psi}', \tag{4.52}$$

and the rotated ensemble of innovation vectors, $\hat{\boldsymbol{D}}' \in \mathbb{R}^{m\times N}$, as

$$\hat{\boldsymbol{D}}' = \hat{\boldsymbol{D}} - \hat{\boldsymbol{H}}\boldsymbol{\Psi} = U^T(\boldsymbol{D} - \boldsymbol{H}\boldsymbol{\Psi}), \tag{4.53}$$

then the alternative EnKF analysis scheme can be written as

$$\boldsymbol{\Psi}^a = \boldsymbol{\Psi}\left(\boldsymbol{I} + \hat{\boldsymbol{S}}^T\left(\hat{\boldsymbol{S}}\hat{\boldsymbol{S}}^T + (N-1)\boldsymbol{P}_{\hat{\epsilon}^o}\right)^{-1}\hat{\boldsymbol{D}}'\right), \tag{4.54}$$

where the rotation is defined so that $\hat{\boldsymbol{S}}\hat{\boldsymbol{S}}^T = \Sigma\Sigma^T$ becomes diagonal. Equation (4.54) can be rewritten as

$$\begin{aligned}
\boldsymbol{\Psi}^a &= \boldsymbol{\Psi}\left(\boldsymbol{I} + \hat{\boldsymbol{S}}^T\left(\Sigma\Sigma^T + (N-1)\boldsymbol{U}^T\boldsymbol{P}_{\epsilon^o}\boldsymbol{U}\right)^{-1}\hat{\boldsymbol{D}}'\right) \\
&= \boldsymbol{\Psi}\left(\boldsymbol{I} + \boldsymbol{S}^T\boldsymbol{U}\left(\Sigma\Sigma^T + (N-1)\boldsymbol{U}^T\boldsymbol{P}_{\epsilon^o}\boldsymbol{U}\right)^{-1}\boldsymbol{U}^T\boldsymbol{D}'\right) \\
&= \boldsymbol{\Psi}\left(\boldsymbol{I} + \boldsymbol{S}^T\left(\boldsymbol{U}(\Sigma\Sigma^T + (N-1)\boldsymbol{U}^T\boldsymbol{P}_{\epsilon^o}\boldsymbol{U})\boldsymbol{U}^T\right)^{-1}\boldsymbol{D}'\right) \\
&= \boldsymbol{\Psi}\left(\boldsymbol{I} + \boldsymbol{S}^T\left(\boldsymbol{U}\Sigma\Sigma^T\boldsymbol{U}^T + (N-1)\boldsymbol{P}_{\epsilon^o}\right)^{-1}\boldsymbol{D}'\right) \\
&= \boldsymbol{\Psi}\left(\boldsymbol{I} + \boldsymbol{S}^T\left(\boldsymbol{S}\boldsymbol{S}^T + (N-1)\boldsymbol{P}_{\epsilon^o}\right)^{-1}\boldsymbol{D}'\right) \\
&= \boldsymbol{\Psi}\left(\boldsymbol{I} + \boldsymbol{S}^T\boldsymbol{C}^{-1}\boldsymbol{D}'\right),
\end{aligned} \tag{4.55}$$

and we see that the assimilation of m rotated measurements in Equation (4.55) is equivalent to the original EnKF scheme in Equation (4.15).

4.5.2 Efficient Subspace EnKF Algorithm

In this section we present the derivation of the efficient subspace EnKF algorithm. The motivation for the proposed subspace EnKF algorithm is based on the efficient square root algorithm presented by Evensen [9] and the alternative EnKF expression. Now assume that $\boldsymbol{S}$ has rank $p \leq \min(m, N-1)$. Some of the singular values in Σ may then be zero. We therefore partition Σ into a submatrix Σ_p of p nonzero singular values and several zero matrices as

$$\Sigma = \begin{bmatrix} \Sigma_p & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \tag{4.56}$$

where Σ_p is a $p \times p$ diagonal matrix. The decomposition then becomes

$$\boldsymbol{S} = \boldsymbol{U}\Sigma\boldsymbol{V}^T = \boldsymbol{U}_p\Sigma_p\boldsymbol{V}_p^T, \tag{4.57}$$

where $\boldsymbol{U}_p \in \Re^{m\times p}$ and $\boldsymbol{V}_p \in \Re^{N\times p}$ consist of the first p columns of $\boldsymbol{U}$ and $\boldsymbol{V}$, respectively. From the decomposition in Equation (4.57) the matrix $\boldsymbol{S}\boldsymbol{S}^T$ can be written as

$$\boldsymbol{S}\boldsymbol{S}^T = \boldsymbol{U}_p\Sigma_p\Sigma_p^T\boldsymbol{U}_p^T, \tag{4.58}$$

where the projection operator $\boldsymbol{U}_p$ span the ensemble space $\mathcal{L}_S$.

We now define the projection of the measurements, $\hat{\boldsymbol{d}}_p \in \Re^p$, as

$$\hat{\boldsymbol{d}}_p = \boldsymbol{U}_p^T\boldsymbol{d}_p, \tag{4.59}$$

and the projected measurement operator $\hat{\boldsymbol{H}}_p \in \Re^{p\times q}$ as

$$\hat{\boldsymbol{H}}_p = \boldsymbol{U}_p^T\boldsymbol{H}. \tag{4.60}$$

The new observation equation will then be given by

$$\hat{\boldsymbol{d}}_p = \hat{\boldsymbol{H}}_p\boldsymbol{\psi} + \hat{\boldsymbol{\epsilon}}_p^o \qquad\qquad \hat{\boldsymbol{\epsilon}}_p^o \sim \mathcal{N}(\mathbf{0}, \boldsymbol{P}_{\hat{\epsilon}_p^o}). \tag{4.61}$$

Here the projected measurement errors are defined by $\hat{\boldsymbol{\epsilon}}_p^o = \boldsymbol{U}_p^T\boldsymbol{\epsilon}_p^o$, and the error covariance matrix is given by $\boldsymbol{P}_{\hat{\epsilon}_p^o} = \boldsymbol{U}_p^T\boldsymbol{P}_{\epsilon^o}\boldsymbol{U}_p$.

Define the projected measurements of the ensemble perturbations, $\hat{\boldsymbol{S}}_p \in \Re^{p\times N}$ as

$$\hat{\boldsymbol{S}}_p = \hat{\boldsymbol{H}}_p\boldsymbol{\Psi}' = \boldsymbol{U}_p^T\boldsymbol{H}\boldsymbol{\Psi}', \tag{4.62}$$

and the projected ensemble of innovation vectors, $\hat{\boldsymbol{D}}_p' \in \Re^{p\times N}$, as

$$\hat{\boldsymbol{D}}_p' = \hat{\boldsymbol{D}}_p - \hat{\boldsymbol{H}}_p\boldsymbol{\Psi} = \boldsymbol{U}_p^T(\boldsymbol{D} - \boldsymbol{H}\boldsymbol{\Psi}). \tag{4.63}$$

The subspace EnKF analysis scheme can then be written as

$$\boldsymbol{\Psi}^a = \boldsymbol{\Psi}\left(\boldsymbol{I} + \hat{\boldsymbol{S}}_p^T\left(\Sigma_p\Sigma_p^T + (N-1)\boldsymbol{U}_p^T\boldsymbol{P}_{\epsilon^o}\boldsymbol{U}_p\right)^{-1}\hat{\boldsymbol{D}}_p'\right). \tag{4.64}$$

If we use the ensemble approximation $\boldsymbol{P}_{\hat{\epsilon}^o}^e = \boldsymbol{E}\boldsymbol{E}^T/(N-1)$ of $\boldsymbol{P}_{\hat{\epsilon}^o}$, the subspace EnKF analysis scheme can be written as

$$\boldsymbol{\Psi}^a = \boldsymbol{\Psi}\left(\boldsymbol{I} + \hat{\boldsymbol{S}}_p^T\left(\Sigma_p\Sigma_p^T + (N-1)\boldsymbol{U}_p^T\boldsymbol{P}_{\epsilon^o}^e\boldsymbol{U}_p\right)^{-1}\hat{\boldsymbol{D}}_p'\right) \tag{4.65}$$

$$= \boldsymbol{\Psi}\left(\boldsymbol{I} + \hat{\boldsymbol{S}}_p^T\left(\Sigma_p\Sigma_p^T + \hat{\boldsymbol{E}}_p\hat{\boldsymbol{E}}_p^T\right)^{-1}\hat{\boldsymbol{D}}_p'\right), \tag{4.66}$$

where $\boldsymbol{U}_p^T\boldsymbol{P}_{\epsilon^o}^e\boldsymbol{U}_p$ and $\hat{\boldsymbol{E}}_p = \boldsymbol{U}_p^T\boldsymbol{E}$ are the projection of the measurement error covariance matrix $\boldsymbol{P}_{\epsilon^o}^e$ and the projection of the ensemble of measurement perturbations $\boldsymbol{E}$ onto the subspace $\mathcal{L}_S$, respectively. This ensures that the measurement perturbations explain the variance within the ensemble space $\mathcal{L}_S$ and we avoid the loss of rank pointed out by Evensen [9] and Kepert [67].

The sub-space pseudo inversion in Equation (4.64) can be computed at a cost of order $\mathcal{O}(Nm^2)$, and when a low rank representation, $\boldsymbol{P}_{\epsilon^o}^e$, is used for the error covariance matrix the cost of the inversion is of order $\mathcal{O}(N^2m)$. This is a significant saving when $m \gg N$.

The interpretation of this subspace EnKF algorithm can be viewed as an assimilation of a set of measurements after they have been projected onto the subspace $\mathcal{L}_S$ as defined by the first p singular vectors of $\boldsymbol{S}$. The subspace EnKF algorithm does not introduce any approximation if $\boldsymbol{P}_{\epsilon^o}$ is diagonal, since the matrix $\boldsymbol{C}$ and $\boldsymbol{S}\boldsymbol{S}^T$ will have the same eigenvectors. However, if $\boldsymbol{P}_{\epsilon^o}$ is non-diagonal they may have different eigenvectors and the projection onto $\mathcal{L}_S$ eliminates the part of $\boldsymbol{C}$ which is orthogonal to the $\mathcal{L}_S$ space. The use of a low rank representation of the measurement error covariance matrix, $\boldsymbol{P}_{\epsilon^o}^e$, allows us to assimilate very large data sets at a low cost, and the result will be the same as obtained by using a full rank $\boldsymbol{P}_{\epsilon^o}$ if $\boldsymbol{U}_p^T\boldsymbol{P}_{\epsilon^o}^e\boldsymbol{U}_p = \boldsymbol{U}_p^T\boldsymbol{P}_{\epsilon^o}\boldsymbol{U}_p$.

To ensure an appropriate representation of the uncertainty of the ensemble, $\boldsymbol{\Psi}$, it is important that the rank of the covariance matrix, $\boldsymbol{P}_{\psi}^e$, defined by Equation (4.6), is maintained during the time integration. The rank of the covariance is maximum equal to $N-1$, and in reservoir applications we typically have $N = 100$. Paper B shows that the results obtained from the subspace EnKF scheme, when either a full rank measurement error covariance matrix, $\boldsymbol{P}_{\epsilon^o}$, or a low rank matrix, $\boldsymbol{P}_{\epsilon^o}^e$, are used, are fairly consistent with the standard EnKF scheme. The experiments also show that there is no loss of rank in the analyzed ensemble. Further discussions can be found in Paper B and Evensen [3, 9].

Chapter 5

Generation of Realizations

In this chapter we discuss how to generate Gaussian random realizations with a specified mean and covariance. In an EnKF technique such simulation methods can be used to generate the model and measurement perturbations, and the initial ensemble members.

In geostatistics the variable under study is called the regionalized variable $\{\boldsymbol{\psi}(\boldsymbol{x}) : \boldsymbol{x} \in \boldsymbol{\Omega} \subseteq \Re^{n_d}\}$, where n_d is the dimension of the system. We assume a second order stationary stochastic process

$$\mathrm{E}(\boldsymbol{\psi}(\boldsymbol{x})) = \mu \tag{5.1}$$

$$\mathrm{E}(\boldsymbol{\psi}(\boldsymbol{x}) - \boldsymbol{\psi}(\boldsymbol{x}+\boldsymbol{r}))^2 = 2\gamma(\boldsymbol{r}) \;\; \forall \boldsymbol{x} \in \Omega, \tag{5.2}$$

where both the expected value and the second moment are location independent. The $\gamma(\boldsymbol{r})$ is called the variogram, see e.g. [68], and is defined as

$$\gamma(\boldsymbol{r}) = P(0) - P(\boldsymbol{r}). \tag{5.3}$$

The covariance function $P(\boldsymbol{r})$ at distance $\boldsymbol{r}$ is given as

$$P(\boldsymbol{r}) = \sigma^2 \rho(\boldsymbol{r}) = \mathrm{Cov}(\boldsymbol{\psi}(\boldsymbol{x}), \boldsymbol{\psi}(\boldsymbol{x}+r)), \tag{5.4}$$

where σ^2 denote the variance and $\rho(\boldsymbol{r})$ denote the correlation function.

Since it is required that the correlation function $\rho(\boldsymbol{r})$ is positive definite, a number of parametric forms are commonly used, such as the exponential correlation function

$$\rho_{\exp}(\boldsymbol{r}) = \exp\left(-\frac{\boldsymbol{r}}{a}\right), \tag{5.5}$$

the Gaussian correlation function

$$\rho_{\mathrm{gauss}}(\boldsymbol{r}) = \exp\left(-\frac{\boldsymbol{r}^2}{a^2}\right), \tag{5.6}$$

and the spherical correlation function

$$\rho_{\text{sphere}} = \begin{cases} 1 - \frac{3}{2}\frac{\boldsymbol{r}}{a} + \frac{1}{2}\frac{\boldsymbol{r}^3}{a^3} & \text{for } 0 \leq \boldsymbol{r} \leq a \\ 0 & \text{for } \boldsymbol{r} > a \end{cases} \tag{5.7}$$

where a defines the range.

In a history matching process we approach the problem in a probabilistic frame, since different models of the reservoir could reproduce the same observed data. Multiple realizations of the reservoir should therefore be generated in order to assess the uncertainty in reservoir performance. Hence, simulation of initial realizations from the prior distribution is required. The simulation algorithm should reproduce the variability of the regionalized variable given from the prior model, and the multiple realizations should be equiprobable. In the following we present a FFT method, a sequential method and a matrix decomposition method, which can be used to generate Gaussian realizations.

5.1 Efficient Gaussian Simulation with Fast Fourier Transform

Let $\psi_z = \psi_z(x, y)$ be a continuous field with zero mean, variance equal to one and a Gaussian correlation. An efficient method for generating a two dimensional pseudo random fields was presented by Evensen [44], and in the following an outline of the method is given.

A continuous field $\psi_z(x, y)$ can be described by its Fourier transform

$$\psi_z(x, y) = \int_{-\infty}^{\infty} \hat{\psi}_z(\boldsymbol{k}) e^{i\boldsymbol{k}\boldsymbol{x}} \mathrm{d}\boldsymbol{k}. \tag{5.8}$$

Let now $n_x \times n_y$ be the dimension of the grid, and define the wave numbers in x and y directions as, $\boldsymbol{k} = (\kappa_l, \lambda_p)$, where l and p are counters. The discrete form of Equation (5.8) can be given as

$$\psi_z(x_n, y_m) = \sum_{l,p} \hat{\psi}_z(\kappa_l, \lambda_p) e^{i(\kappa_l x_n + \lambda_p y_m)} \Delta\boldsymbol{k}, \tag{5.9}$$

where we have $x_n = n\Delta x$, $y_m = m\Delta y, \kappa_l = \frac{2\pi l}{n_x \Delta x}$, $\lambda_p = \frac{2\pi p}{n_y \Delta y}$ and $\Delta k = \Delta\kappa\Delta\lambda = \frac{(2\pi)^2}{n_x n_y \Delta x \Delta y}$. Define the following Gaussian form for the Fourier coefficients

$$\hat{\psi}_z(\kappa_l, \lambda_p) = \frac{c}{\Delta k} e^{-(\kappa_l^2 + \lambda_p^2)/\varsigma^2} e^{2\pi i \phi_{l,p}}. \tag{5.10}$$

In Equation (5.10) the variable $\phi_{l,p} \in [0, 1]$ denote a uniformly distributed random number that introduces a random phase shift. By inserting Equation (5.10) into Equation (5.8) the pseudo random field can be written as

$$\psi_z(x_n, y_m) = \Delta k \sum_{l,p} \frac{c}{\sqrt{\Delta k}} e^{-(\kappa_l^2+\lambda_p^2)/\varsigma^2} e^{2\pi i \phi_{l,p}} e^{i(\kappa_l x_n + \lambda_p y_m)}. \tag{5.11}$$

Typically the values for c and ς are determined by using a numerical algorithm, and when these values are found the Equation (5.11) can be used to generate an ensemble of pseudo random fields. An efficient method for finding the inverse transform in Equation (5.11) is to use a two dimensional fast Fourier transform (FFT).

If this method is used to generate realizations in a three dimensional model, we may need to impose a vertical correlation ρ_v between two independent realizations $\psi_{z_i}(\boldsymbol{x})$ and $\psi_{z_{i-1}}(\boldsymbol{x})$, and a following method can be used [54]

$$\psi_{z_i}(\boldsymbol{x}) = \sqrt{(1-\rho_v^2)}\psi_{z_i}(\boldsymbol{x}) + \rho_v \psi_{z_{i-1}}(\boldsymbol{x}), \tag{5.12}$$

A back transform of a standarized field ψ_z to a field ψ with mean, μ, and standard deviation, σ, can be performed by using the formula

$$\psi = \mu + \sigma\psi_z. \tag{5.13}$$

A more comprehensive description of the method can be found in Evensen [3].

5.2 Sequential Gaussian Simulation

The sequential Gaussian simulation (SGS) algorithm is based on the kriging equations [41]. The kriging equations give an estimate of both the mean and the variance of the regionalized variable at each grid cell, so that the variable can be represented at each grid cell as a random variable following a Gaussian distribution. Instead of choosing the mean as the estimate at each grid cell, SGS chooses a sample from this Gaussian distribution. Kriging is a class of linear estimators, and the Simple Kriging (SK) is given as

$$\hat{\psi}_{\mathrm{SK}}(\boldsymbol{x}) = \mu + \sum_{i=1}^{n(\boldsymbol{x})} \beta_i^{\mathrm{SK}} [\psi(\boldsymbol{x}_i) - \mu], \tag{5.14}$$

where $n(\boldsymbol{x})$ is the number of data points in local neighborhood used for estimation of $\hat{\psi}_{\mathrm{SK}}(\boldsymbol{x})$ and β_i^{SK} is the kriging weight. The estimate in Equation (5.14) is unbiased, since $\mathrm{E}[\psi(\boldsymbol{x}_i) - \mu] = 0$, and we have $\mathrm{E}[\hat{\psi}_{\mathrm{SK}}(\boldsymbol{x})] = \mu = \mathrm{E}[\psi(\boldsymbol{x})]$.

The kriging variance is given by

$$\sigma^2_{\mathrm{SK}} = P(0) - \sum_{i=1}^{n(\boldsymbol{x})} \beta_i^{\mathrm{SK}} P(\boldsymbol{x} - \boldsymbol{x}_i), \tag{5.15}$$

where $P(\boldsymbol{x} - \boldsymbol{x}_i)$ is the point-to-point covariance between point $\boldsymbol{x}$ and $\boldsymbol{x}_i$.

The kriging weights $\beta_i^{\mathrm{SK}}, i = 1, \ldots, n(\boldsymbol{x})$, are determined from the following system of equations

$$\sum_{j=1}^{n(\boldsymbol{x})} \beta_j^{\mathrm{SK}} P(\boldsymbol{x}_i - \boldsymbol{x}_j) = P(\boldsymbol{x}_i - \boldsymbol{x}) \ \ \forall \ \ i = 1, \ldots, n(\boldsymbol{x}) \tag{5.16}$$

obtained by requiring that $\mathrm{Var}[\hat{\psi}_{\mathrm{SK}}(\boldsymbol{x}) - \psi(\boldsymbol{x})]$ is minimal. The basic steps in the SGS algorithm are:

Algorithm 5.2

1. Generate a random path through the grid cells.
2. Visit the first cell along the path and use the kriging Equations (5.14) and (5.15) to estimate the mean and the variance for the variable at that cell based on data values in the local neighbourhood.
3. Sample a value from the corresponding Gaussian distribution and set the variable value at that cell to that number.
4. Visit each successive cell in the random path and repeat the process, including previously simulated cells as data values in the kriging process.

The SGS algorithm uses a random path to avoid artifacts, which may be introduced by moving through the grid in a regular way, and in order to preserve the correlation structure between the simulated values the previously simulated grid cells are included as data in the method.

5.3 Simulations via Cholesky Decomposition

The Cholesky method is a Monte Carlo simulation method for generating Gaussian realizations [69]. In the Cholesky method the covariance matrix, $\boldsymbol{P}$, is written as the product of a lower-triangular matrix and its transpose:

$$\boldsymbol{P} = \boldsymbol{\Gamma}\boldsymbol{\Gamma}^T. \tag{5.17}$$

Then, if $\boldsymbol{\psi}_{z_i} \sim \mathcal{N}(\mathbf{0}, \mathbf{1})$ is a independent standard normal realization vector , the model

$$\boldsymbol{\psi}_i = \boldsymbol{\mu} + \boldsymbol{\Gamma}\boldsymbol{\psi}_{z_i} \tag{5.18}$$

is a realization from the pdf of interest, i.e

$$\boldsymbol{\psi}_i \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{P}), \tag{5.19}$$

since

$$\mathrm{E}[(\boldsymbol{\psi} - \boldsymbol{\mu})(\boldsymbol{\psi} - \boldsymbol{\mu})^T] = \boldsymbol{\Gamma}\mathrm{E}(\boldsymbol{\psi}_z \boldsymbol{\psi}_z^T)\boldsymbol{\Gamma}^T = \boldsymbol{\Gamma}\boldsymbol{\Gamma}^T = \boldsymbol{P}. \tag{5.20}$$

When the Cholesky decomposition of $\boldsymbol{P}$ has been performed, the generation of new realizations will be inexpensive. However, the disadvantage of the Cholesky method is that the dimension of the covariance matrix is equal to the number of points for which simulated values should be calculated, and the computational effort of the method increases with the number of points.

Chapter 6

Conditioning to Seismic Data

In a reservoir characterization perspective it will be important to utilize all the available data to make an improvement of the reservoir model and provide more reliable predictions of the performance of future production efforts.

Traditionally, in a history matching process reservoir properties have been conditioned to dynamic production data from wells only. Such data can be water cut (WCT), gas-oil ratio (GOR) and bottom-hole pressure (BHP). Typically the well data are recorded at high frequency in time but are of low frequency in space, and they can be measured with a relatively low uncertainty.

Recently, time-lapse seismic data have been used as an additional data set in the history matching process. Time-lapse seismic, commonly known as 4D seismic, is the comparison of an initial 3D seismic survey, called base-line, to subsequent repeat surveys, called monitor-lines, over the same geographical location. The seismic signature of a reservoir is mainly dependent on two primary elements, the static rock properties of the reservoir such as lithology, cementation and porosity, and the time varying dynamic properties such as pore pressure and fluid saturation. The use of 3D seismic data was first introduced in the early 1980's in the geophysical community, while time-lapse seismic, which involves the acquisition, processing and interpretation of 3D seismic data, was first introduced in the late 1990's. To date, in addition to traditional marine surface seismic, the ability to use Ocean Bottom Seismometer (OBS) data in 4D becomes increasingly important. The data can be obtained from permanent acquisition systems, and will both have a higher reliability and time frequency.

The 4D data will be of importance in a reservoir-description process, since they have potential to contain information about pressure changes and fluid movement in the reservoir due to depletion or injection. Thus, by monitoring fluid fronts and discriminating between pressure and fluid saturation changes over time, we may be enable to identify water encroachment, gas-cap formations and by-passed oil reserves. Compared to the spatial variations in model properties the

changes in seismic properties due to time lapse effects will often be small, and it is possible to use this advantage to construct more efficiently repeated seismic modelling [70].

The uncertainties of the time-lapse seismic measurements are normally much larger than for the well data. In the space domain seismic data have a high lateral resolution, and a medium to low vertical resolution. Contrary to the well data, the seismic data have a low frequency in time.

Time-lapse seismic has until now mainly served as a monitoring tool to map changes in reservoir pressure and saturation. Recently, several publications have discussed the problem of using real 4D seismic data in history matching of reservoir simulation models, to improve the characterization of permeability and porosity heterogeneities; see for example [71, 72, 73, 74, 75, 76]. However, most of the publications incorporate the 4D seismic qualitatively [74]. Lately, different publications have discussed how time-lapse seismic data can be incorporated in a computer-aided history matching process in a quantitative way [73, 75, 77], and results from history matching cases based on this approach on real field data, can be found in Aanonsen et al. [78] and Haverl et al. [79].

A challenging issue when using seismic data is how to compare the measured data to the model data. In principle, seismic data may be included in a number of different ways when conditioning reservoir models. One possibility is to use seismic amplitudes, another is to use inverted and/or processed data, and finally pressures and saturations derived from a rock physics model can be used. The uncertainty, amount of data, and information content are very different and related to the way seismic data are incorporated. Furthermore the efficiency of the conditioning process will also vary.

Measuring the mismatch between the model and data in a consistent way will always be required when conditioning a model to data [76]. To make the model data and the measured data comparable, both a seismic forward model and a seismic inversion process are needed. An illustration of the different seismic mismatch levels is shown in Figure 6.1. To incorporate measured seismic amplitude data seismic modelling is required, so that synthetic seismograms, adherent pre- and post stack seismic sections, and seismic attributes such as AVO (angle versus offset) parameters can be produced. Input to seismic modelling is a static description of the overburden and reservoir, and for dynamic modelling, fluid saturation and pressure. The static reservoir data may be taken from a geological model of the reservoir or from a reservoir simulation model, while the dynamic data have to be achieved from a reservoir simulation model.

A mismatch in the elastic properties will involve a seismic inversion and a rock physics forward model. A seismic inversion process converts measured seismic amplitude data, an interface property, to the elastic parameter of the layers, see [80]. Similarly those elastic quantities, such as acoustic impedance and Pois-

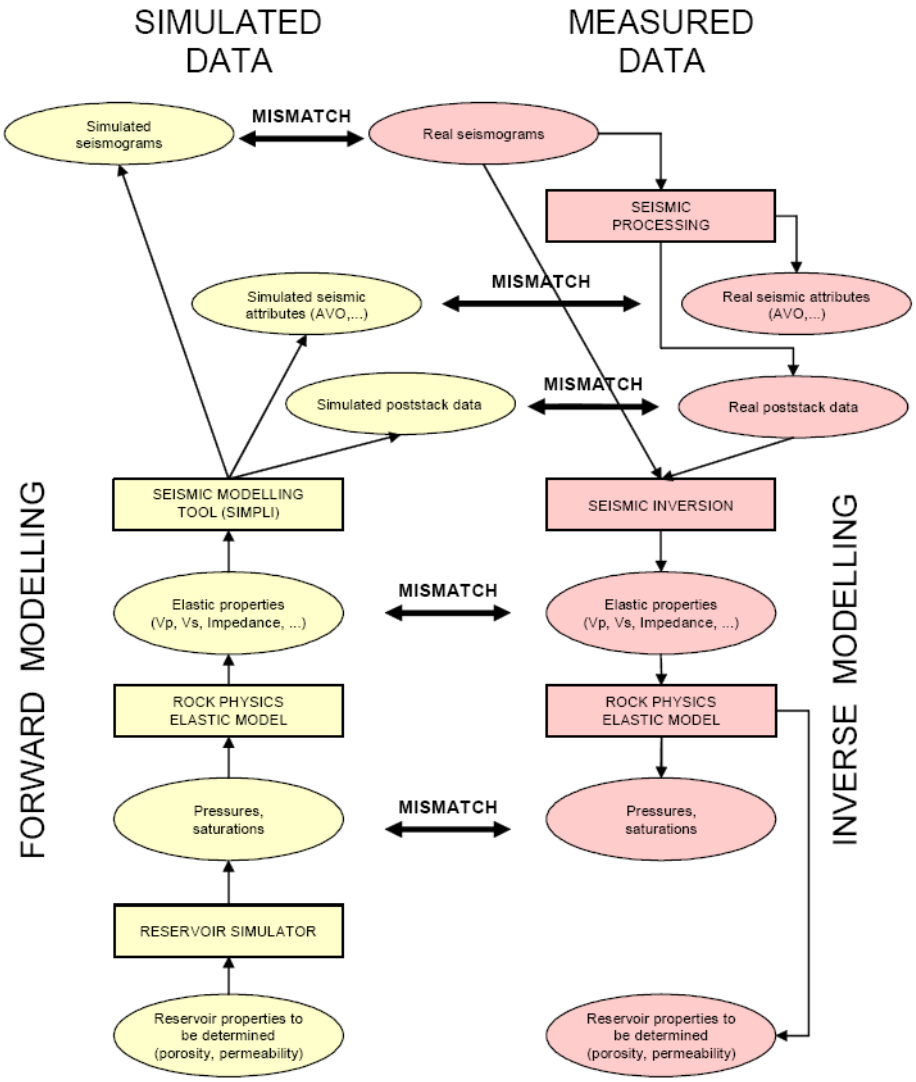


Figure 6.1: Forward and inverse modelling. The left column illustrates the coupled modelling tool, while the right column illustrate various levels of data processing/inversion. In a model updating process or when conditioning a model to time-lapse data, the conditioning may be introduced at different levels corresponding to where the mismatch between simulated and measured data is evaluated.

son's ratio can be constructed by a rock physics model, from dynamic data (pressures and saturations) obtained by a reservoir simulator [76, 78, 79].

Inversion of the elastic properties to saturation and pressure is another possibility for comparing seismic observations to the simulated data. In the literature

several publications have discussed this conversion; see for example [81, 82, 83].

To obtain consistency between the matched data a method for coupling the simulation grid and the seismic grid, which will be independent of the seismic mismatch level, is needed. Seismic data have a much higher areal resolution than the simulation model, while the opposite is the case for the vertical resolution. In addition, there are well log data and the geological model, all with different degrees of spatial resolution. When coupling the models, this effect has to be taken into account, and in a history matching process it will be important to incorporate the various types of data in a consistent way. This includes downscaling and upscaling of uncertainties and properties, as well as data correlations [84]. Behrens et al. [11] proposed a downscaling algorithm based on the sequential Gaussian simulation with Block kriging (SGSBK), and a simulation study of this method can be found in Report 2.

6.1 Rock Physics Model

A rock physics model is needed to link the elastic parameters which govern wave propagation, and the rock and fluid properties which govern fluid flow. The rock physics model calculates the elastic properties (e.g. acoustic impedance and Poisson's ratio) from seismic velocities and densities, based on the reservoir state variables (saturations and pressures). The relationship to perform forward modelling from the reservoir to the elastic domain, is related on principles from a general theory, individually applied for one reservoir field to another.

The rock physics model is often composed of empirical laws calibrated to laboratory measurements and analytical formulas. A comprehensive overview of rock physics can be found in Mavko et al. [6].

In many applications an important part of the rock physics model is Gassmann's equation [5], where the effective bulk modulus of a fluid-saturated rock, K, is connected to the bulk modulus of the mixture of the pore fluids, K_{mix}, the dry rock frame, K_{dry}, and the matrix mineral, K_m:

$$K = K_{dry} + \frac{\left(1 - \frac{K_{dry}}{K_m}\right)^2}{\frac{\phi}{K_{mix}} + \frac{(1-\phi)}{K_m} - \frac{K_{dry}}{(K_m)^2}}, \tag{6.1}$$

where ϕ is the porosity. Gassmann's equation assumes that both the solid and fluid phase are homogeneous and isotropic, and that the frequencies of the seismic wave are so low that the fluid pressures have time to equilibrate throughout the pore space.

Further, according to the assumption that the fluids are not too viscous, the effective shear modulus can be set equal to the shear modulus of the dry rock

frame

$$G = G_{dry}. \tag{6.2}$$

The effective density ρ may be defined as a volume average of the mineral density and the fluid density

$$\rho = \phi(S_w\rho_w + S_g\rho_g + S_o\rho_o) + (1-\phi)\rho_m, \tag{6.3}$$

where S_o, S_g, S_w are respectively the oil, gas and water saturations. If K, G and ρ are obtained from the Equations (6.1)-(6.3), the isotropic compressional and shear wave velocities can be computed from

$$V_p = \sqrt{\frac{K + (4G/3)}{\rho}} \tag{6.4}$$

$$V_s = \sqrt{\frac{G}{\rho}}, \tag{6.5}$$

and the compressional and shear acoustic impedance are given as

$$I_p = \rho V_p \tag{6.6}$$

$$I_s = \rho V_s. \tag{6.7}$$

The Poisson's ratio, ν, may be calculated from

$$\nu = \frac{\gamma - 2}{2\gamma - 2} \tag{6.8}$$

where γ is given as

$$\gamma = \left(\frac{V_p}{V_s}\right)^2. \tag{6.9}$$

The bulk modulus of the pore fluid, K_{mix}, can be obtained by using Wood's law [6]

$$\frac{1}{K_{mix}} = \frac{S_o}{K_o} + \frac{S_w}{K_w} + \frac{S_g}{K_g}, \tag{6.10}$$

where the bulk moduli of the individual fluids, (K_o, K_w, K_g), are derived from the Batzle & Wang equations [85].

The dry rock bulk modulus, K_{dry}, and shear modulus, G_{dry}, are assumed to be given from Equations (6.4) and (6.5) using dry rock density, ρ_{dry}, and velocity values, $(V_{p,dry}, V_{s,dry})$. The dry rock wave velocities and densities can be estimated from empirical functions of effective pressure and porosity, for instance the critical porosity model or from Krief's relation [6] or laboratory core measurements.

In the petroleum industry Gassmann's equation is widely used to calculate fluid saturation effects. However, in many applications the properties of the reservoir rocks will not be compatible with Gassmann, i.e. there may be several fluids forming a non-homogeneous mixture, there may be more than one solid phase, the frequencies of the seismic waves may be too high, and the medium may be anisotropic due to the alignment of flat pores. In carbonate rocks for instance, the main contribution to the compressibility is coming from micro-fractures or cracks since they are the most compliant part of the pore space. The distribution and form of the pore spaces and the way in which the fluid phases are mixed are decisive for the effective elastic properties of the medium and, hence, for the seismic response.

Alternative methods not limited to the Gassmann assumptions are for example: the self-consistent method (SCA), the differential effective medium theory (DEM) (e.g. Kuster-Toksöz model), see Johansen et al. [86] for further discussion, and the so called T-matrix approach by Jakobsen et al. [8]. Common for the SCA, DEM and T-matrix approach, is that a more detailed description of the microstructure of the medium is required. In such approaches the shape of the pore space is typically limited to ellipsoids which are defined by their aspect ratios (ratio between the short and long axis) and their volume fractions (concentrations). Although pores usually are not ellipsoids, it is assumed that any structure may be approximated by a suitable blend of different aspect ratios. Cracks may be modelled as cavities with a very small aspect ratio. Information about the micro-structure of the medium is rarely known. Thus, in order to compute the fluid substitution effect, one is forced to calibrate the model so that a given set of aspect ratios and volume fractions match the observed data.

6.2 Seismic Modelling

Seismic reflection surveying is a widely used geophysical technique, and has gained popularity because raw data can be processed to create a seismic section which is an image of the structure of the subsurface. To achieve such images controlled sources of seismic energy are deployed to generate elastic waves in the underground. These energy waves are reflected, and refracted at the layer boundaries, or diffracted by the heterogeneities in the sub-surface. Hence, large datasets are created that give information on the geological structure.

The details in the seismic sections of the geological structure are on a scale from tens of meteres to the whole lithosphere. A seismic section is in a way similar to a depth section of the geology, but they are also fundamentally different from each other. Raw data generally have a complex appearance and suitable processing techniques are necessary to recover the signature of the seismic signal. In order to interpret seismic sections reliably, we need to understand seismic wave phenomena, acquisition footprints and processing artefacts.

Seismic modelling involves the calculation of synthetic seismic traces [87]. A synthetic seismic trace represents the combined reflection response of the layered ground (i.e. the output for a spike input) and the recording system to a seismic pulse. Consider for example a normal incident compressional ray on an interface separating two media with different density and velocity. Incident energy will be partitioned according to the contrast in the acoustic impedance at the interface, i.e. partly reflected to the surface and partly transmitted.

The normal incidence reflection coefficient, r, is given by

$$r = \frac{\rho_2 V_{p_2} - \rho_1 V_{p_1}}{\rho_2 V_{p_2} + \rho_1 V_{p_1}} = \frac{I_{p_2} - I_{p_1}}{I_{p_2} + I_{p_1}} \tag{6.11}$$

where ρ_1, V_{p_1}, I_{p_1} and ρ_2, V_{p_2}, I_{p_2} are the density, P-wave velocity and acoustic impedance of the first and second layer, respectively. The reflection coefficient defined in Equation (6.11) takes values between $-1 \leq r \leq +1$, where a negative value of r implies a phase change of π in the reflected ray. Notice that the velocity changes as a function of depth, since the physical properties may be different for the individual layers. The velocity may also vary in the horizontal direction, due to lithological variation within the individual layers.

Figure 6.2 shows the correspondence between the geological sections, the alternation in the acoustic impedance and the reflection coefficient as a function of the depth. The reflected pulses are received by the detector at the surface, and the size of the amplitude of the signal is determined by the travelling distance and the reflection coefficient at the boundary of the layers. The arrival times of the pulses are given by the depths to the layer boundaries and the propagation velocities between them. Hence, the sum of the individual reflections are enclosed in the synthetic seismogram in their correct travel-time relationship.

A synthetic seismogram $y(t)$ may be considered in its simplest form as the assumed source function $s(t)$ convolved with a reflectivity function $r(t)$ representing the contrasts in acoustic impedance in the layered model

$$y(t) = s(t) \star r(t). \tag{6.12}$$

This process is also called a 1D convolution. However, in general the geological structure in both overburden and reservoir may be more complex. Hence, an accurate modelling is required, and typically finite-difference modelling (FDM) or

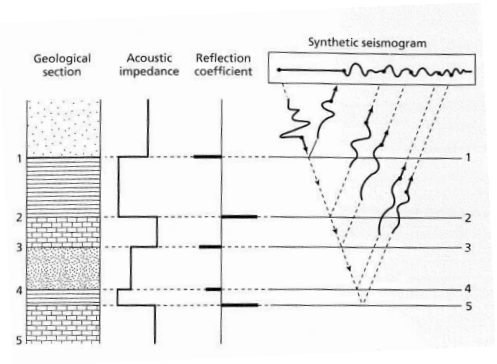


Figure 6.2: Output of seismic traces from different geological sections.

ray-tracing methods are applied. FDM is a standard seismic modelling technique that models the complete elastic wave field on a grid of depth points by using the elastodynamic wave equation [88]. Ray-tracing methods are based on the concept that seismic energy of infinitely high frequency follows a path determined by the ray-tracing equations. These equations explain physically how energy continues in the same direction up to the point where the rays are refracted by velocity variations [89].

To perform a seismic modelling of complex models the so called Simulated Prestack Local Imaging (SimPLI) method may be used [7]. The method is based on theories related to seismic resolution and seismic imaging, and is a method for modelling the seismic response of oil reservoirs. The results are prestack migrated images, in time or depth, simulated directly from angle dependent reflectivity grids, without needing prestack synthetic traces. The process is fast, and is an extension of 1D convolution techniques in time-domain, but with 2D/3D convolution directly in space-domain. The method is able to predict 2D/3D effects of acquisition, resolution and illumination, where these effects take both the overburden properties and the lateral and vertical survey characteristics into account. A coupling between the reservoir simulator and this seismic modelling tool is developed in Paper C.

6.3 Assimilation of Seismic Data

Assimilation of seismic data require that a corresponding error covariance matrix is established. The error covariance matrix includes the noise level in the seismic data and may also include a correlation structure in the measurement error. In synthetic cases the error covariance matrix used in the EnKF scheme should be

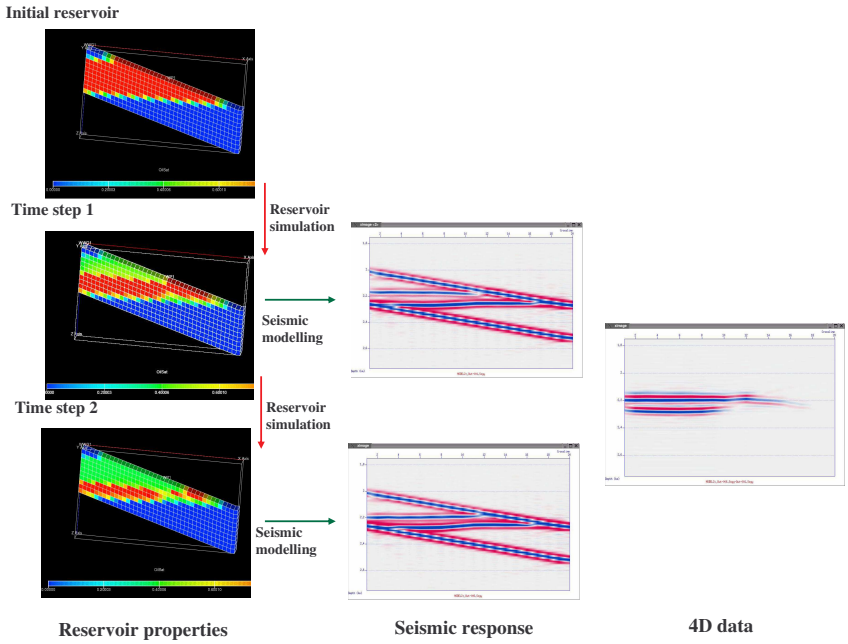


Figure 6.3: Illustration of a coupling process between a reservoir simulation model, with grid size $n_x = 40$ and $n_z = 10$, and a seismic forward model. The reservoir has a gas injector in the gas cap, located in grid cell (1,1), an oil producer, located in grid cell (22,1), penetrating three layers, and an aquifer at the bottom. The figure shows the initial oil saturation and at two simulation time-steps. At each time-step the saturation, pressure and porosity at each grid cell are used as an input to the seismic forward model, and a synthetic seismic response is obtained. Further a difference between the seismic response at time step 2 and time step 1 is computed, called 4D data. The 4D seismic response indicates a movement of both the gas-oil and water-oil contact.

equal to the error covariance matrix used in the sampling of the perturbed noise. For real field reservoirs the error covariance matrix should be estimated from the observed seismic data and then used in the EnKF scheme. The variance can be defined as an absolute value or as a percentage of the seismic signal and may vary from grid cell to grid cell. Commonly the correlation is given as exponential, gaussian or spherical with a given range (see Equations (5.5), (5.6) and (5.7)).

The results from the EnKF runs, where seismic data are used, show that there is a close relationship between the improvement of the characterization and the error level used for the seismic data. However, as seen in Paper A, a positive

impact could be found regardless of the high noise level in the real field time-lapse seismic data. A proper representation of the seismic measurement error covariance matrix may also be important to achieve satisfactory EnKF results.

When using 3D seismic and time-lapse seismic in a data assimilation process a coupling method for combined modelling of reservoir fluid flow and seismic properties has to be applied. The method must control the seismic forward model and the reservoir simulator, as well as the data flow between those two. A normal simulator run will then in addition generate simulated seismic data at the time where data from seismic surveys are available. Figure 6.3 illustrates the coupling process, where the reservoir simulator is coupled to a seismic modelling tool.

In this project we have developed different couplings to rock physics models so that elastic parameters like acoustic impedance and Poisson's ratio can be assimilated. The rock physics models used in the papers are either based on a poro elastic model (e.g. the Gassmann equation [5]), an inclusion based model (e.g. the Kuster-Toksöz method [6]), or the T-matrix approach [8]. In Paper C we have also developed a coupling to a fast seismic modelling tool which simulates prestack depth migrated sections such that seismic waveform data can be assimilated. The results in this work show that incorporating both production data and seismic data have a positive impact on the reservoir evaluation. However, the use of an incorrect rock physics model, as may be the case in Report 1, shows that large modifications in the model variables are obtained. Such large updates can lead to model instabilities and make the ensemble corrupt. The seismic dataset may contain outliers, and we have implemented a filter to ensure that such problematic data are handled in a proper way. A criterion is used where the distance between the observed and simulated data is compared with the sum of the measurement and the predicted standard deviations. If the distance is too large the measurement standard deviations for the outliers are increased and the influence of these data are decreased in the update.

In an EnKF method an efficient coupling between the reservoir simulator, the seismic forward simulator and the data assimilation will be of importance, since normally a size of 100 ensemble members are used in the EnKF analysis. Real field reservoir models typically include a large number of grid cells and the forward simulation of both the fluid flow model and the seismic model then becomes very time consuming. Hence, to reduce the computational time we have developed a parallel implementation of the system so that the fluid flow simulation and the forward simulation of the seismic data are performed in parallel for all the members. This implementation has been used in Paper A and Report 1, where we have performed an update of two different real field reservoir models from the North Sea.

Incorporating time-lapse seismic data in sequential assimilation methods are complicated since such data are time difference data dependent on the model state

of the reservoir at two different time steps. However, time-lapse seismic data can be assimilated properly in the EnKF when an EnKS term is included as discussed in Section 4.2. This formulation is based on the general sequential assimilation method presented in Section 3.2, where data related to more than one time instant are assimilated. Notice that the filter solution of the parameters in the reservoir model is equal to the smoother solution. Hence, to reduce the computational time, only the smoother solution of the dynamic variables has to be calculated.

In reservoir applications both 3D seismic and time-lapse seismic data may introduce a large amount of data in the assimilation. In all the papers except from Paper C, the size of the seismic cube is similar to the number of grid cells in the reservoir simulation model. Assimilation of seismic waveform data, as performed in Paper C, introduced even more data in the update. Thus, to assimilate such large datasets the use of efficient EnKF schemes are required. In Section 4.5 and Paper B this problem is discussed and an efficient global EnKF scheme is presented.

Incorporation of large seismic datasets may introduce problems related to an underestimation of the uncertainty of the model variables when few ensemble members are used in the update. Thus, an assimilation of seismic data with a local analysis scheme has been investigated in Papers D and E. Here we assume that only seismic data within a certain distance from a state variable will impact the analysis of this state variable. Promising results of the reservoir charachterization and the uncertainty estimation are obtained when the combined global and local scheme is applied. The local analysis can be implemented in parallel, and each local update will involve relatively small matrices in the assimilation, so that the method inverts many small matrices instead of one large. Hence, with the local scheme, large seismic datasets can be assimilated at a low computational effort.

Chapter 7

Conclusions and Further Work

We have presented a method based on the ensemble Kalman filter (EnKF) for continuous model updating with respect to the combination of production data, 3D seismic data and time-lapse seismic data. The method has been implemented and applied to synthetic models and real field cases from the North Sea. In this project a coupled reservoir-seismic model is developed for simulation of production data as well as different kinds of seismic data: Inverted data (acoustic impedance and Poisson's ratio) and prestack depth migrated sections (seismic waveform data). The model system consists of a commercial reservoir simulator coupled to existing rock physics and seismic modelling software. In the estimation both static parameters (permeability, porosity) and dynamic variables (pressures and saturations) are updated by the EnKF method. Furthermore, we have updated a lithology parameter (clay ratio) which is linked to a rock physics model, and the fracture density which is updated in a synthetic fractured reservoir. In reservoir applications seismic data may introduce a large amount of data in the assimilation schemes, and the computational time becomes expensive. In this work efficient EnKF schemes are used to handle such large datasets, where challenging aspects such as the inversion of a large covariance matrix and potential loss of rank are considered.

Time-lapse seismic data may be difficult to assimilate since they are time difference data, i.e. data which are related to the model variable at two time instances. Here we have presented a general sequential Bayesian formulation which incorporates time difference data, and we show that the posterior pdf includes both a filter and a smoother solution. Further, we show that when time difference data are used in the EnKF, a combination of the ensemble Kalman filter and the ensemble Kalman smoother has to be applied. However, the method is still completely recursive, with little additional cost compared to the standard EnKF.

