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

Assimilating Time-lapse Seismic Data in Reservoir Engineering Simulations

A thesis submitted to the
Delft Institute of Applied Mathematics
in partial fulfillment of the requirements

for the degree

MASTER OF SCIENCE
in
APPLIED MATHEMATICS

by

Joseph Majkut
Delft, the Netherlands
July 2008

TU Delft

MSc THESIS APPLIED MATHEMATICS

"Assimilating Time-lapse Seismic Data in Reservoir Engineering Simulations"

JOSEPH MAJKUT

Delft University of Technology

Daily supervisor	**Responsible professor**
Dr. Remus Hanea	Prof. dr. ir. Arnold Heemink

Committee members

Prof. dr. ir. Arnold Heemink	Prof. dr. ir. Jan Dirk Jansen
Prof. dr. ir. Rob Arts	Dr. ir. Sippe Douma
Dr. Dorota Kuriwicka	Dr. Remus Hanea

July 2008 Delft, the Netherlands

For Mr. Rosewater

Contents

Chapter 1

Introduction

The drive behind this project is the desire to be better able to model the processes within an oil reservoir undergoing production using a combination of simulation software and available observations. This is a desired ability because of the increasing availability of smart reservoir management technologies, which allow engineers to control many parameters within production and injection wells and better monitor the reservoir. The hope is that optimization techniques could be applied to the production parameters at existing wells within a field and the development plan of a reservoir. With the application of optimization techniques it will be possible to recover more of the hydrocarbons within a reservoir after the start of water injection and during the life-time of production. Such optimization routines will require reservoir monitoring schemes which will be able to provide accurate estimates of the fluid and rock properties of the reservoir and quantitative predictions of dynamic properties at multiple time scales.

Software-based monitoring schemes for producing reservoirs typically rely on some data assimilation technique to combine knowledge of the reservoir coming from simulation and observation. In this project, we focus on the use and evaluation of the Ensemble Kalman Filter (EnKF). Examples reporting positive history matching results for a reservoir using the EnKF are found in the following works: [12], [6] and [15]. This sequential data assimilation method provides a probabilistic way to look at reservoir model and improve it using observations such that the estimates of the filter are consistent with the production history of the reservoir. This is a process known as history matching. At the analysis step, the point in simulated time at which data is assimilated into the filter, a set of variables from the model are changed to make the results better match the true dynamics of the system. These variables can be the predicted variables of interest, such as the ratio of oil to water in a region of the reservoir, or a parameter of the model which is unknown, such as a rock property in that region. The quantaties subject

to analysis at the filtering step are generally; porosity, permeability, saturation and pressure. There is quite a bit of data available in the time domain at a reservoir's wells, but data between the wells is rare. One example is data taken from seismic surveys of a reservoir, which provide spatially dense information at a certain point in time.

In practice, not much is known about the dynamics in a reservoir as it is undergoing production, injection or other processes. There is a significant amount of information available to engineers at the wells and usually some sparse geological information, but otherwise engineers are blind to the internal processes of the reservoir. Seismic imaging allows the engineer to get somewhat of a picture of the reservoir away from wells. The technique uses shear and pressure waves to measure the impedance of the reservoir, which can be inverted to estimate fluid and rock properties. Field engineers often combine their experience reading maps of seismic properties to make educated guesses about the properties of a reservoir. These techniques seem to be reliable for production purposes and determining the location of new wells. Optimization procedures, however, require a quantitative estimate of some of the reservoir properties for seismic data to be of use. Currently, the data is usually processed through a full inversion. The seismic properties measured in a single survey of a field are inverted to the fluid properties of the reservoir. This inversion is non-linear and requires sufficient computational effort. Researchers in the field are actively pursuing alternatives to the full inversion.

The introduction of time-lapse seismic data, exploiting the difference in seismic measurements over time, to improve the historical strength of these methods is a new subject. One example is found in the paper by Skjervheim et. al. [16]. In that paper, the authors combine 4D seismic data with production data to realize a 'slight reduction' in production mismatch and significantly different porosity and permeability estimates from the production data base case. This is an important point, as one of the possible causes of gross mismatch with production data is a history matching scheme that updates the permeability and porosity in a poor way. It is hoped that seismic data will force better estimations of these parameters, allowing the model to perform better. The authors of [16] refer to the use of time-lapse data in their paper, but they are using a slightly different definition than I do in this report. They used fully-inverted data, meaning a one pass seismic survey is inverted into pressures and saturations, taking the difference between two runs after performing the inversion. I use the term time-lapse data to mean differences in measured impedances (I) and 2-way travel times (T_2).

1.1 Project Outlook

The goal for this project is to investigate ways to include 4-D seismic data into history matching schemes without having to perform a full inversion of the seismic data. Through discussions with my colleagues from the reservoir engineering section, we have highlighted two possible ways to realize this goal. Both methods will utilize the differences in impedence and time-travel for multiple seismic surveys to supply data to an ensemble kalman filtering scheme. The difference is where the necessary geologic modeling, which bridges seismic properties with reservoir properties, takes place within the filtering scheme.

In the first idea, shown graphically in figure 1.1, I will use a forward rock physics model to estimate the seismic properties corresponding to the physical changes in each member of the ensemble. This forward model is refered to as the Rock Physics Framwork and abbreviated RPF. The RPF predicts the seismic measurements we would expect, given the model state. These forecasts will then be combined with the instrumental data in a filtering step.

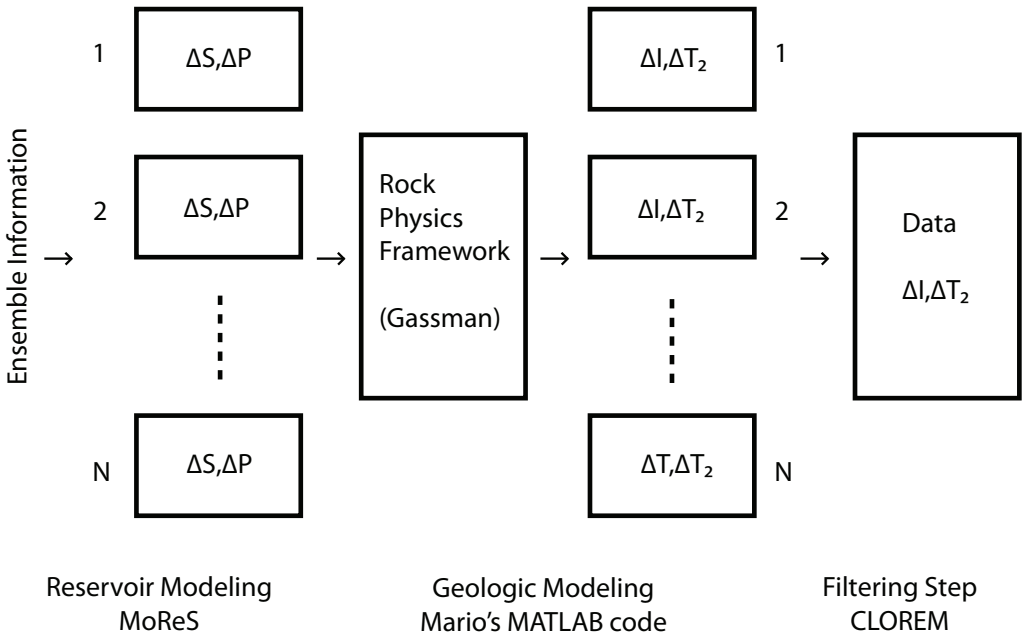

Figure 1.1: Graphical Representation of Method 1

The second method will use an inverted rock physics framework, Mario Trani's Modified Landro Method, to transform the measured changes in the seismic properties into changes in the reservoir properties. The inversion process is complicated and all of the processing takes place outside of the CLOREM-controlled filtering step, but it is computationally preferable to the full inversion previously

discussed.

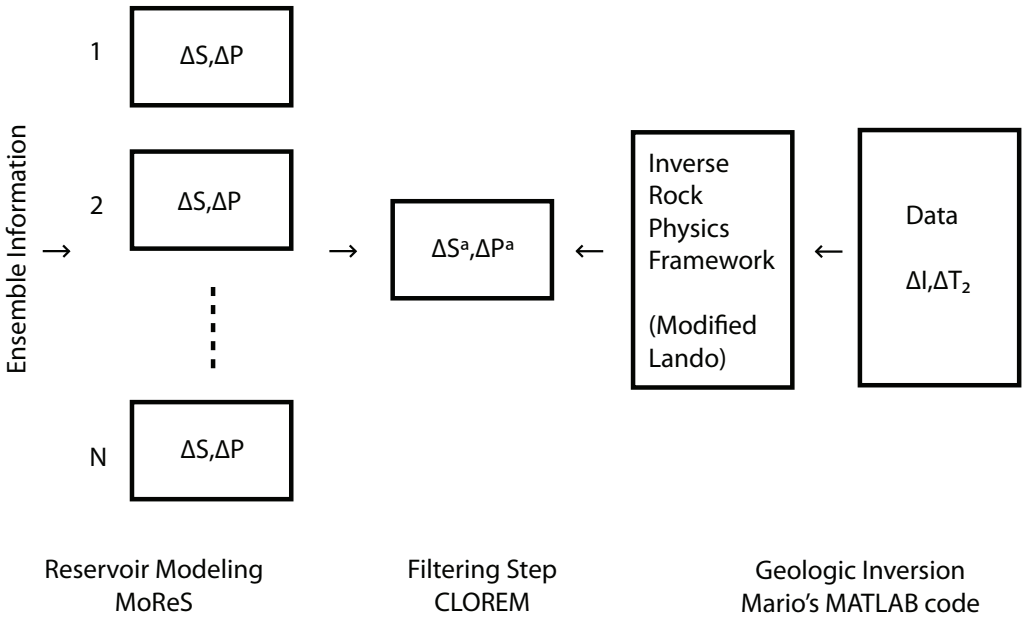

Figure 1.2: Graphical Representation of Method 2

This project report will focus on the use of time lapse seismic data, the difference in seismic properties between two or more successive surveys of a field, in assimilation. The difference in seismic measurements can be inverted into differences in reservoir properties using several methods based in geological physics. We cooperated with Mario Trani in the Faculty of Civil Engineering to implement his inversion technique, known as a Modified Landrø method.

1.2 Reservoir Simulation

Reservoir simulation is meant to provide field and research engineers with a quantitative estimate of hydrocarbon volume in a reservoir undergoing production. The quality of geological data for a particular reservoir and the quality of the modeling technique both contribute to the uncertainties in the estimates provided by simulation software. The field of reservoir simulation is quite large, but there are some important concepts to understand for this application. I will discuss these here by describing both reservoir structure and the methods by which they are represented.

A reservoir is typically represented on a grid structure, where each cell has it's own values of porosity, permeability, pressure, relative saturations of water and hydrocarbons, temperature and others. The definitions for these properties

follow. I have included the symbology that I will attempt to maintain throughout the report. They are interpreted from [3] and [7].

- Porosity, ϕ, The non-rock fraction of volume which can be occupied by liquids.
- permeability, k, A constant which describes the ease with which fluids flow through a given material.
- Pressure, P, The internal pressure of the reservoir.
- Saturation, S, The percentage, within a control volume, of a particular liquid. For example, S_w denotes water saturation. In practice, these saturations are bounded to an interval well within (0,1). It is not possible to force all of the oil out of a reservoir.

Likewise, the grid has a set of spatial properties.

- Grid Block Height, h, The physical height of a gridblock.
- Grid Block Length/Width, $\Delta x, /\Delta y$, Size of the gridblock in the lateral directions.

The simulation of a reservoir undergoing production is based on the conservation of momentum and mass within the reservoir. The equations for an isothermal reservoir with two-phase, oil/water, flow, as they are presented in [7] are based on Darcy's Law,

$$\begin{aligned} \bar{v}_w &= -\tfrac{k_{rw}}{\mu_w}\bar{K}(\nabla p_w - \rho_w g \nabla d), \\ \bar{v}_o &= -\tfrac{k_{ro}}{\mu_o}\bar{K}(\nabla p_o - \rho_o g \nabla d), \end{aligned} \tag{1.1}$$

Darcy's equations provide the conservation of momentum within a control volume. μ is the fluid viscosity, $\bar{K}$ is the permeability tensor, g is the acceleration of gravity, and d is the depth. The subscripts o and w refer to oil and water and k_{rw} and k_{ro} are the relative permiabilities of oil and water. The latter two quantities represent additional resistance to the flow of each liquid in the presence of the other. Typically, the coordinate system of the simulation is setup with regard to the geologic layering such that the permeability tensor is a diagnol matrix. Darcy's Law is an empirical relationship developed through labratory experiments. The Darcy Velocity, $\bar{v}$ is the hypothetical velocity of the fluid, absent the restricted volume within the pore space. The true velocity within the pores is $\bar{v}/\phi$. The mass-balance within a control volume can be written,

$$\nabla \bullet (\alpha \rho_w \bar{v}_w) + \alpha \frac{\partial(\rho_w S_w \phi)}{\partial t} - \alpha \rho_w q_w = 0, \tag{1.2}$$

$$\nabla \bullet (\alpha \rho_o \bar{v}_o) + \alpha \frac{\partial(\rho_o S_o \phi)}{\partial t} - \alpha \rho_o q_o = 0, \tag{1.3}$$

where $\alpha = 1$ (in three dimensions) and q is a source term which represents flow rate per unit time due to production and injection. This term is of importance near wells. $\nabla\bullet$ is the divergence operator. The equations given by mass-balance calculations and Darcy's Law are subject to the constraints

$$S_o + S_w = 1 \text{ and } p_o - p_w = p_c(S_w).$$

The first limits the model constitutents to oil and water. In the second constraint, the term p_c represents the capillary pressure at a given water saturation and represents the pressure drop at oil-water interfaces. In the cited example, Jansen [7], the capillary pressure is neglected. This is justifiable for that case because the capillary pressure causes a low-order diffusion, which is often unimportant on the scale of a reservoir and dwarfed in scale by transitive effects. It is also common to use numerical diffusion, resulting from a particular discretization scheme, to approximate the capillary diffusion. With capillary effects considered negligable, the constraint becomes,

$$p_o = p_w.$$

Lastly, the isothermal assumption gives equations of state which relate pressure, density and compressability.

$$c_o = \frac{1}{\rho_o}\frac{\partial \rho_o}{\partial p_o} \tag{1.4}$$

$$c_o = \frac{1}{\rho_w}\frac{\partial \rho_w}{\partial p_w} \tag{1.5}$$

$$c_r = \frac{1}{\phi}\frac{\partial \phi}{\partial p_o} \tag{1.6}$$

$$\tag{1.7}$$

The third line is the definition of rock compressability. Combining the above, we arrive at a reasonable set of governing equations,

$$-\nabla\left[\frac{\alpha\rho_w k_{rw}}{\mu_w}\bar{K}(\nabla p - \rho_w g\nabla d)\right] + \alpha\rho_w\phi\left[S_w(c_w + c_r)\frac{\partial p}{\partial t} + \frac{\partial S_w}{\partial t}\right] = \alpha\rho_w q_w, \tag{1.8}$$

$$-\nabla\left[\frac{\alpha\rho_o k_{ro}}{\mu_o}\bar{K}(\nabla p - \rho_o g\nabla d)\right] + \alpha\rho_o\phi\left[(1 - S_w)(c_o + c_r)\frac{\partial p}{\partial t} - \frac{\partial S_w}{\partial t}\right] = \alpha\rho_o q_o \tag{1.9}$$

These two equations do not differentiate between oil and water pressure, due to the lack of capillary pressure. The system allows us to see the kind of problem that a reservoir simulator is asked to solve, even though it is a simplification of the dynamics we will later model. The problem is parabolic with respect to pressure and approachs elliptic as compressability in the seccond term goes to zero. Saturation is described by mixed parabolic-hyperbolic terms, but is totally hyperbolic here because of the lack of capillary pressure. Thus, a simulation method should be designed to represent fast changes in pressure and more gradual changes in saturation.

In a reservoir simulator equations 1.8 and 1.9 would be combined with a well model, a detailed model of the geology in a neighborhood of a well and the conditions at the top of a well, which calculates the production of the well due to the pressure drop between the top and bottom. Such a model provides estimates of output rate (input rate in the case of a injector) and the water saturation of the produced (injected) fluid. These models are a field of study themselves and are trusted in comparison to porous media models like the one described above. They are not considered to be sources of high uncertainty in this project, so we will not describe them in detail. Jansen gives a simple introduction to well models to go along with the above physical modeling.

1.3 Seismic Reservoir Data

The goal of this project is to investigate how the information provided through the seismic study of a reservoir can be incorporated in a monitoring scheme. For an offshore reservoir, a seismic survey is carried out by a boat, which sends seismic waves into the reservoir and listens for them to return. Figure 1.3 shows a basic diagram of the process, where the first layer is the overburden of the reservoir and the second is the reservoir itself.

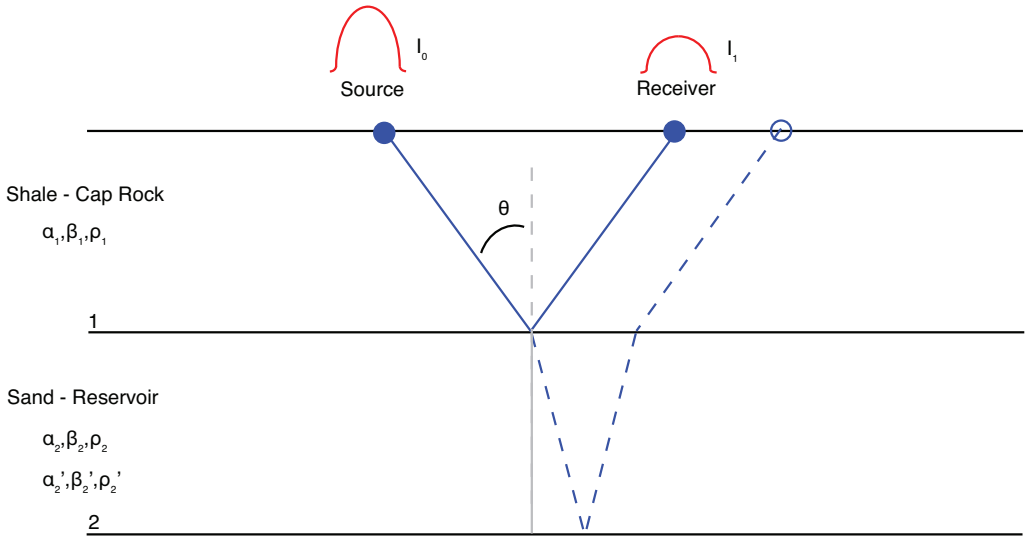

Figure 1.3: A sample path of a seismic wave, from the source to the reservoir bottom and back.

We consider two kinds of seismic data. The first is forward seismic data, often refered to as the seismic attributes, which is a certain type of data that is measured or easily calculated directly during a seismic survey. Section 1.3.1 shows how we generate this type of data for our synthetic studies by relating them to the fluid and geological properties of the reservoir. The second type of

seismic data we considered in known as time-lapse seismic data, modified Landrø data or inverted difference seismic data. Such data is in terms of changes in pressure and water saturation in the reservoir, ΔP and ΔS_w, and is calculated from the changes in the seismic attributes between two seismic surveys. The modified Landrø method is the work of Mario Trani, a phd student in the TU Delft faculty of Civil Engineering. The method that he modified is explained in section 1.3.2 and his modification is briefly motivated.

1.3.1 Seismic Attributes

The Seismic Attributes of a reservoir are those which can be directly measured or result from simple calculations on the measurements. In this project, we refer to the seismic attributes known at AVO data and the two-way travel time of P-waves in the reservoir. This section outlines how we model these attributes from the state of the reservoir for assimilation and as input for the time-lapse inversion. We solve for wave speeds in the reservoir in terms of geological and fluid properties and then translate wave speeds into reflectivity constants. The specific calculations can be found in [14].

We assume a reservoir like the one shown in figure 1.3, where the overburden has known seismic properties. We then use the Gassman equation to calculate the compressability of a control volume in the reservoir, including the rocks and fluids.

$$K_{sat} = K_{dry} + \frac{\left(1 - \frac{K_{dry}}{K_0}\right)^2}{\frac{\phi}{K_{fl}} + \frac{1-\phi}{K_0} - \frac{K_{dry}}{K_0^2}}. \tag{1.10}$$

K is the bulk modulus of the control volume, or its resistence to pressure changes, and the subscripts indicate a material for which K is known.

- K_{sat} is the bulk modulus of the reservoir, including fluid and rock effects.
- K_{dry} is the bulk modulus of the rock structure without fluid.
- K_0 is the bulk modulus of the minerals in the geological structure.
- K_{fl} is the bulk modulus of the fluid in the control volume. This changes with saturation and pressure changes.

A control volume also has a shear modulus, or resistance to shear waves, which does not depend on the fluid. Defining μ as the shear modulus,

$$\mu_{sat} = \mu_{dry}.$$

Given the moduli, and the density of the saturated volume, ρ_{Sat}, we calculate the wave speeds for P and S waves in the reservoir by,

$$\begin{aligned}\alpha &= \sqrt{\frac{K_{sat}+\frac{4}{3}\mu_{sat}}{\rho_{sat}}} \\ \beta &= \sqrt{\frac{\mu_{sat}}{\rho_{sat}}}.\end{aligned}$$

Keeping to the notations used in the seismic field, I define α as the P-wave velocity and β the velocity of S-waves. Additionally, I is the intensity of a signal, the square of the amplitude.

