4D Seismic History Matching Using the Ensemble Kalman Filter (EnKF): Possibilities and Challenges

PhD Thesis

Abul Fahimuddin

Department of Mathematics
University of Bergen

March, 2010

Summary

This research endeavor presents a 4D seismic history matching work flow based on the ensemble Kalman filter (EnKF) methodology. The objective of this work is to investigate the sensitivity of different combinations of production and seismic data on EnKF model updating. In particular, we are interested to quantify the performance of EnKF-based model updating experiments with respect to production and seismic data matching as well as to estimate uncertain reservoir parameters, e.g., porosity and permeability. The reservoir-seismic model system used consists of a commercial reservoir simulator coupled to an implemented rock physics model and a forward seismic modeling tool based on 1D convolution with weak contrast reflectivity approximation. One of the challenging issues of using 4D seismic data into reservoir history matching is to compare the measured data to the model data in a consistent way. Based on our realistic synthetic reservoir characterization case, time-difference impedance data generally performed better than time-difference amplitude data, and the matching of seismic data mostly improved with the inclusion of seismic data. In estimating posterior porosity and permeability, seismic difference data provided better estimate than using only production data, especially in aquifer region and also in areas that might be considered for in-fill wells. We experienced that the integration of seismic data in the elastic domain mostly provided better results than using seismic data at the amplitude level. This may be due to the measurement error used, and hence, further investigations are suggested to ascertain the appropriate level of seismic data integration.

The reservoir simulation model used is a sector model based on a full field North sea reservoir. The prior ensemble used consists of 100 model realizations. For computational efficiency, we have used efficient subspace-based EnKF implementations to handle the effects of large data sets such as 4D seismic. It may be difficult to assimilate 4D seismic data since it is related to the model variable at two or more time instances. Hence, we have used a combination of the EnKF and the ensemble Kalman smoother

(EnKS) to condition the reservoir with seismic data.

We performed a thorough study on the effects of using large number of measurements in EnKF by considering a single update of a very simple linear model. The sensitivity of EnKF update for several parameters, e.g., model dimension, correlation length, and measurement error variance also presented. We investigated the accuracy of the traditional covariance estimate with a large number of measurements. We demonstrated that the ensemble size has to be much larger than the number of measurements in order to obtain an accurate solution, and that the problem becomes more severe when the measurement uncertainty decreases, indicating that some kind of localization may have to be applied more often than previously believed.

In the real field case study, we have focused on matching the inverted acoustic impedance ratio (monitor survey/base survey) data between two time steps of several years of production. Note that for this real field case, there is a long period of production before the seismic data was assimilated. Hence, the porosity and permeability fields had a large influence induced by production data before they were actually updated with seismic data. Global and local analysis schemes assimilate production data and seismic data respectively. In our implementation of local analysis, we used three significant regions and seismic data within a given local analysis region is influenced by only variables in the same region. The posterior ensemble of models showed good match to both production data and seismic data. In most of the cases of reservoir characterization, the combined use of 4D seismic with production data improved history matching for the wells and also improved posterior impedance ratio data matching. In addition, 4D seismic data provided more information related to permeability update in the aquifer and in-fill areas. The results indicate that the local analysis reduced the amount of spurious correlations and tendencies to ensemble collapse seen with global analysis.

Preface

The dissertation is submitted as a partial fulfillment of the requirements for the degree of Doctor of Philosophy at the University of Bergen. The research works leading to the submission of this thesis started in March 2007, and have been carried out at the Center for Integrated Petroleum Research (CIPR), University of Bergen and also at the Statoil ASA Research Center in Bergen. This work is a part of the PETROMAKS project titled "Dynamic reservoir characterization and monitoring - Reservoir Monitoring and Dynamic Reservoir Characterization with Production, Seismic and Electromagnetic Data". Principal Researcher and Adjunct Professor Sigurd Ivar Aanonsen has been the main supervisor for this PhD project. Principal Researcher and Adjunct Professor Trond Mannseth and Associate Professor Morten Jakobsen have been co-advisers for this work. Senior Researcher Jan-Arild Skjervheim at Statoil ASA has provided important supervision to this PhD project. The Ensemble Reservoir Tool developed at Statoil ASA has been used in this work.

This dissertation is divided into two parts: Part I focuses on the theoretical background and methodology used for performing 4D seismic history matching. In Part II a set of research papers and scientific reports developed as the part of this work is presented. A brief outline of the PhD thesis is described in the following section.

Part I: General Background

Chapter 1 provides the context of this research work by describing the scope of 4D seismic history matching in the EnKF framework. Furthermore, the main research objectives of the research work are outlined and also an overview of the papers and the results are presented.

In chapter 2 we have described a comparative study of both qualitative and quantitative integration of 4D seismic data in reservoir engineering. Also we have discussed the definition of the 4D seismic history matching

with an example work flow. At the end, we have outlined the challenges of 4D seismic data integration.

In chapter 3 the concept of reservoir characterization is presented. A general theory for reservoir history matching problem described in a Bayesian framework is outlined, followed by the description of different history matching techniques. The chapter end with a comparison between EnKF and traditional history matching methods.

Chapter 4 provides a through description of all the forward models used in this work. In particular, we have described theories for flow in porous media and a detailed analysis of the rock physics models for pressure-dependent unconsolidated sand reservoir is presented next. And finally, generalized theory for 1D convolution model based on weak contrast approximation of reflection coefficient is delineated.

Chapter 5 contains an overview of the ensemble Kalman filter methodology as a data assimilation technique. Starting from basic Kalman filter methodology, we have shown detailed derivation of EnKF algorithm. We have described the basic steps for EnKF-based reservoir history matching technique. The modeling error issue with EnKF is introduced as well. The chapter ends with the major challenging aspects of EnKF method for reservoir engineering problems.

Chapter 6 provides the mathematical derivation of the algorithm for performing 4D seismic history matching in EnKF; particularly we have shown the ensemble smoother concept in this section. A comprehensive work flow for EnKF-based history matching for time-lapse seismic data is presented. We have also proposed a methodology to incorporate seismic modeling error in EnKF history matching. Finally, we have described the main challenges of this research work in the context of 4D history matching in EnKF framework.

In chapter 7, we have touched the important issue of integration of large number of measurements that is relevant to seismic data integration in EnKF. We have briefly described the efficient subspace EnKF algorithm that has been used in this work. Also, different strategies to handle spurious correlations and filter divergence, e.g., local analysis and covariance localization are described in this chapter.

Finally, we have described the main research findings of this work in the chapter 8. We have finished this introduction part by delineating some further extensions based on the experience and findings of this research task.

Part II: Included Papers and Reports

Paper A: "Ensemble Based 4D Seismic History Matching: Integration of Different Levels and Types of Seismic Data", Abul Fahimuddin, Sigurd Aanonsen and Jan-Arild Skjervheim. Proceeding of the to be held 72nd EAGE/SPE Conference & Exhibition incorporating SPE EUROPEC 2010, Barcelona, Spain, 2010. SPE 131453.

Paper B: "4D Seismic History Matching of a Real Field Case with EnKF: Use of Local Analysis for Model Updating", Abul Fahimuddin, Sigurd Aanonsen and Jan-Arild Skjervheim. Proceeding of the to be held SPE Annual Technical Conference & Exhibition (ATCE 2010), Florence, Italy, 2010. SPE 134894-PP. (To be submitted to SPE Journal)

Paper C: "Effect of Large Number of Measurements on the Performance of EnKF Model Updating", Abul Fahimuddin, Sigurd Aanonsen and Trond Mannseth. Proceedings of the 11th European Conference on the Mathematics of Oil Recovery, Bergen, Norway, 2008.

Report 1: "Petro-elastic Modeling of a North Sea Field: Rock Physics Recipe and ECLIPSE Simulator", Abul Fahimuddin. Scientific/Technical Report.

Report 2: "Forward Seismic Modeling Using 1D Convolution with Weak Contrast Approximation of Reflectivity", Abul Fahimuddin. Scientific/Technical Report.

Acknowledgements

I would like to acknowledge the financial support obtained during the PhD period as part of the Norwegian Research Council PETROMAKS project and also would like to acknowledge the support received from all the project sponsors. Special notes of appreciation are directed towards the members of the EnKF group at Statoil ASA Research Center in Bergen for their spontaneous support and assistance during the last year of this study. I want to express my sincere gratitude to all of my colleges at the Center for Integrated Petroleum Research (CIPR) in Bergen.

I would like to sincerely acknowledge the guidance and supervision of Sigurd Aanonsen. A special note of gratitude is also made to Jan-Arild Skjervheim for his continual support and cooperation. The discussions with Trond Mannseth have always been an important impetus for the timely completion of this project. Special thanks to Joakim Hove and Jon-Gustav Vabø at Statoil ASA for providing assistance to use the Ensemble Based Reservoir Tool. I would also like to express my appreciation to Philippe Doyen at CGGVeritas for his valuable comments and suggestions.

Finally, a deep note of gratitude to my parents and siblings in Bangladesh. Without their support, guidance, prayer and love, it is inconceivable for me to ponder to move towards the goal of my life.

Abul Fahimuddin
March, 2010

Contents

Part I

General Background

Chapter 1

Introduction

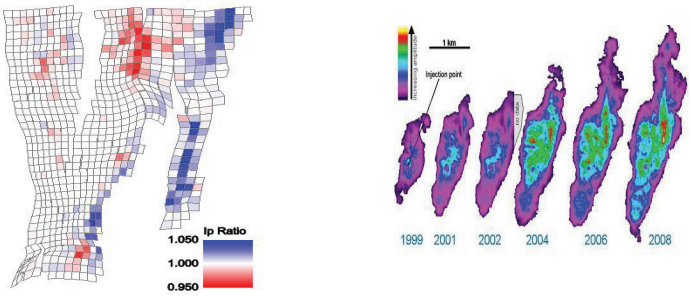
With an ever-increasing rate of depletion of conventional fossil fuel resources, one recurring question that resonates among versatile stake holders is "When do we run out of fossil fuels"? Intuitively, the hydrocarbon reserves should decrease due to continued increasing consumption. There are several reasons why the known reserves increase rather than decrease. With the help of advanced technology deeper targets can be reached, new frontiers deep offshore can be explored, and from existing fields more hydrocarbons can be produced. Due to the complexity of the reservoirs, the current average recovery is quite low, and with advanced technology, it should be possible to increase recovery, perhaps to an average of 70% in the future [107]. Advances in technology are continuously occurring at several fronts, e.g., the introduction of 4D seismic data. 4D seismic refers to repeating seismic acquisition over time. It captures the dynamic behavior of the reservoir and aids reservoir management, allowing increased recovery. Reservoir management also benefits from other technological advances in e.g. stochastic modeling and uncertainty quantification of the model's production forecast. The implementation of the above mentioned and other new developments requires further improvements as well as new advances in the associated disciplines [27].

The main value of 4D seismic is the additional information to constrain or update a model of the reservoir, localization of undrained oil, detailed well planning etc. Knowledge of reservoir connectivity, flow barriers or bypassed hydrocarbons is the kind of information that is expected from 4D data [85]. Such knowledge helps to optimize reservoir investment decisions. The reservoir models are, in general, constrained by history matching with well production data. As this solution is most often not unique, the combination of 4D seismic (large spatial distribution, but with non-unique interpretation) and well data (single point location, but

regarded as accurate) enormously reduces uncertainty. Using this information, subsurface uncertainties are also reduced. This naturally leads to cost reductions, better well placements and (sometimes drastic) changes in field development plans. Today much of the 4D seismic application is qualitative, or at best semi-quantitative, i.e., the 4D seismic data is used to identify areas of changes in saturation and pressure distributions between seismic survey times. The need to be more quantitative is already here, i.e. to estimate not only *what kind* of changes but also *how large* are these changes in saturation and pore pressure [123]. In the rapid work flows of the future, the quantitative interpretation methods will be an integral part.

The simplest and the most direct method of using time-lapse seismic data is to qualitatively monitor reservoir changes due to production. In this approach, one simply identifies regions in which the amplitude or impedance has changed with time and attributes these changes to changes in saturation, pressure, or temperature. In **Fig.** 1.1a, the increase in acoustic impedance ratio between monitor and base survey shows the part of the reservoir which has been drained by injected water. This type of qualitative use of 4D seismic data has been used for numerous oil and gas fields around the world. For example, 4D seismic has been very useful on the Gullfaks field to identify areas where significant gas saturation changes have occurred and to locate fluid communication paths [91]. At Meren field in Nigeria, the primary objective of 4D seismic data integration was to identify pathways of injected water, sealing faults, and compartments that may contain bypassed oil [96]. For the Gannet C oil and gas field in the UK central North sea, 4D data revealed major extensions of reservoir units previously presumed to be absent or thin over much of the reservoir [136]. As a paradigm shift, 4D seismic monitoring can also be extended to quantify the amount of injected CO_2 and any changes that subsequently occur due to leakage or dissolution [31]. At the North sea Sleipner site, six repeat surveys over 12 years have revealed both an expansion and compaction of the CO_2 plume, see Fig. 1.1b which is taken from Sandø et al. [123]. Distribution estimates of the CO_2 saturation have been made from 4D seismic data and have increased the understanding of CO_2 flow, with reasonable accuracy.

The development of assisted history matching techniques with 4D seismic data is a very active domain which represents a significant improvement towards a quantitative use of 4D seismic data in reservoir modeling [119]. Seismic history matching is a process of improving reservoir simulation models by constraining these model with both production and 4D seismic data in order to improve the characterization of permeability and porosity heterogeneity. Gosselin et al. [70] proposed a gradient based op-



(a) Water drained reservoir (b) CO_2 sequestration

Figure 1.1: Use of 4D seismic data in reservoir management.

timization method to simultaneously minimize the mismatch of all types of measured, including 4D seismic, data to the simulated ones. Waggoner et al. [151] used the 4D seismic results to constrain an optimized history matching procedure along with production data for a condensate gas reservoir in the Gulf of Mexico. Stephen et al. [139] [138] applied a method of multiple-model history matching based on simultaneous comparison of spatial data offered by seismic as well as individual well-production data for the UKCS Schiehallion reservoir. The work of Roggero et al. [119] focused on the advanced parametrization technique to constrain fine scale geo-statistical model by means of gradual deformation method in the framework of 4D seismic history matching of the Girassol field. Dong et al. [45] used an adjoint method to compute the gradient of the data mismatch and a quasi-Newton method to compute the search direction in the context of automatic history in order to incorporate 4D seismic data to a reservoir in the Gulf of Mexico.

In general, conditioning reservoir model to seismic data is a difficult undertaking. One of the challenging issues using seismic data is to compare measured data to the model data [134]. Different types of 4D seismic data, e.g., amplitude data or inverted acoustic impedance difference data are in use. The amount of data, uncertainty and information content may be very different and dependent the way seismic data are incorporated. Measuring of mismatch between the model and data in a consistent way is always required. Additional difficulties are related to the choice of appropriate weighting of seismic attributes compared to production data in the objective function used in history matching. Moreover, as 4D attributes provide much more spatial information compared to well or production data, more

flexibility is required in the model parametrization. In the 4D work flow, several differences in scales have to be taken into account. Fluid flow simulations are generally performed on a coarse grid in comparison with the stratigraphic geological grid. The petro-elastic model (PEM) is used at the fine scale, and finally, simulated seismic attributes include high-frequency information due to the heterogeneity of the geological model. Hence, one cannot use the modeled seismic parameters directly to real seismic data which are characterized by a limited frequency bandwidth. Even though a challenging task, there is a growing interest and necessity to incorporate 4D seismic data quantitatively in the work flow for reservoir management [123].

The ensemble Kalman filter (EnKF) is well suited to combine widely different types of data, like, e.g. production history and seismic data [3]. The EnKF method was introduced by Evensen in 1994 [58] and is a Monte Carlo type sequential Bayesian inversion method, suitable for history matching a reservoir simulation model. The EnKF method provides a solution of the combined state and parameter estimation problem where the result is an ensemble of solutions approximating the posterior distribution defined from the likelihood and the prior model [61]. The use of EnKF for history matching reservoir simulation model was first proposed by Naevdal et al. [102]. They used the EnKF to update permeability fields for near-well reservoir models. Applications of the method to real field cases are discussed in several recent works. The work of Bianco et al. [19] is one of the first application which successfully used EnKF to history match a real field reservoir model and showed how the EnKF can be used to evaluate the uncertainty in the production forecast for a given development plan for a real field model. The field at hand was an on-shore saturated oil reservoir. Haugen et al [76] applied the EnKF to history match a North Sea field model and found that the EnKF estimate improved the match to the production data. More recently, Cominelli et al. [34] have used EnKF to history match the simulation model and to assess the remaining uncertainty on production forecast for a deep-water under saturated oil reservoir.

Extending the EnKF method to assimilate seismic data, especially 4D seismic on a real field case is first shown by Skjervheim et al. [135]. They used a subspace EnKF inversion scheme to integrate interpreted seismic data into simulation model for both a 2D synthetic model and a real field case. For the real field case, Skjervheim et al. [135] assimilated Poisson's ratio difference data together with production data to improve a base case model for the reservoir. Dong and Oliver [46] incorporated 4D impedance data in EnKF for a small synthetic case. The observed data were matched

very well, and the true model features were recovered. They also showed that the estimated porosity field was better than the estimated permeability field because seismic data are directly sensitive to porosity but only indirectly sensitive to permeability. As an example of using seismic waveform data in EnKF, Haverl and Skjervheim [78] used 4D stacked amplitude data in EnKF to estimate reservoir properties such as gas-oil-contact. In spite of having several challenges, e.g., issues of integration of large amount of 4D data and localization, the initial EnKF applications for 4D seismic data provided promising results, see Aanonsen et al. [3].

1.1 Main Research Objectives

There is a recognized need to combine the skills of geoscientists and engineers to build quantitative reservoir models that incorporate all available reservoir data [20]. Available reservoir data include conceptual geologic models, seismic, cores, well logs, and production. The challenge is to integrate all these disparate data into a unified, self-consistent reservoir characterization model. The primary objective of history matching (conditioning reservoir models to dynamic data) is to modify a prior model for the reservoir such that the updated model reflects the available production data and the uncertainties in production forecasts are reduced. The incorporation of 4D seismic data to constrain the reservoir models to update fine-scale geological model, has the potential to improve the overall reservoir characterization. It is a challenging endeavor to perform 4D seismic history matching in the EnKF framework, especially for a real field case with interpreted real 4D seismic data. Our work addresses various aspects of this difficult task. For history matching, we are interested to quantify the performance of the wells, also to match the assimilated seismic data and to estimate the reservoir parameters, e.g., porosity and permeability. We would like to investigate the sensitivity of different combinations of production and 4D seismic data on EnKF model updating for a realistic synthetic case based on a full field reservoir model from North sea.

In a model updating process or when conditioning a reservoir model to 4D seismic data, the conditioning may be introduced at different levels corresponding to where the mismatch between simulated and measured data is evaluated. An illustration of the different seismic mismatch levels is shown in **Fig.** 1.2 which is taken from the work of Skjervheim [131]. In order to calculate synthetic seismic data, one has to use a PEM model to convert reservoir simulator output, e.g., pressure and saturation, to seismic signatures (acoustic impedance or amplitude data). At the highest

level, real field 4D amplitude data, e.g., seismograms, AVO attributes or post-stack data, are compared to synthetic amplitudes. This is the amplitude domain. At the next level, we assume that real seismic data have been inverted to elastic properties such as acoustic impedance or Poisson's ratio. These interpreted seismic data are compared with the corresponding simulated impedance or Poisson's ratio data. At the bottom most level, seismic data are assumed to be further inverted for pressure and saturation, which then are directly compared with the flow simulator outputs. We refer this level of integration as fluid changes domain. As it is seen that

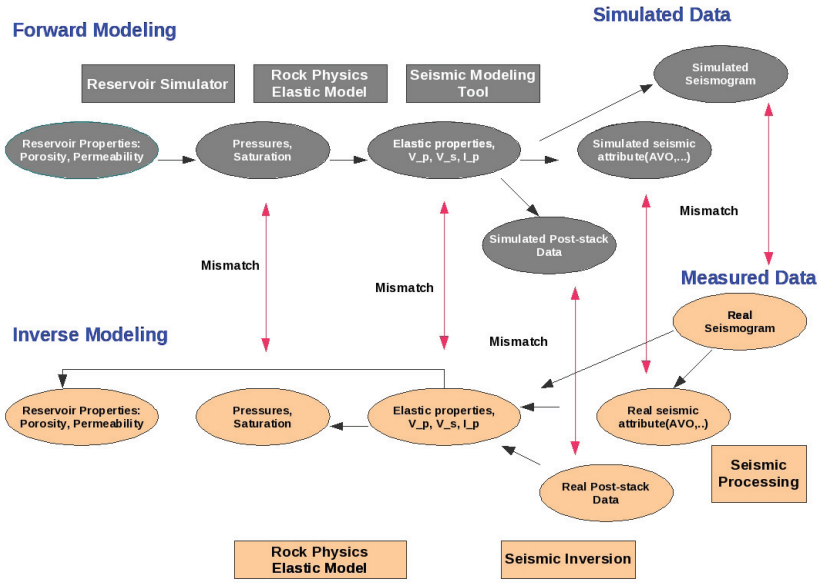


Figure 1.2: Different levels of 4D seismic data integration.

seismic data can be integrated at various levels, i.e., amplitude, elastic and fluid changes domain. In this work, we are interested to ascertain the appropriate level of 4D data integration for our reservoir characterization. For this purpose, we have made EnKF-based history matching experiments with different types of 4D seismic data on a realistic synthetic case.

Because EnKF is typically applied with a limited ensemble size there are not sufficient degrees of freedom to match all seismic data, especially when we effectively have seismic data in every grid cells. As a consequence of

finite ensemble size, an under estimation of ensemble variance may occur. This variance reduction can be seen as spurious correlations over long spatial distances or between variables known to be uncorrelated [62]. Several possible methods based on localization, e.g., using local analysis [132] or distance-based covariance localization [40] have been proposed to handle large seismic data sets in EnKF. For the field case study with inverted real 4D seismic data, we are interested to improve the EnKF update results by using a methodology based on a combination of a global and a local analysis scheme. In addition, we intended to perform a thorough study on the theoretical aspects of using large number of measurements in EnKF updating formula.

When we integrate different types of data for reservoir characterization, we need different forward models to simulate synthetic data to compare with the corresponding field observation. In this work, we have implemented a PEM model for a pressure dependent sandstone North sea reservoir. We have used an initial manually history matched reservoir model and a rock physics recipe from Statoil ASA. Both for the synthetic and real case studies, we have used a sector model based on the full field reservoir simulation model. For the calculation of synthetic amplitude data, we implemented a forward seismic modeling tool based on 1D convolution and weak contrast approximation of reflectivity. The mathematical formulation of this forward model is based on the work of Buland and Omre [24]. Next objective of our work has been the integration of the different forward models in an EnKF framework. In this project, we mostly used the Ensemble Reservoir Tool (ERT) developed at Statoil ASA to run EnKF experiments. For the real field 4D seismic data, CGGVeritas provided us with the inverted acoustic impedance data.

1.2 Summary of Papers and Reports

Paper A is mainly concerned with a methodology to choose an appropriate level of 4D seismic data integration in history matching loop. The main objective of this endeavor is to investigate the effects of different types of seismic data on EnKF model updating. In this work, we have considered a realistic synthetic reservoir sector model based on the full field model by including the top 18 producing layers. We mainly considered two types of synthetic time-difference seismic data, i.e., acoustic impedance difference data and the amplitude difference data. We have introduced a simplified method for seismic forward modeling for amplitude data in depth based on 1D convolution allowing a detailed study of the impact of vertical resolu-

tion in seismic data, while neglecting other sources of seismic uncertainty. We showed that the vertical resolution may have a significant effect on the model update using EnKF. In order to handle time-difference data in EnKF, we used ensemble Kalman smoother (EnKF) in a combination with EnKF. An efficient subspace EnKF implementation was used to handle large number of seismic measurements. In most of the cases of reservoir characterization, time-difference impedance data performed better than time-difference amplitude data, and considerably reduced posterior ensemble spread. The matching of seismic data generally improved with the inclusion of time-difference seismic data. In estimating posterior porosity and permeability, seismic difference data provided better estimate than using only production data, especially in aquifer region and also in areas that might be considered for in-fill wells. Thus, in our realistic synthetic case based on a full field reservoir model, we experienced that the integration of seismic data in the elastic domain mostly provided better results than using seismic data at the amplitude level. The results for seismic data matching also indicated that the introduction of large number of seismic data, probably, introduced spurious correlations, and that some kind of localization should be used. Still, without localization, the results indicate an improvement in the estimate of porosity and permeability when adding seismic data compared to using only production data.

In **paper B**, the ensemble Kalman filter (EnKF) with local analysis was applied to match real, inverted 4D seismic data and production data for a sector of a North Sea oil reservoir. In history matching process, we have focused on matching the acoustic impedance I_p ratio, between two time steps of several years of production. Note that for this real field case, there is a long period of production before the seismic data was assimilated. Hence, the porosity and permeability fields had a large influence induced by production data before they were actually updated with seismic data. Global and local analysis schemes assimilate production data and seismic data respectively. In our implementation of local analysis, we used three significant regions and seismic data within a given local analysis region is influenced by only variables in the same region. The posterior ensemble of models showed good match to both production data and seismic data. In most of the cases of reservoir characterization, the combined use of 4D seismic with production data improved history matching for the wells and also improved posterior impedance ratio data matching. In addition, 4D seismic data provided more information related to permeability update in the aquifer and in-fill areas. The results indicate that the local analysis reduced the amount of spurious correlations and tendencies to ensemble collapse seen with global analysis.

In **paper C**, we have investigated the performance of EnKF updating for smaller ensemble members relative to a large number of measurements. This is done by considering a single update of a very simple linear model and compare the EnKF update with the traditional Kalman filter, which in this case is exact. Because for linear models with Gaussian prior, minimum variance estimation as such the Kalman filter update is the best approximation. The sensitivity of EnKF update for several parameters, e.g., model dimension, correlation length, and measurement error variance is presented. The quality of the EnKF update is assessed by considering the mean and variance of the updated state variable, as well as various error norms e.g., relative root mean square error (RMS) norm, scaled Frobenious matrix norms and the eigen-spectrum of the covariance matrix. In particular, we have varied the number of measurements, N_d alongside ensemble size, N_e. This has provided us the opportunity to observe the spurious correlation effect with large N_d. Even for this simple model and straight forward EnKF implementation, spurious long-range correlation, ensemble collapse, etc. are clearly seen as number of measurements increase for a given N_e. The problems occur for $N_d < N_e$, and if the number of measurements are large, a very large N_e is needed for an accurate solution.