An underestimation of the uncertainty of the model variables may become a problem when few ensemble members and a large amount of data are used during

the assimilation. Thus, to improve the reservoir model updating we have proposed a methodology based on a combination of a global and a local analysis scheme. In this context, the global scheme assimilates the production data, while the local scheme assimilates the seismic data. The local scheme assumes that only seismic data within a given range from a model variable will influence on the analysis of this particular model variable. By solving the global and the local schemes in a sequential process, we obtain the global analysis ensemble conditioned to both production and seismic data.

In the examined reservoir cases the developed EnKF provides satisfactory characterization of the reservoir and history matching results. Furthermore, a proper representation and prediction of uncertainty in the reservoir simulation models are obtained.

Promising results have been achieved and further improvements are suggested. Especially the effect of including seismic data at different mismatch levels, see Figure 6.1, when conditioning reservoir models should be further investigated. The quantification of the uncertainties from the acquisition stage to the final processed seismic data will then become important.

In a history matching process it will be important to incorporate the various types of seismic data in a consistent way. That is, how to assimilate data, which are given on a wide range of scale; from well log data, via geological models to reservoir models and seismic data. In future work improved methods for resolving the scale problem, which includes upscaling and downscaling of reservoir properties and uncertainties, should be developed. In Report 2 a simulation study of one specific downscaling algorithm (SGSBK) was performed, and good results were obtained. However, more extensive research on different upscaling/downscaling algorithms is required on real field seismic data. The EnKF might be further developed with a coupling to an upscaling/downscaling interface so that measured and simulated seismic data, represented on different scales, can be assimilated in a proper way.

In Papers D and E we show that an improvement of the reservoir characterization can be obtained by introducing an uncertainty connected to the rock physics model, and in this study lithology was updated as an additional static parameter to permeability and porosity. Possibilities of updating uncertain parameters in rock physics and seismic modelling should be further investigated. This issue may become important, when the mismatch between the simulated and measured seismic data is big. As seen in Report 1, such a large mismatch may cause large modifications of the reservoir parameters in the data assimilation. An ability to perform an adjustment of the rock physics and seismic modelling parameters may then prevent undesirable updates in the EnKF analysis scheme.

An underestimation of the uncertainty of the model variables is of concern when large amount of data and few ensemble members are applied in the EnKF.

Regarding this issue, the combined global and local scheme proposed in Papers D and E gave promising results, but the method should be examined on real field cases before final conclusions are drawn. A potential disadvantage of the local scheme is that non-physical modes may occur in the analysis fields, because the updates are performed independently in each local region. Especially when observations with a fairly high white noise (typically real field seismic) are assimilated. It is then required to use a large influence region in the local update in order to preserve the smoothness of the analysis fields.

In past work on ensemble Kalman filters [63] it was found that in order to compensate for the tendency of a small ensemble to underestimate uncertainty, it may be suitable to inflate the forecast error covariance before each analysis step. The standard variance inflation approach is to multiply the forecast error covariance by a constant factor slightly larger than one. The variance inflation approach is not used in our experiments, and in further work we should investigate the improvement of inflating the covariance when assimilating seismic data.

A proper treatment of non-Gaussian and multi-modal facies distributions in the EnKF is an unresolved issue. This is a difficult task since then the underlying pdf for the facies variables can not be described properly by the mean and the covariance which is used in the EnKF analysis scheme. In such cases, typically the predicted measurements will also be multimodal. In particular the pluri-Gaussian and truncated Gaussian methods are promising approaches for handling the problem. Liu [90] showed, by providing good results, that, if underlying Gaussian variables can be used to define the facies distribution, then the Gaussian variables could be updated instead.

Estimation of structural parameters is another issue to explore. This is also a difficult task because the EnKF analysis steps combine the ensemble members, and this requires that all the members of the ensemble are using the same simulation grid.

Additional model parameters such as initial fluid contacts, fault transmissivity multipliers and vertical transmissivity multipliers may also be estimated in future EnKF applications, and Evensen [3] has shown that an update of such parameters may have a positive impact on the results in real field cases.

Bibliography

[1] A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation*. Philadelphia: SIAM, 2005.

[2] H. Omre, "Stochastic reservoir models conditioned to non-linear production history observations," in *Geostatistics 2000, Cape Town, pp. 166-175*, April 2000. Proceedings of the Sixth International Geostatistical Congress held in Cape Town, South Africa.

[3] G. Evensen, *Data Assimilation, The Ensemble Kalman Filter*. Berlin: Springer, 2006.

[4] R. E. Kalman, "A new approach to linear filter and prediction problems," *J. Basic. Eng.*, vol. 82, pp. 35–45, 1960.

[5] F. Gassmann, "Über die Elastizität poröser Medien," *Vier. der Natur. Gesellschaft in Zurich*, vol. 96, pp. 1–23, 1951.

[6] G. Mavko, T. Mukerju, and J. Dvorkin, *The Rock Physics Handbook*. Cambridge University Press, 2003.

[7] I. Lecomte, H. Gjøystal, and Å. Drottning, "Simulated Prestack Local Imaging: a robust and efficient interpretation tool to control illumination, resolution, and time-lapse properties of reservoirs," *Expanding Abstracts, SEG International Exposition and Seventy-Third Annual Meeting, Dallas*, pp. 1529–1532(RCT 7.2), 2003.

[8] M. Jakobsen, J. A. Hudson, and T. A. Johansen, "T-matrix approach to shale acoustics," *Geophys. J. Int.*, vol. 54, pp. 219–246, 2003.

[9] G. Evensen, "Sampling strategies and square root analysis schemes for the EnKF," *Ocean Dynamics*, vol. 54, pp. 539–560, 2004.

[10] J. L. Anderson, "An ensemble adjustment Kalman filter for data assimilation," *Mon. Weather Rev.*, vol. 129, pp. 2884–2903, 2001.

[11] R. A. Behrens, M. K. MacLeod, and T. Tran, "Incorporating Seismic Attribute Maps in 3D Reservoir Models," *SPE Reservoir Evaluation and Engineering, SPE 36499*, 1998.

[12] A. Kirsch, *An Introduction to the Mathematical Theory of Inverse Problems*, vol. 120 of *Applied Mathematical Sciences*. New York: Springer-Verlag, 1996.

[13] L. N. Trefethen and D. Bau, *Numerical linear algebra*. Philadelphia: SIAM, 1997.

[14] H. S. Migon and D. Gamerman, *Statistical Inference: an Integrated Approach*. New York: Arnold, 1999.

[15] H. Omre and H. Tjelmeland, "Petroleum geostatistics, in Baafi and Schofield, eds., Geostatistics Wollongong '96," *Kluwer Academic Publishers*, 1997.

[16] G. R. Gavalas, P. C. Shah, and J. H. Seinfeld, "Reservoir history matching by bayesian estimation," *SPE Journal*, vol. 16, pp. 337–350, 1976.

[17] D. S. Oliver, "Incorporation of transient pressure data into reservoir characterization," *In Situ*, vol. 18(3), pp. 243–275, 1994.

[18] D. S. Oliver, "On conditional simulation to inaccurate data," *Mathematical Geology*, vol. 28, pp. 811–817, 1994.

[19] G. Casella and R. L. Berger, *Statistical inference*. Belmont, California: Duxbury Press, 1990.

[20] D. McLaughlin and L. R. Townley, "A reassessment of the groundwater inverse problem," *Water Resources Research*, vol. 32, pp. 1131–1161, 1996.

[21] R. A. Johnson and D. W. Wichern, *Applied Multivariate Statistical Analysis*. New Jersey: Prentice Hall, 1998.

[22] S. M. Ross, *Introduction to probability models*. San Diego: Academic Press, 1997.

[23] D. S. Oliver, L. B. Cunha, and A. C. Reynolds, "Markov chain Monte Carlo methods for conditioning a permeability field to pressure data," *Mathematical Geology*, vol. 29, pp. 61–91, 1997.

[24] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and T. E., "Equation of state calculations by fast computing machine," *Journal of Chemical Physics*, vol. 21, pp. 1087–91, 1953.

[25] S. German and D. German, "Stochastic relaxtion, Gibbs distribution and the Bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, pp. 721–741, 1984.

[26] A. E. Gelfand and A. M. F. Smith, "Sampling-based approaches to calculating marginal densities," *Journal of the American Statistical Association*, vol. 85, pp. 398–409, 1990.

[27] B. K. Hegstad and H. Omre, "Uncertainty assessment in history matching and forecasting, *in* E. Y. Baafi and N. A. Schofield (eds)," *Geostatistics Wollongong '96*, vol. I: Kluwer Academic Publisher, pp. 585–596, 1997.

[28] H. Tjelmeland, "A note on the Bayesian approach to history matching of reservoir characteristics, in Pawlowsky-Glahn, eds.," in *Proceedings of the Third Annual Conference of the International Association for Mathematical Geology*, (Barcelona, Spain), September 1997. p.201-214.

[29] H. Omre and O. P. Lødøen, "Improved production forecasts and history matching using approximate fluid-flow simulators," *Society of Petroleum Engnieers Journal*, vol. 9(3), pp. 339–351, 2004.

[30] P. K. Kitanidis, "Quasi-linear geostatistical theory for inversing," *Water Resource Research*, vol. 31, pp. 2411–2419, 1995.

[31] R. Li, A. C. Reynolds, and D. S. Oliver, "Sensitivity coefficients for three-phase flow history matching," *Journal of Canadian Petroleum Technology*, vol. 42, pp. 70–77, 2003.

[32] J. Nocedal, "Updating quasi-Newton matrices with limited storage," *Math. Comp.*, vol. 35, pp. 773–782, 1980.

[33] F. Zhang and A. C. Reynolds, "Optimization algorithms for automatic history matching of production data," *Proceedings of 8th European Conference on the Mathematics of Oil Recovery*, 2002.

[34] G. Gao and A. Reynolds, "Quantifying the uncertainty for the PUNQ-S3 problem in a bayesian setting with the RML and EnKF," *SPE reservoir simulation symposium (SPE 93324)*, 2005.

[35] S. Subbey, M. Christie, and M. Sambridge, "Uncertainty reduction in reservoir modeling," *SIAM Con. Math.*, no. 295, pp. 457–467, 2002.

[36] P. J. van Leeuwen and G. Evensen, "Data assimilation and inverse methods in terms of a probabilistic formulation," *Mon. Weather Rev.*, vol. 124, pp. 2898–2913, 1996.

[37] G. Evensen and P. J. van Leeuwen, "An ensemble Kalman smoother for nonlinear dynamics," *Mon. Weather Rev.*, vol. 128, pp. 1852–1867, 2000.

[38] P. S. Maybeck, *Stochastic models, estimation, and control*. New York: Academic Press, 1979.

[39] H. W. Sorenson, "Least-Squares estimation: from Gauss to Kalman," *IEEE Spectrum*, vol. 7, pp. 63–68, 1970.

[40] M. West and J. Harrison, *Bayesian Forecasting and Dynamic Models*. New York: Springer, 1997.

[41] N. A. C. Cressie, *Statistics for Spatial Data*. Iowa State University: Wiley, 1991.

[42] G. Evensen, "Using the extended kalman filter with a multilayer quasi-geostrophic ocean model," *J. Geophys. Res.*, vol. 97, pp. 17905–17924, 1992.

[43] H. Madsen and R. Cañizares, "Comparison of extended and ensemble Kalman filters for data assimilation in coastal area modelling," *Int. J. Numer. Meth. Fluids*, vol. 31, pp. 961–981, 1999.

[44] G. Evensen, "Sequential data assimilation with nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics," *J. Geophys. Res.*, vol. 99, pp. 10143–10162, 1994.

[45] I. Szunyogh, E. J. Kostelich, G. Gyarmati, D. J. Patil, B. Hunt, E. Kalnay, E. Ott, and J. A. Yorke, "Assessing a Local Ensemble Kalman Filter: Perfect Model Experiments with the NCEP Global Model," *Tellus*, vol. 57A, pp. 528–545, 2005.

[46] L. J. Natvik and G. Evensen, "Assimilation of ocean colour data into a bio-chemical model of the North Atlantic. part 1. Data assimilation experiments," *J. Marine. Sys.*, vol. 40-41, pp. 127–153, 2003.

[47] G. Nævdal, T. Mannseth, and E. Vefring, "Near well reservoir monitoring through ensemble Kalman filter," *Proceeding of SPE/DOE Improved Oil recovery Symposium (SPE 75235)*, 2002.

[48] G. Nævdal, G. L. Johnsen, S. I. Aanonsen, and E. Vefring, "Reservoir monitoring and continuous model updating using the ensemble Kalman filter," *SPE Annual Technical Conference and Exhibition (SPE 84327)*, 2003.

[49] Y. Gu and D. S. Oliver, "History Matching of the PUNQ-S3 Reservoir Model Using the Ensemble Kalman Filter," *SPE Annual Technical Conference and Exhibition, SPE 89942*, 2004.

[50] R. J. Lorentzen, G. Nævdal, B. Vallés, A. M. Berg, and A.-A. Grimstad, "Analysis of the ensemble Kalman filter for estimation of permeability and porosity in reservoir models," *SPE Annual Technical Conference and Exhibition, SPE 96375*, 2005.

[51] V. Haugen, L. J. Natvik, G. Evensen, A. Berg, K. Flornes, and G. Nævdal, "History Matching Using the Ensemble Kalman Filter on a North Sea Field Case," *SPE Annual Technical Conference and Exhibition, SPE 102430*, 2006.

[52] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. San Diego, Calif.: Academic Press, 1970.

[53] L. Bertino, G. Evensen, and H. Wackernagel, "Sequential data assimilation techniques in oceanography," *International Statistical Review*, vol. 71, pp. 223–241, 2003.

[54] G. Evensen, "The ensemble Kalman filter: Theoretical formulation and practical implementation," *Ocean Dynamics*, vol. 53, pp. 343–367, 2003.

[55] O. Leeuwenburg, G. Evensen, and L. Bertino, "The impact of ensemble filter definition on the assimilation of temperature profiles in the Tropical Pacific," *Q. J. R. Meteorol. Soc., to appear*, 2006.

[56] M. K. Tippett, J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, "Ensemble square-root filters," *Mon. Weather Rev.*, vol. 131, pp. 1485–1490, 2003.

[57] C. H. Bishop, B. J. Etherton, and S. J. Majumdar, "Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects," *Mon. Weather Rev.*, vol. 129, pp. 420–436, 2001.

[58] E. Ott, B. Hunt, I. Szunyogh, A. V. Zimin, E. Kostelich, M. Corazza, E. Kalnay, D. J. Patil, and J. A. Yorke, "A local ensemble Kalman filter for atmospheric data assimilation," *Tellus Ser. A*, vol. 56A, pp. 415–428, 2004.

[59] J. S. Whitaker and T. M. Hamill, "Ensemble data assimilation without perturbed observations," *Mon. Weather Rev.*, vol. 130, pp. 1913–1924, 2002.

[60] V. E. Haugen and G. Evensen, "Assimilation of SLA and SST data into an OGCM for the Indian ocean," *Ocean Dynamics*, vol. 52, pp. 133–151, 2002.

[61] K. Brusdal, J. Brankart, G. Halberstadt, G. Evensen, P. Brasseur, P. J. van Leeuwen, E. Dombrowsky, and J. Verron, "An evaluation of ensemble based assimilation methods with a layered OGCM," *J. Marine. Sys.*, vol. 40-41, pp. 253–259, 2003.

[62] M. Zafari and A. C. Reynolds, "Assessing the uncertainty in reservoir description and performance predictions with the ensemble Kalman filter," *SPE Annual Technical Conference and Exhibition, SPE 95750*, 2005.

[63] J. L. Anderson and S. Anderson, "A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts," *Mon. Weather Rev.*, vol. 127, pp. 2741–2758, 1999.

[64] D. T. Pham, "Stochastic methods for sequential data assimilation in strongly nonlinear systems," *Mon. Weather Rev.*, vol. 129, pp. 1194–1207, 2001.

[65] G. A. Kivman, "Sequential parameter estimation for stochastic systems," *Nonlinear Processes in Geophysics*, vol. 10, pp. 253–259, 2003.

[66] A. Doucet, N. de Freitas, and N. Gordon (Eds.), *Sequential Monte Carlo Methods in Practice*. New York: Springer-Verlag, 2001.

[67] J. D. Kepert, "On ensemble representation of the observation-error covariance in the ensemble Kalman filter," *J. Basic. Eng.*, vol. 82, pp. 35–45, 2004.

[68] H. Wackernagel, *Multivariate Geostatistics*. Springer-Verlag, 1998.

[69] M. W. Davis, "Production of conditional simulations via the LU decomposition of the covariance matrix," *Mathematical Geology*, vol. 19, pp. 91–98, 1987.

[70] A. Kirchner and S. A. Shapiro, "Fast repeat modelling of time-lapse seismograms," *Geophys. Prosp.*, vol. 49(5), pp. 557–569, 2001.

[71] E. Arenas, C. van Kruijsdijk, and T. Oldenziel, "Semi-Automatic History-Matching Using the SPE Pilot Point Method Including Time-Lapse Seismic Data," in *Proceedings of the SPE Annual Technical Conferance and Exhibition*, (Orleans, Lousiana), September-October 2001.

[72] X. Huang, L. Meister, and R. Workman, "Reservoir Caracterization by Integration of Time-Lapse Seismic and Production Data," in *Proceedings of the 1997 SPE Annual Technical Conferance and Exhibition*, (San Antonio, Texas), October 1997. SPE 38695.

[73] J. Waggoner, A. Comminelli, and R. Seymour, "Improved Reservoir Modelling with Time-Lapse Seismic in a Gulf of Mexico Gas Condensate Reservoir," in *Proceedings of the SPE Asia Paciffic Oil an Gas Conferance and Exhibition*, October 2002. SPE 77956.

[74] K. Fagervik et al., "History Matching of Reservoir Flow models Using 4D Seismic," in *Proceedings of EAGE 63th Conferance and Technical Exhibition*, (Amsterdam), June 2001.

[75] O. Gosselin, S. van den Berg, and A. Cominelli, "Integrated History-matching of Production and 4D Seismic Data," in *Proceedings of the SPE Annual Technical Conference and Exhibition*, (New Orleans, Louisiana), 30 Sept.-3 Oct. 2001. SPE 71599,.

[76] O. Gosselin, S. Aanonsen, I. Aavatsmark, A. Cominelli, R. Gonard, M. Kolasinski, F. Ferdinandi, and K. Kovacic, L. andNeylon, "History Matching Using Time-Llapse Seismic (HUTS)," in *Proceedings of the SPE Annual Technical Conference and Exhibition*, (Denver, Colorado), 30 Sept.-3 Oct. 2003. SPE 84464.

[77] S. Aanonsen, A. Cominelli, O. Gosselin, I. Aavatsmark, and T. Barkve, "Integration of 4D Data in the History Match Loop by Investigating Scale dependent Correlations in the Acustic Impedance Cube," in *Proceedings of the 8th European Conferance on the Mathematics of Oil Recovery*, (Freiberg, Germany), 3-6 Sept. 2002.

[78] S. Aanonsen, I. Aavatsmark, T. Barkve, A. Cominelli, R. Gonard, O. Gosselin, M. Kolasinski, and H. Reme, "Effect of Scale Dependent Data Correlations in an Integrated History Matching Loop Combining Production Data and 4D Seismic Data," in *Proceedings of the SPE Reservoir Simulation Symposium*, (Houston, Texas), Feb. 2003. SPE 79665.

[79] M. C. Haverl, M. Aga, and E. Reiso, "Integrated workflow for quantitative use of time-lapse seismic data in history matching: A North Sea field case," *Presented at the* 14^{th} *Europec Biennial Conference, Madrid, 13-16 June*, 2005.

[80] O. Yilmaz and S. Doherty, eds., *Seismic data analysis: processing, inversion, and interpretation of seismic data*. Tulsa: Society of Exploration Geophysicists, 2001.

[81] M. Landrø and Ø. Kvam, "Pore Pressure Estimation - what can we learn from 4D," *CSEG Recorder*, September 2002.

[82] M. Lygren, K. Fagervik, T. Valen, A. Hetlelid, G. Berge, G. Dahl, L. Sønneland, H. Lie, and I. Magnus, "A method for performing history matching of reservoir flow models using 4D seismic data," *Petroleum Geoscience*, vol. 9, pp. 85–90, 2003.

[83] C. Ribeiro and C. MacBeth, "Inversion for reservoir pressure and saturation changes in the Foinaven field UK," *SEG International Exposition and* 75^{th} *Annual Meeting, Houston, USA*, 2005.

[84] P. M. Doyen, D. E. Psaila, L. D. den Boer, and D. Jans, "Reconciling Data at Seismic and Well Log Scales in 3-D Earth Modelling," *SPE Annual Technical Conference and Exhibition, SPE 39698*, 1997.

[85] M. Batzle and Z. Wang, "Seismic properties of pore fluids," *Geophys.*, vol. 57, pp. 1396–1408, 1992.

[86] T. A. Johansen, Å. Drottning, I. Lecomte, and H. Gjøystal, "An approach to combined rock physics and seismic modelling of fluid substitution effects," *Geophys. Prosp.*, vol. 50, pp. 119–137, 2002.

[87] P. Kearey, M. Brooks, and I. Hill, *An Introduction to Geophysical Exploration*. Blackwell Science, 2002.

[88] J. Virieux, "P-SV wave propagation in heterogeneous media: Velocity-stress finite-difference method," *Geophysics*, vol. 51, pp. 889–901, 1986.

[89] V. Červený, *Seismic Ray Theory*. Cambridge, UK: Cambridge University Press, 2000.

[90] N. Liu and D. S. Oliver, "Ensemble Kalman filter for automatic history matching of geologic facies," *J. Petrolium Sci. and Eng.*, vol. 47, pp. 147–161, 2005.

Part II

Papers and Reports

Paper A

Incorporating 4D Seismic Data in Reservoir Simulation Models Using Ensemble Kalman Filter *

* Proceeding at the 2005 *SPE* Annual Technical Conference and Exhibition, Dallas, U.S.A, October 2005. Accepted for publication in the SPE Journal.

Skjervheim, J. A.; Evensen, G.; Aanonsen, S. I.; Ruud B. O. and T. A. Johansen, Incorporating 4D Seismic Data in Reservoir Simulation Models Using Ensemble Kalman Filter. Proceeding at the 2005 Annual Technical Conference and Exhibition held in Dallas, Texas, U.S.A, October 2005. Accepted for publication in the SPE Journal.

Abstract only. Full-text not available due to publisher restrictions.

SPE 95789

Incorporating 4D Seismic Data in Reservoir Simulation Models Using Ensemble Kalman Filter

J.-A. Skjervheim, CIPR U. of Bergen; G. Evensen, Norsk Hydro; and S.I. Aanonsen, SPE, B.O. Ruud and T.A. Johansen, CIPR U. of Bergen

Copyright 2005, Society of Petroleum Engineers

This paper was prepared for presentation at the 2005 SPE Annual Technical Conference and Exhibition held in Dallas, Texas, U.S.A., 9 – 12 October 2005.
This paper was selected for presentation by an SPE Program Committee following review of information contained in a proposal submitted by the author(s). Contents of the paper, as presented, have not been reviewed by the Society of Petroleum Engineers and are subject to correction by the author(s). The material, as presented, does not necessarily reflect any position of the Society of Petroleum Engineers, its officers, or members. Papers presented at SPE meetings are subject to publication review by Editorial Committees of the Society of Petroleum Engineers. Electronic reproduction, distribution, or storage of any part of this paper for commercial purposes without the written consent of the Society of Petroleum Engineers is prohibited. Permission to reproduce in print is restricted to a proposal of not more than 300 words; illustrations may not be copied. The proposal must contain conspicuous acknowledgment of where and by whom the paper was presented. Write Librarian, SPE, P.O. Box 833836, Richardson, TX 75083-3836, U.S.A., fax 01-972-952-9435.

Summary

A method based on the ensemble Kalman filter (EnKF) for continuous model updating with respect to the combination of production data and 4D seismic data is presented. When the seismic data are given as a difference between two surveys, a combination of the ensemble Kalman filter and the ensemble Kalman smoother has to be applied. Also, special care has to be taken because of the large amount of data assimilated. Still, the method is completely recursive, with little additional cost compared to the traditional EnKF. The model system consists of a commercial reservoir simulator coupled to a rock physics and seismic modelling software. Both static variables (porosity, permeability, rock physic parameters, etc.) and dynamic variables (saturations and pressures) may be updated continuously with time based on the information contained in the assimilated measurements. The method is applied to a synthetic model and a real field case from the North Sea. In both cases, the 4D seismic data are different variations of inverted seismic. For the synthetic case, it is shown that the introduction of seismic data gives a much better estimate of reservoir permeability. For the field case, the introduction of seismic data gives a very different permeability field than using only production data, while retaining the production match.

Paper B

Ensemble Kalman filter with time difference data *

* Submitted to *Computational Geosciences*, April 2006.

Ensemble Kalman filter with time difference data

Jan-Arild Skjervheim[1], Sigurd Ivar Aanonsen[1] and Geir Evensen[2]

[1] *Centre for Integrated Petroleum Research (CIPR), University of Bergen, Norway*

[2] *Hydro Research Centre, Bergen, Norway*

21st April 2006

Abstract

An overview of the ensemble Kalman filter technique for continuous model updating with respect to time difference data i.e, data related to a model state at different times, is presented. The method is a Monte Carlo type sequential Bayesian inversion, and provides an approximate solution to the combined state and parameter estimation problem. The use of time difference data involves a combination of the ensemble Kalman filter and the ensemble Kalman smoother, and the result is a sequential simulation from the a posteriori distribution. In reservoir application time lapse seismic data may introduce a large amount of data in the assimilation schemes. Special challenges are involved concerning rank issues and inversion of a large covariance matrix, and a subspace EnKF algorithm, based on the sampling strategy used in the existing efficient square root scheme, is proposed. A comparison between the traditional and the subspace EnKF scheme has been performed, where both a diagonal and a correlated measurement error covariance matrix has been used. The method is applied to a 3D synthetic reservoir model assimilating production and seismic data, where the model consist of a rock physics modelling software coupled to a reservoir simulator. In the estimation both dynamic variables (pressures and saturations) and static variables (permeability, porosity) are updated based on the information contained in the assimilated measurements. The synthetic case shows that the combined assimilation of production data, 3D seismic data and time difference seismic data is performed in a statistical consistent manner and have a positive impact in the EnKF update.

1 Introduction

In this paper we focus on a data assimilation method based on the ensemble Kalman filter (EnKF) for estimation of oil reservoir parameters and state variables given the 3D

seismic data, the time difference seismic data and the production history data. The EnKF method was introduced by Evensen (1994) for updating non-linear ocean models, and it has attained popularity in reservoir history matching because of it's relatively simple conceptual formulation. Several publications have discussed the use of EnKF with oil reservoir models such as Nævdal et al. (2002a,b, 2004), Gu and Oliver (2004), Gao and Reynolds (2005), Liu and Oliver (2005), Wen and Chen (2005) and Skjervheim et al. (2005), and most of them have shown promising results. This paper is based on the work performed by Skjervheim et al. (2005), and here we present in more details the theory when a large amount of time difference data are incorporated in the EnKF.

Data assimilation in reservoir simulation models can be applied to a combined parameter and state estimation problem, where both static variables (permeability, porosity, etc.) and dynamic variables (pressures and saturations) may be estimated. The estimation problem can be described in a Bayesian inversion framework, and a general introduction is presented by Evensen (2005). The Bayesian framework provides a methodology to incorporate a prior and a likelihood model, where the ultimate solution of the inverse problem is the posterior probability density function which is uniquely defined by the prior and the likelihood.

The EnKF method is a Monte Carlo type sequential Bayesian inversion, and provides an approximate solution to the combined parameter and state estimation problem. The result is an ensemble of solutions approximating the posterior probability density function for the model input parameters, state variables and other output data conditioned to measured, dynamic data.

To quantify the uncertainty in the reservoir models it is important to utilize all available measurements, and the motivation for the presented work is to perform a continuously reservoir model update by assimilating both production data, 3D seismic data and time difference seismic data in a statistical consistent manner using an ensemble Kalman filter method, where time difference seismic is defined as a difference between two 3D seismic surveys. Such time difference data may for instant be the result of a 4D inversion where the seismic difference between two surveys are inverted to differences in elastic parameters. Elastic parameters for the individual surveys may then not be available.

When using time difference data the observations at a particular time depend on the model state at more than one time instant. In the Bayesian formulation of the data assimilation the posterior distribution of the combined parameter and state estimation problem then include both a filter and a smoother solution, and a combination of the ensemble Kalman filter and ensemble Kalman smoother has to be applied for updating a reservoir simulation model.

3D seismic and time difference seismic may introduce a large amount of data in the assimilation, and the paper is concerned with the use of efficient EnKF algorithms which can handle such a large amount of data, where special challenges, such as potential loss of rank

and the inversion of a large covariance matrix are involved. Evensen (2004) proved that the rank problem can be avoided by proper sampling of measurement perturbations, and an efficient square root algorithm was presented. In this paper we present a subspace EnKF algorithm based on the sampling strategy given by Evensen with the same efficiency and approximations as the efficient square root algorithm.

The combined state and parameter estimation problem is described in a Bayesian framework in the following section. In Section 3 we derive a Bayesian formulation which incorporate time difference data, and we show that the posterior distribution include both a filter and a smoother solution. In Section 4 we use time difference data and present a sampling strategy for the posterior distribution assuming Gaussian statistics and linear dynamics, while in Section 5 we utilize time difference data and a nonlinear evolution model to present the ensemble Kalman filter sampling scheme.

In Section 6 we present the derivation of the subspace EnKF scheme. Finally, some results when using a large amount of time difference data is presented in Section 7. Here we have used a synthetic 3D reservoir model involving production data, seismic 3D data and seismic difference data. Comparison between the traditional and the subspace EnKF algorithm using different combinations of analysis schemes and measurement error covariance matrix representations is then discussed.

2 Bayesian framework

To improve forecasts the data assimilation methods incorporate observations into a dynamical model, where the state space model is in principle a hidden Markov chain with a finite set of states. In the model each of the hidden states is associated with a probability distribution, and transitions among the states are governed by a set of probabilities called transition probabilities. The probabilistic framework is necessary to quantify the uncertainty, and in this article we have used a sequential data assimilation method based on a Bayesian formalism, introduced by van Leeuwen and Evensen (1996) and Evensen and van Leeuwen (2000).

The sequential data assimilation method estimate the state and the parameters of the system sequentially by propagating information only forward in time using the Bayesian probability theory. In a Bayesian formalism the stochastic model consists of a prior model for the model state variables u and for the model parameters α, and a likelihood function model for the available observations d. The model state variables termed $u{=}u(\mathbf{x},t)$ represent an unknown model state in space and time, and $\alpha(\mathbf{x})$ represent a set of poorly known parameters of the model, which are assumed to be constant in time.

Let now $\psi(\mathbf{x}, t)$ denote the combination of the state vector and the parameters, i.e.,

$$\psi(\mathbf{x}, t) = (u(\mathbf{x}, t), \alpha(\mathbf{x}))^T. \tag{1}$$

The joint prior probability density function (pdf) for the model state and the parameters is given by $g(\psi)$. The observations, d, are also treated as stochastic variables, and the likelihood model for the observations is represented by the distribution $g(d|\psi)$.

The posterior pdf of interest is $g(\psi|d)$, i.e., the stochastic model for ψ given the available observations d, and from Bayesian theory, e.g Tarantola (2005), it follows that the posterior pdf of this combined state and parameter estimation problem can be written as

$$g(\psi|d) \propto g(d|\psi)g(\psi). \tag{2}$$

Let now $\psi_l = \psi(\mathbf{x}, t_l) \in \Re^q$ and $d_l \in \Re^{m_l}$ represent respectively the model variables and the observations at time t_l, and assume that we want to estimate ψ at a discrete set of times, $t_0, t_1, \ldots, t_l$, given observations at the same times (except at t_0) $d_1, d_2, \ldots, d_l$.

In the following we denote the sequence $\{u_l, \ldots, u_1, u_0, \alpha\}$ by $\psi_{l:0}$. It was shown in Evensen and van Leeuwen (2000) that a general smoother and filter could be derived from the Bayesian formalism, and for the time interval $t \in [t_0, t_l]$ the general smoother in Eq. (2) can now be expressed by

$$g(\psi_{l:0}|d_{l:1}) \propto g(d_{l:1}|\psi_{l:0})g(\psi_{l:0}), \tag{3}$$

where the prior probability function, $g(\psi_{l:0})$, can be represented by

$$g(\psi_{l:0}) = g(\psi_l|\psi_{l-1:0})g(\psi_{l-1:0}). \tag{4}$$

The likelihood distribution of the data can be written as

$$\begin{aligned} g(d_{l:1}|\psi_{l:0}) &= g(d_l|\psi_{l:0}) \ldots g(d_1|\psi_{l:0}) \\ &= g(d_l|\psi_{l:0})g(d_{l-1}|\psi_{l-1:0}) \ldots g(d_1|\psi_{1:0}), \end{aligned} \tag{5}$$

where we assume independence between the data available at different times, and that the data at a given time are independent on states at future times. The smoother defined by Eq. (3) is then represented by

$$\begin{aligned} g(\psi_{l:0}|d_{l:1}) &\propto g(d_l|\psi_{l:0})g(d_{l-1:1}|\psi_{l-1:0})g(\psi_l|\psi_{l-1:0})g(\psi_{l-1:0}) \\ &\propto g(d_l|\psi_{l:0})g(\psi_l|\psi_{l-1:0})g(\psi_{l-1:0}|d_{l-1:1}). \end{aligned} \tag{6}$$

Generally it is common to assume that the model evolution ψ_{l+1} is a first-order Markov process, where ψ_{l+1} depends only on ψ_l but not on previous ones $\psi_{l-1:0}$. Similarly, the methodology normally assume that the observations d_l depend only on the model variables at that time, ψ_l, according to Figure 1. From Figure 1 and Eq. (6) the smoother can now

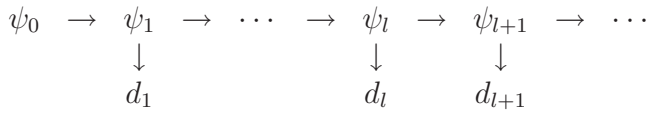


Figure 1: Data assimilation scheme where d_l depend only on ψ_l.

be written as

$$g(\psi_{l:0}|d_{l:1}) \propto g(d_l|\psi_l)g(\psi_l|\psi_{l-1})g(\psi_{l-1:0}|d_{l-1:1}). \tag{7}$$

Eqs. (6) and (7) are both recursive formulas, and thus the solution may be calculated sequentially and forward in time. The forecast density in the middle of the expressions corresponds to integration of the solution forward in time from t_{l-1} to t_l, and the multiplication with the last two probability densities introduce the prior distribution at time t_l. From now on, we use the following notation

$$\int g(*)\mathrm{d}\psi_{l:0} = \int \cdots \int g(*)\mathrm{d}\psi_l \mathrm{d}\psi_{l-1} \cdots \mathrm{d}\psi_0. \tag{8}$$

The filter solution for Eqs. (6) and (7) can be found by integrating over the solutions $\psi_{l-1:0}$, and for Eq. (7) we obtain

$$\begin{aligned} g(\psi_l|d_{l:1}) &\propto \int g(d_l|\psi_l)g(\psi_l|\psi_{l-1})g(\psi_{l-1:0}|d_{l-1:1})\,\mathrm{d}\psi_{l-1:0} \\ &\propto g(d_l|\psi_l)\int g(\psi_l|\psi_{l-1})\left[\int g(\psi_{l-1:0}|d_{l-1:1})\,\mathrm{d}\psi_{l-2:0}\right]\mathrm{d}\psi_{l-1} \\ &\propto g(d_l|\psi_l)\int g(\psi_l|\psi_{l-1})g(\psi_{l-1}|d_{l-1:1})\,\mathrm{d}\psi_{l-1} \\ &\propto g(d_l|\psi_l)g(\psi_l|d_{l-1:1}), \end{aligned} \tag{9}$$

where the forecast step, which compute the current prior for ψ_l, is defined by

$$g(\psi_l|d_{l-1:1}) = \int g(\psi_l|\psi_{l-1})g(\psi_{l-1}|d_{l-1:1})\,\mathrm{d}\psi_{l-1}. \tag{10}$$

This is the framework that has been outlined by Evensen and van Leeuwen (2000).

3 Bayesian framework using time difference data

In the smoother (7) and filter (9) equations from Figure 1, we assumed that the data at a specific time depend only on the model variables at that time. In this section we will

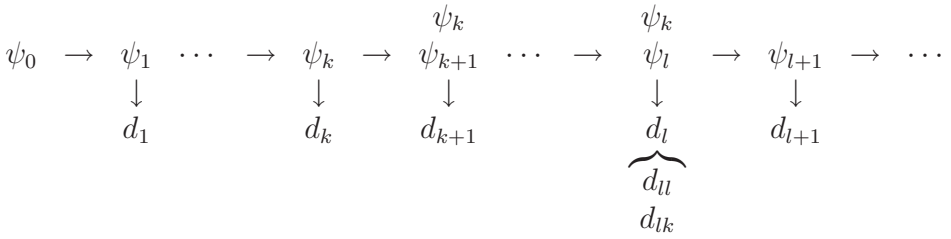


Figure 2: Data assimilation scheme where d_l depend both on ψ_l and ψ_k.

investigate the update procedure for the smoother and filter solutions when data $d_l = [d_{ll}, d_{lk}]^T$ at a particular time t_l depend both on ψ_l and ψ_k, where $t_k < t_l$. The likelihood function is then represented by $g(d_l|\psi_l, \psi_k)$. From now on d_{ll} denote an observation that depend only on ψ_l, and d_{lk} denote an observation that depend both on ψ_l and ψ_k. I our model we let d_{lk} be the time difference observation between ψ_l and ψ_k. We still assume a Markov process and Figure 2 introduces the model, where the updating procedure is as follows:

1. For the time interval $t_i \in (t_0, t_k]$ the data, $d_i = [d_{ii}]$, depend only on ψ_i.
2. For the time interval $t_i \in (t_k, t_l)$ the update is based on the data, $d_i = [d_{ii}]$. Here ψ_k will be included in the update procedure as a smoother update, because at time t_l the observation d_l will include time difference data between ψ_l and ψ_k.
3. When the time is equal t_l the update based on the data, $d_l = [d_{ll}, d_{lk}]^T$, depend on ψ_l and ψ_k.
4. When $t_i > t_l$ the update is based on the data, $d_i = [d_{ii}]$, and the observation depend only on ψ_i.

From Figure 2 and Eq. (6) the smoother at time t_l can be written as

$$g(\psi_{l:0}|d_{l:1}) \propto g(d_l|\psi_l, \psi_k)g(\psi_l|\psi_{l-1})g(\psi_{l-1:0}|d_{l-1:1}), \tag{11}$$

and the filter solution for Eq. (11) is expressed by

$$\begin{aligned} g(\psi_l|d_{l:1}) &\propto \int g(d_l|\psi_l,\psi_k)g(\psi_l|\psi_{l-1})g(\psi_{l-1:0}|d_{l-1:1})\,\mathrm{d}\psi_{l-1:0} \\ &\propto \int g(d_l|\psi_l,\psi_k)\left[\int g(\psi_l|\psi_{l-1})g(\psi_{l-1},\psi_k|d_{l-1:1})\mathrm{d}\psi_{l-1}\right]\mathrm{d}\psi_k \\ &\propto \int g(d_l|\psi_l,\psi_k)g(\psi_l,\psi_k|d_{l-1:1})\mathrm{d}\psi_k \\ &\propto \int g(\psi_l,\psi_k|d_{l:1})\,\mathrm{d}\psi_k, \end{aligned} \tag{12}$$

where the forecast step defining the prior distribution is given by

$$g(\psi_l,\psi_k|d_{l-1:1}) = \int g(\psi_l|\psi_{l-1})g(\psi_{l-1},\psi_k|d_{l-1:1})\mathrm{d}\psi_{l-1}, \tag{13}$$

and the posterior distribution at time t_l is defined by

$$g(\psi_l,\psi_k|d_{l:1}) \propto g(d_l|\psi_l,\psi_k)g(\psi_l,\psi_k|d_{l-1:1}). \tag{14}$$

Notice that the forecast step in Eq. (13) requires the filter solution of ψ_{l-1} and the smoother solution of ψ_k obtained from $g(\psi_{l-1},\psi_k|d_{l-1:1})$. The Bayesian formulation of the filter solutions in Eqs. (9) and (12) now gives the possibility to compute an update of the posterior distribution in a statistical consistent manner, when we include time difference data.

4 The Kalman Filter and Linear Dynamics

The probabilistic framework presented above can be difficult to solve when the dimension of the state space is high and with general probability density functions. To investigate closer the influence of using a model which include time difference data, we consider the case of linear dynamics where we assume Gaussian statistics, see Bertino et al. (2003). From Eq. (12) we have the general filter expression and the dependency of the observations, $d_l = [d_{ll}, d_{lk}]^T \in \Re^{m_l}$, is given from the likelihood function $g(d_l|\psi_l,\psi_k)$. Defining the joint model variables as $\tilde{\psi}_l = [\psi_l,\psi_k]^T \in \Re^{\tilde{q}}$, the state space model can be written as

$$\tilde{\psi}_l = \tilde{F}_l\tilde{\psi}_{l-1} + \epsilon_l^m \qquad \epsilon_l^m \sim \mathcal{N}_{\tilde{q}}(0, P_{\epsilon_l^m}) \tag{15}$$

$$d_l = \tilde{H}_l\tilde{\psi}_l + \epsilon_l^o \qquad \epsilon_l^o \sim \mathcal{N}_{m_l}(0, P_{\epsilon_l^o}), \tag{16}$$

where $\tilde{F}_l \in \Re^{\tilde{q}\times\tilde{q}}$ defined by

$$\tilde{F}_l = \begin{bmatrix} F_l & 0 \\ 0 & I \end{bmatrix} \tag{17}$$

is the dynamic model operator, and $\tilde{H}_l \in \Re^{m_l \times \tilde{q}}$ is the observation operator on the form

$$\tilde{H}_l = \begin{bmatrix} H_{ll} & 0 \\ H_{lk} & -H_{lk} \end{bmatrix}. \tag{18}$$

This form is obtained as a result of including time difference data, where the matrices $H_{ll} \in \Re^{m_{ll} \times q_l}$ and $H_{lk} \in \Re^{m_{lk} \times q_l}$ are respectively the observation matrices for d_{ll} and for the time difference data d_{lk}. The terms ϵ_l^m and ϵ_l^o are respectively model and observation random errors, and their covariance matrices are $P_{\epsilon_l^m}$ and $P_{\epsilon_l^o}$. Notice that the model errors are only added to the terms where F_l is operating on ψ_{l-1} and is zero elsewhere.

Eq. (12) shows that the filter solution at time t_l can be found from the posterior distribution in Eq. (14), and as a result we now want to find a Gaussian expression for the posterior distribution $g(\tilde{\psi}_l|d_{l:1})$. Define the conditional expectations and the conditional error covariance matrices for the forecast and the analysis respectively as

$$\tilde{\psi}_l^f = \mathrm{E}(\tilde{\psi}_l|d_{l-1:1}) \tag{19}$$

$$P_{\tilde{\psi}_l}^f = \mathrm{E}((\tilde{\psi}_l - \mathrm{E}(\tilde{\psi}_l|d_{l-1:1}))(\tilde{\psi}_l - \mathrm{E}(\tilde{\psi}_l|d_{l-1:1}))^T), \tag{20}$$

and

$$\tilde{\psi}_l^a = \mathrm{E}(\tilde{\psi}_l|d_{l:1}) \tag{21}$$

$$P_{\tilde{\psi}_l}^a = \mathrm{E}((\tilde{\psi}_l - \mathrm{E}(\tilde{\psi}_l|d_{l:1}))(\tilde{\psi}_l - \mathrm{E}[\tilde{\psi}_l|d_{l:1}))^T). \tag{22}$$

Assume that we know the posterior distribution $g(\tilde{\psi}_{l-1}|d_{l-1:1})$ at time t_{l-1}

$$\tilde{\psi}_{l-1}|d_{l-1:1} \sim \mathcal{N}_{\tilde{q}}(\tilde{\psi}_{l-1}^a, P_{\tilde{\psi}_{l-1}}^a). \tag{23}$$

The prior distribution $g(\tilde{\psi}_l|d_{l-1:1})$ in Eq. (13) can then be written as

$$\tilde{\psi}_l|d_{l-1:1} \sim \mathcal{N}_{\tilde{q}}(\tilde{\psi}_l^f, P_{\tilde{\psi}_l}^f), \tag{24}$$

where the two forecast moments are found by using basic results from conditional expectation

$$\tilde{\psi}_l^f = \mathrm{E}(\mathrm{E}(\psi_l|\psi_{l-1})|d_{l-1:1}) = \mathrm{E}(\tilde{F}_l\tilde{\psi}_{l-1}|d_{l-1:1}) = \tilde{F}_l\tilde{\psi}_{l-1}^a \tag{25}$$

$$\begin{aligned} P_{\tilde{\psi}_l}^f &= \mathrm{E}(\mathrm{Cov}(\psi_l|\psi_{l-1})|d_{l-1:1}) + \mathrm{Cov}(\mathrm{E}(\psi_l|\psi_{l-1})|d_{l-1:1}) \\ &= \mathrm{E}(P_{\epsilon_l^m}|d_{l-1:1}) + \mathrm{Cov}(\tilde{F}_l\tilde{\psi}_{l-1}|d_{l-1:1}) \\ &= P_{\epsilon_l^m} + \tilde{F}_l P_{\tilde{\psi}_{l-1}}^a \tilde{F}_l^T. \end{aligned} \tag{26}$$

The likelihood distribution $g(d_l|\tilde{\psi}_l)$ is found using the observation equation (16)

$$d_l|\tilde{\psi}_l \sim \mathcal{N}_{m_l}(\tilde{H}_l\tilde{\psi}_l, P_{\epsilon_l^o}), \tag{27}$$

and from Gaussian theory the posterior distribution $g(\tilde{\psi}_l|d_{l:1})$ can be written as

$$\tilde{\psi}_l|d_{l:1} \sim N_{\tilde{q}}(\tilde{\psi}_l^a, P_{\tilde{\psi}_l}^a), \tag{28}$$

where the analyzed moments are

$$\tilde{\psi}_l^a = \tilde{\psi}_l^f + P_{\tilde{\psi}_l}^f \tilde{H}_l^T(\tilde{H}_l P_{\tilde{\psi}_l}^f \tilde{H}_l^T + P_{\epsilon_l^o})^{-1}(d_l - \tilde{H}_l\tilde{\psi}_l^f) \tag{29}$$

$$P_{\tilde{\psi}_l}^a = P_{\tilde{\psi}_l}^f - P_{\tilde{\psi}_l}^f \tilde{H}_l^T(\tilde{H}_l P_{\tilde{\psi}_l}^f \tilde{H}_l^T + P_{\epsilon_l^o})^{-1}\tilde{H}_l P_{\tilde{\psi}_l}^f. \tag{30}$$

To generate a sample from the filter distribution $g(\psi_l|d_{l:1})$ in Eq. (12) we can either sample directly from the Gaussian distribution given in Eq. (28)

$$\tilde{\psi}_l^{a*} = [\psi_l^{a*}, \psi_k^{a*}]^T \sim N_{\tilde{q}}(\tilde{\psi}_l^a, P_{\tilde{\psi}_l}^a) \tag{31}$$

where ψ_l^{a*} represents a sample from the filter solution, or sample by using conditional simulation with kriging (see e.g., Cressie (1991) , Sect. 3.6.2). That is,

1. Sample from the posterior distribution $g(\tilde{\psi}_{l-1}|d_{l-1:1})$:

$$\tilde{\psi}_{l-1}^{a*} \sim \mathcal{N}_{\tilde{q}}(\tilde{\psi}_{l-1}^a, P_{\tilde{\psi}_{l-1}}^a)$$

2. Forecast the sample $\tilde{\psi}_{l-1}^{a*}$ using Eq.(15):

$$\tilde{\psi}_l^{f*} = \tilde{F}_l\psi_{l-1}^{a*} + \epsilon_l^{m*}, \quad \epsilon_l^{m*} \sim \mathcal{N}_{\tilde{q}}(0, P_{\epsilon_l^m})$$

   - $\tilde{\psi}_l^{f*}$ represent a sample from the prior distribution $g(\tilde{\psi}_l|d_{l-1:1})$.

