We P018

Bayesian Inversion of Time-lapse Seismic Waveform Data Using an Integral Equation Method

K. S. Eikrem* (The National IOR Centre of Norway/IRIS), M. Jakobsen (UiB/IRIS) & G. Nævdal (IRIS)

SUMMARY

In the last couple of decades, we have witnessed an increased use of time-lapse seismic data. Interpretation of time-lapse seismic data can give a better understanding of the oil saturation in the reservoir, leading to identification of the water-flooded areas and pockets of remaining oil, and an improved understanding of compartmentalization of the reservoir. Within the context of dynamic reservoir characterization or seismic history matching, where one performs a quantitative integration of time-lapse seismic and production data, the covariance matrix (quantifying the uncertainty) of the seismic data needs to be specified. Usually, this is done in a very ad-hoc manner, for example by using a diagonal covariance matrix where the uncertainty is given in percentage of the measurement values. Eikrem et al. (2016) has recently demonstrated that a more accurate and complete dynamic reservoir characterization can be obtained if one performs a Bayesian seismic waveform inversion for the seismic parameters and use the full covariance matrix when updating permeability and porosity. In that paper a simple linear Born inversion was used, and it is of interest to investigate whether similar results hold for a more advanced seismic inversion method. The present work will focus on Bayesian nonlinear full waveform inversion (FWI) to get an estimate of the uncertainty in the seismic inversion. In contrast with the main stream of researchers within the FWI community, we develop a direct iterative nonlinear Bayesian inversion method based on an explicit representation of the data sensitivity function in terms of Green functions, rather than the indirect optimization approach based on the adjoint state method. Our method is based on the T-matrix approach by Jakobsen and Ursin (2015).
Introduction

The use of time-lapse seismic data has increased over the last couple of decades. As many oil fields become more mature, having a good understanding of the reservoir is important in order to extract the remaining oil. The time-lapse seismic data can be used to identify water-flooded areas and pockets of remaining oil, and gives an improved understanding of compartmentalization of the reservoir. Having good reservoir models is important when choosing the optimal injection strategies and other IOR methods.

Updating reservoir models using time-lapse seismic data is still considered a difficult task. Ensemble based methods have become more popular for history matching reservoir models in recent years (Oliver and Chen (2011)), and is a natural choice for a joint inversion of time-lapse seismic and production data. The ensemble-based methods are generally formulated in a Bayesian framework. That means that we need to provide a prior distribution for the parameters to be estimated, and also the uncertainty of the data in the form of a covariance matrix. Most seismic inversion methods do not provide information about the uncertainty in the inverted values. Usually the uncertainty is estimated in a very ad-hoc manner, for example by using a diagonal covariance matrix where the uncertainty is given as a percentage of the measurement values. However, in Eikrem et al. (2016) it was recently shown that information about the off-diagonal terms of the covariance matrix in the inverted elastic parameters can in some cases lead to improvements in the dynamic reservoir characterization results. In that paper a simple linear Born inversion was used, and it is of interest to investigate whether similar results hold for a more advanced seismic inversion method.

This paper will focus on Bayesian nonlinear full waveform inversion (FWI) to get an estimate of the uncertainty in the seismic inversion. In our integral operator formalism as well as the scattering-integral approach to seismic FWI introduced by Tao and Sen (2013), the sensitivity matrix is formed explicitly via the use of Green functions. This is in contrast with the conventional (indirect) optimization approach to seismic FWI, where the sensitivity matrix is calculated implicitly via the use of the adjoint-state method. Although the adjoint-state method for computing sensitivity matrices (via back-propagation) is commonly used in exploration seismology (Tarantola (1984); Virieux and Operto (2009)), it is not obvious that the adjoint-state method will always be more efficient than a direct scattering-integral approach. In a comparative analysis of the scattering-integral and adjoint-state methods reported by Chen et al. (2007), it was concluded that the scattering-integral approach outperformed the adjoint-state method in a regional tomography study involving many sources (see also Tao and Sen (2013)). The relative computational cost of the adjoint and scattering-integral approach will depend on the number of unknown model parameters as well as the source-receiver configuration (Chen et al. (2007)). Tao and Sen (2013) argue that the scattering-integral approach is particularly attractive for frequency domain waveform inversion and will normally outperform the adjoint approach if the goal is to perform a fast sensitivity (uncertainty) analysis.