Report 1 presents a detailed description of the rock physics model that we have implemented as the part of this research work. The objective of this work was to implement the PEM model so that we could integrate this in our overall EnKF-based 4D seismic history matching framework. The PEM model uses conventional existing models calibrated to well log data measurements. In particular, we have shown how this PEM model can be implemented in the PEM modeling tool in Eclipse simulation software [53]. The detailed mathematical formulation of the PEM model for pressure-dependent unconsolidated sandstone reservoir and also the the necessary input for the Eclipse PEM model are described. In our example, we have used Archie's formula to calculate a realistic water saturation profile based on standard resistivity log parameter values [118]. We have also performed some synthetic studies on the predicted seismic behavior.

In **report 2**, we have made a detailed description of the forward modeling for synthetic seismic amplitude data that has been used in the EnKF experiments. The model is based on 1D convolution and weak contrast approximation of reflection coefficients. The mathematical formulation of this forward model is based on the work of Buland and Omre [24]. Normally, the velocity or reflection coefficients calculated from the reservoir simulation model are depth converted and sampled into a regular time grid before performing the convolution. Since, the velocities and layer thicknesses normally are different for each layer in the reservoir model, this will

always introduce a sampling, or scaling error. Here we have applied a different procedure, where all the calculations, including the convolution, are performed in depth. This will of course not be possible in a real case. However, in this way we are able to generate a synthetic problem where these errors are minimized, and also it simplifies calculations. For the reservoir sector model, we have shown the use of this seismic modeling tool in this report. Detailed analysis of the input parameters necessary for this forward model is also illustrated with an example.

Chapter 2

4D Seismic Data for Reservoir Management

The objective of reservoir management is to produce each reservoir optimally according to economic, political, technical, and environmental constraints. Reservoir management is a complex task heavily depending on the reservoir model. This reservoir model allows analyzing behavior of the reservoir, but more importantly to forecast future production behavior. A reservoir model has to be constrained to the observed static and dynamic data. Static data comprise core data, logs, geological data, 2D and 3D seismic data etc. Dynamic or historical data comprise production, well test, and time-lapse seismic data. Termed "four-dimensional seismology" by Nur [104], 4D seismic data comprise the set of 3D seismic data acquired at different times over the same area, with the objective of monitoring changes occurring in a producing hydrocarbon reservoir over time [29]. It was originally observed from rock physics experiments that the change of saturation and fluid pressure can lead to detectable changes in seismic attributes forms the physical basis of time-lapse seismic monitoring [26].

4D seismic data record two types of changes: changes in reservoir properties due to production, and changes in external variables such as ambient noise, recording equipment, etc. Changes in reservoir properties due to production are recorded by 4D seismic data since seismic velocities and impedance depend on changes in pore fluids, pressure and temperature [105] [154]. The overall elastic modulus of a rock change with the type of fluid in the pores, the effective pressure acting on the rock, and the temperature the rock is subjected to. Due to the change in elastic modulus, the rock becomes more or less resistant to wave-induced deformations; therefore, seismic velocities experience an increase or decrease in magnitude. These observations are the basis for using 4D seismic data in predicting

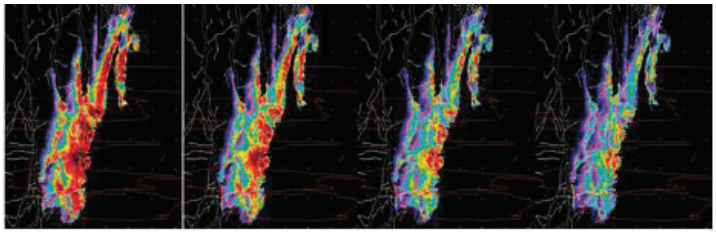


Figure 2.1: Oil saturation map in Tarbert formation from Gullfaks reservoir model: Attribute maps from 4D inversion showing the average oil saturation values, extracted from 4D inversions from surveys in 1996, 1999, 2003 & 2005. Warm colors indicate high oil saturation.

fluid saturation and pressure changes in the reservoir. Typically, the difference between two 3D seismic data sets recorded at different times allows mapping those areas in the reservoir where pressure and/or the distribution of fluids have changed. Therefore, 4D seismic data not only provides information about the dynamic process occurring in the reservoir while production takes place; but also provides information about the spatial lithology heterogeneity where dynamic changes occur. In the context of 4D reservoir management, the reservoir production data can be utilized along with the seismic data to improve the reservoir model through history matching. This technique allows a quantitative, rather than a visual, interpretation of 4D seismic results. The resulting improvement provides a direct linkage to reservoir management tools, so that the knowledge gained from the 4D study can be better used to manage the reservoir future performance.

4D seismic started in the early 1980's and became commercial in the late 1990's. In the North Sea, 4D seismic was investigated on a full field scale in 1995 in a joint Statoil-Schlumberger project at the Gullfaks field [123]. On the Gullfaks field, time lapse (4D) seismic data has played a significant role in management of tail end production [103]. 4D seismic data have increased understanding of both the static properties and dynamic behavior of the field, challenging the production strategy and driving the reservoir management decision process, see **Fig.** 2.1 which is taken from Sandø et al. [123]. The data have identified bypassed volumes in un-swept reservoir compartments, improved the estimate of remaining reserves and the associated uncertainties, and have helped to identify drilling hazards that have evolved over the life time of the field. Since the first repeat seismic survey in 1995, 4D seismic data have been integrated into a multi-

disciplinary increased oil recovery (IOR) effort that aims to recover 70% of the in-place reserves from a structurally complex and heterogeneous reservoir sequence [5].

The arrival of time-lapse seismic data has forced different disciplines to intensify their working relationship in order to optimally benefit from the dynamic information content of the data [50]. These disciplines comprise, but are not limited to, geology, petro-physics, rock physics, reservoir engineering, and seismic acquisition and processing, see **Fig.** 2.2 which is taken from the work of Oldenziel [107]. To allow communication and integration between disciplines, modifications have to be made within each discipline. Existing theories, algorithms, and models have to be revised or

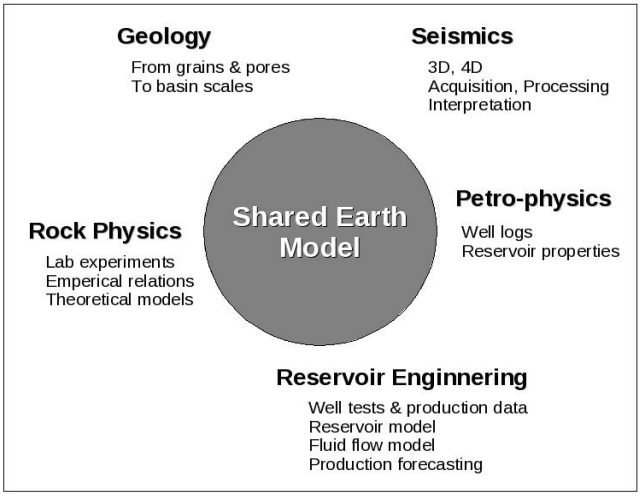


Figure 2.2: Concepts of shared earth model.

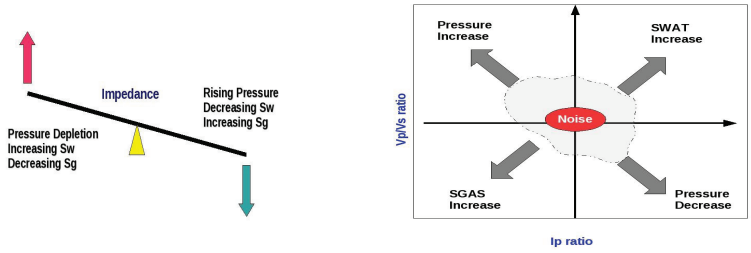
improved to suit time-lapse seismic interpretation. Some are appropriate for 3D data handling, but might not provide the answers for time-lapse seismic. For example, processing of 3D seismic is adapted to include cross-equalization in order to allow comparison of different time-lapse seismic data sets. Within rock physics the focus has to be on the combined effect of the changes in the reservoir rather than the effect of a pressure or a saturation change [153]. To integrate the huge amount of data and information generated by time-lapse seismic, reservoir engineering practices have to be adapted [13].

2.1 Qualitative Use of 4D Seismic Data

The simplest, most direct method of using time-lapse seismic data is to qualitatively monitor reservoir changes due to production. In this approach, one simply identifies regions in which the amplitude or impedance has changed with time and attributes these changes to changes in saturation, pressure, or temperature [43]. Time-lapse seismic is not a new topic in petroleum engineering and geophysics. The pioneering work of time-lapse seismic can be traced back to late 1980s and early 1990s, e.g., Wayland and Lee [155]. Similar studies have been reported by Cooper et al. at the Foinhaven field [35] and by Lumley et al. [96] at the Meren field in Nigeria. The primary objectives at Foinhaven were simply to map fluid movements and to identify by passed oil. The authors of the study concluded that the time-lapse signal qualitatively agreed with the expected reservoir performance. At Meren, the goal was to identify pathways of injected water, sealing faults, and compartments that may contain bypassed oil. 4D seismic has been very useful on the Gullfaks field to identify areas where significant gas saturation changes have occurred and to locate fluid communication paths [145]. Also for the Gullfaks field, 4D seismic has been used to ascertain depleted areas and so far, 14 infill wells have been drilled based on 4D studies [141] [55]. For the Gannet C oil and gas field in the UK central North sea, 4D data revealed major extensions of reservoir units previously presumed to be absent or thin over much of the reservoir [136]. More recently, for the Heidrun field in the North sea, 4D seismic data improved the understanding of reservoir fluid flow and communication across faults [120].

2.2 Information Content of 4D Seismic Data

A change in pressure or saturation within a rock gives rise to a change both in its bulk density and sonic velocity. The magnitude of the changes is controlled by the physical properties of the rock frame and the filling pore fluids. **Fig.** 2.3a illustrates the effect of these changes on acoustic impedance ($\mathbf{I}_p$) which is a function of density and p-wave velocity ($\mathbf{V}_p$) and therefore has a combined response to pressure and saturation change. The polarity of the response depends on whether the pressure is increasing or decreasing and the difference between the fluid properties at the start and end of the period. The effect of the changes on amplitude seen at the top of a reservoir depends on the contrast between the $\mathbf{I}_p$ of the reservoir and the overlying cap rock, so an increase in reservoir impedance will



(a) Qualitative 4D: Discrimination of pressure and saturation

(b) Quantitative 4D interpretation

Figure 2.3: Interpretation of 4D seismic data.

lead to an increase in the seismic amplitude at the top of the zone if the overlying zone is acoustically softer. Most cap rock in the North Sea are harder than the reservoir, so an increase in impedance leads to a dimming of amplitude [97].

When cross plotting the change in $\mathbf{V}_p/\mathbf{V}_s$ versus the change in $\mathbf{I}_p$ (time-lapse change = monitor/base), each quadrant in the cross plot is associated with effects from production (see, Fig. 2.3b). If both $\mathbf{I}_p$ and $\mathbf{V}_p/\mathbf{V}_s$ are reduced over time, then the associated effect may represent a gas flooding. An increase in $\mathbf{I}_p$ and a decrease in $\mathbf{V}_p/\mathbf{V}_s$ may correspond to effects from pressure decrease (depletion), while a decrease in $\mathbf{I}_p$ and an increase in $\mathbf{V}_p/\mathbf{V}_s$ may correspond to pressure increase. An increase in both $\mathbf{I}_p$ ratio and $\mathbf{V}_p/\mathbf{V}_s$ ratio may correspond to effects from water flooding. As an application of this quantitative 4D interpretation in a real filed case, we can mention the work of Andersen and Zachariassen et al. [7]. Paper B attached in the publication section of this thesis utilizes similar way of incorporating 4D seismic data in history matching.

2.2.1 Different Types of 4D Seismic Data

There are few options available to incorporate 4D seismic data quantitatively in the history matching loop. Landa and Horne [90] estimated reservoir parameters assuming that water saturation changes could be derived from the time-lapse seismic. They included dynamic data observed from wells. Huang et al. [81] used time-lapse seismic amplitude data and the finite perturbation method to calculate required derivatives. In a recent work, Waggoner et al. [151] used acoustic impedance difference derived from time-lapse seismic. Kretz et al. [89] matched water fronts extracted

from time-lapse seismic surveys. Mezghani et al. [100] used time-lapse seismic acoustic impedance in history matching, together with production data. The finite perturbation method was used to compute the required derivatives. Dong and Oliver [44] matched both seismic impedance change data and production data in a medium scale problem. The work of Gosselin et al. [69] used different types of 4D seismic data for history matching; Stephen et al. employed the Neighborhood algorithm in the context of 4D seismic history matching [139] and Roggero et al. used gradual deformation technique to constrain the geological models which were used in seismic history matching [119]. For 4D seismic history matching of large scale field cases in the North sea, we can mention the works of Kjelstadli et al. [88] for the Vallhall field, the work of Van Ditzhuijzen [147] for the Stratfjord field, and the work of of Haverl et al. [77].

The use of seismic amplitude data to characterize a reservoir quantitatively, is a not a trivial task. Vasco and Datta-Gupta et al. [149] used a trajectory-based methodology to infer reservoir flow parameters such as permeability from time-lapse seismic data. Another application of quantitative seismic amplitude data is the monitoring of water saturation fronts showed by Jin and Sen et al. [83]. Skjervheim and Ruud applied a simple seismic modeling approach to simulate waveform data for a synthetic 2D reservoir model and used the waveform data together with production data in EnKF setting [133]. Dahashpour et al [37] used a synthetic time-lapse seismic amplitudes and synthetic time-lapse zero-offset amplitude data with a 2D model and minimized the difference with simulated data. They calculated the sensitivities to unknown grid parameters, e.g., porosity using finite differences and a Gauss-Newton scheme.

2.3 Seismic History Matching

The possibility of incorporating 4D seismic information into history matching as additional dynamic data is an attractive proposition as it provides images of fluid movements between wells [49]. It also reduces the extent of non-uniqueness of traditional history matching. In the inverse problem theory based on the least-square formulation (see, Tarantola [142]), the objective function is generally expressed as follows [119]

$$F = \frac{1}{2}\left(\mathbf{d}^{obs} - \mathbf{d}^{sim}(\mathbf{m})\right)^T \mathbf{C}_\mathbf{d}^{-1}\left(\mathbf{d}^{obs} - \mathbf{d}^{sim}(\mathbf{m})\right) + \frac{1}{2}\left(\mathbf{m}_{prior} - \mathbf{m}\right)^T \mathbf{C}_\mathbf{m}^{-1}\left(\mathbf{m}_{prior} - \mathbf{m}\right), \tag{2.1}$$

where F is the objective function value, $\mathbf{m}$ is the reservoir model parameters, and $\mathbf{d}^{obs}$ is the vector of observed values to be matched, and $\mathbf{d}^{sim}(\mathbf{m})$ is the vector of simulated values which depends on $\mathbf{m}$. The inverse of the covariance matrix $\mathbf{C_d}$, characterizes the errors between observed and simulated values and $\mathbf{m}_{prior}$ is a vector of *a priori* mean parameter values and finally, $\mathbf{C_m}$ is the inverse of the covariance matrix which defines the *a priori* probability distribution for the parameters. The first term computes a distance between observed ,and simulated data, using a norm defined by the inverse of the data covariance matrix $\mathbf{C_d}$. The *a priori* term accounts for differences between current parameter estimates and a priori means. The covariance matrix $\mathbf{C_m}$ and the mean values $\mathbf{m}_{prior}$ define *a priori* Gaussian probability distribution on the parameters. A more simplified form of Eq. (2.1) is commonly used for assisted history matching applications. As correlations between observed data are difficult to evaluate, only the diagonal terms are often accounted for in the covariance matrices [119]. When matching production and seismic data together without *a priori* information, Eq. (2.1) takes the following form as

$$F = \frac{1}{2}\sum_{j=1}^{n_{prod}} \frac{w_j^P}{n_j^t} \sum_{i=1}^{n_j^t} \left(\frac{\mathbf{P}_{i,j}^{obs} - \mathbf{P}_{i,j}^{sim}(\mathbf{m})}{\sigma_{i,j}^P} \right)^2 + \frac{1}{2}\sum_{j=1}^{n_{seis}} \frac{w_j^S}{n_j^s} \sum_{i=1}^{n_s^t} \left(\frac{\mathbf{S}_{i,j}^{obs} - \mathbf{S}_{i,j}^{sim}(\mathbf{m})}{\sigma_{i,j}^S} \right)^2 \quad (2.2)$$

where n_{prod} is the number of production data series to be matched defined by measurement type for a given well and n_j^t is the number of measurement times for the data series j; $\mathbf{P}_{i,j}^{obs}$ is an observation production data at time i for the data series j, $\mathbf{P}_{i,j}^{sim}$ is the simulated data from reservoir simulator at the same time with $\sigma_{i,j}^P$ is the standard deviation on production data errors. Similarly n_{prod} is the number of seismic data series to be matched defined by a seismic attribute over a given reservoir region and n_j^s, is the total number of observed seismic data values for the data series j with $\mathbf{S}_{i,j}^{obs}$ is an observed seismic data at time i for the data series j. $\mathbf{S}_{i,j}^{sim}$ is the simulated data from petro-elastic (or seismic modeling) model at the same time with $\sigma_{i,j}^S$ is the standard deviation on seismic data errors. The standard deviations $\sigma_{i,j}^P$ and $\sigma_{i,j}^S$ are used to normalize errors between simulated and observed responses by uncertainty ranges. As a result, different types of data can be combined to a single objective function F. This normalized form of standard deviation term can also be viewed as a weighting term, i.e., the larger the measurement error, the less is the contribution of the mismatch to the overall objective function. However, it is often necessary to adjust the relative contributions of the production and

seismic data to obtain a comparable impact in the optimization process. The weighting coefficients w_j^P and w_j^S applied on production and seismic data series, are used to balance the influence of the different terms. In order to integrate seismic impedance change data into the objective function in Eq. (2.2), it is required to have some equations to compute seismic impedance changes using outputs from the reservoir simulator. This task is achieved by inserting a series of rock physics models, e.g., Gassmann equation [68] in the reservoir simulator.

A key step in the development of a methodology for the integration of 4D seismic data in the history matching is the setup of an efficient method for the minimization of the objective function, F. The main computational bottleneck of any optimization method used in this case is given by the forward model step. Hence, local iterative methods with good convergence properties are usually preferred to slowly convergent global methods, even if the former ones are usually trapped by local minima [70]. Local methods require the gradient ∇F, and/or the hessian of the objective functions to construct iteratively a better approximation of the true solution for the minimization problem above.

It is important to note that production and the pressure data are characterized by a very low resolution in space and a high resolution in time. On the other hand, 4D seismic data has a very low resolution in time (several years between two surveys) and a medium to low resolution in space vertically, and a high resolution in space horizontally. Consequently, a correct formulation of the objective function F, not only requires a proper weighting between production/pressure data and seismic data, but also involves up-scaling (horizontally) and down-scaling (vertically) the seismic data in order to compare with the corresponding simulated data [70].

Another important aspect is the estimation of the data noise covariance matrix. When only production data are used, the noise covariance matrix is usually assumed to be diagonal with its entries equal to the noise variance of each individual datum because the production data sampling interval is generally larger than their correlation length in time space [2], so data noise is uncorrelated. It is probably not valid, however, to assume that the noise in the seismic data are not correlated. For the synthetic case study, a diagonal noise covariance matrix is a valid approximation as both observed and computed seismic impedance change data are generated at each grid block using only pressure and saturation in that grid block, but for the real case studies, it is more appropriate to use a non-diagonal banded error covariance matrix [43].

2.4 Challenges of 4D Seismic Data Integration

In general terms, the main challenge in linking time-lapse seismic measurements to dynamic reservoir models can be expressed as "How to benefit optimally from time-lapse seismic ?" Achieving this, undoubtedly results in a wider acceptance of time-lapse seismic as a standard technique. The challenge is divided into two main categories, see Table 2.1 which is taken from the work of Oldenziel [107]. The first is to link the seismic measurement directly to fluid-flow properties. The second is to fully integrate the time-lapse seismic data with reservoir engineering.

Link of 4D seismic data to reservoir properties	Integration of 4D seismic data with reservoir engineering
1. Repeatability - acquisition - re-processing - cross-equalization 2. Interpretation - rock physics - quantitative application - inversion 3. Lack of calibration data - validation of different methods 4. Decoupling of properties 5. Definition of time-lapse attribute	1. Integration - huge amount of data - incommensurable data - different scale 2. Quicken integration loop to increase benefit of data 3. Parametrization 4. Non-uniqueness 5. Automated history matching - misfit function - optimization algorithm - stopping criteria 6. Coupled reservoir-to-seismic simulator

Table 2.1: Main challenges of 4D seismic data integration.

Acquisition and processing of time-lapse seismic is a challenge. Even when both base and repeat surveys are shot for time-lapse purpose with identical

techniques, one needs to address whether to work with the surveys separately or utilize their difference. Special reprocessing is required for the latter to ensure the surveys can be compared to each other. It is not obvious for the reservoir engineers to interpret time-lapse induced fluid properties in same manner as the geophysicists; one challenge is to overcome this issue of paradigm shifts among different professionals. Rock physics is mainly based on models describing laboratory experiments or empirical relations. It does not describe the complex physics in case a rock is excited by a seismic wave. Each rock is different in texture on the smallest scale, which determines the actual behavior for the rock as a whole. Thus it is nearly impossible to have a perfect rock physics model. Abundant experience is available on how to obtain a better fit for the production and well test data. With regard to time-lapse seismic data, a trial-and-error approach is used as experience is lacking. It takes a long process to integrate time-lapse data with reservoir engineering applications from the time seismic data are collected. The benefit and information content of time-lapse seismic information is much higher when it could be made available in a shorter time span.

Chapter 3

Reservoir Characterization: History Matching

The reservoir characterization is the process of building a model by integration of all data available at different stages of development of a petroleum reservoir. The dynamic reservoir model permits us to evaluate different possible exploration scenarios. In order to evaluate the economic potential of a petroleum reservoir, a forecast of the hydrocarbon production from the reservoir under various recovery strategies is of paramount importance. In order to minimize the uncertainty in the reservoir characterization, and thereby also in the forecast, all available data should be conditioned to in a mathematically consistent framework. The available data collected from the reservoir typically consists of acoustic seismic surveys, well logs and a production history data. To forecast the hydrocarbon production, these data are used to estimate the parameters and state in a mathematical model of the reservoir, which in turn is fed into a numerical reservoir simulator to forecast the hydrocarbon production under a given recovery strategy, see **Fig.** 3.1.

Despite the seemingly abundance of data, estimating the state and parameters in the mathematical model of the reservoir is not a well-posed problem. It might not be possible to find a single model that match the observed data exactly. On the contrary, there might be multiple models that match the observed data. Clearly the estimation of the state and parameters in reservoir model must be treated as an inverse problem, see Tarantola [142]. Unfortunately, most inverse problems are highly under-determined, which means that the number of system parameters is much greater than the number of system outputs. In this case, there are an unlimited number of combinations of the system parameters, which, if this information is fed back to the system, cannot reproduce the same system outputs. This

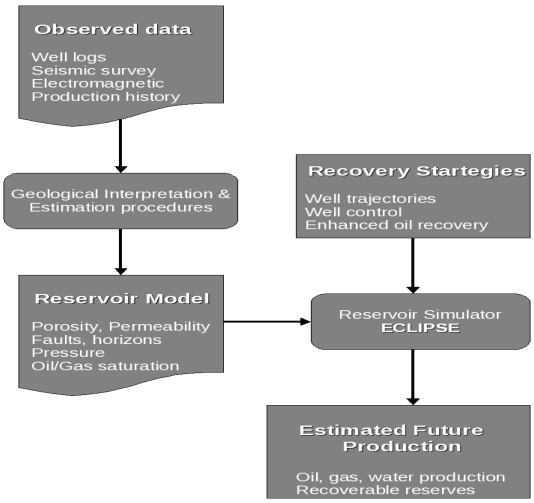


Figure 3.1: Basic steps of reservoir characterization and forecasting.

non-unique feature makes it necessary to find an appropriate solution to an inverse problem with aid of some particular criteria.

3.1 General Statement of History Matching

When conditioning a reservoir model to dynamic data, one needs to take the effect of producing the reservoir into account. If the dynamic data to be conditioned on is the production history, this is commonly known as history matching. The primary objective of history matching is to modify a prior model for the reservoir such that the updated model reflects the available production data and the uncertainties in production forecasts are reduced. This process mainly consists of the following steps [158]:

- Identifying model parameters that could be modified to affect history matching and perform proper parametrization.
- Defining a suitable objective function for optimization.
- Proper selection and design of an optimization technique in order to reduce the objective function to minimum.

- Tracking computational cost associated with flow simulations and within the selected optimization technique.

In the classical history match loop, the reliability of numerical reservoir model is improved through the minimization of a similar weighted least-squares type objective function like Eq. (2.2), a new objective function F_{prod} can be defined which measures the misfit between computed and observed well production data, as

$$F_{prod} = \frac{1}{2} \sum_{j=1}^{n_{prod}} \frac{w_j^P}{n_j^t} \sum_{i=1}^{n_j^t} \left(\frac{\mathbf{P}_{i,j}^{obs} - \mathbf{P}_{i,j}^{sim}(\mathbf{m})}{\sigma_{i,j}^P} \right)^2 . \tag{3.1}$$

The parameters in Eq. (3.1) already defined in the chapter 2, see Eq. (2.2). As assisted history matching can be thought of as a minimization problem, efficient optimization algorithm is necessary to find the optimal parameters from the Eq. (3.1). A plethora of optimization methods are in use in reserver engineering community, and hence, various types of optimization-based history matching techniques will be discussed in the subsequent sections in this chapter.

3.2 Bayesian Framework

The most general theory of inverse problems can be achieved when using a probabilistic point of view [87]. 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 measurements, and an uncertain relation between the data and the model parameters, a posterior distribution can be established like the following statement as

$$\textbf{Posterior} \quad \propto \quad \textbf{Likelihood} \times \textbf{Prior}$$

and represents the solution of the inverse problem. In order 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 [110]. The motivation for taking a Bayesian approach to the history matching problem is to consistently honor prior geological

knowledge, accounting for the observation error in the production data and obtain multiple history matched models so that the forecast uncertainty can be assessed.

Suppose that the reservoir parameters, e.g., porosity and permeability are random variables [43]. The probability of occurrence of any particular configuration of flow properties can then be characterized using a probability density function $g(\mathbf{m})$. In our applications the random field is assumed to be Gaussian. Under this assumption, the PDF of the reservoir model can be written as

$$g(\mathbf{m}) \propto a \exp\left(-\frac{1}{2}(\mathbf{m} - \mathbf{m}_{prior})^T \mathbf{C}_M^{-1} (\mathbf{m} - \mathbf{m}_{prior})\right), \tag{3.2}$$

where $\mathbf{m}$ is the vector of model parameters, a is a constant and $\mathbf{m}_{prior}$ is the best estimate for parameters based on prior information about the field, and $\mathbf{C}_M$, is the model variable covariance matrix. The prior information contains general knowledge about the reservoir, such as expected porosity and permeability. The matrix $\mathbf{C}_M$, is usually constructed through geostatistical tools. Its diagonal entries are variances of all model parameters.