After calculating the wave speeds within the reservoir, we can move to the calculation of the seismic attributes. The two-way travel time of P-waves within the reservoir, T_2, is given by

$$T_2 = \frac{2 * D}{\alpha}$$

for a vertical shot. With an angle of incidence, the transform comes from trigonometric calculations. To relate α and β to the reflectivity at the interface we use the Zoeppritz equation,

$$R = \frac{\left(b\frac{\cos\theta_1}{\alpha_1} - c\frac{\cos\theta_2}{\alpha_2}\right)F - \left(a + d\frac{\cos\theta_1}{\alpha_1}\frac{\cos\psi_2}{\beta_2}\right)Hp^2}{D}. \tag{1.11}$$

Where θ is the P-wave angle of incidence at the interface and ψ is the S-wave angle of reflection at the interface.

$$\begin{aligned}p &= \frac{\sin\theta_1}{\alpha_1} \\ a &= \rho_2(1-2\beta_2^2p^2) - \rho_1(1-2\beta_1^2p^2) \\ b &= \rho_2(1-2\beta_2^2p^2) + 2\rho_1\beta_1^2p^2 \\ c &= 2\rho_2\beta_2^2p^2 + \rho_1(1-2\beta_1^2p^2) \\ d &= 2(\rho_2\beta_2^2 - \rho_1\beta_1^2) \\ F &= b\frac{\cos\psi_1}{\beta_1} + c\frac{\cos\psi_2}{\beta_2} \\ H &= a - d\frac{\cos\theta_2}{\alpha_2}\frac{\cos\psi_1}{\beta_1}\end{aligned}$$

The most important property of the relation in equation 1.11 is that the reflectivity depends on the angle of incidence θ. For small values of θ, the relationship of R to θ is approximately linear and we use a linear fit to define

$$R \approx R_0 + G\theta.$$

for several calculations of the Zoeppritz equation. R_0 is the reflectivity for a square incident wave and G represents how the reflectivity changes as the angle

of incidence increases. G should have negative sign. R_0, G and T_2 are the seismic attributes we use for direct assimilation in chapter 5 and as inputs for the modified Landrø method. These quantities are typically available from seismic surveys of a reservoir.

1.3.2 Landro Method

The following description of the estimation of change in pore pressure and water saturation in a reservoir using time-lapse seismic data is based on the method presented by Landrø[10]. The difference in wave intensities measured at the source and the receiver can be used to calculate the reflection coefficient. This measured property can be related to the initial properties at the interface by,

$$R_0(\theta) = \frac{1}{2}\left(\frac{\Delta\rho}{\rho} + \frac{\Delta\alpha}{\alpha}\right) - \frac{2\beta^2}{\alpha^2}\left(\frac{\Delta\rho}{\rho} + \frac{2\Delta\beta}{\beta}\right)\sin^2(\theta) + \frac{\Delta\alpha}{2\alpha}\tan^2(\theta). \quad (1.12)$$

In the above equation $\Delta\rho = \rho_2 - \rho_1$ and $\rho = (\rho_1 + \rho_2)/2$. These relationships are valid for α and β as well. The reflection coefficient can be written for the second imaging time and combined with eq. 1.12 to yield

$$\begin{aligned} R_1(\theta) = {} & \frac{1}{2}\left(\frac{\Delta\rho}{\rho} + \frac{\Delta\alpha}{\alpha}\right) - \frac{2\beta^2}{\alpha^2}\left(\frac{\Delta\rho}{\rho} + \frac{2\Delta\beta}{\beta}\right)\sin^2(\theta) + \frac{\Delta\alpha}{2\alpha}\tan^2(\theta) \\ & + \frac{1}{2}\left(\frac{\Delta\rho'}{\rho} + \frac{\Delta\alpha'}{\alpha}\right) - \frac{2\beta^2}{\alpha^2}\left(\frac{\Delta\rho'}{\rho} + \frac{2\Delta\beta'}{\beta}\right)\sin^2(\theta) + \frac{\Delta\alpha'}{2\alpha}\tan^2(\theta), \end{aligned}$$

with $\Delta\rho' = \rho_2' - \rho_2$, and so on. This combination ignores all second-order terms and relies on the assumptions

$$\frac{\Delta\alpha}{\alpha}, \frac{\Delta\alpha'}{\alpha} \ll 1 \quad (1.13)$$

$$\frac{\beta'}{\alpha'} \approx \frac{\beta}{\alpha}. \quad (1.14)$$

The detailed derivation is included in the appendix of [10]. At this point in the derivation, Landrø separates the effects of fluid substitution, a change in saturation, and pressure changes. Under fluid substitution the shear modulus remains constant, meaning that the speed of S-waves are unaffected, and thus $\Delta\beta^2\rho = 0$. This zeros the fifth term in R_1 and we see

$$R_1(\theta) = R_0(\theta) + \frac{1}{2}\left(\frac{\Delta\rho'}{\rho} + \frac{\Delta\alpha'}{\alpha}\right) + \frac{\Delta\alpha'}{2\alpha}\tan^2(\theta). \quad (1.15)$$

Expressed as the change in reflectivity due to fluid substitution,

$$\Delta R(\theta) = \frac{1}{2}\left(\frac{\Delta\rho'}{\rho} + \frac{\Delta\alpha'}{\alpha}\right) + \frac{\Delta\alpha'}{2\alpha}\tan^2(\theta). \quad (1.16)$$

Under a change in pore pressure, density should not change, because changes in porosity due to pressure changes are negligible. Mathematically,

$$\frac{\Delta\rho}{\rho} \approx 0.$$

This leads to the change in reflectivity due to a change in pore pressure,

$$\Delta R(\theta) = \frac{1}{2}\frac{\Delta\alpha'}{\alpha} - \frac{4\beta^2}{\alpha^2}\frac{\Delta\beta'}{\beta}\sin^2(\theta) + \frac{\Delta\alpha'}{2\alpha}\tan^2(\theta). \tag{1.17}$$

Equations and are accurate to first-order changes in the wave velocities and density and they are combined linearly to arrive at the total change in reflection constant. It is given by

$$\Delta R(\theta) = \frac{1}{2}\frac{\Delta\rho'}{\rho} + \frac{\Delta\alpha'}{\alpha} - \frac{4\beta^2}{\alpha^2}\frac{\Delta\beta'}{\beta}\sin^2(\theta) + \frac{\Delta\alpha'}{\alpha}\tan^2(\theta). \tag{1.18}$$

The relative changes in wave speed and density can be estimated by combinations of the changes in saturation and pressure. Landrø uses linear changes in saturation and quadratic changes in pressure. The relationships,

$$\frac{\Delta\alpha'}{\alpha} \approx k_\alpha\Delta S + l_\alpha\Delta P + m_\alpha\Delta P^2, \tag{1.19}$$

$$\frac{\Delta\beta'}{\beta} \approx k_\beta\Delta S + l_\beta\Delta P + m_\beta\Delta P^2, \text{ and} \tag{1.20}$$

$$\frac{\Delta\rho'}{\rho} \approx k_\rho\Delta S, \tag{1.21}$$

need to be estimated from data according to some regression technique. These relationships can be inserted into eq. 1.18

$$\begin{aligned}\Delta R \approx \frac{1}{2}&(k_\rho\Delta S + k_\alpha\Delta S + l_\alpha\Delta P + m_\alpha\Delta P^2)\\ &+\frac{1}{2}(k_\alpha\Delta S + l_\alpha\Delta P + m_\alpha\Delta P^2)\tan^2(\theta)\\ &-\frac{4\beta^2}{\alpha^2}(l_\beta\Delta P + m_\beta\Delta P^2)\sin^2(\theta).\end{aligned}$$

Generally, in these types of analysis, it is assumed that θ is sufficiently small, such that $\sin^2(\theta) \approx \tan^2(\theta)$. Then, each vertical transect of the reservoir can be imaged from multiple values of θ (all small) and we can perform a linear regression for the model $R = R_c + G\sin^2(\theta)$. Multiple surveys allows us to find,

$$\Delta R_c \approx \tfrac{1}{2}(k_\rho\Delta S + k_\alpha\Delta S + l_\alpha\Delta P + m_\alpha\Delta P^2) \tag{1.22}$$

$$\Delta G \approx \tfrac{1}{2}(k_\alpha\Delta S + l_\alpha\Delta P + m_\alpha\Delta P^2) \tag{1.23}$$

$$-\tfrac{4\beta^2}{\alpha^2}(l_\beta\Delta P + m_\beta\Delta P^2). \tag{1.24}$$

The above equations are solvable for ΔS and ΔP, using an estimate for the ratio, β/α.

In our implementation, we use a Trani's modified version of the Landrø method for inverting time-lapse seismic attributes. The modified version includes some of the higher order terms dropped in the above derivation and includes the change in ΔT_2. We relate the change in travel time to the fluid changes by

$$\begin{aligned} \Delta T_2 &= \frac{2D}{\alpha_2'} - \frac{2D}{\alpha_2} \\ &= \frac{2D}{\alpha'}\frac{\Delta\alpha'}{\alpha'}, \end{aligned}$$

and the relative change in α can be estimated by the fluid changes according to the relationships seen previously. The system that we use to relate the changes in the seismic attributes to the fluid changes is shown below. All of the constants come from pervious data. This is a system of three equations on two unknowns. We approximate a solution using a Gauss-Newton method.

$$\Delta R_0 \sim a\Delta S^2 + b\Delta S + c\Delta P^2 + d\Delta P \tag{1.25}$$

$$\Delta R_0 \sim e\Delta S^2 + k\Delta S + g\Delta P^2 + h\Delta P \tag{1.26}$$

$$\Delta T_2 \sim \frac{2D}{\alpha'}(i\Delta S^2 + j\Delta S + k\Delta P^2 + l\Delta P). \tag{1.27}$$

Chapter 2

Kalman Filtering Techniques

Kalman filtering is a powerful technique used in filtering and prediction problems where high uncertainty is a characteristic of both the modeling and observation of a dynamic system. The general technique is named after Rudolf E. Kalman, who presented it in his 1960 paper [8], and has been extended and modified over the last 50 years to allow for use on larger classes of problems. The traditional Kalman filter is applied to a linear system that can be described by

$$X_{k+1} = F(k)X_k + B(k)U_k + G(k)W_k \tag{2.1}$$

$$Z_k = M(k)X_k + V_k \tag{2.2}$$

where

- X_k denotes the state of the system,
- k is the time index,
- Z_k is the set of observables,
- U_k represents system input,
- F_k, B_k, M_k are matrices,
- W_k is a Gaussian noise with mean zero and covariance matrix Q, and
- V_k is a Gaussian noise with mean zero and covariance matrix R.

Equation 2.1 describes the propogation of the system state through time and equation 2.2 describes the relationship between the system and the available observations. W_k and V_k are taken as independent, indicating independence in the modeling uncertainty in eq. 2.1 and the observation uncertainty in eq. 2.2.

2.1 Kalman Filter

The Kalman filtering technique solves for the distribution of the state X_k conditioned on the measurements $Z_1, \ldots Z_l, l \leq k$. This conditional density function is Gaussian, because of the conditions imposed on the noise processes V and W. Therefore it is completely characterized by the mean and covariance matrix. Moreover, for a Gaussian distribution the mean is an optimal estimate of the state in the minimum variance sense and also in the maximum likelihood sense. Recursive equations to obtain the mean, $X(k|l)$, and covariance matrix, $P(k|l)$, of the probability density of the state X_k at time k conditioned on the history of the measurements $Z_1, \ldots, Z_l$ up to and including time l are formulated as follows:
Initial condition:

$$\begin{aligned} X(0|0) &= X_0, \\ P(0|0) &= P_0. \end{aligned}$$

Time update:

$$\begin{aligned} X(k|k-1) &= F(k)X(k-1|k-1), \\ P(k|k-1) &= F(k)P(k-1|k-1)F(k)^T + G(k)Q(k)G(k)^T. \end{aligned}$$

Measurement update:

$$\begin{aligned} X(k|k) &= X(k|k-1) + K(k)\left(Z_k - M(k)X(k|k-1)\right), \\ P(k|k) &= (I - K(k)M(k))P(k|k-1), \end{aligned}$$

where I is an identity matrix and $K(k)$ is the Kalman gain,

$$K(k) = P(k|k-1)M(k)^T \left((k)P(k|k-1)M(k)^T + R(k)\right)^{-1}. \tag{2.3}$$

The term

$$(Z_k - M(k)X(k|k-1)) =: I_k \tag{2.4}$$

is called the innovation of the filter and can be useful in evaluating the performance of a filter.

The direct use of the standard Kalman Filter algorithm for reservoir simulation is impossible for several reasons. Primarily, the model which moves the reservoir forward in time is non-linear, requiring the model matrix $F(k)$ be replaced by an operator F. Secondly, because the sytem is very large, the calculation of $F(k)P(k|k-1)F(k)^T$ is very expensive. The filtering technique shown in the next section is an extension of the traditional Kalman Filter shown above which has relaxed conditions.

2.2 Ensemble Kalman Filter

The Ensemble Kalman Filter (EnKF) was originally proposed by Evensen [4] and later altered by Burgers [1]. A recent paper by Evensen [5] provides a good summary of the theory and implementation of the method, which formed the basis for the description given in this section.

The EnKF is a recursive filter suitable for problems with a large number of variables, such as reservoir modeling, and it is relatively easy to implement. EnFK methods apply a Monte-Carlo method to estimate the forecast system state, based on an estimation of the probability density of the state estimate by a finite number, N, randomly generated system states $\psi = \{\psi_i\},\ i = 1, \ldots, N$. The forcast density is then adjusted, with the available observations, using Baye's Rule. The data then have a higher likelihood for the analyzed state than for the forecast state.

The EnKF algorithm is presented below, with a following explanation.
The initial ensemble of state vectors:

$$\psi_0 \sim N(X_0, P_0)$$

Forecast Step:

$$\begin{aligned} \psi^f_{k+1} &= \mathcal{M}(\psi_k), \\ \mathbf{P}^f_{k+1} &= E\left[(\psi^f - \bar{\psi^f})(\psi^f - \bar{\psi^f})^T\right] \end{aligned}$$

Analysis Step:

$$\mathbf{K}_{k+1} = \mathbf{P}^f_{k+1}\mathbf{H}^T(\mathbf{H}\mathbf{P}^f_{k+1}\mathbf{H}^T + \mathbf{R})^{-1} \tag{2.5}$$

$$\psi^a_{k+1} = \psi^f_{k+1} + \mathbf{K_{k+1}}(\mathbf{D}_{k+1} - \mathbf{H}\psi^f) \tag{2.6}$$

The sub-index k refers to time, $\mathcal{M}$ is the model operator, ψ^f is called the forecasted ensemble, P^f is the covariance of the forecasted ensemble, the bar operator indicates an average over the ensembles, $\mathbf{K}$ is the Kalman gain, the matrix H describes the relationship between the ensemble state and the measured quantites in the ensemble of measurements

$$\mathbf{D} = [d_1, \ldots, d_N],\ d_j = d + v_j,\ v_j \sim N(0, R).$$

The analyzed estimate of the ensemble state is represented by ψ^a and $\mathbf{W}$ and $\mathbf{V}$ are matricies where each column is a random vector with covariances $\mathbf{Q}$ and $\mathbf{R}$.

2.2.1 Measurement Formulation

One of the strengths of the EnKF is its ability to utilize a general operator, $\mathcal{M}$, for the forward predictions rather than a strictly linear dynamical system or a

linearized approximation. However, the method does require a linear relationship, H, between the observation matrix and ensemble, according to the formulation in eqns. 2.5 and 2.6. The creation and storage of H poses a challenge in practical applications of the EnKF for several reasons.

The matrix H is often very large and sparse. In a practical application the state vector can contain on the order of 10^9 members. The number of measurements will certainly be lower, for any given time step. The number of measurements available is typically on the order of 10^1 or 10^2, where each is either a direct measurement of a member of the state vector or linearly related to several members of the state. In such a case, H includes many zeros. If a measurement is a complicated function of many variables in the state vector then the matrix H can become quite difficult to construct. Likewise, if the measurements available change with time H must be reconstructed with every new measurement. It is desirable to use an operator $h(x)$ to define the state-measurement relationship according to

$$h(\psi_i) = \mathbf{H}\psi_i + f,$$

or

$$H\psi = [h(\psi_1) - f, \ldots, h(\psi_N) - f].$$

Thus, h predicts the measurements that would be provided by the product $\mathbf{H}\psi$ up to some error vector f.

Under a slightly different formation, the analysis step can be performed with h instead of $\mathbf{H}$. The description of this method that is included here is based on the work of Mandel [13]. The derivation of the observation matrix free kalman equations begins with a rewriting of the predicted ensemble covariance,

$$P = \frac{AA^T}{N-1} \tag{2.7}$$

where

$$A = (\psi^f - \bar{\psi}^f).$$

The substitution of eqn. 2.7 into eqns. 2.6 and 2.5, the analysis step can be rewritten

$$\psi^a = \psi^f + \frac{1}{N-1}A(HA)^T G^{-1}(D - H\psi^f)$$

with

$$HA = H\psi^f - \overline{H\psi^f}, \text{ and}$$
$$G = \frac{1}{N-1}HA(HA)^T + R.$$

If the analysis step is calculated using the above equations, it is only necessary to calculate $\mathbf{H}A$ once. Considering a single column of $\mathbf{H}A$,

$$\begin{aligned}
[\mathbf{H}A]_j &= \mathbf{H}\psi_j - \mathbf{H}\bar{\psi} \\
&= \mathbf{H}\psi_j - \mathbf{H}\frac{1}{N}\sum_{l=1}^{N}\psi_l \\
&= (\mathbf{H}\psi_j + f) - \frac{1}{N}\sum_{l=1}^{N}(\mathbf{H}\psi_l + f) \\
&= h(\psi_j) - \frac{1}{N}\sum_{l=1}^{N}h(\psi_l).
\end{aligned}$$

Thus, the single calculation of $h(\psi_i)$ for all ensemble members $i = 1, \ldots, N$ can be used to find $\mathbf{H}A$ and calculate the ensemble update without storage or construction of the matrix H. This method, using the operator h, is employed in the applications udertaken for this project, see chp. 3, in order to utilize the Modified Landro method for inverting seismic data.

2.2.2 Twin Experiments

In practice, an EnKF is supplied with instrumental measurement of the actual physical process and the true state of the world is easily defined. In a synthetic case, the definition of the true state of the world is less obvious. To give some order to a synthetic experiment with an EnKF, we perform a twin experiment. In such an experiment, an extra ensemble member is propogated forward in time by the same modeling software and this serves as the base for the synthetic measurements. The extra state vector is not included in any of the calculations of ensemble statistics. The influence of the true case on the ensemble is communicated through the measurements of the true state. Thus, as this report proceeds, when we refer to the true state or make calculations using the true value we are refering to the extra state vector. This allows us to measure the performance of the filter. Measures which require the true values would not be available in practice.

2.3 Filter Performance

Given the construction of the EnKF in the preceding section, the second part of this chapter focuses on measuring the performance of a filter after a number of analysis steps. Measurement of a filter's performance allows for comparison between filter configurations and can provide confidence in a filter's applicability. The measurement schemes are seperated into two groups. The first measures the filter's performance as parameter estimation tool. The second type measures the

reliability of the predictions made by the filter after an initial training period, which consists of a number of analysis steps.

Typically, uncertain parameters from the forward modeling scheme are included as variables in the state vector. These modeling parameters, as well as variables in the system state which do not vary significantly on the time scale of simulation, are the subject of parameter estimation measurement schemes. The theory states that as the analysis steps are performed, the estimated distribution for the time-constant variables becomes more representative of reality. In reservoir modeling, engineers are usually interested in the ability of the EnKF to estimate the geological properties of the reservoir. As previously described, quality information regarding the rock properties of the reservoir is spatially sparse and the EnKF is known to resolve some information about these parameters. Refinement of the estimates of the geological properties of the reservoir allow engineers to develop the reservoir by strategically adding producing and injecting wells or altering the controls on the existing infastructure.

The second type of performance measurement assesses the quality of the predictions made by the EnKF after time integration. From the predictive standpoint, the assimilation of data improves the filter's predictions of the system for advanced times by training the filter for the specific problem. Practically, it is important that a mature filtering scheme performs reliable predictions. In reservoir engineering, it is useful to be able to predict the fluid properties of the reservoir in the future. It is especially desireable to estimate the total production and injection rates for the entire field, but it is certainly important that the filter predict fluid motion within the reservoir.

This section introduces several measurements of performance and discusses how they are used to evaluate the schemes used in the experiments, see 4 and 5.

2.3.1 Graphical Method

The Graphical Method asks the engineer to compare plots of variables as estimated by the EnKF and the simulated true state to evaluate how well the ensemble represents the truth. This technique is typically used to evaluate the parameter estimates provided by the filter, especially regarding geological properties. The true reservoir state and the ensemble estimate are plotted in color over a horizontal map of the reservoir. Figure 2.1 is an example of the type of plots which are considered using this technique. The figure shows the porosity, ϕ, in a layer of the example reservoir for the true state of the system and three example estimates. The first, in the upper right-hand side, shows an estimate in which each grid point takes the average value of ϕ for the layer with a slight random perturbation. Graphically, this is not an impressive estimate. The plot shows none of the structure that is seen in the true case. The next two examples provide more structure.

It falls to the analyst to interpret which filter is providing a better estimate of ϕ. In such a case, the evaluation of an estimate or the comparison of estimates becomes subjective. The subjectivity of this method is it's weakness. Additionally, it can become very labor intensive on the part of the analyst to evaluate these plots for large or complicated reservoirs. In 3 dimensional reservoirs it is especially difficult as at least each geological layer should be checked, but it is better to check each vertical discretization layer. However, the graphical method does provide a good way to check that the filter is not estimating unreasonable values for porosity or permiability.

2.3.2 Error Measurements

In order to counter the subjectivity inherent in the graphical evaluation of filter performance it can be useful to measure the absolute error of the estimate provided by the EnKF with the true state. The root mean squared error is a common measurement for the distance between the estimate and the truth. For a vector variable, X,

$$RMS(\hat{X}) = \left[\frac{1}{\dim(X)} \sum_{\dim(X)} (\hat{X} - X_{true})^2 \right]^{-1/2}.$$

The RMS provides the analyst with a quantitative measurement of the filter performance. Like the graphical method, it is typically applied to measure a filter algorithm's ability to perform parameter estimation. The RMS values for the example porosity estimates in figure 2.1 are shown in the following table.