Our method is based on the T-matrix approach by Jakobsen and Ursin (2015). In that paper a Tikhonov method was used to get an estimate of the inverted velocities. Here we will use a Levenberg-Marquardt version of the iterated extended Kalman filter that will also give an estimate of the uncertainty in the inversion. We test our method on a synthetic two-dimensional model and perform two surveys. We add a small change in the second model to illustrate a time-lapse effect, and compare different choices of prior for the second inversion.

We will first review the modelling of seismic data and construction of the sensitivity matrix in the next section. Then we will describe the Bayesian inversion method, and after that we present the numerical experiments. Finally, the conclusion follows.
Modelling of seismic waveform data residuals

The Lippmann-Schwinger equation

In the frequency domain, the seismic (pressure) wavefield $\psi(x)$ satisfies the scalar Helmholtz equation (Morse and Feshbach (1953))

$$\left( \nabla^2 + \frac{\omega^2}{c^2(x)} \right) \psi(x) = -S(x),$$

(1)

where $c(x)$ is the wave speed at position $x$ and $\omega$ is the angular frequency, and $S(x)$ is the frequency-dependent source density. A general solution to the scalar Helmholtz equation is given by (Morse and Feshbach (1953))

$$\psi(x) = \int dG(x, x') S(x'),$$

(2)

where $G(x, x')$ is the Green’s function for the Helmholtz equation, which gives the wavefield at point $x$ due to a unit point source at point $x'$. The Green’s function $G(x, x')$ is defined by (Morse and Feshbach (1953); Jakobsen (2012); Jakobsen and Ursin (2015)):

$$\left( \nabla^2 + \frac{\omega^2}{c^2(x)} \right) G(x, x') = -\delta(x - x').$$

(3)

Thus, solving the scalar Helmholtz equation (1) is equivalent to first determining the Green’s function by solving equation (3) and then calculating the corresponding wavefield via the scattering-integral (2).

In order to determine the Green’s function, we use a kind of perturbation theory which is valid for both weakly and strongly scattering media. Following Jakobsen (2012) and Jakobsen and Ursin (2015), we define $c_0(x)$ to be the wave speed in an arbitrary heterogeneous background medium, so that we can rewrite equation (3) exactly as

$$\left( \nabla^2 + \frac{\omega^2}{c_0^2(x)} \right) G(x, x') = -\delta(x - x') - \omega^2 \chi(x) G(x, x'),$$

(4)

where

$$\chi(x) = c^{-2}(x) - c_0^{-2}(x)$$

(5)

is the perturbation of the squared slownesses. The second term on the right-hand side of equation (4) represents the so-called contrast-sources. By using equation (2) and (4), we obtain the following volume integral equation for the Green’s function (Morse and Feshbach (1953); Jakobsen (2012); Jakobsen and Ursin (2015); Jakobsen and Wu (2016))

$$G(x, x') = G^{(0)}(x, x') + \omega^2 \int_D dG''(x, x'') \chi(x'') G(x'', x'),$$

(6)

where $D$ is the scattering domain where $\chi(x'')$ is non-zero and $G^{(0)}(x, x')$ is the background medium Green’s function, that satisfies

$$\left( \nabla^2 + \frac{\omega^2}{c_0^2(x)} \right) G(x, x') = -\delta(x - x').$$

If the background medium is homogeneous, then one can use simple analytical expressions for $G^{(0)}(x, x')$. If the background medium is heterogeneous but smooth, then one can estimate the Green’s functions using ray theory (see Cerveny (2001)). In the case of a general heterogeneous background medium, the Green’s functions need to be calculated numerically, for example by using the finite difference method (e.g., Kirchner and Shapiro (2001)) or a volume integral equation method (Jakobsen and Ursin (2015); Jakobsen and Wu (2016)).
Integral operator formalism

To allow for the use of an established integral operator formalism (Jakobsen (2012); Jakobsen and Ursin (2015); Jakobsen and Wu (2016)), we rewrite the Lippmann-Schwinger equation (6) exactly as

\[
G(x, x') = G^{(0)}(x, x') + \int_D dx_1 \int_D dx_2 G^{(0)}(x_1, x_2)V(x_1, x_2)G(x_2, x'),
\]

where

\[
V(x_1, x_2) = \omega^2 \chi(\mathbf{x}_1) \delta(x_1 - x_2),
\]

is the kernel of a local scattering potential operator.