Observed data gathered during exploration and production, such as production data from wells and seismic data from seismic surveys, can be written as $\mathbf{d}_{obs} = \mathbf{d}_{true} + \epsilon$, where ϵ, is measurement noise. The addition of noise term accounts for the fact that the recorded observations are corrupted by noise due to limitations of measurement tools. If the measurement noise is also assumed to be Gaussian with mean equal to zero, then the PDF of the observation noise can be written as

$$g(\mathbf{d}_{obs}) = b \exp\left(-\frac{1}{2}\epsilon^T \mathbf{C}_D^{-1}\epsilon\right), \tag{3.3}$$

$$= b \exp\left(-\frac{1}{2}(\mathbf{d}_{true} - \mathbf{d}_{obs})^T \mathbf{C}_D^{-1}(\mathbf{d}_{true} - \mathbf{d}_{obs})\right), \tag{3.4}$$

where b, is a constant and $\mathbf{C}_D$, is the measurement noise covariance matrix, which defines correlations among noise contained in observed data, diagonal for production data but non-diagonal for real field seismic data [2]. The diagonal entries of $\mathbf{C}_D$ are variances of the measurement noise. On the other hand, if the true model $\mathbf{m}_{true}$, is given to a reservoir simulator, then some "true" observations can be computed as $\mathbf{d}_{true} = g(\mathbf{m}_{true})$. Since the measurement noise ϵ, is random, the observations given true model $\mathbf{m}_{true}$, are also random and can be described using the conditional PDF written as

$$g(\mathbf{d}_{obs}|\mathbf{m}_{true}) = b \exp\left(-\frac{1}{2}(g(\mathbf{m}_{true}) - \mathbf{d}_{obs})^T \mathbf{C}_D^{-1}(g(\mathbf{m}_{true}) - \mathbf{d}_{obs})\right), \tag{3.5}$$

where $g(\mathbf{m}_{true})$, represents forward simulation run. According to Baye's theorem, the conditional PDF for the model parameters $\mathbf{m}$, given observations $\mathbf{d}_{obs}$, can be written as

$$g(\mathbf{m}|\mathbf{d}_{obs}) = \frac{g(\mathbf{d}_{obs}|\mathbf{m})\ g(\mathbf{m})}{g(\mathbf{d}_{obs})}, \tag{3.6}$$

$$= \frac{g(\mathbf{d}_{obs}|\mathbf{m})\ g(\mathbf{m})}{\int g(\mathbf{d}_{obs}|u)g(u)\mathrm{du}}, \tag{3.7}$$

where $g(\mathbf{d}_{obs})$, is the PDF of the observation. Inserting two equations in this formulation, the conditional PDF can be written as

$$g(\mathbf{m}|\mathbf{d}_{obs}) = c\ \exp\left(-\frac{1}{2}(\mathbf{m}-\mathbf{m}_{prior})^T\ \mathbf{C}_M^{-1}\ (\mathbf{m}-\mathbf{m}_{prior})\right) \times\left(-\frac{1}{2}(g(\mathbf{m}_{true})-\mathbf{d}_{obs})^T\ \mathbf{C}_D^{-1}(g(\mathbf{m}_{true})-\mathbf{d}_{obs})\right), \tag{3.8}$$

where c, is the normalizing constant. The PDF in the above equation is called the posterior PDF. For our automatic history matching problem, we intend to generate an estimate of $\mathbf{m}$, that has the maximum probability, i.e., the maximum a posterior (MAP) estimate [14]. Obviously such an estimate can be obtained by minimizing the arguments of the exponential term in the equation which provides us with the objective function Q, in the Bayesian framework as

$$Q(\mathbf{m}) = \frac{1}{2}(\mathbf{m}-\mathbf{m}_{prior})^T\ \mathbf{C}_M^{-1}\ (\mathbf{m}-\mathbf{m}_{prior}) + \frac{1}{2}(g(\mathbf{m})-\mathbf{d}_{obs})^T\ \mathbf{C}_D^{-1}(g(\mathbf{m})-\mathbf{d}_{obs}) \tag{3.9}$$

$$= Q_m(\mathbf{m}) + Q_d(\mathbf{m}), \tag{3.10}$$

where $Q_m(\mathbf{m})$ is the model mismatch term and $Q_d(\mathbf{m})$ is the data mismatch term . The total objective function Q, is conditioned to the observed data and the prior information. Also the model mismatch term provides normalization for the Hessian matrix [43].

From the perspective of reservoir management, values of permeability and porosity at some specific places of the reservoir model are not very meaningful, especially considering that the results of automatic history matching problem are non-unique. The more important problem is to characterize uncertainties of future reservoir performance. To quantify the prediction uncertainty of the model parameters, an assessment of the full posterior distribution $g(\mathbf{m}|\mathbf{d}_{obs})$, 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 [131]. For more efficient sampling algorithm, the Randomized Maximum Likelihood (RML) method can also be used [108].

3.3 Different Methods for History Matching

As automatic history matching can be thought of as a minimization problem, efficient optimization algorithm is necessary. Based on this criterion, history matching techniques can broadly be classified into the following categories:

1. Gradient based methods
2. Global optimization methods
3. Geo-statistical techniques
4. Streamline based methodologies
5. Ensemble-based methods

Some gradient methods require only the gradient of the objective function, while other methods require also the second derivative (the Hessian). To approximate the Hessian, one needs the sensitivity coefficients, which are the derivatives of the data (the sensitivity coefficients are not the gradient of the objective function). Some widely used gradient based minimization methods include Steepest Descent method, Gauss-Newton method, Levenberg-Marquardt algorithm and Conjugate Gradient (CG) method [43]. In order to calculate sensitivity coefficients for reservoir models of the order of thousands of grid cells, efficient strategies are necessary. One such approach for three phase flow problems is the adjoint method [93]. One major drawback of gradient-based methods is that they do not account for the spatial covariance model exhibited by model parameters [158]. Further, these approaches are computationally expensive since they involve the computation of sensitivities. Also, these methods frequently converge to a local minima of the objective function rather than a global minimum.

The main idea of global optimization methods is to transform an initial configuration of a problem into a configuration close to the optimum where the objective function is at a minimum. One such approach is the simulated annealing [39]. Genetic algorithm has been used in stochastic

reservoir modeling [126]. Although non-gradient methods such as genetic algorithm can be used to find the global minimization point, they may require thousands or millions of simulation runs to converge. When processing large scale or even medium scale automatic history matching problems, this intensive demanding on simulation computation makes the global optimization based methods impractical. The combination of global and local optimization has been investigated by Schulze-Riegert et al. [124]. In this approach, one uses global optimization methods to find the initial better matched reservoir parameters. Once a location of the acceptable parameters in search space is found, local optimization methods are used to fine tune the model.

The gradual deformation process causes the reservoir model to evolve while preserving the overall statistical characteristics of the target reservoir model [115]. Deformation is controlled by a limited number of parameters; thereby improve the efficiency of the process. The pilot point method [114] selects a series of optimal points in the field and employs Kriging to interpolate among the pilot points to get field distributions of model parameters. But this method has over-shooting problems at the pilot points [43].

Streamlines-based history matching has been proposed by generalized travel time inversion [157]. Instead of matching the production data directly, a "travel time shift" at each well divided by maximizing the cross correlation between the observed and calculated responses is minimized. An optimal control method is used to calibrate the sensitivity of the travel time with respect to the reservoir parameters.

One of the more successful Bayesian approaches to the history matching problem is the ensemble Kalman filler (abbreviated EnKF). The Ensemble Kalman Filter is a Monte Carlo based sequential Bayesian data assimilation method that relies on repeated random sampling to compute a result. In the context of reservoir simulation, an ensemble of reservoir models is sequentially updated during flow simulation to integrate production data, and if available, seismic data as well. In the chapter 5, we will provide a through introduction of EnKF, and in particular, we will delineate the work flow for reservoir history matching with EnKF.

3.3.1 Traditional vs. EnKF-based History Matching

Traditional history matching methods suffer from one or more from the following drawbacks as mentioned in the work of Wen and Chen [156]. First drawback is the fact that for traditional history matching methods,

production data for the entire history are matched at the same time, and hence, repeated flow simulations of the entire history are needed. This makes history matching extremely time-consuming. In **Fig.** 3.2, we have shown a comparative analysis between EnKF-based history matching and a traditional history matching method. Note the sequential updating fea-

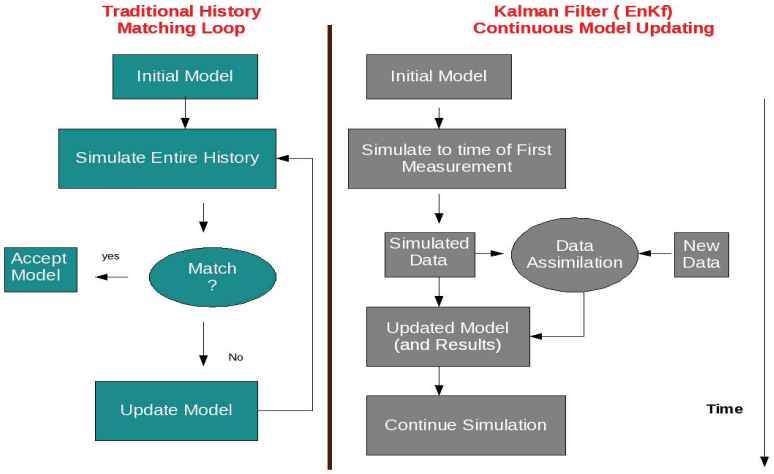


Figure 3.2: Traditional vs EnKF-based history matching.

ture of EnKF compared to the traditional method. Secondly, gradient based history matching methods require sensitivity coefficient calculations and minimization which are complicated, CPU-intensive and often, trapped by local minima. Finally, it is difficult to assess uncertainty with traditional methods and may involve repeating the history matching process with different initial models. As EnKF is a sequential model updating technique, it provides a robust framework for history matching. A limitation of EnKF is the fact that its computations are based on first and second order moments, and there are problems that are difficult to handle, particularly when the probability distributions are multi-modal (e.g., when representing a bi-modal channel facies distribution) [76]. Despite outstanding theoretical and practical questions on the applicability of the EnKF for highly non-linear and non-Gaussian reservoir problems, there have been promising results of EnKF as an history matching tool, see Aanonsen et al [3].

Chapter 4

Forward Models

In order to compare the field observation with simulated data, one needs to perform forward modeling at different stages of reservoir modeling work flow. Given a complete description of a physical system, we can forecast or predict the outcome of some measurement of interest. The method of predicting the results of a measurement can be termed as forward model. Examples of forward modeling can be mentioned as, (a) production forecast in reservoir simulator (fluid flow equations in porous medium) and (b) simulation of synthetic seismograms (seismic modeling).

4.1 Flow Equations in Porous Media

In reservoir simulation, the flow equations governing the multi-phase flow in porous media are obtained by combining Darcy's Law and the equation of mass conservation. In order to describe the flow of the phases, we have to describe their interaction as well. The interaction between the phases is described by saturations and capillary pressure. If water and oil are both present in the pores of the medium, we have a two phase flow system for which the Darcy's equation becomes

$$\text{Darcy's law:} \qquad u_j = -\mathbf{K}\frac{\kappa_{rj}}{\mu_i}(\nabla P_j + \rho_j g \hat{e}_z), \tag{4.1}$$

where u, is average velocity, $\mathbf{K}$, is permeability tensor, $\hat{e}_z$ is the unit vector in vertical direction, ∇P, is the pressure gradient, ρ, fluid density and j corresponds to each of the phases. Each phase will occupy some pore space. This yields a reduction in the permeability due to the presence of other phases. So we introduced relative permeability, κ_{rj}. The relative permeability is generally obtained from experiments but formulas depending on saturations, e.g., Corey-model [36] are also available.

Conservation laws describe how a physical quantity is preserved within a closed system. This means that the accumulation of this quantity must be balanced by fluid flux and source or sink. If, for two phase flow, mass balance is expressed in terms of fluid density, and flux in terms of specific discharge, the differential form of mass conservation law in this case takes the following form

$$\text{Mass conservation:} \qquad \nabla \cdot (\rho_j u_j) + \phi \frac{\partial \rho_j S_j}{\partial t} = 0, \tag{4.2}$$

where ϕ, is the effective porosity of the medium and S_j, denotes the corresponding saturation for each phases. The total saturation for both phases are expressed as:

$$\text{Saturation:} \qquad S_w + S_o = 1. \tag{4.3}$$

The relationship $\rho = \rho(P, T)$, between fluid density, temperature and pressure is called an equation of state. For isothermal processes, i.e. processes with constant temperature, the equation of state becomes

$$\text{Equation of state:} \qquad \rho_j = \rho_j(P_j). \tag{4.4}$$

Fluid viscosity μ, is a measure of the ability of a fluid to flow. Like density, viscosity vary as a function of temperature and pressure, $\mu = \mu(P, T)$. This can be expressed as

$$\text{Viscosity:} \qquad \mu_j = \mu_j(P_j). \tag{4.5}$$

When two fluids with different properties are in contact, forces will act between them. The molecules in each phase are forced closer together at the boundary by surface tension. The surface tension between the phases are balanced by a pressure difference called the capillary pressure P_c. When the capillary pressure increases, larger pores are filled, and the amount of fluid in the porous material increases. The capillary pressure for two phase flow can be written as

$$\text{Capilary pressure:} \qquad P_o - P_w = P_c(S_w). \tag{4.6}$$

The unknowns are u_j, P_j, ρ_j, μ_j, and S_j, for each phase $j = o, w$. From the equations stated above, we can now compute solutions for the equations for two phase flow. For general multi-phase flow equations, we refer to Aziz and Settari [16].

4.2 Rock Physics

By far the most widely used and successful method for determining the effects of production-induced fluid change is the Gassmann substitution method, see [15]. Generally, when a rock is loaded under an increment of compression, such as from a passing seismic wave, a increment of pore pressure change is induced, which resists the compression and therefore stiffens the rock. The low frequency Gassmann theory [68] predicts the resulting increase in effective bulk modulus $\mathbf{K}_{sat}$, and shear modulus $\mathbf{G}_{sat}$, of the saturated rock through the following equation

$$\mathbf{K}_{sat} = \mathbf{K}_{dry} + \frac{(1 - \frac{\mathbf{K}_{dry}}{\mathbf{K}_s})^2}{\frac{\phi}{\mathbf{K}_{fluid}} + \frac{1-\phi}{\mathbf{K}_s} - \frac{\mathbf{K}_{dry}}{\mathbf{K}_s^2}}; \quad \mathbf{G}_{sat} = \mathbf{G}_{dry}, \tag{4.7}$$

where $\mathbf{K}$, is the effective bulk modulus with subscript referring to dry rock frame, minerals and fluid mixture. $\mathbf{G}$, is the corresponding shear modulus with different constituents and ϕ is the porosity. The effective density $\boldsymbol{\rho}_{sat}$, can be defined as a volume average of the mineral density and the fluid density. Gassmann's equation assumes a homogeneous mineral modulus and statistical isotropy of the pore space but is free of assumptions about the pore geometry. Most importantly, it is valid only at sufficiently low frequencies such that the induced pore pressures are equilibrated throughout the pore space, i.e., there is sufficient time for the pore fluid to flow and eliminate wave-induced pore pressure gradients [98].

In order to model changes in seismic response that result from fluid changes, we first need to know the initial and target fluid properties. The effective density ρ_{sat}, may be defined as a volume average of the mineral density and the fluid density, as

$$\rho_{sat} = \phi \left(S_w\rho_w + S_g\rho_g + S_o\rho_o\right) + (1 - \phi)\, \rho_s, \tag{4.8}$$

where S_w, S_g, S_o are respectively water, gas and oil saturations and ρ_w, ρ_g, and ρ_o are the corresponding densities of these three phases. If $\mathbf{K}_{sat}$, $\mathbf{G}_{sat}$ and ρ_{sat} are obtained from the above equations, the acoustic properties of saturated porous media, e.g., isotropic compressional velocity $\mathbf{V}_P$, and shear wave velocity $\mathbf{V}_S$, can be computed as $\mathbf{V}_P = \sqrt{\frac{\mathbf{K}_{sat}+(4\mathbf{G}_{sat}/3)}{\rho_{sat}}}$, and $\mathbf{V}_S = \sqrt{\frac{\mathbf{G}_{sat}}{\rho_{sat}}}$. The acoustic impedance can be defined as $\mathbf{I}_p = \rho_{sat}\mathbf{V}_P$ and the Poisson's ratio is expressed as $\mathbf{PR} = \frac{\mathbf{V}_P^2 - 2\mathbf{V}_S^2}{2\mathbf{V}_P^2 - \mathbf{V}_s^2}$. The effect of reservoir property changes on $\mathbf{I}_p$, has a combined response to pressure and saturation change.

The polarity of the response depends on whether the pressure is increasing or decreasing and the difference between the fluid properties at the start and end of the period.

Although rocks are extremely complicated, their elastic properties depend largely on a combination of connectivity and elasticity and density of what is connected. As a result, many rules of thumb can be explained by such combinations, see the work of Wang [152]. Statistically, elastic properties decrease as porosity increases. From lithology point of view, unconsolidated sands may have high $\mathbf{V}_p/\mathbf{V}_s$ ratio. Full saturation of a liquid in a rock increases the compressional seismic properties and shear impedance but decreases $\mathbf{V}_s$, resulting in increased $\mathbf{V}_p/\mathbf{V}_s$. When gas is introduced to a fully liquid-saturated rock, the compressional seismic properties and shear impedance all decreases, but $\mathbf{V}_s$ increases, resulting in lower $\mathbf{V}_p/\mathbf{V}_s$. Seismic properties in all rocks increase as effective pressure (difference between over-burden and pore pressures) increases [74]. Also better compacted rocks have higher elastic properties due to better connectivity and contact.

4.2.1 Background on Elastic Bounds

Gassmann's model requires homogeneous mixture. If the rock contains more than one mineral phase, for instance clay rich sandstone, the different mineral phases must be mixed and averaged. Many "effective medium" models have been proposed, attempting to describe theoretically the effective elastic modulus of rocks and sediments [15]. Theoretically when trying to predict the effective elastic modulus of a mixture of grains and pores, three aspects must be satisfied: (i) the volume fractions of the various phases, (ii) the elastic modulus of the various phases, and (iii) the geometric details of how the phases are arranged relative to each other. Regardless of the volume fractions, the effective modulus of a mixture of solids will fall between an upper and lower boundary, referred to as the upper and lower bounds. If only the conditions (i) and (ii) are determined, the prediction of the upper and lower bounds is the most exact technique that can be performed. Thus the bounds provide a useful and elegant framework for velocity-porosity relations.

Hashin-Shtrikman Bounds

The Hashin-Shtrikman bounds [75] give the upper and lower limits for elastic constants of isotropic rocks. The Hashin-Shtrikman bounds are

regarded as the best bounds giving the narrowest possible range, without information on the geometries of the constituents. For a mixture of two solids, the Hashin-Shtrikman bounds are given by

$$\begin{aligned}\mathbf{K}_{uHS} &= \mathbf{K}_1 + \frac{f_2}{(\mathbf{K}_2 - \mathbf{K}_1)^{-1} + f_1(\mathbf{K}_1 + 4\mathbf{G}_1/3)^{-1}};\\ \mathbf{K}_{lHS} &= \mathbf{K}_2 + \frac{f_1}{(\mathbf{K}_1 - \mathbf{K}_2)^{-1} + f_2(\mathbf{K}_2 + 4\mathbf{G}_2/3)^{-1}},\end{aligned} \tag{4.9}$$

with $\mathbf{K}_1$, $\mathbf{K}_2$, bulk modulus of individual phases and f_1, and f_2 volume fractions of individual phases. Similar relationships can be obtained for shear modulus as

$$\begin{aligned}\mathbf{G}_{uHS} &= \mathbf{G}_1 + \frac{f_2}{(\mathbf{G}_2 - \mathbf{G}_1)^{-1} + 2f_1(\mathbf{K}_1 + 2\mathbf{G}_1)/[5\mathbf{G}_1(\mathbf{K}_1 + 4\mathbf{G}_1/3)]};\\ \mathbf{G}_{lHS} &= \mathbf{G}_2 + \frac{f_1}{(\mathbf{G}_1 - \mathbf{G}_2)^{-1} + 2f_2(\mathbf{K}_2 + 2\mathbf{G}_2)/[5\mathbf{G}_2(\mathbf{K}_2 + 4\mathbf{G}_2/3)]},\end{aligned} \tag{4.10}$$

where $\mathbf{G}_2$, and $\mathbf{G}_2$, are bulk modulus of individual phases . Here the stiffest material is subscripted 1 in the expressions above. The upper bound realizes a stiff rock that will be less elastically affected by fluid substitution, and geologically this bound refers to cemented rocks. The lower bound can be utilized as a sorting trend for un-cemented or unconsolidated rocks [52]. The separation between upper and lower bounds depends on how elastically the different solids are. The bounds are often fairly similar when mixing solids, since the elastic modulus of common materials are usually within a factor of two of each other. Since many effective-medium models assume a homogeneous mineral modulus, it is often useful (and adequate) to represent a mixed mineralogy with an *average mineral* modulus [15], equal either to one of the bounds computed for the mix of solids or to their average $\mathbf{K}_s = \frac{(\mathbf{K}_{uHS}+\mathbf{K}_{lHS})}{2}$, and $\mathbf{G}_s = \frac{(\mathbf{G}_{uHS}+\mathbf{G}_{lHS})}{2}$. On the other hand, when the constituent solids are very different, e.g., for minerals and pore fluids, then the bounds are quite separated, and we lose some of the predictive values.

4.2.2 Contact Models

In case of rock physics perspective, the contact theories are used to calculate the effective elastic properties of unconsolidated sediments or cemented sediments. In contact models, rocks are assumed to be a collection of separate grains. Their elastic properties are determined by the stiffness and deformability of their grain-to-grain contacts.The majority of contact models are based on the Hertz-Mindlin contact theory.

Hertz-Mindlin Contact Theory

The contact theory of Hertz-Mindlin [101] describes seismic parameter changes caused by pressure changes. This model is used to calculate the elastic modulus of sand at critical porosity. An extended form of this theory is achieved by using the modified Hashin-Shtrikman lower bound where one end member has zero porosity and the modulus of the solid phase; the other end member has critical porosity and a pressure-dependent modulus given by the Hertz-Mindlin contact theory. This Hertz-Mindlin-Hashin-Shtrikman (HMHS) model allows to account for the noticeable pressure dependence normally observed in sands [51].

The elastic modulus of the dry well-sorted end member at critical porosity, ϕ_c = 40% for sands [98], are modeled as an elastic sphere subject to confining pressure by the Hertz-Mindlin theory as follows

$$\mathbf{K}_{HM} = \left[\frac{n^2(1-\phi_c)^2\ \mathbf{G}_s^2}{18\pi(1-\nu)^2}\ P_{\text{eff}} \right]^{\frac{1}{3}}, \tag{4.11}$$

$$\mathbf{G}_{HM} = \frac{5-4\nu}{5(2-\nu)} \left[\frac{3n^2(1-\phi_c)^2\ \mathbf{G}_s^2}{2\pi^2(1-\nu)^2}\ P_{\text{eff}} \right]^{\frac{1}{3}}, \tag{4.12}$$

where $\mathbf{K}_{HM}$, and $\mathbf{G}_{HM}$, are dry rock bulk and shear modulus, respectively, at critical porosity (depositional porosity); P_{eff}, is the effective pressure (i.e., the difference between the overburden pressure and the pore pressure); $\mathbf{G}_s$, and ν are the shear modulus and Poisson's ratio of the solid phase and n, is the coordination number (the average number of contacts per grain).

The elastic modulus of the poorly sorted sands are with porosities between 0 and ϕ_c are "interpolated" between the mineral point and the well-sorted member using the lower Hashin-Shtrikman bound. At porosity ϕ, the concentration of the pure solid phase (added to the sphere pack to decrease porosity) in the rock is $(1-\phi/\phi_c)$ and that of the original sphere-pack phase is ϕ/ϕ_c. Then the bulk ($\mathbf{K}_{dry}$) and the shear ($\mathbf{G}_{dry}$) modulus of the dry unconsolidated sand mixture are defined as

$$\mathbf{K}_{dry} = \left[\frac{\phi/\phi_c}{\mathbf{K}_{HM} + 4\mathbf{G}_{HM}/3} + \frac{1-\phi/\phi_c}{\mathbf{K}_s + 4\mathbf{G}_{HM}/3} \right]^{-1} - \frac{4}{3}\mathbf{G}_{HM}, \tag{4.13}$$

$$\mathbf{G}_{dry} = \left[\frac{\phi/\phi_c}{\mathbf{G}_{HM} + z} + \frac{1-\phi/\phi_c}{\mathbf{G}_s + z} \right]^{-1} - z, \tag{4.14}$$

where

$$z = \frac{\mathbf{G}_{HM}}{6} \left(\frac{9\mathbf{K}_{HM} + 8\mathbf{G}_{HM}}{\mathbf{K}_{HM} + 2\mathbf{G}_{HM})} \right).$$

The modified Hertz-Mindlin-Hashin-Shtrikman (HMHS) method provides good estimates for velocities in un-cemented sands [98].

4.2.3 Effective Pore Fluid Properties

Pore fluids significantly influence the seismic properties of rocks. Pore fluids have properties that vary substantially, but systematically, with composition, pressure and temperature [18]. Gas and oil density and modulus, as well as oil viscosity, increase with molecular weight and pressure, and decrease with temperature. Gas viscosity has a similar behavior, except at higher temperatures and lower pressures, where the viscosity will increase slightly with increasing temperature. Brine modulus, density, and viscosities increase with increasing salt content and pressure. Batzle and Wang [18] has described these individual phase effects on hydrocarbon generation. From an exploration standpoint, pore fluid mixtures of liquid and gas phases are extremely important. The density of a mixture is straightforward. Mass balance requires an arithmetic average of the separate fluid phases. The effective fluid density ρ_f, is the saturation weighted average of the individual fluid component densities, and can be expressed as

$$\rho_f = S_w\rho_w + S_g\rho_g + S_o\rho_o. \tag{4.15}$$

The effective modulus of the mixed phase fluid is easily calculable if we assume that the pressures in the two phases are always equal. We must also assume that there is no mass interchange between the two phases during the passage of a sonic wave [12]. Thus, the bulk modulus of the mixed pore fluid phase $\mathbf{K}_{\text{fl}}$, can be estimated by inverse bulk modulus averaging (also known as Wood's equation). This is simply a saturation weighted harmonic average of the individual fluid bulk modulus, that is

$$\frac{1}{\mathbf{K}_{\text{fl}}} = \frac{S_w}{\mathbf{K}_w} + \frac{S_o}{\mathbf{K}_o} + \frac{S_g}{\mathbf{K}_g}, \tag{4.16}$$

where $\mathbf{K}_{w/o/g}$, is the bulk modulus of the water/oil/gas. Thus, if we know the properties of the individual fluids and their volume fraction, the properties of the mixture can be calculated.