3. Sample from the measurement distribution:

$$d_l^* \sim \mathcal{N}_{m_l}(d_l, P_{\epsilon_l^o})$$

4. Use the kriging formula to generate a sample from the posterior distribution $g(\tilde{\psi}_l|d_{l:1})$:

$$\tilde{\psi}_l^{a*} = \tilde{\psi}_l^{f*} + P_{\tilde{\psi}_l}^f \tilde{H}_l^T(\tilde{H}_l P_{\tilde{\psi}_l}^f \tilde{H}_l^T + P_{\epsilon_l^o})^{-1}(d_l^* - \tilde{H}_l\tilde{\psi}_l^{f*})$$

   - $\tilde{\psi}_l^{a*} = [\psi_l^{a*}, \psi_k^{a*}]^T$ and ψ_l^{a*} represent a sample from the filter distribution $g(\psi_l|d_{l:1})$.

5 The Ensemble Kalman Filter

To continue our study of using an evolution model which include time difference data, we consider now a nonlinear state space model defined by

$$\tilde{\psi}_l = \begin{bmatrix} f_l(\psi_{l-1}) \\ \psi_k \end{bmatrix} + \epsilon_l^m \qquad \epsilon_l^m \sim \mathcal{G}_{\tilde{q}}(0, P_{\epsilon_l^m}) \tag{32}$$

$$d_l = \tilde{H}_l \tilde{\psi}_l + \epsilon_l^o \qquad \epsilon_l^o \sim \mathcal{G}_{m_l}(0, P_{\epsilon_l^o}) \tag{33}$$

where $\tilde{\psi}_l = [\psi_l, \psi_k]^T \in \Re^{\tilde{q}}$, $f_l : \Re^q \to \Re^q$ is a non-linear model operator operating only on ψ_{l-1} with model errors ϵ_l^m, $\tilde{H}_l$ is the linear measurement operator, defined by Eq. (18), relating the model state $\tilde{\psi}_l$ to the observations d_l allowing for measurement errors ϵ_l^o, and $\mathcal{G}(\mu, P)$ denote a partially specified distribution with mean μ and covariance P. Notice that the error statistics are now free to take any form provided they are consistent with their defined first and second-order moments, and that the model errors are only added to the terms where f_l is operating on ψ_{l-1} and is zero elsewhere. Because of the nonlinearity in the state space model the posterior distribution $g(\tilde{\psi}_l|d_{l:1})$ can now longer be explicitly described by a Gaussian distribution, and as a result we need to use a sampling strategy to explore the distribution.

Assume now that the joint distribution of $g(\tilde{\psi}_l, d_l|d_{l-1:1})$ can be partially specified through the mean and the covariance. By applying linear Bayes estimation with respect to a quadratic loss function we achieve the approximations to the posterior moments of the posterior distribution $g(\tilde{\psi}_l|d_{l:1})$ (see e.g., West and Harrison (1997), Sect. 4.9). The linear posterior mean $\tilde{\psi}_l^a$ and linear posterior covariance $P_{\tilde{\psi}_l}^a$ are respectively given by Eq. (29) and Eq. (30), and $g(\tilde{\psi}_l|d_{l:1})$ is represented by

$$\tilde{\psi}_l|d_{l:1} \sim \mathcal{G}_{\tilde{q}}(\tilde{\psi}_l^a, P_{\tilde{\psi}_l}^a). \tag{34}$$

Notice that the expressions for the posterior moments are the same as for the Gaussian case, but an optimal variance minimizing estimate is only obtained when we assume Gaussian statistics and linear dynamics.

To sample from the posterior distribution $g(\tilde{\psi}_l|d_{l:1})$ we are using an Ensemble Kalman filter (EnKF) technique, and to describe the EnKF analysis scheme we now introduce an EnKF notation, which is similar to that used in Evensen (2003). The matrix holding the ensemble members, $\psi_l \in \Re^q$, is defined by

$$\Psi_l = \Psi(\mathbf{x}, t_l) = (\psi_l^1, \ldots, \psi_l^N) \in \Re^{q \times N}, \tag{35}$$

where N is the number of ensemble members, and the ensemble mean is given by

$$\bar{\psi}_l = \frac{1}{N} \sum_{i=1}^{N} \psi_l^i. \tag{36}$$

Consider now a joint ensemble matrix including the ensemble matrices for the different time instants, $\Psi_l = \Psi(\mathbf{x}, t_l), \Psi_k = \Psi(\mathbf{x}, t_k)$

$$\tilde{\Psi}_l = \left[\begin{array}{c} \Psi_l \\ \Psi_k \end{array} \right] \in \Re^{\tilde{q} \times N}. \tag{37}$$

The ensemble perturbation matrix at time t_l can be written as

$$\tilde{\Psi}'_l = \tilde{\Psi}_l - \bar{\tilde{\Psi}}_l = \tilde{\Psi}_l (I - 1_N), \tag{38}$$

where $\bar{\tilde{\Psi}}_l$ is the ensemble mean matrix and 1_N is the $N \times N$ matrix where each element is equal to 1/N. The ensemble covariances, $P^e_{\tilde{\psi}_l} \in \Re^{\tilde{q} \times \tilde{q}}$, can then be estimated as

$$P^e_{\tilde{\psi}_l} = \frac{\tilde{\Psi}'_l (\tilde{\Psi}'_l)^T}{N-1}. \tag{39}$$

Given a vector of measurements, $d_l \in \Re^{m_l}$, we can define the N vectors of perturbed measurements as

$$d^j_l = d_l + \epsilon^{o,j}_l \quad \forall \quad j = 1 : N, \tag{40}$$

which can be stored in the columns of a matrix

$$D_l = (d^1_l, \ldots, d^N_l) \in \Re^{m_l \times N}. \tag{41}$$

The ensemble of perturbations can be stored in the matrix

$$E_l = (\epsilon^{o,1}_l, \ldots, \epsilon^{o,N}_l) \in \Re^{m_l \times N}, \tag{42}$$

and the ensemble representation of the measurement error covariance matrix is given by

$$P^e_{\epsilon^o_l} = \frac{E_l E_l^T}{N-1}. \tag{43}$$

The EnKF sampling may be viewed as a conditional simulation with kriging, and the sampling strategy for a case with time difference data is then as follows:

0. Generate N independent samples, $\psi^{a,j}_k$ from $g(\psi_k | d_{k:1})$.

For $i = k+1, k+2, \ldots, l$

1. Forecast each of the samples, $\psi^{a,j}_{i-1}$, using Eq.(32):

$$\tilde{\psi}^{f,j}_i = \left[\begin{array}{c} f_i(\psi^{a,j}_{i-1}) \\ \psi^{a,j}_{k,i-1} \end{array} \right] + \epsilon^{m,j}_i, \ \epsilon^{m,j}_i \sim \mathcal{G}_{\tilde{q}}(0, P_{\epsilon^m_i}) \ \forall \ j = 1 : N$$

- $\psi_{k,i-1}^{a,j}$ is the smoother update of $\psi_k^{a,j}$ at the previous time step.
- $\tilde{\psi}_i^{f,j}$ represent a sample from the prior distribution $g(\tilde{\psi}_i|d_{i-1:1})$.

2. Use Monte Carlo integration to estimate the moments from the prior density given in Eqs. (19)-(20).

3. Generate N samples from the measurement distribution:

$$d_i^j \sim \mathcal{G}_{m_i}(d_i, P_{\epsilon_i^o}) \ \ \forall \ \ j = 1 : N$$

4. Use the kriging formula to generate a sample from the posterior distribution $g(\tilde{\psi}_i|d_{i:1})$:

$$\tilde{\psi}_i^{a,j} = \tilde{\psi}_i^{f,j} + P_{\tilde{\psi}_i}^{e,f} \tilde{H}_i^T (\tilde{H}_i P_{\tilde{\psi}_i}^{e,f} \tilde{H}_i^T + P_{\epsilon_i^o})^{-1} (d_i^j - \tilde{H}_i \tilde{\psi}_i^{f,j})$$

- where $\tilde{\psi}_i^{a,j} = [\psi_i^{a,j}, \psi_{k,i}^{a,j}]^T$ and $\psi_i^{a,j}$ represent a sample from the filter distribution $g(\psi_i|d_{i:1})$.

The ensemble representation of the analysis step is

$$\tilde{\Psi}_i^a = \tilde{\Psi}_i^f + \tilde{\Psi}_i^{f'} \tilde{\Psi}_i^{f'T} \tilde{H}_i^T (\tilde{H}_i \tilde{\Psi}_i^{f'} \tilde{\Psi}_i^{f'T} \tilde{H}_i^T + (N-1) P_{\epsilon_i^o})^{-1} (D_i - \tilde{H}_i \tilde{\Psi}_i^f). \tag{44}$$

If we now at time t_l define the measurement of the ensemble perturbations, $S_l \in \Re^{m_l \times N}$, as

$$S_l = \tilde{H}_l \tilde{\Psi}_l^{f'}, \tag{45}$$

the ensemble of innovation vectors as

$$D_l' = D_l - \tilde{H}_l \tilde{\Psi}_l^f, \tag{46}$$

and the matrix $C_l \in \Re^{m_l \times m_l}$ as

$$C_l = S_l S_l^T + (N-1) P_{\epsilon_l^o}, \tag{47}$$

we have from Evensen (2003) that the analysis equation (44) can be written as

$$\tilde{\Psi}_l^a = \tilde{\Psi}_l^f X_l, \tag{48}$$

where the matrix $X_l \in \Re^{N \times N}$ is defined as

$$X_l = I + S_l^T C_l^{-1} D_l'. \tag{49}$$

The filter solution is then expressed by

$$\Psi_l^a = \Psi_l^f X_l. \tag{50}$$

6 EnKF and large data sets

In reservoir simulation models spatial data such as 3D seismic and time difference seismic may introduce a large amount of data in the assimilation. In this section we are concerned with the use of efficient EnKF algorithms which handle such large data sets, where special challenges such as inversion of the $m \times m$ matrix C in Eq.(47), and potential loss of rank are involved. To simplify the notation we have in this section dropped the forecast notation f, the time subscripts l and the joint notation $\tilde{\psi}$.

A large amount of data may introduce numerical singularity in C even when using a full rank error covariance matrix P_{ϵ^o}, and the pseudo inverse C^+ of C has to be used in the analysis schemes. The pseudo inverse of C is given as

$$C^+ = Z_m \Lambda_m^+ Z_m^T, \tag{51}$$

where the diagonal matrix Λ_m^+ is defined as

$$\text{diag}(\Lambda_m^+) = (\lambda_1^{-1}, \ldots, \lambda_p^{-1}, 0, \ldots, 0), \tag{52}$$

when $p = \text{rank}(C)$ and the eigenvalues $\lambda_j \geq \lambda_{j+1} > 0$. The computation may be very time demanding when the size of the observation vector gets large, since the factorization of C requires a cost proportional to $\mathcal{O}(m^3)$. In the following we will discuss how the time demanding pseudo inverse can be avoided in the analysis scheme and we will focus on the rank issues presented by Evensen (2004) and Kepert (2004).

Evensen (2004) and Kepert (2004) have pointed out that the use of an ensemble representation, $P_{\epsilon^o}^e$ for P_{ϵ^o} in some cases may lead to loss of rank in the ensemble when $m > N-1$. Evensen (2004) showed that the rank problem can be avoided if the measurement perturbations is defined such that the column vectors of E lies in the $N-1$ dimensional space $\mathcal{L}_S$ spanned by the column vectors of S in Eq. (45). Based on this sampling strategy an efficient square root algorithm was presented, which reduced the factorization of the $m \times m$ matrix C in Eq.(47) to a factorization of a $N \times N$ matrix. The factorization in the efficient square root algorithm can be computed at a cost of $\mathcal{O}(Nm^2)$ floating point operations, and when a low rank representation, $P_{\epsilon^o}^e$, for the error covariance matrix is used the cost is of $\mathcal{O}(N^2 m)$.

In this section a subspace EnKF algorithm, with the same computational cost as given above, is proposed based on a principal component projection of the measurements d, where the projection matrix $U \in \Re^{m \times N-1}$ span the ensemble space $\mathcal{L}_S$. The motivation for the proposed subspace EnKF algorithm is based on the algorithm presented by Evensen (2004). The projection introduces an approximate pseudo inverse of C in a similar way as Evensen (2004) and reduces the factorization of the $m \times m$ matrix C to a factorization of a $N-1 \times N-1$ matrix.

We now define the projection of the measurements, $\hat{d} \in \Re^{N-1}$ as

$$\hat{d} = U^T d, \tag{53}$$

and the projected measurement operator $\hat{H} \in \Re^{N-1 \times q}$ as

$$\hat{H} = U^T \tilde{H}. \tag{54}$$

Assuming a nonlinear evolution model like in Eqs. (32) and (33), where we apply the projection matrix U on the observations d, the new observation equation will then be given by

$$\hat{d} = \hat{H}\psi + \hat{\epsilon}^o \qquad\qquad \hat{\epsilon}^o \sim \mathcal{G}_{N-1}(0, P_{\hat{\epsilon}^o}). \tag{55}$$

Here the projected measurement errors are defined by $\hat{\epsilon}^o = U^T \epsilon^o \in \Re^{N-1}$, and the error covariance matrix is given by $P_{\hat{\epsilon}^o} = U^T P_{\epsilon^o} U$. The likelihood distribution is obtained from Eq. (55):

$$\hat{d}|\psi \sim \mathcal{G}_{N-1}(\hat{H}\psi, P_{\hat{\epsilon}^o}). \tag{56}$$

By using the ensemble approximation, $P^e_{\epsilon^o}$, of P_{ϵ^o} in Eq. (43) the analysis of the ensemble can be written as

$$\Psi^a = \Psi + \Psi' \Psi'^T \hat{H}^T (\hat{H} \Psi' \Psi'^T \hat{H}^T + (N-1) P^e_{\hat{\epsilon}^o})^{-1} \hat{D}', \tag{57}$$

where $\hat{D}'$ is given by

$$\hat{D}' = \hat{D} - \hat{H}\Psi = U^T (D - H\Psi). \tag{58}$$

The matrix $P^e_{\hat{\epsilon}^o} = U^T P^e_{\epsilon^o} U$ is the projection of the measurement error covariance matrix, $P^e_{\epsilon^o}$, onto the subspace $\mathcal{L}_S$ and all possible contributions in $\mathcal{L}_S^{\perp}$ are rejected. This ensures that the measurement perturbations explain the variance within the ensemble space $\mathcal{L}_S$ and we avoid the loss of rank pointed out by Evensen (2004) and Kepert (2004).

From Eq. (45) we have defined $S \in \Re^{m \times N}$ with $\text{rank}(S) = N - 1$, where the columns of S span the subspace $\mathcal{L}_S$ of dimension $N - 1$. The spectral decomposition of the covariance matrix associated with the ensemble measurement perturbations can be written as

$$SS^T = U \Lambda U^T. \tag{59}$$

The diagonal matrix $\Lambda \in \Re^{N-1 \times N-1}$ is given by

$$\text{diag}(\Lambda) = (\lambda_1, \lambda_2, \ldots, \lambda_{N-1}), \tag{60}$$

with the eigenvalues $\lambda_j \geq \lambda_{j+1} \geq 0$. If we now define the projected measurements of the ensemble perturbations, $\hat{S} \in \Re^{N-1 \times N}$ as

$$\hat{S} = \hat{H}\Psi' = U^T H \Psi' = U^T S, \tag{61}$$

we can define the matrix $\hat{C} \in \Re^{N-1 \times N-1}$ as

$$\begin{aligned}\hat{C} &= \hat{S}\hat{S}^T + (N-1)P^e_{\hat{\epsilon}^o} \\ &= U^T S S^T U + (N-1)U^T P^e_{\epsilon^o} U \\ &= U^T U \Lambda U^T U + U^T E E^T U \\ &= \Lambda + \hat{E}\hat{E}^T. \end{aligned} \tag{62}$$

Here the matrix product $U^T U = I_{N-1} \in \Re^{N-1 \times N-1}$ where I_{N-1} is the identity matrix and the projection of the ensemble of perturbations is given as $\hat{E} = U^T E$.

The analysis equation can now be written as

$$\Psi^a = \Psi^f \hat{X}, \tag{63}$$

where the matrix $\hat{X} \in \Re^{N \times N}$ is defined by

$$\hat{X} = I + \hat{S}^T \hat{C}^{-1} \hat{D}', \tag{64}$$

and from Eq.(64) the pseudo inverse of C becomes

$$C^+ = U\hat{C}^{-1}U^T. \tag{65}$$

The interpretation of this subspace pseudo inversion can be viewed as an assimilation of a set of observations after they have been projected onto the subspace $\mathcal{L}_S$ as defined by the first $N-1$ singular vectors of S. The subspace EnKF scheme do not introduce any approximation if P_{ϵ^o} is diagonal, since the matrix C and SS^T will have the same eigenvectors. However, if P_{ϵ^o} is non-diagonal they will have different eigenvectors and the projection onto $\mathcal{L}_S$ space eliminates the part of C which is orthogonal to the $\mathcal{L}_S$ space. The use of a low rank representation for P_{ϵ^o} allows us to assimilate very large data sets to a low cost, and the results will be the same as obtained using a full rank P_{ϵ^o} if $U^T P^e_{\epsilon^o} U = U^T P_{\epsilon^o} U$.

In Appendix A we have presented the derivation of the square root algorithm, where a low rank or a full rank P_{ϵ^o} can be specified.

7 Example

7.1 Combined assimilation of production and seismic data

A synthetic 3D reservoir model with $30 \times 30 \times 5$ grid cells has been used to show the impact of adding seismic data in the assimilations. The reservoir is produced by water injection,

	True					Inital
	L 1	L 2	L 3	L 4	L 5	Layers 1-5
Mean permeability (mD)	2000	1900	1800	1700	1600	1800
Std. of permeability (mD)	100					350
Mean porosity	0.30	0.29	0.28	0.27	0.26	0.28
Std. of porosity	0.01					0.04
Variogram range along x dir. (m)	700 (45^o counter clockwise)					400
Variogram range along y dir. (m)	300 (45^o counter clockwise)					400
Variogram range along z dir. (m)	40					40
Correlation between perm. and poro.	0.5					0.5
Reservoir dimension (m)	$1800 \times 1800 \times 100$					

Table 1: Input data to synthetic reservoir model.

where the water injector and the oil producer are respectively located in the upper right and in the lower left corner. The water is injected in the bottom layer, while oil is produced from the top layer. Injection of water starts after one year of production, allowing for production below bubble-point, and before a higher reservoir pressure is re-established, free gas will evolve around the producer.

The reference ("true") porosity and permeability fields were simulated as Gaussian fields with a trend from the top to the bottom of the reservoir and with a fairly strong variogram range parallel to the flow direction, Figure 3.

The initial ensemble consists of 100 realizations sampled from Gaussian random fields with constant mean permeability and porosity in all layers and a gaussian variogram model. Model data are listed in Table 1, and one realization from the initial ensemble is plotted in Figure 4.

Total simulation time is 9 years, where 6 years are history and 3 years prediction. During the run, 1 seismic survey, 1 seismic difference survey and production data every 3 months are assimilated. The seismic survey is assimilated after one month production, and the seismic difference survey, which is the difference between the surveys after 6 years and 1 month production, is assimilated after 6 years production.

Production data are measurements of water cut (WCT), gas-oil ratio (GOR) and flowing bottomhole pressure (BHP). Seismic data are acoustic impedance calculated from grid cell properties (saturation, pressure and porosity) using a petroelastic model from a real North Sea reservoir, see Haverl et al. (2005). The measurement uncertainties of the well data are presented in Table 2.

While running the EnKF we update the static variables, which are porosity and perme-

Water cut (fraction)	Gas-oil ratio (Sm^3/Sm^3)	Bottomhole pressure (bar)
0.03	10 %	2.0

Table 2: Measurement errors for well data, standard deviations.

ability, and the state variables, which are the saturation of gas and water, the solution gas-oil ratio, and the pressure.

Eight cases of model updating with the EnKF were run. One case assimilates only production data, and seven cases assimilate production data, seismic 3D data and seismic difference data with different combinations of analysis schemes and error covariance matrix representations. The last seven cases have been run to show the impact of adding seismic data in the assimilation and to evaluate the properties of the efficient analysis schemes when $m \gg N$. The cases are listed in Table 3.

The analysis schemes which have been used are the traditional EnKF scheme (Case 1, 2a, 2b), and the subspace EnKF scheme presented in Section 6 where either a full rank prescribed error covariance matrix is used (Case 3a, 4a, 4b), or E is used directly without P_{ϵ^o} being formed (Case 5a, 5b). In case 3a we have used a lower seismic measurement error than in the rest of the cases. The seismic measurement errors used are described in Table 4.

In the naming of the experiments, a denotes that we have used a diagonal error covariance matrix with the observation error variance on the diagonal, and b denotes that we have used an error covariance matrix computed from an exponential variogram model with ranges respectively equal to 7, 7 and 2 blocks in x, y and z-direction. The error term, ϵ^o, added to the "true" observations are Gaussian noise simulated respectively from a diagonal P_{ϵ^o} ("a-cases"), and a P_{ϵ^o} computed from the exponential variogram model ("b-cases").

To evaluate the quality of the solution of the static variables we use the root mean square (RMS) error defined as

$$\Gamma(\alpha) = \sqrt{\frac{1}{N_g}\sum_{j=1}^{N_g}\left(\frac{1}{N}\sum_{i=1}^{N}(\alpha_j^i - \alpha_j^t)^2\right)} \tag{66}$$

where N_g is the number of grid cells, N is the number of ensemble members, α_j^i is the estimate of static variable j in ensemble member i, and α_j^t is the true value of static variable j.

Figure 5 show the mean porosity field for case 1 obtained by averaging the final estimated porosities of all the ensemble members, when only production data are assimilated. As can be seen, there is little information about the trend structure of the porosity in the production data, and almost no features from the true porosity field are recovered. A

Case no.	Scheme	Rank	P_{ϵ^o}	Seismic	Porosity		Permeability	
					RMS	Corr.	RMS	Corr.
Initial					0.044		385.83	
Case 1	trad. EnKF	full	diag.	no	0.041	-0.07	366.22	-0.12
Case 2a	trad. EnKF	full	diag.	yes	0.019	0.52	212.19	0.28
Case 2b	trad. EnKF	full	exp.	yes	0.020	0.37	180.86	0.31
Case 3a	sub. EnKF	full	diag.	yes	0.012	0.88	182.82	0.77
Case 4a	sub. EnKF	full	diag.	yes	0.021	0.48	221.03	0.28
Case 4b	sub. EnKF	full	exp.	yes	0.024	0.39	225.41	0.32
Case 5a	sub. EnKF	low	diag.	yes	0.023	0.45	257.59	0.25
Case 5b	sub. EnKF	low	exp.	yes	0.027	0.30	259.90	0.32

Table 3: Summary of experiments: The first column is the name of the experiment, in the second column is the analysis scheme used, the third column specify if a full or low rank P_{ϵ^o} is used, the fourth column specify if P_{ϵ^o} is diagonal or has a exponential correlation structure and the fifth column indicate if seismic data is included. The four last columns contain the RMS deviation and correlation between estimated and true static variables.

similar result is obtained for the permeability field, but still a good match to the data is achieved. Figure 8 shows the ensemble predictions of water cut and oil production rate, compared with the true solution and a simulation run where the final static ensemble average fields is used to simulate both the history and prediction period. As can be seen the uncertainty increases with time, but compared with the predictions based on the initial ensemble, which are plotted in Figure 7, a significant reduction in the estimated uncertainty is obtained.

The final mean ensemble porosity and the predictions for case 3a, using both production and seismic data in the assimilation, are plotted in Figures 6 and 9. In case 3a both a good history match and a good characterization of the main features of the true porosity and permeability fields are obtained, and clearly the addition of seismic data have a positive impact in the EnKF update.

The RMS values defined by Eq.(66) and the correlation values between the estimated mean ensemble and the true fields at the last assimilation step for both the porosity and the permeability are given for case 1 and 3a in Table 3. As can be seen, fairly high correlation values and low RMS values are obtained for case 3a compared to case 1, and the results shows an improvement of the estimate of the static variables when adding seismic data to the production data measurements. Especially an improvement of the porosity is obtained, and one reason might be that the acoustic impedance measurements do have a stronger relation to the porosity than to the permeability in the reservoir.

Case no.	Acoustic imp.(kg/m^{2}s)	Acoustic imp. difference kg/m^{2}s
Case 1	-	-
Case 3a	$1.5{\cdot}10^5$	$2.0{\cdot}10^3$
Cases 2a,4a,5a	$1.0{\cdot}10^6$	$1.0{\cdot}10^4$
Cases 2b,4b,5b	$1.0{\cdot}10^6$	$3.0{\cdot}10^3$

Table 4: Measurement errors for seismic data, standard deviations.

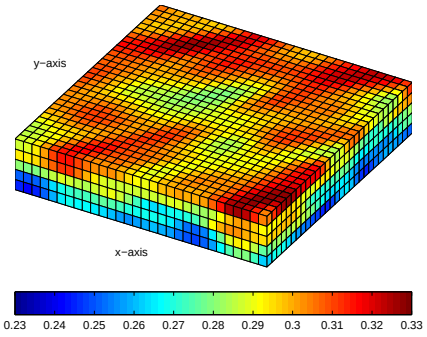


Figure 3: The reference ("true") porosity field.

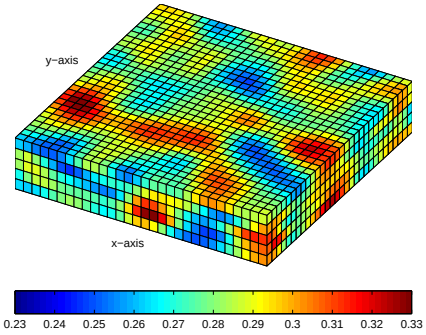


Figure 4: Realization from initial ensemble.

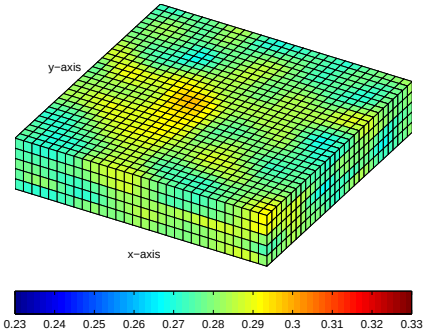


Figure 5: Final mean ensemble porosity case 1 (well data only).

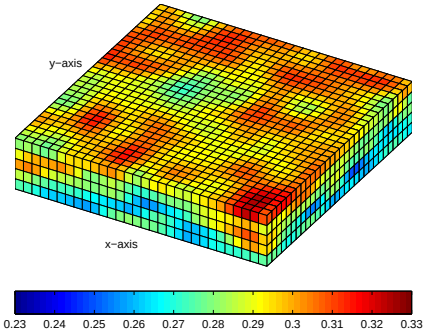


Figure 6: Final mean ensemble porosity case 3a (well data + seismic data).

7.2 Simulation runs when $m \gg N$

To perform a comparison between the subspace EnKF scheme proposed in Section 6 and the traditional EnKF when $m \gg N$, we have been running following six cases 2a, 2b, 4a, 4b, 5a and 5b, where all of them are described in Table 3. For cases 2a, 4a and 5a we have sampled normal independent perturbations for each element of E, while for cases 2b, 4b and 5b we have used a Cholesky decomposition method (see e.g Cressie (1991)) to generate

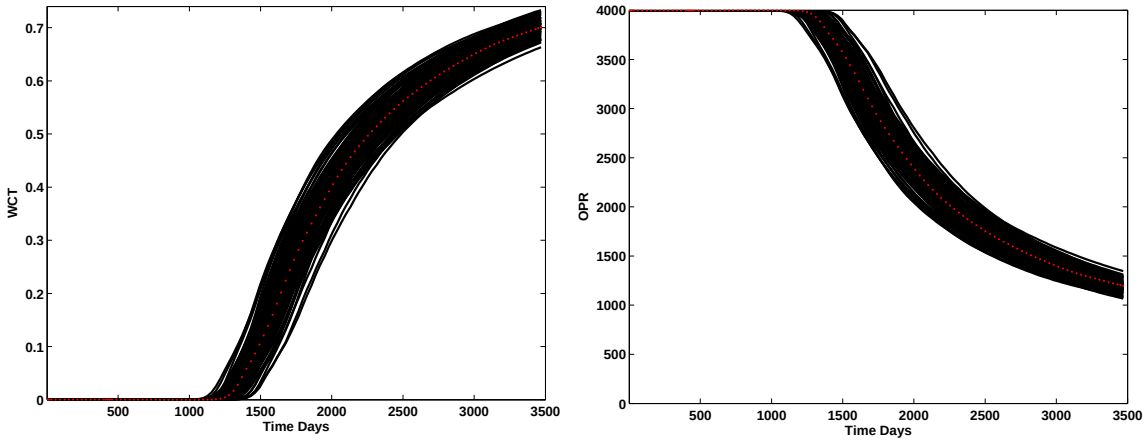


Figure 7: Predictions based on the initial ensemble. Red dots: True.

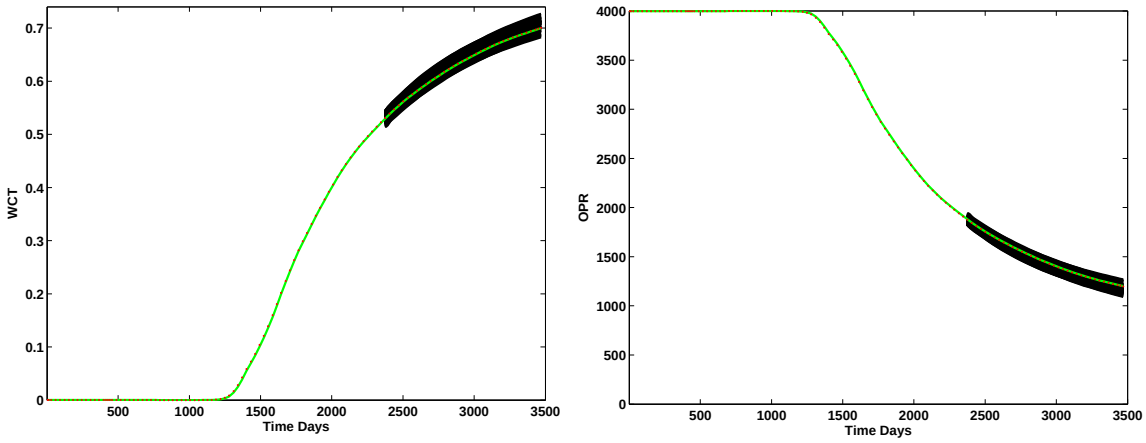


Figure 8: Predictions Case 1, with production data only. Red dots: True. Green line: Simulation based on final ensemble mean porosity and permeability.

the Gaussian noise realizations. Notice that E is sampled with rank equal to $N-1$.

To evaluate the results it is interesting to look at how the conditioning and the rank of the ensemble impacts the analysis, and in Figures 10 and 11 we have plotted the singular value spectra for the ensemble of porosities and permeabilities at the end of the analysis time for each of the six cases. The initial singular spectrum of the ensemble of the static variables is plotted as the upper black line. The lower blue, red and green lines show the singular spectrum at the end of the experiment for each of the cases, and Figures 10 and 11 indicate respectively the reduction of the ensemble variance when either a diagonal error covariance matrix or a error covariance matrix with an exponential correlation structure has been used.

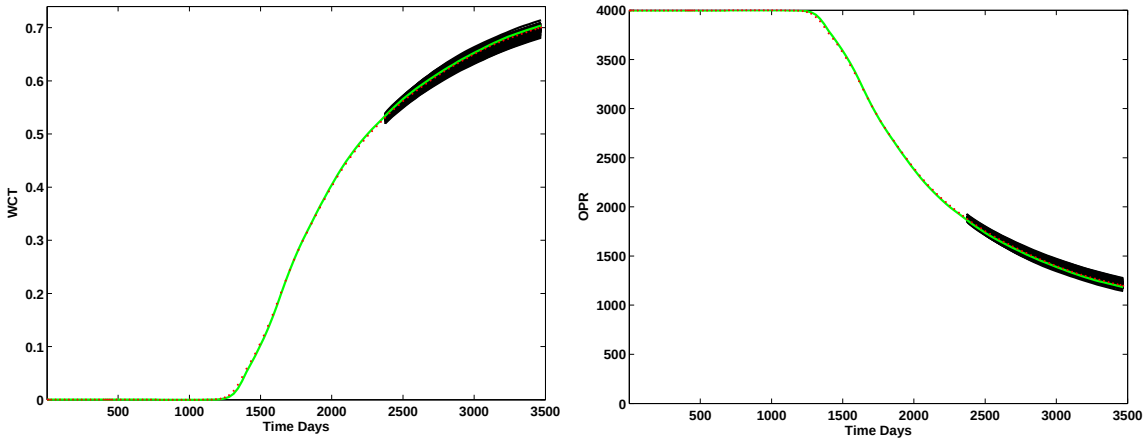


Figure 9: Predictions Case 3a, with production data and seismic data. Red dots: True. Green line: Simulation based on final ensemble mean porosity and permeability.

Further, both the RMS values and the correlation values between the estimated mean ensemble and the true fields for the static variables obtained at the end of each run for all the cases have been evaluated, and they are presented in Table 3.

Comparing cases 2a and 4a we see that they produce singular spectra, RMS and correlation values that are almost identical, which is expected since the pseudo inverse in Eq.(65) will contain the same eigenvector and eigenvalues as the pseudo inverse in Eq.(51) when a full rank diagonal error covariance matrix is used. In case 5a we avoid the formation of the full measurement error covariance matrix, and both the RMS and the correlation values are consistent with the values obtained from cases 2a and 4a. A slightly larger reduction in the variance compared to cases 2a and 4a can be seen, which is in agreement with the results presented by Evensen (2004).

In cases 2b and 4b we have used a full rank error covariance matrix with an exponential correlation structure, and comparing cases 2b and 4b we see from Figure 11 that a lower reduction of the ensemble variance by using the subspace EnKF scheme is obtained. The lower variance reduction is related to the calculation of the approximate pseudo inverse in Eq.(65), and as can be seen, case 4b produces slightly higher RMS values then case 2b. Looking at the correlation values in Table 3, we see that the updated ensemble mean of the static variables from both cases are fairly similar. In case 5b we have used the matrix E generated from a Cholesky decomposition method in the update. The RMS and the correlation values obtained are respectively a bit higher and lower than seen for cases 2b and 4b, and a slightly larger reduction in the variance compared to case 4b is obtained.

From the above experiments it is clear that there is no loss of rank in the analyzed ensemble, when the subspace EnKF scheme has been used. Notice that the subspace EnKF scheme

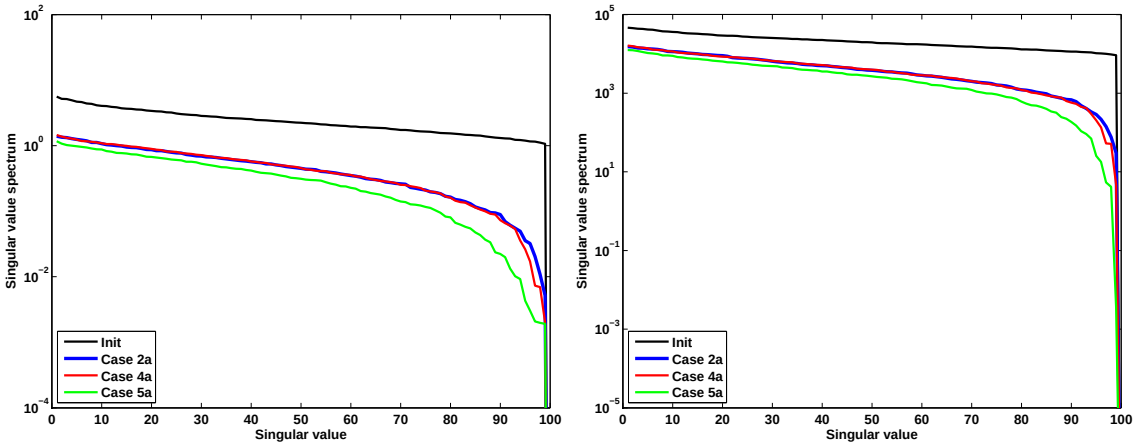


Figure 10: The singular value spectra for the ensemble of porosities (left) and permeabilities (right) for cases $2a$, $4a$ and $5a$.

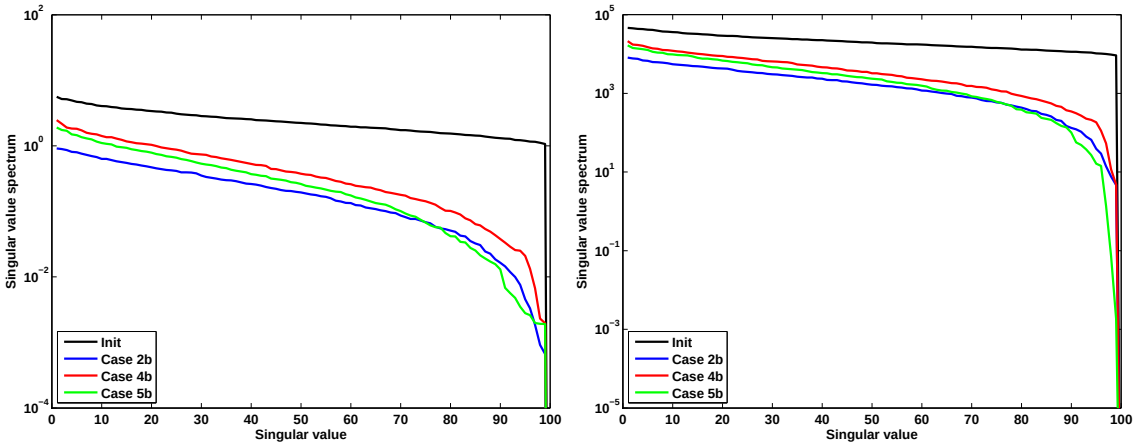


Figure 11: The singular value spectra for the ensemble of porosities (left) and permeabilities (right) for cases $2b$, $4b$ and $5b$.

gives a significant saving in the computational cost compared to the traditional EnKF when $m \gg N$.

8 Conclusion

A methodology to solve the combined parameter and state estimation problem using time difference data is presented in a Bayesian framework. The method is completely recursive,

and thus the solution may be calculated sequentially and forward in time.

A sequential EnKF simulation from the posterior distribution of interest using time difference data and a nonlinear evolution model, has been derived and examined. The EnKF simulation can be viewed as a conditional simulation with kriging, and we have shown that introduction of time difference data in the methodology involves a combination of the ensemble Kalman filter and the ensemble Kalman smoother. The method requires little additional cost compared to the traditional EnKF method.

Special challenges related to a potential loss of rank and the inversion of a large covariance matrix are introduced in the assimilation of large data sets, such as time difference seismic data. This resulted in a subspace EnKF scheme derived in Section 6, based on a principal component projection of the measurements, where the projection matrix U span the ensemble space $\mathcal{L}_S$. In assimilations where $m \gg N$, the projection avoid the loss of rank when a low rank representation of the measurement error covariance matrix is used.

The ensemble Kalman filter for continuous model updating with respect to the combination of production data, 3D seismic data and time difference seismic data is tested on a synthetic reservoir model. For the synthetic reservoir model, the method gives good history matching results, and it is shown that a better estimate of the porosity and the permeability is obtained by adding seismic data to the production data measurements.

Additional simulation runs have been performed to compare the traditional and the subspace EnKF scheme, where different combinations of a diagonal and a correlated measurement error covariance matrix have been used. It is shown from the simulation runs that the results achieved from the subspace EnKF scheme, when a diagonal measurement error covariance matrix is used, are nearly identical to the traditional EnKF scheme. Further, the simulation runs show that the use of a correlated measurement error covariance matrix in the subspace EnKF scheme gives a lower reduction of the ensemble variance compared to the traditional scheme, but results in similar RMS values and correlations values between the estimated and the true porosity and permeability fields. The computation time of the subspace EnKF scheme, when a large amount of data is incorporated in the assimilation, is very small compared to the traditional EnKF algorithm. The results are fairly consistent, although the subspace EnKF scheme introduces an approximate pseudo inverse of the original matrix C, when a correlated measurement error covariance matrix is included.

Appendix A: Square root analysis algorithm

From Section 6 the subspace square root algorithm can be written as

$$\Psi^{a'}\Psi^{a'T} = \Psi'(I - \hat{S}^T\hat{C}^{-1}\hat{S})\Psi'^T, \tag{67}$$

and the analysis equation for the ensemble mean is given by

$$\bar{\Psi}^a = \bar{\Psi} + \Psi'\hat{S}^T\hat{C}^{-1}(\hat{d} - \hat{H}\bar{\Psi}). \tag{68}$$

Notice that because $\hat{C} \in \Re^{N-1\times N-1}$ is a full rank matrix the inverse, $\hat{C}^{-1}$, exists and we can compute the eigenvalue decomposition $Z\hat{\Lambda}Z^T = \hat{C}$, and obtain

$$\hat{C}^{-1} = Z\hat{\Lambda}^{-1}Z^T, \tag{69}$$

where $Z \in \Re^{N-1\times N-1}$ and $\hat{\Lambda} \in \Re^{N-1\times N-1}$. We can now write the square root scheme as follows

$$\begin{aligned}
\Psi^{a'}\Psi^{a'T} &= \Psi'(I - \hat{S}^T\hat{C}^{-1}\hat{S})\Psi'^T \\
&= \Psi'(I - \hat{S}^TZ\hat{\Lambda}^{-1}Z^T\hat{S})\Psi'^T \\
&= \Psi'(I - (\hat{\Lambda}^{-\frac{1}{2}}Z^T\hat{S})^T(\hat{\Lambda}^{-\frac{1}{2}}Z^T\hat{S}))\Psi'^T \\
&= \Psi'(I - Q^TQ)\Psi'^T,
\end{aligned} \tag{70}$$

where $Q \in \Re^{N-1\times N}$ is defined as

$$Q = \hat{\Lambda}^{-\frac{1}{2}}Z^T\hat{S}, \tag{71}$$

and $\text{rank}(Q) = N - 1$. Defining the singular value decomposition of Q as

$$U_Q\Lambda_QV_Q^T = Q, \tag{72}$$

with $U_Q \in \Re^{N-1\times N-1}$, $\Lambda_Q \in \Re^{N-1\times N}$ and $V_Q \in \Re^{N\times N}$, the square root scheme can be written as

$$\Psi^{a'}\Psi^{a'T} = \Psi'(I - (U_Q\Lambda_QV_Q^T)^T(U_Q\Lambda_QV_Q^T))\Psi'^T \tag{73}$$

$$= \Psi'(I - V_Q\Lambda_Q^T\Lambda_QV_Q^T)\Psi'^T \tag{74}$$

$$= \Psi'V_Q(I - \Lambda_Q^T\Lambda_Q)V_Q^T\Psi'^T \tag{75}$$

$$= \left(\Psi'V_Q\sqrt{I - \Lambda_Q^T\Lambda_Q}\right)\left(\Psi'V_Q\sqrt{I - \Lambda_Q^T\Lambda_Q}\right)^T. \tag{76}$$

The analyzed ensemble perturbations is then given by

$$\Psi^{a'} = \Psi'V_Q\sqrt{I - \Lambda_Q^T\Lambda_Q}\Theta, \tag{77}$$

where Θ is a random orthogonal matrix. Similarly to the algorithm presented by Evensen (2004), there are no matrix operations which requires $\mathcal{O}(m^2)$ floating point operations.

References

L. Bertino, G. Evensen, and H. Wackernagel. Sequential Data Assimilation Techniques in oceanography. *International Statistical Review*, 71:2,223–241, 2003.

N. A. C. Cressie. *Statistics for Spatial Data*. Wiley, Iowa State University, 1991.

G. Evensen. Sequential data assimilation with nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, 99:10143–10162, 1994.

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

G. Evensen. Sampling strategies and square root analysis schemes for the EnKF. *Ocean Dynamics 04*, 54:539–560, 2004.

G. Evensen. The combined parameter and state estimation problem. *Submitted to Computational Geoscience*, 2005.

G. Evensen and P. J. van Leeuwen. An Ensemble Kalman Smoother for nonlinear dynamics. *Mon. Weather Rev.*, 128:1852–1867, 2000.

G. Gao and A. C. Reynolds. Quantifying Uncertainty for the PUNQ-S3 Problem in a Bayesian Setting with RML and EnKF. *SPE Reservoir Simulation Symposium, SPE 93324*, 2005.

Y. Gu and D. S. Oliver. History Matching of the PUNQ-S3 Reservoir Model Using the Ensemble Kalman Filter. *SPE Annual Technical Conference and Exhibition, SPE 89942*, 2004.

M. C. Haverl, M. Aga, and E. Reiso. Integrated workflow for quantitative use of time-lapse seismic data in history matching. *SPE 14th Europec Biennial Conference, SPE 94453*, 2005.

J. D. Kepert. On ensemble representation of the observation error covariance in the Ensemble Kalman Filter. *Ocean Dynamics '04*, 6:539–560, 2004.

N. Liu and D. S. Oliver. Critical evaluation of the Ensemble Kalman Filter on history matching of geologic facies. *SPE Reservoir Simulation Symposium, SPE 92867*, 2005.

G. Nævdal, L. M. Johnsen, S. I. Aanonsen, and E. H. Vefring. Reservoir Monitoring and Continuous Model Updating Using Ensemble Kalman Filter. *SPE Annual Technical Conference and Exhibition*, 2004.

G. Nævdal, T. Mannseth, and E. H. Vefring. Instrumented wells and near-well reservoir monitoring through ensemble Kalman filter. *Proccedings of 8th European Conference on the Mathematics of Oil Recovery*, 2002a.

G. Nævdal, T. Mannseth, and E. H. Vefring. Near-well reservoir monitoring through ensemble Kalman filter. *SPE Improved Oil Recovery Symposium*, 2002b.

J. A. Skjervheim, G. Evensen, S. I. Aanonsen, B. O. Ruud, and T. A. Johansen. Incorporating 4D Seismic Data in Reservoir Simulation Models Using Ensemble Kalman Filter. *SPE Annual Technical Conference and Exhibition, SPE 95789*, 2005.

A. Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation.* SIAM, Philadelphia, 2005.

P. J. van Leeuwen and G. Evensen. Data assimilation and inverse methods in terms of a probabilistic formulation. *Mon. Weather Rev.*, 124:2898–2913, 1996.

X.-H. Wen and W. H. Chen. Real-Time Reservoir Model Updating Using Ensemble Kalman Filter. *SPE Reservoir Simulation Symposium, SPE 92991*, 2005.

M. West and J. Harrison. *Bayesian forecasting and Dynamic Models.* 2 edn, New York: Springer, 1997.

Paper C

Combined inversion of 4D seismic waveform data and production data using ensemble Kalman filter *

* 76th Annual International Meeting, *SEG*, Expanded Abstracts, 1776-1780

Combined inversion of 4D seismic waveform data and production data using ensemble Kalman filter

Jan-Arild Skjervheim and Bent Ole Ruud, Centre for Integrated Petroleum Research, University of Bergen*

SUMMARY

In this computational study we combine 4D seismic data with production data to continuously update the reservoir model during production. The inversion is based on the ensemble Kalman filter (EnKF), and the forward method is a combination of a fast seismic modeling tool and a reservoir simulator. The EnKF method is a Monte Carlo approach, and state variables, as fluid saturations and pressure, and model input parameters as the porosity and permeability, are updated in the reservoir model at each assimilation step. The assimilated measurements are seismic waveforms data and production data, which are measured gas-oil ratio, water cut and flowing bottomhole pressure. The updated models allow for improved estimation of the parameters and the state variables during the production, because additional well data and repeated seismic surveys are sequentially included in the inversion. The method is applied to a synthetic 2D reservoir model, and it is shown that introduction of production and seismic waveforms data gives a fairly good estimation of the porosity and permeability fields.