The symmetric form (7) of the Lippmann-Schwinger equation (6) can be interpreted as the real-space coordinate representation of the following integral operator equation (Jakobsen and Wu (2016))

\[
G = G^{(0)} + G^{(0)}V G.
\]

The use of integral operator notation is convenient for formal manipulations (e.g., Jakobsen and Wu (2016)), but all integral operators can simply be represented by matrices of different dimensions after spatial discretization (Jakobsen and Ursin (2015)).

Since the integral equation (7) is similar to the Lippmann-Schwinger equation in quantum scattering theory, we may try to modify the highly developed perturbative or iterative methods developed in that part of scattering theory for use in seismic. Following the standard approach of quantum potential scattering theory (e.g., Taylor (1972); Jakobsen and Ursin (2011); Jakobsen (2012); Jakobsen and Ursin (2015)), we now introduce the so-called transition operator \( T \) by

\[
VG = TG^{(0)}.
\]

In terms of the \( T \) operator, the Lippmann-Schwinger equation can be rewritten exactly as (detailed by Jakobsen and Ursin (2015))

\[
G = G^{(0)} + G^{(0)}TG^{(0)}.
\]

From the Lippmann-Schwinger equation (8), the defining relation (9) and the fact that the background medium is arbitrary, it follows that (e.g., Taylor (1972); Jakobsen (2012); Jakobsen and Ursin (2015) Appendix B)

\[
T = V + VG^{(0)}T.
\]

Thus, the \( T \)-operator satisfies an integral equation of the Lippmann-Schwinger type, independent of the source-receiver configuration; convenient for applications to time-lapse seismic (e.g., Jakobsen and Ursin (2011); Muhumuza (2015)).

The sensitivity matrix

To compute the sensitivity of the seismic waveform data with respect to small perturbations in the model parameter for a surface seismic acquisition geometry, we now assume there are multiple sources and receivers located at positions \( x_s \), where \( s = 1, \ldots, N_s \) and \( x_r \), where \( r = 1, \ldots, N_r \), respectively. We divide the target zone where the scattering potential operator \( V \) is non-zero into a set of \( N \) gridblocks (or pixels), with centroids located at positions \( x_n \), where \( n = 1, \ldots, N \). Following Jakobsen and Ursin (2015), we need to introduce the following matrices of Green’s function values:

\[
G_{RS} = \{ G(x_r, x_s) \}, \quad \text{dim}(G_{RS}) = N_r \times N_s,
\]

\[
G_{VS} = \{ G(x, x_s) \}, \quad \text{dim}(G_{VS}) = N \times N_s,
\]

\[
G_{VV} = \{ G(x, x') \}, \quad \text{dim}(G_{VV}) = N \times N,
\]

\[
G_{RV} = \{ G(x_r, x) \}, \quad \text{dim}(G_{RV}) = N_r \times N,
\]
The sensitivity (Fréchet derivative) matrix is defined to be the linear operator that relates small perturbations in the model parameters to the corresponding perturbations in the observable waveforms. In order to compute the sensitivity matrix that represents the Fréchet derivative in terms of the above Green’s functions, we linearize the forward model around a heterogeneous reference model with scattering potential $V^{(b)}$ (typically associated with the inverted model from a previous iteration) by using the so-called distorted Born approximation (Chew and Wang (1990); Tao and Sen (2013); Jakobsen and Ursin (2015))

$$\delta d_{RS}^{\text{obs}} - G_{RV}^{(b)} T^{(b)} G_{VS}^{(0)} S = G_{RV}^{(b)} \left( V - V^{(b)} \right) G_{VS}^{(b)} S.$$  \hspace{1cm} (10)

where

$$G_{RV}^{(b)} = G_{RV}^{(0)} + G_{RV}^{(0)} T^{(b)} G_{VV}^{(0)}$$

$$G_{VS}^{(b)} = G_{VS}^{(0)} + G_{VV}^{(0)} T^{(b)} G_{VV}^{(0)}$$

and

$$T^{(b)} = \left( I - V^{(b)} G_{VV}^{(0)} \right)^{-1} V^{(b)}$$  \hspace{1cm} (11)