Truth	Appx. Mean
RMS = 0	RMS = 0.036
Est. 1	Est. 2
RMS = 0.025	$RMS = 0.036$

The results in the above table indicate that the estimate presented in the lower-left of the figure is the best. The RMS values of the other estimate reveal a weakness of the RMS as a measure of filter performance. The two estimates are very different in nature, but are rated as nearly equal under RMS. This is due to the fact that the RMS measurement collapses quite a bit of information into a single quantity. Additionally, the RMS measurement is unit-dependent. This can lead to confusion and difficulties when analyzing the parameter estimation for variables with different units valued on different scales.

It is generally advisable to use a combination of the graphical method and RMS scores to evaluate the performance of different filters. In this case, the combination of the two would rate the filter resulting in the estimate on the lower-left as the most suitable. This result is in accord with the construction

of the example estimates. The lower-left estimate is a smoothed version of the truth. The estimate on the right is a 180 degree rotation of the other.

2.3.3 Graphical Prediction Validation

Graphical prediction validation is a subjective method for measuring the viability of the forecasts given by an EnKF. A prediction validation plot typically shows an independent measurement of a variable along with the ensemble estimates for that same quantity. Often this is done for several measurements. Plotting the two together shows where the measurement, or the true value if it is available, falls within the ensemble's predictions. Figure 2.2 shows such a plot for four producers from the test reservoir introduced in chapter 3. Taking the last time available, we see that the predictions for the two producers on the top seem acceptable. The ensemble estimates are scattered around the true value with a slight bias. In the other cases the ensemble is biased or too small and misses the true value. The estimate in the southwest corner seems particularly bad from a graphical standpoint. In this report, we plot all of the ensemble estimates along with the realiztions. In other studies, it is common to see the ensemble mean plotted with the realization and a pair of bounds which are given by the ensemble variance or extreme values.

2.3.4 Reliability Measurement

The graphical method of prediction validation can be formalized statistically to provide a quantification of the prediction reliability, given a set of validation measurements. The statistical definition of prediction reliability is based on the distributions governing the system variables. Thus, a filter's predictions are considered perfectly reliable if,

$$\hat{F}(F) = U[0,1] \; \forall \; F,$$

where F is the distribution of the variables to be validated and $\hat{F}$ is the empirical distribution from the ensemble. The equality will be satisfied when the empirical distribution given by the ensemble is equal to the true distribution of the system variable. If the relationship is approximate, instead of equal, then the predictions can still be considered reliable as equality can usually only be approached at high sample sizes. This is naturally a difficult thing to measure, but this section introduces a statistic, the reduced centered random variable (RCRV), used in evaluating the reliability of weather prediction systems [2], that can be used to gauge the degree to which the empirical distribution matches the truth.

Consider a group of m varification variables for an EnKF, $\{\xi_l, l = 1 \dots m\}$. This is a set of measurements of variables from the state vector of the modeled

system. There is then an associated group of variables,

$$Z_l = \frac{\xi_l - \mu_l}{\sigma_l},$$

where μ_l and σ_l^2 are the ensemble mean and variance for the measured variable. The set $\{Z_l, l = 1, \ldots m\}$ represents samples of the RCRV, Z. Given that the samples of the verification variables are performed independently, we expect to see, for a proper EnKF configuration,

$$E[Z] = 0$$

and

$$\mathrm{Var}(Z) = 1.$$

If the sample mean of Z differs from 0, then there is some bias in the distribution of the ensemble. A positive expectation of Z indicates that the ensemble is underestimating the verification variables and a negative expectation indicates that the ensemble prediction is too high. If the sample variance of Z is lower (higher) than 1, then the empirical distribution represented by the ensemble is too large (small) with respect to the true distribution.

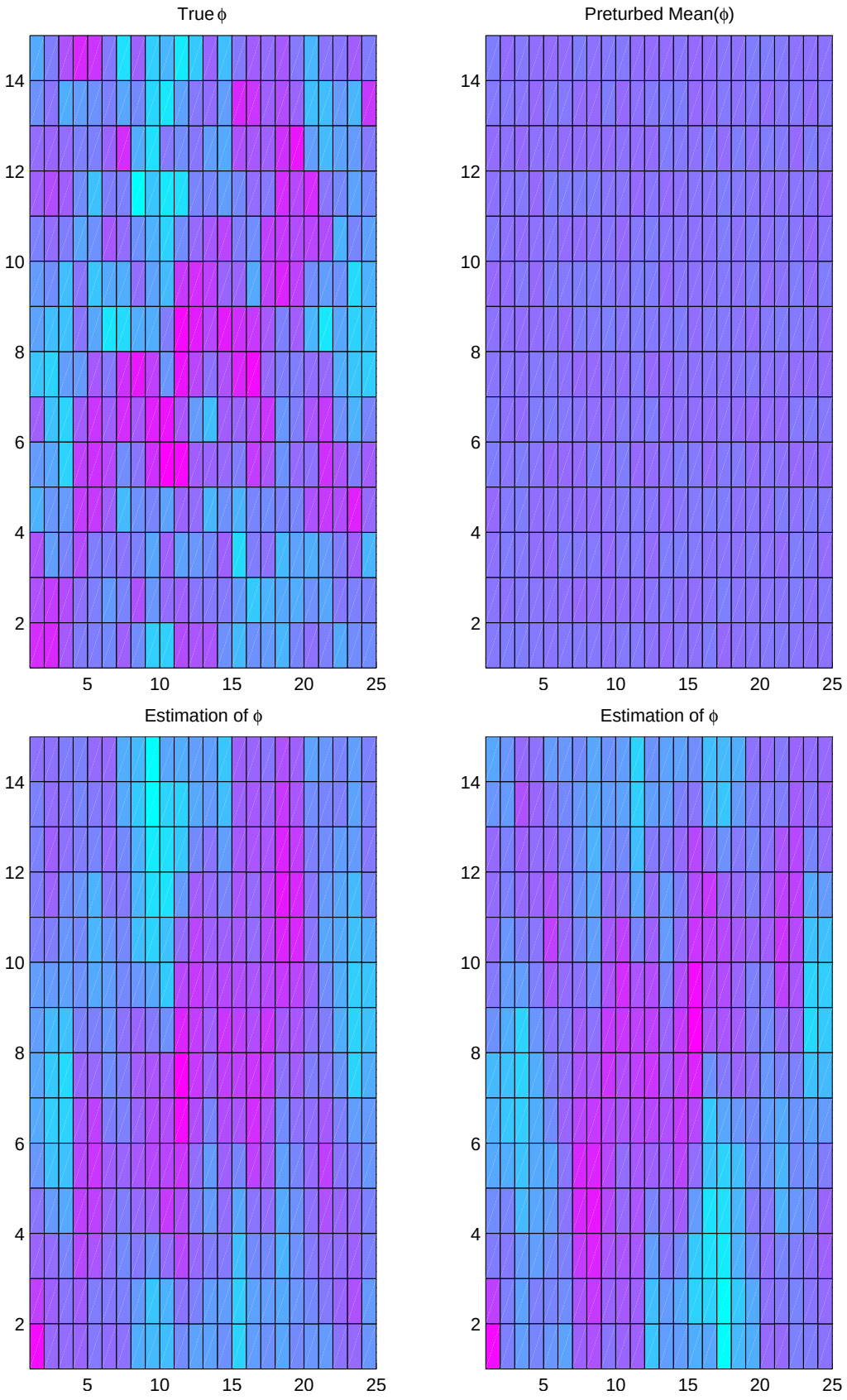

Figure 2.1: This figure shows ϕ for 1 layer of an experimental reservoir. The frame in the upper left is the truth. An almost homogenous realization is plotted in the upper right. Two possilbe ensemble means for these variables are shown on the bottom row.

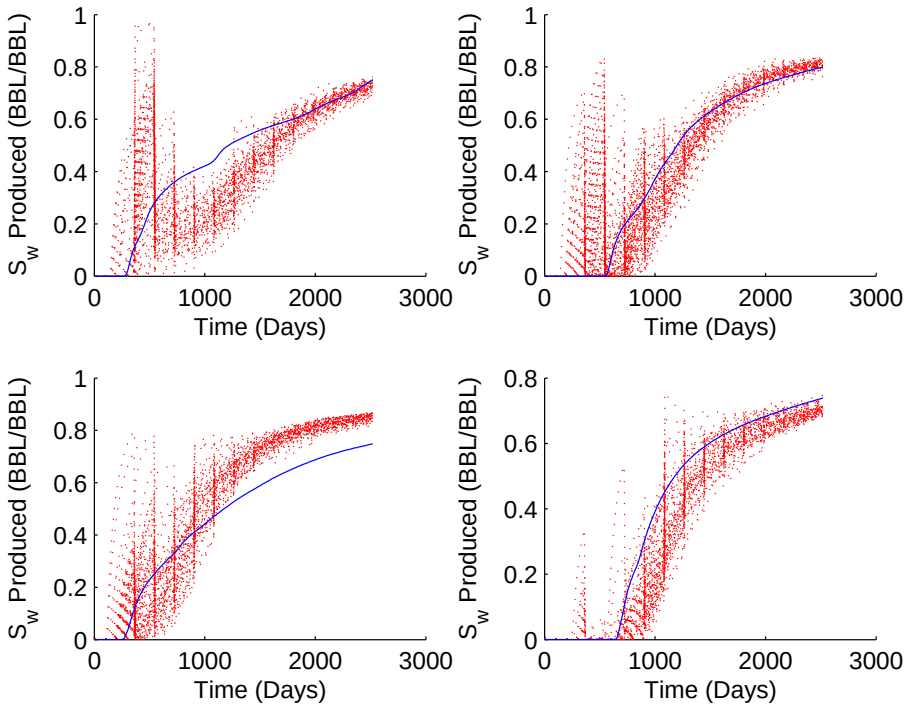

Figure 2.2: Watercut at 4 producers in the test reservoir. The true water cut is plotted in blue, the ensemble estimates are shown in red.

Chapter 3

Experimental Design

In order to be able to compare the efficacy of different data assimilation schemes, we developed a repeatable experiment that offered the ability to include the two types of seismic data, time-lapse and forward modeling, in different simulations of the same reservoir. The selected model is a three-dimensional oil/water reservoir. The production of the reservoir was simulated using the commercial package MoReS, from Royal Dutch Shell. The EnKF analysis was done using the CLOREM package, which was also developed by Shell. In the appendix for this chapter, we show the results of an EnKF configuration which assimilates only production data. That case is presented as a point of reference for the results in chapters 4 and 5.

3.1 The Box Model

We tested the different EnKF schemes on the so-called Box Model. A two-phase, oil/water, reservoir which we represent in three dimensions on a cartesian grid with dimensions 25 x 15 x 11. The gridblocks represent volumes of dimension 30 x 30 x 10 m, giving the reservoir a total horizontal dimension of 750 x 450 m. The depth of the reservoir, D, is 110 m. In field units, which is the unit system used in the simulations, the reservoir has the approximate dimensions 2460 x 1476 x 360 ft. Figure 3.1 shows a 3-D picture of the reservoir and colors the cells according to porosity.

The reservoir features 5 wells, 4 producers and 1 injector. Figure 3.2 shows the locations of the injector and producers. We attempt to mantain the orientation of the horizontal map seen in figure 3.2 in following maps of the reservoir. The injector was placed near the horizontal center of the reservoir and the producers were placed near the corners. The local environment for each well is summarized

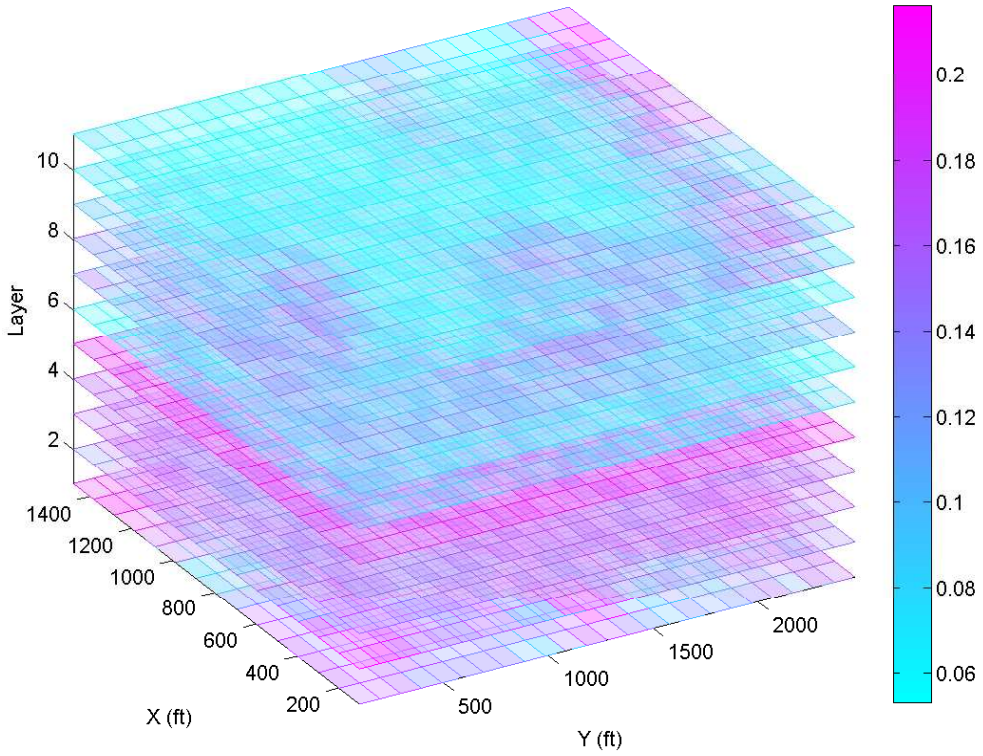

Figure 3.1: The true porosity, ϕ, for the reservoir.

in the following:

- *Injector 1*: Located near the horizontal center of the reservoir, grid point (9,15). The neighborhood of the injector is characterized by an average value of porosity and low permeability.

- *Producer 1*: Located in the lower left of the aerial maps of the reservoir (2,2). The surrounding geological properties display low-mid permeability and high-mid porosity.

- *Producer 2*: Located in the upper right of the aerial maps of the reservoir (13,3). The surrounding geological properties display medium porosity and low permeability.

- *Producer 3*: Located in the upper right of the aerial maps of the reservoir (2,21). This producer is in a region of low-mid permeability and high-mid

porosity, similar to producer 1.

- *Producer 4*: Located in the upper right of the aerial maps of the reservoir (14,24). This producer is in a region of low permeability and high porosity.

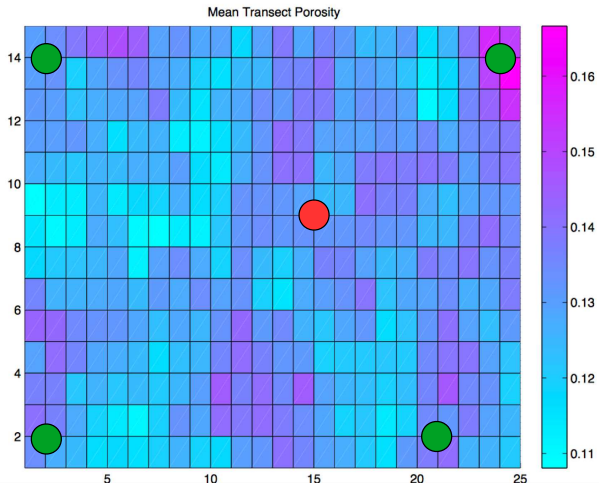

Figure 3.2: This figure shows the vertical mean porosity of the reservoir, from an aerial perspective. The location of the injecting well is denoted by the red dot. Likewise, the green dots inidicate the locations of the 4 producers.

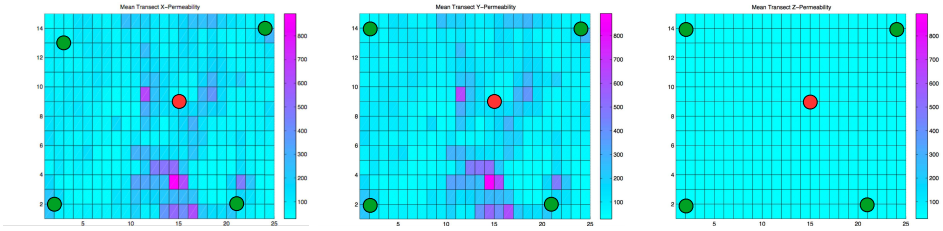

Figure 3.3: The X,Y, and Z-permeabilities, averaged in the vertical direction. The coloring is uniform amidst the plots.

The producers are constrained during simulation to a constant production rate of about 1700 barrels per day (BBL/DAY). Figure 3.4 shows the production rate curves for one experiment. In the experiments, maximum and minimum bottom hole pressures are defined, but these conditions are never imposed over the controlled rate.

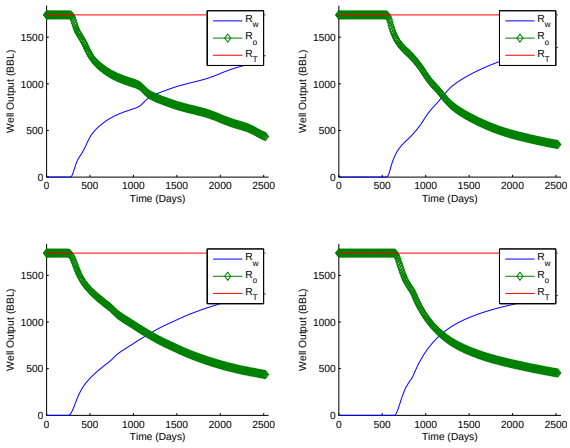

Figure 3.4: The oil, water and total production rates for the four producers for 1 simulation of the reservoir.

3.2 Ensemble Kalman Filter Details

The possible configurations for an EnKF are numerous for any application. It falls to the engineer to decide which parameters should be modeled with uncertainty and which can be considered known, what kind of measurements can be used, the timing of the analysis steps and the use of transformations to maintain physical results. In this section we show some of the parameters we used our experiments.

The data that is considered uncertain by our EnKF is listed below. S_w and S_o were included in the state vector using the Reynold's transformation [18], which constrains the saturations between 0 and 1 using log transformations [9].

- Elements in the state vector: $\log(k_{x,y,z})$, ϕ, P, S_w, S_o

- Measurements Available: {bottom hold pressure (BHP), water rate (R_w), oil rate(R_o)} and $\{\Delta S_w, \Delta P\}$ or $\{R_0, G, T_2\}$

Our state vector has size $O(10^4)$ members and we use ensembles with $N = 80$ members. The production measurements are known for each of the wells, giving us 15 well measurements. There are 375 measurements for each of the seismic attributes and the inverted fluid changes, one for each horizontal gridpoint. Using standard estimates for the uncertainty in the measurements, we define the

measurement uncertainty covariance matricies as

$$R_{\text{wells}} = \begin{pmatrix} 5625I(5) & 0 \\ 0 & 900I(10) \end{pmatrix},$$

$$R_{\text{IS}} = \begin{pmatrix} 0.0036I(375) & 0 \\ 0 & 900I(375) \end{pmatrix} \text{ and}$$

$$R_{\text{AVO}} = \begin{pmatrix} 0.01I(375) & 0 & 0 \\ 0 & 0.09I(375) & 0 \\ 0 & 0 & 5x10^{-13}I(375) \end{pmatrix}.$$

$I(m)$ represents the identity matrix with main diagnol length m. The use of the identity matrix imposes the assumption that the uncertainty in the measurements is independent. It is known that the uncertainty in seismic data cannot be assumed to be fully independent, see [11], but the dependence is small and widely modeled as zero.

Chapter 4

Inverted Seismic Assimilation

This chapter includes results and analysis for configurations of the EnKF assimilating the estimates obtained as Gauss-Newton approximations of the system of equations beginning at 1.25. Results for simulations testing parameter estimation and prediction are shown. As a point of camparison, the appendix for chapter 3 includes results for a EnKF that assimilated only production data.

4.1 Initial Experiment

The initial experiments on configurations featuring inverted time-lapse data showed suspicious results. Investigation showed that there was a mismatch in the measurements calculated by the modified Landrø method and the true values. Figure 4.1 shows the measurements supplied by the modified Landrø equations for ΔS_w at 540 days in a certain simulation. The true change in saturation is shown adjacent to the measurement. At this time, the seismic measurement overestimated the truth. A similar plot of pressure change is shown in the appendix, B.2. It shows that the pressure measurement was also too high. The source of this bias was assumed to come from inaccuracies in the geological parameters from the modified Landrø calculations, which were taken from experimental values and not developed for the Box Model geology specifically.