Rock Physics Work Flow

Starting from the Gassmann equation, we have now discussed the necessary inputs for fluid substitution formula. With that end in view, we can perform the rock physics modeling for unconsolidated sands in the following steps (**Fig.** 4.1 is taken from Davies and Mave [38]):

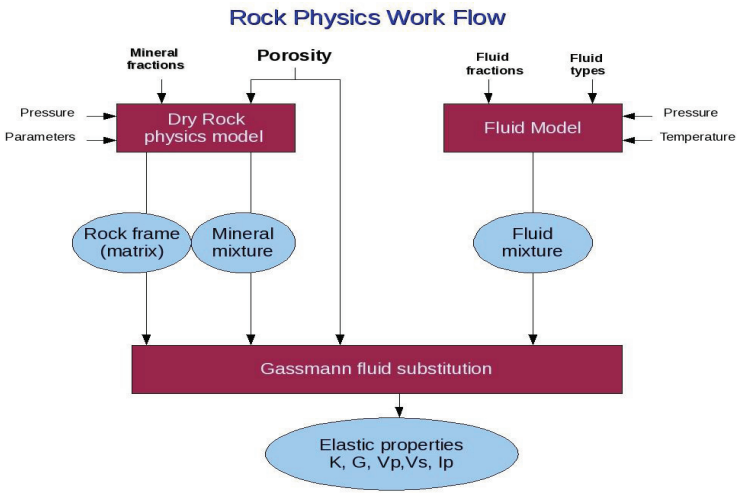


Figure 4.1: Basic steps of rock physics modeling.

1. At first we estimate the mineral properties $\mathbf{K}_s$ and $\mathbf{G}_s$ by using the Hashin-Shtrikman average bounds for solid minerals (Eq. (4.9) and Eq. (4.10)). The volume of mineral fractions can be calculated with the aid of wire line logs, thin section analysis or from empirical correlation between porosity and volume fractions.

2. Second step estimates the dry rock properties ($\mathbf{K}_{HM}$ and $\mathbf{G}_{HM}$) at high porosity member, ϕ_c using the Hertz-Mindlin contact theory (Eq. (4.11) and Eq. (4.12)). We can incorporate the pressure effect on dry frame modulus at this level as well.

3. Next step includes the interpolation for dry rock bulk and shear modulus ($\mathbf{K}_{dry}$ and $\mathbf{G}_{dry}$), before fluid substitution, using the heuristic Hashin-Shtrikman lower bounds (Eq. (4.13) and Eq. (4.14)).

4. At this step, we calculate the effective pore fluid properties ($\mathbf{K}_{\mathrm{fl}}$) based on Batzle-Wang relations and Wood's formulae (Eq. (4.16)).

5. Finally, using the effective pore fluid properties based on Batzle-wang relations, we perform the Gassmann fluid substitution (Eq. (4.7) and Eq. (4.8)) and calculate the seismic signatures, e.g. $\mathbf{V}_p$ and $\mathbf{I}_p$.

4.3 Seismic Modeling

Forward modeling of 4D seismic involves three processes: reservoir simulation, rock physics conversion and seismic modeling. Seismic modeling is done in two main steps. First, reservoir parameters are converted to seismic parameters, e.g., saturated P-wave velocity or density of fluid mixture by using the rock physics models. Then, synthetic seismic sections are calculated based on these seismic parameters. Seismic modeling involves the calculation of synthetic seismic traces. Numerous methods for seismic modeling exist including the finite difference method, reflectivity method and a simple convolution-based method. For simplicity, we restrict ourselves to primary waves and post-stack data, and therefore use simple convolution-based forward modeling [127]. 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. A synthetic seismogram or time series, $E(x,t)$ may be considered as the assumed source function $\omega(t)$, convolved with a reflectivity function $r(x,t)$, representing the contrasts in acoustic impedance in the layered model, and $E(x,t)$ can be expressed as

$$E(x,t) = \omega(t) \otimes r(x,t), \tag{4.17}$$

where x is the lateral location and t is the two-way vertical seismic travel time. This process is also called 1D convolution. For 4D seismic modeling, it is assumed that the wavelet does not change with production time. The reflection coefficients are a function of seismic velocity and density and are therefore dependent on the production time step. Eq. (4.17) is used to generate the synthetic seismic response for different production time steps. The earth's impulse response is based on the combined amplitude effects of transmission losses, reflection strength and (spherical) divergence. Some of the effects that are generally not considered in 1D convolution are: (1) multiple reflections; (2) conversion between seismic phases (S-waves, surface waves); (3) spatial interference due to limited bandwidth; (4) neglecting acquisition footprints, i.e., in terms of time-lapse seismic data this refers to ideal repeatability and so on. However, in general the geological structure in both overburden and reservoir may be more complex. Hence, an accurate modeling is required, and typically finite-difference modeling (FDM) [150] or ray-tracing methods [30] are applied.

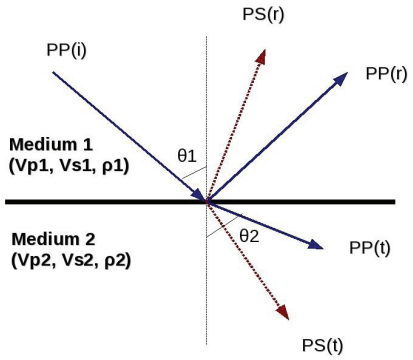


Figure 4.2: Reflections and transmissions at a single interface between two elastic half-space media for an incident plane P-wave, PP(i). There will be both a reflected P-wave, PP(r), and a transmitted P-wave, PP(t). In addition, a reflected S-wave, PS(r), and a transmitted S-wave, PS(t), will be produced.

4.3.1 Seismic Response of Layered Earth Model

In a homogeneous medium, waves propagate as spherical wavefronts, and the wave amplitude decays with distance from the source [127]. On the other hand, plane waves do not have geometric divergences loss in a homogeneous medium, and in an inhomogeneous-layered medium, plane waves undergo reflection and transmission losses affecting their amplitudes. Thus, plane-wave reflection coefficient, which determines partitioning of seismic energy separating two media in welded contact, acts as one of the most important factors affecting seismic amplitudes. The reflection coefficient of plane elastic waves as a function of reflection angle is described by complicated Zoeppritz equation [161]. For analysis of P-wave reflections, a well-known linear approximation is given by Aki and Richards formulation [6], assuming weak layer contrasts and is described as

$$\mathbf{r}_{pp}(\theta_1) \approx \frac{1}{2}\left(1 - 4\mathbf{p}^2\bar{\mathbf{V}}_S^2\right)\frac{\Delta\boldsymbol{\rho}}{\bar{\boldsymbol{\rho}}} + \frac{1}{2\cos^2\theta}\frac{\Delta\mathbf{V}_P}{\bar{\mathbf{V}}_P} - 4\mathbf{p}^2\bar{\mathbf{V}}_S^2\frac{\Delta\mathbf{V}_S}{\bar{\mathbf{V}}_S}, \tag{4.18}$$

where

$$\begin{aligned} &\mathbf{p} = \tfrac{\sin\theta_1}{V_{P1}} && \theta = (\theta_1 + \theta_2)/2 \approx \theta_1 \\ &\Delta\boldsymbol{\rho} = \boldsymbol{\rho}_2 - \boldsymbol{\rho}_1 && \bar{\boldsymbol{\rho}} = (\boldsymbol{\rho}_1 + \boldsymbol{\rho}_2)/2 \\ &\Delta\mathbf{V}_P = \mathbf{V}_{P2} - \mathbf{V}_{P1} && \bar{\mathbf{V}}_P = (\mathbf{V}_{P1} + \mathbf{V}_{P2})/2 \\ &\Delta\mathbf{V}_S = \mathbf{V}_{S2} - \mathbf{V}_{S1} && \bar{\mathbf{V}}_S = (\mathbf{V}_{S1} + \mathbf{V}_{S2})/2. \end{aligned}$$

In the formulations above, $\mathbf{r}_{pp}$, is the vector of reflectivity coefficients at each layer interface, $\mathbf{p}$, is the ray parameter, θ_1, is the angle of incidence, and θ_2, is the transmission angle; $\mathbf{V}_{P1}$, and $\mathbf{V}_{P2}$, are the P-wave velocities above and below a given interface, respectively. Similarly, $\mathbf{V}_{S1}$, and $\mathbf{V}_{S2}$,

are the corresponding S-wave velocities, while $\boldsymbol{\rho}_1$, and $\boldsymbol{\rho}_2$, are densities above and below this interface (see, **Fig.** 4.3 which is taken from Avseth et al. [15]). Also $\bar{\mathbf{V}}_P$, $\bar{\mathbf{V}}_S$, and $\bar{\boldsymbol{\rho}}$, are the average P-wave velocity, S-wave velocity and density over the reflecting interface; $\Delta\,\mathbf{V}_P$, $\Delta\,\mathbf{V}_S$, and $\Delta\,\boldsymbol{\rho}$, are the corresponding contrasts.

The approximation given by Aki and Richards can be further approximated, according to Shuey [130], as

$$\mathbf{r}_{pp}(\theta) \;\approx\; \mathbf{r}_{pp}(0) \;+\; \mathbf{G}\sin^2\theta \;+\; \mathbf{F}\,(\tan^2\theta - \sin^2\theta), \tag{4.19}$$

where,

$\mathbf{r}_{pp}(0) \;=\; \frac{1}{2}\left(\frac{\Delta\mathbf{V}_P}{\bar{\mathbf{V}}_P} + \frac{\Delta\,\boldsymbol{\rho}}{\bar{\boldsymbol{\rho}}}\right)$,
$\mathbf{G} \;=\; \frac{1}{2}\frac{\Delta\mathbf{V}_P}{\bar{\mathbf{V}}_P} - 2\frac{\bar{\mathbf{V}}_S^2}{\bar{\mathbf{V}}_P^2}\left(\frac{2\Delta\mathbf{V}_S}{\bar{\mathbf{V}}_S} + \frac{\Delta\,\boldsymbol{\rho}}{\bar{\boldsymbol{\rho}}}\right)$ and
$\mathbf{F} \;=\; \frac{1}{2}\frac{\Delta\mathbf{V}_P}{\bar{\mathbf{V}}_P}$.

This form can be interpreted in terms of different angular ranges, where $\mathbf{r}_{pp}(0)$ is the normal-incidence reflection coefficient, $\mathbf{G}$ describes the variation at intermediate offsets and is often termed as AVO gradient , whereas $\mathbf{F}$ dominates the far offsets, near critical angle. Normally, the range of angles available for AVO analysis is less than 30^o-40^o. Therefore, we only need to consider the first two terms, valid for angles less than 30^o, and the expression becomes as

$$\mathbf{r}_{pp}(\theta) \;\approx\; \mathbf{r}_{pp}(0) \;+\; \mathbf{G}\sin^2\theta. \tag{4.20}$$

The zero-offset reflectivity $\mathbf{r}_{pp}(0)$, is controlled by the contrast in acoustic impedance across an interface. The AVO gradient, $\mathbf{G}$, is more complex and depends on the contrasts of $\mathbf{V}_P$, and $\mathbf{V}_S$ [28]. Hence, the importance of $\boldsymbol{V_P/V_S}$ ratio (or equivalently the Poisson's ratio) change can be spotted by AVO gradient. Ostrander [111] showed that a gas-filled formation would have a very low Poisson's ratio compared to Poisson's ratios in the surrounding non-gaseous formations. This would cause a significant increase in positive amplitude versus angle at the bottom of the gas layer, and a significant increase in negative amplitude versus angle at the top of the gas layer.

4.3.2 Convolution Model with Weak Contrast Reflectivity

Seismic reflection coefficients depend on material properties of the subsurface. An isotropic, elastic medium is completely described, according to Sheriff and Geldart [129], by three elastic parameters, such as P-wave

velocity, S-wave velocity and density, $\{V_p(t), V_s(t), \rho(t)\}$ where t is the two-way vertical seismic travel time. Several other parametrization can also be used, e.g., based on acoustic impedance. The forward modeling of a seismic time-angle gather used here is based on the matrix-vector formulation in Buland and Omre [24]. The forward model is linear that uses weak contrast expressions for reflection coefficients by Aki and Richards (Eq. (4.18)). The single-interface reflection coefficient in the Eq. (4.18) can be extended to a time-continuous reflectivity function $c_{pp}(t, \theta)$, as follows [140]

$$\begin{aligned} c_{pp}(t,\theta) = \mathbf{a}_{V_p}(t,\theta)\frac{1}{\partial t}\ln V_p(t) \; + \; \mathbf{a}_{V_s}(t,\theta)\frac{1}{\partial t}\ln V_s(t) \\ + \mathbf{a}_{\rho}(t,\theta)\frac{1}{\partial t}\ln \rho(t), \end{aligned} \tag{4.21}$$

where $\mathbf{a}_{V_p}$, $\mathbf{a}_{V_s}$ and $\mathbf{a}_{\rho}$ are the generalizations of the reflection coefficients in expressions (4.18) with time dependent velocities $\bar{V}_P(t)$ and $\bar{V}_s(t)$. We assume that $\bar{V}_p(t)$ and $\bar{V}_s(t)$ can be represented by constant or slowly varying known background model, such that $\bar{V}_p(t)$ and $\bar{V}_s(t)$ are the average or moving average of V_p and V_S in a relatively small time window. In this formulation, reflection angle θ is considered independent variable. However, the seismic data is recorded as a function of source-receiver distance, h (offset). The transformation of the data from (t, h)-domain to (t, θ)-domain depends on the velocity function [23].

The material parameters $V_p(t)$, $V_s(t)$ and $\rho(t)$ are a priori assumed to be log-Gaussian, which implies that the parameters are restricted to take positive values. This assumption is required for later analytical treatment due to expression in Eq. (4.21) [23]. The logarithm of these material parameters defines a continuous Gaussian vector field $m(t)$, and can be written as

$$\mathbf{m}(t) \; = \; \Big[\ln V_p(t) \ln V_s(t) \ln \rho(t) \Big]^T. \tag{4.22}$$

The continuous form of the Gaussian field $\mathbf{m}(t)$, makes it possible to give a proper definition of the time differentiated Gaussian field, $\mathbf{m}'(t)$. Buland [23] also incorporated spatial coupling between the elastic parameters via a covariance model. Lateral and temporal (or vertical) dependencies between model parameters are also described by a factorized scalar function, see Doyen [49]. Because of this assumption of linearity in the forward model with log-Gaussian prior model, Buland and Omre [24] were able to obtain analytical expressions for mean and covariance of the log-Gaussian posterior distribution.

Discretized Form of Reflectivity Function

In a computer program, continuous fields are represented on a grid. The grid density should be determined by the temporal variability of the elastic parameters, and not by the sampling density of seismic data [24]. A discrete representation of $\mathbf{m}(t)$, is Gaussian in a given time interval and expressed as

$$\mathbf{m} \;=\; \left[\, \ln \mathbf{V}_p^T \ln \mathbf{V}_s^T \ln \boldsymbol{\rho}^T \,\right]. \tag{4.23}$$

Similarly, a discrete representation of the differentiated field $\mathbf{m}^{'}(t)$ is also Gaussian and noted as $\mathbf{m}^{'}$. A discrete version of the reflectivity function $\mathbf{c}_{PP}\ (t, \theta)$, for a given time interval and a set of reflection angles, now can be written in the matrix-vector form as

$$\mathbf{r} \;=\; \mathbf{A}\,\mathbf{m}^{'} = \mathbf{ADm}, \tag{4.24}$$

with the matrices defined as

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{V_p}(\theta_1) & \mathbf{A}_{V_s}(\theta_1) & \mathbf{A}_{\rho}(\theta_1) \\ \vdots & \vdots & \vdots \\ \mathbf{A}_{V_P}(\theta_{n_\theta}) & \mathbf{A}_{V_s}(\theta_{n_\theta}) & \mathbf{A}_{\rho}(\theta_{n_\theta}) \end{bmatrix}; \qquad \mathbf{D} = \begin{bmatrix} -1 & 1 & \cdots & 0 \\ & \ddots & & \\ 0 & \cdots & -1 & 1 \end{bmatrix}.$$

In the above formulation $\mathbf{r}$, is the vector of reflectivity coefficients at each layer interface and $\mathbf{A}$ is a sparse matrix. $\mathbf{A}_{V_p}(\theta_i)$, $\mathbf{A}_{V_s}(\theta_i)$, and $\mathbf{A}_{\rho}(\theta_i)$, are $(n_m/3)\times(n_m/3)$ diagonal sub-matrices of $\mathbf{A}$ containing discrete time samples of $\mathbf{a}_{V_p}(t, \theta_i)$, $\mathbf{a}_{V_p}(t, \theta_i)$, and $\mathbf{a}_{\rho}(t, \theta_i)$, respectively. Also n_θ is the number of reflection angles and n_m is the dimension of $\mathbf{m}$ and $\mathbf{m}^{'}$. An example of the discrete version of the differentiated field of the layer properties $\mathbf{D}$, is shown above.

Wavelet Discretization and Convolution

The seismic observations d, are connected to the reflection coefficients c_{pp}, through the convolution model described by Dobrin and Savit [42] as

$$d(t, \theta) \;=\; \int \omega_\theta(t-u)\; c_{pp}(\theta, u)\; \mathrm{d}u + \epsilon(t, \theta), \tag{4.25}$$

where $\omega_\theta(t-u)$, is a seismic wavelet dependent on the angle of incidence θ, and u, is an auxiliary variable and ϵ is the error term. Ideally, the Eq. (4.25) simulates the response to a delta function or a spike-like source because only such a wavelet would enable us to identify individual interfaces [127].

In practice, however, such an ideal source-time function is impossible to achieve. As the Earth acts as a filter in which high frequencies are attenuated as the energy propagates through the earth, the shape of the wavelet changes with time. For all practical purposes, a wavelet is assumed to be stationary and band-limited; often an average wavelet is estimated from seismic data. In seismic data processing, the wavelet is usually removed, and a simpler, so-called zero-phase wavelet is convolved. The most common zero-phase wavelet is a Ricker wavelet [117] described as

$$f(t) \;=\; (1 - 2\pi^2\omega_{max}^2 t^2)\exp(-\pi^2\omega_{max}^2 t^2), \tag{4.26}$$

where $f(t)$, is the amplitude of the wavelet at time t, and ω_{max}, is the peak frequency of the wavelet. In the discretized framework for convolution with reflection coefficient, one needs to perform discretization of the evaluated Ricker wavelet along a trace. The ability of the Ricker wavelet to capture seismic information depends on the choice of the wave number and the associated wave sampling frequencies [127]. Improper choice of wavelet sampling frequencies may lead to numerical dispersions, or even may violate the Shanon information theory relevant to the relationship between wavelet sampling interval and Nyquist frequency [128]. For seismic exploration, a band-limited Ricker wavelet with 25-40 Hz peak frequency and a proper choice of sampling frequency, generally satisfy these conditions.

In the frequency domain, convolution is equivalent to multiplication and hence, according to Buland and Omre [24], the integral expression for convolution can be further discretized such that

$$\mathbf{d}_{obs} = \mathbf{W}\,\mathbf{r} + \mathbf{e}, \tag{4.27}$$

and in this expression $\mathbf{e}$, is an n_d-dimensional error vector, and $\mathbf{W}$, is a block-diagonal matrix containing one wavelet for each reflection angle. In this setting, a seismic time-angle gather is represented by the vector $\mathbf{d}_{obs} \in \mathbf{R}^{n_d}$. In an expanded form, this relationship can be further written as

$$\begin{bmatrix} \mathbf{d}_{obs}(\theta_1) \\ \vdots \\ \mathbf{d}_{obs}(\theta_{n_\theta}) \end{bmatrix} = \begin{bmatrix} \mathbf{W}(\theta_1) & & \\ & \ddots & \\ & & \mathbf{W}(\theta_{n_\theta}) \end{bmatrix} \times \begin{bmatrix} \mathbf{r}(\theta_1) \\ \vdots \\ \mathbf{r}(\theta_{n_\theta}) \end{bmatrix} + \begin{bmatrix} \mathbf{e}(\theta_1) \\ \vdots \\ \mathbf{e}(\theta_{n_\theta}) \end{bmatrix}. \tag{4.28}$$

The wavelets are allowed to be angle dependent, but assumed to be independent of location. The block matrix $\mathbf{W}(\theta_i)$, represents the wavelet for

angle θ_i, and can be expressed in a block-matrix format as

$$\mathbf{W}(\theta_i) = \begin{bmatrix} w_1(\theta_i) & & & & & & \\ w_2(\theta_i) & w_1(\theta_i) & & & & & \\ \vdots & & \ddots & & & & \\ w_{n_s}(\theta_i) & \cdots & \cdots & w_1(\theta_i) & & & \\ & \ddots & & & \ddots & & \\ & & w_{n_s}(\theta_i) & \cdots & \cdots & w_1(\theta_i) & \\ & & & \ddots & & \vdots & \\ & & & & \ddots & \vdots & \\ & & & & & w_{n_s}(\theta_i) & \end{bmatrix}, \tag{4.29}$$

where $w_1(\theta_i), \cdots, w_{n_s}(\theta_i)$, are the samples of the wavelet for angle θ_i. In this case, the sampling of the wavelet is equal to the sampling of the seismic data [121]. If the sampling of $\mathbf{r}$ and $\mathbf{d}_{obs}$ are different, the rows of $\mathbf{W}$, in Eq. (4.29) contains wavelet corresponding to the sampling of $\mathbf{r}$, and the rows are shifted relatively according to the sampling of $\mathbf{d}_{obs}$ [24].

The expressions for the seismic time-angle gather $\mathbf{d}_{obs}$ can also be written, in more compacted form, as

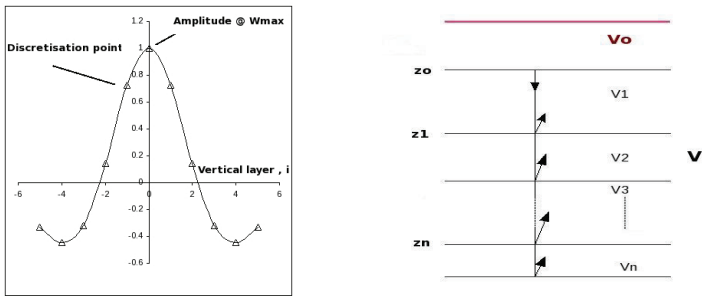
$$\mathbf{d}_{obs} = \mathbf{Gm} + \mathbf{e} = \mathbf{WADm} + \mathbf{e}, \tag{4.30}$$

where $\mathbf{G} = \mathbf{WAD}$, is a linear modeling operator for the seismic time-angle gather. In order to implement this formulation for a practical case, one can use Ricker wavelet with frequency of 20 - 40 Hz and small incident angle of the order of 5-20 degrees. By using seismic properties calculated from rock physics model, we can use this linearized convolution model Eq. (4.30) to simulate synthetic seismic amplitude data at different times of the reservoir production period.

The method of Buland and Omre [24] presented here is extremely fast because all computations are performed in the Fourier domain. However, some limitations result from this simplified approach based on linearity and log-Gaussian approximations [49]. First, the seismic data and the model must both be discretized on the same 3-D grid, regularly sampled in x, y, and time. Furthermore, covariance model for the prior distribution and the noise covariance model must be stationary, i.e., they are not allowed to vary in space. Escobar et al. [57] have recently expanded Buland's approach to work in stratigraphic grid and allow non-stationary covariance structures.

4.4 Vertical Resolution of Seismic Modeling

Normally, the velocity or reflection coefficients calculated from the reservoir simulation model are depth converted and sampled into a regular time grid before performing the convolution. Since the velocities and layer thicknesses normally are different for each layer in the reservoir model, this will always introduce a sampling, or scaling, error. Here we have applied a different procedure, where all the calculations, including the convolution, are performed in depth. We also neglect over- and under-burden effects. This will of course not be possible in a real case. However, in this way we



(a) Ricker wavelet (b) Vertical reflected ray paths

Figure 4.3: Discretized Ricker wavelet and time-to-depth relationship.

are able to generate a synthetic problem where these errors are minimized, allowing us to focus on the differences using inverted and not-inverted data in the history-matching procedure. In addition, it simplifies calculations. The discretized representation of the weak contrast reflectivity signal corresponding to interface j, is then

$$\mathbf{d}_j = \sum_{i=-N_\omega}^{N_\omega} \omega_i \mathbf{r}_{j-1} + e_j, \tag{4.31}$$

where e_j, is an error term and $\mathbf{r}_j$, is the reflectivity coefficient at interface j, and ω_i, a modified Ricker wavelet given by

$$\omega_i = \left(1 - 2\alpha^2(4(N_\omega + 1 - i)^2)\right) \exp\left(-\alpha^2(4(N_\omega + 1 - i)^2)\right), \tag{4.32}$$

with $\alpha = \pi\omega_{max}$, and N_ω, is the number of discretization points on each side of the wavelet center shown by the diamond label in **Fig.** 4.3a. This effectively implies a deformation of the wavelet, which also varies with depth.

However, since the same model is used both for the inverse modeling as well as to generate the synthetic "true" data, we believe that this is an efficient way of removing the sampling error, and at the same time retaining some of the main features related to the seismic modeling. As a next step, the effect of the sampling error can be easily included by comparing this approach with a traditional approach, where the convolution is performed in the time domain. Note that if the layer thicknesses are constant, and applying the common assumption that the velocity variations within the reservoir section are small ($\Delta V_i \ll V_i \approx V$), the time-depth relation for a normal incident wave simplifies to

$$\begin{aligned} t_i &= 2\left[(\frac{1}{V_o} - \frac{1}{V_1})z_0 \; + \; (\frac{1}{V_1} - \frac{1}{V_2})z_1 + \cdots + \; \frac{z_i}{V_i}\right] \\ &\approx \; 2\left[\frac{z_0}{V_o} + \frac{z_i - z_0}{V}\right], \;\; i = 1, \cdots, n. \end{aligned} \tag{4.33}$$

This model assumes each layer to be characterized by an interval velocity V_i, and with z_i, as the thickness of such interval and V_o is the over-burden velocity (see, Fig. 4.3b). Here the horizontal layer interval velocity may be averaged over several depth intervals to yield a time-average velocity V. That is, our approach is exact within this approximation provided the sampling interval $\Delta t = 2\overline{\Delta z}/V$, where $\overline{\Delta z}$ is the average layer thickness. The breadth (distance between each of the two side lobes) of a normal Ricker wavelet in time is given by $B_t = \frac{\sqrt{6}}{\pi\omega_{max}}$. From Eq. (4.33) it follows that the average breadth of the wavelet (in depth) in our approach is equal to $B_d = \frac{10\sqrt{6}}{\pi}\overline{\Delta z}$. Thus, the corresponding average time-wavelet frequency in our approach, is given (in Hz) by

$$f = \frac{\sqrt{6}}{\pi B_t} \; = \; \frac{\sqrt{6}}{\pi}\frac{V}{2B_d} = \frac{V}{20\overline{\Delta z}}. \tag{4.34}$$

If we consider an average velocity V= 2500m/s, and an average cell thickness $\overline{\Delta z}$= 3.9 m, the "effective" wavelet frequency corresponding to a time frequency is close to 30Hz. The corresponding sampling interval in time becomes approximately 3 milliseconds. Thus, sampling of wavelet in depth in our case is a reasonable approximation for a synthetic case. In paper A, we have used this approximated method to check the sensitivity to the vertical resolution applied in the EnKF experiments. The choice of N_ω had a significant effect on the results of history matching as shown in the examples section of paper A.