INTRODUCTION

The purpose of repeated seismic acquisitions (4D) is to reveal changes in the fluid saturations and fluid pressure during production. Such information is crucial when trying to maximize oil and gas recovery from a field. To get the most out of the seismic and production data, they should be combined in order to get a consistent model of the dynamic reservoir. This is usually done by running the inversion of the seismic data and the history matching of the production data as separate processes and then check for consistency in the results, but ideally the inversion of both types of measurements should be performed simultaneously. The latter approach has been reported by Gosselin et al. (2003), Haverl et al. (2005) and Skjervheim et al. (2005) by incorporating the seismic data as elastic parameters (P-wave impedance and Poisson's ratio) in the history matching process. In the present study, we do not rely on a previous inversion of the seismic data to elastic parameters, but directly include the seismic waveform data in the extended inversion process. Before inversion the seismic data should be processed to obtain migrated angle stacks, and the (assumed known) source signature should be removed by deconvolution. Otherwise, only a time integration operator (mainly a phase rotation) is applied before the waveform data are included in the extended inversion process realized as an ensemble Kalman filter.

The ensemble Kalman filter method is a Monte Carlo type sequential Bayesian inversion, and provides an approximate solution to the combined parameter and state estimation problem, see Evensen (2005). The result is an ensemble of solutions approximating the posteriori probability density function (pdf) for the reservoir model input parameters (porosity and permeability), state variables (saturations and pressures) and other output data (well production and seismic data) conditioned to measured, dynamic data. In our study, we are using a three phase (gas,oil and water) black-oil reservoir simulation model.

METHOD

Seismic modelling and processing

The seismic modelling tool used in the study (SimPLI) computes directly prestack depth-migrated seismic angle stacks in 2D or 3D, see Lecomte et al. (2003). The only additional processing we have applied to the data before inclusion the EnKF is a deconvolution and a time integration. The motivation for this preprocessing originates from the simple 1D convolutional model for computation of synthetic seismograms, i.e.,

$$y(t) = S(t) * R(t), \tag{1}$$

where the synthetic trace $y(t)$ is computed by convolving the time series of reflection coefficients $R(t)$ with the source wavelet $S(t)$. The reflection coefficients may be approximated by

$$R(t) = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{dt}} \ln Z(t), \tag{2}$$

where $Z(t)$ is the P-wave impedance, see e.g., Buland and Omre (2003). If we assume that the source wavelet is known, an approximation for the inverse wavelet $S^{-1}(t)$ may be found by a least square method, and $\ln Z(t)$ may be retrieved directly, except for an integration constant C, by a time integration

$$\ln Z(t) = 2\int_0^t S^{-1}(t') * y(t')dt' + C. \tag{3}$$

Here $S^{-1}(t)$ serves as a deconvolution operator. It is the traces denoted $\ln Z(t)$ which are input to EnKF, but we emphasize that we consider the deconvolution and time integration just as filter operations and not as seismic inversion. The seismic forward modelling method we are using is not restricted to 1D media, and the physical meaning of the input data is of no importance to the EnKF as long as the forward modelling method is able to produce the same kind of data. In principle, we may skip these filter operations, but we have experienced that better results are obtain when they are included. Examples of unfiltered and filtered data are shown in Figure 1.

EnKF

Following Tarantola (2005) the general solution of the combined parameter and state estimation problem is given as the posterior pdf from the Bayesian theory. Let now $\psi_n = \psi(x,t_n) \in \Re^{q_n}$ and $d_n \in \Re^{m_n}$ represent respectively the parameters, the state variables and the observations at time t_n, and assume that we want to estimate the parameters and the states at a discrete set of times, $t_0, t_1, \ldots, t_n$, given observations at the same times (except at t_0) $d_1, d_2, \ldots, d_n$. In the following we denote the sequence $\{\psi_n, \ldots, \psi_1, \psi_0\}$ by $\psi_{n:0}$.

In Evensen and van Leeuwen (2000) it was shown that a general smoother and filter could be expressed from the Bayesian formalism, and for the time interval $t \in [t_0, t_n]$ the general smoother can be written as

$$f(\psi_{n:0}|d_{n:1}) \propto f(d_{n:1}|\psi_{n:0})\, f(\psi_{n:0}), \tag{4}$$

where data available at different times are assumed independent, and that the data at a given time are independent on states at future times. In Eq.(4) the prior pdf is represented by $f(\psi_{n:0})$, and the likelihood distribution is given as $f(d_{n:1}|\psi_{n:0})$.

A first order Markov process is normally assumed in the forward model. That is, the solution at a given time step depends only on the solution at the previous time step. The filter solution for Eq. (4) can then be derived by integrating over the solutions $\psi_{n-1:0}$, and we obtain

$$f(\psi_n|d_{n:1}) \propto f(d_n|\psi_n)\, f(\psi_n|d_{n-1:1}). \tag{5}$$

When the seismic data are given as a difference between two surveys at time $t_k < t_n$ and t_n, the estimation problem include both a filter and a smoother solution, see Skjervheim et al. (2005), and the updating step is given as

$$f(\psi_n, \psi_k|d_{n:1}) \propto f(d_n|\psi_n, \psi_k)\, f(\psi_n, \psi_k|d_{n-1:1}). \tag{6}$$

To sample from the posterior distribution $f(\psi_n|d_{n:1})$ we are using an ensemble Kalman filter technique, see Evensen (2003).

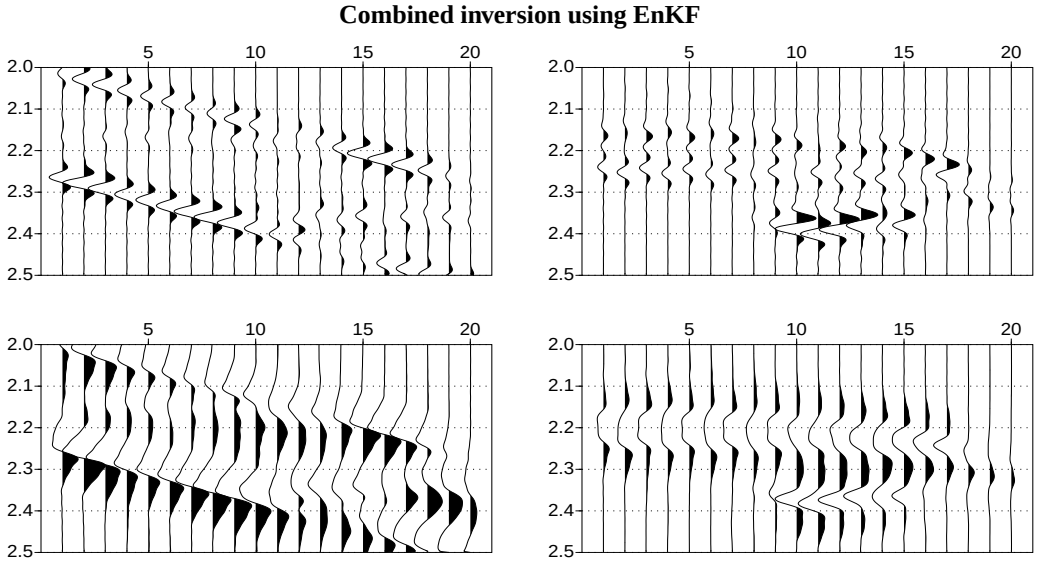


Figure 1: Synthetic depth converted seismic data before (upper part) and after deconvolution and time integration (lower part). The sections to the left are the initial seismic data, and the sections to the right are the difference between the data acquired three years after start of production and the initial data. The difference sections (at the right) have been scaled up by a factor of 2.5 relative to the initial data. The initial seismic sections are dominated by reflections from the top and bottom of the reservoir, while the difference sections clearly shows the changes in the fluid saturations, especially on the deconvolved and time integrated section (see also Figure 2).

The matrix holding the ensemble members, $\psi_n \in \Re^{q_n}$, is defined by

$$\Psi_n = \Psi(\mathbf{x}, t_n) = (\psi_n^1, \ldots, \psi_n^N) \in \Re^{q_n \times N}, \tag{7}$$

where N is the number of ensemble members. The ensemble covariances, $P_{\psi_n} \in \Re^{q_n \times q_n}$, can then be defined as

$$P_{\psi_n} = \frac{(\Psi_n - \bar{\Psi}_n)(\Psi_n - \bar{\Psi}_n)^T}{N-1}, \tag{8}$$

where $\bar{\Psi}_n$ is the ensemble mean matrix. Given a vector of measurements, $d_n \in \Re^{m_n}$, we can define the N vectors of perturbed measurements as

$$d_n^j = d_n + \varepsilon_n^j \quad \forall \quad j = 1:N, \tag{9}$$

which can be stored in the columns of a matrix

$$D_n = (d_n^1, \ldots, d_n^N) \in \Re^{m_n \times N}, \tag{10}$$

where $\varepsilon_n \sim N_{m_n}(0, P_{\varepsilon_n})$ are Gaussian measurement errors. The ensemble representation of the analysis step is

$$\Psi_n^a = \Psi_n^f + P_{\psi_n}^f H_n^T (H_n P_{\psi_n}^f H_n^T + (N-1)P_{\varepsilon_n})^{-1}(D_n - H_n \Psi_n^f), \tag{11}$$

where Ψ_n^f is the forecast (prior) and Ψ_n^a is the analysis (posterior) for all the members in the ensemble at time t_n.

For time difference data, the analysis scheme is similar, except that we also have to update the smoother solution at the time of the previous survey according to Eq(6). The EnKF update in Eq(11) can still be used, but Ψ need to be replaced by the combination of Ψ_n and Ψ_k. Hence, the analysis equation can be written as

$$\tilde{\Psi}_n^a = \tilde{\Psi}_n^f + P_{\tilde{\psi}_n}^f \tilde{H}_n^T (\tilde{H}_n P_{\tilde{\psi}_n}^f \tilde{H}_n^T + (N-1)P_{\varepsilon_n})^{-1}(D_n - \tilde{H}_n \tilde{\Psi}_n^f), \tag{12}$$

where $\tilde{\Psi}_n^a = [\Psi_n, \Psi_k]^T$ and $\tilde{H}_n$ is the observation operator on the form

$$\tilde{H}_n = \begin{bmatrix} H_n^{prod} & 0 \\ H_n^{seis} & -H_k^{seis} \end{bmatrix}, \tag{13}$$

where H^{prod} picks the simulated production data, and H^{seis} picks the simulated seismic data.

RESULTS

The test of combining assimilation of production data and seismic data has been performed on a synthetic 2D vertical reservoir model with 40×10 grid cells. The reservoir is produced by gas injection, where the gas injector is located in grid cell (1,1) and penetrated in one layer, while the oil producer is located in grid cell (30,1) and penetrated in three layers. Total simulation time is $4\frac{1}{4}$ years, where $3\frac{1}{4}$ years are history and 1 year prediction, and during the run, 4 seismic survey and 14 assimilation step of production data are used in the EnKF update. The seismic data from the 3 last surveys were included in the inversion by computing difference data relative to the first survey. The fluid saturations at an early and a late time after the start of production are shown in Figure 2.

Seismic data are waveforms data, and production data are measurements of gas-oil ratio (GOR), water cut (WCT) and flowing bottom-hole pressure (BHP). Measurement errors for the production data are presented in Table 2, and the signal to noise ratio in the seismic data is set to 10. The elastic properties of the dry reservoir rock depends on the porosity through a simple critical porosity sandstone model. The fluid properties are computed from standard relations, and Wood's equation is used for mixed fluids. The effective elastic properties of the saturated rock are then found from Gassmann's relations. The pressure depends mainly on depth in our modeling case, so that the effects of pressure variations with time are negligible.

The reference ("true") permeability and porosity fields (see Figure 3) were simulated as Gaussian fields, and the ensemble consists of 100 members with a spherical variogram and constant means. Model data are listed in Table 1. In evaluating the quality of the estimation of the permeability and the porosity we use the root mean square (RMS) error defined as

$$\Gamma(X) = \sqrt{\frac{1}{N_g}\sum_{j=1}^{N_g}\left(\frac{1}{N}\sum_{i=1}^{N}(X_j^i - X_j^t)^2\right)}, \tag{14}$$

where N_g is the number of grid cells, N is the number of ensemble members, X_j^i is the estimate of static variable j in ensemble member i, and X_j^t is the true value of static variable j.

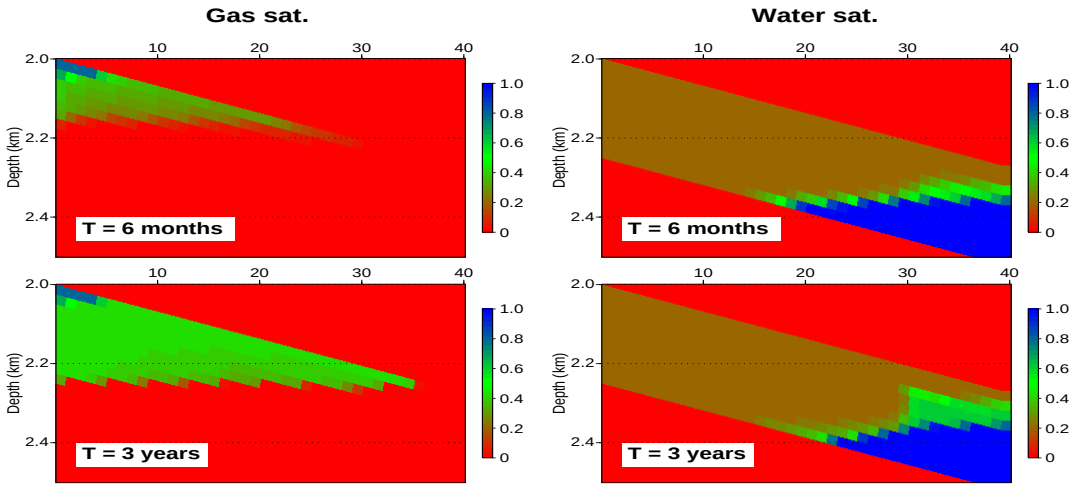


Figure 2: Fluid saturations at two different times (6 months and 3 years) after start of production. A production well is situated at block number 30 (see numbering along top of model), and a gas injection well is found in the upper left corner at block number 1.

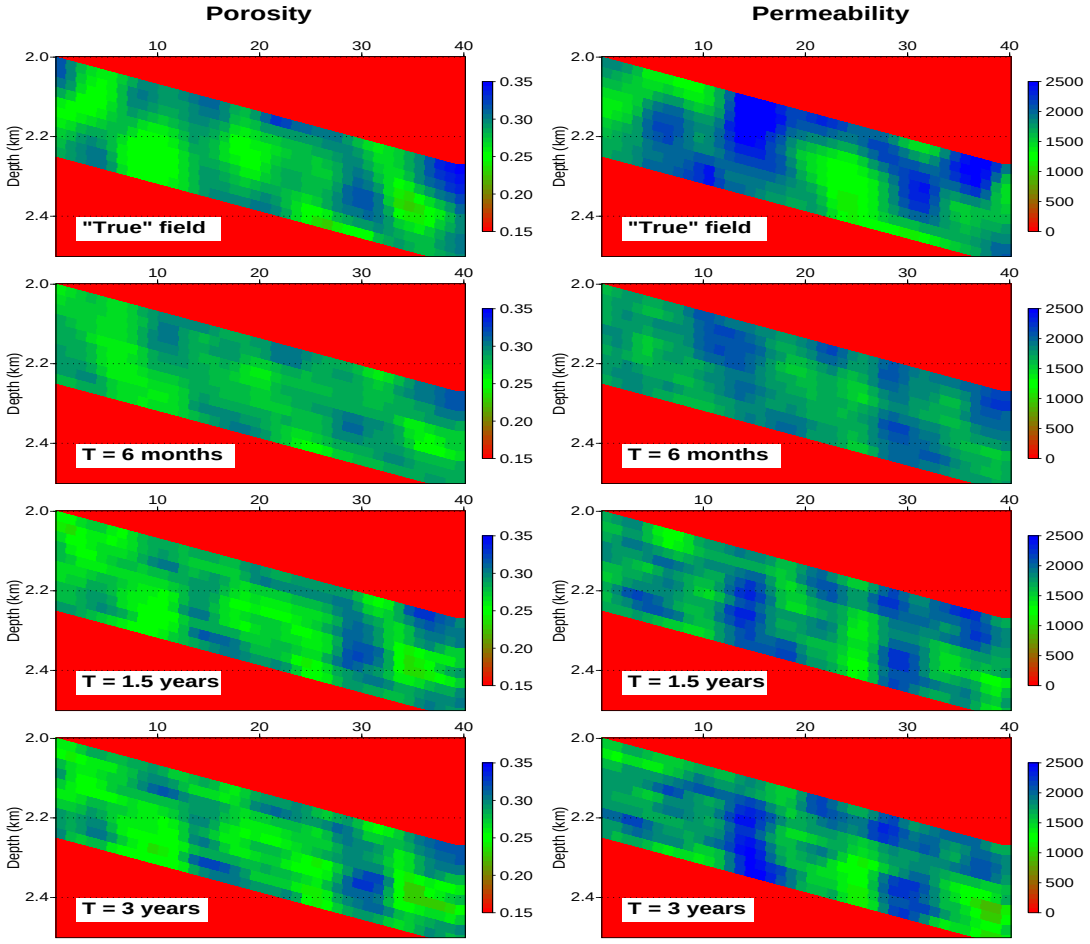


Figure 3: The evolution of the mean ensemble porosity and permeability at three different assimilation steps compared to the true solution. Permeability in millidarcy.

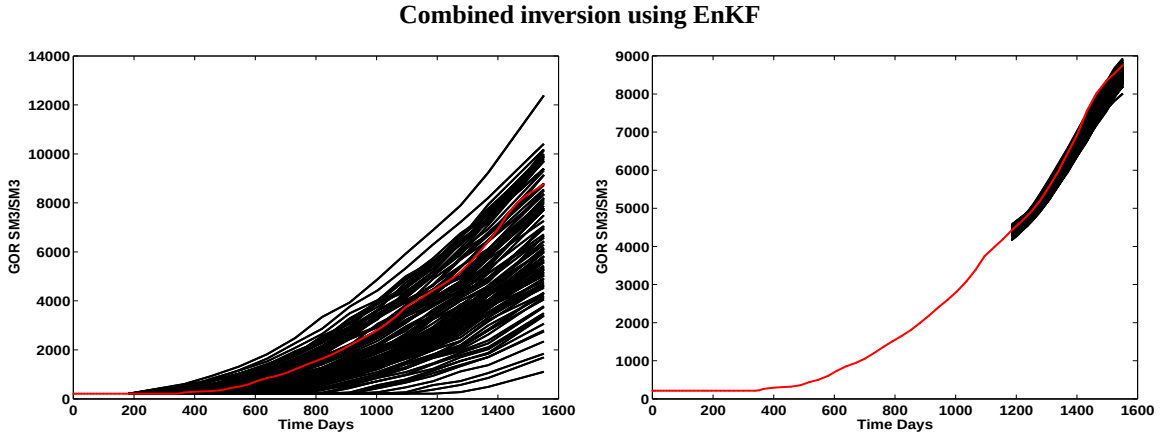


Figure 4: Initial prediction of GOR (left) and GOR predicted from the last assimilation step and one year forward (right). Red line: True production profile of GOR, based on the true porosity and permeability fields given at the top of Figure 3.

	True	Initial
	Layers 1-10	Layers 1-10
Mean perm (mD)	1800	1800
Std. of perm (mD)	200	350
Mean poro.	0.28	0.28
Std. of poro.	0.02	0.03
Variogram range along x dir.	5 blocks	3 blocks
Variogram range along z dir.	5 blocks	3 blocks
Correlation between perm. and poro.	0.7	0.7

Table 1: Input data to synthetic reservoir model.

Water cut (fraction)	Gas-oil ratio (Sm^3/Sm^3)	Bottomhole pressure (bar)
0.03	10 %	2.0

Table 2: Measurement errors for well data, standard deviations.

Figure 4 shows the predictions of the ensemble of gas-oil ratio, which are compared with the true solution. The uncertainty in the predictions cover the true solution quite well, and compared with the predictions based on the initial ensemble, a clearly reduction in the uncertainty is achieved.

Figure 3 shows the averaged permeability and porosity fields at three different assimilation step, where the true fields are plotted at the top. As can be seen, the combined assimilation of production and seismic waveforms data gives a fairly good estimation of the main features of the true permeability and porosity fields.

The initial RMS value of the permeability is 497.27 md, and it is reduced to a value of 355.44 md at the last assimilation step, while the RMS value for the porosity is reduced from 0.036 to 0.018. Clearly, a decrease in the RMS deviations for the reservoir parameters are obtained, and the results show a good relation between the estimated and the true parameter fields. Especially, a strong relation between the estimated and the true porosity field is achieved, which is expected because the seismic signal has a stronger correlation to the porosity than to the permeability in the reservoir.

CONCLUSIONS

The ensemble Kalman filter for continuous model updating of the parameters (permeability and porosity) and the state variables (saturations and pressures) with respect to a combination of seismic waveform data and production data are presented. For the synthetic reservoir model, the EnKF update gives good estimation results of the permeability and porosity fields. The method demonstrate that seismic waveforms data can be introduced in the methodology, and that the estimates of both the porosity and permeability fields can be improved over time with repeated seismic surveys and assimilation of production data. When using seismic waveforms data, it is important that the members of the ensemble correctly reflects the uncertainty both in the reservoir model and the seismic modelling during the assimilation, and this have to be further investigated.

ACKNOWLEDGMENTS

We thank Norsk Hydro and Statoil for sponsoring the Dynamic Reservoir Characterization Project at CIPR.

REFERENCES

Buland, A. and H. Omre, 2003, Bayesian linearized AVO inversion: Geophysics, **68**, 185–198.

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

——— 2005, The combined parameter and state estimation problem: Submitted to Computational Geoscience.

Evensen, G. and P. J. van Leeuwen, 2000, An Ensemble Kalman Smoother for nonlinear dynamics: Mon. Weather Rev., **128**, 1852–1867.

Gosselin, O., S. I. Aanonsen, I. Aavatsmark, A. Cominelli, R. Gonard, M. Kolasinski, F. Ferdinandi, L. Kovacic, and K. Neylon, 2003, History Matching Using Time-Lapse Seismic (HUTS): SPE Annual Technical Conference and Exhibition, SPE 84464.

Haverl, M. C., M. Aga, and E. Reiso, 2005, Integrated workflow for quantitative use of time-lapse seismic data in history matching: SPE 14th Europec Biennial Conference, SPE 94453.

Lecomte, I., H. Gjøystdal, and Å. Drottning, 2003, Simulated Prestack Local Imaging: a robust and efficient interpretation tool to control illumination, resolution, and time-lapse properties of reservoirs: Expanding Abstract, SEG International Exposition and Seventy-Third Annual Meeting, Dallas 26-31 October, 1529–1532,(RCT 7.2).

Skjervheim, J.-A., G., Evensen, S. I. Aanonsen, B. O. Ruud, and T. A. Johansen, 2005, Incorporating 4D Seismic Data in Reservoir Simulation Models Using Ensemble Kalman Filter: SPE Annual Technical Conference and Exhibition, SPE 95789.

Tarantola, A., 2005, Inverse problem theory and methods for model parameter estimation: SIAM.

Paper D

Using the Ensemble Kalman Filter with 4D Data to Estimate Properties and Lithology of Reservoir Rocks *

* Proceeding at *ECMOR X*, Amsterdam, The Netherlands, September 2006.

A028 USING THE ENSEMBLE KALMAN FILTER WITH 4D DATA TO ESTIMATE PROPERTIES AND LITHOLOGY OF RESERVOIR ROCKS

J.A. SKJERVHEIM[1], B.O. RUUD[1], S.I AANONSEN[1], G. EVENSEN[2], T.A. JOHANSEN[1]

[1] *Centre for Integrated Petroleum Research, University of Bergen, Norway*
[2] *Norsk Hydro Research Centre*

Abstract

Improvement of the seismic reservoir characterisation is performed by introducing an uncertainty connected to parameters in the rock physics model. Traditionally the Ensemble Kalman filter (EnKF) method has been used to estimate permeability and porosity.

In this paper we show that when seismic difference data are available also the lithology can be estimated, which is coupled to the effective bulk modulus via the rock physics model. Incorporation of inverted seismic difference data in the EnKF introduces a large amount of data in the assimilation step. Thus, to improve the results a methodology based on a combination of a global and a local analysis scheme is proposed. The global and the local analysis are used to assimilate respectively the production data and the inverted seismic difference data, where the local scheme assume that only seismic data within a certain distance from a state variable will impact the analysis in this state variable. The technique is applied to synthetic 2D and 3D reservoir models, where effects of using local versus global analysis schemes on different inverted seismic difference data, such as acoustic impedance and Poisson's ratio, are investigated. Other evaluated factors are the effects of using an incorrect seismic data error model in the analysis schemes.

Introduction

This paper describes an estimation of both static variables (lithology, porosity and permeability) and dynamic variables (saturations and pressures) in a Bayesian framework based on the information contained in the assimilated reservoir measurements, such as water cut (WCT), flowing bottom hole pressure (BHP), gas-oil ratio (GOR) and seismic difference data. The use of seismic data may introduce a large number of data points in the assimilation step, and problems related to an underestimation of the uncertainty in the low rank covariance matrix representation of the static and dynamic variables may occur when few ensemble members are used during the assimilation.

To avoid this problem we propose a combination of a global and local EnKF. Assimilation of production data is performed with a global scheme, while the seismic difference data are assimilated with a local scheme, where we assume that only seismic data within a certain distance from a state variable will impact the analysis in this state variable. Different formulations and applications of the local scheme can be found in Evensen (2003); Ott et al. (2004); Szunyogh et al. (2005).

The local scheme updates each grid cell independently accounting for the seismic observations only in the local region surrounding the cell. The local schemes can be implemented in parallel, and each local update will involve relatively small matrices in the assimilation. Thus, with the local scheme, large amounts of seismic data can be assimilated at a low computational cost.

The seismic reservoir characterisation can be improved by introducing an uncertainty connected to parameters in the rock physics model. Lithology is such a variable of interest, which for siliciclastic rocks can be quantified by the clay ratio (see e.g., Dræge et al. (2006)). Lithology may have a strong impact on the seismic signal, and an estimation of the clay ratio will be of importance to achieve satisfactory predictions of the elastic parameters.

Combined global and local EnKF scheme

The objective of an ensemble Kalman filter (EnKF) is to simulate samples from the posterior probability density function (pdf) of interest, which from Bayesian theory can be defined as

$$g(\psi|d) \propto g(d|\psi)g(\psi). \tag{1}$$

Here $g(d|\psi)$ is the likelihood pdf for the observations, $d \in \Re^m$, and $g(\psi)$ is the joint prior distribution for the parameters and the model state, $\psi \in \Re^q$.

When working with seismic difference data, the measurements at a given time may depend on the model state at more than one time instant. Such seismic data may be a result of a 4D inversion, where the seismic difference between two surveys are inverted to differences in elastic parameters. The individual surveys of the elastic parameters may then not be available. In Skjervheim et al. (2006) it was seen that the posterior distribution then included both a filter and a smoother solution, and that the EnKF simulation involves a combination of the ensemble Kalman filter and the ensemble Kalman smoother. Further, it was shown that this could be handled in a similar framework as in the traditional EnKF, and for simplicity we here present the equations on a filter form.

Let now $\psi_j = \psi(\mathbf{x}, t_j) \in \Re^{q_j}$ and $d_j \in \Re^{m_j}$ represent the model variables and the measurements at time t_j, where the observations $d_j = [d_{p_j}, d_{s_j}]^T$ at a particular time t_j depend on ψ_j. The likelihood distribution is then represented by $g(d_j|\psi_j)$. Let d_{p_j} denote production observations and d_{s_j} denote seismic difference observations. In a Bayesian formulation the posterior distribution can be represented as

$$g(\psi_j|d_j, \ldots, d_1) \propto g(d_j|\psi_j)g(\psi_j|d_{j-1}, \ldots, d_1), \tag{2}$$

and to sample from $g(\psi_j|d_j,\ldots,d_1)$ a combination of a global and a local ensemble Kalman filter is proposed.

The observation equation in the nonlinear state space model is defined by Evensen (2003), as

$$d_j = H_j\psi_j + \epsilon_j, \tag{3}$$

where $H_j = [H_{p_j}, H_{s_j}]^T$ is a linear observation operator. We define, H_{p_j}, as the linear production measurement operator, and H_{s_j} relates the model state ψ_j to the seismic difference measurements. Both operators allow for measurement errors ϵ_j, which are assumed to be Gaussian:

$$\epsilon_j = \begin{bmatrix} \epsilon_{p_j} \\ \epsilon_{s_j} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} P_{\epsilon_{p_j}} & 0 \\ 0 & P_{\epsilon_{s_j}} \end{bmatrix} \right). \tag{4}$$

The terms ϵ_{p_j} and ϵ_{s_j} are respectively production and seismic measurement errors with covariance matrices $P_{\epsilon_{p_j}}$ and $P_{\epsilon_{s_j}}$. Here we have assumed independence between the production and the seismic measurement errors.

The ensemble mean and the ensemble spread of the model variables at any given time reflects both an estimate of the true reservoir state and an uncertainty of that estimate. Each member of the ensemble are evolved independently forward in time, but when new measurements from the reservoir are available the first and second order moments from the ensemble is used to generate the new reservoir state estimate and the reduced uncertainty introduced by the measurements.

The combined global and local EnKF analysis scheme is described in the following. Let the ensemble matrix holding the members $\psi_j \in \Re^{q_j}$ be defined by

$$\Psi_j = \Psi(\mathbf{x}, t_j) = (\psi_j^1, \ldots, \psi_j^N) \in \Re^{q_j \times N}, \tag{5}$$

where N is the number of ensemble members. All the equations below is evaluated at time t_j, and to simplify the notation we have from now dropped the time subscript j.

The ensemble perturbation matrix can be represented by

$$\Psi' = \Psi - \bar{\Psi} = \Psi(I - 1_N), \tag{6}$$

where $\bar{\Psi}$ is the ensemble mean matrix and 1_N is the $N \times N$ matrix where each element is equal to 1/N. The error covariance matrix of the ensemble can be estimated as

$$P_\psi = (N-1)^{-1}\Psi'(\Psi')^T. \tag{7}$$

Given vectors of production data, $d_p \in \Re^{m_p}$, and seismic difference data, $d_s \in \Re^{m_s}$, we can define the N vectors of perturbed measurements as

$$\begin{aligned} d_p^k &= d_p + \epsilon_p^k \\ d_s^k &= d_s + \epsilon_s^k \end{aligned} \quad \forall\ k = 1:N\ , \tag{8}$$

which can be stored in the columns of the matrices

$$\begin{aligned} D_p &= (d_p^1, \ldots, d_p^N) \in \Re^{m_p \times N} \\ D_s &= (d_s^1, \ldots, d_s^N) \in \Re^{m_s \times N}. \end{aligned} \tag{9}$$

D_p and D_s are respectively the perturbed measurements matrix for the production data and the seismic difference data. The ensemble representation of the analysis step using production data is

$$\Psi_p^a = \Psi^f + P_\psi^f H_p^T (H_p P_\psi^f H_p^T + P_{\epsilon_p})^{-1}(D_p - H_p\Psi^f). \tag{10}$$

The analysis step Ψ_p^a is now used as input to the local EnKF scheme, where seismic difference data are assimilated.

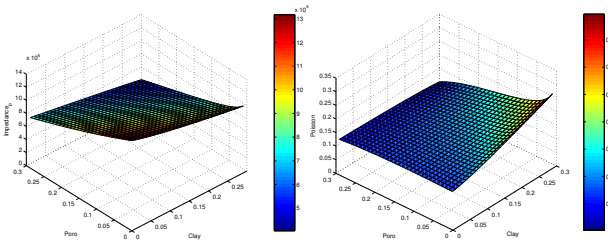


Figure 1: Acoustic impedance and Poisson's ratio, as function of porosity and clay ratio. Evaluated at the saturations: S_w=0.2 and S_o=0.8.

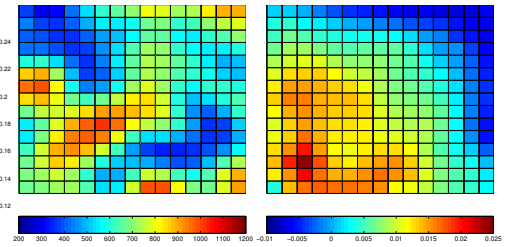


Figure 2: True permeability (left). True Poisson's ratio difference at the last seismic assimilation step (right).

In the local analysis, only data from a local region is applied when updating the state vector in a given grid cell. Thus, the data influencing the analysis for grid cell, l, are given by

$$d_{s_l} = L_l d_s \in \Re^{m_{s_l}}, \tag{11}$$

where the matrix L_l includes only zeros and ones and picks the simulated seismic difference data within this prescribed local region. Correspondingly, the local observation equation can be written as

$$d_{s_l} = H_{s_l}\psi_l + \epsilon_{s_l} \qquad \epsilon_{s_l} \sim N(0, P_{\epsilon_{s_l}}), \tag{12}$$

where $H_{s_l} = L_l H_s$, $\epsilon_{s_l} = L_l \epsilon_s$, and the error covariance matrix is given by $P_{\epsilon_{s_l}} = L_l P_{\epsilon_s} L_l^T$.

Defining the local representation of the seismic difference ensemble perturbations, $S_{s_l} \in \Re^{m_{s_l} \times N}$, and the ensemble of innovation vectors, $D'_{s_l} \in \Re^{m_{s_l} \times N}$, as

$$S_{s_l} = H_{s_l}\Psi_p^{a'} = L_l H_s(\Psi_p^a - \bar{\Psi}_p^a), \qquad D'_{s_l} = D_{s_l} - H_{s_l}\Psi_p^{a'} = L_l(D_s - H_s\Psi_p^{a'}) \tag{13}$$

the local analysis equation at grid point l, i.e., Ψ_l^a, see Evensen (2003), becomes

$$\Psi_l^a = \Psi_{p,l}^a X_l, \tag{14}$$

where X_l can be written as

$$X_l = I + S_{s_l}^T (S_{s_l} S_{s_l}^T + (N-1)P_{\epsilon_{s_l}})^{-1} D'_{s_l}. \tag{15}$$

By solving the local analysis equation (14) for each grid cell in the reservoir, we have determined the global analysis ensemble conditioned to both production and seismic data.

Rock physics model

Several models have been proposed to model the effect of clay in sandstones, see Dræge et al. (2006); Sam and Andrea (2001). The main difference between the various approaches is whether the clay is included as a part of the load-bearing structure or not. In order to choose between alternative models one would in practical cases use geological information, or base the choice on previous experience with the reservoir in question. For the purpose of this numerical experiment we have chosen to use the so-called dispersed clay model, i.e., the clay is considered as dispersed in the pore volume of the sandstone and not as a part of the load-bearing structure. This approach allows the clay to be easily included by any rock physics model, but the clay ratio should not be too high because when the pore-filling clay exceeds about 40% of the pore space, it will become pore-bridging, see Dræge et al. (2006). The rock physics model used here to compute the effective properties of a porous, fluid filled medium is the Kuster-Toksøz model as described by Johansen et al. (2002).

In order to compute the elastic properties for a rock with porosity ϕ and clay ratio c (the volume ratio of clay to the total solid), the pore space must be increased to accommodate the volume of the clay. The pore space volume ϕ' which is used to compute the properties of the load-bearing structure becomes

$$\phi' = \phi + (1-\phi)c. \tag{16}$$

The properties of the pore fluid, including the clay, is computed by the Reuss average. The acoustic impedance and the Poisson's ratio, computed as functions of the porosity and clay ratio, are shown in Fig. 1.

Synthetic 2D example

We consider a synthetic 2D problem with 15×15 grid cells to show the effects of using a local versus a global scheme, when assimilating a large dataset of seismic difference data. The reservoir consists of one oil producer and one water injector located respectively in the upper right and in the lower left corner.

The initial ensemble of clay ratio, porosity and permeability are sampled as Gaussian fields with constant mean and a spherical variogram model, and the reference ("true") fields are generated from the same statistics. The means, variances and other model data are summarised in Table 1.

The production data consist of observations of GOR, BHP and WCT. Seismic difference data are acoustic impedance and Poisson's ratio computed from the rock physics model presented above. The measurement uncertainties are given in Table 2. Here, the seismic measurement errors have the same variogram as the static fields. Two acoustic impedance difference datasets, two Poisson's ratio difference datasets, calculated in all grid cells, and production data every 3 months are assimilated. The simulation time is 8.5 years, where 7 years are history and 1.5 years prediction. The reference permeability field, and the Poisson's ratio difference dataset at the last seismic assimilation step, are plotted in Fig. 2.

The EnKF analysis updates the parameters, which are clay ratio, porosity and permeability, and the state variables, which are saturation and pressure. The clay ratio only appears in the rock physics model, and is therefore updated only at the time steps where seismic difference data are assimilated.

Four experiments of model updating were performed. In Cases 1, 2, and 3 only the global EnKF scheme is applied with respectively 100, 250 and 800 ensemble members. Case 4 uses the combined global and local EnKF scheme, with a local stencil of 5×5 grid cells.

To evaluate the performance of the EnKF update, we have focused on the estimation of permeability. Most of the information affecting this parameter should be in the water flooded areas, propagating forward in time with the water front, and thus the correlation between the seismic difference data and the permeability is expected to be stronger in the lower left part of the reservoir during the assimilation. The quality of the EnKF estimation is measured by the correlation between the ensemble mean and the true field, and by the root mean square (RMS) error defined for the parameter α as

$$\Gamma(\alpha) = \left(N_g^{-1} \sum_{j=1}^{N_g} \left(N^{-1} \sum_{i=1}^{N} (\alpha_j^i - \alpha_j^t)^2 \right) \right)^{1/2}, \tag{17}$$

where α_j^i is the estimate of α in grid cell j in ensemble member i, α_j^t is the corresponding true value, and N_g is the number of grid cells.

The top row of Fig. 3 shows the ensemble of permeability estimates in all grid cells, plotted together with the true field at the last assimilation step for Cases 1 to 4, where the counting of grid cells starts from the lower left to the upper right corner in the reservoir. The middle row

Clay ratio		Porosity		Permeability		Variogram range	Corr. between
Mean	Std.	Mean	Std.	Mean	Std.	along x and y dir.	static fields
0.11	0.03	0.17	0.03	665	170	5 cells	0.0

Table 1: Model data used as input in the EnKF.

WCT (-)	GOR (Sm^3/Sm^3)	BHP(bar)	Acoustic imp. (kg/m^2s)	Poisson's ratio (-)
0.03	10 %	2.0	1.0E+04	1.0E-03

Table 2: Measurement errors for well data and seismic difference data, standard deviations.

Case no.	Init.	Case 1	Case 2	Case 3	Case 4
Scheme		Global 100	Global 250	Global 800	Global and local 100
RMS	240.52	192.04	160.73	155.48	181.79
Corr.		0.37	0.64	0.81	0.78

Table 3: Summary of the permeability estimations for Cases 1 to 4.

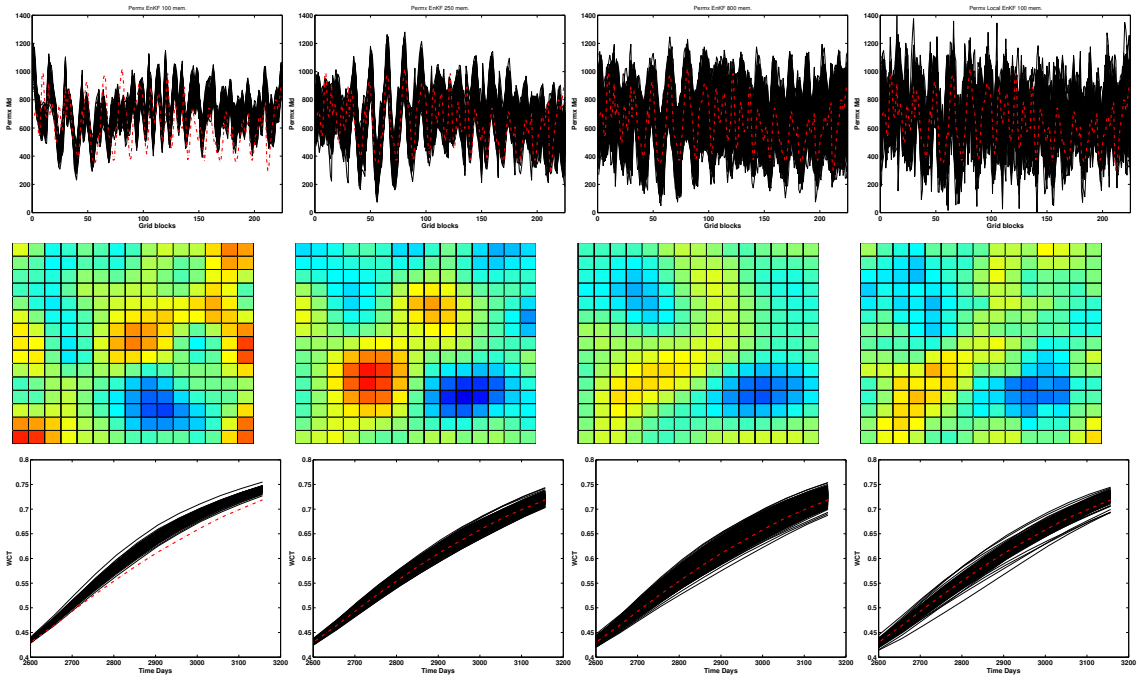
of Fig. 3 shows the final ensemble mean of permeability and the last row of Fig. 3 shows the ensemble predictions of WCT for the different cases.

From the first column in Fig. 3, where we have used a global scheme with 100 members, we clearly see that the ensemble spread does not cover the true field in many grid cells. The global scheme is underestimating the uncertainty, and we have obtained some strong updates of the permeability in parts of the reservoir, where we expected the correlation between the seismic difference data and the permeability to be small. In the final ensemble mean only parts of the true permeability field are recovered, and all the predictions of WCT from the ensemble is higher than the true solution.

The second column of Fig. 3 shows the EnKF results, using a global scheme of 250 members. The ensemble spread now covers the true field in almost all grid cells, but parts of the true field are still close to the boundary of the ensemble spread. The final ensemble mean is better correlated with the true field than in the case with 100 members, and a lower RMS value is obtained (Table 3). The WCT predictions for Case 2 seem to have a reasonable uncertainty spread. However, the permeability variation for many of the ensemble members are quite similar, and a too low uncertainty in the predictions might be achieved.

In column three and four in Fig. 3, we have respectively used a global scheme with 800 members and a combined global and local scheme with 100 members. As can be seen, the true field is covered by the ensemble spread in both cases, and a large uncertainty remains in the upper right part of the reservoir as expected. The final ensemble means characterise the true field well, and Table 3 show that strong correlations are obtained. From Table 3 we see that the lowest RMS is obtained for Case 3. An interesting observation is that Case 4 has a higher RMS than Cases 2 and 3. An explanation can be that the local scheme has introduced stronger local updates, and the ensemble spread for Case 4 shows that some members have taken both lower and higher permeability values than the global schemes. The final ensemble mean for Case 4 is also not as smooth as the final ensemble means obtained from the other cases. The ensemble predictions of WCT for Cases 3 and 4 cover the true solution fairly well. The predicted uncertainties increase faster with time, and are larger than for the predictions based on Cases 1 and 2.

The results of the above experiments show that when the amount of data is large, there are problems related to the low rank representation of the model variable covariance matrix P_ψ, and that a large ensemble size is needed to avoid these. This problem becomes more pronounced if the measurement error is decreased. The proposed global and local scheme seems to obtain a better uncertainty estimation, when a fairly small local region is used, and it acts more as a



Case 1 Case 2 Case 3 Case 4

Figure 3: Top row shows the ensemble of permeability in all grid cells. Middle row shows the final ensemble mean permeability, where the colour scale is equal to the reference plot. Last row shows the WCT predictions from year 7 and 1.5 years forward. The red dashed line shows the true solution.

global scheme with a large ensemble.

Synthetic 3D example

In this simulation study a synthetic 3D model of size $30 \times 30 \times 5$ is used to show the impact of introducing an uncertainty connected to the clay ratio in the rock physics model. The aim is also to illustrate the effects of using an incorrect seismic difference measurement error in the EnKF scheme. The production plan and the ensemble statistics are equal to the 2D model, where the wells now are penetrated in all layers, and the variogram range in the z-direction is equal to 2 grid cells. The local region stencil is set to be $7 \times 7 \times 1$ in this study. Three Poisson's ratio difference datasets, three acoustic impedance difference datasets, and well data every 3 months are assimilated.

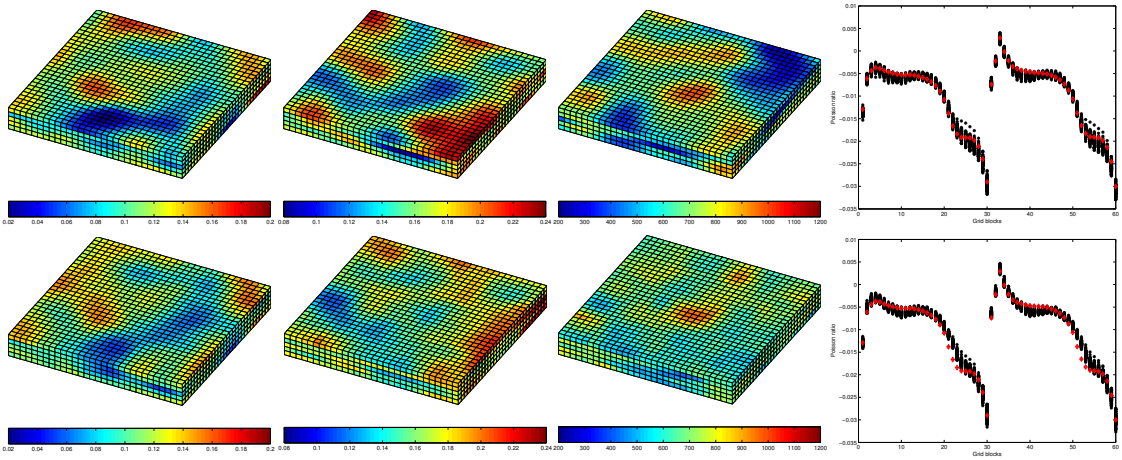
Four cases of model updating were performed. In Case 5 clay ratio, porosity and permeability are updated using the combined global and local scheme with 100 members. Case 6 is similar to Case 5, but uses a constant clay ratio value equal to 0.11. Case 7 is similar to Case 5, but uses a diagonal seismic measurement error covariance matrix. Case 8 uses a global scheme with 250 members.

The reference fields and the final ensemble mean fields for Case 5 are plotted in the three first columns of Fig. 4. The plots show that the combined global and local scheme obtain results that are in good agreement with the true model, and this can also be seen from the RMS and correlation values in Table 4. Especially the estimation of the clay ratio and the porosity are better when the global/local scheme is used. The reason for this is most probably that the rock physics parameters in this case have a stronger correlation to the seismic, than the permeability.