We ran the EnKF, which produced the results shown in figure 4.1 for 2556 days, calculating the error in the ΔS and ΔP measurements given by the modified Landrø method. We discovered an increasing relationship between the mean error, over the 375 points, in ΔS and ΔP and the time of the second seismic survey. This relationship is illustrated in figure 4.2, along with the relationship between the measured ΔS at each update time and the mean error in the inverted measurements. The appendix contains the plots for ΔP. These plots show a clear positive bias in the measurements supplied by the seismic inversion, which

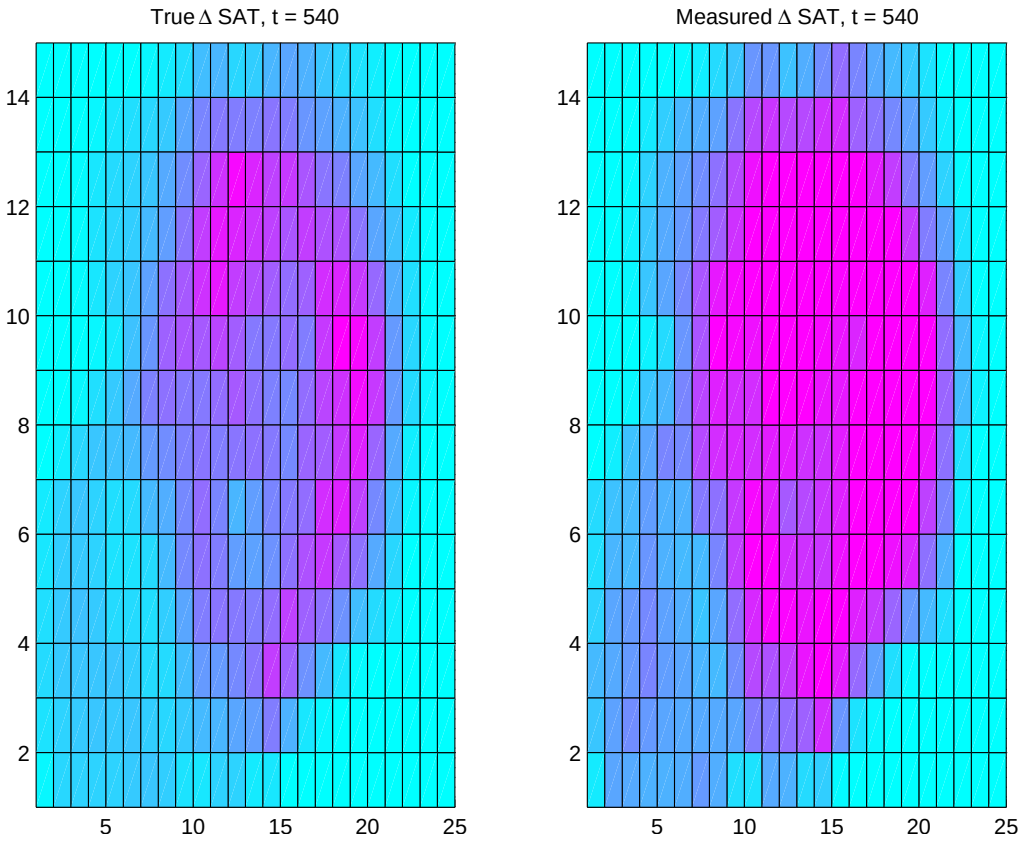

Figure 4.1: The measured and exact change in water saturation between 180 and 540 days in a realization of the true model. The color scheme is the same for both frames.

we chose to estimate using a best linear fit. The fit resulted in the following estimations for the error, or bias, in the modified Landrø inversion, which we removed in subsequent experiments.

$$\epsilon_{\Delta S_w} = 0.2726\hat{\Delta S}_w + 0.0087$$
$$\epsilon_{\Delta P} = 0.9580\hat{\Delta P} + 1.9207$$

The bias correcting factor shown above was not included in the update at the first seismic survey, because it is known that the change in ΔS and ΔP is zero at that point. We took the bias into account for all later measurements, which were truely time-lapse measurements.

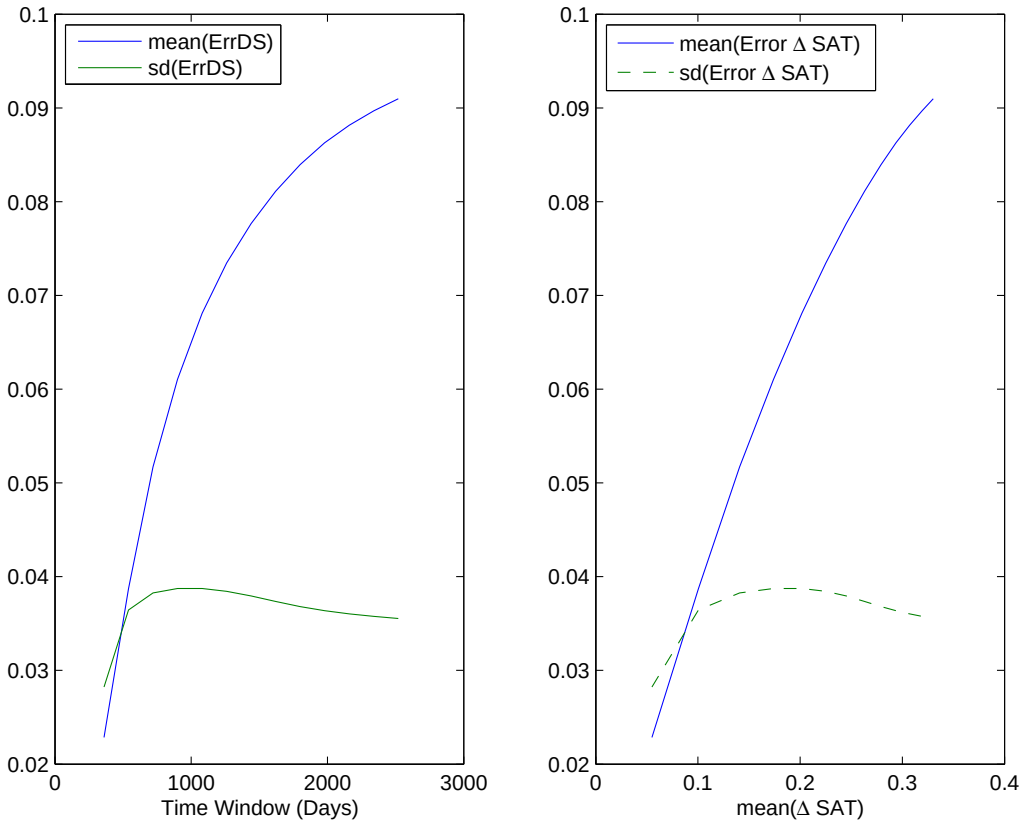

Figure 4.2: The mean error in the ΔS_w measurement is plotted against both the time of the second seismic survey used in the time-lapse inversion and the mean, inverted, ΔS values for the associated time windows.

4.2 Bias Removal

We performed an EnKF simlation of the box model reservoir, removing the bias we observed in the initial experiments. As in the initial experiment, the EnKF simulated the production of the reservoir for 2556 days. At every 180 days of simulation, we simulated a seismic survey of the reservoir and assimilated modified Landrø data with the bias removed. Figure 4.3 shows the ensemble and true water cut at the 4 producers. The curves show mixed performance of the filter on the prediction of water breakthrough. The ensemble predictions for producers 1 and 4 predicted water breakthrough late, while the others predicted it early.

The assimilation of the time-lapse seismic data improved the parameter estimation properties of the EnKF, over the assimilation of only production data. Figure 4.4 shows the time evolution of the X-permeability in layer 7 of the reservoir. Plots of the other geological parameters, similar to those in figure 4.4, are shown in the appendix. Unlike in the production-only EnKF, we saw that the EnKF revealed the structure of the field when the time-lapse seismic data was

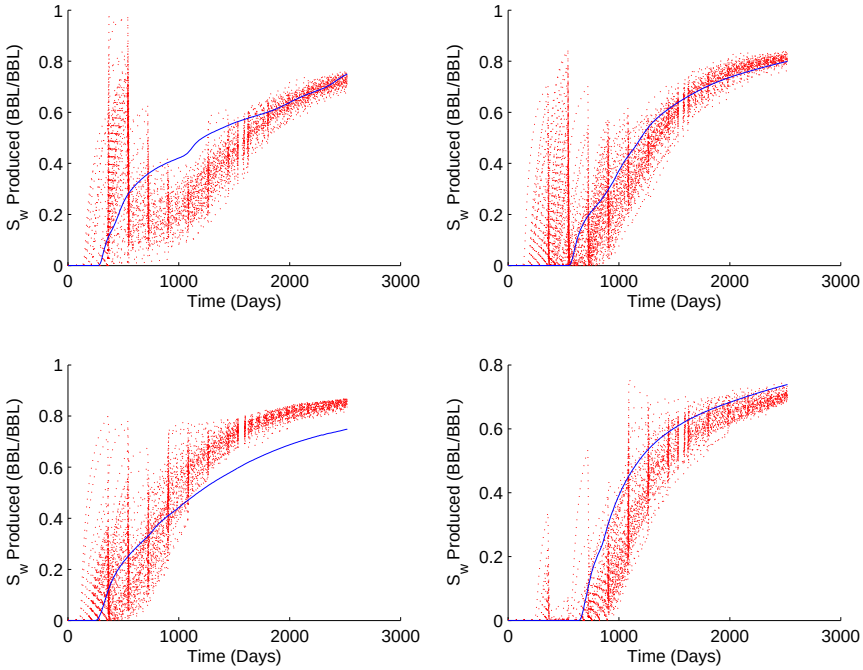

Figure 4.3: The water cut curves for each of the producers, under the assimilation of bias-removed modified Landrø data. The ensemble is shown in red and the truth in blue.

used in the analysis step. The evolving estimate of the layer 7 X-permeability, k_x, starts to approximate the true structure. The area of high permeability near producer 2 (upper left) is resolved, especially in the analyzed version at 540 days, which has experienced 2 inverted seismic updates. As the filter moved the ensemble forward in time and assimilates new data, we observed overshooting of the true values of k_x. The final estimate shows a large region (approximately 50 grid blocks) of high permeability near the injector, located in the middle of the map.

The global RMSE in the ensemble estimates of k_x, shown in 4.5, shows the overshooting seen in the permeability maps. After the update at 1020 days, the RMSE of the estimate increases with subsequent updates. We observed the same overshooting behavior in the other geological parameters with this experiment, see the appendix for the corresponding plots.

Moving to the predictive tests for this configuration of the EnKF, we measured the reliability of the cumulative oil production forecasts graphically and using RCRV statistics. Figure 4.6 shows the cumulative oil output for the 4 producers along with the value from the truth model. It shows that the ensemble underestimates the cumulative oil production in the 3rd producing well. The underestimation is in accord with the overestimation of S_w at this producer seen

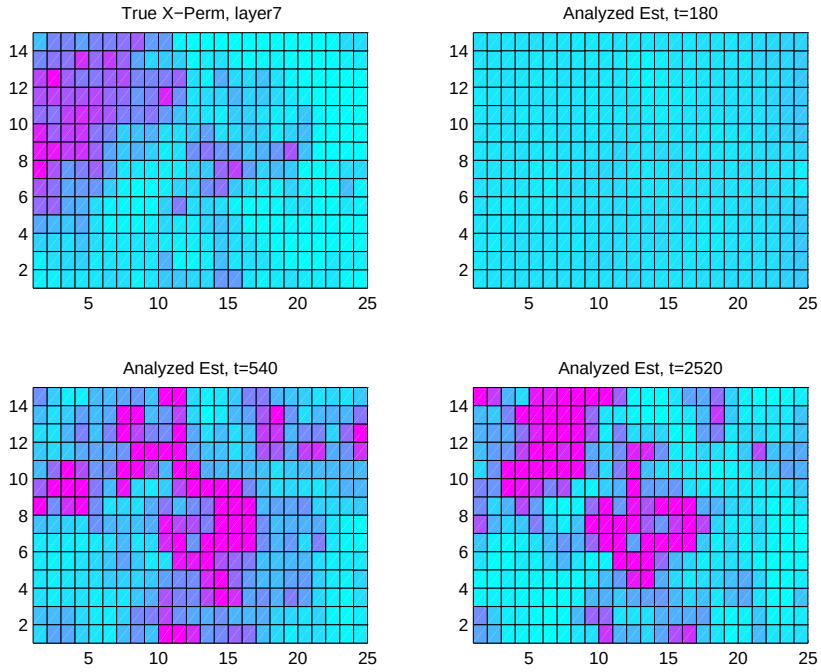

Figure 4.4: The X-permeability in layer 4 of the box model, truth (upper left) and three ensemble estimates, under assimilation of bias-removed modified Landrø data.

in figure 4.3, given the controlled total fluid rate. The other producers show a slight overestimation of the cumulative oil production.

Though the estimates of cumulative oil production mentioned above are biases, the overall agreement of the ensemble with the truth is satisfactory. The true value is typically in the body of the ensemble, without the assimilation of production data. This reliability is confirmed by the RCRV statistics shown in figure 4.7. To attain the Z corresponding to the pictured statistics, we measured the cumulative oil output for the producers at 30 day time intervals starting at a certain time and ending at the end of the simulation. The construction of Z explains the instability of the means for high starting times, plotted in the upper window of the figure, as there are fewer measurements available to average. The estimates for producer 2 were the most reliable, but the mean and variance of Z for producers 1 and 3 are also reasonable. The variance of Z for the 4th producer is much higher than the others indicating that the ensemble is too small.

4.3 Predictive Experiments

The overshooting in the parameter estimation we saw in the EnKF configuration in the previous section led us to consider a difference updating scheme, with an

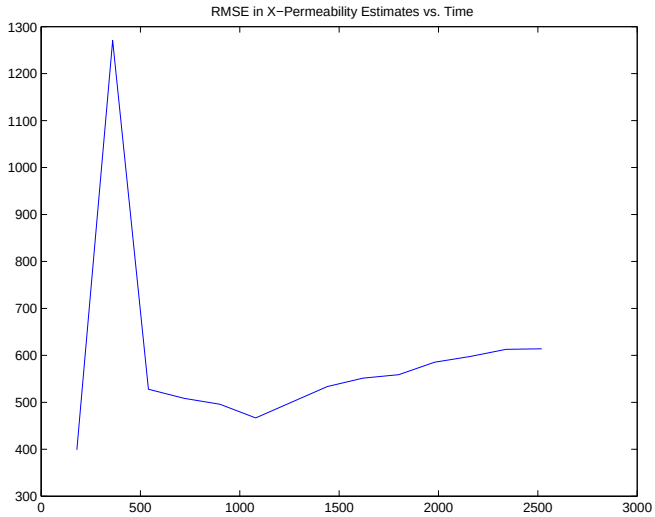

Figure 4.5: The RMSE for the post-analysis X-permeability estimates after the assimilation of modified Landrø data.

otherwise identical filter setup. Instead of assimilating the bias-adjusted modified Landrø data every 180 days until the end of simulation, we stopped regularly assimilating data after the update at 1260 days and performed a final assimilation at the end of simulation (2551 days). There were several motivations for experimenting with this update timing. First, we observed an increasing RMSE in the geological parameter estimates after the first few assimilations of time-lapse data. Additionally, we sought to test the predictive power of the ensemble after a long period of uninterrupted simulation. From a practical standpoint, it is also desirable to see if the benefits of inverted seismic assimilation can be acheived with fewer surveys. In practice, it is very expensive to perform a seismic survey and field operators would like to maximize their returns with respect to the number of surveys performed on a given field.

The filter configuration with a break in seismic assimilation for the second half of the simulation time yielded results quite similar to those of the configuration with regular seismic surveys. Figure 4.8 shows the evolving ensemble estimate of k_x for layer 7 of the reservoir. Again, the EnKF picks up much of the reservoir structure in the estimations, but suffers from overshooting. Comparing the estimates in this figure to the results from the filter configuration in section 4.2, we see a similar structure to the estimate at 540 days, but a higher degree of overshooting at 1260 days in the current experiment than we saw in the final estimate from the previous configuration.

The RMSE for the estimates of k_x given by the predictive configuration is shown in figure 4.9. There, we see that the RMSE again increases after the first couple assimilations of time-lapse seismic data. In this case, the final assimilation

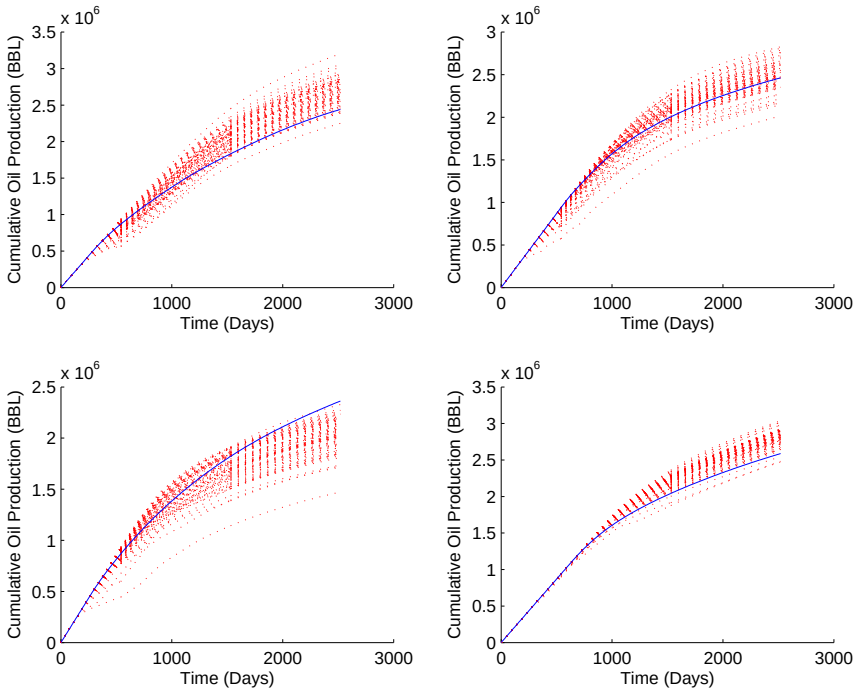

Figure 4.6: The cumulative oil curves for each of the producers, under the assimilation of bias-removed modified Landrø data. The ensemble values are shown in red and the true value is shown in blue.

at the end of simulation exhibits a very high RMSE, indicating that the final analysis step causes a large overshooting phenomena. The appendix shows the RMSE plots for the other geological parameters and each shows a similar final increase. Porosity is unique in that the RMSE of the estimate in increasing throughout the EnKF simulation, something which we observed in all of the EnKF schemes. The final increase in the RMSE of the parameter estimates, combined with the high level of overshooting after the analysis step at 1020 days led us to believe that the regular assimilation of time-lapse seismic is preferable from a parameter estimation standpoint.

While the parameter estimates given by this configuration of the EnKF were less desireable than those generated with regular assimilation of data, the reliability measures seemed highlight the advantage of assimilating less modified Landrø data. Figure 4.10 shows the produced watercut at the 4 producers. Again, the ensemble underestimates the time of water breakthrough at producer 3, but the other predictions appear to be acceptable. The ensemble mean might overestimate the time of water breakthrough at producers 2 and 4, but the true S_w curves are well within the bounds of the ensemble throughout the simulation. Figure 4.11 shows improved predictive ability over the results from the more data-laden configuration. The ensemble contains the true value of the cumulative oil pro-

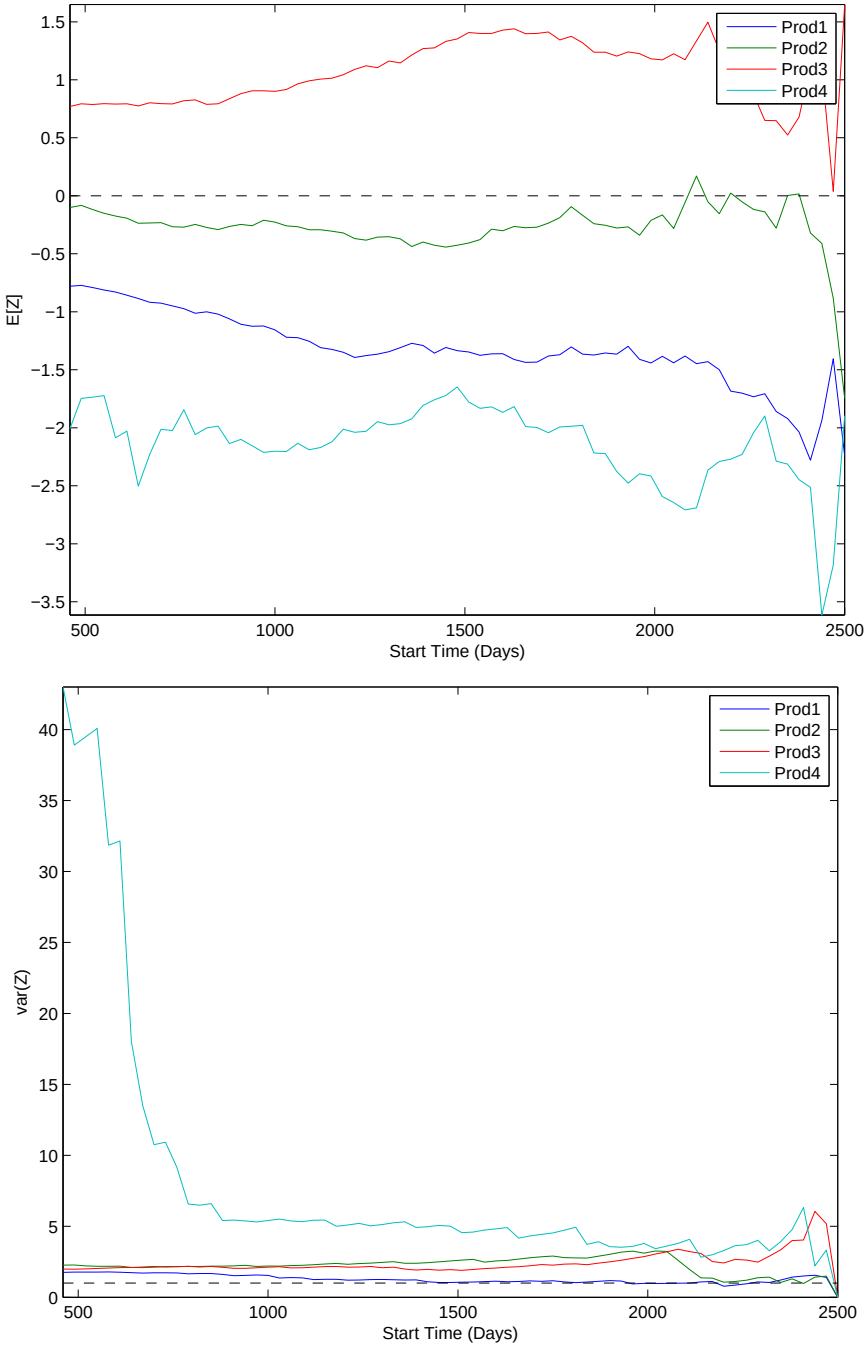

Figure 4.7: The statistics of the RCRV, calculated for cumulative oil production at all of the producers, under the assimilation of bias-removed modified Landrø data.

duction for each of the producers throughout the simulation. It still holds that the ensemble overestimates the oil production at the 4th producer.