Chapter 5

Ensemble Based Data Assimilation

In data assimilation, one aims at merging the information from observations into a numerical model, typically of a geophysical system. The analysis and forecasts in data assimilation techniques are best thought of as probability distributions. The analysis step is an application of the Baye's theorem. Advancing the probability distribution in time, would be done exactly in the general case by the Fokker-Planck equation, but that is unrealistically expensive, so various approximations operating on simplified representations of the probability distributions are used instead [59]. If the probability distributions are normal, they can be represented by their mean and covariance, which gives rise to the Kalman filter [84]. However it is not feasible to maintain the covariance because of the large number of degrees of freedom in the state, so various approximations are used instead. One variant of this type of approximation is the use of ensemble Kalman filter (EnKF). The probability distribution in EnKF is represented by an ensemble of simulations and the covariance is approximated by sample covariance [54]. In reservoir history matching problem, the assumption is often made that the initial state of the reservoir is known and the joint probability of the reservoir parameters before assimilation of data can be characterized [109]. It is often valid to assume that the conditional probability distribution of future states of the reservoir, given the present state and all past states, depends only upon the current state and not on any past states. The data (or observations) depends only upon the current state of the reservoir, and not on previous states. If these assumptions are valid, the probability density for the parameters and the state variables can be defined recursively. Once we sample the posterior distribution, using the prior and likelihood term by Baye's theorem, we then have conditioned our

reservoir model simultaneously with the most recent data available and the prior information. Hence, EnKF can be used as a recursive sequential data assimilation method to update the reservoir models. This chapter will briefly outline the Kalman filter algorithm, and then we will describe the mathematical background of EnKF. Finally, we will provide a work flow for EnKF-based history matching for reservoir simulation models.

5.1 Kalman Filter Algorithm

The Kalman filter [84] can be viewed as an optimal recursive data processing algorithm, and it is a technique for assessing the posterior distribution. 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 and the posterior covariance [95]. 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 mathematical derivation of Kalman filter and EnKF algorithm presented here closely follows the work of Skjervheim [131].

Let $\mathbf{\Psi}(x,t)$, denote the unknown *model variable* which constituted by the model parameter, $\mathbf{m}(x)$ and model state, $u(x,t)$. In the standard state-space formulation, observations $\mathbf{d}_k$, are related to model variables $\mathbf{\Psi}_k$, by a linear observation-to-state equation

$$\mathbf{d}_k = \mathbf{H}_k \mathbf{\Psi}_k + \epsilon_k^o \qquad \epsilon_k^o \backsim N_{m_k}(0, \mathbf{C}_{\epsilon_k^o}), \tag{5.1}$$

where $\mathbf{H}_k$, is the observation operator, and ϵ_k^o, is a Gaussian model noise sequence with error covariance matrix $\mathbf{C}_{\epsilon_k^m}$. The model state evolution equation is defined by

$$\mathbf{\Psi}_k \;=\; \mathbf{F}_k \mathbf{\Psi}_{k-1} + \epsilon_k^m \qquad \epsilon_k^m \backsim N_q(0, \mathbf{C}_{\epsilon_k^m}), \tag{5.2}$$

where $\mathbf{F}_k$, is a linear model operator and ϵ_k^m, is a Gaussian model noise sequence with error covariance matrix $\mathbf{C}_{\epsilon_k^m}$. The initial prior distribution has expectation $E(\mathbf{\Psi}_0 = \mathbf{\Psi}_0^{init})$, and covariance $\mathrm{Cov}(\mathbf{\Psi}_0) = \mathbf{C}_{\epsilon_0}^{init}$, and is assumed to be Gaussian, $\mathbf{\Psi}_0 \backsim N_q(\mathbf{\Psi}_0^{init}, \mathbf{C}_{\epsilon_0}^{init})$. Let us assume that the posterior distribution at time t_{k-1} defined by $g(\mathbf{\Psi_{k-1}}|\mathbf{d}_{k-1:1})$, is known,

$$\Psi_{k-1}|\mathbf{d}_{k-1:1} \;\backsim N_q(\mathbf{\Psi}_{k-1}^a, \mathbf{C}_{\Psi_{k-1}}^a). \tag{5.3}$$

Combining Eq. (5.3) and the system Eq. (5.2), we can write the prior distribution $g(\mathbf{\Psi_k}|\mathbf{d}_{k-1:1})$, at time t_k as

$$\Psi_k|\mathbf{d}_{k-1:1} \backsim N_q(\mathbf{\Psi}_k^f, \mathbf{C}_{\Psi_k}^f), \tag{5.4}$$

where $\mathbf{\Psi}_k^f$, is the prior mean and $\mathbf{C}_{\Psi_k}^f$, is the prior covariance. Both of these can be defined as

$$\mathbf{\Psi}_k^f = E(\mathbf{\Psi}_k|\mathbf{d}_{k-1:1}), \tag{5.5}$$

$$\mathbf{C}_{\Psi_k}^f = E\left((\mathbf{\Psi}_k - \mathbf{\Psi}_k^f)(\mathbf{\Psi}_k - \mathbf{\Psi}_k^f)^T|\mathbf{d}_{k-1:1}\right). \tag{5.6}$$

For the linear state space model in Eq. (5.2), the prior moments can be further written as

$$\mathbf{\Psi}_k^f = \mathbf{F}_k\mathbf{\Psi}_{k-1}^a, \tag{5.7}$$

$$\mathbf{C}_{\Psi_k}^f = \mathbf{F}_k\mathbf{C}_{\Psi k-1}^a\mathbf{F}_k^T + \mathbf{C}_{\epsilon_k^m}. \tag{5.8}$$

The derivation of the posterior distribution $g(\mathbf{\Psi_k}|\mathbf{d}_{k:1})$, at time, t_k is found from the joint distribution of $g(\mathbf{\Psi_k}, \mathbf{d}_k|\mathbf{d}_{k-1:1})$. The joint distribution is established by observing from Eq. (5.1) that

$$\text{Cov}(\mathbf{\Psi_k}, \mathbf{d}_k|\mathbf{d}_{k-1:1}) = \text{Cov}(\mathbf{\Psi_k}, \mathbf{H}_k\mathbf{\Psi_k} + \epsilon_k^o|\mathbf{d}_{k-1:1}) = \mathbf{C}_{\Psi_k}^f\mathbf{H}_k^T, \tag{5.9}$$

and hence the joint distribution can be expressed as

$$\begin{pmatrix} \mathbf{\Psi_k} \\ \mathbf{d}_k \end{pmatrix}_{|\mathbf{d}_{k-1:1}} \backsim N_{q+m_k}\left(\begin{pmatrix} \mathbf{\Psi}_k^f \\ \mathbf{d}_k^f \end{pmatrix}, \begin{bmatrix} \mathbf{C}_{\Psi_k}^f & \mathbf{C}_{\Psi_k}^f\mathbf{H}_k^T \\ \mathbf{H}_k\mathbf{C}_{\Psi_k}^f & \mathbf{C}_{\Psi_k}^f \end{bmatrix}\right). \tag{5.10}$$

By applying the formulas to the joint distribution in the above equation as mentioned by Skjervheim [131], the posterior distribution $g(\mathbf{\Psi_k}|\mathbf{d}_{k:1})$, can be expressed as

$$\Psi_k|\mathbf{d}_{k:1} \backsim N_q(\mathbf{\Psi}_k^a, \mathbf{C}_{\Psi_k}^a). \tag{5.11}$$

Here the analysis mean and covariance are given by

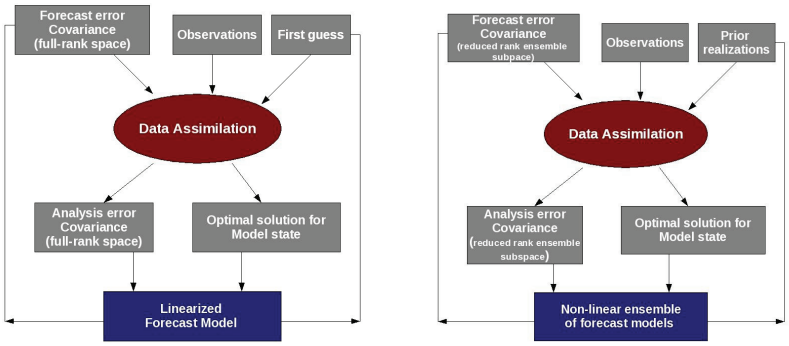
$$\mathbf{\Psi}_k^a = \mathbf{\Psi}_k^f + \mathbf{K}_k(\mathbf{d}_k - \mathbf{H}_k\mathbf{\Psi}_k^f), \tag{5.12}$$

$$\mathbf{C}_{\Psi_k}^a = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{\Psi}_k^f, \tag{5.13}$$

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

$$\mathbf{K}_k = \mathbf{C}_{\Psi_k}^f\boldsymbol{H}_k^T(\boldsymbol{H}_k\mathbf{C}_{\Psi_k}^f\boldsymbol{H}_k^T + \mathbf{C}_{\epsilon_k^o})^{-1}. \tag{5.14}$$

From the Kalman filter theory the state space models are assumed linear with Gaussian statistics and the posterior distribution $g(\mathbf{\Psi_k}|\mathbf{d}_{k:1})$, is then Gaussian and subject to analytical evaluation. The objective is to sample from the posterior pdf $g(\mathbf{\Psi_k}|\mathbf{d}_{k:1})$, and to generate a sample from this distribution, different sampling strategies, e.g., Rejection sampling and Markov chain Monte Carlo (MCMC), can be applied [131].



(a) Kalman filter (b) Ensemble Kalman filter

Figure 5.1: Comparative analysis of different filtering techniques.

5.2 The Ensemble Kalman Filter

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 [58]. 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 [61]. The result is an ensemble of solutions approximating the posterior pdf. The essential difference between the Kalman filter and EnKF is that the covariance matrix in EnKF is replaced by a sample covariance computed from the ensemble, see **Fig.** 5.1. The ensemble is operated with as if it were a random sample, but the ensemble members are really not independent - the EnKF ties them together. One advantage of EnKF is that advancing the PDF in time is achieved by simply advancing each member of the ensemble.

Let us Consider the non-linear state space model:

$$\mathbf{\Psi}_k = \mathbf{f}_k(\mathbf{\Psi}_{k-1}) + \epsilon_k^m \qquad \epsilon_k^m \backsim N_q(0, \mathbf{C}_{\epsilon_k^m}), \tag{5.15}$$

$$\mathbf{d}_k = \mathbf{h}_k(\mathbf{\Psi}_k) + \epsilon_k^o \qquad \epsilon_k^o \backsim N_{m_k}(0, \mathbf{C}_{\epsilon_k^o}), \tag{5.16}$$

where $\mathbf{f}_k : R^q \rightarrow R^q$ is a non-linear model operator with model errors ϵ_k^m, and $\mathbf{h}_k : R^q \rightarrow R^{m_k}$ is the non-linear measurement operator, relating the model variable $\mathbf{\Psi}_k$, to the observations $\mathbf{d}_k$, allowing for measurement

errors ϵ_k^o. One option to solve for the non-linearity in the context of state-space formulation is to use the so called Extended Kalman filter (EKF) [99]. But EKF is not an optimal filter, but rather it is constructed on the basis of a set of approximations. The EnKF was introduced as a Monte Carlo alternative to the traditional EKF. Evensen [58] 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. He also showed that EnKF is a stochastic alternative to using the approximate error covariance equation in the EKF.

In the EnKF setting the nonlinear state space model is reformulated, so that the observation equation can be treated as linear [59]. In order 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

$$\mathbf{d}_k = \mathbf{H}_k \widetilde{\mathbf{\Psi}}_k + \epsilon_k^o, \tag{5.17}$$

where $\widetilde{\mathbf{\Psi}}_k = (\mathbf{\Psi}_k, \mathbf{h}_k(\mathbf{\Psi}_k))$, and $\mathbf{H}_k$, being the linear measurement operator that picks data from $\widetilde{\mathbf{\Psi}}_k$. For simplicity we omit the augmented notation and let $\mathbf{\Psi}_k$, from now, include the parameters, the states and the model prediction of the measurements.

Ensemble Representation of $\mathbf{C}_\Psi$

Define the matrix holding the ensemble members $\Psi_k \in R^q$, at time t_k, as

$$\mathbf{\Psi}_k = \mathbf{\Psi}(x, t_k) = (\Psi_k^1, \Psi_k^2, \cdots, \Psi_k^N) \in R^{q \times N_e}, \tag{5.18}$$

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

$$\overline{\mathbf{\Psi}}_k = \mathbf{\Psi}_k \mathbf{1}_{N_e}, \tag{5.19}$$

where $\mathbf{1}_{N_e} \in R^{N_e \times N_e}$, is the matrix where each element is equal to $1/N_e$. Based on the above equations, the ensemble perturbation matrix $\mathbf{\Psi}_k^{'}$, can be expressed as

$$\mathbf{\Psi}_k^{'} = \mathbf{\Psi}_k - \overline{\mathbf{\Psi}}_k = \mathbf{\Psi}_k(\mathbf{I} - \mathbf{1}_{N_e}). \tag{5.20}$$

The ensemble covariance matrix $\mathbf{C}_{\Psi_k}^e \in R^{q \times q}$, can then be defined as

$$\mathbf{C}_{\Psi_k}^e = \frac{\mathbf{\Psi}_k^{'}(\mathbf{\Psi}_k^{'})^T}{N_e - 1}. \tag{5.21}$$

Measurement Perturbations

Given a vector of observations $\mathbf{d}_k \in R^{m_k}$, the N_e vectors of perturbed observations can be expressed as

$$\mathbf{d}_k^i = \mathbf{d}_k + \epsilon_k^{o,i}, \quad i = 1, 2, \cdots, N_e \tag{5.22}$$

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

$$\mathbf{D}_k = (\mathbf{d}_k^1, \mathbf{d}_k^2, \cdots, \mathbf{d}_k^{N_e}) \in R^{m_k \times N_e}. \tag{5.23}$$

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

$$\mathbf{E}_k = (\epsilon_k^{o,1}, \epsilon_k^{o,2}, \cdots, \epsilon_k^{o,N_e}), \tag{5.24}$$

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

$$\mathbf{C}_{\epsilon_k^o}^e = \frac{\mathbf{E}_k \mathbf{E}_k^T}{N_e - 1}. \tag{5.25}$$

The EnKF can be viewed as a sequential simulation with Kriging update and is based on a variance minimizing scheme [61]. The ensemble representation of the analysis equation can now be written as

$$\widehat{\mathbf{\Psi}}_k^a = \widehat{\mathbf{\Psi}}_k^f + \mathbf{C}_{\Psi_k}^{e,f} \mathbf{H}_k^T (\mathbf{H}_k \mathbf{C}_{\Psi_k}^{e,f} \mathbf{H}_k^T + \mathbf{C}_{\epsilon_k^o})^{-1} (\mathbf{d}_k^i - \mathbf{H}_k \widehat{\mathbf{\Psi}}_k^f). \tag{5.26}$$

However, the EnKF is strictly a variance minimizing scheme only when the assumption is made that the prior is Gaussian. The complete EnKF update scheme in Eq. (5.26) can also be seen as combination of a nonlinear prediction and and a linear update.

Ensemble Representation

At time t_k, let us define the ensemble of innovation vectors $\mathbf{D}_k^{'} \in R^{m_k \times N_e}$, as

$$\mathbf{D}_k^{'} = \mathbf{D}_k - \mathbf{H}_k \widehat{\mathbf{\Psi}}_k^f, \tag{5.27}$$

the measurement of the ensemble perturbations $\mathbf{S}_k \in R^{m_k \times N_e}$, as

$$\mathbf{S}_k = \mathbf{H}_k \widehat{\mathbf{\Psi}}_k^{f'}, \tag{5.28}$$

and the matrix $\mathbf{G}_k \in R^{m_k \times m_k}$, as

$$\mathbf{G}_k = \mathbf{S}_k \mathbf{S}_k^T + (N_e - 1) \mathbf{C}_{\epsilon_k^o}. \tag{5.29}$$

By using Eqs. (5.27)-(5.29) and the ensemble error covariance matrix definition in Eq. (5.21), the analysis equation already stated in the Eq. (5.26) can further be expressed as the following:

$$\begin{aligned}\widehat{\mathbf{\Psi}}_k^a &= \widehat{\mathbf{\Psi}}_k^f + \widehat{\mathbf{\Psi}}_k^{f'}\widehat{\mathbf{\Psi}}_k^{f'^T}\mathbf{H}_k^T(\mathbf{H}_k\widehat{\mathbf{\Psi}}_k^{f'}\widehat{\mathbf{\Psi}}_k^{f'^T}\mathbf{H}_k^T + (N-1)\mathbf{C}_{\epsilon_k^o})^{-1}\mathbf{D}_k^{'} \\ &= \widehat{\mathbf{\Psi}}_k^f + \widehat{\mathbf{\Psi}}_k^f(\mathbf{I}-\mathbf{I}_N)\mathbf{S}_k^T\mathbf{G}_k^{-1}\mathbf{D}_k^{'} \\ &= \widehat{\mathbf{\Psi}}_k^f(\mathbf{I} + (\mathbf{I}-\mathbf{I}_N)\mathbf{S}_k^T\mathbf{G}_k^{-1}\mathbf{D}_k^{'}) \\ &= \widehat{\mathbf{\Psi}}_k^f(\mathbf{I} + \mathbf{S}_k^T\mathbf{G}_k^{-1}\mathbf{D}_k^{'}) \\ &= \widehat{\mathbf{\Psi}}_k^f\mathbf{X}_k.\end{aligned} \tag{5.30}$$

Here we have used the ensemble perturbation matrix in Eq. (5.20) and $\mathbf{1}_N\mathbf{S}_k^T = 0$. The analyzed ensemble can be considered as a combination of the forecast ensemble. Thus, the updated ensemble can be considered as a weakly nonlinear combination of the forecast ensemble members. In applications, where we are trying to estimate, e.g. the permeability, this implies that we can only expect to find corrections to the permeability estimates which can be represented in the space spanned by the initial permeability ensemble. This is, however, only a practical limitation. Because this effect can be reduced by increasing the ensemble size and by choosing the initial ensemble carefully [64].

5.3 Reservoir History Matching Using EnKF

The EnKF provides an ideal setting for real-time updating and prediction in reservoir simulation models. Every time new observations area available and are assimilated, there is an improvement of the model parameters, and the associated model saturations and pressure. Thus, the analyzed ensemble provides optimal realizations which are conditioned on all previous data, and these can also be used in a prediction of future production [61]. In spite of outstanding theoretical and practical questions on the applicability of the EnKF for highly non-linear and non-Gaussian reservoir problems, there have been promising results of EnKF as an history matching tool. EnKF has provided good history matching results for several real field case studies. For some references related to real field applications of EnKF, we can mention the works of Cominelli et al. [34], Seiler at al. [125], Haugen et al. [76], Bianco et al. [19], Evensen et al. [64] and Skjervheim et al. [135]. An exhaustive analysis of most of these works together with a comparison of the EnKF and related ensemble methods in other industries can be found in the review paper by Aanonsen et al. [3].

The history matching work flow involves three major steps as discussed in Seiler et al. [125]: *parametrization* where the parameters that are uncertain and characterize the major certainty of the model solution are identified, thereafter a *a prior error model* is specified for the selected parameters based on an initial uncertainty analysis, and finally a *solution method* needs to be selected. All three steps are equally important and hence, the quality of the EnKF-based history matching will depend how accurately these steps are performed.

Parametrization

For traditional assisted history matching methods, only a restricted set of reservoir parameters can be included for optimization. These parameters must be selected on the basis of sensitivity and effectiveness. The pilot point method [13] and gard-zoning process [22] are two methods that are in use to select important parameters for updating with history matching. The EnKF is not limited by number of model parameters, because the dimension of the inverse problem is reduced to the number of realizations included in the ensemble. Thus, the solution is searched for in the space spanned by the ensemble members rather than the high dimensional parameter space [64]. In reservoir characterization process, variations in porosity ϕ, have effects on the modeled oil in place. The permeability κ_h determines how well fluids flow through the reservoir and need to be adjusted to match the observed production rate as well as the timing of water breakthrough. With a large number of faults and only few pressure measurements, there is a large uncertainty in the assumed fault transmissibilities and hence, fault transmissibility multiplier $multflt$, is a parameter to be estimated. Another major set of parameters to be estimated is the water-oil contact (WOC) and gas-oil contact (GOC) in different regions of the model. The reason for including these contacts is that this determines the volume of oil in the reservoir as well as the optimal vertical location of horizontal production wells. We can also include vertical transmissibility multipliers *multz*, which modify the effective vertical communication, as parameters to be estimated.

State Vector

For the combined parameter and state estimation problem, the *state vector* updated by EnKF contains both the dynamic variables and static parameters. It is useful to augment the state vector with a vector of predicted

or simulated data. Typical predicted measurements are the oil, water and gas rates for each well, as well as the gas-oil ratio and the water cut. These data may be, in general, nonlinear functions of the state variables. Thus, for a typical ensemble member j, the EnKF state vector takes the following form [61]:

$$\underset{\text{Update}}{\begin{Bmatrix} P \\ S_w \\ \kappa_h \\ \phi \\ \text{multz} \\ \text{WOC} \end{Bmatrix}_j} = \underset{\text{Forecast}}{\begin{Bmatrix} P \\ S_w \\ \kappa_h \\ \phi \\ \text{multz} \\ \text{WOC} \end{Bmatrix}_j} + \sum_i \alpha_{ij} \underset{\text{Covariances}}{\begin{Bmatrix} \mathbf{C}(P, \mathbf{d}_i) \\ \mathbf{C}(S_w, \mathbf{d}_i) \\ \mathbf{C}(\kappa_h, \mathbf{d}_i) \\ \mathbf{C}(\phi, \mathbf{d}_i) \\ \mathbf{C}(\text{multz}, \mathbf{d}_i) \\ \mathbf{C}(\text{WOC}, \mathbf{d}_i) \end{Bmatrix}_j}, \tag{5.31}$$

where i is a counter for predicted measurements, $\mathbf{d}_i$. Thus, the updated state vector consists of the forecast plus a combination of covariances between the predicted measurements and the updated variables. The coefficients, α_{ij} define the impact each measurement has on the update of an ensemble member, j. It is seen that the different dynamic and static variables are updated by adding weighted covariances between the modeled measurements and the variables, one for each measurement. The reason why it is possible to update the parameters given only rate information from the wells, is that the rates are dependent on the properties of the reservoir as given by the parameter set defined above. Thus, there will exist correlations between reservoir properties and the observed production rates [64].

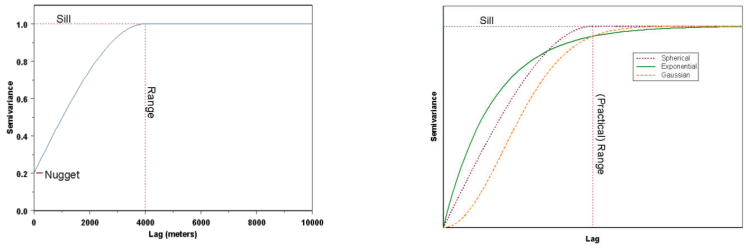
Prior Error Models

An initial uncertainty analysis leads to a quantification of the prior uncertainties of the parameters, which is then presented using probability density functions (PDF). The specified PDFs then represent our prior belief concerning the uncertainty of each particular parameter selected by parametrization. Stochastic simulations are used to produce multiple realizations of the porosity and permeability fields based on the geological models. The uncertain parameters are normally characterized by a Gaussian distribution with mean equal to the best estimate and a standard deviation reflecting the uncertainty spread. There are generally large uncertainties associated with the fault fluid flow properties and fault transmissibility parameters. Treating fault multiplier uncertainty as Gaussian distribution may lead to inaccurate or even insufficient update values for these parameters. One way to circumvent this issue of non-Gaussian parameters is to

us an appropriate transformations, e.g. normal score transformations of the non-Gaussian variables to Gaussian variables [71].

Geological Continuity and Prior Realizations

Statistical information from geological data is used when sampling the initial ensemble such that the sample ensemble is consistent with the parameters and uncertainties given in the geo-model. The geological model should integrate all available prior information from 3D seismic surveys, core and well log data, outcrop studies, and conceptual models. Stochastic simulation is a means for generating multiple equi-probable realizations of the property in question, rather than simply estimating the mean. The two most commonly used forms of simulation for reservoir modeling applications are sequential Gaussian simulation for continuous variables like porosity and sequential indicator simulation for categorical variables like facies. Variograms are used to describe the geological continuity of "homogeneously heterogeneous" properties, and hence, variograms are typically best suited for establishing continuity of porosity and permeability within layers or facies bodies [25]. The experimental variogram is a quantitative measure of spatial correlation and the correlation length or *range* indicates the average extent of continuity/correlation along various directions. In



(a) Range, sill and nugget effect

(b) Different variogram functions

Figure 5.2: Experimental varigrams in use.

most reservoirs, one expects the range in vertical/depth direction to be shorter compared to horizontal direction. The plateau is termed as the *sill* and corresponds to the statistical variance of the property. Often, there is an apparent discontinuity on the y-axis of variogram plot (**Fig.** 5.2a), termed *nugget effect* which reflects geological variability at scales smaller than the smallest experimental distance $|h|$. Since it is required the correla-

tion function used in calculating experimental variogram is positive definite, a number of parametric forms, e.g., exponential, Gaussian or spherical correlation functions are commonly used, see Fig. 5.2b. Generally, the initial ensemble consists of 100 realizations of a Gaussian random field with constant mean parameter value (e.g., porosity) and a chosen variogram function; the popular choice in reservoir engineering is the spherical variogram function. The porosity distribution can also be constrained at the observed porosities in the appraisal wells. The dynamic variables, pressure and saturation grid-cell values, are included in the initial ensemble through an initialization using the flow simulator.

Measurement Error Model

During the ensemble analysis, the prior realizations will become closer to the observations, e.g., production data and thus decrease the ensemble spread. In the present EnKF setup for history matching, the model predicted well variables such as bottom hole pressure (BHP), oil production rate (OPR), well water (WCT) and gas/oil ratio (GOR) are needed to update the model state at analysis step. The production measurement errors are drawn from a Gaussian distribution with a mean zero and standard deviations (for example):

- Bottom hole pressure (BHP): 10%
- Oil production rate (OPR): 15%
- Water cut (WCT): 20%
- Gas/oil ratio: 15%

The standard deviations given above are relative to the actual value of the measurements.