Notice that the porosity RMS value is higher than the initial RMS for Case 6. This is caused by an over- and underestimation of the porosity values in parts of the reservoir during the assimilation, and seems to occur because a significant information in the seismic is related

Case no.	Scheme	Clay ratio		Porosity		Permeability	
		RMS	Corr.	RMS	Corr.	RMS	Corr.
Initial		0.041		0.041		222.83	
Case 5	Local 100	0.020	0.82	0.026	0.71	190.55	0.29
Case 6	Local 100, clay ratio=0.11	0.041	0.0	0.048	0.45	228.97	0.03
Case 7	Local 100, P_{ϵ_s} = diagonal	0.029	0.67	0.036	0.61	239.57	0.07
Case 8	Global 250	0.025	0.66	0.031	0.52	225.81	0.20

Table 4: Summary of experiments from the synthetic 3D model.



(a) Clay ratio (b) Porosity (c) Permeability (d) Poisson's ratio

Figure 4: Columns (a,b,c) show the true static fields (top) and the EnKF fields for Case 5 (bottom). Column (d) show the seismic difference ensemble predictions (black stars), 0.5 year after the last assimilation step, for Case 5 (top) and Case 6 (bottom), for the first 60 cells in layer 1. The red diamonds show the true solution.

to a variation in the clay ratio. The correlation value obtained for the porosity is much lower than for Case 5, and this is due to the fact that the signature from the true clay ratio is found in the update of the porosity. The use of a constant clay ratio value also affects the estimation of the permeability, and non-satisfactory RMS and correlation values are obtained. In the last column in Fig. 4 the Poisson's ratio difference predictions for Cases 5 and 6 are plotted. As can be seen, the seismic ensemble predictions for Case 5 cover the true model much better than for Case 6, and clearly, including the lithology in the update improves the prediction of the elastic parameters.

To investigate the importance of using a correct seismic measurement error covariance matrix, we used in Case 7 a diagonal P_{ϵ_s} instead of the correct spherical covariance matrix. The RMS and correlation values in Table 4 show that the EnKF performance is then significantly reduced, and the results indicate the necessity of working with a proper P_{ϵ_s}.

Case 8 is performed as a comparison run with the proposed global/local scheme. The RMS and correlation results for Cases 8 and 5 in Table 4 show that Case 5 has obtained a better estimation result.

Discussion and conclusions

A combined global and local EnKF scheme for the estimation of lithology, porosity and permeability by an assimilation of production data and seismic difference data has been presented.

The benefit of using a local scheme is that the ratio between the dimension of the model state space and the ensemble size is smaller than for a standard global analysis in the EnKF

update, and this allows for a larger flexibility to obtain different model solutions, since the analysis will use a different combination of ensemble members for each grid point. The local scheme might also be more advantageous when many observations, especially with long range correlated errors, are assimilated.

A potential disadvantage of the local scheme is that non-physical modes may occur in the analysis fields, because the updates are performed independently in each local region. Especially when observations with a fairly high white noise are assimilated, this is a concern. It is then required that a large influence region is used in the local update to preserve the smoothness of the analysis fields.

It is demonstrated that taking into account uncertainty in lithology by updating the clay ratio may be important. A proper representation of the seismic data error model is also shown to be of importance during the assimilation.

Large amount of data in a global scheme with few members may cause a too large reduction in the uncertainty of the model variables. Here, the combined global and local scheme gave promising results regarding the uncertainty estimation, even with few members applied, and should be investigated further.

Acknowledgements

We thank Statoil and Norsk Hydro for financial support of this research and Schlumberger for providing the reservoir simulation software.

References

Dræge, A., Johansen, T. A., Brevik, I. and Dræge, C. T. [2006]. A strategy for modelling diagenetic evolution of seismic properties in sandstones. Accepted for publication in Petroleum Geoscience.

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

Johansen, T. A., Drottning, Å., Lecomte, I. and Gjøystdal, H. [2002]. An approach to combine rock physics and seismic modelling of fluid substitution effects. Geophysical Prospecting, 50, 119–137.

Ott, E., Hunt, B. R., Szunyogh, I., Zmin, A. V., Kostelich, E. J., Corazza, M., Kalnay, E., Patil, D. J. and Yorke, J. A. [2004]. A local ensemble Kalman filter for atmospheric data assmilation. Tellus, 56A, 415–428.

Sam, M. S. and Andrea, M. [2001]. The effect of clay distribution on the elastic properties of sandstones. Geophysical Prospecting, 49, 128–150.

Skjervheim, J. A., Aanonsen, S. I. and Evensen, G. [2006]. Ensemble Kalman filter with time difference data. Submitted to Computational Geoscience.

Szunyogh, I., Kostelich, E. J., Gyarmati, G., Patil, D. J., Hunt, B. R., Kalnay, E., Ott, E. and Yorke, J. A. [2005]. Assesing a Local Ensemble Kalman Filter: Perfect Model Experiments with the NCEP Global Model. Tellus, 57A, 528–545.

Paper E

Estimating Lithology and other Reservoir Properties from 4D Seismic Data Using the Ensemble Kalman Filter *

* To be submitted to the *Journal of Petroleum Science and Engineering*.

Estimating Lithology and other Reservoir Properties from 4D Seismic Data Using the Ensemble Kalman Filter

Jan-Arild Skjervheim [a,*], Sigurd Ivar Aanonsen [a] Bent Ole Ruud [a] Geir Evensen [b] Tor Arne Johansen [c]

[a] *Centre for Integrated Petroleum Research (CIPR), University of Bergen, Norway*
[b] *Hydro Research Centre, Bergen, Norway*
[c] *Department of Earth Science, University of Bergen*

Abstract

Traditionally in reservoir models the Ensemble Kalman filter (EnKF) method has been used to estimate permeability and porosity. An improvement of the reservoir characterisation can be obtained by introducing an uncertainty connected to parameters in the rock physics model. In this paper we show that when seismic data are available also the lithology can be estimated, which is coupled to the effective bulk modulus via the rock physics model. Incorporation of inverted seismic data in the EnKF introduces a large amount of data in the assimilation step. Thus, to improve the results a methodology based on a combination of a global and a local analysis scheme is proposed. The global and the local analysis are used to assimilate the production data, the inverted 3D seismic and time-lapse seismic data, where the local scheme assumes that only seismic data within a certain distance from a state variable will impact the analysis of this particular state variable. The technique is applied to synthetic 2D and 3D reservoir models, where effects of using local versus global analysis schemes on different inverted seismic data, such as acoustic impedance and Poisson's ratio, are investigated. Other evaluated factors are the effects of using different stencil sizes in the local scheme, and the sensitivity of the EnKF performance related to the seismic measurement error quantity. An improved sampling algorithm is also shown to influence the quality of the EnKF results.

Key words: EnKF, time lapse seismic, lithology, local analysis

1. Introduction

Conditioning petroleum reservoir simulation models to dynamic data, also called history matching, is a challenging task in reservoir characterisation. History matching can be considered as an ill-conditioned inverse problem, where the characterisations of the reservoir are represented in a high-dimensional space, and where the relation between the reservoir model and the reservoir performance data is represented by a highly non-linear fluid flow model which require a large computational effort.

The inverse problem can be formulated in a Bayesian context, e.g Tarantola (2005), which provides a stochastic framework for generating realizations of reservoir variables conditional to both dynamic and static data and consistent with a prior knowledge of the geology. Based on the Bayes rule the uncertainty in the description of the reservoir and thereby the description of future model predictions can be characterised by the posterior probability density function, and the aim is to perform a sampling of the posterior pdf, such that the properties of the posterior distribution can be assessed.

Different conditional sampling methods have been proposed to explore the posterior pdf. Markov Chain Monte Carlo (MCMC) sampling algorithms are used by Oliver et al. (1997) and Hegstad and Omre (2001) in uncertainty evaluation of reservoirs. In many applications samples representing the uncertainty are obtained by reformulating the history matching problem as a minimization problem, where the objective function takes into account both the data mismatch and the prior term. The Randomized Maximum Likelihood (RML) is such a method, and for additional discussion see e.g Liu and Oliver (2003) and Gao and Reynolds (2005).

In this paper we describe an estimation of both static variables (permeability, porosity and lithology) and dy-

* Corresponding author.
Email addresses: jan-arild.skjervheim@cipr.uib.no (Jan-Arild Skjervheim), sigurd.aanonsen@cipr.uib.no (Sigurd Ivar Aanonsen), bent.ruud@cipr.uib.no (Bent Ole Ruud), geir.evensen@hydro.com (Geir Evensen), TorArne.Johansen@geo.uib.no (Tor Arne Johansen).

namic variables (pressures and saturations) by an ensemble Kalman filter (EnKF) method. The update is based on the information contained in the assimilated reservoir measurements, such as water cut (WCT), flowing bottom hole pressure (BHP), gas-oil ratio (GOR), 3D seismic data and time-lapse seismic data.

The EnKF, developed by Evensen (1994), is a Monte Carlo type sequential Bayesian inversion method which assimilate observations as they becomes available. The EnKF is related to the Kalman filter (KF), see Kalman (1960), which provides the optimal state estimate of a linear dynamical system, when the error statics are assumed Gaussian. The main difference between the Kalman Filter and the EnKF is that an ensemble of non-linear forecasts are used to estimate the error covariance. At the analysis step, each member is then updated such that the EnKF generate an ensemble of approximate random samples from the posterior pdf.

The analysis schemes used in the simulation studies are based on the EnKF formulation with perturbed observations (see Burgers et al. (1998) and Houtekamer and Mitchell (1998)), and some recent studies and publications discussing the combined parameter and state estimation problem within petroleum reservoir applications using EnKF can be found in Nævdal et al. (2003); Gu and Oliver (2004); Skjervheim et al. (2005); Wen and Chen (2005); Haugen et al. (2006); Evensen (2006).

The use of seismic data may introduce a large number of data points in the assimilation step, and problems related to an underestimation of the uncertainty in the low rank covariance matrix representation of the static and dynamic variables may occur when few ensemble members are used during the assimilation (see e.g Anderson (2001)). A reason which may affect the estimation of the variance is that a finite ensemble size will introduce spurious long-range correlations and cause a reduction of the estimated variance. Another reason is that the rank of the ensemble covariance matrix is limited by the ensemble size, and the variance of the system can only be represented within the space spanned by the ensemble.

Anderson (2001) and Hamill et al. (2001) pointed out that spurious correlations between uncorrelated variables would be suppressed by using localization (local analysis). A localization may also improve the underestimation of the variance since it involves less data and the ratio between the dimension of the model state space and the ensemble size will become smaller.

To obtain an improvement regarding the underestimation problem we propose a combination of a global and local EnKF scheme in our reservoir study. Assimilation of production data is performed with a global scheme, while the seismic data are assimilated with a local scheme, where we assume that only seismic data within a certain distance from a state variable will impact the analysis in this state variable. The local scheme updates each grid cell independently accounting for the seismic observations only in the local region surrounding the cell. The local schemes can be implemented in parallel, and each local update will involve relatively small matrices in the assimilation, such that the method inverts many small matrices instead of one large. Thus, with the local scheme, large amounts of seismic data can be assimilated at a low computational cost.

Recently, several formulations and applications of the local EnKF have been reported, mostly within ocean and atmospheric sciences. Haugen and Evensen (2002) applied a local analysis to assimilate sea-surface temperature and sea-level anomalies of a Indian ocean model, and a general formulation of the local analysis was presented by Evensen (2003).

Local analysis is also applied in the data assimilation of Topaz (topaz.nersec.no), which is an operational ocean prediction system for the Atlantic and Arctic oceans. In atmospheric data assimilations Ott et al. (2004) introduced a local EnKF approach, where the scheme is based on an ensemble square root filter formulation (see e.g Tippett et al. (2003)). The accuracy and computational efficiency of the local EnKF was investigated on a operational numerical weather prediction model by Szunyogh et al. (2005).

The seismic reservoir characterisation can be improved by introducing an uncertainty connected to parameters in the rock physics model. Lithology is such a variable of interest, which for siliciclastic rocks can be quantified by the clay ratio (see e.g., Dræge et al. (2006)). Lithology may have a strong impact on the seismic signal, and an estimation of the clay ratio will be of importance to achieve satisfactory predictions of the elastic parameters. For the purpose of this numerical experiment we have chosen to use the so-called dispersed clay model, i.e., the clay is considered as dispersed in the pore volume of the sandstone and not as a part of the load-bearing structure. This approach allows the clay to be easily included by any rock physics model.

This paper proceeds as follows. In Section 2 we define the notation, describe the global/local EnKF method and discuss the complexity of the algorithm. Then in Section 3 we present the rock physics model used to compute the inverted seismic data. Section 4 presents a 2D synthetic example and examines the impact of using a global versus the local scheme. In Section 5 we discuss the simulation results obtained by introducing clay ratio as an lithology parameter. Finally in Section 6 some conclusions are given.

2. Combined global and local EnKF scheme

The objective of an ensemble Kalman filter is to simulate samples from the posterior distribution of interest, which from Bayesian theory is given by

$$g(\psi|d) \propto g(d|\psi)g(\psi). \tag{1}$$

Here $g(\psi)$ is the joint prior distribution for the parameters and the model state, $\psi \in \Re^q$, and $g(d|\psi)$ is the likelihood pdf for the observations, $d \in \Re^m$, see Tarantola (2005).

In a reservoir history matching process both production data and seismic data becomes available at different time steps during the production history. Thus, a sequential

processing will be of importance, and in Evensen and van Leeuwen (2000) it was shown that the general smoother in Eq. (1) could be formulated as a sequential method.

When working with time-lapse seismic data, the measurements at a given time may depend on the model state at more than one time instant. Such seismic data may be a result of a 4D inversion, where the seismic difference between two surveys are inverted to differences in elastic parameters. The individual surveys of the elastic parameters may then not be available. In Skjervheim et al. (2006) it was shown that the posterior pdf then included both a filter and a smoother solution, and that the EnKF simulation involves a combination of the ensemble Kalman filter and the ensemble Kalman smoother. Further, it was shown that this could be handled in a similar framework as in the traditional EnKF, and for simplicity we here present the equations on a filter form.

Let now $\psi_j = \psi(x, t_j) \in \Re^{q_j}$ and $d_j \in \Re^{m_j}$ represent the model variables and the observations at time t_j, where the measurements d_j at a particular time t_j depend on ψ_j. The likelihood distribution is then expressed by $g(d_j|\psi_j)$. Let $d_j = [d_{p_j}, d_{s_j}]^T$, where d_{p_j} denote production observations and d_{s_j} denote seismic observations. By using the prior distribution, $g(\psi_j|d_{j-1}, \ldots, d_1)$, obtained from the evolution equation in the state space model, and the likelihood distribution, $g(d_j|\psi_j)$, the posterior distribution can be represented by a sequential Bayesian formulation

$$g(\psi_j|d_j, \ldots, d_1) \propto g(d_j|\psi_j)g(\psi_j|d_{j-1}, \ldots, d_1), \quad (2)$$

where a first order Markov process and independence between the data available at different times are assumed (see e.g Evensen (2006)).

If we augment the model state with the model prediction of measurements, see Evensen (2003), the observation equation in the nonlinear state space model can be defined as

$$d_j = H_j\psi_j + \epsilon_j, \quad (3)$$

where $H_j = [H_{p_j}, H_{s_j}]^T$ is a linear observation operator, H_{p_j} relates the model state ψ_j to the production data, and H_{s_j} to the seismic measurements. Both operators allow for measurement errors, ϵ_j, which are assumed to be Gaussian:

$$\epsilon_j = \begin{bmatrix} \epsilon_{p_j} \\ \epsilon_{s_j} \end{bmatrix} \sim \text{Gauss}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} P_{\epsilon_{p_j}} & 0 \\ 0 & P_{\epsilon_{s_j}} \end{bmatrix} \right). \quad (4)$$

The terms ϵ_{p_j} and ϵ_{s_j} are respectively production and seismic measurement errors with covariance matrices $P_{\epsilon_{p_j}}$ and $P_{\epsilon_{s_j}}$. Here we have assumed independence between the production and the seismic measurement errors.

To sample from the posterior pdf, $g(\psi_j|d_j, \ldots, d_1)$, a combination of a global and a local ensemble Kalman filter is proposed. The ensemble mean and the ensemble spread of the model variables at any given time reflect both an estimate of the true state of the reservoir and an uncertainty of that estimate. Each member of the ensemble is evolved independently forward in time, but when new observations from the reservoir are available the first and second order moments from the ensemble are used to generate the new reservoir state estimate and the reduced uncertainty introduced by the observations.

Let the ensemble matrix holding the members $\psi_j \in \Re^{q_j}$ be defined by

$$\Psi_j = \Psi(\mathbf{x}, t_j) = (\psi_j^1, \ldots, \psi_j^N) \in \Re^{q_j \times N}, \quad (5)$$

where N is the number of ensemble members. All the equations below are evaluated at time t_j, and to simplify the notation we will now on drop the time subscript j.

The ensemble perturbation matrix, $\Psi' \in \Re^{q \times N}$, can be expressed by

$$\Psi' = \Psi - \bar{\Psi} = \Psi(I - 1_N), \quad (6)$$

where $\bar{\Psi}$ is the ensemble mean matrix and 1_N is the $N \times N$ matrix where each element is equal to 1/N. From the ensemble perturbations matrix the error covariance matrix of the ensemble, $P_\psi \in \Re^{q \times q}$, can be estimated as

$$P_\psi = \frac{\Psi'(\Psi')^T}{N-1}. \quad (7)$$

The rank of the product in Eq. (7) cannot exceed the smallest rank of the ensemble perturbation matrix Ψ'. Hence, the rank of the covariance matrix is given by

$$\text{rank}(P_\psi) \leq \min(q, N-1). \quad (8)$$

Note that $N-1$ appear because of the mean subtraction in Eq. (6).

Given vectors of production data, $d_p \in \Re^{m_p}$, and seismic data, $d_s \in \Re^{m_s}$, we can define the N vectors of perturbed measurements as

$$\begin{aligned} d_p^k &= d_p + \epsilon_p^k \\ d_s^k &= d_s + \epsilon_s^k \end{aligned} \quad \forall \ k = 1:N \ . \quad (9)$$

Define now ensemble matrices for the measurement perturbations and the perturbed measurements

$$E_p = (\epsilon_p^1, \ldots, \epsilon_p^N) \in \Re^{m_p \times N} \quad E_s = (\epsilon_s^1, \ldots, \epsilon_s^N) \in \Re^{m_s \times N} \quad (10)$$

$$D_p = (d_p^1, \ldots, d_p^N) \in \Re^{m_p \times N} \quad D_s = (d_s^1, \ldots, d_s^N) \in \Re^{m_s \times N}. \quad (11)$$

The ensemble representation of the production and the seismic measurement error covariance matrices, denoted respectively as, $P_{\epsilon_p}^e \in \Re^{m_p \times m_p}$ and $P_{\epsilon_s}^e \in \Re^{m_s \times m_s}$, are then given by

$$\begin{aligned} P_{\epsilon_p}^e &= \frac{E_p E_p^T}{N-1} \\ P_{\epsilon_s}^e &= \frac{E_s E_s^T}{N-1}. \end{aligned} \quad (12)$$

Using the equations above the ensemble representation of the analysis step conditioned only to the production data, $\Psi_p^a \in \Re^{q \times N}$, can be written as (see e.g Evensen (2006)

$$\Psi_p^a = \Psi^f + P_\psi^f H_p^T (H_p P_\psi^f H_p^T + P_{\epsilon_p})^{-1}(D_p - H_p\Psi^f), \quad (13)$$

where $\Psi^f \in \Re^{q \times N}$ is the forecast ensemble matrix. The analysis step Ψ_p^a is now used as input to the local EnKF scheme, where seismic data are assimilated.

In the local scheme, only data from a local region is applied when updating the state vector in a given grid cell. Hence, the the seismic data influencing the analysis for grid cell, l, is denoted $d_{s_l} \in \Re^{m_{s_l}}$, and extracted by a matrix L_l which includes only zeros and ones and picks the seismic data within this prescribed local region. Correspondingly, the local observation equation can be written as

$$d_{s_l} = H_{s_l}\psi_l + \epsilon_{s_l} \qquad \epsilon_{s_l} \sim \text{Gauss}(0, P_{\epsilon_{s_l}}), \tag{14}$$

where $H_{s_l} = L_l H_s$, $\epsilon_{s_l} = L_l \epsilon_s$, and the error covariance matrix is given by $P_{\epsilon_{s_l}} = L_l P_{\epsilon_s} L_l^T$.

Defining the local representation of the seismic measurements of the ensemble perturbations, $S_{s_l} \in \Re^{m_{s_l} \times N}$, as

$$S_{s_l} = H_{s_l}\Psi_p^{a'} = L_l H_s(\Psi_p^a - \bar{\Psi}_p^a), \tag{15}$$

the ensemble of innovation vectors, $D'_{s_l} \in \Re^{m_{s_l} \times N}$, as

$$D'_{s_l} = D_{s_l} - H_{s_l}\Psi_p^{a'} = L_l(D_s - H_s\Psi_p^{a'}), \tag{16}$$

and the matrix $C_{s_l} \in \Re^{m_{s_l} \times m_{s_l}}$ as

$$C_{s_l} = S_{s_l}S_{s_l}^T + (N-1)P_{\epsilon_{s_l}}, \tag{17}$$

the local analysis equation at grid point l, i.e., $\Psi_l^a \in \Re^{q_l \times N}$, see Evensen (2006), becomes

$$\begin{aligned}
\Psi_l^a &= \Psi_{p,l}^a + \Psi_{p,l}^{a'}\Psi_{p,l}^{a'T}H_s^T \\
&\quad \cdot (H_s\Psi_{p,l}^{a'}\Psi_{p,l}^{a'T}H_s^T + (N-1)P_{\epsilon_s})D'_{s_l} \\
&= \Psi_{p,l}^a + \Psi_{p,l}^a(I - 1_N)S_{s_l}^T C_{s_l}^{-1}D'_{s_l} \\
&= \Psi_{p,l}^a(I + (I - 1_N)S_{s_l}^T C_{s_l}^{-1}D'_{s_l}) \\
&= \Psi_{p,l}^a(I + S_{s_l}^T C_{s_l}^{-1}D'_{s_l}) \\
&= \Psi_{p,l}^a X_l,
\end{aligned} \tag{18}$$

where the ensemble matrix conditioned to production data at the same grid point, $\Psi_{p,l}^a$, is used as the forecast. The matrix $X_l \in \Re^{N \times N}$ can be written as

$$X_l = I + S_{s_l}^T C_{s_l}^{-1}D'_{s_l}, \tag{19}$$

where we have used that $1_N S_{s_l}^T \equiv 0$.

By solving the local analysis equation (18) for each grid cell in the reservoir, we obtain the global analysis ensemble conditioned to both production and seismic data.

2.1. *Computational cost*

When computing the analysis ensemble Ψ_l^a at a seismic assimilation step, the corresponding local region includes q_l model variables together with m_{s_l} seismic observations and an ensemble size of N members. The computations are repeated N_g times, where N_g denotes the total number of local regions. In reservoir applications N_g is typical equal to the number of active grid cells in the reservoir.

For each local region the calculation of the inverse of the matrix C_{s_l} defined in Eq. (17) is required. The inverse of C_{s_l} is found from an eigenvalue decomposition and is given as

$$C_{s_l}^{-1} = Z_{s_l}\Lambda_{s_l}^{-1}Z_{s_l}^T, \tag{20}$$

where $Z_{s_l} \in \Re^{m_{s_l} \times m_{s_l}}$ contains the eigenvectors and the diagonal matrix $\Lambda_{s_l} \in \Re^{m_{s_l} \times m_{s_l}}$ holds the eigenvalues. The inverse in Eq. (20) will be the pseudo inverse if the matrix is singular. If C_{s_l} is a matrix of full-rank a computational cost of order $\mathcal{O}(Nm_{s_l}^2)$ is required to form it. The eigenvalue decomposition requires $\mathcal{O}(m_{s_l}^3)$ operations.

In our applications we normally have $m_{s_l} < N$ in the local scheme. Thus, based on the construction of the inverse matrix $C_{s_l}^{-1}$ the dominant cost is related to the forming of X_l which is $\mathcal{O}(m_{s_l}N^2)$ and the final computation in Eq. (18) which is of order $\mathcal{O}(q_l N^2)$. The main cost for a sequential process is then $\mathcal{O}(q_l N_g N^2)$ since $q_l > m_{s_l}$. Notice that the data assimilation on each local region is performed independently in the local scheme. Hence each assimilation can be implemented in parallel, which may reduce the computational time significantly.

For a global analysis of the seismic data q model variables and m_s seismic observations are included in the equations. In this case we have $m_s \gg N$, and the eigenvalue decomposition of the matrix C_s will be the dominating part of the computational time. Hence, an alternative efficient algorithm is used, which reduces the factorization of the $m_s \times m_s$ matrix C_s to a factorization of a $N \times N$ matrix (see e.g Evensen (2004)). The dominant cost of forming X is of order $\mathcal{O}(m_s^2 N)$, and the final computation will be of order $\mathcal{O}(qN^2)$. In a real field case we typically have $m_s = N_g$ and $q = 10N_g$ and the main cost for the global analysis require then $\mathcal{O}(N_g^2 N)$ operations. Comparing the computational cost for the local scheme, $\mathcal{O}(q_l N_g N^2)$, and the global scheme, $\mathcal{O}(N_g^2 N)$, we see that as N_g increase the global scheme will become the most expensive scheme to compute.

A reduction in the computational cost for the global analysis in the case when $m_s \gg N$ can be obtained by using a low rank representation of the seismic measurement error covariance matrix, $P_{\epsilon_s}^e$, given in Eq. (12). The dominant cost is then of order $\mathcal{O}(N_g N^2)$. If $m_{s_l} \gg N$ this efficient algorithm can also be used in the local scheme. A more detailed description of the efficient EnKF schemes when a large amount of data are introduced in the analysis can be found in Evensen (2004) and Skjervheim et al. (2006).

3. Rock physics model

Several methods have been proposed to model the effect of clay in sandstones, e.g Dræge et al. (2006); Sam and Andrea (2001). The main difference between the various approaches is whether the clay is included as a part of the load-bearing structure or not. In order to choose between alternative models one would in practical cases use geological information, or base the choice on previous experience with the reservoir in question. In this study we have used a dispersed clay model, i.e., the clay is considered as dispersed in the pore volume of the sandstone and not as a part of the load-bearing structure. In such a model the clay ratio should not be too high because when the pore-filling

clay exceeds about 40% of the porespace, it will become pore-bridging, see Dræge et al. (2006). If this happens, the additional amount of clay should be included as part of the load-bearing matrix. In such cases we compute new effective properties of the matrix material by the Voigt-Reuss-Hill average, see Mavko et al. (2003). The rock physics model used here to compute the effective properties of a porous, fluid filled medium is the Kuster-Toksöz model as described by Johansen et al. (2002). The Kuster-Toksöz method is an effective medium theory in which the fluid filled pores are modelled as ellipsoidal inclusions in a homogeneous background medium. Thus, it is not necessary to use Gassmann's relation in order to calculate the effect of the fluids. The resulting moduli are slightly higher than if the moduli were calculated for empty pores (dry rock) and the fluid effect were accounted for by Gassmann's relation. This is because the Kuster-Toksöz method assumes that the pore are unconnected while Gassmann assumed connected pores. When clay partially fills up the pore-space, the assumption of unconnected pores may be preferable.

In order to compute the elastic properties for a rock with porosity ϕ (volume fraction of fluids) and clay ratio c (the volume ratio of clay to the total volume of solids), the pore-space must be increased to accommodate the volume of the clay. The pore-space volume ϕ_p which is used to compute the properties of the load-bearing structure becomes

$$\phi_p = \phi + (1-\phi)c. \tag{21}$$

However, the pore-space volume should stay below the critical porosity ϕ_c of sandstones, typically 40%, i.e.

$$\phi_p < \phi_c. \tag{22}$$

If this limit is exceeded, the excessive clay is again included in the load-bearing matrix material by the Voigt-Reuss-Hill

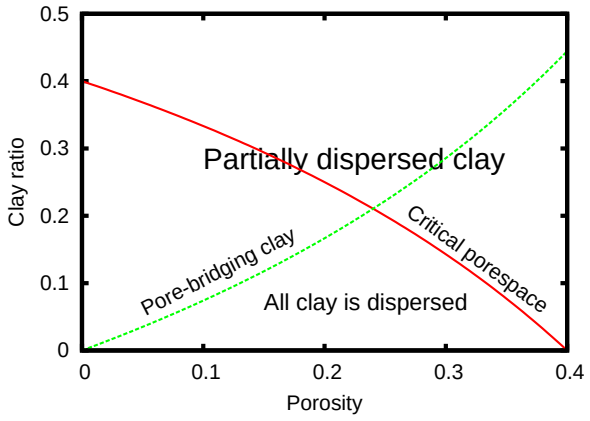


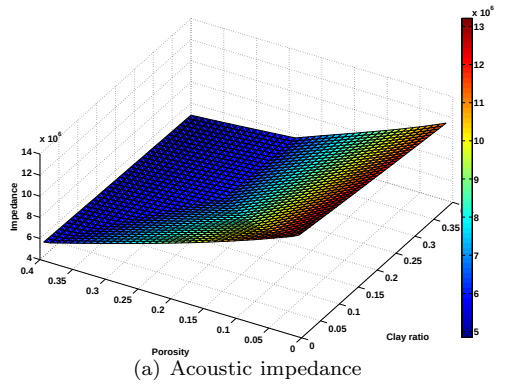
Fig. 1. The clay will be fully or partially dispersed in the pore-space depending on the combination of porosity and clay ratio. The limits shown here were computed from equations (24) and (25) for a critical pore-space porosity $\phi_c = 0.4$ and a pore-bridging limit $c_p = 0.4$. Both requirements must be fulfilled if all the clay should be dispersed, so it is only in the lower triangular area that the clay is fully dispersed. For other combinations of porosity and clay ratio, the remaining clay will be included in the matrix.

Table 1
Pore structure model. The relative concentration for a given aspect ratio is the fraction of the total porosity. Hence, the relative concentration for all aspect ratios sum up to one.

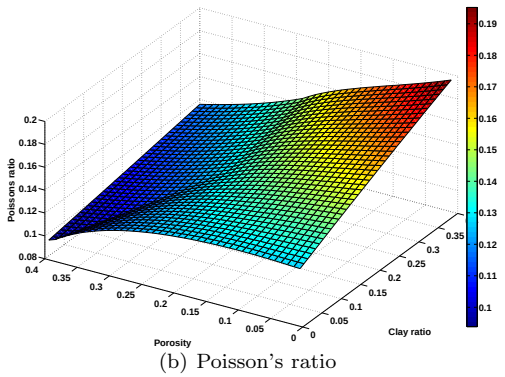
Aspect ratio	1	0.5	0.1	0.05	0.01
Relative concentration	0.40	0.38	0.16	0.05	0.01

Table 2
Fluid components used in the modelling, computed at 140 bar and 109 deg C.

	Bulk modulus (GPa)	Shear modulus (GPa)	Density (kg/m^3)
Brine	2.52	-	990
Oil	0.384	-	683
Gas	0.025	-	87.7
Sandstone matrix	30.0	30.0	2650
Clay	20.0	10.0	2600



(a) Acoustic impedance



(b) Poisson's ratio

Fig. 2. Acoustic impedance and Poisson's ratio as a function of porosity and clay ratio. Evaluated at a pressure of 140 bar and saturations: $S_w = 0.2$ and $S_o = 0.8$.

average. The limitation on c imposed by pore-bridging may be expressed as

$$\frac{(1-\phi)c}{\phi_p} < c_p. \tag{23}$$

where c_p is the limit for when the dispersed clay becomes pore-bridging. By inserting (21) into (22) and (23) and solving for c, we get the following two requirements for the dispersed clay:

$$c < \frac{\phi_c - \phi}{1-\phi} \tag{24}$$

$$c < \frac{c_p}{1-c_p}\frac{\phi}{1-\phi}. \tag{25}$$

These two requirements are plotted as functions of porosity (fluid content) in Figure 1.

The effective bulk modulus of the pore fluid, including the dispersed clay, is computed by the Reuss average.

The pore structure in the rock physics model is given in Table 1. The model contain both spherical and ellipsoidal inclusions defined by the aspect ratio, $\beta = b/a$, where b and a are the smallest and largest semi-axis, respectively. Five different aspect ratios are used in the model, where the relative concentration for each aspect ratio is given as a fraction of the total porosity. The material properties of the solid and the fluids are given in Table 2. Typical acoustic impedance and the Poisson's ratio, computed as functions of the porosity and clay ratio, are shown in Fig. 2. If this way of including the clay (dispersed) were compared to the more common practice of including all the clay in the load-bearing matrix, we would find that the dispersed clay method has a higher sensitivity, i.e., small quantities of clay have a larger effect on the elastic properties of the rock.

4. Synthetic 2D example

To evaluate the effects of using a global versus a local scheme when assimilating a large dataset of time-lapse seismic data, we consider a synthetic 2D problem with 15×15 grid cells. The reservoir consists of one water injector and one oil producer located respectively in the lower left and upper right corner. Injection of water starts after one year of production, allowing for production below bubble-point, and before a higher reservoir pressure is re-established, free gas will evolve around the producer.

The initial ensemble of permeability, porosity and clay ratio are assumed independent and sampled as Gaussian fields with constant mean and a gaussian variogram model.

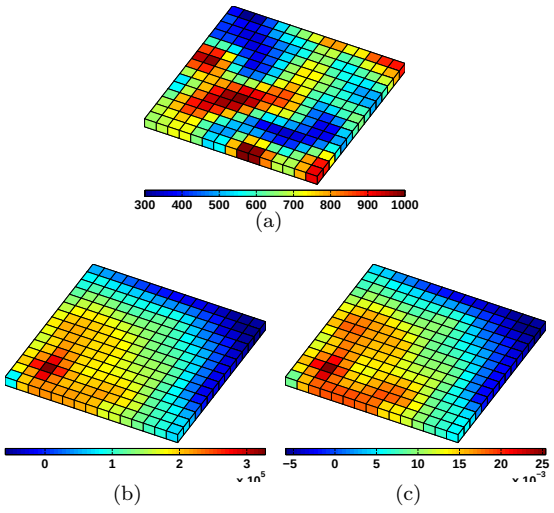


Fig. 3. (a): True permeability, (b) and (c): True acoustic impedance difference and Poisson's ratio difference at the last seismic assimilation step.

Table 3
Model data used as input in the EnKF.

	Mean	Standard deviation
Clay ratio	0.11	0.03
Porosity	0.17	0.03
Permeability (mD)	665	170

Table 4
Measurement errors for production data, standard deviations.

Water cut	Gas-oil ratio	Bottomhole pressure
(-)	(Sm^3/Sm^3)	(bar)
0.01	10 %	2.0

Table 5
Measurement errors for time-lapse seismic data, standard deviations.

	"a-cases"	"b-cases"
Acoustic impedance (kg/m^2s)	1.0E+04	5.0E+04
Poisson's ratio (-)	1.0E-03	5.0E-03

Table 6
Summary of the permeability estimations at the last assimilation step for Cases 1 to 8.

Cases	Scheme	"a-cases"		"b-cases"	
		Corr.	RMS	Corr.	RMS
Case 1	Global 100	0.04	200.1	-0.12	203.3
Case 2	Global 250	0.80	110.2	0.28	173.5
Case 3	Global 800	0.86	89.6	0.31	160.2
Case 4	Local 100, 3×3	0.78	112.6	0.13	177.3
Case 5	Local 100, 5×5	0.79	106.2	0.17	180.0
Case 6	Local 100, 7×7	0.69	121.9	0.14	189.1
Case 7	Global 100 Imp.			0.23	193.0
Case 8	Local 100 Imp., 5×5			0.40	154.7

The reference ("true") fields are generated from the same statistics. The variogram range is set to be 5 grid cells in x and y direction. The means and variances are summarised in Table 3.

The production data consist of observations of WCT, GOR and BHP. Time-lapse seismic data are acoustic impedance and Poisson's ratio computed from the rock physics model presented in Section 3. The measurement uncertainties for the production data are given in Table 4. In our study we have used two different measurement error levels for the time-lapse seismic data, denoted by "a-cases" and "b-cases", and the standard deviation values can be found in Table 5. The seismic measurement errors have a spherical variogram with same range as the static fields.

From the study of the rock physics model it was seen that the elastic parameters had different sensitivity to the parameters in the model. Thus, to take in to account the complementary information both the acoustic impedance and the Poisson's ratio differences are used simultaneously at a given seismic assimilation step. The measurement errors in the acoustic impedance and Poisson's ratio differences are assumed independent, and the assimilation of the elastic parameters is performed sequentially one dataset at a time. This approach is used to avoid that a poor scaling between the acoustic impedance and the Poisson's ratio may affect the results, see Evensen (2003). The validity of this ap-

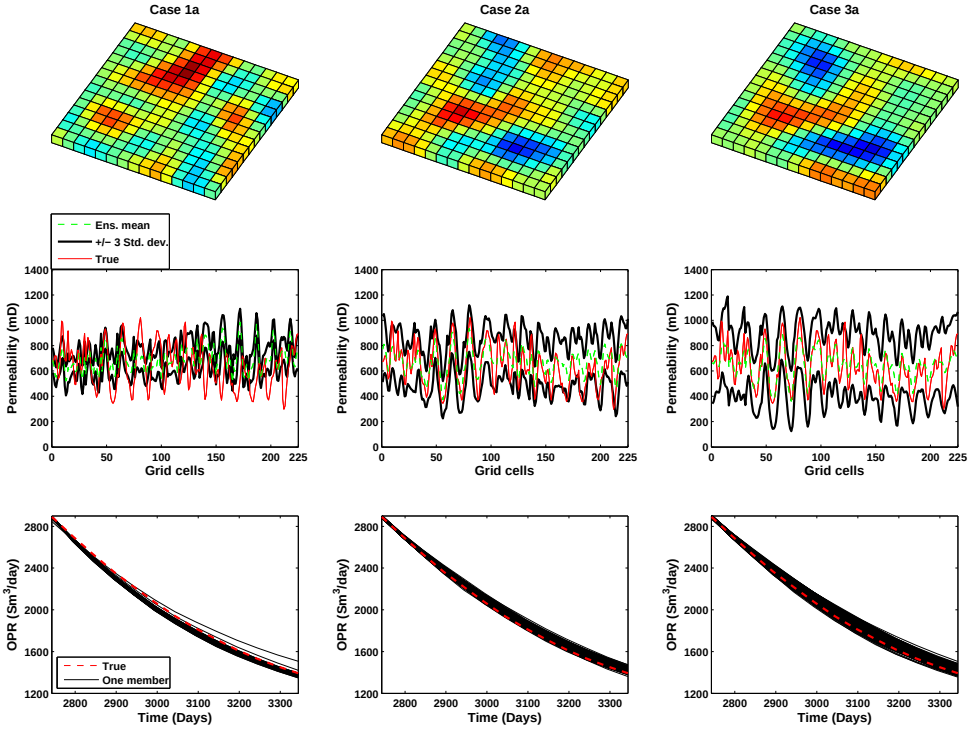


Fig. 4. EnKF results for Cases 1a-3a. The top row in (a) and (b) shows the final ensemble mean permeability, where the colour scale is equal to the reference plot. Middle row shows the ensemble mean plus/minus three standard deviation plotted together with the true field. Last row shows the OPR predictions from year 7.5 and 1.5 year forward, where the red dashed line shows the true solution.

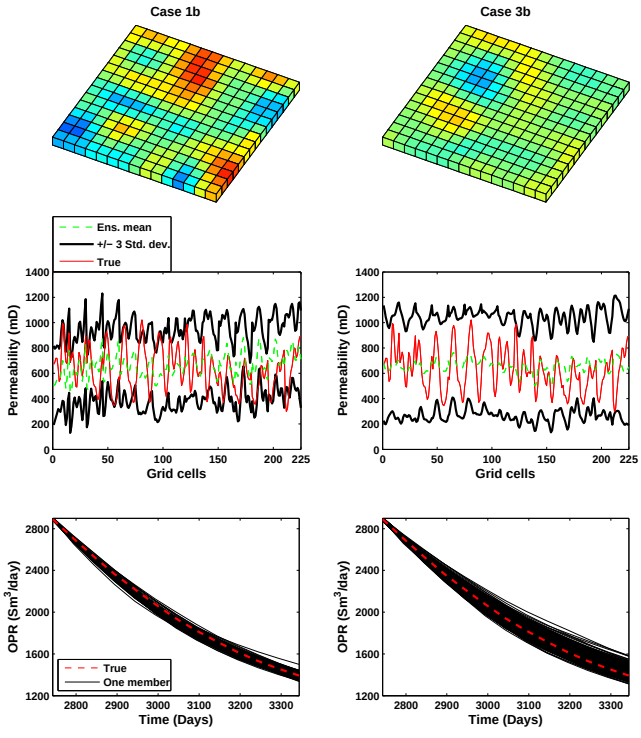


Fig. 5. EnKF results for Cases 1*b* and 3*b*. The explanation of the plots are given in Fig. 4.

proach can be found in Evensen and van Leeuwen (2000). The time-lapse seismic datasets are assimilated at years 1, 3.5, 6 and $7\frac{1}{4}$, and are the differences between the monitor surveys at these times and the base survey, computed at time zero. Production data are assimilated every 3 month.

The simulation time is 9 years, where 7.5 years are history and 1.5 years prediction. The reference permeability field is plotted in Fig. 3 together with the computed acoustic impedance and the Poisson's ratio differences datasets at the last seismic assimilation step.

The EnKF updates the parameters, which are permeability, porosity and clay ratio, and the state variables, which are pressure and saturation. The clay ratio only appears in the rock physics model, and is therefore updated only at the time steps where time-lapse seismic data are assimilated.

Eight cases of model updating were performed. In Cases 1, 2, and 3 only the global EnKF scheme is applied with respectively 100, 250 and 800 ensemble members. When using the local scheme it is important that the size of the local stencil is large enough to include all the information from the surrounding observations, which will significantly impact the analysis of state variable at the center of the stencil. On the other hand, an underestimation of the variance in the low rank ensemble covariance matrix may occur if the stencil is too large. Thus, in Cases 4, 5, and 6 we vary the size of the local region, using 3×3, 5×5 and 7×7 stencils, and keeping the ensemble size fixed at 100 members. Each of the six cases above have been run by the two different seismic measurement error levels, given in Table 5. In Cases 7 and 8 we have used an improved sampling algorithm presented by Evensen (2004), where a start en-

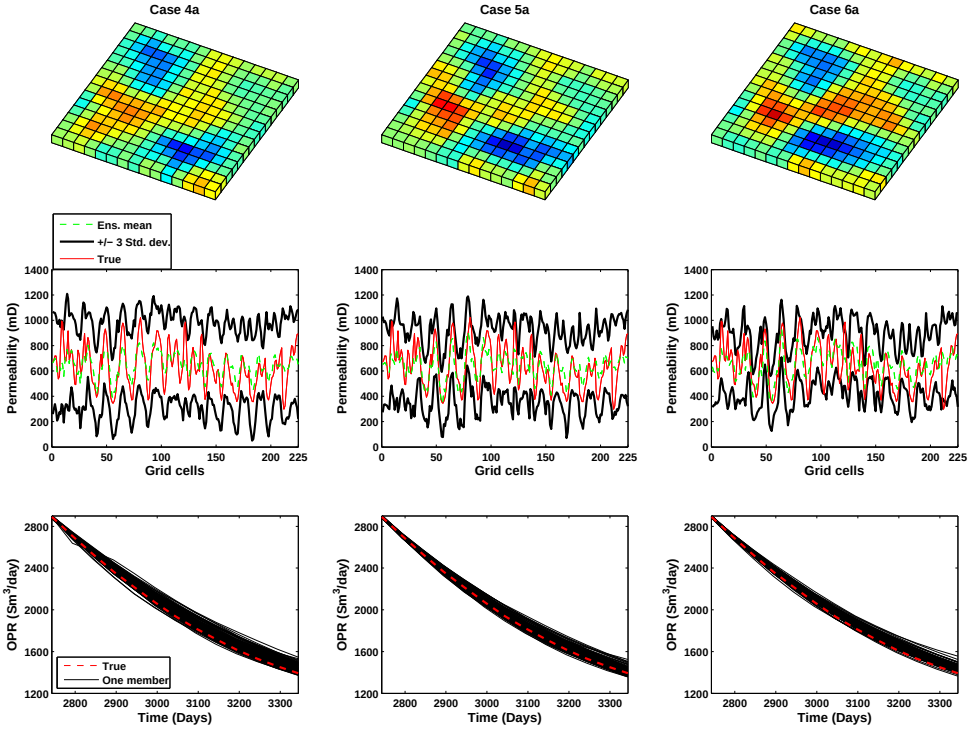


Fig. 6. EnKF results for Cases 4a-6a. The explanation of the plots are given in Fig. 4.

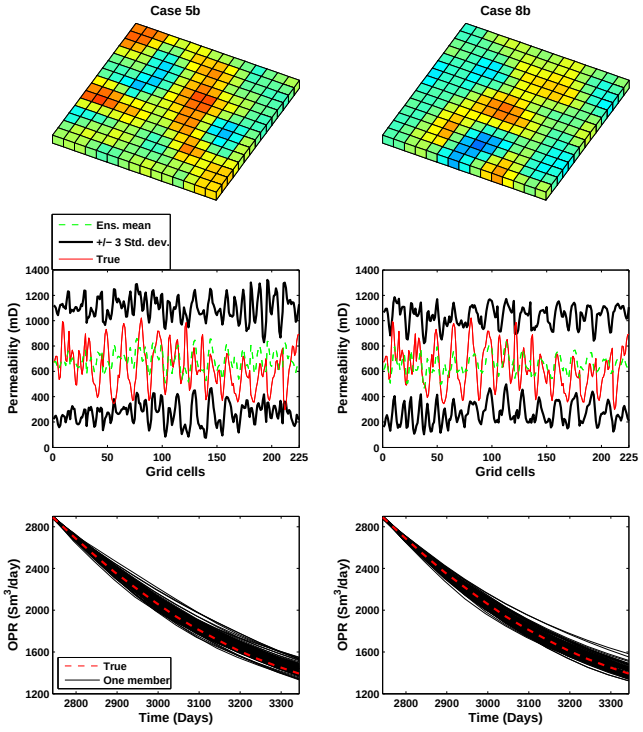


Fig. 7. EnKF results for Cases 5*b* and 8*b*. The explanation of the plots are given in Fig. 4.

semble of 800 members are used to generate the initial 100 member ensemble. Case 7 uses a global scheme and Case 8 uses a local scheme with a 5×5 stencil. Both cases use the largest seismic measurement error in Table 5.

To evaluate the quality of the EnKF solution, we have focused on the estimation of permeability. Most of the information affecting this parameter should be in the water flooded areas, propagating forward in time with the water front, and thus the correlation between the permeability and the time-lapse seismic data is expected to be stronger in the lower left part of the reservoir during the assimilation. The performance of the EnKF update is measured by the correlation between the true field and the ensemble mean, and by the root mean square (rms) error of the ensemble mean, J, defined for the parameter α as

$$J = \sqrt{\frac{1}{N_g}\sum_{j=1}^{N_g}\left(\frac{1}{N}\sum_{i=1}^{N}\alpha_j^i - \alpha_j^t\right)^2}, \qquad (26)$$

where α_j^i is the estimate of α in grid cell j in ensemble member i, α_j^t is the corresponding true value, and N_g is the number of grid cells. As a measure of how well the members in the ensemble represent the true solution, we will use the ensemble spread of permeability by the envelope of plus/minus three standard deviations. The rms and correlation results for the different cases are summarized in Table 6.

In Fig. 4 the experimental results for Cases 1*a*-3*a* are plotted. For the global scheme with 100 members, shown in the left column, only few parts of the true permeability field is recovered in the final ensemble mean. The ensemble spread is small, and the obtained low rank covariance matrix of the permeability is clearly underestimating the uncertainty. Large rms and small correlation values are

achieved in Table 6, and the predictions of OPR from the ensemble are all higher than the true solution except from some few members.