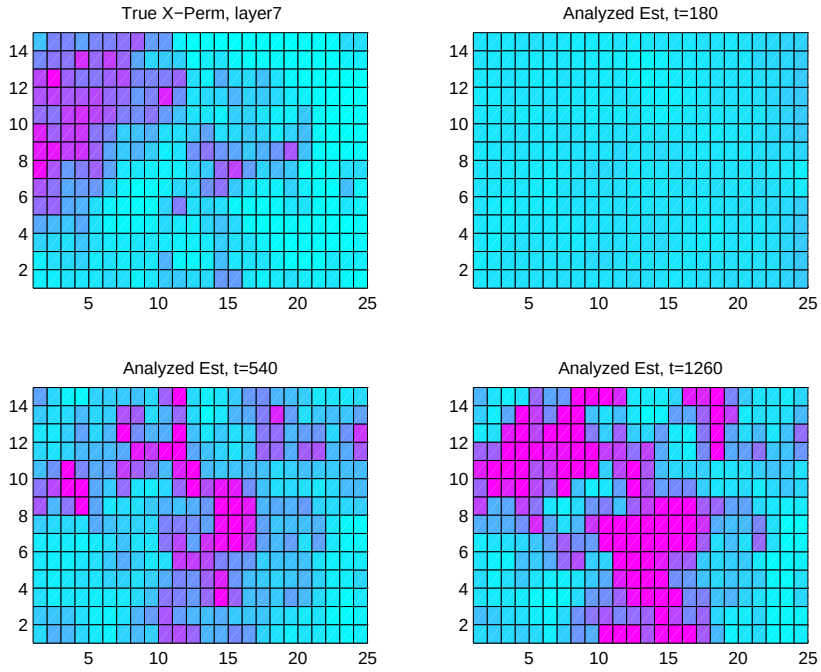

Figure 4.8: The X-permeability in layer 7 of the box model, truth (upper left) and three ensemble estimates, under assimilation of bias-removed modified Landrø data up to 1260 days.

The statistics of the RCRV for cumulative oil production, sampled at 30 day intervals with 10% error, are included in figure 4.12. By these measures, the EnKF under this configuration performs quite well. While there are biases in the means of Z for producers 3 and 4, the means for producers 1 and 2 are close to zero. Likewise, the variance of Z for producers 1 and 2 are close to 1 as is the variance of Z for producer 3. Producer 4 still exhibits a high variance of Z.

4.4 Summary

This chapter exposed typical results for EnKF configurations which assimilated modified Landrø data, taken with a time-lapse intervals of 180 days. We observed that the assimilation of time-lapse data improved the parameter estimation of the EnKF over the production-only base case. The improvement was seen in both the graphical estimates and the RMSE of the geological parameters. This result indicates that there is a strong covariance between the geological parameters, particularly permeability, and the changes in saturation and pressure after production. This is a nice result, as previously permeability has been a difficult variable for EnKFs to estimate. A caveat to this result is that the reservoir under consideration was fairly homogeneous and the method might have trou-

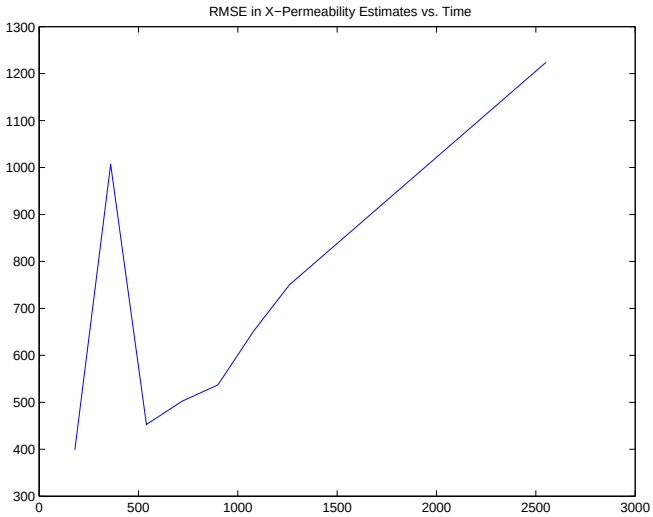

Figure 4.9: The RMSE for the post-analysis X-permeability estimates after the assimilation of modified Landrø data up to 1260 days with a final assimilation near the end of simulation.

ble with a more segmented reservoir. This potential weakness is especially true for reservoirs in which the geology changes significantly in the vertical direction. From a predictive standpoint, the forecasts made after an initial training with inverted time-lapse data seemed to be fine. Though there was some bias in the forecasts, the ensemble had an appropriate spread, and the bias could be quickly removed with the added assimilation of production data from several points in time.

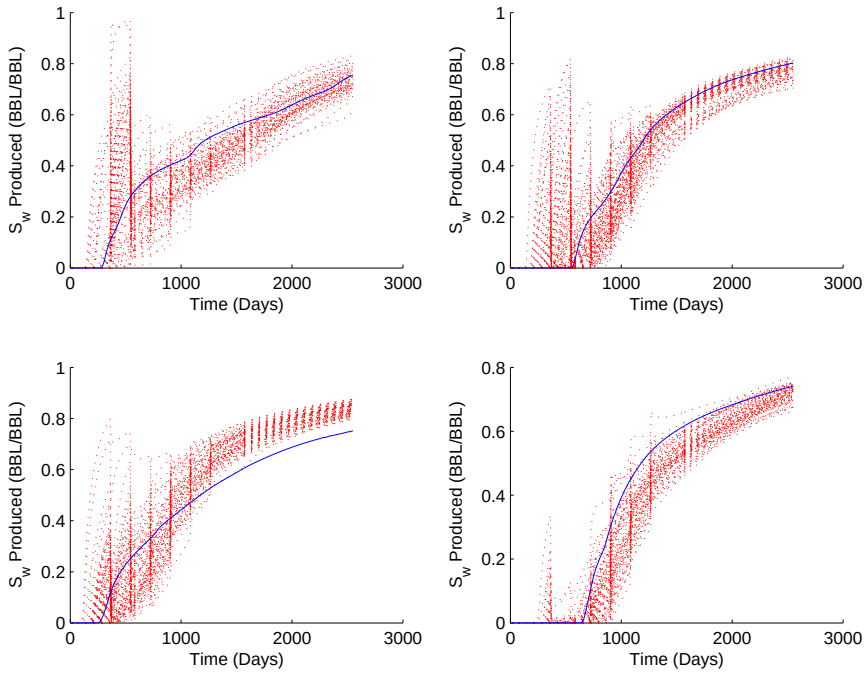

Figure 4.10: The water cut curves for each of the producers, under the assimilation of bias-removed modified Landrø data up to 1260 days. The ensemble is shown in red and the truth in blue.

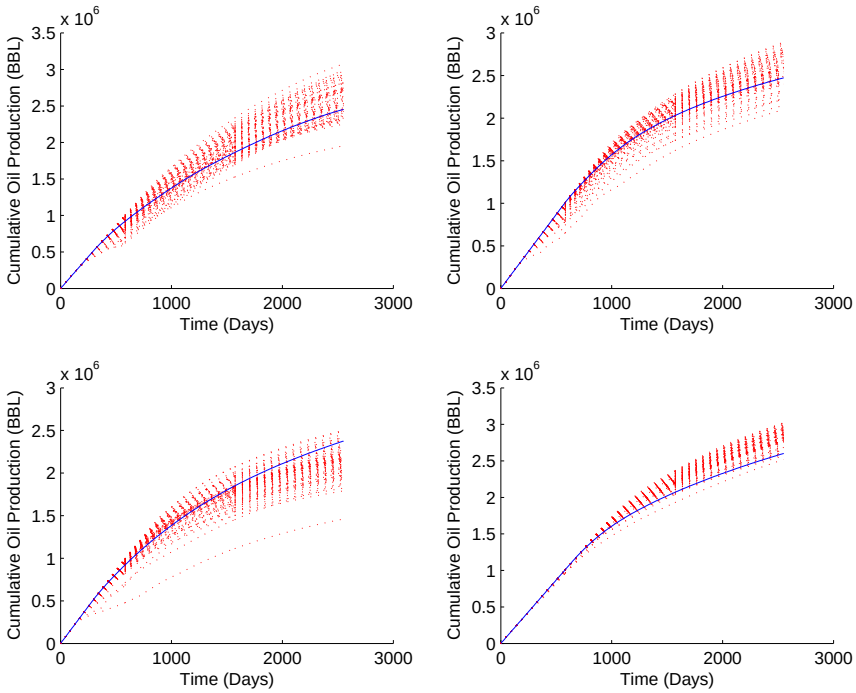

Figure 4.11: The cumulative oil curves for each of the producers, under the assimilation of bias-removed modified Landrø data up to 1260 days. The ensemble values are shown in red and the true value is shown in blue.

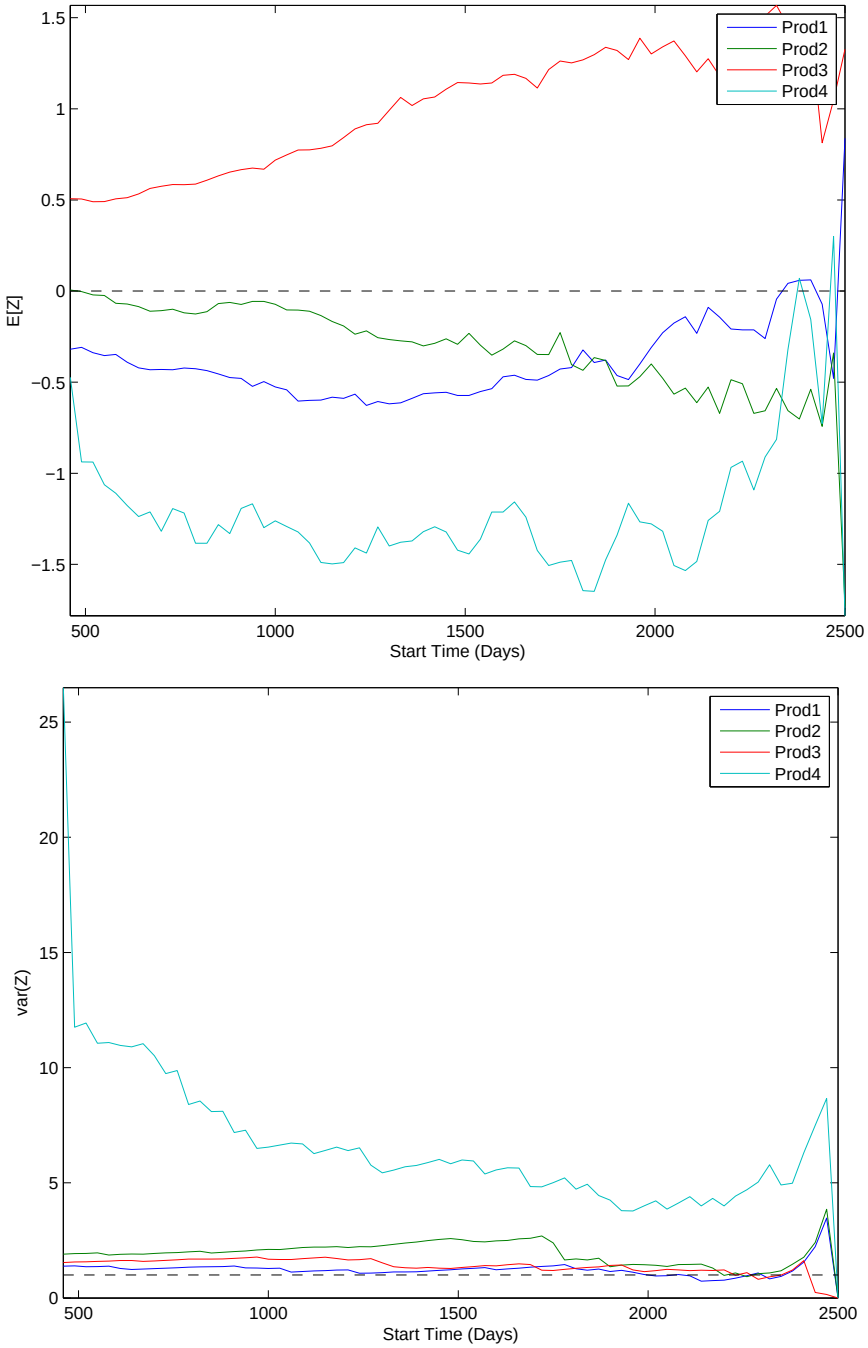

Figure 4.12: The statistics of the RCRV, calculated for cumulative oil production at all of the producers, under the assimilation of bias-removed modified Landrø data up to 1260 days.

Chapter 5

Seismic Attribute Assimilation

In this chapter we present results from an experiment where the seismic attributes introduced in section 1.3.1: G, R_0 and T_2, are assimilated by an EnKF simulating the test reservoir. Recall that these attributes are calculated at each assimilation time and are not considered time-lapse data. The configuration for the EnKF in this section is similar to those featured in the preceeding chapter. The reservoir is simulated for 2556 days, 7 years, and seismic data is assimilated and an analyzed state is calculated every 180 days of simulation. The data is generated by the rock physics framework and the Zoeppritz equation. An example of the measurements provided by the measurement operator are shown in figure 5.1. The figure shows a clear relationship between the seismic parameters and the average porosity.

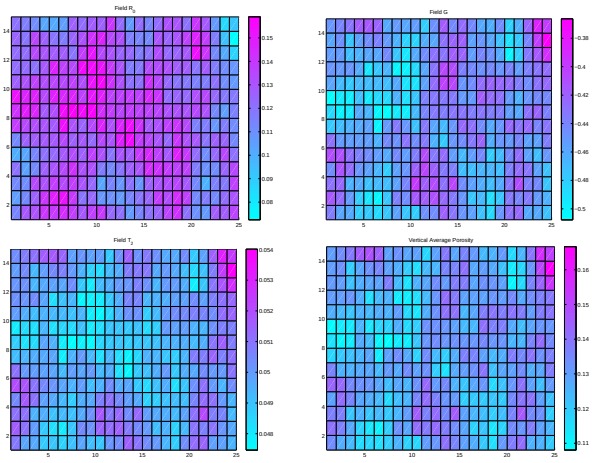

Figure 5.1: The Seismic Attributes of the reservoir, calculated for the analysis step at 540 days, and the vertically averaged true porosity. A correlation between mean porosity and the Seismic attributes is evident.

The EnKF configuration, with the direct assimilation of the seismic parameters, performed poorly by both methods of parameter estimation. From the graphical perspective, figure 5.2 shows the evolving X-permeability estimate for the 7th layer. It is clear from the figure that the estimate misrepresents the structure of k_x in this layer and gets worse through time. Unlike in the configurations that assimilated inverted seismic data, the results for this simulation do not exhibit overshooting. Instead the estimates diverge from the true values. This divergence is evident in the RMSE plot for X-permeability seen in figure 5.3. The values of the RMSE are high and increase quickly with the assimilation of more seismic data. From both perspectives, graphical and quantitative, this EnKF configuration did not effectively estimate the geological parameters of the reservoir. The appendix includes example plots for porosity and the other permeabilities, but none looks satisfactory. This is an unexpected result for porosity, as we saw such a clear relationship between the seismic parameters and ϕ maps in figure 5.1.

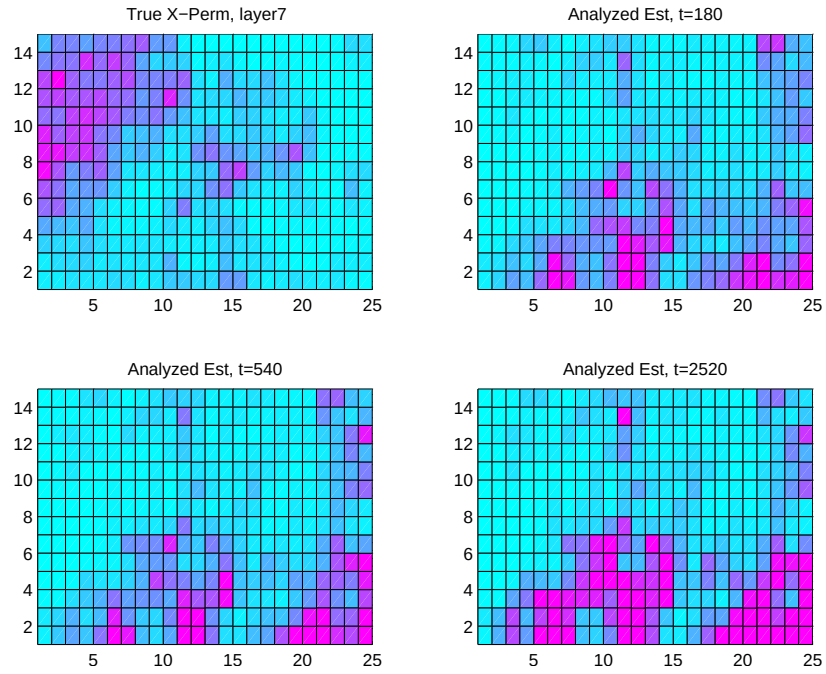

Figure 5.2: The X-permiability in layer 4 of the box model, truth (upper left) and three ensemble estimates, under assimilation of seismic attributes.

Given that the assimilation of seismic data does not provide reasonable parameter estimates, we still consider the predictive ability of the method. Figures 5.4 and 5.5 show the ensemble estimates for the watercut and cumulative oil production at the 4 producing wells. These plots show mixed results and some strange behavior in the ensemble. At producers 2 and 3, the ensemble overesti-

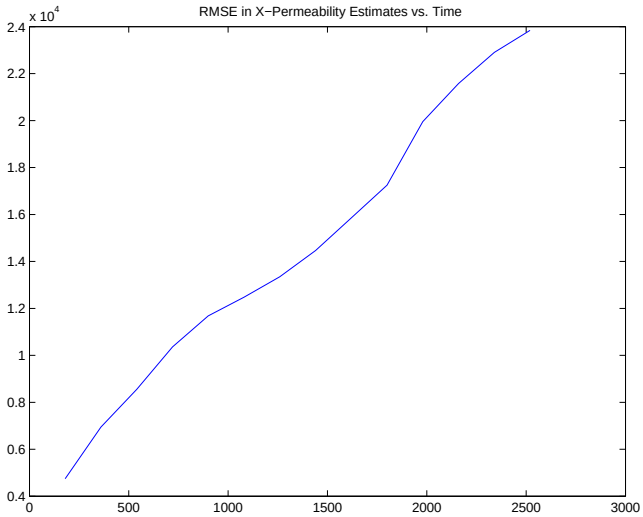

Figure 5.3: The RMSE for the post-analysis X-permiability estimates after the assimilation of seismic parameters.

mates S_w and underestimates the cumulative production. The opposite holds for producers 1 and 4, as can be seen in the statistics of Z shown in figure 5.6. In the upper frame of the figure, the sample mean of Z is away from 0 for all of the producers and almost all of the starting times for the collection of verification values of cumulative oil. As was previously described, we collected the cumulative oil production at the producers every 30 days until the end. Interestingly, this configuration led to variances of Z less than 1. This indicates that the ensemble is too large, which is an unexpected behavior. Due to the biases in each of the producers and the large spread of the ensemble, this EnKF configuration is unreliable compared to the performace that we saw in the inverted seismic and production assimilating configurations.

5.1 Summary

This chapter showed our measurements of the performance of an EnKF configuration which assimilated the seismic parameters known as AVO data into the Box Model simulation. We were expecting to see improved parameter estimation compared with the production case, especially for porosity, as there is a known relationship between the geologic and seismic parameters. For porosity, this relationship is pictured in figure 5.1 and can also be seen in the equations from the rock physics model, see section 1.3.1. We did not observe an improvement and in fact observed quite a bit of instability when this data was assimilated. Many simulations would eventually fail as the state vector became inplausible after a number of analysis steps. As we know there is a correlation between the

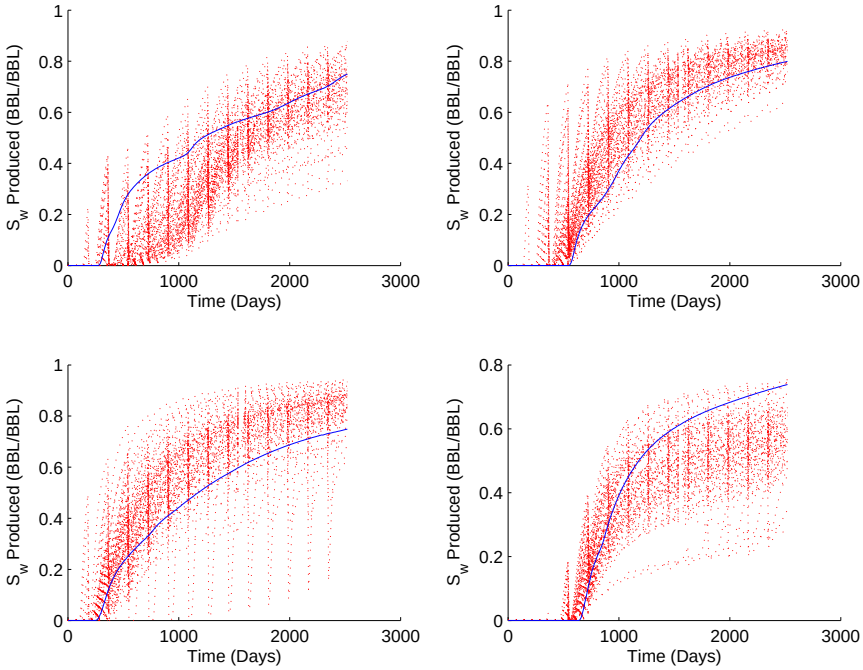

Figure 5.4: The water cut curves for each of the producers, under the assimilation of seismic parameters. The ensemble is shown in red and the truth in blue.

geology and seismic measurements, it could be that the source of the instability may come from a modeling error in our seismic modeling. It also could be that the permeability and porosity relationship is just difficult to expose using the one-shot seismic attributes. The data here is not conclusive enough to say which of these is more consistent with what we observed.