Any data set may contain contaminated data (outliers) that could result in inconsistent updates, possibly leading to model instabilities. Thus it is recommended that appropriate filters are employed to get rid of the possible data outliers. For example, the following criteria proposed by Haugen et al. [76] can be considered. Here, the distance between the model and the data is compared with the sum of the predicted measurement standard deviation:

$$|\mathbf{d} - \mathbf{H}\Psi| \; < \; \zeta(\sigma_{\mathbf{d}} + \sigma_{\Psi}), \tag{5.32}$$

where $\sigma_{\mathbf{d}}$ and σ_{Ψ} are the standard deviations of the observations and the predictions, respectively. Furthermore, ζ is a specified constant (i.e., chosen

to set the strength of the criteria). Normally, ζ will be chosen between 1 and 3, reflecting that the distributions should overlap within 1 to 3 standard deviations.If the equation (5.35) is not satisfied, $\sigma_{\mathbf{d}}$ is updated as follows:

$$\sigma_{\mathbf{d}} = |\mathbf{d} - \mathbf{H}\Psi|/\zeta - \sigma_{\Psi}. \tag{5.33}$$

That is, the error of the observation is increased and is given a smaller weight in the EnKF update.

Work Flow of EnKF-based History Matching

When an initial ensemble of reservoir models is generated, the EnKF is used to update the ensemble sequentially in time to honor the new production observations at the time they arrive. The EnKF consists of a forward integration generate the *forecast* followed by the updating of state variable to generate the *analysis*, see **Fig.** 5.3. An each assimilation step, the model is updated according to the Eq. (5.26) and there is no need to run the simulator from time zero again. The assimilated observations are considered as random variables having a Gaussian distribution with mean equal to

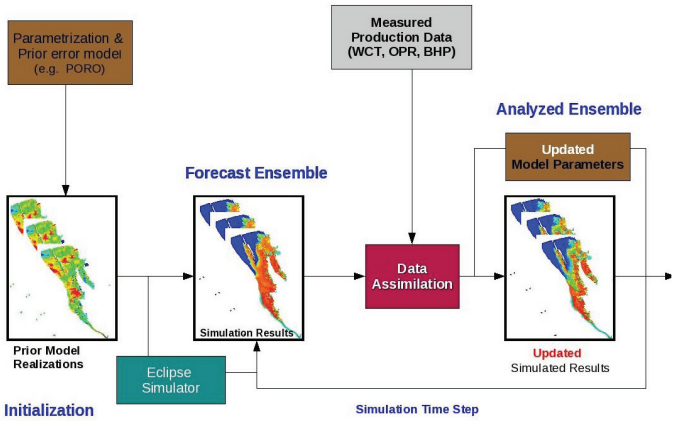


Figure 5.3: EnKF history matching work flow for production data.

observed value and an error covariance reflecting the accuracy of the measurement. The updated ensemble is then integrated until the next update time. The result is an updated ensemble of realizations, conditioned on all previous production data, and thus provides the optimal starting point for predictions of future production [125].

EnKF and Model Predictability

The target of history matching is to prepare the ground for an evaluation of the expected performance of the field. This should be integrated with a technical risk analysis which highlights the expected bounds in the performances according to the remaining uncertainty. This process can be termed as the predictability of the history matched model [34]. At the final update time in the production history, it is possible to use the updated simulation models to compute a prediction of the future production. In principle for a sufficiently large ensemble, the updated model ensemble is adequately sufficient to approximate all the statistical integrals based on posterior distribution. If one has a method to rank the quality of the different realizations, it is also possible to integrate, e.g., the first, second and third quartiles to obtain predictions with approximate uncertainty estimates [64]. The recent work of Cominelli et al. [34] has shown an example of using EnKF to predict the future production scenarios for a deep water reservoir.

5.4 Different Variations of EnKF Algorithm

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 deterministic analysis is an alternative algorithm which avoids the perturbation of observations, and the scheme is based on a square-root formulation [131]. Here, the updates of the ensemble mean and the ensemble perturbations are performed separately [92]. 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 [144]. 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 [159]. 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 randomized likelihood method (RML) and the EnKF and can be referred to as an RML/EnKF algorithm [66]. The ensemble particle filter relies on the assumption that a relatively small ensemble size represents the uncertainty in a high dimensional model system [11]. The particle filter, more generally known as a sequential Monte-Carlo method [47], 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 [131].

5.5 Model Error in EnKF History Matching

Successful reservoir prediction requires an accurate estimation of parameters to be used in the reservoir model. The history matching process may produce poor estimates for parameter space, as simulation errors are introduced from using an approximate model [143]. Thus it is necessary to develop error models for flow simulation error within the petroleum industry, enabling accurate parameter estimation. In the work of O' Sullivan et al. [112], the model error used is based on using simulation results from an approximate model together with statistical error data (e.g., mean error and error covariance).

In most of the EnKF applications, modeling error is commonly neglected, assuming that all error is due to uncertainty in the estimated parameters. While a good history match may be often be obtained by tuning "wrong" parameters when traditional methods are applied, this is not always possible when using the EnKF method. The reason is that EnKF requires a large solution space spanned by prior uncertainty with adequate degrees of freedom. This is apparently a disadvantage of EnKF; however, in such a situation a match obtained by traditional method may hide a potential problem in the model. The work of Aanonsen [1] introduced an explicit model error term in the formulation of the dynamical system. The model error is added to the state vector and estimated as a part in the EnKF update; thus correlations between model error and other variables developed. The applicability of this approach has been tested for a single phase flow model by assimilating well pressure data where an unknown model error due to grid coarsening has introduced.

5.6 Challenging Aspects of EnKF

The EnKF works quite well when the prior probability distributions are Gaussian and when the relationships between the model parameters, state variables, and observation variables ar approximately linear. Generally, there are two major approximations made in the EnKF which may have consequences in practical cases: (1) the update is based on covariances only,

and (2) these covariances are computed from a finite ensemble size. The first approximation implies that third and other higher order moments of the joint PDF of the model variables (including parameters and simulated observations) are neglected, which makes it difficult to maintain a priori imposed non-Gaussian distributions. Thus EnKF update neglects any non-Gaussian contributions in the predicted pdf, when the update increments are computed. On the other hand, the updated ensemble will inherit the non-Gaussian contributions already present in the forecast ensemble. Thus, the analyzed ensemble can not be considered as a re-sampling of a Gaussian posterior pdf. This issue becomes very important when EnKF is applied to complex reservoir models that are characterized by a distribution of facies which cannot be represented by Gaussian distributions. The approach currently taken is to attempt to identify a suitable manner to "transform" the different facies types into intermediate random variables having Gaussian distributions [3]. The use of truncated pluri-Gaussian method [4] and the Gaussian mixture models [48] can be mentioned as possible ways to handle non-Gaussian prior models in EnKF.

Methods for deling with nonlinearity in the state variables for ensemble members can be categorized into two main approaches: parametrization of the state variables and iterative filters. Chen et al. [32] re-parameterize the state vector, replacing a non-Gaussian variable (water saturation) with another related state variable (saturation arrival time) that is approximately Gaussian. The introduction of the use of the discrete cosine transform (DCT) parameterizations has been proposed by Jafarpour and McLaughlin [82]. Most of the iterative forms of the EnKF can be viewed as algorithms for minimizing a stochastic objective function. One such example is the work of Reynolds et al. [116] that can be viewed as ensemble approximations of the randomized maximum likelihood (RML) method [3].

The second approximation, i.e., the use of a limited number of model realizations, while making the problem computationally tractable, introduces errors in the covariance estimates, leading to incorrect updates. Furthermore, it has been shown that such errors tend to lead to systematic under-estimation of model error variances, and ultimately to filter divergence [73], which in ensemble filters is accompanied by ensemble collapse. Practical experiences also shown that similar problems can be encountered when large number of relatively accurate data are assimilated. Using local analysis and covariance localization, one can handle this issue of large number of measurements, see Aanonsen et al. [3].

Chapter 6

4D Seismic History Matching Using EnKF

Conditioning reservoir model to seismic data is a difficult task. Application of EnKF to 4D seismic data in a real case was first proposed by Skjervheim et al. [135]. They used a subspace EnKF inversion scheme to integrate interpreted seismic data into simulation model for both a 2D synthetic model and a real field case. As an example of using seismic waveform data in EnKF, [78] used 4D stacked amplitude data in EnKF to estimate reservoir properties such as gas-oil-contact. In spite of having several challenges, e.g., issues of integration of large amount of 4D data and localization, the initial EnKF applications for 4D seismic data provided promising results [3].

6.1 Integration of 4D Seismic Data in EnKF

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. Hence, traditional EnKF formulations, described in chapter 5, must be modified to incorporate 4D seismic data. When the seismic data are given as a difference between two surveys, a combination of EnKF and the ensemble Kalman smoother (EnKS) has to be applied, see [135]. When we assimilate 4D seismic data, or interpreted seismic data such as acoustic impedance or Poisson's ratio, we effectively have a data measurement corresponding to each grid block. Thus, in assimilating seismic data, the matrix that must be inverted to apply Kalman gain to the data mismatch term is extremely large, see Aanonsen et al. [3]. Moreover, to introduce a sufficient number of degrees of freedom

to match seismic data accurately, one may require a significantly large ensemble size. These practical issues related to seismic data integration are important in the EnKF-based history matching context.

$$\boldsymbol{\Psi}_0 \longrightarrow \underset{\mathbf{d}_1}{\boldsymbol{\Psi}_1} \cdots \rightarrow \underset{\mathbf{d}_j}{\boldsymbol{\Psi}_j} \underset{\mathbf{d}_{j+1}}{\overset{\boldsymbol{\Psi}_j}{\rightarrow \boldsymbol{\Psi}_{j+1}}} \cdots \underset{\mathbf{d}_k}{\overset{\boldsymbol{\Psi}_j}{\rightarrow \boldsymbol{\Psi}_k}} \underset{\mathbf{d}_{k+1}}{\rightarrow \boldsymbol{\Psi}_{k+1}} \rightarrow \cdots \tag{6.1}$$

We have defined a sequence of data assimilation scheme in Eq. (6.1) based on a model state sequence $\boldsymbol{\Psi}_i$, and the corresponding available measurement vectors $\mathbf{d}_i$. Here the counter i, varies from the initial state at 0, to the base survey time j, and then to the monitor survey time k, and next to $k+1$, and so on. Let the non-linear state space model for times, $t_i \in (t_j, t_k)$ be defined as

$$\widetilde{\boldsymbol{\Psi}}_i = \begin{bmatrix} \mathbf{f}_i(\boldsymbol{\Psi}_{i-1}) \\ \boldsymbol{\Psi}_j \end{bmatrix} + \begin{bmatrix} \epsilon_i^m \\ 0 \end{bmatrix} \qquad \epsilon_i^m \backsim N_{\bar{q}}(0, \mathbf{C}_{\epsilon_i^m}), \tag{6.2}$$

$$\mathbf{d}_i = \widetilde{\mathbf{H}}_i \widetilde{\boldsymbol{\Psi}}_i + \epsilon_i^o \qquad \epsilon_i^o \backsim N_{m_i}(0, \mathbf{C}_{\epsilon_i^o}), \tag{6.3}$$

where $\widetilde{\boldsymbol{\Psi}}_i = [\boldsymbol{\Psi}_i, \boldsymbol{\Psi}_j]^T \in R^{\bar{q}}$, $\mathbf{f}_i$, is a non-linear model operator operating only on $\boldsymbol{\Psi}_{i-1}$ with model errors ϵ_i^m. The model errors are added only to the terms where $\mathbf{f}_i$ is operating on $\boldsymbol{\Psi}_{i-1}$, and are zero elsewhere. The linear measurement operator $\widetilde{\mathbf{H}}_i$, is relating the model states $\widetilde{\boldsymbol{\Psi}}_i$, to the observations $\mathbf{d}_i$, in accordance with the data assimilation scheme in Eq. (6.1). Here $\mathbf{d}_k$, includes time difference observations between $\boldsymbol{\Psi}_k$ and $\boldsymbol{\Psi}_j$. Also $\widetilde{\mathbf{H}}_i$, allows for the measurement errors ϵ_i^o.

Let us consider now a joint ensemble matrix including the ensemble matrices of the time instants $\boldsymbol{\Psi}_i = \boldsymbol{\Psi}(x, t_i)$, and $\boldsymbol{\Psi}_j = \boldsymbol{\Psi}(x, t_j)$ as follows

$$\widetilde{\boldsymbol{\Psi}}_i = \begin{bmatrix} \boldsymbol{\Psi}_i \\ \boldsymbol{\Psi}_j \end{bmatrix} \quad \in R^{\bar{q} \times N_e}. \tag{6.4}$$

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

$$\bar{\widetilde{\boldsymbol{\Psi}}}_i = \widetilde{\boldsymbol{\Psi}}_i \mathbf{1}_{N_e}, \tag{6.5}$$

$$\widetilde{\boldsymbol{\Psi}}_i' = \widetilde{\boldsymbol{\Psi}}_i - \bar{\widetilde{\boldsymbol{\Psi}}}_i = \widetilde{\boldsymbol{\Psi}}_i(\mathbf{I} - \mathbf{1}_{N_e}), \tag{6.6}$$

where $\mathbf{1}_N \in R^{N_e \times N_e}$, is the matrix where each element is equal to $1/N_e$. The ensemble covariances, $\mathbf{C}^e_{\widetilde{\Psi}_i} \in R^{\bar{q} \times \bar{q}}$, can then be defined as

$$\mathbf{C}^e_{\widetilde{\Psi}_i} = \frac{\widetilde{\boldsymbol{\Psi}}_i'(\widetilde{\boldsymbol{\Psi}}_i')^T}{N_e - 1}. \tag{6.7}$$

If we define at time t_i, the measurement of the ensemble perturbations $\mathbf{S}_i \in R^{m_i \times N_e}$, as

$$\mathbf{S}_i = \widetilde{\mathbf{H}}_i \widetilde{\mathbf{\Psi}}_i^{f'}, \tag{6.8}$$

and also the ensemble of innovation vectors as

$$\mathbf{D}_i^{'} = \mathbf{D}_i - \widetilde{\mathbf{H}}_i \widetilde{\mathbf{\Psi}}_i^{f}, \tag{6.9}$$

and the matrix $\mathbf{G}_i \in R^{m_i \times m_i}$, written as

$$\mathbf{G}_i = \mathbf{S}_i \mathbf{S}_i^T + (N_e - 1)\mathbf{C}_{\epsilon_i^o}, \tag{6.10}$$

then by following the similar analysis like in Eq. (5.30), we can define the equation of the ensemble representation of the analysis step for both smoother and filter solution as

$$\widetilde{\mathbf{\Psi}}_i^a = \widetilde{\mathbf{\Psi}}_i^f \boldsymbol{X}_i. \tag{6.11}$$

Hence, the EnKF update equation for 4D data, in this case time-difference data, can be expressed as

$$\begin{bmatrix} \mathbf{\Psi}_i^a \\ \mathbf{\Psi}_{j,i}^a \end{bmatrix} = \begin{bmatrix} \mathbf{\Psi}_i^f \boldsymbol{X}_i \\ \mathbf{\Psi}_{j,i-1}^a \boldsymbol{X}_i \end{bmatrix}, \tag{6.12}$$

where $\mathbf{\Psi}_{j,i}^a$ 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

$$\begin{bmatrix} \mathbf{\Psi}_k^f \\ \mathbf{\Psi}_{j,k-1}^a \end{bmatrix}. \tag{6.13}$$

At time t_k, we use the time difference measurement operator $\widetilde{\mathbf{H}}_i$ which relates the measurements to both $\mathbf{\Psi}_k^f$ and $\mathbf{\Psi}_{j,k-1}^a$ and we compute a standard EnKF analysis step like in Eq. (5.26). Note that the filter and the smoother solution coincide at the time of the last measurement, so the predictions from the smoother and filter solutions will be identical [3]. This method is well suited for 4D data when measured at two different time instants. However, special challenges, e.g., inversion of large covariance matrix and potential loss of rank, are involved in the assimilation of large amount of data coming from 4D seismic [131].

6.2 Steps of 4D History Matching with EnKF

As discussed in chapter 5, any assisted history matching work flow involves three major steps such as *parametrization* where sensitive reservoir parameters are identified for model updating, *a prior error model* is specified for the selected parameters based on an initial uncertainty analysis, and finally a *solution method* needs to be selected. In the case of 4D seismic history matching, besides these three major steps, we also need a work flow for integrating seismic data mismatching term. Additionally, a proper uncertainty description of 4D seismic data is necessary.

Parametrization

Like the case with traditional history matching work flow previously discussed, we can choose a range of parameters that can be estimated in EnKF by conditioning the reservoir models with 4D seismic data together with production data. Let us define a state vector which is comprised of dynamic variables, e.g., pressure and gas saturation values in every grid-cells and we have also included porosity and permeability values in all of the active grid-cells. If we want to condition this model with 4D seismic data, a typical stencil for the state vector takes the following form:

$$\begin{matrix} \textbf{Update} \\ \begin{Bmatrix} P \\ S_g \\ \kappa_h \\ \phi \end{Bmatrix}_j \end{matrix} = \begin{matrix} \textbf{Forecast} \\ \begin{Bmatrix} P \\ S_g \\ \kappa_h \\ \phi \end{Bmatrix}_j \end{matrix} + \sum_i \beta_{ij} \begin{matrix} \textbf{Covariances} \\ \begin{Bmatrix} \mathbf{C}(P, \mathbf{d}_i^{4D}) \\ \mathbf{C}(S_g, \mathbf{d}_i^{4D}) \\ \mathbf{C}(\kappa_h, \mathbf{d}_i^{4D}) \\ \mathbf{C}(\phi, \mathbf{d}_i^{4D}) \end{Bmatrix}_j \end{matrix}, \tag{6.14}$$

where i is a counter for predicted 4D measurements $\mathbf{d}_i^{4D}$. Thus, the updated state vector consists of the forecast plus a combination of covariances between the predicted measurements and the updated variables. The coefficients β_{ij}, define the impact each 4D measurement has on the update of an ensemble member j. In addition, we can integrate predicted production data values that can be included in this state vector as well.

Forward Models

Following the approach proposed by Bachrach [17], it is possible to link the elastic parameters, $\mathbf{V}_P$, $\mathbf{V}_S$ and $\boldsymbol{\rho}$ with the reservoir properties, e.g., porosity, ϕ and water saturation, s_w through a non-linear function, $\mathbf{g} : \mathbf{R}^3 \rightarrow \mathbf{R}^3$, by using the Biot-Gassmann theory [68] and the empirical results

of Batzle and Wang [18]. In order to generate synthetic seismograms, one needs to run the flow simulator to predict dynamic behavior of the reservoir. In our study, we have used the Eclipse reservoir simulator extensively [53] to predict the dynamic behavior of the reservoir. Rock

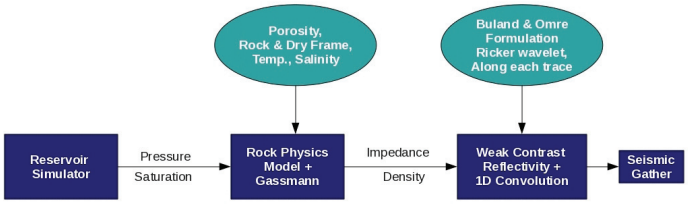


Figure 6.1: Example forward modeling work flow used in EnKF.

physics model use the outputs from the Eclipse simulator as the input for the PEM model. Once we have calculated density and impedance values in each of the active grid cell, we can use, e.g., 1D convolution model with approximated reflection coefficients (discussed in chapter 4) to calculate seismic gather data at each of this cells, see **Fig.** 6.1. The difference of the calculated seismic properties between base and monitor survey times provide the time-difference impedance data or amplitude data. Note that in the context of EnKF-based 4D history matching, the prior realizations used should be consistent with the observed 4D effects of the base case model. Otherwise, it will be very difficult to match the real field seismic data with the simulated posterior seismic data.

Ensemble Smoother Solution and Model Updating

For time difference data, the procedure for updating the state is exactly same as the case with only production data, except that we need also to update the smoother solution at the time of previous survey (base survey). The smoother solution is opposite to the filter solution as it is comprised of the solution at all times conditioned to all data. The filter solution depend only on measurements taken at previous times. The smoother solution can be obtained by sequential procedure such as the ensemble Kalman smoother [63]. The filter and smoother solution coincides at the time of the last measurement, so the predictions from the smoother and filter solutions will be identical [3]. The modified version of the update equation for both

smoother and filter solution becomes:

$$\widetilde{\mathbf{\Psi}}_i^{a,l} = \widetilde{\mathbf{\Psi}}_i^{f,l} + \mathbf{C}_{\widetilde{\Psi}_i}^{e,f}\widetilde{\mathbf{H}}_i^T(\widetilde{\mathbf{H}}_i\mathbf{C}_{\widetilde{\Psi}_i}^{e,f}\widetilde{\mathbf{H}}_i^T + \mathbf{C}_{\epsilon_i^o})^{-1}(\mathbf{d}_i^l - \widetilde{\mathbf{H}}_i\widetilde{\mathbf{\Psi}}_i^{f,l}), \tag{6.15}$$

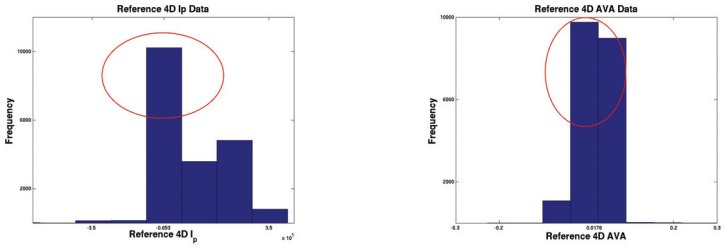
where $\widetilde{\mathbf{\Psi}}_i = [\mathbf{\Psi}_i, \mathbf{\Psi}_k]^T$, and $\widetilde{\mathbf{H}}_i$ picks data from both $\mathbf{\Psi}_i$ and $\mathbf{\Psi}_k$. That is, if the measurements to be assimilated at time t_i is a combination of well data at time, t_i and the change in some seismic property from time t_k to t_i, thus $\widetilde{\mathbf{H}}_i$ will be of the form

$$\widetilde{\mathbf{H}}_i = \begin{pmatrix} \mathbf{H}_i^{prod} & 0 \\ \mathbf{H}_i^{seis} & \mathbf{H}_k^{seis} \end{pmatrix}, \tag{6.16}$$

where $\mathbf{H}_i^{prod}$ picks the simulated well data and $\mathbf{H}_i^{seis}$ picks the simulated seismic data. In practical applications, smoother solution concept is implemented by using same porosity values for both base and monitor surveys to calculate seismic properties using rock physics model, see the attached paper A and paper B.

4D Seismic Measurement Error Model

Results from EnKF studies, where the 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, a positive impact should be found regardless of the high noise level in the real field seismic data, see [132]. There is not any concrete established criteria to



(a) ΔI_p data (b) $\Delta\mathbf{d}$ data

Figure 6.2: 4D seismic measurement error level used in EnKF.

determine the exact level of seismic data uncertainties. The challenge of specifying uncertainties in 4D seismic data is resolved by using results from Bayesian inversion techniques as reported by El Ouair et al. [56]. For

the realistic synthetic case studies in the paper A, the level of noise in the seismic data is taken as 30% of the average reference case time-difference seismic data value over all active cells. In **Fig.** 6.2, the highlighted areas show the average value of the corresponding reference case 4D seismic data values. We have taken 30% of this reference case value as the standard deviation for 4D seismic data noise level in EnKF. For the synthetic case study, one may use a diagonal seismic error covariance matrix. But diagonal covariance matrix is not necessarily valid for synthetic data, since the pressures and saturations in neighboring cells may be correlated. Thus, it is more appropriate to use a non-diagonal seismic data covariance matrix [43].

6.3 EnKF Work Flow for 4D History Matching

The EnKF work flow for 4D seismic history matching is shown in **Fig.** 6.3. The history matching problem is addressed by first defining the model parameters and their uncertainties. Then, an ensemble of N realizations for each parameter is generated by sampling from the initial geo-model with given statistics. This makes up what we call the *prior distribution*. We

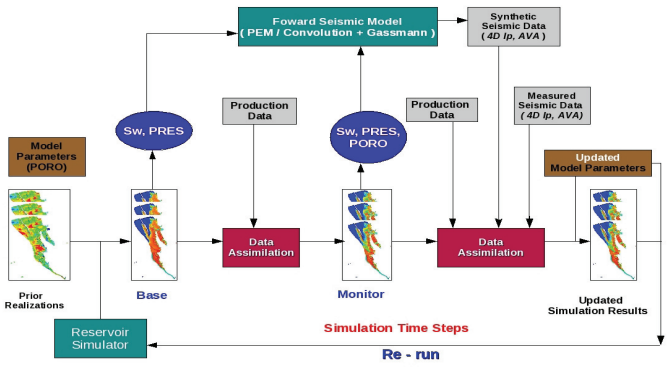


Figure 6.3: 4D seismic history matching work flow for EnKF.

run the reservoir simulator up to the time where the first measurements are available. Typically these are production data including oil production rate, water-cut and gas-oil ratio. At that time the *forecasted* model state

vector of dynamic and static variables are updated by including the measurements and their uncertainties giving the *analyzed* model state. At each assimilation step the model is updated and there is no need to run the simulator from time zero again. At some certain time steps, seismic data may be available in addition to production data. When the seismic data are given as a difference between two surveys, a combination of EnKF and the ensemble Kalman smoother (EnKS) has to be applied, see [135]. As an implementation of EnKS, one can use the same porosity for both base and monitor survey times to calculate respective seismic signatures (see, Fig. 6.3). In the paper A and paper B, we used this approach to incorporate EnKS when we integrated 4D seismic data.

6.4 Challenges of 4D History Matching in EnKF

The inclusion in 4D seismic data in history matching process is a demanding task and when this integration is made in EnKF framework, this becomes even more difficult because of the challenging issues of of EnKF itself. In the following section, we have described some of the demanding tasks in EnKF-based 4D history matching.

Seismic Modeling Error

4D seismic data offers spatial and dynamic information about changes in reservoir fluid properties and can be used to constrain flow simulation models thereby improving confidence in the reservoir characterization. But this process is affected by scale and process dependent model errors [137]. To speed up reservoir simulations, some form of up-scaling is required to capture two-phase flow properties such as relative permeability and also to represent geological heterogeneity. The up-scaling may be over-simplified or ignored. In addition, simplifications to the flow processes may be made, for example by using streamline methods. Finally, the petro-elastic transform contributes to the model errors due to assumptions about saturation distributions and cross-scaling is required because modeled and observed seismic are obtained for different volumes. Also the model errors in 4D seismic integration depends on the rock physics parameters as well as the underlying geo-model. These can be important, however, as they can affect the convergence rate, and may lead to incorrect models or alter the posterior uncertainty estimate [33].