The middle and right column show a significantly improvement in the EnKF performance, when the number of ensemble members is increased. Especially, the global scheme with 800 members obtain a final ensemble mean that has a close resemblance to the true field. The spread of the ensemble covers the true field except a few grid cells, and a rather large uncertainty remains in grid cells close to the boundary of the reservoir as expected since the water front has not reached this area. The ensemble predictions of OPR also cover the true solution well.

Fig. 5 shows the corresponding plots for Cases 1*b* and 3*b*, where the time-lapse seismic data have a higher uncertainty level. The representation of the truth is clearly reduced for all cases, and they have obtain a much larger ensemble spread, both for the permeability and the OPR predictions. However, for Case 1*b*, the true field is still outside the envelope for some grid cells. Similar as for the *a*-cases the largest ensemble size in the *b*-cases obtain the best permeability estimation and highest spread.

By using a large number of members in the EnKF the computational effort will be demanding when working with real reservoir cases. Thus, to avoid the large ensemble size and at the same time achieve an improvement regarding the underestimation problem, a combined global and local scheme is investigated.

In Fig. 6 the EnKF results for Cases 4*a*-6*a* are presented, where the left, middle and right column show respectively the results when a stencil of size 3×3, 5×5 and 7×7 is used. As can be seen from the top row the final ensemble means are in close agreement to the truth, even with the smallest stencil. The local scheme use the information in the time-lapse seismic data from each local region, and is able to produce permeability estimation results that are consistent with the global scheme using a large ensemble size. The plot show that the ensemble spread from the local schemes are all higher than the global scheme with 100 and 250 members, and the uncertainty level for both the permeability and the OPR predictions are similar to the global scheme using 800 members. Further examination show that final ensemble means produced by the local scheme are a bit sparser than for the cases using the global scheme, and an explanation can be that the small influence region has a stronger impact of the center update. The final ensemble mean becomes gradually smoother and the ensemble spread decrease as the size of the stencil increase. Notice that if a stencil of size 15×15 had been used in the local scheme the results should be essentially indistinguishable with the global scheme using members of 100.

Corresponding plot for Case 5*b* is plotted in the left column of Fig. 7, where a higher measurement error in the seismic is used. By increasing the measurement error the correlation between the final ensemble mean and the truth is reduced. However, comparing the estimated mean for Cases 1*b*, 3*b* and 5*b*, similar features in part of the reservoir for Cases 3*b* and 5*b* can be seen, but some stronger local updates seems to be introduced in the local scheme.

The uncertainty results for the ensemble spread and the OPR predictions are significantly higher for Case 5*b* than for 1*b*, and have a closer relationship with the results obtained for Case 3*b*.

Inspecting Table 6, we see a slightly better performance of the local scheme when using a 5×5 stencil. Further, we see that Cases 4*b*-6*b* have lower rms values and higher correlation values than the global scheme with the same ensemble size, but the results are not as good as the global scheme with 250 and 800 members. Thus, to make the schemes with 100 ensemble members more comparable with the global scheme using 800 members, we have used an improved sampling strategy where an ensemble of 800 members is used to generate the initial 100 members.

Two cases using the improved sampling strategy were run. Case 7*b* use a global scheme and Case 8*b*, plotted in right column of Fig. 7, use a local scheme with a 5×5 stencil. An improvement can be seen for both the global and the local scheme, and Case 8*b* has obtained results that are at the same level as the global scheme using 800 members. The improvement of the results are in an agreement with the results observed by Evensen (2004), where the EnKF simulations were based on a linear advection model to demonstrate the impact of the improved sampling strategy.

Compared to Case 5*b*, a small reduction in the uncertainty of the permeability is seen for the improved sampling, and this can be explained by that a stronger correlation between the seismic signal and the permeability is obtained.

5. Synthetic 3D example

A synthetic 3D model of size $30 \times 30 \times 5$ is used to show the impact of introducing an uncertainty connected to the lithology in the rock physics model, and in this study we have quantified the lithology by the clay ratio parameter. The production plan is equal to the 2D model, where the wells now are penetrated in all layers. The realizations of permeability, porosity and clay ratio are generated and updated similarly as in Section 4. The ensemble statistics can be found in Table 7, where the variogram range in x and y direction is now set to be 10 grid cells and 2 grid cells in z-direction.

During the run 1 3D seismic dataset and 3 time lapse seismic datasets are assimilated, where each dataset include both the acoustic impedance and the Poisson's ratio. The 3D seismic dataset is assimilated after 1 year production, and the time lapse seismic datasets, which are the difference between the monitor and base survey at years 3 and 1, 5 and 3, and 7 and 5, are assimilated after 3, 5 and 7 years of production. The seismic measurement errors are presented in Table 8. Here we have used a spherical variogram with range of 5 grid cells in x and y direction and 2 grid cells and z-direction. The measurement errors in the acoustic

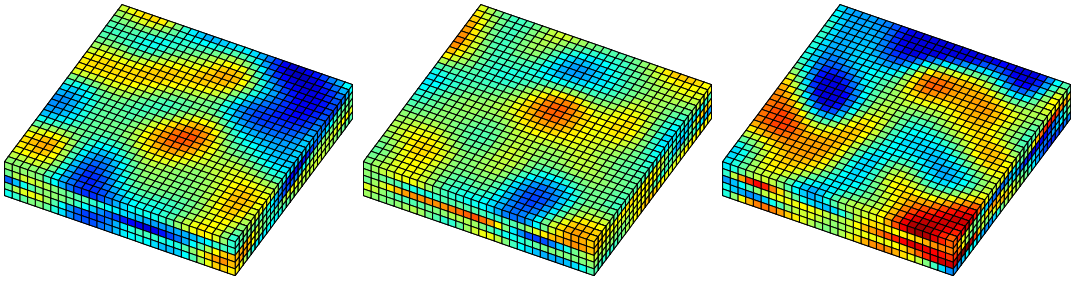


Fig. 8. Reference fields. Left: Permeability, Middle: Porosity, and Right: Clay ratio.

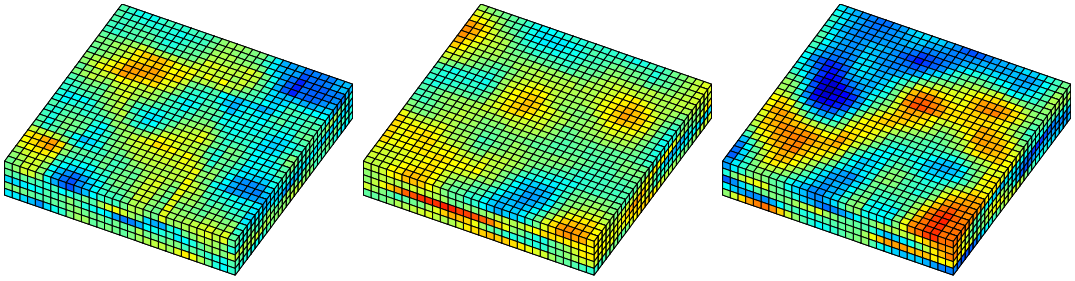


Fig. 9. Final estimated ensemble means for Case 9. Left: Permeability, Middle: Porosity, and Right: Clay ratio

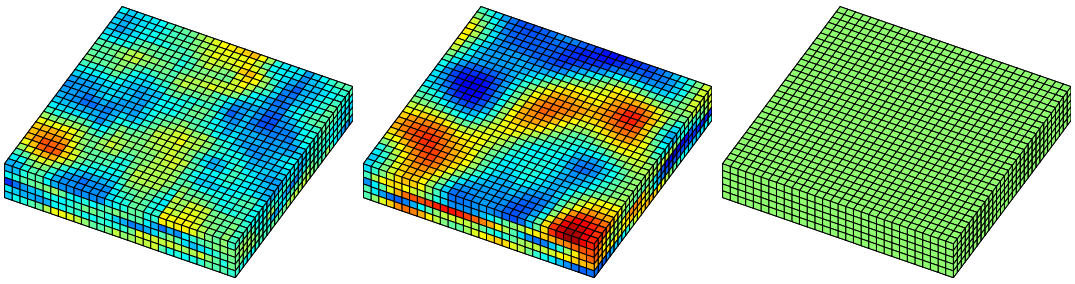


Fig. 10. Final estimated ensemble means for Case 10. Left: Permeability, Middle: Porosity, and Right: Clay ratio

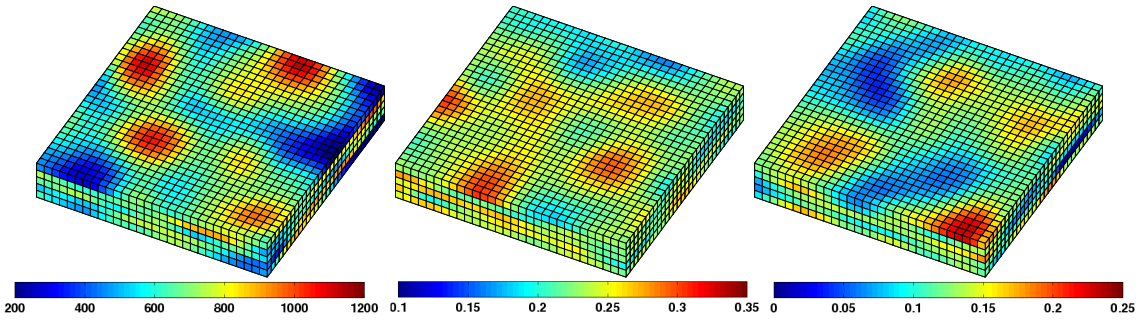


Fig. 11. Final estimated ensemble means for Case 11. Left: Permeability, Middle: Porosity, and Right: Clay ratio

Table 7
3D model data used as input in the EnKF.

	Mean	Standard deviation
Clay ratio	0.13	0.04
Porosity	0.23	0.03
Permeability (mD)	665	170

impedance and the Poisson's ratio are assumed to be uncorrelated. The well data are assimilated every 6 month, and the measurement errors can be found in Table 4.

Three cases of model updating were performed. In Case 9 clay ratio, porosity and permeability are updated using the combined global and local scheme with 100 members and a stencil of size $9 \times 9 \times 1$. Case 10 is similar to Case 9, but uses a constant clay ratio value equal to 0.13. Case 11 uses a global scheme with 200 members.

Table 8
Measurement errors for seismic data, standard deviations.

	3D data	4D data
Acoustic impedance (kg/m^{2}s)	5.0E+05	1.0E+04
Poisson's ratio (-)	5.0E-02	1.0E-03

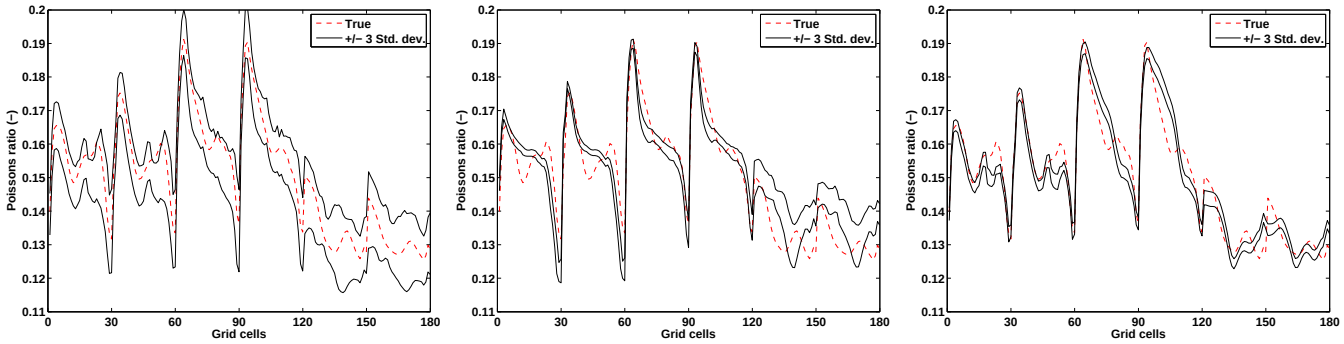


Fig. 12. Poisson's ratio predictions, Left: Case 9, Middle: Case 10, Right: Case 11. The plots show the plus/minus three standard deviation around the ensemble mean plotted together with the true solution. The predictions are calculated at year 9, where following grid cells are plotted: 1-60 in the upper layer, 1861-1920 in the third layer, and 4441-4500 in the bottom layer. The counting starts from the lower left grid cell in the first layer to the upper right grid cell in the fifth layer.

Table 9
Summary of experiments from the synthetic 3D model.

Case no.	Scheme	Clay ratio		Porosity		Permeability	
		RMS	Corr.	RMS	Corr.	RMS	Corr.
Case 9	Local 100	0.027	0.79	0.021	0.72	143.09	0.40
Case 10	Local 100, const.	0.045	0.0	0.037	0.59	170.65	0.30
Case 11	Global 200	0.046	0.37	0.033	0.32	220.84	0.11

The reference fields are plotted in Fig. 8 and the final ensemble mean fields for Case 9 are plotted in Fig. 9. The plots show that the combined global and local scheme obtain results that are in good agreement with the true model, and this can also be seen from the rms and correlation values in Table 9. Especially the estimation of the clay ratio and the porosity are better when the global/local scheme is used. The reason for this is most probably that the rock physics parameters in this case have a stronger correlation to the seismic, than the permeability.

Fig. 10 shows the final ensemble mean fields for Case 10. The plot shows that an over- and underestimation of the porosity values in parts of the reservoir during the assimilation are achieved. This seems to occur because a significant information in the seismic is related to a variation in the clay ratio. The rms and correlation values obtained for the porosity are respectively higher and lower than for Case 9, and this is due to the fact that the signature from the true clay ratio is found in the update of the porosity. The use of a constant clay ratio value also affects the estimation of the permeability shown by the values in Table 9.

Case 11 is performed as a comparison run with the proposed global/local scheme, and the final ensemble means are plotted in Fig. 11. The features in the estimated fields recover only some parts of the truth, and compared with Case 9 the values in Table 9 show that the global/local scheme has obtained a much better estimation result.

To examine the importance of introducing an uncertainty connected to the clay ratio parameter in the rock physics model the spread of the predictions of the elastic parameters will be of interest. Thus, a forward run from year 7.5 to year 9 is performed, and at the end time the Poisson's ration for each member is calculated. In Fig. 12 the Poisson's ratio predictions for Cases 9, 10 and 11 are plotted, where we have chosen the grid cells 1-60 in the upper layer, 1861-1920 in the third layer, and 4441-4500 in the bottom layer. The counting starts from the lower left grid cell in the first layer to the upper right grid cell in the fifth layer. As can be seen, the seismic ensemble predictions for Case 9 cover the true model significantly better than for Case 10, and clearly, including the lithology in the update improves the prediction of the elastic parameters. Fig. 12 also show the performance of the global scheme, and compared with the results in Case 9, a too low uncertainty in the seismic predictions are obtained. For the local scheme, the grid cells in the fifth layer show a higher uncertainty in the predictions, and a reason of this behavior may be related to the lack of information in the conditioning, since the the waterfront had not reached those cells at the last time lapse seismic assimilation step.

6. Discussion and conclusions

In this study we have presented a combined global and local EnKF scheme to estimate the permeability, porosity and lithology by an assimilation of production data and seismic data. An investigation of the performance of the local scheme versus the global scheme, when different ensemble sizes and stencil sizes are used in the EnKF method, is also performed.

The EnKF performance of the experiments discussed here were sensitive to the ensemble size, the sampling scheme, the error level of the seismic data, and in the local scheme an increase of the stencil size gave a gradually smoother ensemble mean and a decrease in the ensemble spread. All of these sensitivities affected the characterization of the reservoir and influenced the covariance estimation of the model variables, and thus the uncertainty in the predictions of the production data.

The results of the above experiments show that when the amount of data is large, there are problems related to the low rank representation of the model variable covariance matrix P_ψ, and that a large ensemble size is needed to

avoid these. This problem becomes more pronounced if the measurement error is decreased. Here, the combined global and local scheme gave promising results regarding the uncertainty estimation, even with few members applied.

The benefit of using a local scheme is that the ratio between the dimension of the model state space and the ensemble size is smaller than for a standard global analysis in the EnKF update, and this allows for a larger flexibility to obtain different model solutions, since the analysis will use a different combination of ensemble members for each grid point. The local scheme might also be more advantageous when many observations, especially with long range correlated errors, are assimilated.

A potential disadvantage of the local scheme is that non-physical modes may occur in the analysis fields, because the updates are performed independently in each local region. Especially when observations with a fairly high white noise are assimilated, this is a concern. It is then required that a large influence region is used in the local update to preserve the smoothness of the analysis fields.

It is demonstrated that taking into account uncertainty in lithology by updating the clay ratio may be important. A dispersed clay model was chosen to model the effect of clay in sandstones, and in such a model the lithology may have a strong impact on the seismic signal. The simulations show an improvement of the reservoir characterisation, when an estimation of the clay ratio was performed. The results also show the importance of including the clay ratio to obtain satisfactory uncertainty predictions of the elastic parameters.

Acknowledgements

We thank Statoil and Norsk Hydro for financial support of this research and Schlumberger for providing the reservoir simulation software.

References

Anderson, J., 2001. An ensemble adjustment filter for data assimilation. Monthly Weather Review 129, 2884–2903.

Burgers, G., van Leeuwen, P. J., Evensen, G., 1998. Analysis scheme in the ensemble Kalman filter. Monthly Weather Review 126, 1719–1724.

Dræge, A., Johansen, T. A., Brevik, I., Dræge, C. T., 2006. A strategy for modelling diagenetic evolution of seismic properties in sandstones. Accepted for publication in Petroleum Geoscience.

Evensen, G., 1994. Sequential data assimilation with nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J. Geophys. Res. 99, 10143–10162.

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

Evensen, G., 2004. Sampling strategies and square root analysis schemes for the EnKF. Ocean Dynamics 04 54, 539–560.

Evensen, G., 2006. Data Assimilation, Ensemble Kalman filter. Springer.

Evensen, G., van Leeuwen, P. J., 2000. An Ensemble Kalman Smoother for nonlinear dynamics. Mon. Weather Rev. 128, 1852–1867.

Gao, G., Reynolds, A. C., 2005. Quantifying the uncertainty for the PUNQ-S3 problem in a Bayesian setting with the RML and EnKF. SPE reservoir simulation symposium (SPE 93324).

Gu, Y., Oliver, D. S., 2004. History Matching of the PUNQ-S3 Reservoir Model Using the Ensemble Kalman Filter. SPE Annual Technical Conference and Exhibition, SPE 89942.

Hamill, T., Whitaker, J., Snyder, C., 2001. Distance-dependent filtering of background error covariance estimates in an ensemble kalman filter. Monthly Weather Review 129, 2776–2790.

Haugen, V., Evensen, G., 2002. Assimilation of SLA and SST data into an OGCM for the Indian ocean. Ocean Dynamics 52, 133–151.

Haugen, V., Natvik, L. J., Evensen, G., Berg, A., Flornes, K., Nævdal, G., 2006. History Matching Using the Ensemble Kalman Filter on a North Sea Field Case. SPE Annual Technical Conference and Exhibition, SPE 102430.

Hegstad, B. K., Omre, H., 2001. Uncertainty in production forecasts based on well observations, seismic data, and production history. Society of Petroleum Engineers Journal, 409–424.

Houtekamer, P., Mitchell, H., 1998. A sequential ensemble Kalman filter for atmospheric data assimilation. Monthly Weather Review 126(3), 796–811.

Johansen, T. A., Drottning, Å., Lecomte, I., Gjøystdal, H., 2002. An approach to combine rock physics and seismic modelling of fluid substitution effects. Geophysical Prospecting 50, 119–137.

Kalman, R. E., 1960. A new approach to linear filtering and prediction problems. Transactions of the ASME-Journal of Basic Engineering 82 (Series D), 35–45.

Liu, N., Oliver, D. S., 2003. Evaluation of Monte Carlo methods for assessing uncertainty. SPE Annual Technical Conference and Exhibition (SPE 84936).

Mavko, G., Mukerji, T., Dvorkin, J., 2003. The Rock Physics Handbook. Cambridge University Press.

Nævdal, G., Johnsen, L. M., Aanonsen, S. I., Vefring, E. H., 2003. Reservoir Monitoring and Continuous Model Updating Using Ensemble Kalman Filter. SPE Annual Technical Conference and Exhibition, SPE 84732.

Oliver, D. S., Cunha, L. B., Reynolds, A. C., 1997. Markov chain Monte Carlo methods for conditioning a permeability field to pressure data. Mathematical Geology 29, 61–91.

Ott, E., Hunt, B. R., Szunyogh, I., Zmin, A. V., Kostelich, E. J., Corazza, M., Kalnay, E., Patil, D. J., Yorke, J. A., 2004. A local ensemble Kalman filter for atmospheric data assmilation. Tellus 56A, 415–428.

Sam, M. S., Andrea, M., 2001. The effect of clay distribution on the elastic properties of sandstones. Geophysical Prospecting 49, 128–150.

Skjervheim, J. A., Aanonsen, S. I., Evensen, G., 2006. Ensemble Kalman filter with time difference data. Submitted to Computational Geoscience.

Skjervheim, J. A., Evensen, G., Aanonsen, S. I., Ruud, B. O., Johansen, T. A., 2005. Incorporating 4D Seismic Data in Reservoir Simulation Models Using Ensemble Kalman Filter. SPE Annual Technical Conference and Exhibition, SPE 95789.

Szunyogh, I., Kostelich, E. J., Gyarmati, G., Patil, D. J., Hunt, B. R., Kalnay, E., Ott, E., Yorke, J. A., 2005. Assesing a Local Ensemble Kalman Filter: Perfect Model Experiments with the NCEP Global Model. Tellus 57A, 528–545.

Tarantola, A., 2005. Inverse Problem Theory and Methods for Model Parameter Estimation. SIAM, Philadelphia.

Tippett, M. K., Anderson, J. L., Bishop, C. H., Hamill, T. M., Whitaker, J. S., 2003. Ensemble square root filters. Monthly Weather Review 131, 1485–1490.

Wen, X.-H., Chen, W. H., 2005. Real-Time Reservoir Model Updating Using Ensemble Kalman Filter. SPE Reservoir Simulation Symposium, SPE 92991.

Paper F

Characterization of fractured reservoirs by effective medium modelling and joint inversion of seismic and production data *

* Presented at the 12th International Workshop on Seismic Anisotrophy (12IWSA). Submitted to the *Journal of Seismic Exploration*.

Characterization of fractured reservoirs by effective medium modelling and joint inversion of seismic and production data

Morten Jakobsen[a,b], **Jan-Arild Skjervheim**[a] **and Sigurd Ivar Aanonsen**[a]

[a] *Centre for Integrated Petroleum Research, University of Bergen*

[b] *Department of Earth Science, University of Bergen*

(15 December 2006)

ABSTRACT

This paper proposes a method for characterization of natural fractures in hydrocarbon reservoirs by quantitative integration of production data and (anisotropic) seismic attributes. The method is based on a unified model for the effective elastic and hydraulic properties of fractured porous media (which takes into account the effects of fracture geometry and fracture-fracture interaction in a consistent manner) and an Ensemble Kalman Filter (data assimilation) method (which provides uncertainties as well as mean values). In principle, our method can deal with fairly complex models of fractured reservoirs (e.g., involving multiple sets of vertical fractures that are embedded in a heterogeneous matrix). However, the initial inversion results presented here are based on a simplified model (involving a single set of horizontal fractures that are embedded in a homogeneous matrix). At the same time, the simplified model is both heterogeneous and anisotropic, since the fracture density was allowed to vary from grid block to grid block, in accordance with a (Gaussian) geostatistical model. An application to synthetic data suggests that one may obtain a significantly better estimate of the fracture density and permeability distributions within a fractured reservoir, by using time-lapse measurements of seismic attributes (the vertical P-wave acoustic impedance of each grid block) in addition to reservoir production data (bot-

tomhole pressure, water cut, gas-oil ratio), in the dynamic reservoir characterization (history matching) process.

INTRODUCTION

As conventional clastic reservoirs deplete, exploration and production of tight sandstones and carbonates becomes an increasingly important business for the petroleum industry. Its success is largely determined by our ability to identify and characterize fracture networks that often provide pathways for hydrocarbon flow. Heterogeneities related to fractures between wells cannot be reliably inferred from borehole data, logs or cores alone. However, they, may be obtained from remote physical measurements, especially seismic measurements. Seismic methods may play a more central role in helping to simulate the flow of fluids in fractured reservoirs, but first we need a better understanding of the relevant physics and scaling-issues.

When a fluid flows in a fractured porous medium the scale-size of pressure variations (the size of a typical grid block in a numerical reservoir simulator) is often much larger than the scale-size of the fractures (e.g., the distance between them) so that the flowing fluid cannot 'see' the individual fractures, but only a homogenized (or upscaled) structure. For the purpose of numerical reservoir simulation, therefore, the (heterogeneous) fractured porous medium may be replaced by an effective homogeneous porous medium.

Similarly, when a seismic wave propagates through a fractured porous medium the seismic wavelength is often much larger than the scale-size of the fractures, so that the propagating wave cannot 'see' the individual fractures, but only an averaged, or smeared-out structure. For the purpose of seismic modelling, therefore, the (heterogeneous) fractured porous medium may be replaced by a long-wavelength equivalent homogeneous porous medium.

The effective hydraulic and elastic properties of a fractured porous medium will

be anisotropic if the fractures have a preferred orientation and/or if they are not distributed in space in accordance with isotropic correlation functions. Curies principle implies that the tensors of effective elastic stiffnesses and hydraulic permeabilities are at least as symmetric as the underlying fracture pattern (which is responsible for the anisotropy). In other words, there may be strong correlations between the effective permeability and seismic anisotropy of a fractured porous medium.

In this study, we are interested in the physics of heterogeneous media like fractured reservoirs (composite porous media) in the domains of effective hydraulic and (poro)elastic properties. As discussed by Gueguen and Palciauskas (1994), there exist a formal mathematical relationship between the physical properties of heterogeneous mediarocks in these two domains. This analogy is quite useful because a calculation of an effective property in one domain can more or less directly be carried over to another. This was recently demonstrated by Jakobsen (2006), who carried the T-matrix approximations of Jakobsen et al. (2003a) over from the (poro)elastic domain for use in the hydraulic domain.

The correlations between the effective hydraulic and (poro)elastic properties of fractured porous media discussed above suggests that one could in principle use 3D seismic data to determine the effective permeability of fractured reservoirs. However, this task generally represents an ill-posed (nonlinear) inverse problem, which can only be solved (by linearization) if we know the initial fluid-saturations. Temporal changes in the dynamic variables (fluid-saturations, pore pressure) can be predicted by using a numerical reservoir simulator, provided that we know the effective (absolute) permeability tensor of each grid block. However, the effective permeabilities are exactly what we are trying to determine within the framework of fractured reservoir characterization. This is not so easy, because the properties (porosity, permeability) of the matrix material (in which the fractures are embedded) can also vary with position. In general, the use complementary data types may help us to reduce the uncertainty associated with an inverse problem. For many reasons, therefore, it may

be a good idea to try to perform some kind of (iterative) joint inversion of production data and seismic attributes (derived from the effective elastic properties of each grid block). The historically first attempt to integrate production data and (anisotropic) seismic attributes for characterization of naturally fractured reservoirs (using discrete-feature-network models) is represented by the paper of Will et al. (2005). Essentially, the present study represents an attempt to improve on the work of Will et al. (2005).

The outline of this paper is as follows. First, we discuss how to estimate the effective permeability and undrained stiffness tensors of fractured porous media on the basis of rigorous integral equation methods (eg., Jakobsen et al., 2003a; Jakobsen and Hudson, 2003; Jakobsen, 2006), rather than the empirical methods of Oda (1985) and Schoenberg and Sayers (1995) used by Will et al. (2005). Second, we discuss how to perform a joint inversion of seismic data (vertical P-wave acoustic impedances) and production data (gas-oil ratios, water cuts, bottomhole pressures) on the basis of an Ensemble Kalman Filter (Bayesian data assimilation) method (see Evensen, 2006), rather than the gradient-based method of optimization used by Will et al. (2005). Numerical examples as well as short discussions about the advantages and disadvantages of the different methods of effective medium modelling and joint inversion will be presented as we go along. Finally, we provide some concluding remarks and recommendations for future work.

EFFECTIVE MEDIUM MODELLING OF FRACTURED POROUS MEDIA

We depict a fractured porous medium as being composed of a porous matrix with a population of oblate spheroidal inclusions representing the fractures. The elastic and hydraulic properties of the matrix is assumed to be known and represented by the permeability and stiffness tensors $\mathbf{k}^{(0)}$ and $\mathbf{C}^{(0)}$, respectively. The population of fractures is divided into families of fractures having the same shape/orientation, and volume concentration $v^{(n)}$, labelled by $n = 1, 2, ..., N$. The fracture volume

concentration (or porosity) $v^{(n)}$ is related with the fracture density $\epsilon^{(n)}$ by $v^{(n)} = (4/3)\epsilon^{(n)}\alpha^{(n)}$, where $\alpha^{(n)}$ is the aspect ratio for fractures of type n. The aspect ratio $\alpha^{(n)}$ of a spheroidal inclusion with short (long) axis $c^{(n)}$ ($a^{(n)}$) is given by $\alpha = c^{(n)}/a^{(n)}$.

Effective hydraulic properties

The effective permeability tensor tensor $\mathbf{k}^*$ of the fractured porous medium described above is given by (Jakobsen, 2006)

$$\mathbf{k}^* = \mathbf{k}^{(0)} + \mathbf{k}_1 \left(\mathbf{I}_2 + \mathbf{k}_1^{-1}\mathbf{k}_2\right)^{-1}. \tag{1}$$

Here, $\mathbf{I}_2$ is the identity for second-rank tensors, $\mathbf{k}_1$ is a second-rank tensor of first-order corrections (for the effects of isolated fractures), and $\mathbf{k}_2$ is a second-rank tensor of second-order corrections (for the effects of fracture-fracture interaction).

The first-order correction is given by (Jakobsen, 2006)

$$\mathbf{k}_1 = \sum_r v^{(r)}\boldsymbol{\tau}^{(r)}, \tag{2}$$

where

$$\boldsymbol{\tau}^{(n)} = \left(\mathbf{k}^{(n)} - \mathbf{k}^{(0)}\right)\left[\mathbf{I}_2 - \mathbf{g}^{(n)}\left(\mathbf{k}^{(n)} - \mathbf{k}^{(0)}\right)\right]^{-1}. \tag{3}$$

Here, $\mathbf{k}^{(n)}$ is a second-rank tensor of effective permeability coefficients for fractures of type n, $\mathbf{g}^{(n)}$ is a second-rank tensor given by the pressure gradient Green's function (for a material with properties given by $\mathbf{k}^{(0)}$) integrated over a characteristic (oblate) spheroid having the same shape as fractures of type n (see Jakobsen, 2006).

The second-order correction is given by (Jakobsen, 2006)

$$\mathbf{k}_2 = \sum_r \sum_s v^{(r)}\boldsymbol{\tau}^{(r)}\mathbf{g}_d^{(rs)}\boldsymbol{\tau}^{(s)}v^{(s)}. \tag{4}$$

Here, $\mathbf{g}_d^{(rs)}$ is a second-rank tensor given by the pressure gradient Green's function (for a material with properties given by $\mathbf{k}^{(0)}$) integrated over a characteristic (oblate)

spheroid having the same aspect ratio as $p^{(s|r)}(\mathbf{x}-\mathbf{x}')$ which, in turn, gives the probability density for finding an inclusion of type s at point $\mathbf{x}'$ given that there is an inclusion of type r at point $\mathbf{x}$ (see Jakobsen et al., 2003a; Jakobsen, 2006).

Figure 2 shows the results of a numerical investigation of the effective hydraulic properties of a porous medium containing a single set of horizontal fractures. The ratio of the isotropic permeabilities of the fractures and matrix was 10^4. The aspect ratio of the fractures was taken to be 1/100. The figure shows that the horizontal permeability of the fractured porous medium can be 60 percent higher than the matrix permeability, whereas the vertical permeability is hardly affected by the fractures. These results are consistent with those of Pozdniakov and Tsang (2004). Jakobsen (2006) compared the prediction of the T-matrix approach with numerical simulations performed by Pozdniakov and Tsang (2004), and obtained a good match between theory and numerical experiments.

Effective elastic properties

The dry case.—The effective stiffness tensor $\mathbf{C}^{d*}$ of the fractured porous medium described above is given (for the dry case) by (Jakobsen et al., 2003a; Jakobsen and Hudson, 2003)

$$\mathbf{C}^{d*} = \mathbf{C}^{(0)} + \mathbf{C}_1 \left(\mathbf{I}_4 + \mathbf{C}_1^{-1}\mathbf{C}_2\right)^{-1}, \tag{5}$$

Here, $\mathbf{I}_4$ is the identity for second-rank tensors, $\mathbf{C}_1$ is a fourth-rank tensor of first-order corrections (for the effects of isolated fractures), and $\mathbf{C}_2$ is a fourth-rank tensor of second-order corrections (for the effects of fracture-fracture interaction).

The first-order correction is given by (Jakobsen et al., 2003a)

$$\mathbf{C}_1 = \sum_r v^{(r)}\mathbf{t}^{(r)} \tag{6}$$

where

$$\mathbf{t}^{(r)} = -\mathbf{C}^{(0)} \left[\mathbf{I}_4 + \mathbf{G}^{(n)} \mathbf{C}^{(0)} \right]^{-1}. \tag{7}$$

Here, $\mathbf{G}^{(n)}$ is a fourth-rank tensor given by the strain Green's function (for a material with properties given by $\mathbf{C}^{(0)}$) integrated over a characteristic (oblate) spheroid having the same shape as fractures of type n (see Jakobsen et al., 2003a).

The second-order correction is given by (Jakobsen et al., 2003a)

$$\mathbf{C}_2 = \sum_r \sum_s v^{(r)} \mathbf{t}^{(r)} \mathbf{G}_d^{(rs)} \mathbf{t}^{(s)} v^{(s)} \tag{8}$$

Here, $\mathbf{G}_d^{(rs)}$ is a fourth-rank tensor given by the strain Green's function (for a material with properties given by $\mathbf{C}^{(0)}$) integrated over a characteristic (oblate) spheroid having the same aspect ratio as the two-point correlation function $p^{(s|r)}(\mathbf{x} - \mathbf{x}')$ (described earlier).

The saturated case.—In order to calculate the effects of fluid-saturation on the effective elastic properties of a fractured porous medium, one can use the (anisotropic Gassmann) relation of Brown and Korringa (1975):

$$S^*_{ijkl} = S^{d*}_{ijkl} + \frac{\left(S^{d*} - S^{(m)} \right)_{ijuu} \left(S_{d*} - S^{(m)} \right)_{vvkl}}{\phi^0 \left(S^{(m)}_{uuvv} - 1/\kappa_f \right) - \left(S^*_d - S^{(m)} \right)_{uuvv}}. \tag{9}$$

Here, $\mathbf{S}^{(m)} = (\mathbf{C}^{(m)})^{-1}$, $\mathbf{S}^{d*} = (\mathbf{C}^{d*})^{-1}$ and $\mathbf{S}^* = (\mathbf{C}^*)^{-1}$ is the compliance tensor of the solid mineral component, the dry fractured porous medium and the saturated fractured porous medium (in the undrained case), respectively. ϕ is the total porosity; that is, the sum of the storage porosity associated with the homogeneous matrix material, and the fracture porosity (which is normally extremely small) and κ_f is the bulk modulus of the saturating fluid.

Brown and Korringa (1975) originally assumed that the anisotropic porous system was fully saturated with a single fluid constituent. In the case of a fractured porous medium which is partially saturated with oil (o), gas (g) and water (w), κ_f may be regarded as the bulk modulus of an effective fluid, and we may use the well-known formula of Wood:

$$\frac{1}{\kappa_f} = \frac{S_w}{\kappa_w} + \frac{S_o}{\kappa_o} + \frac{S_g}{\kappa_g}, \tag{10}$$

where

$$S_w + S_o + S_g = 1. \tag{11}$$

Woods equation (10) corresponds with an iso-stress calculation, which is generally believed to be useful in the low-frequency regime, where there is enough time for the pore pressure to equalize throughout the pore space (as assumed by Brown and Korringa, 1975).

Jakobsen et al. (2003b) have developed a theory of wave-induced fluid flow, which is (consistent with the Brown-Korringa relations) based on the higher-order T-matrix approximations of Jakobsen et al. (2003a). As discussed by Aagersborg et al. (2007), the theory of Jakobsen et al. (2003a,b) includes (but is more general) than the theory for frequency-dependent anisotropy due to mesoscopic fractures in the presence of equant porosity developed by Chapman (2003). By ignoring the higher-order terms in the theory of Jakobsen et al. (2003a,b) and assuming that the ratio of the characteristic times for wave-induced fluid flow at the scale of the mesoscopic fractures and the pore scales is given by the ratio of the corresponding scale-sizes (as suggested by Chapman, 2003), Aagersborg et al. (2007) obtained a complete match between these two theories. This is very encouraging, but we decided to use neither of these two (squirt flow) theories in this study. The reason for this is two-fold: (1) they are restricted to a single fluid constituent only, and (2) they become single-scale models in the low-frequency limit. By single-scale, we mean that they becomes independent of the fracture-size in the limit of very low frequencies (and in the dry case), because the fractures in these models were embedded in the solid matrix material rather than in the porous matrix (unfractured reservoir).

Figure 3 shows T-matrix estimates of the effective elastic properties of a fractured porous medium which is partially saturated with a water and gas. Displayed are the

anisotropy parameters of Thomsen (1986), as well as the speed of vertical P- and S-waves. If we look at the effects of fluid-saturation at the highest fracture densities, we see that a small amount of gas leads to a large drop in the vertical P-wave speed, and a significant increase of Thomsens anisotropy parameters ϵ and δ. We note that the anellipticity is positive $(\epsilon - \delta) < 0$, which means that the slowness surface bulges outward from the ellipse which is connecting the horizontal and vertical slownesses. This observation may be used as a constraint within the context of seismic inversion. Figure 4 is similar to Figure 3 but for a fractured porous medium which is partially saturated with water and oil. We note that the surfaces are now much more planar, and the anellipticity is still positive.

JOINT INVERSION OF SEISMIC AND PRODUCTION DATA

General considerations

In the traditional (seismic) history matching system of Will et al. (2005) (Figure 6), the model parameters of a discrete-feature-network model (e.g., fracture density, fracture strike) are adjusted until the synthetic production data and seismic anisotropy attributes matches the observations. This procedure requires repeated flow simulations and seismic attribute generations, because it is formulated in the form of a minimization problem in which the mismatch between measurements and computed values is minimized. Will et al. (2005) used a gradient-based method for this minimization problem, which require the computation of sensitivity coefficients. Such computations are very time-consuming, and so there is a limit on the number of model parameters that one can practically use. Furthermore, the assessment of uncertainty (not performed by Will et al., 2005) can only be done through repeated history matching with different initial models, which makes the process even more time-consuming. In addition, traditional methods for joint inversion of seismic and

production data, like the one used by Will et al. (2005), does not allow for continuous model updating.

The Kalman filter (KF) method (which is compared with the conventional method in Figure 5) was originally developed to continuously update the states of linear systems to account for available measurements. When the system is non-linear, the extended Kalman Filter was proposed to linearize the non-linear system. For very large models and/or highly non-linear systems, the extended Kalman filter fails (see Evensen, 2006).

The ensemble Kalman filter (EnKF) (which is illustrated in Figure 7) was introduced to overcome some of the problems of the extended Kalman filter (Nævdal et al., 2003; Evensen, 2006). Particularly, instead of evaluating the necessary statistics (e.g., correlations between model parameters and responses) based on linear assumptions, EnKF uses an ensemble of model representations from which all necessary statistics can be directly computed (Nævdal et al., 2002; Evensen, 2006). As discussed by Skjervheim et al. (2005), the EnKF method represents a data assimilation method based on a sequential use of Bayes's theorem (e.g., Tarantola, 2005; Evensen, 2006).

Workflow

In the EnKF approach used in this paper, the state vectors $\boldsymbol{\psi}$ include four types of variables: (1) static parameters (related with the fractures and the porous matrix in which they are embedded) (2) dynamic state variables (the fluid saturation and pore pressure within each grid block), (3) production data (observations in boreholes) and (4) seismic attributes (derived from the effective density and elastic properties of each grid block). The process of seismic history matching in fractured reservoirs using the EnKF approach consists of the following stages:

1. Initialization of the state vector.

2. Forecasting.

3. Updating of the state vector by data assimilation.

4. Evaluate and repeat from stage 2.

Each step of this process will be detailed in what follows.

Initialization of the state vector

The filter is initialized by generating initial ensembles of static and dynamic vectors (Nævdal et al., 2003; Wen and Chen, 2005). There is no production data available at the starting time t_0, but 3D seismic data may be available. The initial state should be conditioned to the initial seismic data.

In our system, a geostatistical model is used to generate multiple realizations of the fracture density, with given statistical parameters (histogram and variogram) to represent the initial uncertainty in the fracture distribution before the sequential data assimilation process is started. All other parameters related with the fractures and the matrix are kept constant. If other static variables related with the fractures (e.g., fracture orientations, shapes) and/or the matrix (e.g., the background permeability, storage porosity) are not known they should also be represented by ensembles.

Initial dynamic variables (i.e., initial saturations and pore pressures) are sometimes assumed known without uncertainty, so that they are the same for each realization. If the initial dynamic variables are not known then they should also be represented by ensembles.

Forecasting

The forecast step calls a numerical reservoir simulator and a seismic attribute generator for each of the realizations of the uncertain fracture distribution, until the next point in time where production and/or seismic data are available and to be assimilated (e.g., t_1). The state vector for ensemble member i after running the

forecasting is denoted by $\boldsymbol{\psi}_i^f$. The determination of $\boldsymbol{\psi}_i^f$ (which includes the old static variables, the new dynamic variables and forecasted production data and seismic attributes) is a two-stage process that can be described as follows.

Fluid flow simulation.—For each realization of the fracture density, the effective absolute permeability tensor of each (fractured) grid block was estimated by using the T-matrix approach. For simplicity, the relative permeabilities of each grid block (which depends on the corresponding fluid-saturations) were in this study assumed approximately equal to the relative permeabilities of the unfractured matrix material. In the future, however, we shall try to perform a steady-state upscaling (or homogenization) of the relative permeabilities based on a repeated use of single-phase upscaling methods (the T-matrix approach), in the limit of capillary equilibrium (see Ekran and Aasen, 2000).

An ensemble of forecasted dynamic variables and production data were subsequently generated by using a single-porosity formulation associated with the Eclipse reservoir simulator. Will et al. (2005) used a dual-porosity formulation, but both single- and double-porosity simulators have been used for fractured reservoirs (see Gurpinar and Kossack, 2000).

It is well-known that the pore pressure and the fluid saturations changes during production of hydrocarbones from a naturally fractured reservoir. The absolute permeabilities are independent of the saturations, but generally depends on the pore pressure, which controls the deformation (closure) of the fractures, when they are subjected to a constant overburden stress.

A combination of the T-matrix approximations of Jakobsen (2006) with the formula of Jakobsen and Johansen (2005) for the deformation of a population during drained loading may in principle be used to estimate the effects of pore pressure on the effective absolute permeability tensors of fractured porous grid block. However, in this paper we followed Will et al. (2005) in ignoring such geomecanical effects.

Seismic attribute generation.—The T-matrix approach was used to transform maps of fracture densities and dynamic variables into corresponding maps of effective elastic parameters which, in turn, where used to construct synthetic seismic data. In this study, we treated the vertical P-wave acoustic impedance of each grid block as seismic data. As discussed below, however, the cost function (describing the difference between synthetic and observed seismic data) may be defined at various levels.

Within the context of seismic history matching (in unfractured media), it is common to treat the acoustic impedance of each grid block as "seismic data". This is probably because standard (isotropic) methods for inverting the seismic sections with respect to the acoustic impedance and the v_p/v_s ratio (or other combinations of the two independent elastic parameters for an isotropic medium) of each grid block are commonly available (see Skjervheim et al., 2005). A proper joint inversion of seismic and production data is possible if one includes (the effects of overburden) seismic sections and/or higher-level seismic attributes (travel time and/or amplitude information) in the (cost function) data assimilation process (Skjervheim and Ruud, 2006).

Updating of the state vector by data assimilation

The presentation of the EnKF given in this section follows the presentation of Nævdal et al. (2002). To take into account the measurements we use the covariance matrix of the ensemble around the ensemble mean. The mean value of the ensemble is given by

$$\langle \boldsymbol{\psi} \rangle = \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{\psi}_j, \tag{12}$$

and the ensemble covariance matrix is given by (Nævdal et al., 2002)

$$\mathbf{R} = \frac{1}{N-1} \sum_{i=1}^{N} (\boldsymbol{\psi}_i^f - \langle \boldsymbol{\psi} \rangle)(\boldsymbol{\psi}_i^f - \langle \boldsymbol{\psi} \rangle)^T, \tag{13}$$

where N is the number of members in the ensemble. We now assume that we have an ensemble of observations given in the following form (Nævdal et al., 2002):

$$\mathbf{d}_i = \mathbf{d} + \boldsymbol{\eta}_i, \tag{14}$$

where $\mathbf{d}$ is the actual observation and $\boldsymbol{\eta}$ is often drawn from a Gaussian distribution with zero mean and covariance matrix $\boldsymbol{\Sigma}$. The observation vector contains both production and seismic data and is related to the state vector $\boldsymbol{\psi}$ through the linear relation

$$\mathbf{d} = \mathbf{H}\boldsymbol{\psi}. \tag{15}$$

Here, the matrix $\mathbf{H}$, which contains only zeros and ones, pick the production data and the seismic attributes from the state vector $\boldsymbol{\psi}$). We consider the time lapse production and seismic data to be properly assimilated when they have been used to evaluate the Kalman gain matrix (Nævdal et al., 2002):

$$\mathbf{K} = \mathbf{R}\mathbf{H}^T(\mathbf{H}\mathbf{R}\mathbf{H}^T + \boldsymbol{\Sigma})^{-1}. \tag{16}$$

The state vectors in the ensemble are updated using the following equation (Nævdal et al., 2002):

$$\boldsymbol{\psi}_j^a = \boldsymbol{\psi}_j^f + \mathbf{K}(\mathbf{d}_j - \mathbf{H}\boldsymbol{\psi}_j^f). \tag{17}$$

The second part on the right-hand side of the above equation is the difference between simulated and observed production data and seismic attributes; the larger the difference, the large update will be applied to the initial/previous state vector. With this updating, the state vector of each realization in the ensemble is considered to reflect the most current production data and seismic attributes, and we can proceed to the next time step, where new production data and seismic attributes are available for assimilation.

Example

For simplicity, we consider a flay-lying reservoir consisting of a uniform porous matrix with embedded horizontal fractures. The fractures are assumed to be spheroidal in shape and characterized by a single aspect ratio. The different grid blocks will then have different tensors of effective permeabilities, because the fracture density is varying from grid block to grid block.

Figure 8 shows a 2-D geostatistical reference fracture density field ($15 \times 15 \times 1$) grid with cell size $120 \times 120 \times 100$ meter. The reference field is generated using a Gaussian Fast Fourier Transform (FFT) method. The fracture density ϵ has a Gaussian distribution with a mean and variance of 0.08 and 0.03, respectively. The (spherically symmetric) variogram is Gaussian with a range of 5 grid blocks. We assume an injection well (I) at the lower left corner and a producer (P) at the upper right corner. The reference field is considered as the true model, and our goal is to reconstruct it using a combination of production data and seismic attributes.

The reservoir was assumed to be initially fully saturated with oil at constant pore pressure of 140 bar. The injection well has a constant injection rate of 7000 Sm^3/day, and the producer is producing a constant volume of 6000 Sm^3/day. The relative permeability curves we have used in this study are taken from the real field case of Skjervheim et al. (2006). The flow simulation is run for 2500 days and results for the bottomhole pressure (BHP) as well as the water cut (WCT) and gas-oil ratio (GOR) are shown in Figure 2b-d. We can see in Figure 2b-d that the bottomhole pressure starts to increase after 365 days. This corresponds to the start of the injection process.