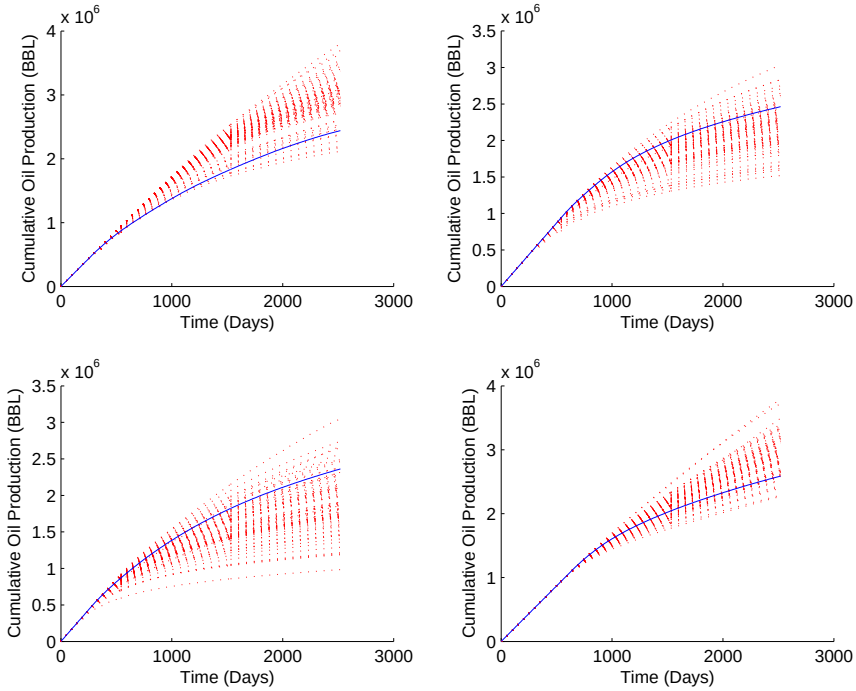

Figure 5.5: The cumulative oil curves for each of the producers, under the assimilation of seismic parameters. The ensemble values are shown in red and the true value is shown in blue.

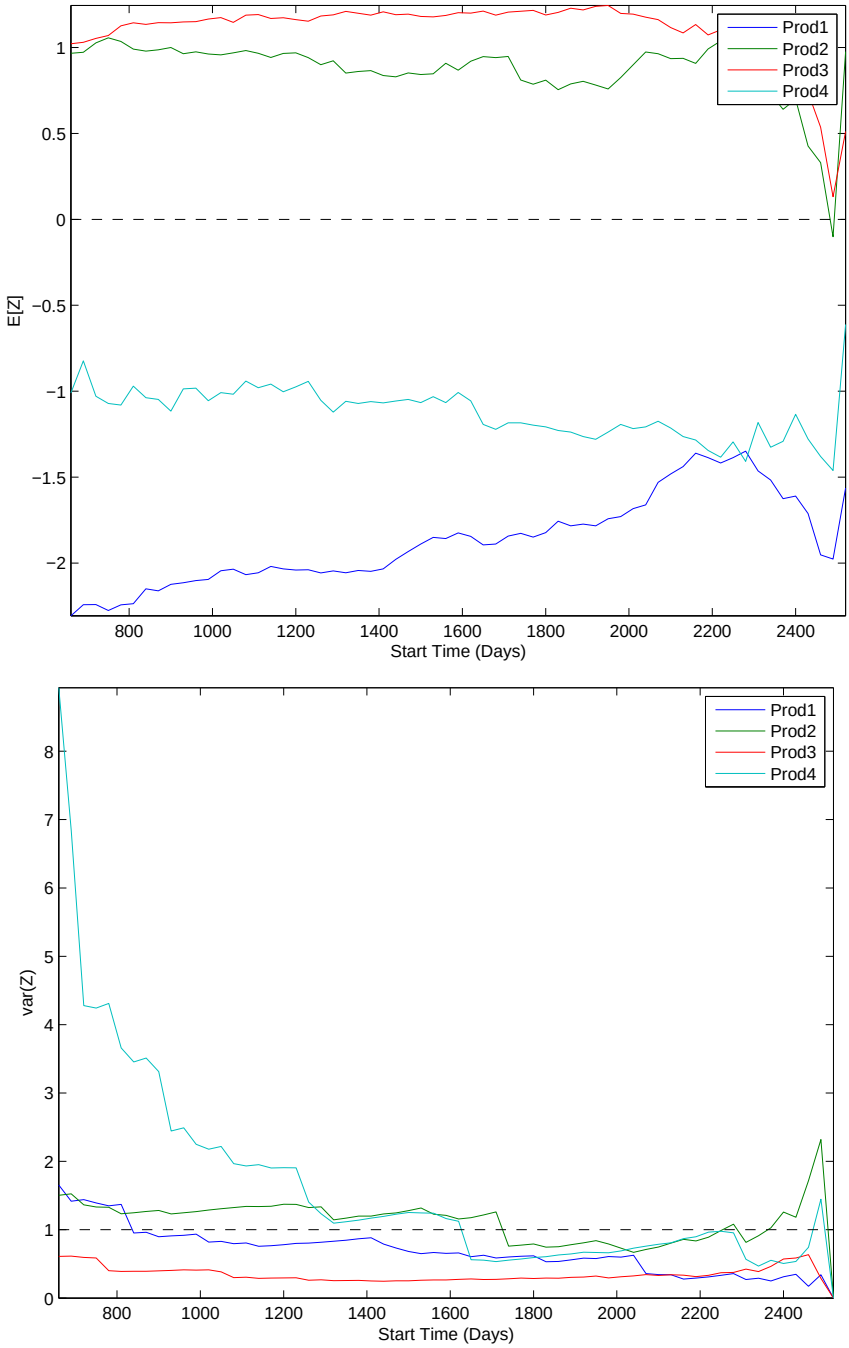

Figure 5.6: The statistics of the RCRV, calculated for cumulative oil production at all of the producers, under the assimilation of forward seismic data.

Chapter 6

Conclusions

In this project, we developed a data assimilation technique for the monitoring and forecasting of an oil reservoir based on the Ensemble Kalman Filter (EnKF). Specifically, we extended the Shell-owned software CLOREM to include the ability to assimilate synthetic seismic data. The purpose was to use a twin experiment to investigate seismic assimilation with respect to the information that it could supply to the EnKF and the elements of filter configuration necessary to make seismic data relevent. The idea of the data being relevent led to an investigation of the various ways to evaluate the performance of an EnKF describing a reservoir undergoing fluid injection and extraction.

One of the fundamental considerations for this project was the type of seismic data which we would attempt to assimilate using the EnKF. Seismic data is essentially an instrumental measurement of impedence within the reservoir and travel times to and from the interfaces within the reservoir. The data can then be processed to estimate the properties of the reservoir. Different estimates can be made according to the data present and the amount to computational power available. We chose to implement the modified Landrø inversion scheme, provided by Mario Trani [17], which estimates the change in saturation and pressure between two seismic surveys of a reservoir. This is much easier, computationally, than the full inversion from seismic to exact pressures and saturations. We also experimented with the assimilation of the seismic properties of the reservoir, known as AVO data, which are given by a minimumal amount of computation of the instrumental data. The difference between the assimilation of these two can be seen in the difference between figures 1.1 and 1.1. The results for the two methods were shown in the chapters 4 and 5.

We found that the assimilation of the inverted, time-lapse, seismic data resulted in better performance of the EnKF, in parameter estimation and predictive ability, than did the assimilation of the seismic attributes directly. We believe

the better performance results from a stronger correlation between the changes in seismic attributes and the changes in fluid properties than between the seismic attributes and fluid properties at a given time. It could also be that there was some modeling error in the seismic modeling. Those calculations include many co-efficients and other model parameters which come from experimental values and legacy calculations. In the case of the modified Landrø data, we had to remove a bias in the seismic results to see improved performance of the EnKF. This was not possible for the AVO data, as the true seismic parameter is not something which we can take from the forcasts of the true state and it must be modeled. Inaccuracies in the rock physics framework could have caused problems with the assimilation of the seismic parameters.

In evaluating the assimilation of the modified Landrø data, we also compared the performance of that EnKF configuration with one which assimilated the production and injection at the wells. The comparison between the two showed that the EnKF with time-lapse seismic data resolved the rock properties of the reservoir better than the production-only EnKF, particularly with regard to the structure of the permeability and porosity fields. Though further work is required to determine the number of seismic shots necessary to get the best characterization of the rock properties possible. It is likely that future observation schemes will combine the assimilation of production and seismic data and an interesting course of study would be to look at using several seismic surveys early in the monitoring and then continue to use production data in updates as the EnKF moves the ensemble forward. The RCRV statistics may even give a way to monitor the performance of the EnKF and request further seismic or geological data when the predictive power of the ensemble with respect to fluid production degrades below some acceptable level.

The use of statistical measures to monitor the performance of an EnKF is one of the most interesting ideas to come from this thesis from a mathematical standpoint. We introduce the reduced centered random variable (RCRV) as a measure of filter predictive power. Specifically, we tracked the statistics of the that variable to assess the ensemble's bias and spread for production variables. In our experiments, the RCRV for cumulative oil production seemed to give a new way to evaluate the efficacy of a EnKF configuration's predictions. The RCRV showed when the ensemble was too small or large and also indicated bias and the statistics of the RCRV seemed stable over a large number of samples. From a practical standpoint, it would be easy to implement an RCRV in real time, especially with regard to production data. The benefits from knowing the statistical strength and informativeness of the production predictions with regard to optimization strategies should be further investigated.

In the end, we concluded that the assimilation of inverted time-lapse seismic data is a promising technique for reservoir engineers. We found that after a few

assimilations of modified Landrø data the ensemble parameter estimates were greatly improved. From the reservoir structure standpoint, this was especially true. This could prove to be very useful for engineers charged with making the field development plans and further study should be performed to get a better idea of the practicalities of using modified Landrø data in reservoir EnKFs.

Bibliography

[1] Gerrit Burgers, J. van Leeuwen, and Gier Evensen. Analysis scheme in the ensemble kalman filter. *Mon. Wea. Rev.*, (126):1719–1724, 1998. [cited at p. 15]

[2] G. Candille, C. Cote, P.L. Houtekamer, and G. Pellerin. Verification of an ensemble prediction system against observations. Technical report, Environment Canada, 2006. [cited at p. 20]

[3] L.P. Drake. *Fundamentals of Reservoir Engineering*. Elsevier, 1978. [cited at p. 5]

[4] Gier Evensen. Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. *J.G.R.*, 99(C5):10143–10162, 1994. [cited at p. 15]

[5] Gier Evensen. Sampling strategies and square root analysis schemes for the enkf. *Ocean Dynamics*, (54):539–560, 2004. [cited at p. 15]

[6] Y. Q. Gu and D. S. Oliver. History matching of the punq-s3 reservoir model using the ensemble kalman filter. *SPE J.*, 10(2):217–224, JUN 2005. [cited at p. 1]

[7] Jan Dirk Jansen. Systems theory for reservoir management. Lecture Notes, TU Delft, AES1490, 2008. [cited at p. 5, 6]

[8] R.E. Kalman. A new approach to linear filtering and prediction problems. *Trans. AMSE*, (82):35–45, 1960. [cited at p. 13]

[9] M.V. Krymskaya. Parameter estimation in reservoir engineering models via data assimilation techniques. Master's thesis, Technical University of Delft, 2007. [cited at p. 28]

[10] Martin Landro. Discrimination between pressure and fluid saturation changes from time-lapse seismic data. *Geophysics*, 66(3):836–844, May-June 2001. [cited at p. 10]

[11] Martin Landro. Uncertainties in quantitative time-lapse seismic analysis. *Geophysical Prospecting*, (50):527–538, 2002. [cited at p. 29]

[12] Rolf J. Lorentzen, Geir Naeval, Brice Valles, Aina M. Berg, and Alv-Arne Grimstad. Analysis of the ensemble kalman filter for estimation of permiability and porosity in reservoir models. *SPE J.*, OCT 2005. [cited at p. 1]

[13] Jan Mandel. Efficient implementation of the ensemble kalman filter. Technical Report CCM 231, University of Colorado, 2006. [cited at p. 16]

[14] G. Mavko, T. Mukerji, and Jack Dvorkin. The rock physics handbook. 1998. [cited at p. 8]

[15] G. Naevdal, L. M. Johnsen, S. I. Aanonsen, and E. H. Vefring. Reservoir monitoring and continuous model updating using ensemble kalman filter. *SPE J.*, 10(1):66–74, MAR 2005. [cited at p. 1]

[16] J.A. Skjervheim, G. Evenson, S.I. Aanonsen, B.O. Ruud, and T.A. Johnson. Incorporating 4d seismic data in reservoir simulation models using ensemble kalman filter. *SPE J.*, OCT 2005. [cited at p. 2]

[17] M. Trani, R.J. Arts, O. Leeuwenburgh, and J.H. Brouwer. Estimation of pressure and saturation changes from 4d seismic and time-shift analysis. In *70th EAGE Conference and Exhibition Incorporating SPE Europec*, 2008. [cited at p. 51]

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

Appendices

Appendix A

Addendum to Chapter 3

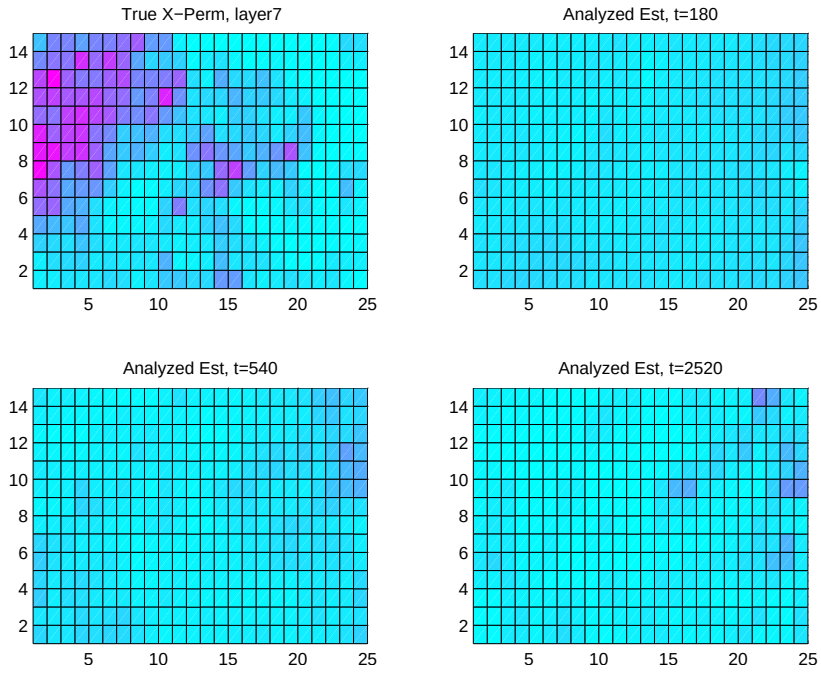

Figure A.1: The X-permeability in layer 4 of the box model, truth (upper left) and three ensemble estimates, under assimilation of production data.

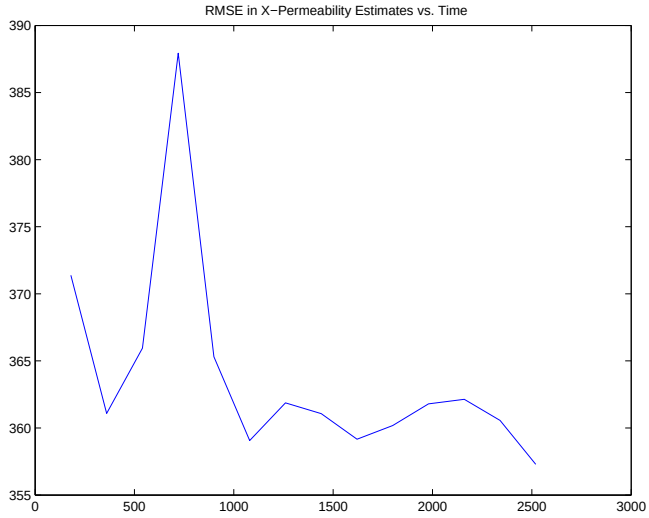

Figure A.2: The RMSE for the post-analysis X-permeability estimates after the assimilation of production data.

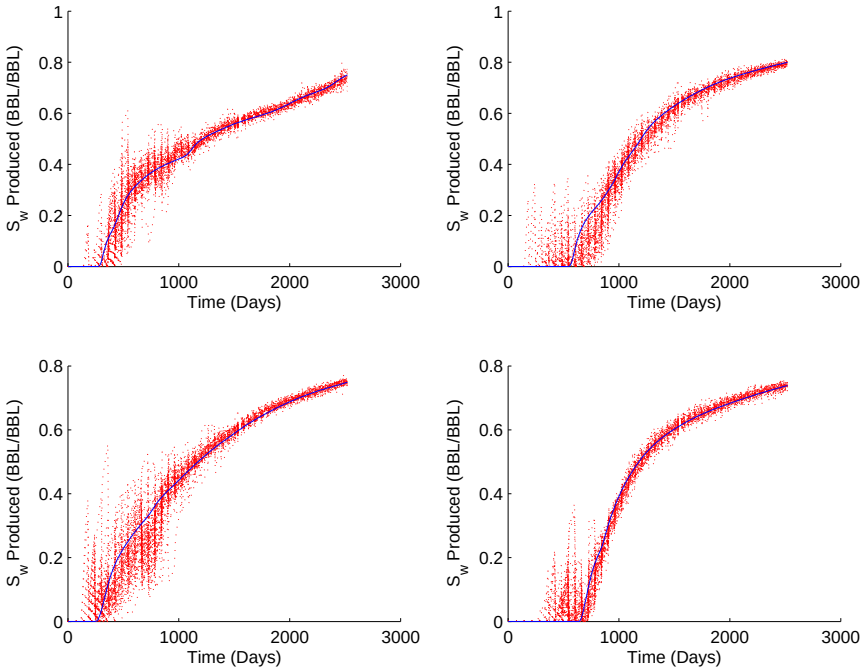

Figure A.3: The cumulative oil curves for each of the producers, under the assimilation of production data. The ensemble values are shown in red and the true value is shown in blue.

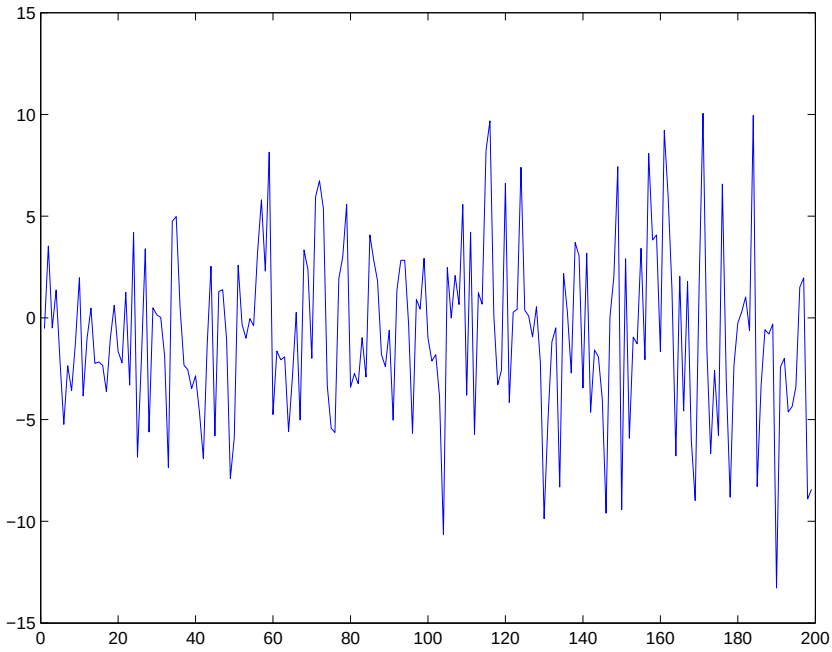

Figure A.4: The water cut curves for each of the producers, under the assimilation of production data. The ensemble is shown in red and the truth in blue.