Proper Estimation of Rock physics Parameters

Estimation of rock physics parameters, e.g., lithology (clay ratio) or fracture density parameter for fractured reservoirs may also be crucial if an incorrect rock physics model is introduced in the seismic history matching scheme. Such a model may introduce large modifications in the model variables, and such large updates can lead to model instabilities and make the prediction corrupt [132]. Hence, sensitivity studies related to rock physics parameter should be performed.

Appropriate Level of 4D Seismic Data Integration

A challenging issue when using seismic data is how to compare the measured data to the model data (explained in chapter 1). In principle, seismic data may be included in a number of different ways when conditioning reservoir simulation models[134]. One possibility is to use seismic amplitude data, another is to use inverted and/or processed data (such as acoustic impedance and Poisson's ratio). Finally saturations and pressures derived from the rock physics model can be used as well to compare with inverted saturation and pressure data. The amount of data, uncertainty, and information content may be very different and dependent on the way seismic data are incorporated. But it is not always straight forward to determine the appropriate level of 4D seismic data integration as shown in paper A.

Spurious Correlations and Rank Issues

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 in EnKF. This may contribute to the long-range spurious correlation effects in the updated fields, and a result of this sampling error is that the updated ensemble variance is under estimated. The rank of the error covariance matrix $\mathbf{C}^{e}_{\epsilon^{e}_{k}}$ is less than or equal to the number of ensemble members, and consequently, the number of perfect data that can be assimilated in a standard EnKF is severely limited. The problem of rank deficiency is most significant for problems with large amount of independent data, such as 4D seismic [3]. In the next chapter, we have described both of these issues in more details. In Paper B, we have formulated a localization strategy to handle large amount of 4D seismic data in EnKF for a real field case study.

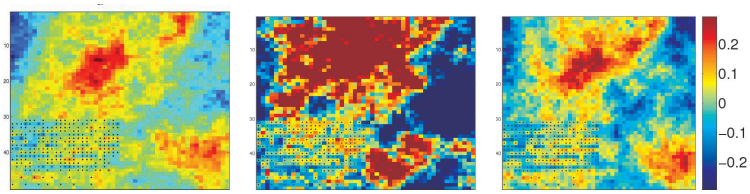
Chapter 7

EnKF and Large Number of Measurements

When we assimilate 3D or 4D seismic data, or interpreted seismic data such as acoustic impedance or Poisson's ratio, we effectively have a data measurement corresponding to each simulation grid block. Because of practical limitations, EnKF is typically applied with an ensemble size of the order of 100 and each updated (analyzed) model must be a linear combination of the initial ensemble of models [64], there are not sufficient degrees of freedom to match all seismic data. The loss of rank in the updated covariance matrix, due to finite ensemble size, further exacerbates the necessary degrees of freedom [86]. Moreover, even if a full rank version of measurement error covariance matrix is used, the matrix that must be inverted in the data assimilation step is $N_d \times N_d$, where N_d, is the number of data to be assimilated at a particular step [160]. When assimilating seismic data, N_d is often so large that it is not feasible to solve a $N_d \times N_d$, matrix problem. Hence, there are certain features, e.g., filter divergence and ensemble collapse are associated with EnKF when it is to assimilate large data sets such as 4D seismic data. In this chapter we would investigate these limitations of EnKF and would focus on the mitigation of these problems by using localization [106].

7.1 Spurious Correlations and Rank Issues

The use of a finite ensemble size to approximate the error covariance matrix introduces sampling errors that are seen as spurious correlations over long spatial distances or between variables known to be uncorrelated [62]. A result of this sampling error is that the updated ensemble variance is under



(a) True field (b) N_e is small (c) N_e is very large

Figure 7.1: EnKF update & spurious long-range correlation effect.

estimated. In the **Fig** 7.1, we observe that for a relatively large number of measurements (shown as black dots), we clearly see that the field far from the measurement locations has been incorrectly updated due to the fact that we have used smaller number of ensemble (Fig. 7.1b). As soon as we have very large ensemble, the true field is recovered (Fig 7.1c). In paper C, we concluded that the number of ensemble members has to be much larger (ten times or more) than the number of measurements to avoid the spurious correlations and obtain a good solution. Furthermore, it has been shown by van Leeuwen [148], that such errors tend to lead to systematic under estimation of model error variances, and ultimately to filter divergence, which in ensemble filter is accompanied by ensemble collapse [8] [73].

The power of EnKF is a direct consequence of the ability to approximate the entire covariance matrix from an ensemble that is often orders of magnitude smaller than the number of the state variables. As it has been shown in the Eq. (5.25), the ensemble representation of the error covariance matrix of observations can be written as:

$$\mathbf{C}^{e}_{e_o} = \frac{\mathbf{E}\,\mathbf{E}^T}{N_e - 1}. \tag{7.1}$$

The rank of the matrix $\mathbf{C}^{e}_{e^o_k}$, is less than or equal to the number of ensemble members, and consequently, the number of perfect data that can be assimilated in a standard EnKF is severely limited. The updates of the standard EnKF are also restricted to the subspace spanned by the members of the forecast ensemble (rank less or equal to N_e), and the assimilation of a perfect observation removes one degree of freedom from the ensemble [94]. The problem of rank deficiency is most significant for problems with large amount of independent data, such as 4D seismic [3].

7.2 Efficient Sub-space EnKF Formulation

One suggestion for dealing with the computational efficiency and rank issues is to use a subspace EnKF inversion scheme [60] [131], where data mismatches and simulated measurement errors are projected onto the subspace, L_s spanned by the principal left singular values of the measurement of the ensemble perturbations $\mathbf{S}$, shown in Eq. 5.28. First, we compute the singular value decomposition of $\mathbf{S}$ as

$$\mathbf{S} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T, \tag{7.2}$$

with $\boldsymbol{U} \in R^{m\times m}$, $\boldsymbol{\Sigma} \in R^{m\times N_e}$, and $\boldsymbol{V} \in R^{N_e\times N_e}$. If there are p, non-zero singular values and $\boldsymbol{\Sigma}_p$, is the $p \times p$, diagonal matrix with non-zero singular value λ_j, as its j-th diagonal entry, then the matrix $\mathbf{S}\mathbf{S}^T$, can be written as:

$$\mathbf{S}\mathbf{S}^T = \boldsymbol{U}_p\boldsymbol{\Sigma}_p^2\boldsymbol{U}_p^T, \tag{7.3}$$

where $\boldsymbol{U}_p$, is an $N_e\times p$, matrix with j-th column equal to the j-th left singular vector. In order to avoid loss of rank $\mathbf{E}$, is projected onto the subspace spanned by columns of $\boldsymbol{U}_p$, and data mismatch terms are also projected onto this space as shown by Zhao et al. [160]. With this approach and also by using pseudo-inverse, the analysis equation in Eq. (5.26) can be written as:

$$\widehat{\boldsymbol{\Psi}}^a = \widehat{\boldsymbol{\Psi}}^f\left(\mathbf{I} + \widehat{\boldsymbol{S}_p}^T(\boldsymbol{\Sigma}_p\,\boldsymbol{\Sigma}_p^T + \widehat{\boldsymbol{E}_p}\widehat{\boldsymbol{E}_p}^T)^{-1}\widehat{\boldsymbol{D}_p'}\right), \tag{7.4}$$

where $\widehat{\boldsymbol{E}_p}$, is the projection of the ensemble of the measurement perturbation $\boldsymbol{E}$ onto the subspace L_s, and $\widehat{\boldsymbol{S}_p}$, is the projected measurements of the ensemble perturbations and $\widehat{\boldsymbol{D}_p}$, is the projected ensemble of innovation vectors [131]. This formulation ensures that the measurement perturbations explain the variance within the ensemble space L_s, and we avoid the loss of rank pointed out by Evensen [60]. However, there is no assurance that when using this method, one will obtain realizations of model parameters that honor all observed data, and in this case, the estimate of model parameters may be considerably less accurate than the estimate that could be obtained by using a larger ensemble size [3].

7.3 Localization Methods

The sensitivity of the EnKF to ensemble size has been extensively investigated in the literature, with varied results [80] [8]. It is observed that

because the increments derived from ensemble-based schemes, without localization, are a linear combination of ensemble perturbations, it is essential for a successful ensemble-based assimilation that the ensemble adequately spans the model sub-space [59]. A possible solution to situations when only a small ensemble is feasible is the use of a technique called localization [106]. Localization is typically achieved either by masking of covariances between distant elements of the model state vector [73] [79] or by applying filters locally in physical space [113] [61]. The primary benefit of localization is often considered to be the reduction of spurious long-range covariances occurring when a small ensemble is used. More importantly, localization effectively increases the rank of the system. By increasing the rank of the system, an analysis can be computed that fits the observations better.

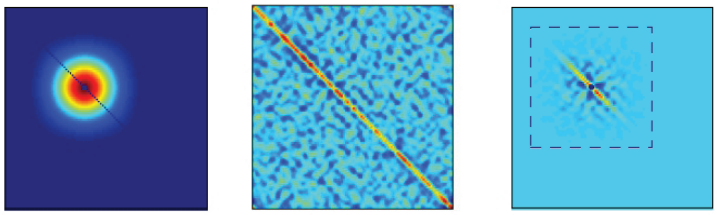
7.3.1 Local Updating (Analysis)

An effective approach to reduce the influence of spurious estimate of correlations on the update is the local analysis (local updating) [9] [132]. In the simplest form of the local analysis, one would update all the components of the state vector associated with simulator grid lock i using all data in some local neighborhood N_i of grid block. If the neighborhood includes all grid blocks in the simulation model, the local updating is equivalent to the standard EnKF. The degree of approximation introduced by local analysis is dependent on the range of influence defined for the observations. In the limit when this range becomes large enough to include all the data, for all the grid points, the solution becomes identical to the standard global analysis [62].

The actual algorithm can be explained as follows [61]: we first construct the input matrices to the global EnKF, i.e., the measured ensemble perturbations $\mathbf{S}$, the innovations $\mathbf{D}^{'}$, the measurements perturbations $\mathbf{E}$, or the measurement error covariance matrix, $\mathbf{C}_{ee}$. We then loop through the model grid, and for each grid point, e.g. (i, j) for a two dimensional model, we extract the rows from these matrices corresponding to measurements that will be used in the current update, and then compute the matrix $\mathbf{X}_{(i,j)}$ which defines the update for grid point (i, j). The analysis at grid point (i, j), i.e. $\mathbf{A}^{a}_{(i,j)}$, becomes

$$\begin{aligned}\mathbf{\Psi}^{a}_{(i,j)} &= \mathbf{\Psi}_{(i,j)}\mathbf{X}_{(i,j)} \\ &= \mathbf{\Psi}_{(i,j)}\mathbf{X} + \mathbf{\Psi}_{(i,j)}(\mathbf{X}_{(i,j)} - \mathbf{X}),\end{aligned} \tag{7.5}$$

where $\mathbf{X}$ is the global solution, while $\mathbf{X}_{(i,j)}$ becomes the solution for a local



(a) Local update stencil (b) $\mathbf{C}^e_\psi$ (c) Local update of $\mathbf{C}^e_\psi$

Figure 7.2: Local updating procedure for $\mathbf{C}^e_\psi$.

analysis corresponding to grid point (i, j) where only the nearest measurements are used in the analysis. Thus, it is possible to compute the global analysis first and then add the corrections from the local analysis if these are significant [61]. In **Fig.** 7.2a, we observe a possible local update stencil in the parameter grids; for us, the region of interest is the circular area. Local scheme assumes that only observation data in the designated region would impact state variable within the same region. The effect of this choice of local area can be seen on the update of $\mathbf{C}^e_\psi$, see Fig. 7.2c. The Fig. 7.2 and also the Fig. 7.3 are taken from the work of Pavel Sakov [122]. In case for the global analysis scheme, we have only one region, and hence, at the limit when this range of local analysis becomes large enough to include all the data, for all the grid points, the solution becomes identical to the standard global analysis. In paper B, we have introduced a localization strategy based on a combination of a local and a global analysis scheme. The global and the local analysis are used to assimilate the production and 4D seismic data respectively. The local scheme assumes that seismic data, within a given local analysis region, is influenced by only variables in the same region. This approach is different than the distance-based localization as proposed by Skjervheim et al. [132], in the sense that we are using a certain region of influence, instead of using local update centered on a single grid cell.

Modified EnKF Analysis Scheme Based on Local Update

From basic EnKF formulation in Eq. (5.28), the ensemble representation of the analysis equation can be written as:

$$\widehat{\mathbf{\Psi}}^a_k = \widehat{\mathbf{\Psi}}^f_k + \mathbf{C}^{e,f}_{\Psi_k}\mathbf{H}^T_k(\mathbf{H}_k\mathbf{C}^{e,f}_{\Psi_k}\mathbf{H}^T_k + \mathbf{C}_{e^o_k})^{-1}(\mathbf{d}^i_k - \mathbf{H}_k\widehat{\mathbf{\Psi}}^f_k). \tag{7.6}$$

This equation of analysis step is 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 seismic data influencing the analysis for grid cell l, is denoted by $\mathbf{d}_{sl} \in \mathbf{R}^{m_{s_l}}$, and extracted by matrix $\mathbf{L}_l$, which includes only zeros and ones and picks the seismic data which within this prescribed local region. Correspondingly, the local observation equation can be written as following Skjervheim et al. [132]

$$\mathbf{d}_{sl} = \widetilde{\mathbf{H}}_{sl}\widetilde{\mathbf{\Psi}}_l + \epsilon_{sl} \qquad \epsilon_{sl} \backsim N_{m_k}(0, \mathbf{C}_{\epsilon_{sl}}), \tag{7.7}$$

where $\widetilde{\mathbf{H}}_{sl} = \mathbf{L}_l\widetilde{\mathbf{H}}_s$, $\epsilon_{sl} = \mathbf{L}_l\epsilon_s$, and the error covariance matrix is given by $\mathbf{C}_{\epsilon_{sl}} = \mathbf{L}_l\mathbf{C}_{\epsilon_s}\mathbf{L}_l^T$.

Defining the local representation of the seismic measurements of the ensemble perturbations, $\mathbf{S}_{sl} \in \mathbf{R}^{m_{sl}\times N_e}$, as

$$\mathbf{S}_{sl} = \widetilde{\mathbf{H}}_{sl}\widehat{\mathbf{\Psi}}_p^{a'} = \mathbf{L}_l\widetilde{\mathbf{H}}_s(\widehat{\mathbf{\Psi}}_p^a - \bar{\widehat{\mathbf{\Psi}}}_p^{a'}), \tag{7.8}$$

and the ensemble innovation vectors $\mathbf{D}_{sl}^{'} \in \mathbf{R}^{m_{sl}\times N_e}$ as

$$\mathbf{D}_{sl}^{'} = \mathbf{D}_{sl} - \widetilde{\mathbf{H}}_{sl}\widehat{\mathbf{\Psi}}_p^{a'} = \mathbf{L}_l(\mathbf{D}_s - \widetilde{\mathbf{H}}_s\widehat{\mathbf{\Psi}}_p^{a'}), \tag{7.9}$$

and the matrix $\mathbf{G}_{sl} \in R^{m_{sl}\times m_{sl}}$ as

$$\mathbf{G}_{sl} = \mathbf{S}_{sl}\mathbf{S}_{sl}^T + (N-1)\mathbf{C}_{\epsilon_{s_l}}, \tag{7.10}$$

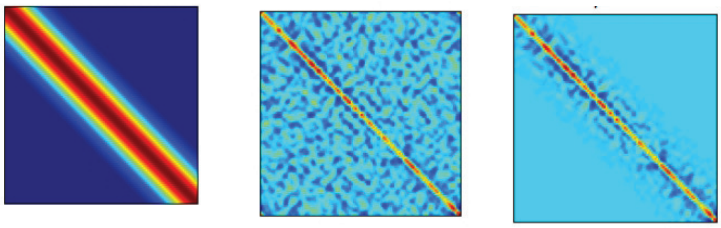
finally, the local analysis equation at grid point l becomes

$$\begin{aligned}
\widehat{\mathbf{\Psi}}_l^a &= \widehat{\mathbf{\Psi}}_{p,l}^a + \widehat{\mathbf{\Psi}}_{p,l}^{a'}\widehat{\mathbf{\Psi}}_{p,l}^{a'T}\widetilde{\mathbf{H}}_s^T(\widetilde{\mathbf{H}}_s\widehat{\mathbf{\Psi}}_{p,l}^{a'}\widehat{\mathbf{\Psi}}_{p,l}^{a'T}\widetilde{\mathbf{H}}_s^T + (N-1)\mathbf{C}_{\epsilon_s})^{-1}\mathbf{D}_{sl}^{'} \\
&= \widehat{\mathbf{\Psi}}_{p,l}^a + \widehat{\mathbf{\Psi}}_{p,l}^a(\mathbf{I} - \mathbf{1}_N)\mathbf{S}_{sl}^T\mathbf{G}_{sl}^{-1}\mathbf{D}_{sl}^{'} \\
&= \widehat{\mathbf{\Psi}}_{p,l}^a(\mathbf{I} + (\mathbf{I} - \mathbf{1}_N)\mathbf{S}_{sl}^T\mathbf{G}_{sl}^{-1}\mathbf{D}_{sl}^{'}) \\
&= \widehat{\mathbf{\Psi}}_{p,l}^a(\mathbf{I} + \mathbf{S}_{sl}^T\mathbf{G}_{sl}^{-1}\mathbf{D}_{sl}^{'}) \\
&= \widehat{\mathbf{\Psi}}_{p,l}^a\mathbf{X}_l,
\end{aligned} \tag{7.11}$$

where the ensemble matrix conditioned to production data at the same grid $\widehat{\mathbf{\Psi}}_{p,l}^a$, is used as forecast. The matrix $\mathbf{X}_l \in \mathbf{R}^{N_e\times N_e}$, can be written as

$$\mathbf{X}_l = \mathbf{I} + \mathbf{S}_{sl}^T\mathbf{G}_{sl}^{-1}\mathbf{D}_{sl}^{'}, \tag{7.12}$$

where we have used that $\mathbf{1}_N\mathbf{S}_{sl}^T \equiv 0$. By solving the local analysis equation in Eq. (7.11) for each grid cell in the reservoir, we obtain the global analysis ensemble conditioned to both production and seismic data.



(a) Correlation matrix, Ω (b) $\mathbf{C}^e_\psi$ (c) $\mathbf{C}^e_{\psi,\omega} = \Omega \circ \mathbf{C}^e_\psi$

Figure 7.3: Covariance localization of $\mathbf{C}^e_\psi$.

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 [132]. 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 nonphysical 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. It is then required that a large influence region is used in the local update to preserve the smoothness of the analysis fields [62].

7.3.2 Covariance Localization

An effective implementation of the localization can also be achieved by multiplying the ensemble-based estimate of the covariance element wise by a compactly supported positive definite matrix Ω, to produce a localized covariance estimate $\mathbf{C}^e_{\psi,\Omega} = \Omega \circ \mathbf{C}^e_\psi$. The element-wise product of matrices are known as Schur or Hadamard product. Houtekamer and Mitchell [79] simply applied a distance cutoff to the Kalman gain so that only model parameters within a critical distance of the observation were updated, see **Fig.** 7.3. In actual EnKF applications, the covariance matrix for the model and state variables is never directly computed or even stored, and hence,

the localization is typically preformed on components:

$$[(\Omega \circ \mathbf{C}^e_\psi)\mathbf{H}^T][\mathbf{H}(\Omega \circ \mathbf{C}^e_\psi)\mathbf{H}^T + \mathbf{C}^d_\psi]^{-1} = [(\Omega\mathbf{H}) \circ (\mathbf{C}^e_\psi\mathbf{H}^T)][(\mathbf{H}\Omega\mathbf{H}^T) \circ (\mathbf{H}\mathbf{C}^e_\psi\mathbf{H}^T) + \mathbf{C}^d_\psi]^{-1}. \tag{7.13}$$

Although the above relationship is not valid for arbitrary $\mathbf{H}$, it is valid for the block identity matrices typically used in the EnKF for augmented state vectors. Most of the practitioners who use localization appear to have chosen to use the fifth-order compactly supported correlated function of Gaspari and Cohen [67] for Ω to eliminate spurious correlations. Instead of using one particular correlation function for localization, Furrer and Bengtsson [65] consider the problem of estimating an optimal distance-dependent localization function. This approach involves the minimization of the difference between true covariance matrix and the localized ensemble estimate of the covariance. Devegowda et al. [40] proposed a covariance localization method for reservoir flow that uses sensitivities from the streamline simulation method to quantify the region of influence of model parameters on the observed data. Trani et al [146] used this approach based on covariance localization and obtained promising results for their synthetic case and they have updated parameters conditioned to seismic data as well. However, it is not clear how well covariance localization can work when the prior covariance has long-range correlations, see Aanonsen et al. [3].

Inflation and Adaptive Localization

A covariance inflation procedure [12] can be used to counter the variance reduction observed due to the impact of spurious correlations as well as other effects leading to under estimation of ensemble variance. The inflation factor is used to replace the forecast ensemble according to the equation shown in Evensen [62]:

$$\mathbf{\Psi}_j = \gamma(\mathbf{\Psi}_j - \overline{\mathbf{\Psi}}) + \overline{\mathbf{\Psi}}, \tag{7.14}$$

with the inflation parameter γ, slightly greater than one (typically 1.01) and $\overline{\mathbf{\Psi}}$, is the mean of the ensemble. The inflation parameter is a tuning parameter and optimally it is best estimated adaptively. The need for inflation depends on the use of a local versus global analysis scheme, and the use of a local scheme can, to a large scale, reduce the necessity of using an inflation parameter [62].

In adaptive localization methods, the assimilation system itself is used to determine the localization strategy. Such algorithms are useful since the dynamical covariance functions change in space and time, and the spurious correlations depend on ensemble size [62]. The hierarchical appro¡ch by Anderson [10] is based on the online computation of a flow-dependent moderation function and Monte Carlo method is used to dump long-range and spurious correlations. Although this approach is statistically consistent, it is optimized for small ensemble and may become sub-optimal when used with the full ensemble including all realizations. An alternative online localization scheme is known as SENCORP (Smoothed Ensemble Correlations raised to a Power) [21] which computes a flow-dependent moderation function to dump the spurious correlations. The idea is that the moderation functions can be generated from a smoothed covariance function, which when raised to a power, will dampen small correlations [62].

Chapter 8

Conclusions and Further Investigations

We have presented a seismic history matching work flow based on the ensemble Kalman filter (EnKF) for continuous model updating with both production and 4D seismic data. One objective of this project is to perform a through sensitivity study of the effects of different types of 4D seismic data in EnKF for a realistic synthetic reservoir model. We have also investigated the issues of large number of measurements and spurious correlations in EnKF model updating. For the real field case study, we have proposed a localization strategy based on a combined local and global analysis schemes. It is already mentioned that 4D seismic history matching in EnKF framework is a challenging task. Each of the preceding chapters provided the testimony of the relevant challenges for each of the different paradigm of field of applications that has to be dealt with. Still a daunting endeavor, we made an effort to tackle various aspects of history matching for our real field case study. The research findings show promising results, but to reach concrete conclusions, further investigations are necessary. In the following, we will describe the summary and accomplishments of this work first and then, we will provide directions for further research activities based on the findings and experiences from this PhD work.

8.1 Summary and Results

In spite of several challenges, e.g., issues of integration of large amount of 4D data and localization, the initial EnKF applications for 4D seismic data provided promising results. Still, examples of 4D seismic data integration

in EnKF framework for real field cases are virtually non-existent except few mentioned recent works. In this regard, this work makes an endeavor to analyze some of the important issues relevant to real field case seismic history matching using EnKF. In pursuit of this goal, we have implemented the rock physics model and forward seismic modeling tool for the realistic synthetic case. For the real field case, the additional task of interpretation of inverted seismic data, and sensitivity study of the rock physics model with respect to production effects were also performed. The next important task was to integrate all these disparate modeling tools in the EnKF framework which in our case was the Ensemble Reservoir Tool developed at Statoil ASA. At the end of this project, we have an integrated EnKF framework to update our reservoir model with both production and inverted seismic data.

One of the primary objectives of this project was to ascertain an optimal level of 4D seismic data integration in the history matching loop for the synthetic case. We made a good progress in this direction; for most of the cases of reservoir characterization, we experienced that the integration of seismic data in the elastic domain provided better results than using seismic data at the amplitude level. This may be due to the measurement error used, and hence, further investigations are necessary to ascertain the appropriate level of seismic data integration. The effect of vertical resolution of seismic modeling on the quality of history matching was an insightful finding. The linearity assumption of the forward modeling, and the idea of working only in depth simplified our seismic modeling tool to a great extent. Initial attempts were also made to consider fault multipliers and vertical transmissibilities as uncertain parameters. But because of large number of faults in the reservoir, and lack of related proper sensitivity studies, it was difficult to analyze the effects of these parameters.

For the real field case study, we intended to history match our complicated reservoir model with inverted impedance data. In spite of simplifications related to scaling and filtering of PEM models, the initial matching for seismic data are promising. All thorough the EnKF experiments, we used an ensemble of 100 realizations, and used an efficient subspace EnKF implementation. The localization strategy proposed for this field should be regraded as a preliminary attempt to overcome the spurious correlations effects of using finite ensemble size. Even with a very simple local analysis scheme, we were managed to reduce some spurious correlation effects. This showed that localization is of great importance when we integrate 4D seismic data in EnKF. Further investigations are necessary to determine a robust local analysis scheme for this field.

8.2 Further Investigations

In all of the EnKF experiments presented in this work, we have used an efficient subspace EnKF implementation. The relationship between reservoir parameters and measurements, especially seismic data, are non-linear, and hence, the linearity assumption in standard EnKF introduces approximations. A natural extension is to use some form of iterative Ensemble filters. The iterative EnKF based on RML (randomized maximum likelihood) proposed by Gu and Oliver [72] can be considered as a good start to explore the possibilities of using iterative EnKF in case of using seismic data.

For the coupled reservoir-seismic framework, we have considered same reservoir grid for both forward seismic modeling and geo-modeling. In reality, this is not appropriate as all these three paradigms have different grid resolutions. Seismic attributes are computed using the PEM on the fine grid includes high frequency term. On the other hand, seismic data are inherently acquired in a band limited frequency domain. Hence, synthetic attributes computed by PEM should be filtered, see Roggero et al. [119]. In order to be consistent to condition reservoir models with different data types, we should use appropriate up-/down-scaling techniques.

For seismic modeling, we have applied a different procedure, where all the calculations, including the convolution, are performed in depth. We also neglected over- and under-burden effects. This will of course not be possible in a real case. Hence, a proper depth-to-time conversion should be made in order to perform wavelet convolution in time. For a complicated reservoir like ours, one should use more accurate finite-difference modeling (FDM) [150] or ray-tracing methods [30].