is the T-matrix of the heterogeneous reference model, which in the iterative Bayesian inversion method described in the next section is equal to the inverted model from the previous iteration. In equation (10), the two first terms on the left hand side denotes observed and calculated values of the scattered wavefield data; that is, the difference between the total wavefield and the wavefield in a static background model for which the Green’s functions are known. For our synthetic example the observed data is calculated as

$$\delta d_{RS}^{\text{obs}} = G_{RV}^{(0)} T^{(\text{true})} G_{VS}^{(0)} S + \eta_{RS},$$  \hspace{1cm} (12)

where $\eta_{RS}$ is a stochastic noise term. This implies that our noise model (to be discussed later) is also for the scattered wavefield rather than the total wavefield.

Jakobsen and Ursin (2015) derived an explicit formula for the sensitivity matrix, depending only on the source function $S$ and the two Green’s matrices $G_{RV}^{(b)}$ and $G_{VS}^{(b)}$. In this study, we have essentially used the same data sensitivity matrix as Jakobsen and Ursin (2015). However, we have also introduced an additional coordinate transformation, which allows us to use the natural logarithm of the velocity field, rather than the contrast function defined in equation (5), as the fundamental model parameter field. This parametrization fits better with our Bayesian approach that will be presented later.

The explicit representation of the sensitivity matrix in terms of the Green’s functions for the reference medium will be used to develop a Bayesian version of the distorted Born iterative method introduced by Chew and Wang (1990). The most expensive part of our direct iterative T-matrix inversion algorithm is the updating of the T-matrix after each nonlinear iteration based on solving equation (11), whose computational solution time scales like $N^3$. The computational cost may be reduced by the use of a new domain decomposition method for larger 2D and 3D models, which is currently under development (see Jakobsen and Wu (2016); Wang et al. (2017)). The asymptotic computational cost will then be $N^{2.808}$, consistent with Strassen’s method (Strassen (1969)).

**Bayesian inversion**

We aim for finding a Bayesian solution to the inverse problem of finding the velocity field given the observed seismic traces. An important reason for choosing a Bayesian solution, is the fact that the solution of the seismic inverse problem will be part of the input for joint history matching of seismic and production data. In such an inversion, an ensemble based approach seems a natural choice. The ensemble based approach requires information about the uncertainty in the seismic data. We expect that a proper uncertainty quantification will be even more important in a case where two very different data...
types as seismic and production data are available. Moreover, the better the uncertainty quantification of the seismic data, the better uncertainty quantification would be expected after history matching the reservoir model. We will first outline the Bayesian approach in general terms, but the concrete example the reader might have in mind would be that our prior distribution consists of a spatially varying log-normally distributed velocity field, and the available data are seismic data corresponding to a set of frequencies. For the observed seismic signal, we assume that the observation noise has a multinormal distribution. Using a log-normal distribution as the prior distribution for the velocities was also done in Buland and Omre (2003).

A Bayesian solution to the inverse problem starts out by determining a prior density \( p(m) \) for the model parameters \( m \). Given an observation \( Y \) and a likelihood function \( p(Y|m) \), the sought posterior distribution is given as

\[
p(m|Y) = \frac{p(Y|m)p(m)}{\int p(Y|m)p(m)dm}.
\]  

(13)

In our case, we will assume that both the prior distribution, \( p(m) \), and the observation noise are Gaussian. With these assumptions, we follow (Tarantola, 2005, Section 3.2.1) to obtain

\[
p(m|Y) \propto \exp \left( -\frac{1}{2} ((y - h(m))^T R^{-1} (y - h(m)) + (m - m_{\text{prior}})^T P^{-1} (m - m_{\text{prior}})) \right).
\]  

(14)

Here \( m_{\text{prior}} \) denotes the mean and \( P \) the covariance matrix which describes the Gaussian prior distribution \( p(m) \). The measurement function \( y = h(m) + e \) relates the model parameter \( m \) to the available data. Here \( e \) is the observation error, which is assumed to be normally distributed with zero mean and covariance matrix \( R \).