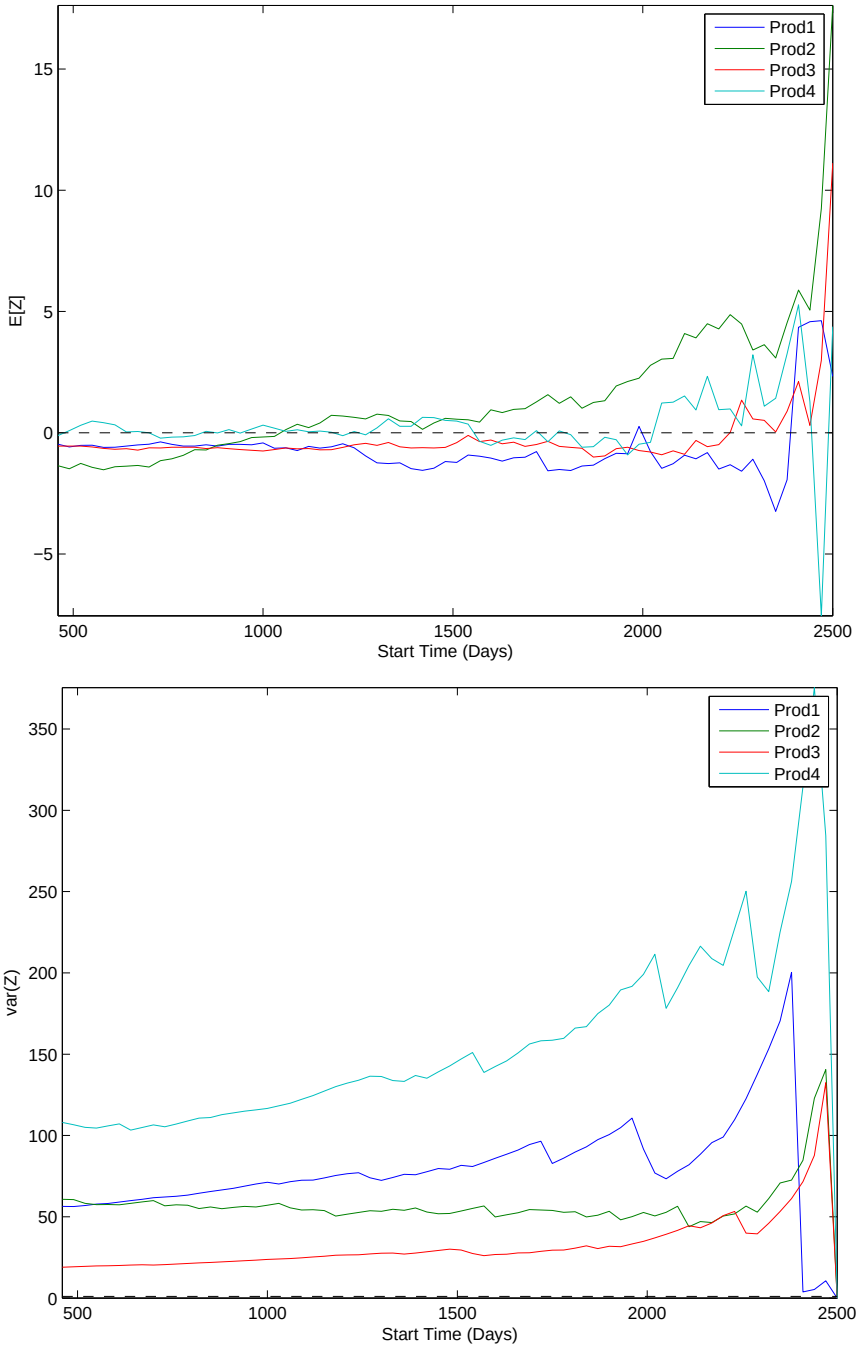

Figure A.5: The statistics of the RCRV, calculated for cumulative oil production at all of the producers, under the assimilation of production data.

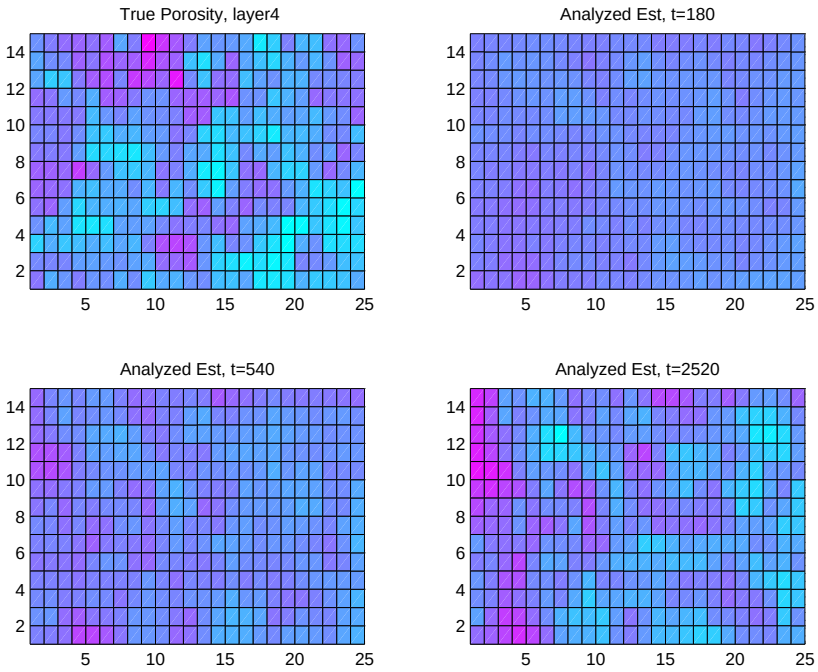

Figure A.6: The porosity truth (upper left) and three ensemble estimates, under assimilation of production data.

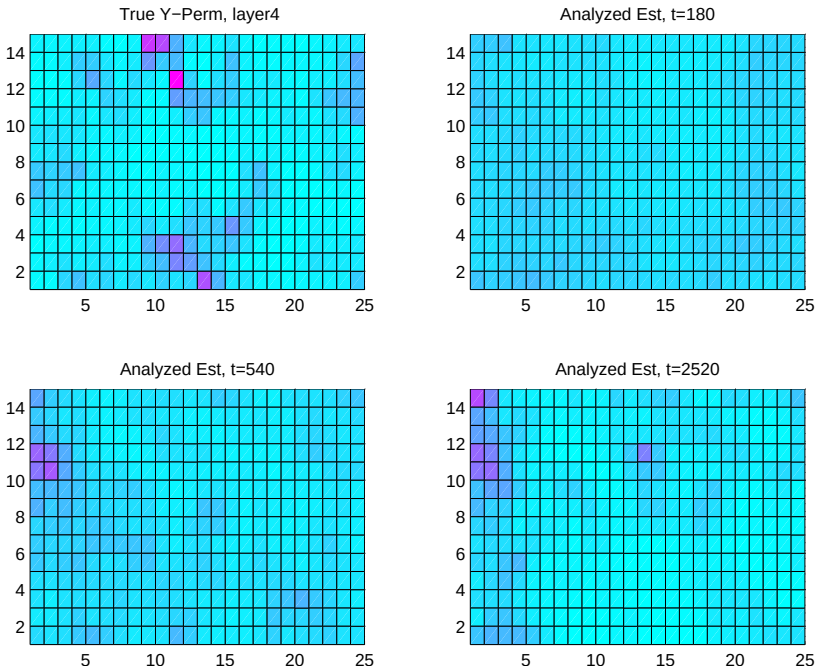

Figure A.7: The Y-permeability truth (upper left) and three ensemble estimates, under assimilation of production data.

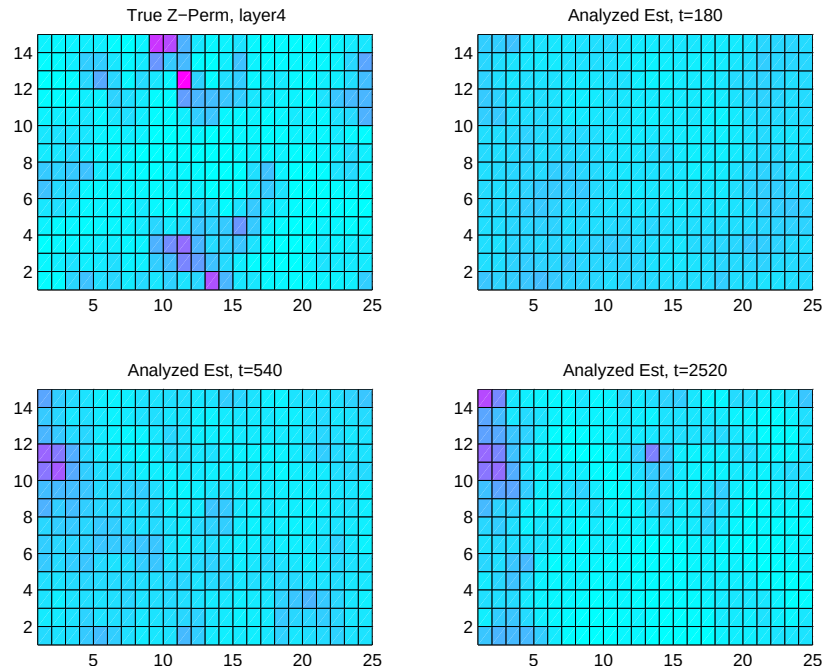

Figure A.8: The Z-permeability truth (upper left) and three ensemble estimates, under assimilation of production data.

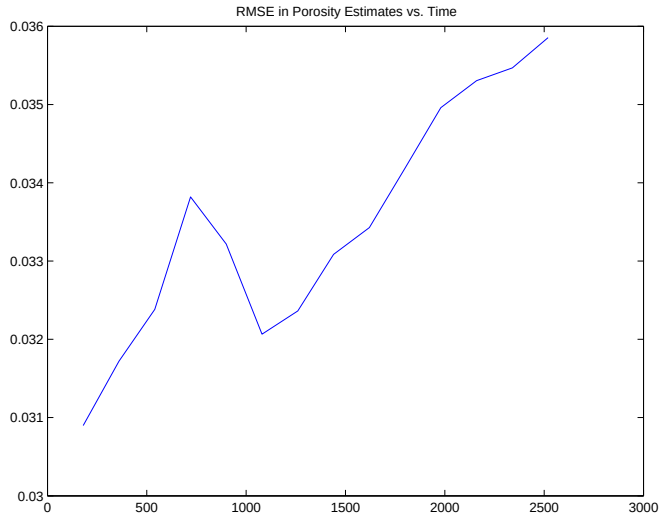

Figure A.9: The RMSE for the post-analysis porosity estimates after the assimilation of production data.

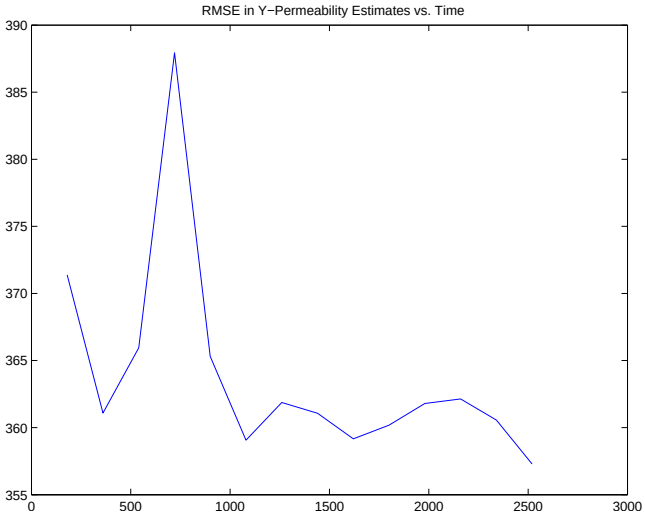

Figure A.10: The RMSE for the post-analysis Y-permeability estimates after the assimilation of production data.

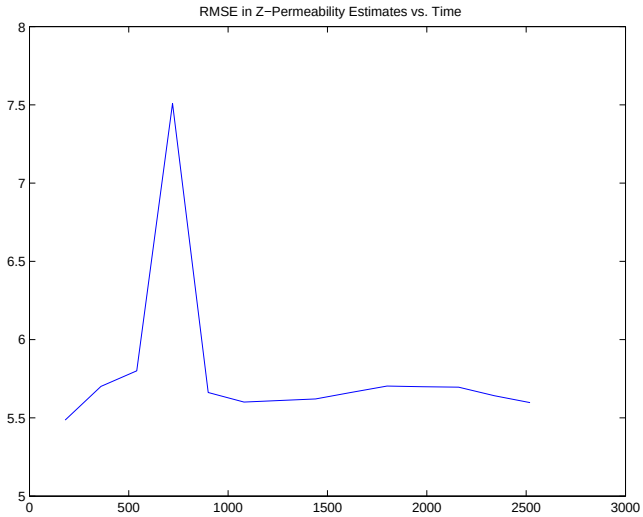

Figure A.11: The RMSE for the post-analysis Y-permeability estimates after the assimilation of production data.

Appendix B

Addendum to Chapter 4

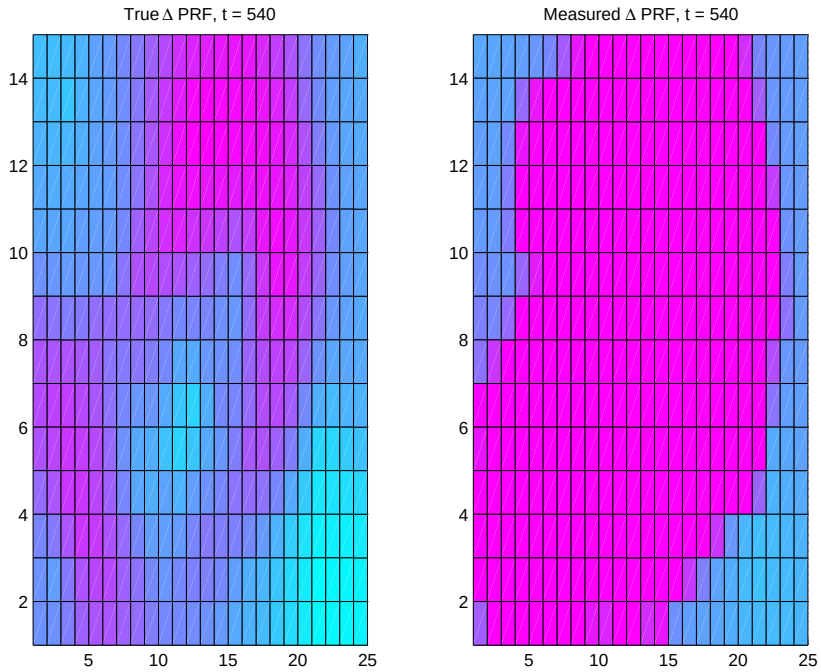

Figure B.1: The change in pressure between 180 and 540 days in a realization of the true model. The color scheme is the same for both frames.

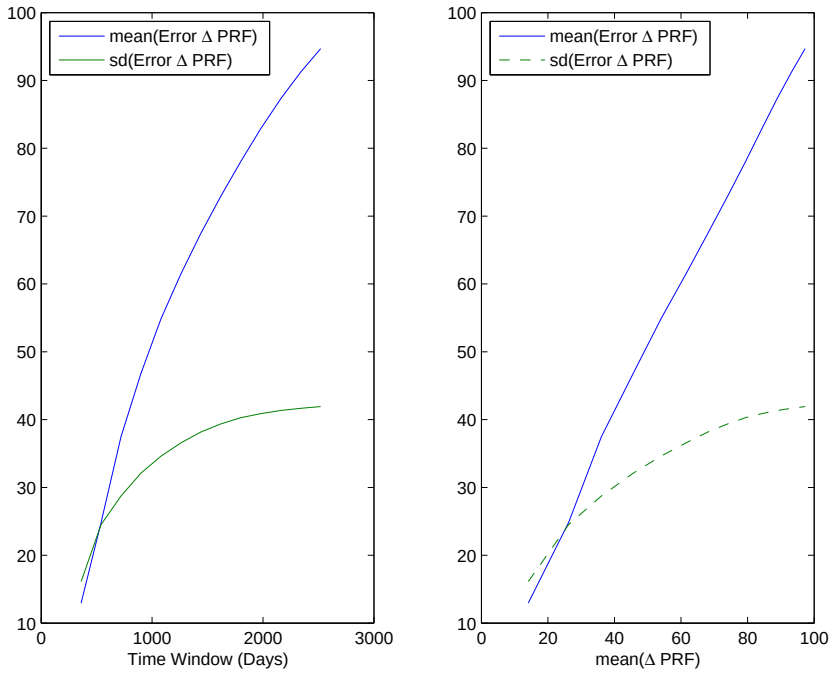

Figure B.2: The mean error in the ΔP measurement is plotted against both the time of the second seismic survey used in the time-lapse inversion and the mean, inverted, ΔP values for the associated time windows.

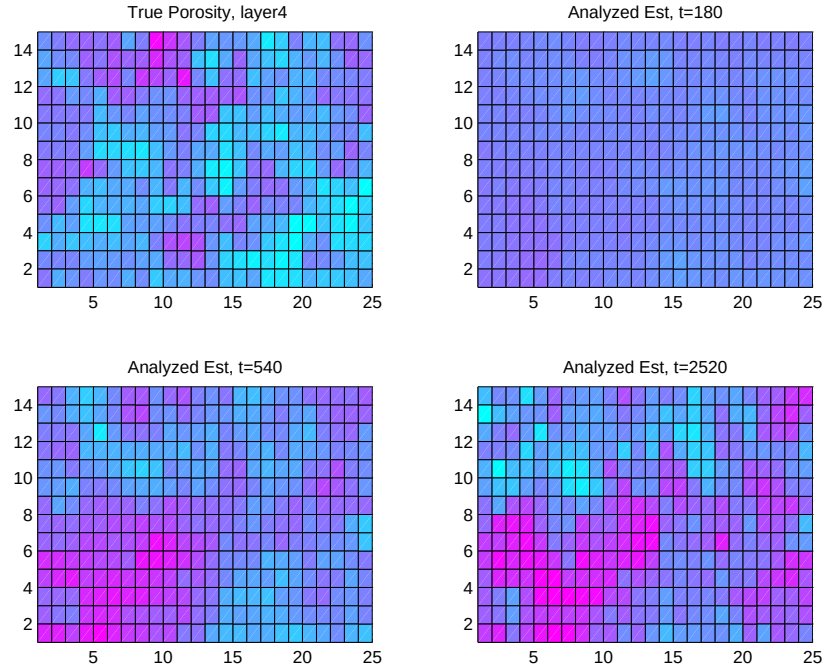

Figure B.3: The porosity truth (upper left) and three ensemble estimates, under assimilation of bias-removed modified Landrø data.

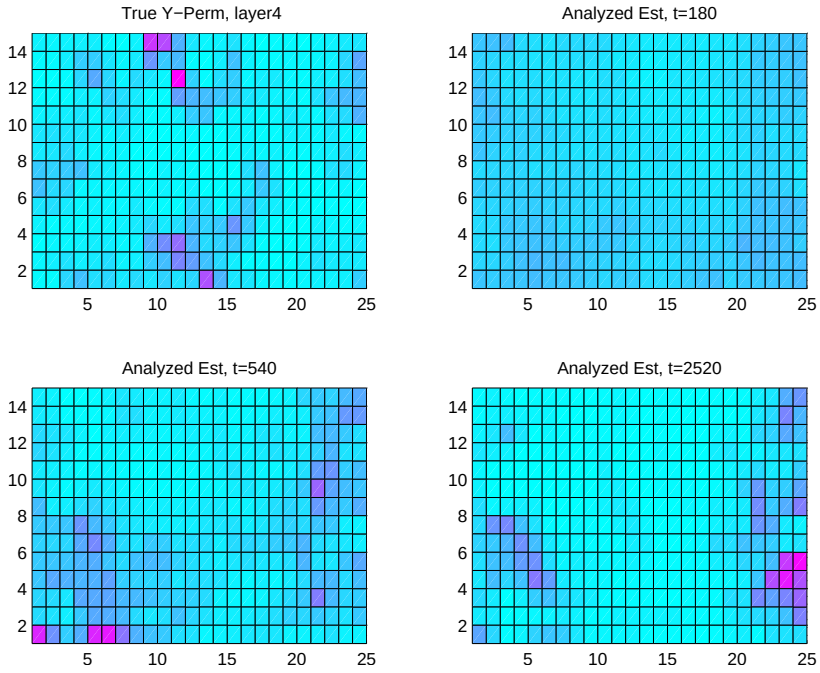

Figure B.4: The Y-permeability truth (upper left) and three ensemble estimates, under assimilation of bias-removed modified Landrø data.

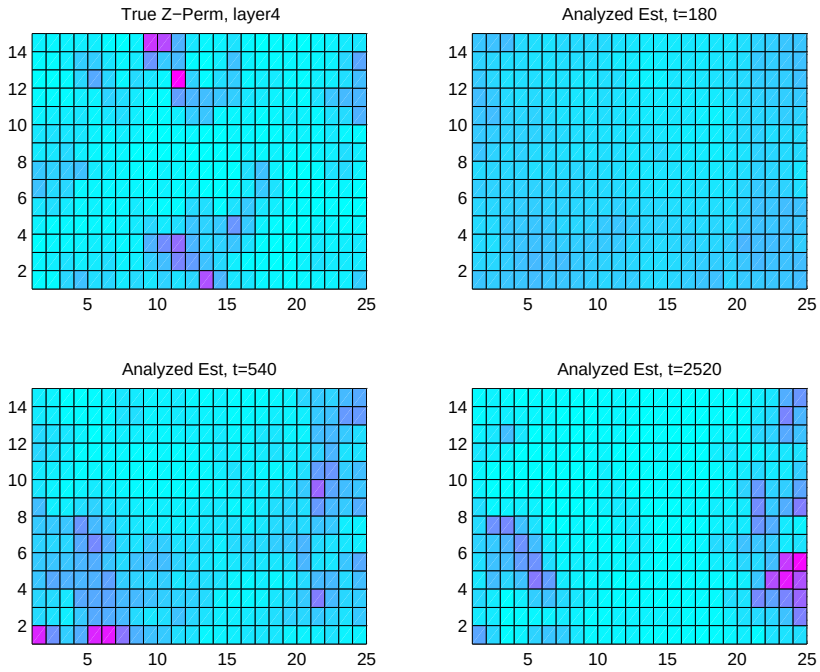

Figure B.5: The Z-permeability truth (upper left) and three ensemble estimates, under assimilation of bias-removed modified Landrø data.