Figure 9 shows the acoustic impedances (with noise added) at the reference time $t_0 = 0.25$ years, the difference between the acoustic impedances at the first repetition time $t_1 = 4.5$ years and the reference time, and the difference between the acoustic impedances at the second repetition time $t_2 = 6.5$ years and the first repetition time. The assumed noise level (standard deviation of the acoustic impedances) for

the initial time (3D seismics) was 2×10^5, and 10^4 for the repetitions (4D seismics). The elastic properties of each grid block were estimated from the true fracture density field and the fluid-saturations predicted by the reservoir simulator for each time step. In the difference at time t_1, we can see changes in the acoustic impedances near the producer, even though the water front has not yet reached this area. This can be explained by a production-induced decrease in the pore pressure, leading to the formation of gas-bubbles in the oil phase (near the producer), and a consequent drop of the vertical P-wave speed (see Figure 3). The difference at time t_2 is mainly related to the movement of the waterfront.

Figure 10 shows a comparison of the true fracture density field with the reconstructed fields obtained by using production data only or a combination of seismic and production data. Clearly, a significantly better estimate of the fracture density is obtained by including time-lapse measurements of changes in the vertical P-wave acoustic impedance into the fractured reservoir history matching process.

As stated earlier, one of the main advantages of using the EnKF approach, rather than a conventional method of seismic history matching, is that we can estimate uncertainties as well as mean values. Figure 11 shows that the effect of adding seismic attributes to the history matching loop is to reduce the uncertainty of the estimated fracture densities.

Figure 12 shows a comparison of the true horizontal permeability field with the one reconstructed using both seismic and production data. Clearly, the concept of joint inversion is working even for complicated systems like naturally fractured reservoirs.

Having determined the permeability tensor of each grid block, the reservoir simulator can be used to a relatively high confidence for such tasks as predicting primary reservoir performance (Figure 13), well planning and EOR programs.

CONCLUDING REMARKS

The material presented here was the result of initial synthetic tests on a simplified model. The following conclusions and comments are made:

- The addition of seismic data to the data assimilation process improved our estimates of the fracture densities as compared with the use of production data only.

- The vertical P-wave acoustic impedance of each grid block were treated like "seismic data" in this study. It could be interesting to see if good results can be obtained if we ignore the anisotropy when we estimate the acoustic impedance of each grid block. After all, amplitude-based seismic inversion in anisotropic media does not represent an easy task.

- In principle, one could use all the five independent elastic parameters of each VTI grid block as seismic data. As discussed by Michelena et al. (1995), elastic constants in heterogeneous transversely isotropic media can in principle be determined by traveltime-based crosswell tomography.

- In any case, the assumption of a transversely isotropic medium with a vertical symmetry axis (VTI medium) means that the present system can not only be applied to reservoirs containing horizontal fractures but also to related anisotropic systems like sand-shale formations (see Desbarats, 1987; Pozdniakov and Tsang, 2004).

- It is possible to perform a proper joint inversion of seismic and production data on the basis of the present method, if full seismic sections are treated as seismic data and assimilated in the workflow of Ensemble Kalman Filtering. However, it could then be that the EnKF method will compensate for errors in

the model of the reservoir overburden by introducing unrealistic heterogeneities in the reservoir target zone.

- The next step may be to develop a similar system for reservoirs containing (multiple sets of) vertical fractures. A theory for calculating pressure-dependent seismic attributes of such azimuthally anisotropic media (in the case of a single fluid-constituent) was developed by Aagersborg et al. (2007).

- The work of Boulenko et al. (2005) suggests that it is also possible to use controlled source electromagnetic (CSEM) data in the characterization of fractured reservoirs by joint inversion of geophysical and production data.

REFERENCES

Aagersborg, R., Jakobsen, M., Ruud, B.O. and Johansen, T.A., 2007. The seismic response of a fractured carbonate reservoir. Stud. Geophys. et Geodaet., in press.

Brown, R.J.S, Korringa, J., 1975. On the dependence of elastic properties of a porous rock on the compressibility of the pore fluid. Geophysics **40**, 608-616.

Boulenko, M., Jakobsen, M. and Hesthammer, J., 2005. EM in reservoir monitoring - Response of a fractured hydrocarbon reservoir. 67th EAGE conference and exhibition, Madrid.

Chapman, M., 2003. Frequency-dependent due to mesoscale fractures in the presence of equant porosity. Geophys. Prosp., **51**, 369-379.

Desbarats, A.J., 1987. Numerical estimation of effective hydraulic conductivity in sand-shale formations. Water Resour. Res. **23**, 273-286.

Ekran, S. and Aasen, J.O., 2000. Steady-state upscaling. Transport in porous media, **41**, 245-262.

Evensen, G., 2006. Data assimilation - Ensemble Kalman Filter. Springer.

Gueguen, Y. and Palciauskas, V., 1994. Introduction to the physics of rocks. Princeton University Press.

Gurpinar, O. and Kossack, C., 2000. Realistic numerical models for fractured reservoirs. Paper 68268, SPE International Petroleum Conference and Exhibition, Villahermosa, Mexico.

Jakobsen, M., Hudson, J.A. and Johansen, T.A., 2003a. T-matrix approach to shale acoustics. Geophys. J. Int., **154**, 533-558.

Jakobsen, M., Johansen, T.A. and McCann, C., 2003b. The acoustic signature of fluid flow in complex porous media. J. Appl. Geophys. **54**, 219-246.

Jakobsen, M. and Hudson, J.A., 2003. Visco-elastic waves in rock-like composite. Stud. Geophys. Geodaet., **41**, 793-826.

Jakobsen, M. and Johansen, T.A., 2005. The effects of drained and undrained loading on visco-elastic waves in rock-like composites. Int., J. Solids and Structures, **42**, 1597-1611.

Jakobsen, M., 2006. The effective permeability of fractured reservoirs and composite porous media. Expanded abstract, 76th SEG meeting, New Orleans, US.

Michelena, R.J., Harris, J.M. and Muir, F., 1995. Crosswell tomographic estimation of elastic constants in heterogeneous transversely isotropic media. Geophysics, **60**, 774-783.

Nævdal, G., Mannseth, T. and Vefring, E.H., 2002. Near-well reservoir monitoring through ensemble kalman filter. SPE 75235.

Nævdal, G., Johnsen, L.M., Aanonsen, S.I. and Vefring, E., 2003. Reservoir monitoring and continuous model updating using the ensemble Kalman filter. SPE Annual Technical Conference and Exhibition (SPE 84372).

Oda, M., 1985. Permeability tensor for discontinuous rock masses. Geotech., **35**, 483.

Pozdniakov, S. and Tsang, C.T., 2004. A self-consistent approach for calculating the effective hydraulic conductivity of a binary, heterogeneous medium. Water Resour. Res., **40**, 1-13.

Schoenberg, M. and Sayers, C.M., 1995. Seismic anisotropy of fractured rock. Geophysics, **60**, 204.

Skjervheim, J.A., Evensen, G., Aanonsen, S., Ruud, B.O. and Johansen, T.A., 2005. Incorporating 4D seismic data in reservoir simulation models using ensemble kalman filter. SPE 95789.

Skjervheim J.A. and Ruud, B.O., 2006. Combined inversion of 4D seismic waveform data and production data using ensemble Kalman filter. Expanded abstract, 76th SEG meeting, New Orleans, US.

Thomsen, L., 1986. Weak elastic anisotropy. Geophysics, **51**, 1954-1966.

Tarantola, A., 2005. Inverse problem theory and methods for model parameter estimation. Society for Industrial and Applied Mathematics (SIAM).

Wen, X.H. and Chen, W.H., 2005. Real-time reservoir model updating using ensemble kalman filter. SPE 92991.

Will, R., Archer, R. and Dershowitz, B., 2005. Integration of seismic anisotropy and reservoir-performance data for characterization of naturally fractured reservoirs using discrete-feature-network models. SPE Reservoir Evaluation and Engineering, SPE 84412, April.

FIGURES

FIG. 1. The hydraulic response of an interacting fracture due to an applied pressure gradient at infinity. Reproduced from Pozdniakov and Tsang (2004).

FIG. 2. The effective hydraulic properties of a fractured reservoir.

FIG. 3. Estimated effective elastic parameters as functions of the fracture density and water-saturation for a fractured porous medium which is partially saturated with water and gas.

FIG. 4. Same as Figure 3 but for a fractured porous medium which is partially saturated with water and oil.

FIG. 5. The conventional (gradient-based) workflow for joint inversion of seismic and production data used by Will et al. (2005).

FIG. 6. A comparison between the workflows of conventional and seqential (data assimilation) methods of (seismic) history matching.

FIG. 7. An illustration of the Ensemble Kalman Filter.

FIG. 8. The reference fracture density field and production data.

FIG. 9. Acoustic impedances (with noise added) at 0.25 years (left), the difference between the acoustic impedances at 4.5 years and 0.25 years (middle), the difference between the acoustic impedances at 6.5 years and 4.5 years (right).

FIG. 10. The true fracture density field (left) vs the estimated fracture density field

using production data only (middle) and using both seismic and production data (right).

FIG. 11. The uncertainty of the fracture density (represented by the standard deviation) obtained by using the EnKF method with production data only (top) and both seismic and production data (bottom).

FIG. 12. The true horizontal permeability field (left) vs the estimated horizontal permeability field using production data only (middle) and using seismic and production data (right).

FIG. 13. Forcasting changes in the future oil production.

;

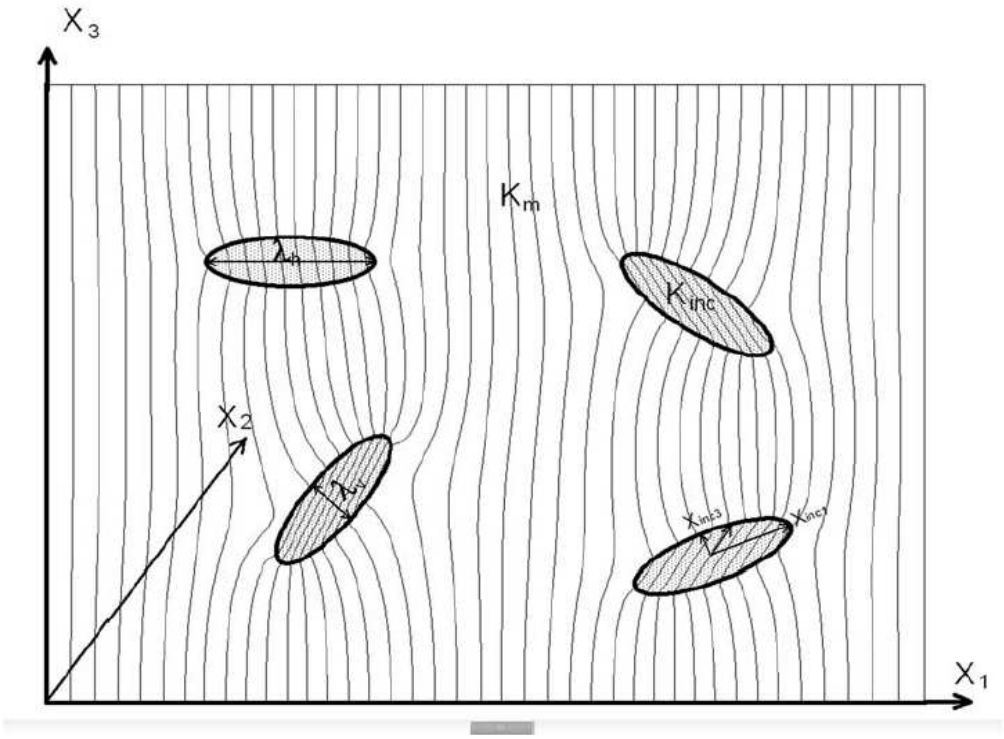


FIG. 1. The hydraulic response of an interacting fracture due to an applied pressure gradient at infinity. Reproduced from Pozdniakov and Tsang (2004).

;

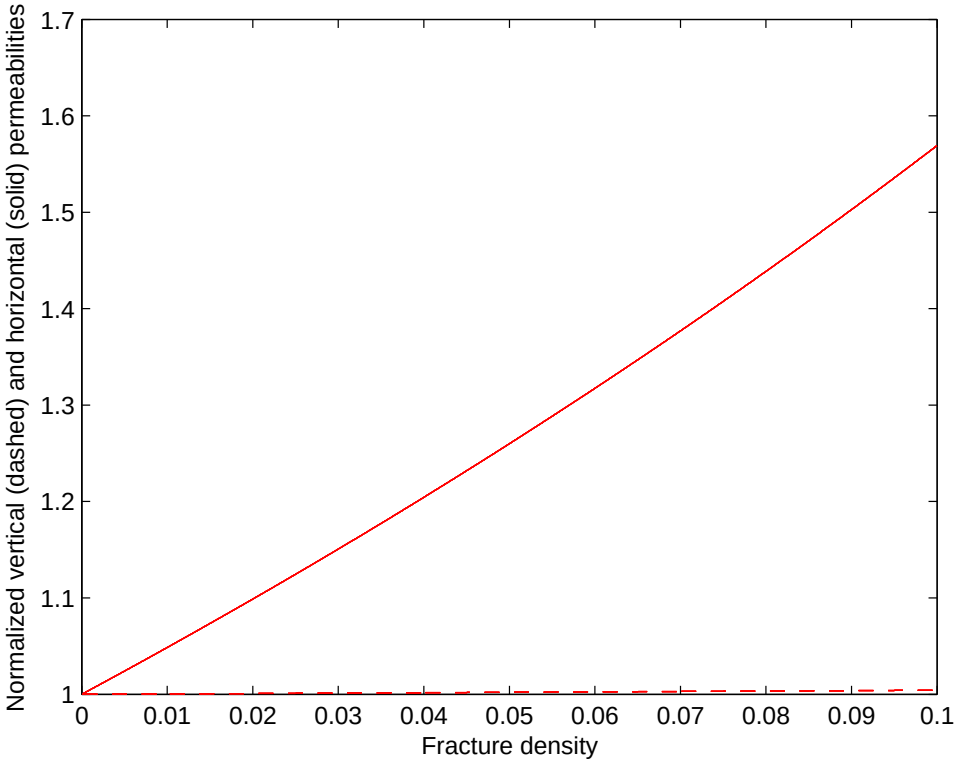


FIG. 2. The effective hydraulic properties of a fractured reservoir.

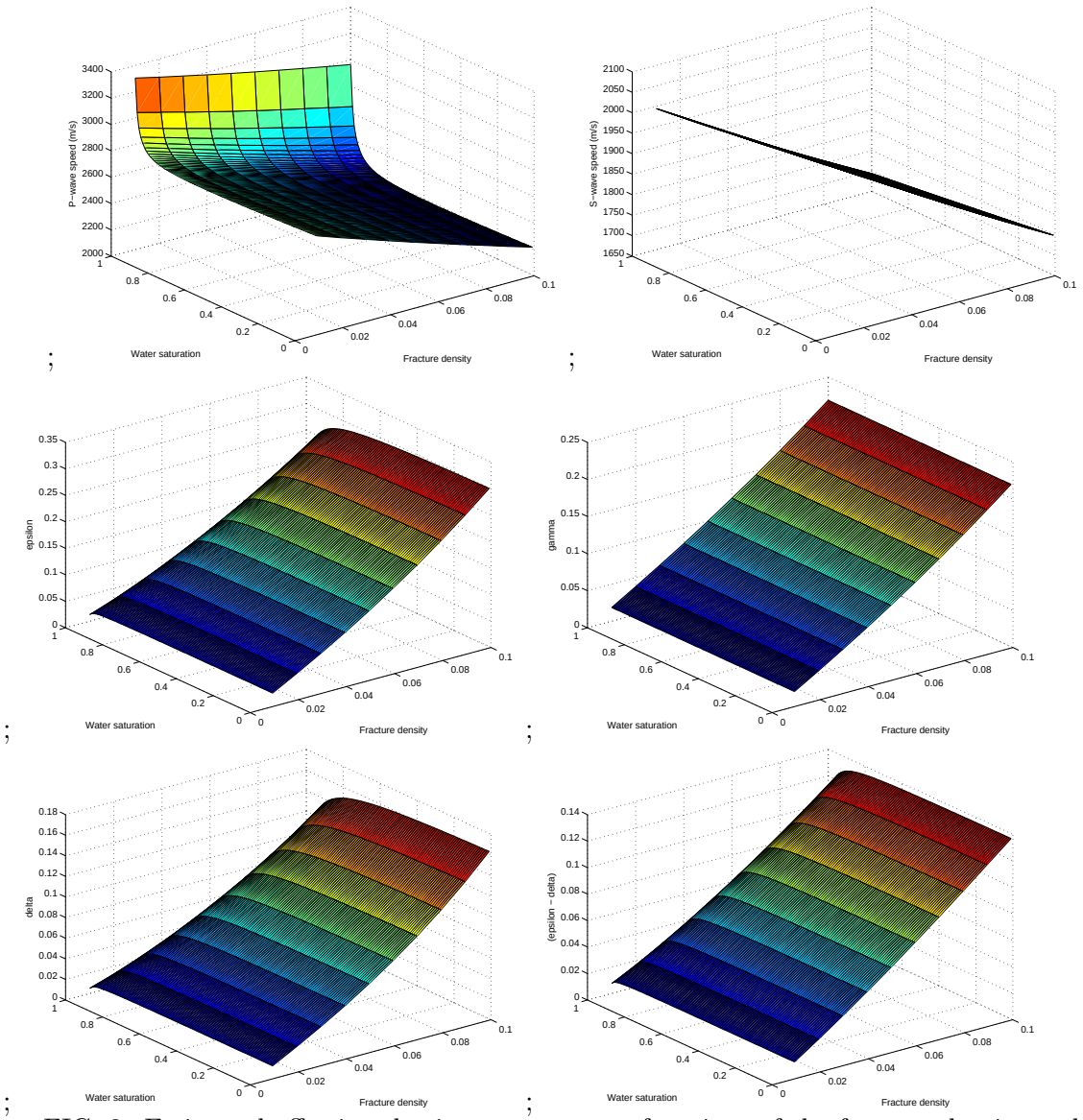


FIG. 3. Estimated effective elastic parameters as functions of the fracture density and water-saturation for a fractured porous medium which is partially saturated with water and gas.

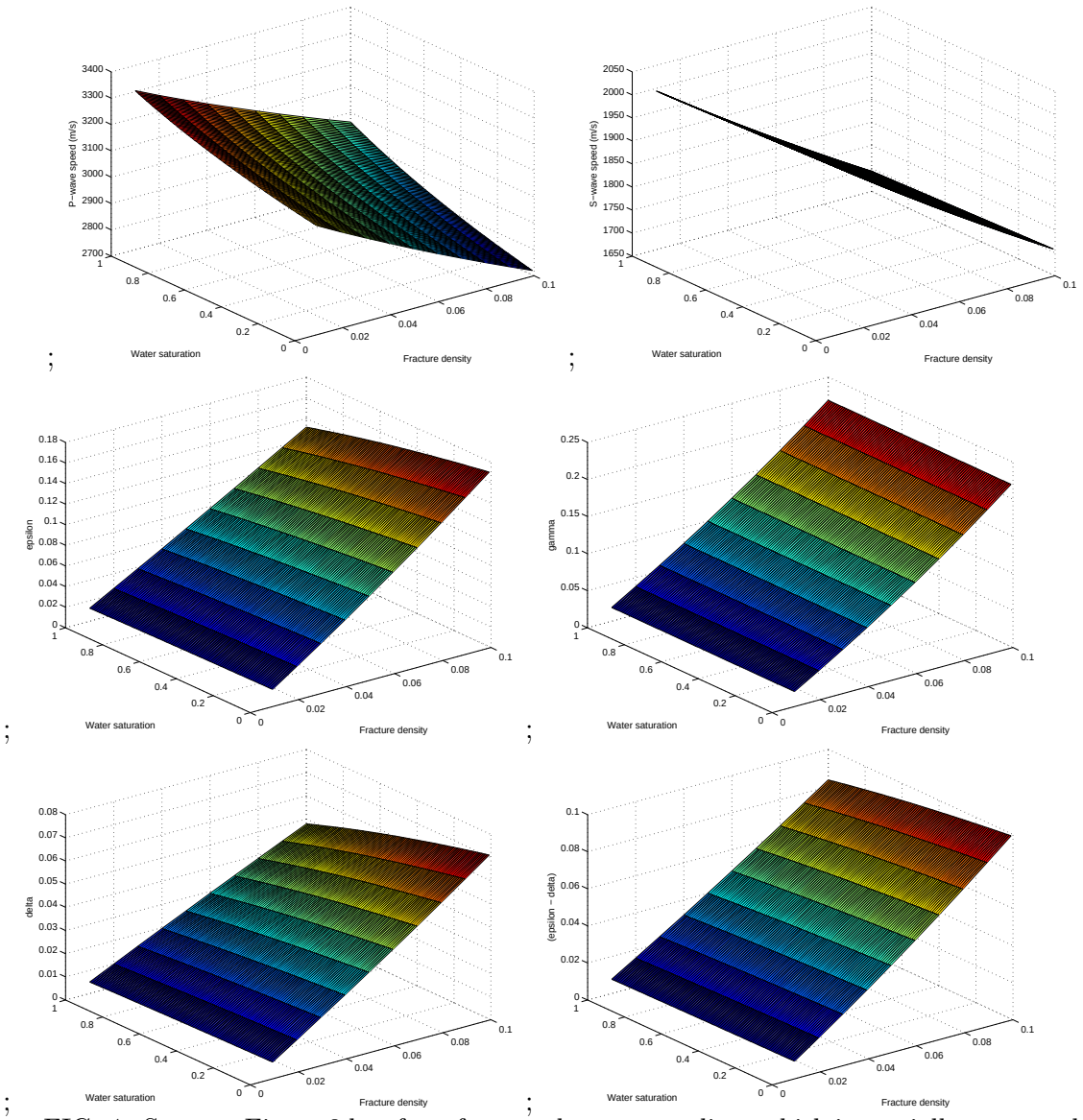


FIG. 4. Same as Figure 3 but for a fractured porous medium which is partially saturated with water and oil.

;

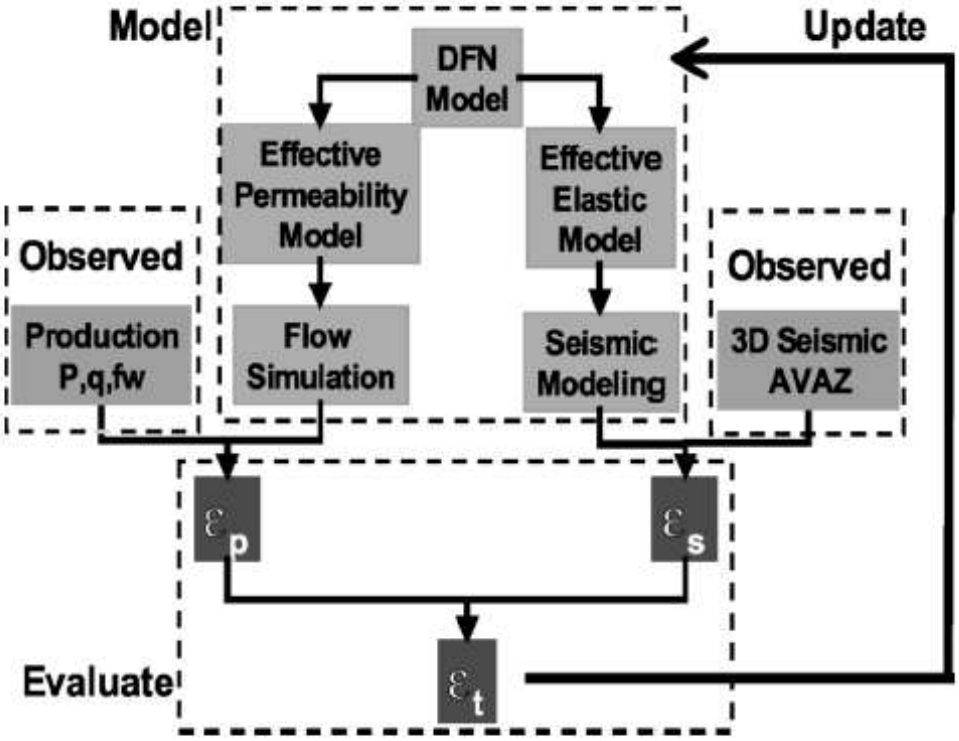


Fig. 1—Workflow schematic.

FIG. 5. The conventional (gradient-based) workflow for joint inversion of seismic and production data used by Will et al. (2005).

;

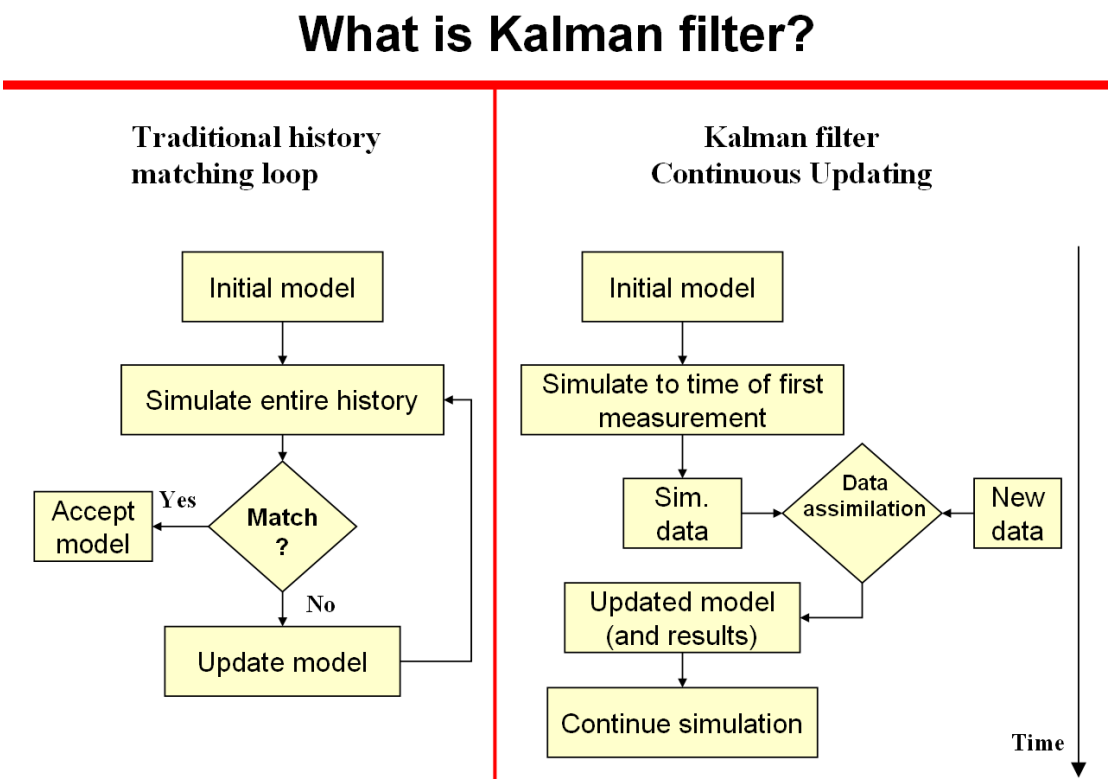


FIG. 6. A comparison between the workflows of conventional and seqential (data assimilation) methods of (seismic) history matching.

;

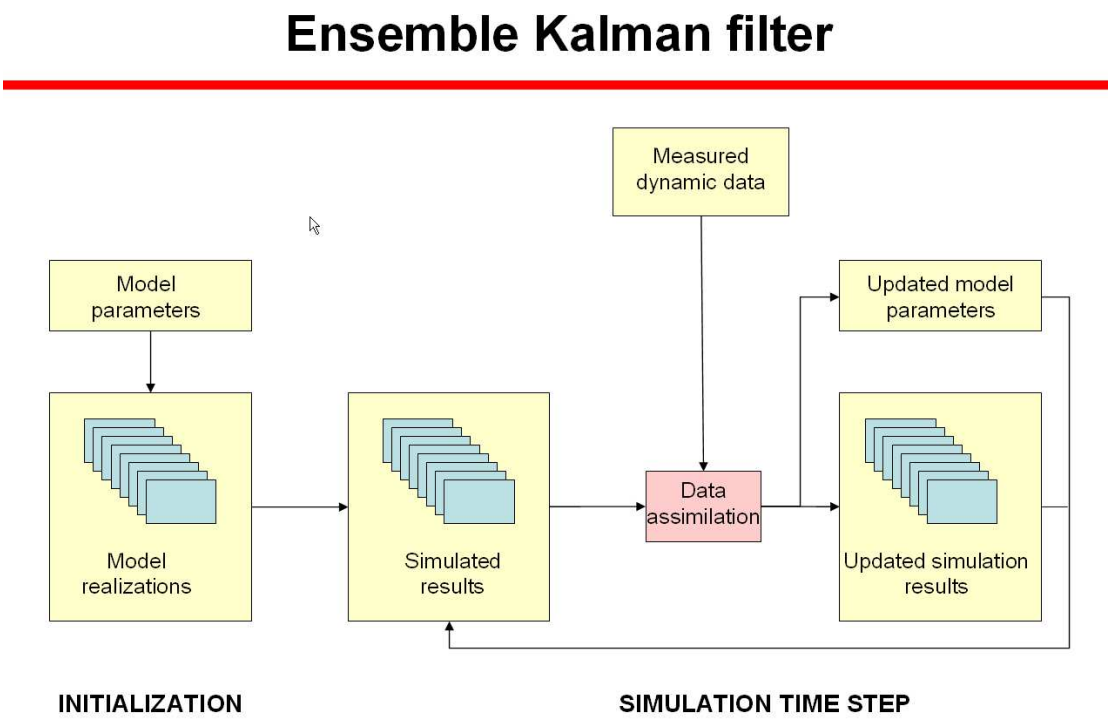


FIG. 7. An illustration of the Ensemble Kalman Filter.

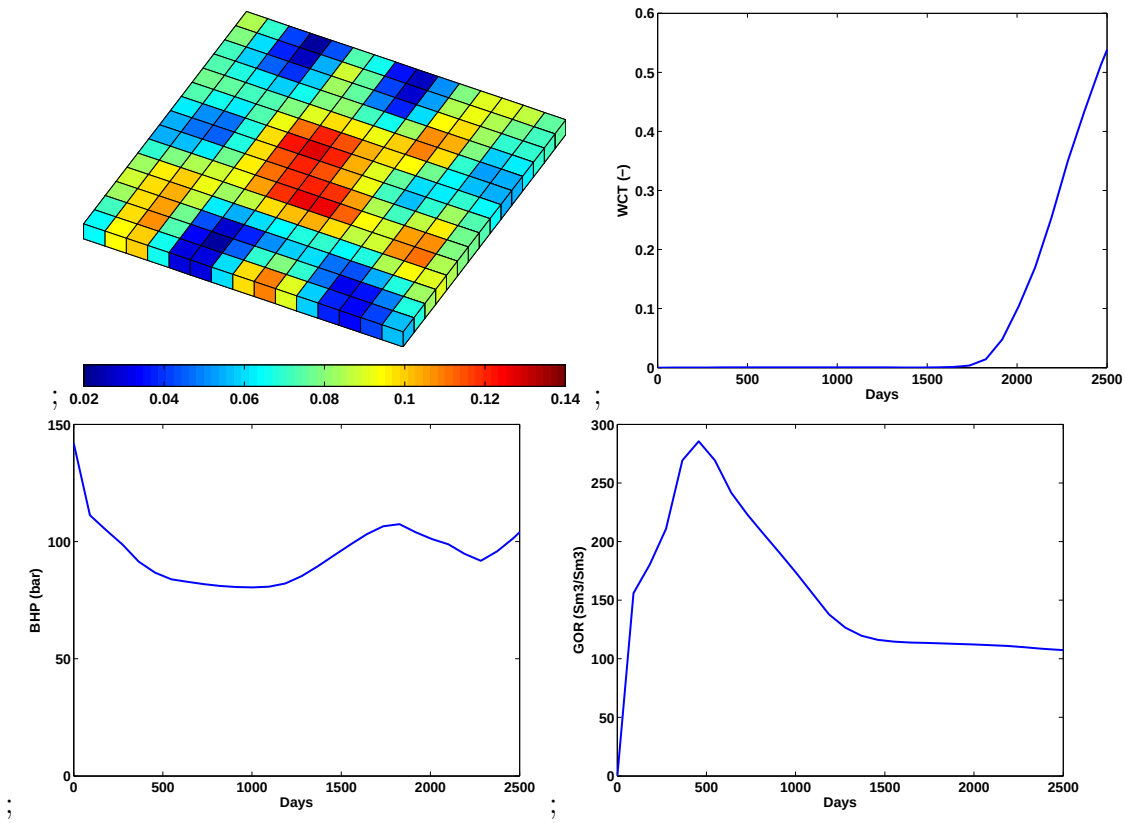


FIG. 8. The reference fracture density field and production data.

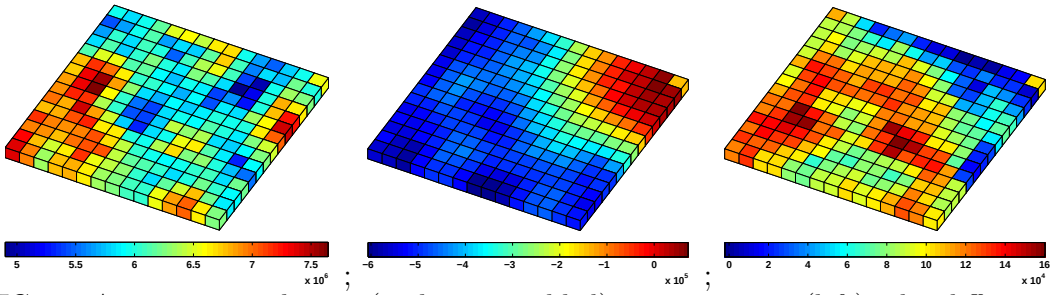


FIG. 9. Acoustic impedances (with noise added) at 0.25 years (left), the difference between the acoustic impedances at 4.5 years and 0.25 years (middle), the difference between the acoustic impedances at 6.5 years and 4.5 years (right).

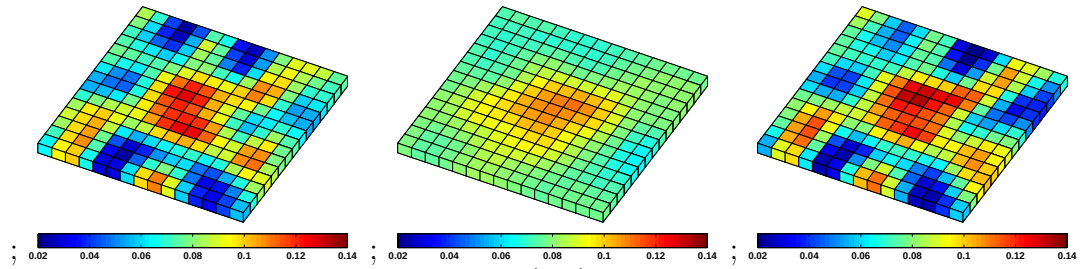


FIG. 10. The true fracture density field (left) vs the estimated fracture density field using production data only (middle) and using both seismic and production data (right).

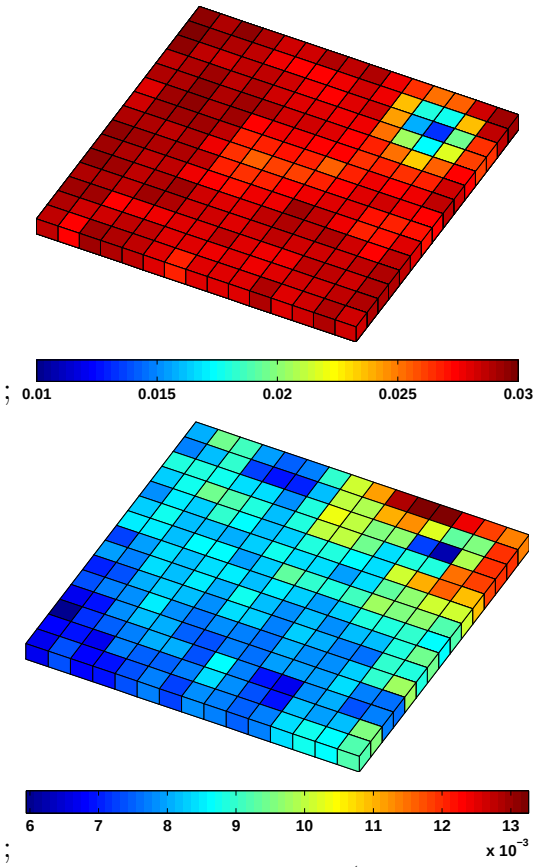


FIG. 11. The uncertainty of the fracture density (represented by the standard deviation) obtained by using the EnKF method with production data only (top) and both seismic and production data (bottom).

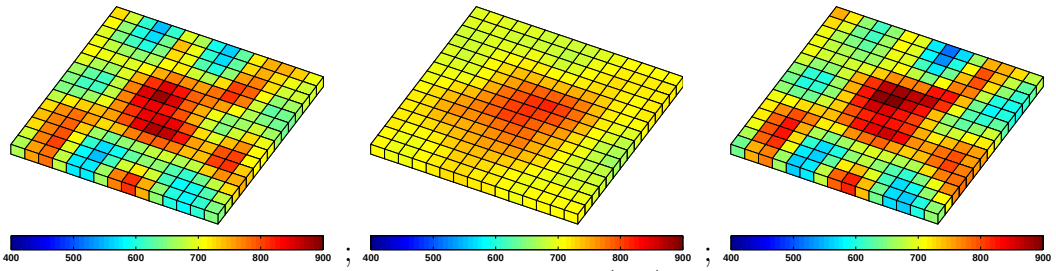


FIG. 12. The true horizontal permeability field (left) vs the estimated horizontal permeability field using production data only (middle) and using seismic and production data (right).

;

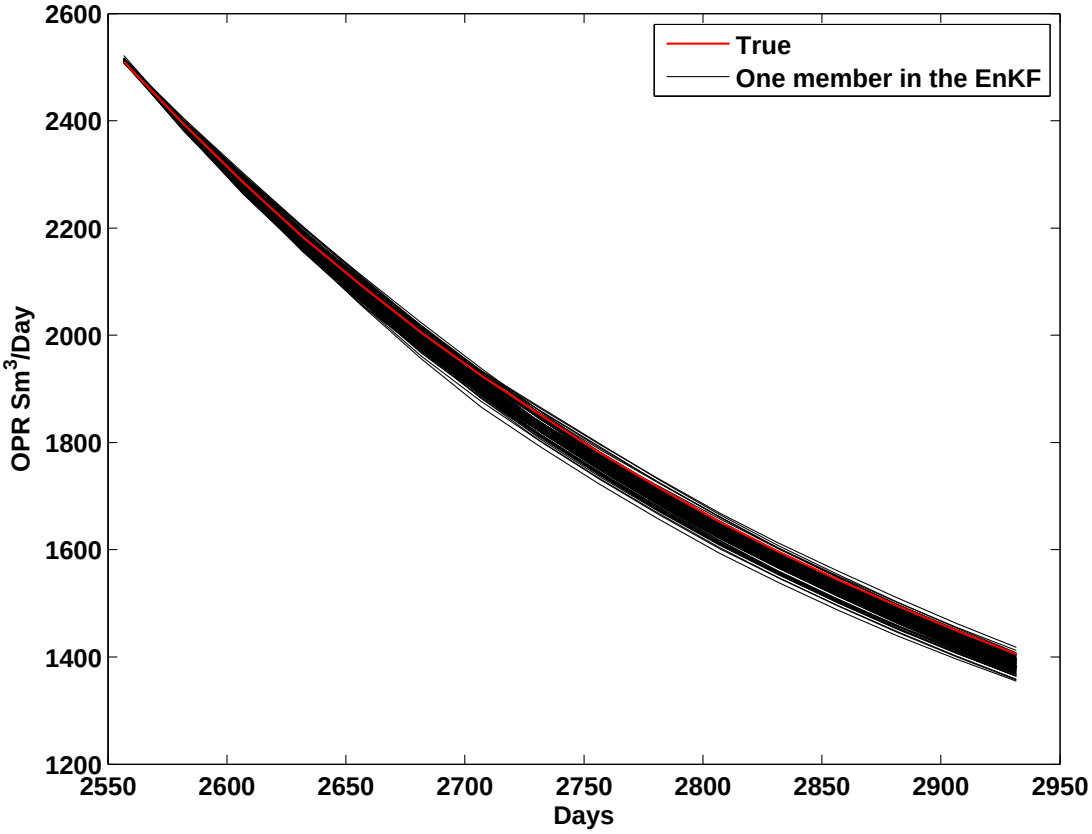


FIG. 13. Forcasting changes in the future oil production.

Report 1

Use of EnKF on a North Sea Field Case *

* Scientific/Technical Report.

Using the Ensemble Kalman Filter on a North Sea field case

Jan-Arild Skjervheim[1], Sigurd Ivar Aanonsen[1], Marco Haverl[2], Geir Evensen[2]

[1] *Centre for Integrated Petroleum Research (CIPR), University of Bergen, Norway*
[2] *Hydro Research Centre, Bergen, Norway*

January 9, 2007

Abstract

This report presents a study of a North Sea field case, where real production data and seismic time-lapse data have been assimilated using the ensemble Kalman filter (EnKF) methodology.

1 Introduction

The goal of this study is to improve the history matching process of a North Sea reservoir by incorporating both production data and 4D seismic data. The following procedure is based on the ensemble Kalman filter (EnKF) methodology developed by Evensen (1994, 2003). The EnKF methodology has shown promising results in reservoir applications and has been reported in several papers, see Nævdal et al. (2003); Gu and Oliver (2004); Gao and Reynolds (2005).

The EnKF methodology has previously been applied to other fields in the North Sea, and it was demonstrated that an improvement of the match to seismic data could be obtained by adjusting the model permeability and porosity fields, see Skjervheim et al. (2005).

2 The EnKF

In this work an ensemble square root formulation is used in the data assimilation, and in the following paragraph the ensemble square root filter is described. The notation follows the one used in Evensen (2004).

2.1 Ensemble representation for P_ψ

We define the matrix holding the ensemble members $\psi_j \in \Re^q$ as,

$$\Psi = (\psi_1, \psi_2, \ldots, \psi_N) \in \Re^{q \times N}, \tag{1}$$

where q is the size of the model state vector and N is number of ensemble members. The ensemble mean is stored in each column of $\overline{\Psi}$ which can be defined as

$$\overline{\Psi} = \Psi 1_N, \tag{2}$$

where 1_N is the $N \times N$ matrix where each element is equal to $1/N$. The ensemble perturbation matrix can then be defined as

$$\Psi' = \Psi - \overline{\Psi} = \Psi(I - 1_N). \tag{3}$$

The ensemble covariance matrix P_ψ can be defined as

$$P_\psi^e = \frac{\Psi'(\Psi')^T}{N-1}. \tag{4}$$

2.2 Measurement perturbations

Given a vector of measurements $d \in \Re^m$, where m is the size of the measurement vector, we can define the N vectors of perturbed measurements as

$$d_i = d + \epsilon_i \;\; i = 1, \ldots, N \tag{5}$$

where ϵ is Gaussian distributed with zero mean and covariance P_ϵ. The perturbed measurements can be stored in the columns of a matrix

$$D = (d_1, d_2, \ldots, d_N) \in \Re^{m \times N}, \tag{6}$$

and the ensemble of perturbations can be stored in the matrix

$$E = (\epsilon_1, \epsilon_2, \ldots, \epsilon_N) \in \Re^{m \times N}. \tag{7}$$

The ensemble representation of the measurement error covariance matrix can then be constructed as

$$P_\epsilon^e = \frac{EE^T}{N-1}. \tag{8}$$

2.3 Ensemble matrices

We now introduce the ensemble of innovation vectors defined as

$$D' = D - H\Psi, \tag{9}$$

where H is a linear measurement operator. The matrix holding the measurements of the ensemble perturbations, $S \in \Re^{m \times N}$, is given as

$$S = H\Psi', \tag{10}$$

and we define the matrix $C \in \Re^{m \times m}$ as

$$C = SS^T + (N-1)P_\epsilon. \tag{11}$$

2.4 Analysis equation

The analysis equation for the standard EnKF is then given by

$$\Psi^a = \Psi + P_\psi H^T (HP_\psi H^T + P_\epsilon)^{-1} D'. \tag{12}$$

The square root algorithm for the EnKF analysis is derived from the traditional analysis equation for the covariance update in the Kalman filter (Kalman (1960)).

$$P_\psi^a = P_\psi^f - P_\psi^f H^T (HP_\psi^f H^T + P_\epsilon)^{-1} HP_\psi. \tag{13}$$

Using the ensemble representation for the error covariance matrix, P_ψ, as defined in Eq. (4), and the definitions of S and C in Section 2.3, the above equation can be expressed as

$$\begin{aligned} \Psi^{a'}\Psi^{a'T} &= \Psi'\Psi'^T - \Psi'\Psi'^T H^T (H\Psi'\Psi'^T H^T + (N-1)P_\epsilon)^{-1} H\Psi'\Psi'^T \\ &= \Psi'(I - S^T C^{-1} S)\Psi'^T. \end{aligned} \tag{14}$$

The analyzed ensemble mean is computed from the standard Kalman filter equation, which can be achieved by multiplication of Eq. (12) from the right with 1_N. Each column is the resulting equation for the mean and is given as

$$\overline{\psi^a} = \overline{\psi^f} + \Psi' S^T C^{-1} (d - H\overline{\psi^f}). \tag{15}$$

2.5 Derivation of the EnSRF algorithm

We now assume that C, defined in Eq. (11), is of full rank. The inverse of C is found from an eigenvalue decomposition and is given as

$$C^{-1} = Z\Lambda^{-1} Z^T, \tag{16}$$

where $Z \in \Re^{m \times m}$ contains the eigenvectors and the diagonal matrix $\Lambda \in \Re^{m \times m}$ holds the eigenvalues. The Eq. (14) can then be written as

$$\Psi^{a'}\Psi^{a'T} = \Psi'(I - S^T Z\Lambda^{-1} Z^T S)\Psi'^T \tag{17}$$

$$= \Psi' \left(I - (\Lambda^{-\frac{1}{2}} Z^T S)^T (\Lambda^{-\frac{1}{2}} Z^T S) \right) \Psi'^T \tag{18}$$

$$= \Psi'(I - Y^T Y)\Psi'^T, \tag{19}$$

$$\tag{20}$$

where we have defined the $m \times N$ matrix Y as

$$Y = \Lambda^{-\frac{1}{2}} Z^T S. \tag{21}$$

Computing the singular value decomposition (SVD) of Y,

$$U \Sigma V^T = Y, \tag{22}$$

where $U \in \Re^{m \times m}$, $\Sigma \in \Re^{m \times N}$ and $V \in \Re^{N \times N}$, the Eq. (19) can be expressed as

$$\begin{aligned} \Psi^{a'} \Psi^{a'T} &= \Psi^{'}(I - (U\Sigma V^T)^T (U\Sigma V^T))\Psi^{'} && (23)\\ &= \Psi^{'}(I - V\Sigma^T \Sigma V^T)\Psi^{'T} && (24)\\ &= \Psi^{'} V (I - \Sigma^T \Sigma) V^T \Psi^{'T} && (25)\\ &= \left(\Psi^{'} V \sqrt{I - \Sigma^T \Sigma}\right) \left(\Psi^{'} V \sqrt{I - \Sigma^T \Sigma}\right)^T. && (26) \end{aligned}$$

Hence, a solution for the analysis ensemble perturbations is

$$\Psi^{a'} = \Psi^{'} V \sqrt{I - \Sigma^T \Sigma} \Theta^T, \tag{27}$$

where Θ is a random orthogonal matrix. A random redistribution is in some cases necessary, and a further discussion can be found in Evensen (2004). In this work an efficient subspace ensemble square root implementation is used when seismic data are assimilated, and a more detailed discussion of the subspace method can be found in Evensen (2004) and Skjervheim et al. (2006). Notice that when the seismic data are given as a difference between two surveys, a combination of the ensemble Kalman filter and the ensemble Kalman smoother has to be applied, see Skjervheim et al. (2006).