In order to reduce the spurious correlation effects and collapse of the updated ensemble, different localization strategies should be explored. One option is to use a covariance localization method for reservoir flow that uses sensitivities from the streamline simulation method to quantify the region of influence of model parameters on the observed data as proposed by Devegowda et al. [41]. A covariance inflation procedure [12] can also be used to counter the variance reduction observed due to the impact of spurious correlations.

Bibliography

[1] Aanonsen S, *Using model error with the ensemble Kalman filter for more accurate model updates and predictions*, SPE 119096-MS, SPE Reservoir Simulation Symposium, The Woodlands, Texas, (2009).

[2] Aanonsen S, Avatsmark I, Gosselin O and Cominelli A, *Effect of scale dependent data corrections in an integrated history matching loop combining production and 4D seismic data*, SPE Reservoir Simulation Workshop, SPE 79655, (2003).

[3] Aanonsen S, Nævdal G, Oliver D, Reynolds A, and Vallès B, *The ensemble Kalman filter in reservoir engineering - a review*, SPE Journal, SPE 117274, (2009).

[4] Agbalaka C and Oliver D, *Application of the EnKF and localization to automatic history matching of facies distribution and production data*, Mathematical Geosciences, 40(4), (2008).

[5] Agustsson H and Grinestaff G, *A study of IOR by co_2 injection in the Gullfaks field, offshore Norway*, SPE/DOE Fourteenth Symposium on Improved Oil Recovery held in Tulsa, Oklahoma, (2004).

[6] Aki K and Richards P, *Quantitative seismology*. W. H. Freeman & Co., 1980.

[7] Andersen T, Zachariassen E, Otterlei C , Hatland K and Liestolt F, *Method for conditioning the reservoir model on 3D and 4D elastic inversion data applied to a fluvial reservoir in the North sea*, SPE 100190, SPE Europec/EAGE Annual Conference and Exhibition, Vienna, Austria, (2006).

[8] Anderson J, *An ensemble adjustment Kalman filter for data assimilation*, Monthly Weather Review, 129(2), (2001).

[9] ——, *A local least squares framework for ensemble filtering*, Monthly Weather Review, 131(4), (2003).

[10] Anderson J, *Exploring the need for localization in the ensemble data assimilation using a hierarchical ensemble filter*, Physica D, vol. 230, (2007).

[11] Anderson J and Anderson S, *A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts*, Mon. Weather Rev., vol. 127, (1999).

[12] ——, *A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts*, Mon. Weather Rev., vol. 127, (1999).

[13] Arenas E, van Kruijsdijk C, Oldenziel T, *Semi-automatic history matching using the pilot point method including time-lapse seismic data*, SPE 71634, SPE 2001 Annual Technical Conference, New Orleans, (2001).

[14] Aster R, Borchers B and Thurber C, *Parameter estimation and inverse problems*. Elsevier Academic Press, 2005.

[15] Avseth P, Mukerji T, and Mavko G, *Qualitative seismic interpretation: Applying rock physics tools to reduce interpretation risk*. Cambridge University Press, 2005.

[16] Aziz K and Settari A, *Petroleum reservoir simulation*. Applied Science Publishers, 1979.

[17] Bachrach R, *Joint estimation of porosity and saturation using stochastic rock-physics modeling*, Geophysics, vol. 71, (2006).

[18] Batzle M and Wang Z, *Seismic properties of pore fluids*, Geophysics Prospecting, vol. 57, (1992).

[19] Bianco A, Cominelli A, Dovera L, Nævdal G, Valles B, *History matching and production forecast uncertainty by means of the ensemble Kalman filter: A real field application*, SPE Europec/EAGE Annual Conference & Exhibition, London, (2007).

[20] Biondi B, Mavko G, Mukerji T, Lumley D et al., *Reservoir monitoring: A multidisciplinary feasibility study*, The Leading Edge, (1998).

[21] Bishop C and Hodyss D, *Flow-adaptive moderation of spurious ensemble correlations and its use in ensemble-based data assimilation*, Q. J. R. Meteorol. Soc., vol. 133, (2007).

[22] Bissel R, *Calculating optimal parameters for history matching*, 4th European Conference on the Mathematics of Oil Recovery, Roros, (1994).

[23] Buland A, *Bayesian seismic AVO inversion*. PhD Thesis, Norwegian University of Science & Technology, 2002.

[24] Buland A and Omre H, *Bayesian linearized AVO inversion*, Geophysics, Vol.68, No.1, (2003).

[25] Caers J, *Petroleum geostatistics*. Society of Petroleum Engineers, 2005.

[26] Calvert R, *Insights and methods for 4D reservoir monitoring and characterization society of exploration geophysicists*. SEG/EAGE Distinguished Instructor Short Course, 2005.

[27] Calvert R and Bakulin A, *New 4D seismic monitoring techniques as enablers for effective smart fields*, SPE 108207-MS, Digital Energy Conference and Exhibition, Texas, U.S.A, (2007).

[28] Castagna, J and Backus M, *Offset dependent reflectivity: Theory and practice of AVO analysis*. Soc. Expl. Geophys., Invest. Geophys, 8, 1993.

[29] Castro S, *A probabilistic approach to jointly integrate 3D/4D seismic, production data and geological information for building reservoir models*. PhD Thesis, Stanford University, USA, 2007.

[30] Cerveny V, *Seismic ray theory*. Cambridge University Press, 2000.

[31] Chadwick, R, Arts R, Eiken O, Kirby G, Lindeberg E, Zweigel P, *Dseismic imaging of an injected CO2 plume at the Sleipner field, central North sea*, In Davies, Richard J, (ed.) 3D seismic technology : application to the exploration of sedimentary basins, Geological Society of London, UK, (2004).

[32] Chen Y, Oliver D and Zhang D, *Data assimilation for nonlinear problems by ensemble Kalman filter with re-parameterizations*, J. Petrol. Science and Engineering, 66, (2009).

[33] Christie M, Glimm J, Grove J, Higdon, Wood-Schultz M et al., *Error analysis and simulations of complex phenomena*, Los Alamos Science, 29, (2005).

[34] Cominelli A, Dovera L and Naevdal G, *Benchmark study of ensemble Kalman filer methodology: History matching and uncertainty quantification for a deep water oil reservoir*, International Petroleum Technology Conference, Doha, Qatar, (2009).

[35] Cooper ME, Thorogood A et al., *Foinhaven active reservoir management: The time-lapse signal*, SEG Extended Abstracts, (1999).

[36] Corey T, *The interrelation between gas and oil relative permeabilities*, Prod. Monthly 19 (1), (1954).

[37] Dadashpor M, Landrø M, Kleppe J, *Nonlinear inversion for estimating reservoir parameters from time-lapse seismic data*, Journal of Geophysics and Engineering, 5(2008), (2006).

[38] Davies D and Maver K, *4D time-lapse studies and reservoir simulation to seismic modeling*, Offshore Technology Conference in Houston, Texas, (2004).

[39] Deutsch C and Journel A, *Stochastic reservoir modeling using simulate annealing and genetic algorithm*, SPE Formation Evaluation, 10(1), (1995).

[40] Devegowda D, Arroyo E, Datta-Gupta A, *Efficient and robust reservoir model updating using ensemble Kalman filter with sensitivity based covariance localization*, SPE 106144 presented at the SPE Reservoir Simulation Symposium, Houston, TX, (2007).

[41] Devegowda D, Arroyo E and Datta-Gupta A, *Flow-relevant covariance localization during dynamic data assimilation using EnKF*, Advances in Water Resources 33, (2), (2010).

[42] Dobrin M and Savit C, *Introduction to geophysical prospecting*. McGraw-Hill Companies, 1988.

[43] Dong Y, *Integration of time-lapse seismic data into automatic history matching*. PhD Thesis, The University of Oklahoma, USA, 2005.

[44] Dong Y and Oliver D, *Quantitative use of 4D seismic data for reservoir description*, SPE Journal, SPE 84571, (2005).

[45] ——, *Reservoir simulation model updates via automatic history matching with integration of seismic impedance change and production data*, International Petroleum Technology Conference, 3-5 December 2008, KualaLampur, Malaysia, (2008).

[46] Dong Y, Gu Y, Oliver D, *Sequential assimilation of 4D seismic data for reservoir description using the ensemble Kalman filter*, Journal of Petroleum Science and Engineering, 51(4), (2006).

[47] Doucet A, Freitas N and Gordon N (Eds.), *Sequential Monte Carlo methods in practice*. Springer-Verlag, New York, 2001.

[48] Dovera L and Della Rossa E, *Ensemble Kalman filter for Gaussian mixture models*, Petroleum Geostatistics, Cascais, Portugal, (2007).

[49] Doyen P, *Seismic reservoir characterization: An earth modeling perspective*. EAGE Publications, The Netherlands, 2007.

[50] Dubrule O, Escobar I, Kuznetsov D, *Constraining 3D earth models by multidisciplinary data- inversion, geostatistics, regularization and data assimilation*, EAGE Petroleum Geostatistics, Cascais, Portugal, (2007).

[51] Dvorkin J and Nur A, *Elasticity of high-porosity sandstones: Theory for two North sea data sets*, Geophysics, vol. 61, (1996).

[52] Dvorkin J, Mavko G, Nur A, *The effect of cementation on the elastic properties of granular material*, Mechanics of Material, vol. 12, (1991).

[53] ECLIPSE, *Eclipse Technical Description*. Schlumberger - GeoQuest, 2007.

[54] Ehrendorfer M, *A review of issues in ensemble-based Kalman filtering*, Metereologische Zeitschrift Vol. 16(6), (2007).

[55] El Ouair Y and Strønen L, *Value creation from 4D seismic at the gullfaks field: Achievements and new challenges*, SEG 2006 Annual Meeting, New Orleans, USA, (2006).

[56] El Ouair Y, Lygren and Springer M, *Integrated reservoir management approach: From time-lapse acquisition to reservoir model update at the Norne field*, International Petroleum Technology Conference, Doha, Qatar, (2005).

[57] Escobar I, Williamson P, Doyen P, Moyen R and Crozat T, *Fast geostatistical stochastic inversion in a stratigraphic grid*, 76th SEG Annual International Meeting, (2006).

[58] Evensen G, *Sequential data Assimilation with a nonlinear qasi-geostrophic model using Monte Carlo methods to forecast error statistics*, Journal of Geophisics Research, vol. 99, (1994).

[59] ——, *The ensemble Kalman filter: Theoretical formulations and practical implementation*, Ocean Dynamics, vol. 53, (2003).

[60] ——, *Sampling strategies and square root analysis schemes for the EnKF*, Ocean Dynamics, vol. 54, (2003).

[61] ——, *Data Assimilation: The ensemble Kalman filter*. Springer, Berlin, 2007.

[62] ——, *The ensemble Kalman filter for combined state and parameter estimation*, IEEE Control Systems Magazine, (2009).

[63] Evensen G and van Leeuwen P, *An ensemble Kalman smoother for nonlinear dynamics*, Mon. Weather Rev., 128, (2000).

[64] Evensen G, Hove J, Rieso E, Seim K and Espelid O, *Using the enkf for assisted history matching of a North sea reservoir model*, SPE Reservoir Simulation Symposium, SPE 89942-PA, (2007).

[65] Furrer R and Bengtsson T, *Estimation of high-dimensional prior and posterior covariance matrices in Kalman filter variants*, Journal of Multivariate Analysis 98(2007), (2006).

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

[67] Gaspari G and Cohn S, *Construction of correlation functions in two and three dimensions*, Quarterly Journal of Royal Meteorological Society, 125, (1999).

[68] Gassmann F, *Uber die elastizitat poroser medien*, Vier. der Natur. Gesellschaft in Zurich, vol. 96, (1951).

[69] Gosselin O, van de Berg S, Cominelli A, *Integrated history matching of production and seismic data*, SPE Annual Technical Conference and Exhibition, New Orleans, Louisiana, (2001).

[70] Gosselin O, Aanonsen S, Aavatsmark I, Cominelli A et al., *History matching using time-llapse seismic (HUTS)*, SPE 84464, SPE Annual Technical Conference and Exhibition, Colorado, USA, (2003).

[71] Gu Y and Oliver D, *The ensemble Kalman filter for continuous updating of reservoir simulation models*, Journal of Energy Resources Technology, (2006).

[72] ——, *An iterative ensemble Kalman filter for multiphase fluid flow data assimilation*, SPE Journal, SPE 108438-PA, (2007).

[73] Hamil T and Whitaker J, *Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter*, Monthly weather Review, 129, (2001).

[74] Hansen H, Gommesen L, Ouair Y El, *Using RFT data for elastic rock properties' dependency on pore pressure in 4D quantitative analysis*, SPE Europec/EAGE Annual Conference & Exhibition, Vienna, Austria, (2006).

[75] Hashin Z and Shtrikman S, *A variational approach to the elastic behavior of multi-phase materials*, Journal of Mech. Phys. Solids, vol. 11, (1963).

[76] Haugen V, Nævdal G.et al., *History matching using the ensemble Kalman filter on a North sea field case*, SPE Journal, SPE 102430, (2008).

[77] Haverl M, Aga M, Reiso E, *Integrated work flow for qualitative use of time-lapse seismic data in history matching: A North sea field case*, SPE Europec/EAGE Annual Conference, Madrid, Spain, (2005).

[78] Haverl M and Skjervheim J-A, *4D seismic modeling integrated with the ensemble Kalman filter method for history matching of reservoir simulation model*, 11th European Conference on the Mathematics of Oil Recovery, Bergen, Norway, (2008).

[79] Houtekamer P and Mitchell H, *A sequential ensemble Kalman filter for atmospheric data assimilation*, Monthly Weather Review, 129(1), (2001).

[80] Houtemaker P and Mitchell H, *Data Assimilation using an estimates in an ensemble Kalman filter technique*, Monthly weather Review, 126, (1998).

[81] Huang X, Meister L and Workman R, *Reservoir characterization by integration of time-lapse seismic and production data*, SPE 38695, SPE Annual Technical Conference, San Antonio, Texas, (1997).

[82] Jafarpour B and McLaughlin D, *Efficient permeability parameterization with the discrete cosine transform*, SPE-106453, SPE Reservoir Simulation Symposium, Houston, (2007).

[83] Jin L, Sen M Stoffa P and Seis R, *Time-lapse seismic attribute analysis for a water-flooded reservoir*, Journal Geophysical Engineering, 5, (2008).

[84] Kalman R, *A new approach to linear filtering and prediction problems*, Transactions of the ASME-Journal of Basic Engineering, 82(Series D), (1960).

[85] Kawar R, Calvert R, Khan M, *The work flow of 4D seismic*, SPE 81527, SPE 13th Middle East Oil Show & Conference, (2003).

[86] Kepert J, *On ensemble representation of the observation-error covariance in the ensemble Kalman filter*, J. Basic. Eng., vol. 82, (2004).

[87] Kirsch A, *Introduction to the mathematical theory of inverse problems*. Springer-verlag, New York, 1996.

[88] Kjelstadli R, Lane B, Barkved O, Buer K, and Kristiansen T, *Quantitative history match of 4D seismic response and production data in the Valhall field*, Offshore Europe, Aberdeen, United Kingdom, (2005).

[89] Kretz V, Le Ravalec-Dupin M and Roggero F, *An integrated reservoir characterization study matching production and 4D seismic*, SPE 77516, SPE Annual Technical Conference and Exhibition, San Antonio, USA, (2002).

[90] Landa J and Horne R, *A procedure to integrate well test data, reservoir performance history and 4D seismic information into a reservoir description*, SPE 38653,SPE Annual Technical Conference and Exhibition, San Antonio, USA, (1997).

[91] Landrø M, Solheim O, Hilde E, Ekren B, Strøne L, *The Gullfaks 4D seismic study*, Petroleum Geoscience, Volume 5 No. 3, (1999).

[92] Leeuwenburg O, Evensen G, and Bertino L, *The impact of ensemble filter definition on the assimilation of temperature profiles in the Tropical Pacific*, Journal of Royal Meteorological Society, (2006).

[93] Li R, Reynolds A and Oliver D, *History matching of three phase flow production data*, SPE Journal, 8(4), (2003).

[94] Lorenc A, *The potential of ensemble Kalman filter for NWP - a comparison with 4D-var*, Quarterly Journal of the Royal Meteorological Society, 129(595), (2003).

[95] Luenberger D, *Optimization by vector space methods*. John Wiley & Sons, Inc., New Yok, 1967.

[96] Lumley D, *Meren field Nigeria: A 4D case study*, SEG Extended Abstracts, 1628, (1999).

[97] Marsh M, *4D in reservoir management - successes and challenges*, IOR Views, Issue 8, (2004).

[98] Mavko G, Mukerji T, and Dvorkin J, *The rock physics handbook: Tools for seismic analysis in porous media*. Cambridge University Press, 2003.

[99] Maybeck P, *Stochastic models, estimation and control*. Academic Press, New Yok, 1979.

[100] Mezghani M, Fornel A, Langlais V, and Lucet N, *History matching and quantitative use of 4D seismic data for an improved reservoir characterization*, SPE 90420, SPE Annual Technical Conference and Exhibition, Houston, USA, (2004).

[101] Mindlin R, *Compliance of elastic bodies in contact*, Journal of Appl. Mech., vol. 16, (1949).

[102] Nævdal G, Johnsen L, Aanonsen S, Verfring E, *Reservoir monitoring and continuous model updating using ensemble Kalman filter*, SPE Journal 10, (2005).

[103] Najjar N, Strønen L, Alsos T, *Time-lapse seismic program at Gullfaks: value and the road ahead*, Petroleum Geoscience, Vol. 9, (2003).

[104] Nur A, *Four-dimensional seismology and (true) direct detection of hydrocarbon: the petrophysical basis*, The Leading Edge, 8(9), (1989).

[105] Nur A and Simmons G, *The effect of saturation on velocity in low porosity rocks*, Earth and Planetary Sciences Letters, 7, (1967).

[106] Oke P, Sakov P and Corney S, *Impacts of localization in the EnKF and EnOI: experiments with a small model*, Ocean Dynamics, vol. 57, (2007).

[107] Oldenziel T, *Time-lapse seismic within reservoir engineering*. PhD Thesis, Delft University of Technology, The Netherlands, 2003.

[108] Oliver D, *Multiple realizations of the permeability field from well-test data*, SPE Journal, 1(2), (1996).

[109] Oliver D, Reynolds A, Liu Ning , *Inverse theory for petroleum reservoir characterization and history matching*. Cambridge University Press, Cambridge, 2008.

[110] Omre H, *Stochastic reservoir models conditioned to non-linear production history observations*, Geostatistics 2000, Cape Town, (2000).

[111] Ostrander W, *Plane-wave reflection coefficients for gas sands at non-normal angles of incident*, Geophysics, V. 49, (1984).

[112] O'Sullivan A and Christie M, *Error models for reducing history match bias*, Computational Geosciences 9, (2005).

[113] Ott E, Hunt B, Zimin A et al., *A local ensemble Kalman filter for atmospheric data assimilation*, Tellus 56A, (2004).

[114] Ramarao B, Marsily G et al., *Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissibility fields, 1. Theory and computational experiments*, Water Resources Research, 31(3), (1995).

[115] Reis L, Marsily L et al., *Production data integration using a gradual deformation approach: Application to an oil field (offshore brazil)*, SPE 63064, (2000).

[116] Reynolds A, Zafari M and Li G, *Iterative forms of the ensemble Kalman filter*, Proceedings of the 10th European Conference on the Mathematics of Oil Recovery, Amsterdam, (2006).

[117] Ricker N, *The form and laws of propagation of seismic wavelets*, Geophysics, vol. 18, (1953).

[118] Rider M, *The geological interpretation of well logs (second edition)*. Rider-French Consulting Ltd, 2002.

[119] Roggero F, Ding D, Berthet P, Lerat O, *Matching of production history and 4D seismic data - application to the Girassol field, offshore Angola*, SPE Annual Technical Conference & Exhibition, Anaheim, California, SPE 109929, (2007).

[120] Roste T, Husby O et al., *Using 4D seismic to monitor fluid flow in a heterogeneous and compartmentalized reservoir - cases from the Heidrun field*, EAGE Annual Conference & Exhibition, Vienna, Austria, (2006).

[121] Saetrom J, *Hierarchical ensemble Kalman filter for observations of production and 4D seismic data*. Master's Thesis, Norwegian University of Science & Technology, 2007.

[122] Sakov P and Oke P, *A deterministic formulation of the ensemble Kalman filter: an alternative to ensemble square root filters*, Tellus 60A, (2008).

[123] Sandø I, Munkvold OP and Elde R, *4D geophysical data*, Geo ExPro,Issue NO 5, (2009).

[124] Schulze-Riegert R, Haase O, Nekrassov A, *Combined global and local optimization techniques applied to history matching*, SPE 59668, (2003).

[125] Seiler A, Evensen G, Skjervheim JA, Hove J, *Advanced reservoir management work flow using an EnKF based assisted history matching method*, SPE Reservoir Simulation Symposium, SPE 118906, (2009).

[126] Sen M, Gupta A, Stoffa L et al., *The application of simulated annealing to stochastic reservoir modeling*, SPE Applied Technology Series, 2(2), (1994).

[127] Sen M, *Seismic inversion*. Society of Petroleum Engineers, U.S.A., 2006.

[128] Shannon C, *Communication in the presence of noise*, Proc. Institute of Radio Engineers, vol. 37, (1949).

[129] Sheriff R and Geldart L , *Exploration seismology*. Cambridge University Press, 1999.

[130] Shuey R, *A simplification of Zoeppritz equations*, Geophysics, V. 50, (1985).

[131] Skjervheim J-A, *Continuous Updating of a coupled reservoir-seismic model using ensemble Kalman filter technique*. PhD Thesis, University of Bergen, Norway, 2007.

[132] Skjervheim J-A , Ruud B, Aanonsen S , Evensen G , Johansen T, *Using the ensemble Kalman filter with 4D data to estimate properties and lithology of reservoir rocks*, 10th European Conference on the Mathematics of Oil Recovery – Amsterdam, The Netherlands, (2006).

[133] Skjervheim J-A and Ruud B, *Combined inversion of 4D seismic waveform data and production data using ensemble Kalman filter*, 76th Annual International Meeting, SEG, Expanded Abstracts, (2006).

[134] Skjervheim J-A, Evensen G, Aanonsen S, Hove J , *A methodology to assimilate seismic data in addition to production data on real field applications using the ensemble Kalman filter*, Petroleum Geostatistics, Cascais, Portugal, (2007).

[135] Skjervheim, J-A, Evensen G, Aanonsen S, Ruud B, Johansen T, *Incorporating 4D seismic data in reservoir simulation models using ensemble Kalman filter*, SPE Journal, SPE 958789-PA, (2007).

[136] Staples R, Cook A, Braisby J, Mabillard A and Rowbotham P, *Integration of 4D seismic data and the dynamic reservoir model - revealing new targets in Gannet C*, EAGE Annual Conference & Exhibition, Vienna, Amsterdam, (2006).

[137] Stephen K, *Scale and process dependent model errors in seismic history matching*, Oil and Gas Science Technology- Rev. IFP, Vol 62, (2007).

[138] Stephen K and Macbeth C, *Seismic history matching in the Schiehallion UKCS field*, First Break, 24, (2006).

[139] Stephen K, Soldo J, Macbeth C and Christie C, *Multiple-model seismic and production history matching: A case study*, SPE 94173, SPE Journal, Volume 11, Number 4, (2006).

[140] Stolt R and Weglein A, *Migration and inversion of seismic data*, Geophysics, V. 50, (1985).

[141] Talukdar S, Instefjord R, *Reservoir management of the Gullfaks main field*, SPE Europec/EAGE Annual Conference & Exhibition, Madrid, Spain, (2008).

[142] Tarantola A, *Inverse problem theory and methods for model parameter estimation*. SIAM, Philadelphia, 2005.

[143] Tavassoli Z, Carter Z and King P, *Errors in history matching*, SPE 86883, SPE Journal, (2004).

[144] Tippett M, Anderson J, Bishop C, Hamill T, and Whitaker J, *Ensemble square-root filters*, Mon. Weather Rev., vol. 131, (2003).

[145] Todnem C, Arnesen L and Gassø R, *4D seismic and through tubing drilling and completion wells extend life on the Gullfaks field*, SPE / IADC drilling Conference, Amsterdam, The Netherlands, (2005).

[146] Trani M, Leeuwenburgh O, Arts R, Brouwer J, Douma S, *The value of streamlines based localization in the assimilation of 4D seismic data with the ensemble Kalman filter*, EAGE Annual Conference & Exhibition, Amsterdam, Netherlands, (2009).

[147] Van Ditzhuijzen R, Oldenziel T, and Van Kruijsdijk C, *Geological parametrization of a reservoir model for history matching incorporating time-lapse seismic based on a case study of the Stratfjord field*, SPE 71318, SPE Annual Technical Conference and Exhibition, New Orleans, USA, (2001).

[148] Van Leeuwen P, *Comment on data assimilation using an ensemble Kalman filter technique*, Mon. Weather Rev., 127, (1999).

[149] Vasco D, Datta-Gupta A, Behrens R et al., *Seismic imaging of reservoir flow properties: Time-lapse amplitude changes*, Geophysics, (2004).

[150] Virieux J, *P-S wave propagation in heterogeneous media: Velocity-stress finite-difference method*, Geophysics, vol. 51, (1985).

[151] Waggoner J, Cominelli A, and Seymour R, *Improved reservoir modeling with time-lapse seismic in a Gulf of Mexico gas condensate reservoir*, SPE 77514, SPE Annual Technical Conference and Exhibition, San Antonio, USA, (2002).

[152] Wang Zee, *Fundamentals of seismic rock physics*, Geophysics, vol. 66, NO.2, (2001).

[153] Wang Z, *Feasibility of time-lapse seismic reservoir monitoring: The physical basis*, The Leading Edge, 16 (9), (1997).

[154] Wang Z and Nur A, *Effect of temperature on wave velocities in sands and sandstones with heavy hydrocarbons*, SPE 15646, SPE Reservoir Engineering, (1988).

[155] Wayland J and Lee D, *Seismic mapping of EOR processes*, The Leading Edge, (1986).

[156] Wen X and Chen W, *Some practical issues on real-time reservoir model updating using ensemble Kalman filter*, SPE Journal 14, (2007).

[157] Wu Z and Datta-Gupta A, *Rapid history matching using a generalized travel time inversion method*, SPE 66352, (2001).

[158] Yadav S, *History matching using probabilistic approach in a distributed computing environment*. PhD Thesis, The University of Texas at Austin, USA, 2005.

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

[160] Zhao Y, Reynolds A and Li G, *Generating facies maps by assimilating production data and seismic data with the ensemble Kalman filter*, SPE 113990, Proceedings of the 16th SPE Improved Oil Recovery Symposium, (2009).

[161] Zoeppritz K, *Erdbebenwellen viiib, ueber reflexion and durchgang seismischer wellen durch unstetigkeitsflaechen*, Gottinger Nachrichten, (1919).

Part II

Included Papers and Reports