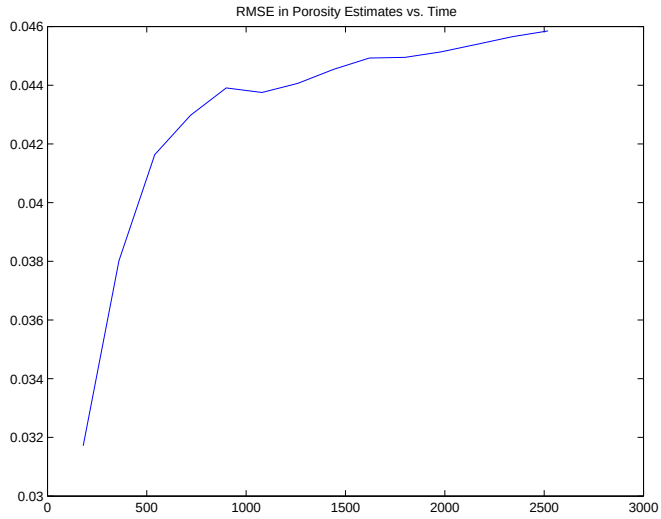

Figure B.6: The RMSE for the post-analysis porosity estimates after the assimilation of modified Landrø data.

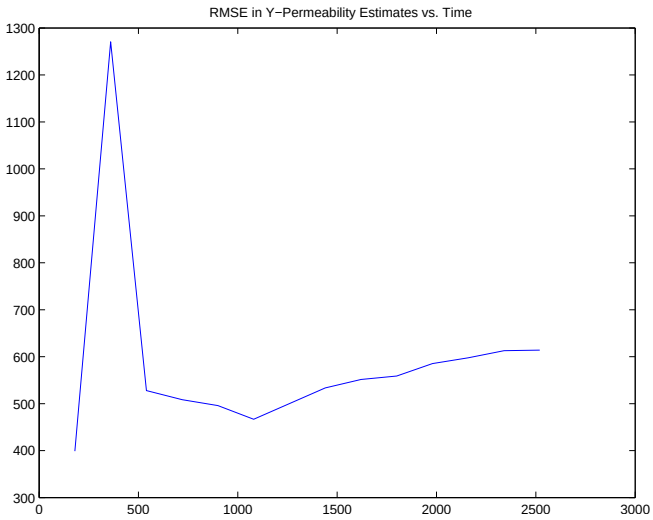

Figure B.7: The RMSE for the post-analysis Y-permeability estimates after the assimilation of modified Landrø data.

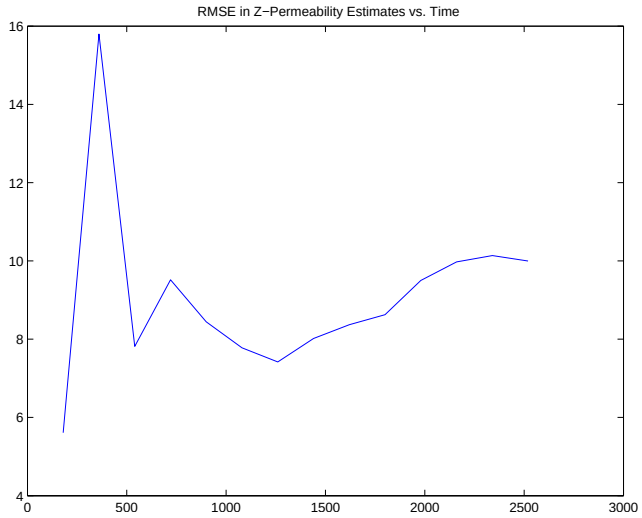

Figure B.8: The RMSE for the post-analysis Y-permeability estimates after the assimilation of modified Landrø data.

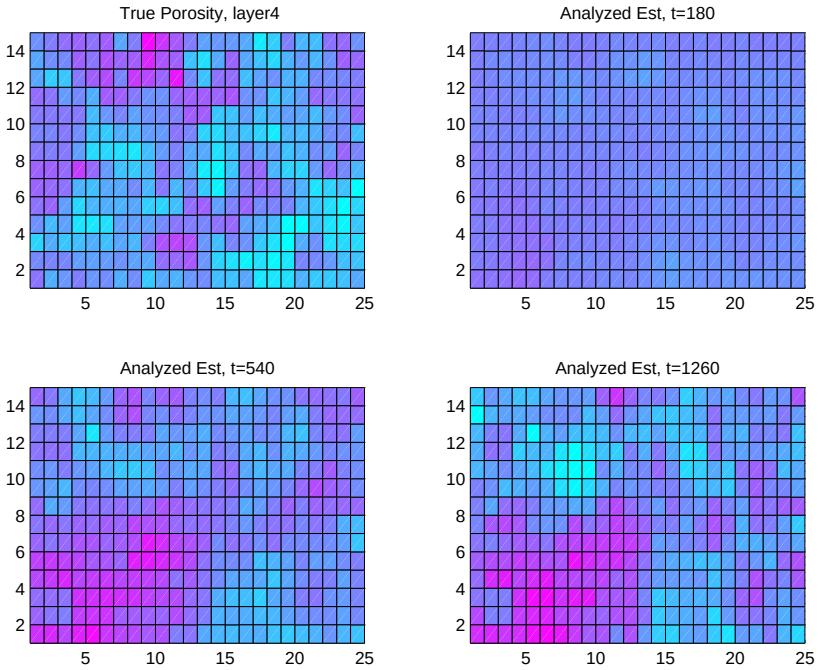

Figure B.9: The porosity truth (upper left) and three ensemble estimates, under assimilation of bias-removed modified Landrø data until 1260 days.

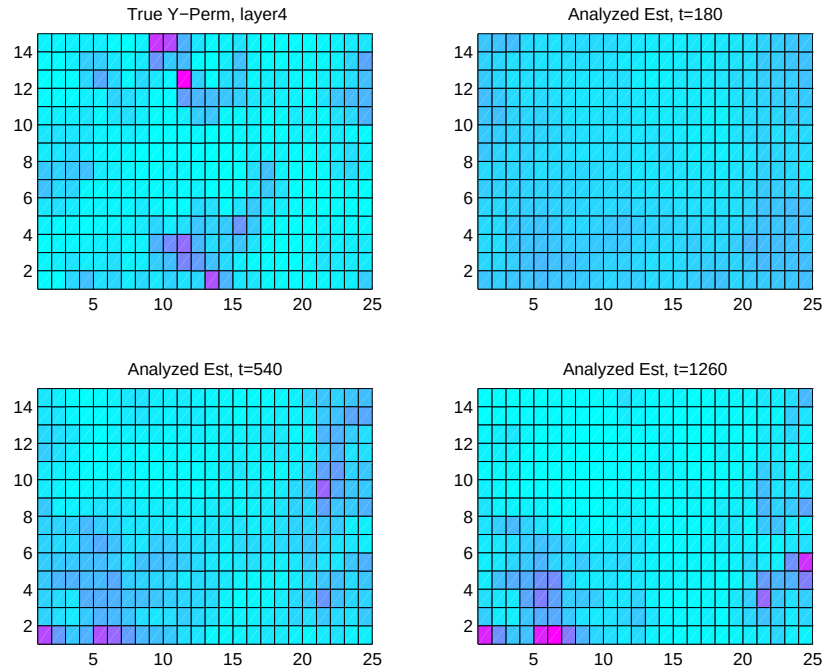

Figure B.10: The Y-permeability truth (upper left) and three ensemble estimates, under assimilation of bias-removed modified Landrø data until 1260 days.

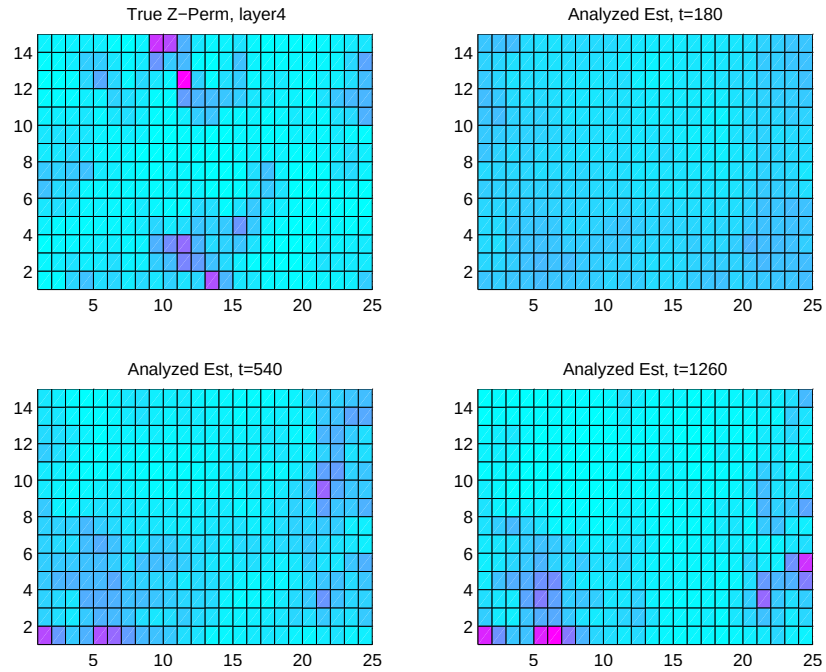

Figure B.11: The Z-permeability truth (upper left) and three ensemble estimates, under assimilation of bias-removed modified Landrø data until 1260 days.

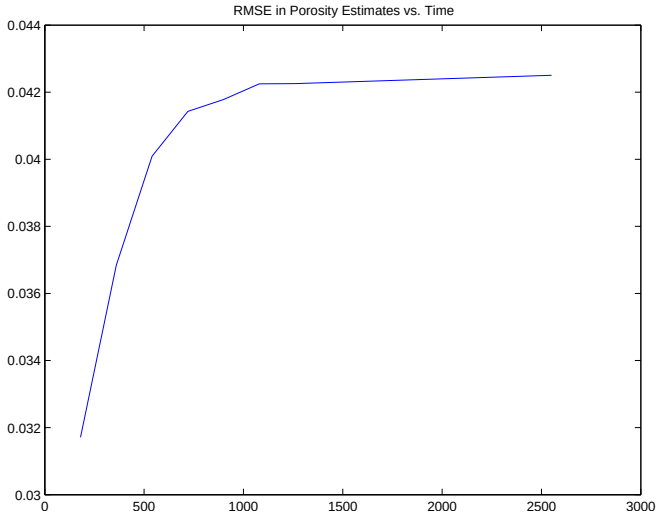

Figure B.12: The RMSE for the post-analysis porosity estimates after the assimilation of modified Landrø data until 1260 days.

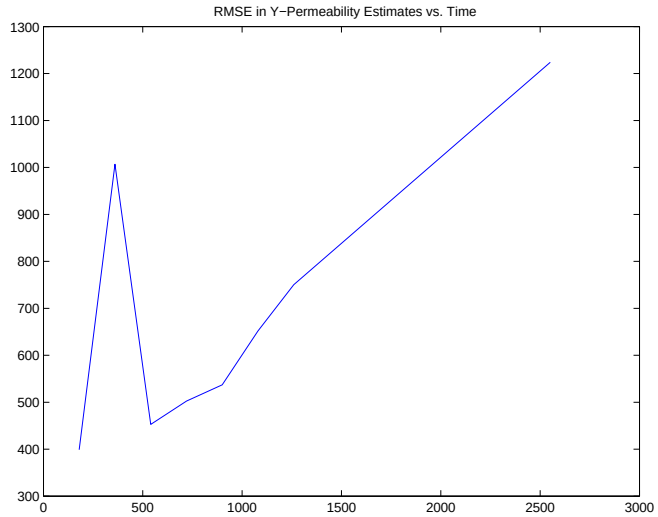

Figure B.13: The RMSE for the post-analysis Y-permeability estimates after the assimilation of modified Landrø data until 1260 days.

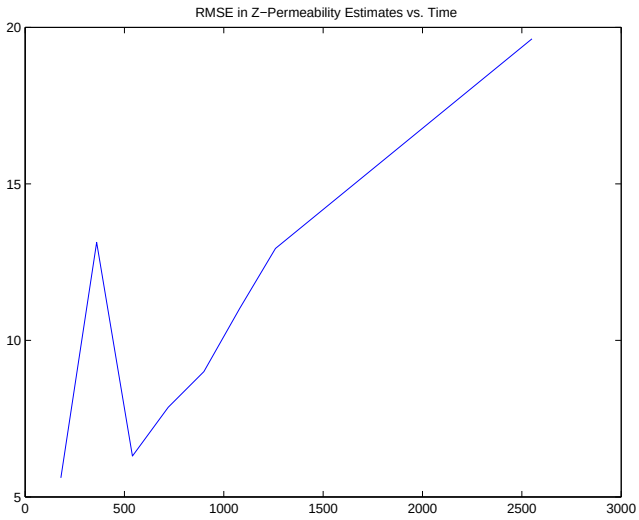

Figure B.14: The RMSE for the post-analysis Y-permeability estimates after the assimilation of modified Landrø data until 1260 days.

Appendix C

Addendum to Chapter 5

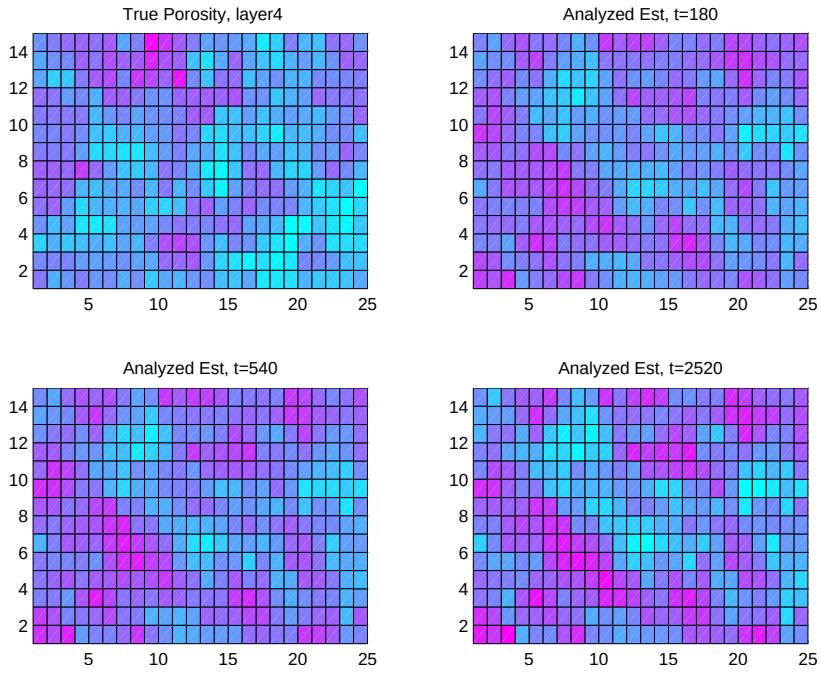

Figure C.1: The porosity truth (upper left) and three ensemble estimates, under assimilation of the AVO seismic attributes.

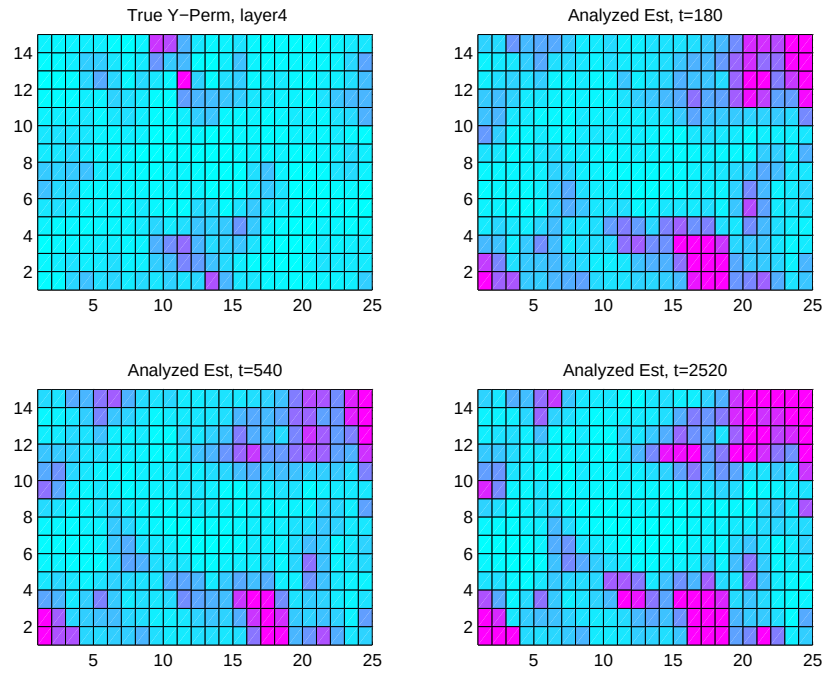

Figure C.2: The Y-permeability truth (upper left) and three ensemble estimates, under assimilation of the AVO seismic attributes.

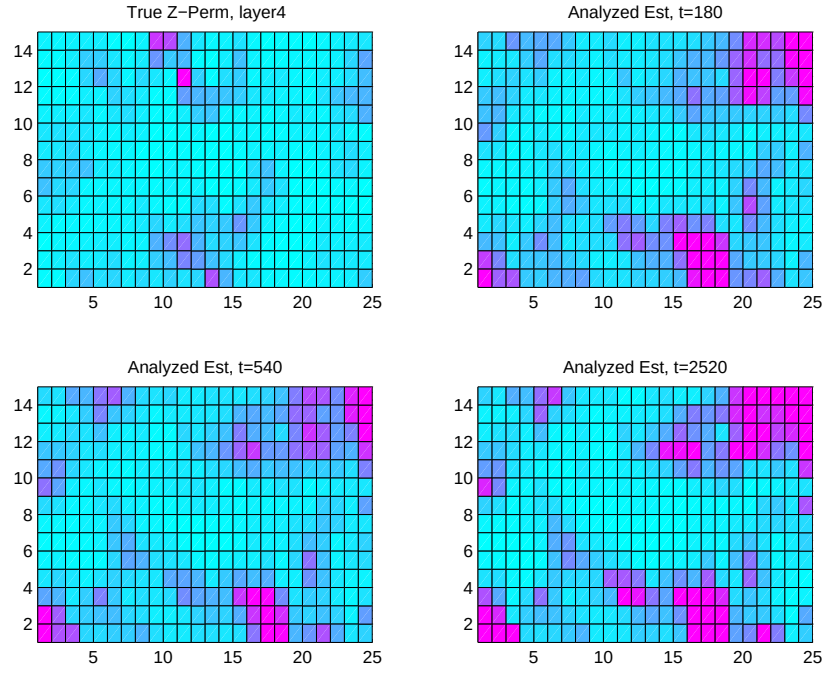

Figure C.3: The Z-permeability truth (upper left) and three ensemble estimates, under assimilation of the AVO seismic attributes.

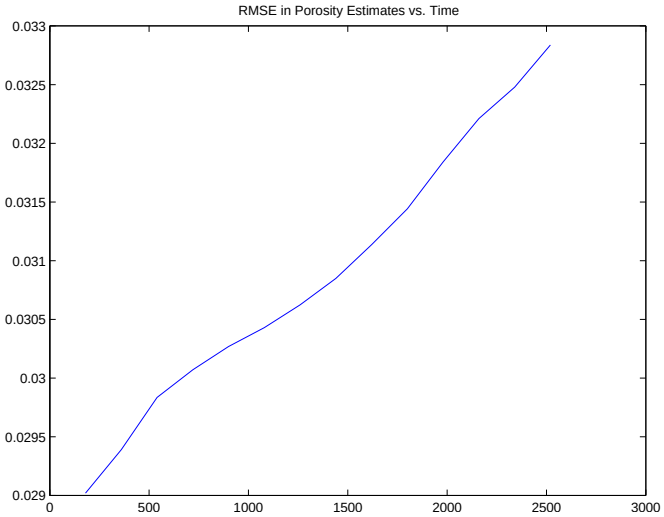

Figure C.4: The RMSE for the post-analysis porosity estimates after the assimilation of the AVO seismic attributes.

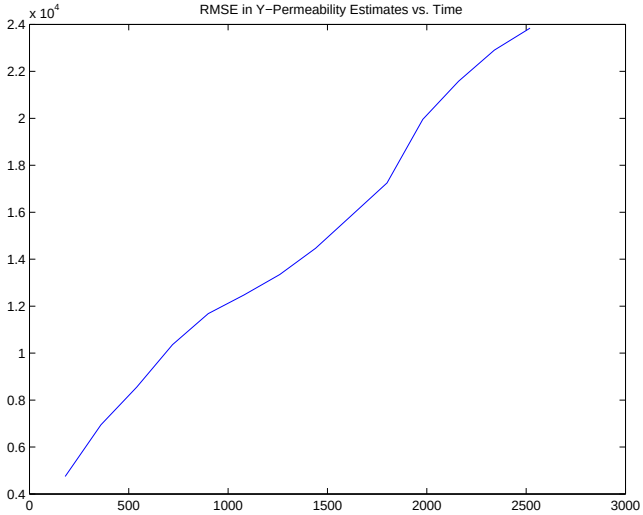

Figure C.5: The RMSE for the post-analysis Y-permeability estimates after the assimilation of the AVO seismic attributes.

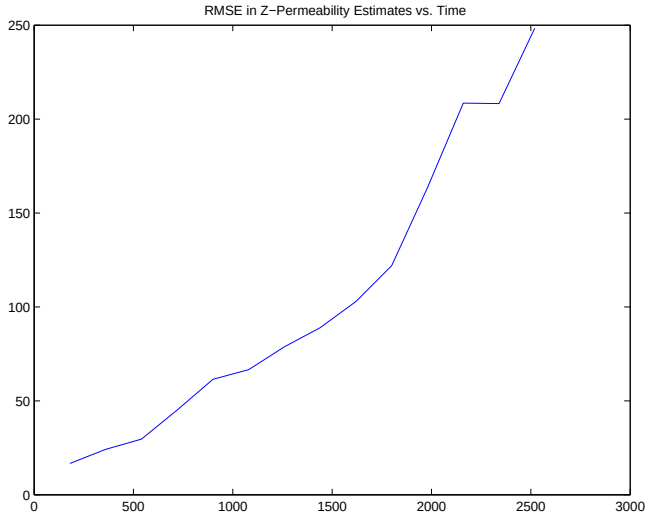

Figure C.6: The RMSE for the post-analysis Y-permeability estimates after the assimilation of the AVO seismic attributes.