3 Field case

3.1 Simulation model

The study is based on an eclipse simulation model of a field located in the North Sea. The reservoir model consists of 17×39 cells areally and 13 layers. In this model layers 1-6 is called Upper Part, layers 7-12 is called Middle Part, and layer 13 is called Lower Part. There are three oil producers and one water injector in the model. The producers P-1 (start 17 Oct. 1993), P-2 (start 20 Sept. 1994), and P-3 (start 14 Nov. 2000); and the water injector I-1 (start 23 Apr. 1994).

The global average rate of water circulation has been relatively large, and this, together with the good reservoir properties, has made the oil production and the reservoir pressure sensitive to the daily injection rates. Starting at an initial pressure of 245 barsa, the reservoir pressure has been varying between 190 and 260 barsa (well above the bubble point pressure), and these fluctuations, together with spatial pressure variations, may have influenced the time-lapse seismic survey.

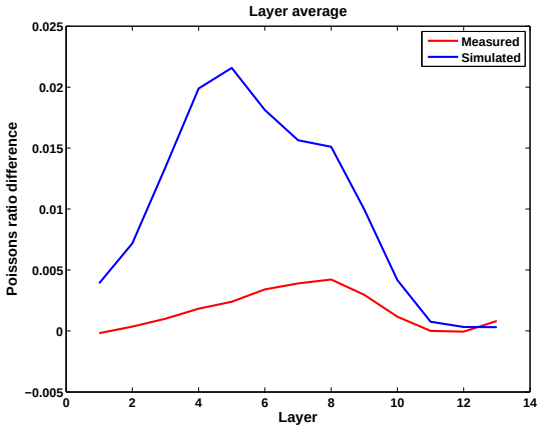


Figure 1: Measured and simulated Poisson's ratio averaged over each simulation model layer.

3.2 Rock physics model and inverted seismic data

In the history matching process we have focused on matching the Poisson's ratio difference between two time steps of several years of production. Thus, the purpose of the rock physics model is to convert simulated saturations and pressures to Poisson's ratio, in order to compared them to the observed counterparts.

The rock physics model, developed by others, is applied in the built-in petroelastic module in Eclipse, consisting mainly of the Gassmann equation to derive the effective rock properties and the Batzle & and Wang equations to derive fluid properties. However, dependencies of elastic parameters, and optionally fluid parameters have to be included and represent a crucial task. These dependencies, which are the dry-rock bulk and shear modulus versus porosity, are derived from laboratory core plug measurements. In the measured data there is a very good consistency between the inverted seismic data and the initial water oil contact.

Difference in Poisson's ratio from the base line survey to the monitor survey averaged over each simulation model layer is plotted in Figure 1. Measured and simulated Poisson's ratio differences in one of the model layers is plotted in Figure 2, which shows that the size of the seismic signal for the simulated data are much larger than for the measured data.

Theoretically, only pressure changes should influence the data below the original oil-water contact, and since pressure changes only have a limited influence on the Poisson's ratio, the water-flooded areas are expected to have an increase in Poisson's ratio.

When Poisson's ratio reflect water movement, we anticipate that water has mainly flooded the lower layers of the oil zone (i.e Middle Part) before it has moved up to the producers, which are only perforated in the Upper Part. However this is not the case in the current model, where both the Middle and Upper Part are flooded. Thus, the seismic data may indicate that the communication between the Middle and Upper Part, or vertical communication in general, is lower than the current model. However, because of the uncertainties associated

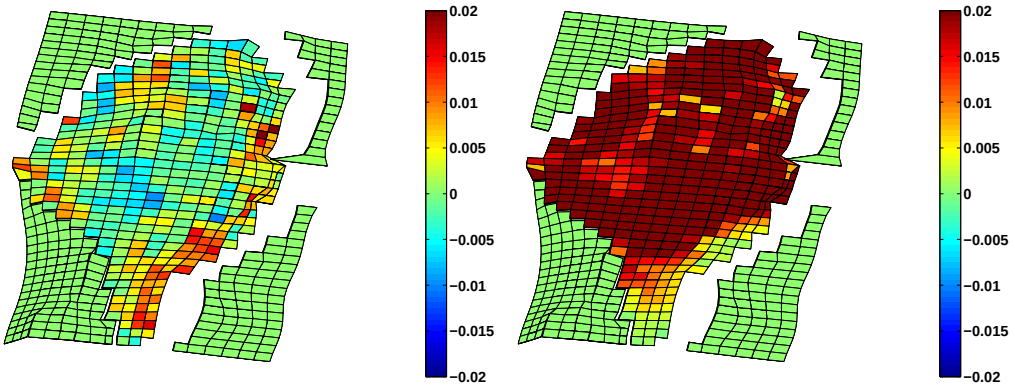


Figure 2: Measured (left) and simulated (right) Poisson's ratio difference in layer 3 of simulation grid.

with seismic data, we should be cautious when drawing conclusions.

3.3 Data assimilation

Statistical information from the geological model is used to generate the initial ensemble of 100 realizations of a Gaussian random field. In this model, information about correlation lengths and orientation, vertical layer correlation and correlation between permeability and porosity are given.

In this field case, the dynamic variables pressures and saturations are updated, as well as the static parameters porosity and permeability. The well variables BHP and WCT predicted by the model are used in the data assimilation to update the model state at the analysis times.

Two cases of EnKF updating were run: i) With production data only, i.e., BHP and WCT for 3 wells over a 12 years period with measurements approximately every month, called "EnKF prod"; and ii) Poisson's ratio difference between the base line and monitor surveys in addition to the production data, called "EnKF prod 4D". The simulation run of the base case model is called "Base case".

Well data errors (standard deviations) used in this study are 4 bars for the bottom-hole pressure and 0.025 for the water cut. A constant standard deviation of 0.007 was applied for the Poisson's ratio. No spatial correlations were applied for the seismic data. The measurement errors are drawn from a Gaussian distribution with mean zero and the above standard deviations.

	Base case	EnKF mean prod	EnKF mean prod 4D
RMS	3.06	3.16	2.65

Table 1: RMS values for the seismic data (Poisson's ratio difference).

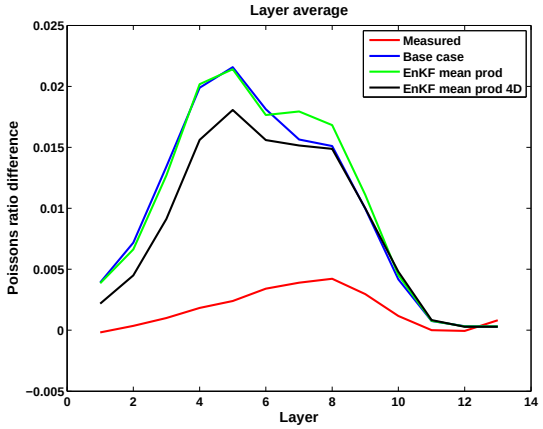


Figure 3: Measured and simulated Poisson's ratio for the two EnKF runs averaged over each simulation model layer.

3.4 Results

RMS deviation and correlation between simulated and measured Poisson's ratio for the base case and the two EnKF runs are listed in Table 1. Difference in Poisson's ratio averaged over each simulation model layer for the two EnKF runs is plotted in Figure 3. The RMS results and the results in Figure 3 are based on one simulation of the entire history using the final estimated mean permeability and porosity, therefore called "EnKF mean prod" and "EnKF mean prod 4D". As can be seen, a minor reduction in the RMS value and a slightly smaller mismatch for the seismic data is obtained in the EnKF run when both production data and seismic data are assimilated.

Considering the vertical trend in the simulated Poisson's ratio of the base case model, the main increase occurs from layer 2 to 4, then the trend flattens out, and decreases again from layer 5-6. For the measured data, there is a more gradual increase all the way down to the water-oil contact (see Figure 3). It is clear that one way of matching the measured data is to adjust the model such that the top part of the reservoir becomes less flooded than the lower part. However, it is also possible that the overall level of either the simulated or measured data is wrong; the simulated data because of an incorrect rock physics model, and the measured data because of the uncertainties in the inversion process. The reason that the simulated data start to decrease is, that some layers are below the initial water-oil contact. These parts increase in size, and although there is a good correlation between the measured Poisson's ratio difference and the initial water-oil contact, this correlation is not large enough to achieve the same trend as in the model.

Since water is injected in the lower part of the reservoir (Middle Part), and the producers

are perforated in the upper part of the reservoir (in Upper Part), the main water flood may be "shifted downwards" by modifying the permeability or the communication between the Upper and Lower Part. However, by doing that, the effective volume flooded in the early period will be reduced, and the water will break through earlier. So by tuning the model to match the seismic data, the match to production data is expected to be worse, and this also turned out to be the case in the EnKF run where both production data and seismic data were assimilated.

Production data for the two EnKF runs are compared with the base case in Figure 4. As can be seen the EnKF run with production data only is matching the history better than the base case. For the EnKF run using both production data and 4D data, a clear increase in the mismatch term of the production data is seen at the analysis time where the seismic data are assimilated. This is as expected, and it seems to be difficult to obtain a good match of both production data and seismic data simultaneously. As mentioned, a major uncertainty in the process is the rock physics model, especially the pressure effects, which are not properly understood. Still, it is not believed that using a different rock physics model will change the results significantly. However, in future work it would be interesting to introduce an uncertainty to parameters in the rock physics model which are sensitive to the seismic signal.

The final permeability and porosity fields after 12 years of assimilation for the two EnKF runs were used to re-initialize the base case simulation. The results are plotted in Figure 5, and show the effect of the updated permeability and porosity fields on a single integration. The forward run of the final estimated fields from the EnKF run with production data only shows a better match than the base case model. Especially an improvement in well P-3 is obtained. Matching the seismic data results in a too early water break through in the wells P-1 and P-2, and the EnKF method has tried to compensate for this by increasing the pore volume of the lower parts of the reservoir. This is seen in Figure 6, where we have plotted final ensemble mean permeability and porosity for the two EnKF runs averaged over each simulation model layer.

Figures 7 and 8 show the final ensemble mean of permeability and porosity in two layers for the two EnKF runs. The results in Figures 5-6 show that relatively large modifications have been obtained during the assimilation, and the estimation results are quite different for the two EnKF runs.

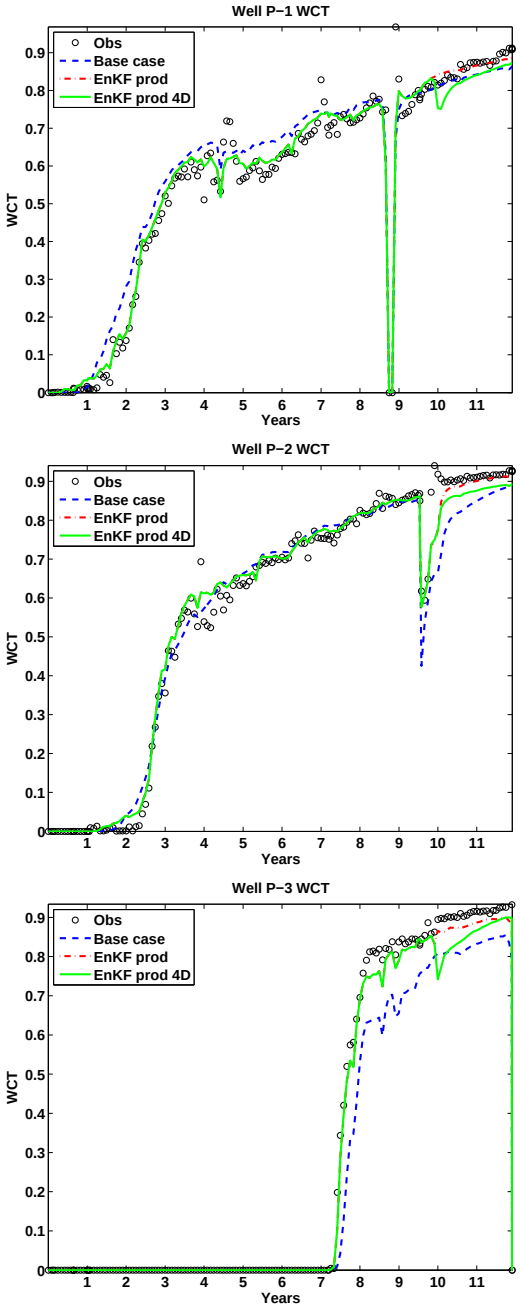


Figure 4: Production profiles for Wells P-1 (top), P-2 (middle), and P-3 (bottom). For base case, EnKF with production data, and EnKF with production data and 4D data.

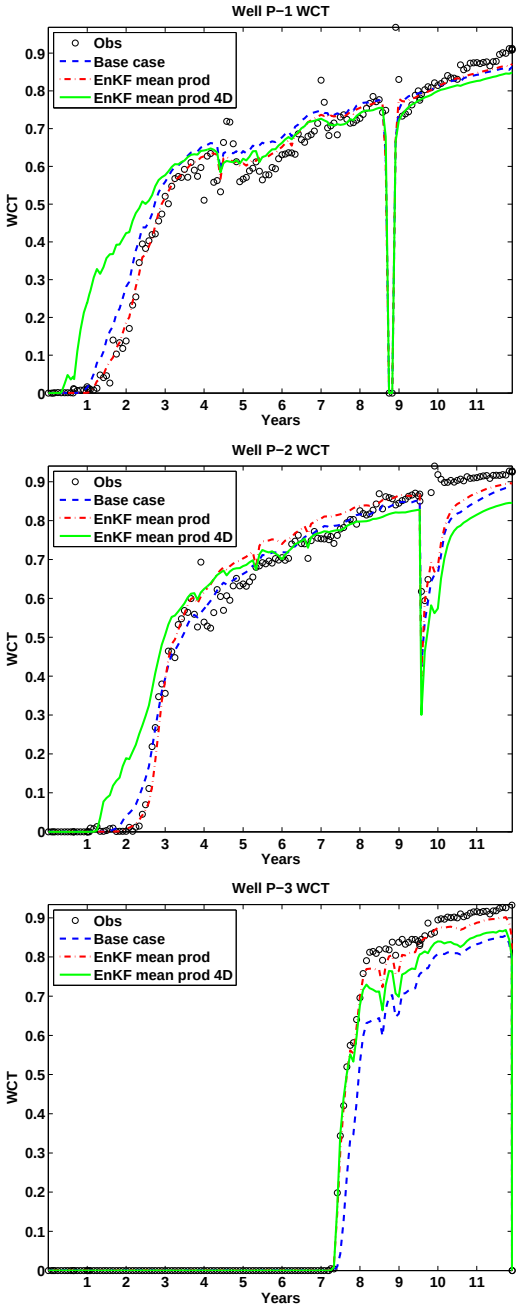


Figure 5: Production profiles for Wells P-1 (top), P-2 (middle), and P-3 (bottom). For base case, EnKF with production data, and EnKF with production data and 4D data. The results are based on the final estimated permeability and porosity fields.

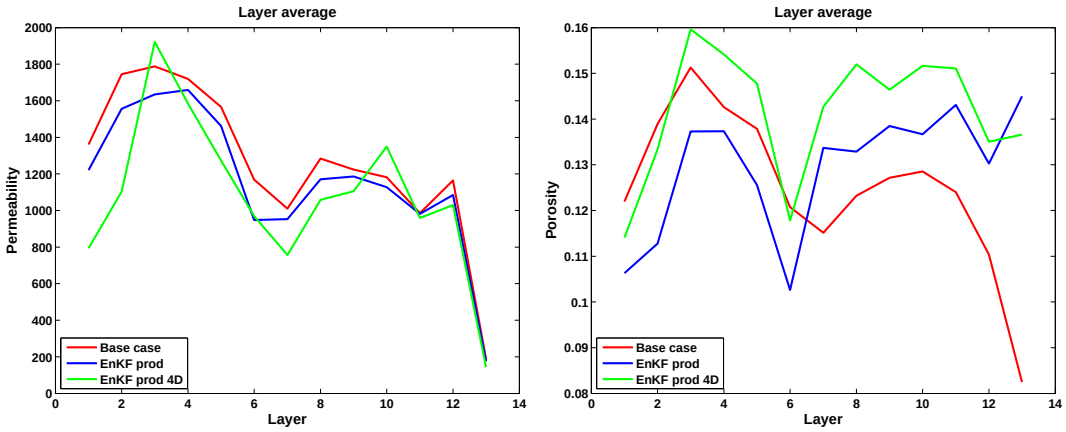


Figure 6: The final ensemble mean of permeability (left) and porosity (right) for the two EnKF runs averaged over each simulation model layer.

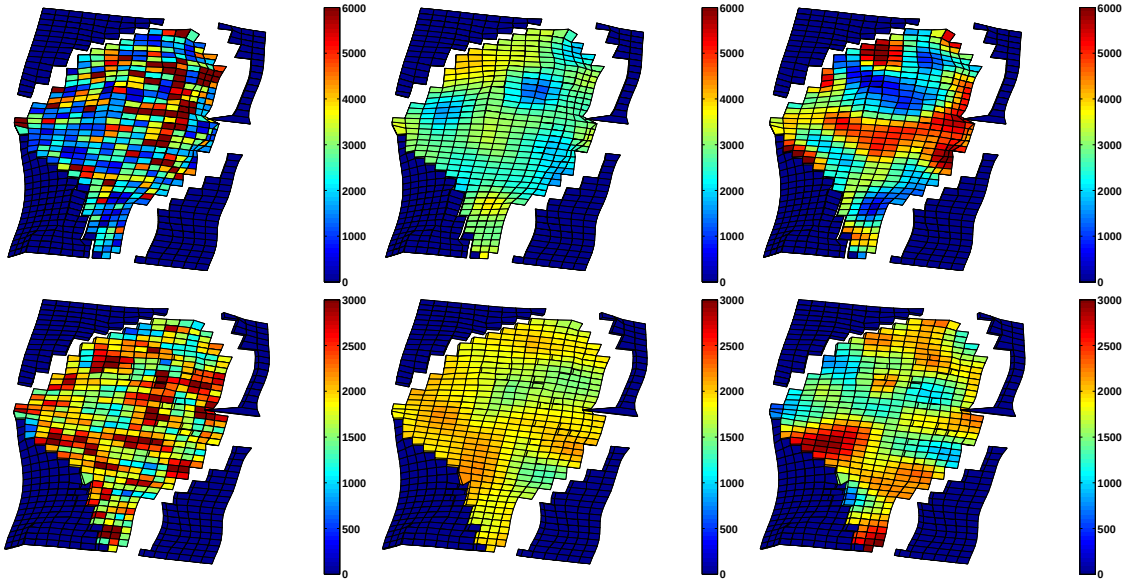


Figure 7: EnKF results for permeability in layers 3 (top) and 12 (bottom). Base case (left), EnKF prod (middle) and EnKF prod 4D (right).

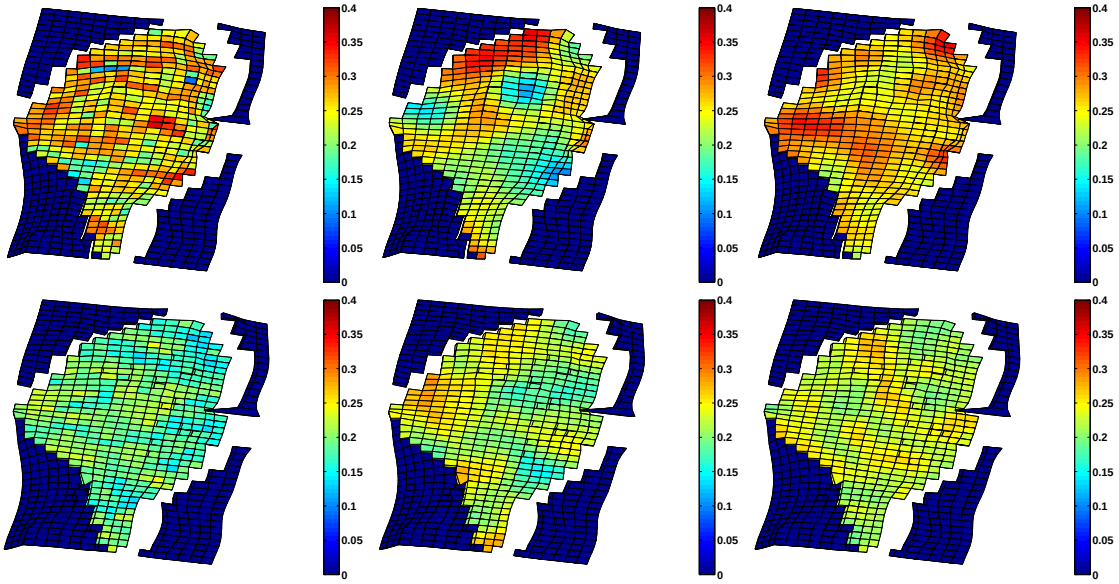


Figure 8: EnKF results for porosity in layers 3 (top) and 12 (bottom). Base case (left), EnKF prod (middle) and EnKF prod 4D (right).

4 Conclusion

A method based on the ensemble Kalman filter was used on a real field reservoir to improve model parameters such as porosity and permeability by assimilating both production data (WCT,BHP) and seismic data (Poisson's ratio).

To match both production data and the time-lapse seismic data the EnKF method posed a combination of permeability modifications and an increased pore volume in the lower part of the reservoir. In this history matching study it has been difficult to obtain a good match of both production data and seismic data simultaneously, and the uncertainty in data and validity of the obtained results should be investigated further.

References

G. Evensen. The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dynamics*, 53:343–367, 2003.

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

G. Evensen. Sequential data assimilation with nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, 99:10143–10162, 1994.

G. Gao and A.C Reynolds. Quantifying the uncertainty fo the PUNQ-S3 problem in a bayesian setting with the RML and EnKF. *SPE reservoir simulation symposium (SPE 93324)*, 2005.

Y. Gu and D. S. Oliver. History Matching of the PUNQ-S3 Reservoir Model Using the Ensemble Kalman Filter. *SPE Annual Technical Conference and Exhibition, SPE 89942*, 2004.

R. E. Kalman. A new approach to linear filter and prediction problems. *J. Basic. Eng.*, 82: 35–45, 1960.

G. Nævdal, G. L. Johnsen, S. I. Aanonsen, and E. Vefring. Reservoir monitoring and continuous model updating using the ensemble Kalman filter. *SPE Annual Technical Conference and Exhibition (SPE 84327)*, 2003.

J. A. Skjervheim, S. I. Aanonsen, and G. Evensen. Ensemble Kalman filter with time difference data. *Submitted to Computational Geosciences*, 2006.

J.A. Skjervheim, G. Evensen, S.I Aanonsen, B.O Ruud, and T.A. Johansen. Incorporating 4D Seismic Data in Reservoir Simulation Models Using Ensemble Kalman Filter. *SPE ATCE (SPE 95789)*, 2005.

Report 2

Sequential Gaussian Simulation with Block Kriging *

* Scientific/Technical Report.

Sequential Gaussian Simulation With Block Kriging

Jan-Arild Skjervheim[1], Sigurd Ivar Aanonsen[1]

[1] *Centre for Integrated Petroleum Research (CIPR), University of Bergen, Norway*

July 4, 2005

1 Introduction

Seismic data have a much higher areal resolution than the simulation model, while the opposite is the case for the vertical resolution. In addition, the geological model and well log data show different degrees of spatial resolution. When coupling the models in an ensemble Kalman filter (EnKF) methodology, this effect has to be taken into account. This includes upscaling and downscaling of properties and uncertainties, as well as data correlations. In this report we have performed a simulation study of a specific downscaling algorithm called Sequential Gaussian Simulation with Block Kriging (SGSBK), see Behrens et al. (1998) and Tran et al. (2001).

2 Sequential Gaussian Simulation With Block Kriging (SGSBK) method

Let Z be a multivariate Gaussian random variable with

$$\mathrm{E}(Z) = 0 \tag{1}$$

$$\mathrm{Var}(Z) = 1. \tag{2}$$

In a discretized 3D geologic model, let V denote a coarse block that comprises n point values z_i, $i = 1, \ldots n$, and let $\bar{z}_V$ represent the linear block average

$$\bar{z}_V = \sum_{i=1}^{n} \alpha_i z_i, \quad \text{where} \quad \sum_{i=1}^{n} \alpha_i = 1, \tag{3}$$

where the weights are set to $1/n$, such that each fine-scale grid cell contributes equally to the linear coarse-block average.

In this study the aim is to simulate a realization of Z in a geostatistically consistent manner that honor the block average constraint in Eq. (3) in addition to prescribed histogram and spatial covariance models and any conditioning known data. The simulation is performed by a Sequential Gaussian Simulation With Block Kriging method.

From the sequential Gaussian simulation paradigm we have that the joint probability density function of a random variable $X = [X_1, \ldots, X_m]$ can be expressed by

$$g(x) = g(x_1, \ldots, x_m) = g(x_m|x_{m-1}, \ldots, x_1) \cdots g(x_2|x_1)g(x_1), \tag{4}$$

where we can simulate the joint distribution as

- $x_1 \sim g(x_1)$
- $x_2 \sim g(x_2|x_1)$
- Continue until
- $x_m \sim g(x_m|x_{m-1}, \ldots, x_1)$

In the special case when X is a Gaussian process it can be shown that the conditional distributions $g(x_2|x_1), \ldots, g(x_m|x_{m-1}, \ldots, x_1)$ are Gaussian processes too and the mean and the variance can be expressed by kriging systems.

Let now z_o be the value to be simulated at the current grid cell, and let $\{z_{SGS}\}$ be the set of previously simulated values. The mean and the variance of the conditional distribution $g(z_o|\{z_{SGS}\}, \bar{z}_V)$ is then given by the kriging system, see Behrens et al. (1998):

$$\mu_{o,BK} = \sum_{i \in \{SGS\}} \lambda_i z_i + \lambda_V \bar{z}_V \tag{5}$$

$$\sigma^2_{o,BK} = 1 - \sum_{i \in \{SGS\}} \lambda_i C_{oi} - \lambda_V \bar{C}_{oV} \tag{6}$$

where

- λ_i is the kriging weight corresponding the the previously simulated value z_i
- λ_V is the kriging weight given to the coarse-block average constraint $\bar{z}_V$
- C_{oi} is the point to point covariance between the values at cell o and cell i.
- C_{oV} is the point to block covariance between the value at cell o and the block value at V
- $\{SGS\}$ denotes the set of previously simulated cells

The kriging weights in Eqs (5) and (6) are the unknowns in the following equations

$$\sum_{j\in\{SGS\}} \lambda_j C_{ij} + \lambda_V \bar{C}_{iV} = C_{oi} \text{ for } i \in \{SGS\} \tag{7}$$

$$\sum_{j\in\{SGS\}} \lambda_j \bar{C}_{jV} + \lambda_V \bar{\bar{C}}_{VV} = \bar{C}_{oV} \tag{8}$$

The point-to-block and block-to-block covariances can be approximated by

$$\bar{C}_{iV} = \text{Cov}(z_i, \bar{z}_V) = \text{Cov}(z_i, \frac{1}{n}\sum_{j=1}^{n} z_j) = \frac{1}{n}\sum_{j=1}^{n} \text{Cov}(z_i, z_j) \tag{9}$$

$$\bar{\bar{C}}_{VV} = \text{Cov}(\bar{z}_V, \bar{z}_V) = \text{Cov}(\frac{1}{n}\sum_{j=1}^{n} z_j, \frac{1}{n}\sum_{k=1}^{n} z_k) = \frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n} \text{Cov}(z_j, z_k) \tag{10}$$

The Sequential Gaussian Simulation with Block Kriging (SGSBK) algorithm is then given as follows:

Algorithm

1. Transform the known fine-scale data to normal-score data with zero mean and unit variance; transform the coarse-scale data to normal-score data with zero mean and variance of $\bar{\bar{C}}_{VV}$ as given in Eq. (10)
2. Define a random path visiting all grid nodes to be simulated
3. At each node, construct the Gaussian conditional density function whose mean,$\mu_{o,BK}$, and variance, $\sigma_{o,BK}$, are given by Eqs. (5) and (6), using all known fine-scale normal-score data and all previously simulated values that are within the search neighborhood
4. Draw a value from the above Gaussian conditional density functions: $z_o \sim \mathcal{N}(\mu_{o,BK}, \sigma^2_{o,BK})$
5. Return to Step 3. until all nodes have been visited
6. Back-transform to the simulated normal-score values to original data space

3 2D Example

In this example we assume that the true field is Gaussian distributed given by

$$Y_t \sim \mathcal{N}(100, 10^2 C), \tag{11}$$

with mean, $\mu = 100$, variance, $\sigma^2 = 10^2$, and where C contains a spherical covariance structure. The two dimensional true field consists of 64×64 grid cells, and the correlation range

is set to 20 grid cells in both directions. The true field is plotted in Figure 2 and Figures 3 and 4 shows the estimated histogram and variogram, respectively. The statistics of the true random field is given in Table 1.

The true field is then transformed (linear averaging) to a 8×8 coarse-scale, plotted in Figure 1. From now on we assume that we only know the coarse-scale random field and the fine-scale covariance C. Before the SGSBK method is applied we transform the coarse-scale random field to the normal-score domain with zero mean and variance $\bar{\bar{C}}_{VV}$. Figure 2 shows the random field generated by the SGSBK algorithm, and Figures 3 and 4 shows the estimated histogram and variogram respectively. The statistics is given in Table 1. As can be seen the SGSBK algorithm reproduce the features and the statistical properties of the true 2D field fairly good, and the correlation coefficients between the coarse-grid constraints and the corresponding averages from fine-grid values are 0.99.

4 3D Example

In this example we assume that the true field is a three dimensional cube of size $16 \times 16 \times 16$ that is Gaussian distributed with zero mean and unit variance

$$Y_t \sim \mathcal{N}(0, C), \tag{12}$$

where C is a spherical covariance structure with range equal 10 grid cells in all three directions. The true field is plotted in Figure 6 and Figures 7 and 8 shows the estimated histogram and variogram respectively. The statistics of the true field is shown in Table 2.

The true field is then transformed (linear averaging) to two different coarse-scale fields of size $8 \times 8 \times 8$ and $4 \times 4 \times 4$, plotted in Figure 5. These two fields are then normal-score transformed and used as input to the SGSBK algorithm.

Figure 6 shows the random fields generated by the SGSBK algorithm, when either the $8 \times 8 \times 8$ coarse-scale field or the $4 \times 4 \times 4$ coarse-scale field is used the as the constraint. The estimated histogram and variogram are shown in Figures 7 and 8, and the statistics of the fields is given in Table 2.

The results show that the SGSBK algorithm reproduce the features and the statistical properties of the true 3D field satisfactory. The correlation coefficients between the coarse-grid constraints $8 \times 8 \times 8$ and $4 \times 4 \times 4$ and the corresponding averages from fine-grid values are 0.98 and 0.97, respectively .

5 Conclusion

A simulation study of a specific downscaling algorithm, called Sequential Gaussian Simulation with Block Kriging (SGSBK), is performed.

	Mean	Variance
True	100.82	101.08
SGSBK	100.82	98.22

Table 1: True statistics compared with the statistics estimated from the SGSBK field for the 2D case

	Mean	Variance
True	0.00	1.00
SGSBK constraint to $4 \times 4 \times 4$	0.03	0.91
SGSBK constraint to $8 \times 8 \times 8$	0.01	0.95

Table 2: True statistics compared with the statistics estimated from the SGSBK field for the 3D case

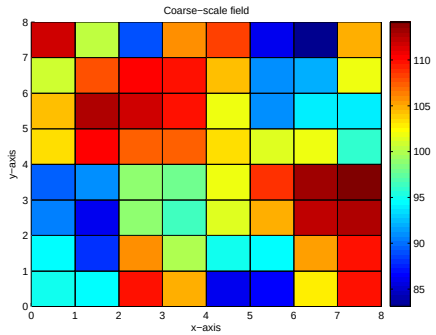


Figure 1: Coarse-scale field

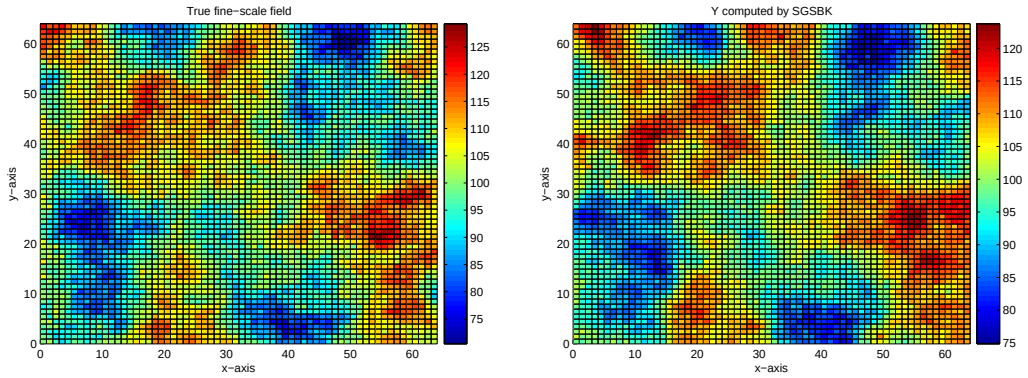


Figure 2: True field (left) compared with the field generated by the SGSBK method (right).

The SGSBK results show that the method generate fine-scale geostatistical models constrained to coarse-scale average constraint in a properly way, that honor the prescribed histogram and the spatial covariance model.

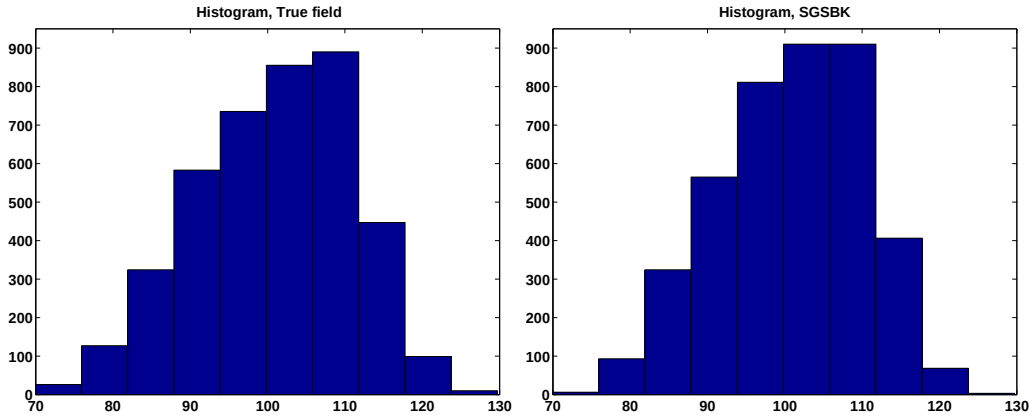


Figure 3: Estimated histogram of the true field (left) compared with the estimated histogram of the field generated by the SGSBK method (right).

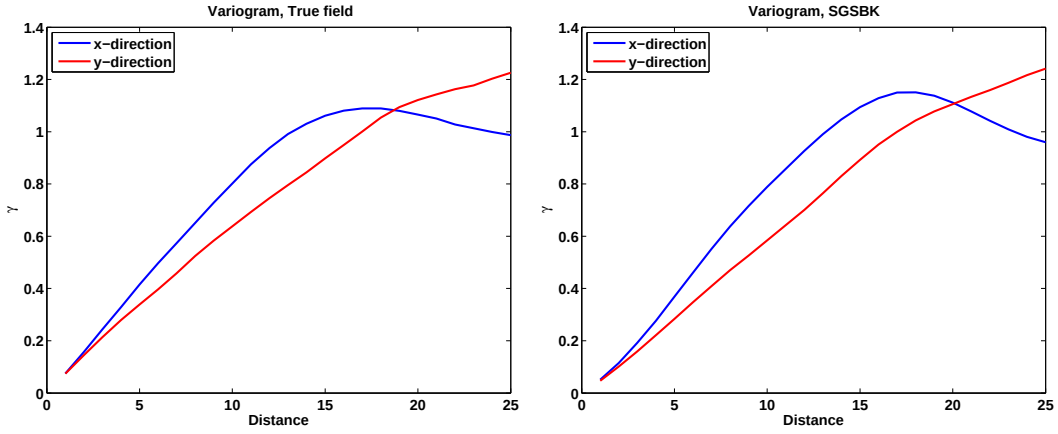


Figure 4: Estimated variogram from the true field (left) compared with the estimated variogram of the field generated by the SGSBK method (right).

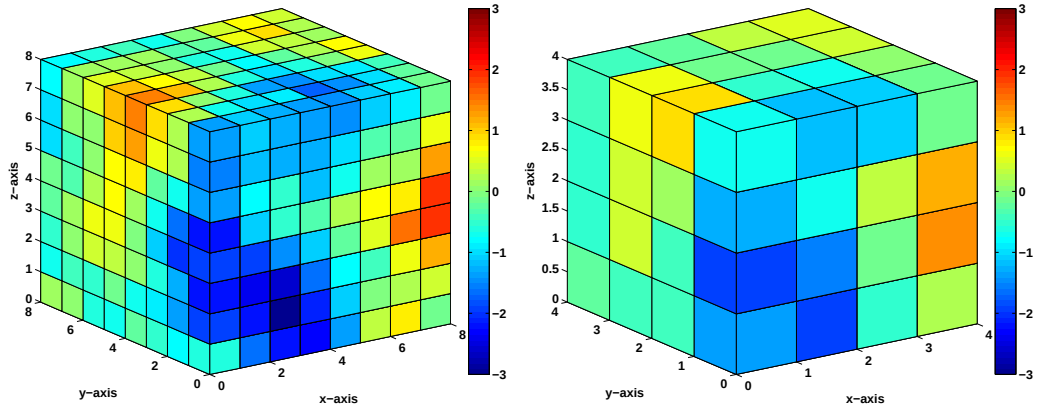


Figure 5: Coarse-scale fields, Left: $8 \times 8 \times 8$, Right:$4 \times 4 \times 4$.

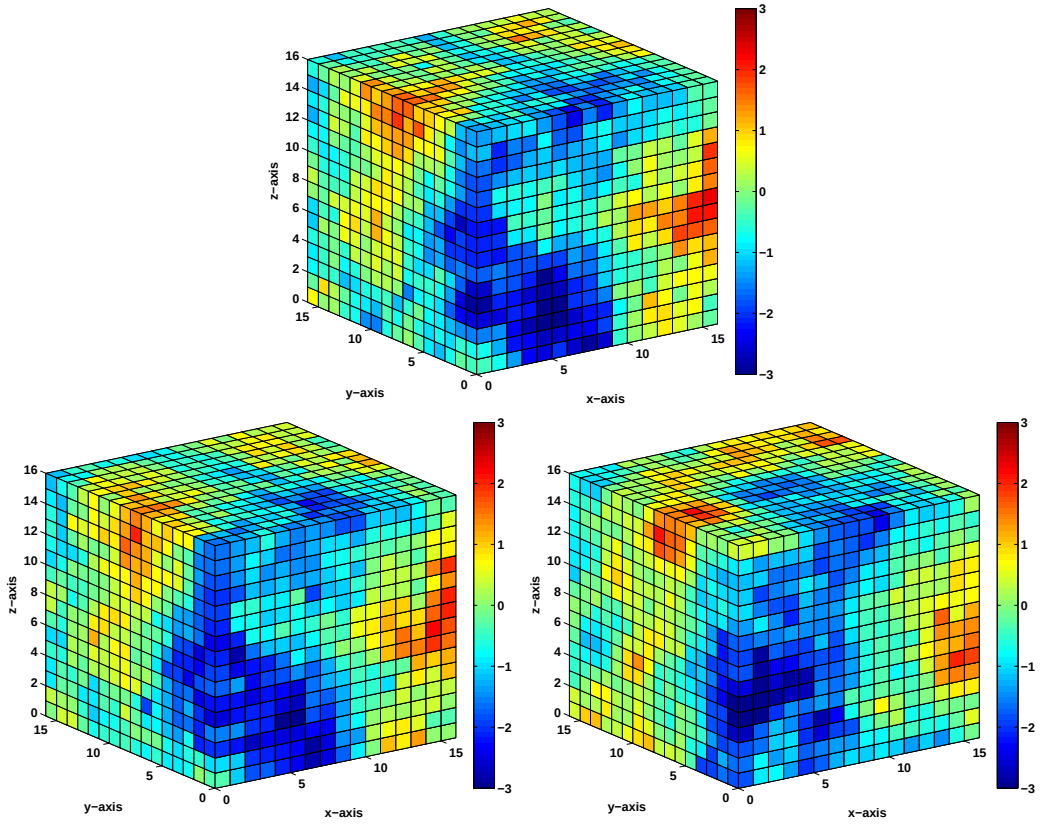


Figure 6: True field (top) compared with two fields generated by the SGSBK method. Bottom Left: SGSBK constraint to a coarse-grid of size $8 \times 8 \times 8$. Bottom Right: SGSBK constraint to a coarse-grid of size $4 \times 4 \times 4$.

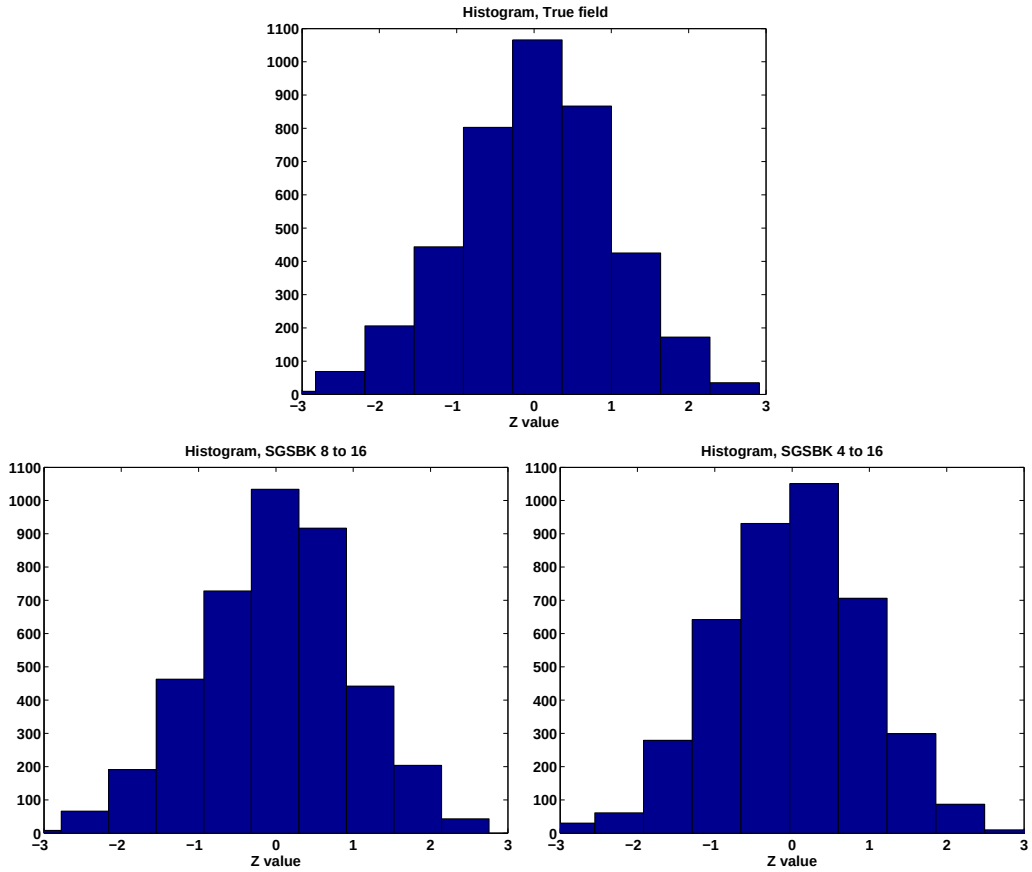


Figure 7: Estimated histogram from the true field (top) compared with the estimated histogram of the fields generated by the SGSBK method Bottom Left: SGSBK constraint to a coarse-grid of size $8 \times 8 \times 8$. Bottom Right: SGSBK constraint to a coarse-grid of size $4 \times 4 \times 4$.

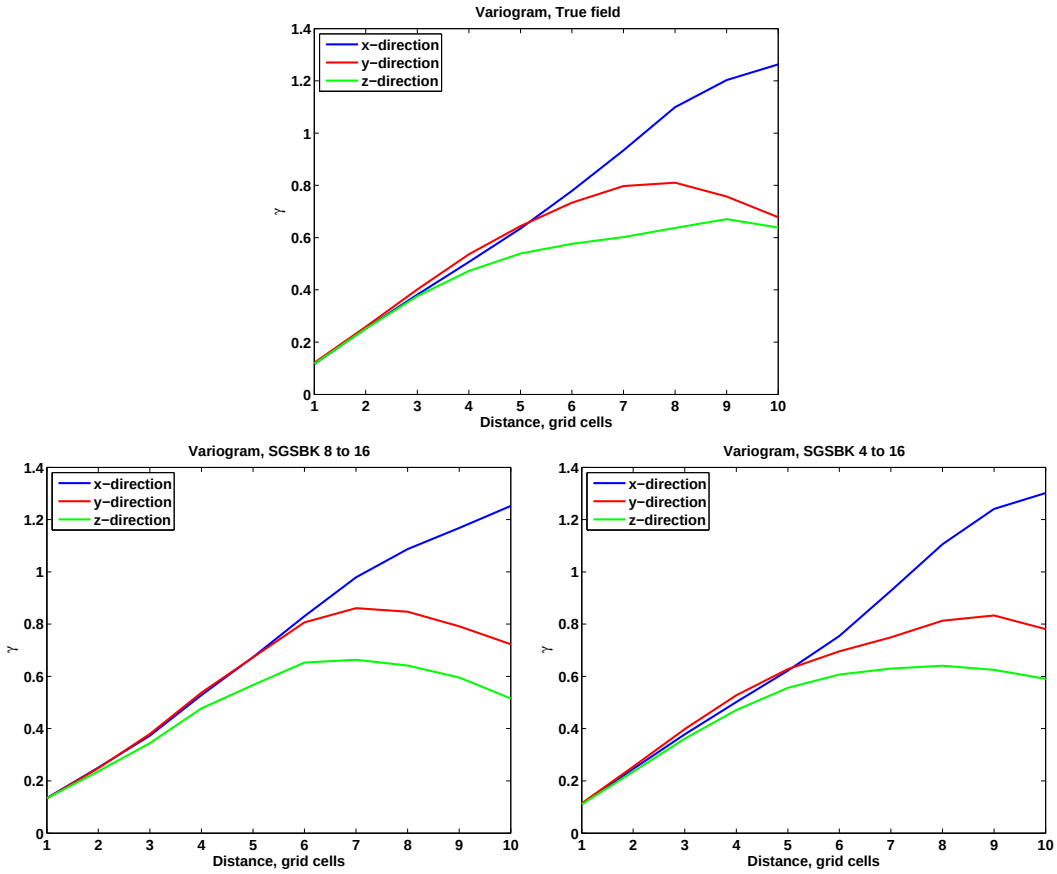


Figure 8: Estimated variogram from the true field (top) compared with the estimated variogram of the fields generated by the SGSBK method Bottom Left: SGSBK constraint to a coarse-grid of size $8 \times 8 \times 8$. Bottom Right: SGSBK constraint to a coarse-grid of size $4 \times 4 \times 4$.

References

R. A. Behrens, M. K. MacLeod, and T.T Tran. Incorporating Seismic Attribute Maps in 3D Reservoir Models. *SPE Reservoir Evaluation and Engineering, SPE 36499*, 1998.

T.T Tran, X. H. Wen, and R. A. Behrens. Efficient Conditioning of 3D Fine-Scale Reservoir Model To Multiphase Production Data Using Streamline-Based Coarse-Scale Inversion and geostatistical downscaling. *SPEJ, SPE 74708*, 2001.