For a linear measurement function \( h(m) \), the posterior distribution will also be Gaussian and a closed form solution exists. In our case, there is a nonlinear relationship between the model parameters \( m \) and the observations, i.e. \( h(m) \) is nonlinear. Then a closed form of the posterior distribution \( p(m|Y) \) is out of reach. However, assuming that a reasonable linear approximation to \( h(m) \) can be made, we again follow Tarantola (2005) and assume that the posterior distribution \( p(m|Y) \) can be approximated by a Gaussian with mean, \( \hat{m} \), equal to the maximum a posteriori (MAP) solution which maximizes the quadratic form in equation (14) and covariance matrix \( \hat{P} = (H^T R^{-1} H + P^{-1})^{-1} \) where \( H = \frac{\partial h}{\partial m}|_{m=\hat{m}} \) is the sensitivity matrix at \( \hat{m} \). The suitability of this approximation of course depends on the non-linearity of \( h(m) \).

The typical workflow for deterministic seismic full waveform inversion is to process the frequencies sequentially, as the problem is less non-linear for low frequencies, and the risk for getting trapped in a local minimum is thereby reduced. Also for the Bayesian case there is support for a sequential approach, although in a slightly different setting, see Fossum and Mannseth (2014). In a sequential setting, the observation, \( Y \), is divided in batches \( Y_1, Y_2, \ldots, Y_n \).

An approach for solving a linear problem with data arriving in batches in time is the Kalman filter, see Kalman (1960). Although, our data is obtained simultaneously in time, we can still process them in batches and rely on the theory developed for the Kalman filter and its variants. Since we have a nonlinear problem, we will use a nonlinear generalization of the Kalman filter. The extended Kalman filter (EKF), which is the most basic nonlinear version of the Kalman filter, approximates \( h \) by a first order Taylor expansion and uses the same update formulas as the Kalman filter. If \( h \) is strongly nonlinear the approximation can be poor and the filter may diverge.

In our case, we have chosen to use an iterated extended Kalman filter (IEKF), see for example Bell and Cathey (1993); Skoglund et al. (2015). Starting out with our prior distribution \( p(m) \), we can use equation (13) repeatedly, where \( p(m|Y) \) serves as the prior while processing the next set of data (i.e. a new set of seismic frequencies). In Bell and Cathey (1993) it was shown that the IEKF measurement update is the same as a Gauss-Newton update, and the EKF is a special case of the IEKF with only one Gauss-Newton iteration. Here we will use a Levenberg-Marquardt version of the iterated extended Kalman filter, see for example Skoglund et al. (2015).
Let $Y_{k,...,1}$ denote the collection of the $k$ first batches of data. A recursive solution of the Bayesian problem can then be obtained by using the solution $p(m|Y_{k-1,...,1})$ obtained after utilizing the $k-1$ first batches of data to find the a new solution $p(m|Y_{k,...,1})$ that also accounts for the information available by adding the $k$'th batch of data. This solution is given

$$p(m|Y_{k,...,1}) \propto p(y_k|m)p(m|Y_{k-1,...,1}),$$

where terms not depending on $m$ have been dropped, see Skoglund et al. (2015). Assuming that $p(m|Y_{k-1,...,1})$ can be approximated by a normal distribution with mean $\hat{m}_{k-1}$ and covariance $P_{k-1}$, we get

$$p(m|Y_{k,...,1}) \propto \exp\left(-\frac{1}{2} \left( (y_k - h(m))^T R_k^{-1} (y_k - h(m)) + (\hat{m}_{k-1} - m)^T P_{k-1}^{-1} (\hat{m}_{k-1} - m) \right) \right).$$

Then we can find the maximum a posteriori (MAP) estimate as

$$\hat{m}_k = \arg\min_m \left( (y_k - h(m))^T R_k^{-1} (y_k - h(m)) + (\hat{m}_{k-1} - m)^T P_{k-1}^{-1} (\hat{m}_{k-1} - m) \right).$$

Let $H_i = \frac{\partial h}{\partial m}|_{m=m_i}$ be the sensitivity matrix at $m = m_i$. For fixed $k$, we iterate using the following Levenberg-Marquardt update until some stopping criteria is met:

$$m^{t+1} = m^t + (H_i^T R_k^{-1} H_i + P_{k-1}^{-1} + \lambda I)^{-1} (H_i^T R_k^{-1} (d - H m^t) + P_{k-1}^{-1} (\hat{m}_{k-1} - m^t)).$$

(15)

Here, $\lambda$ is a damping factor that is adjusted at each iteration, and $I$ is the identity matrix. Usually one starts with a large $\lambda$ to make the first step small for increased stability. After each iteration the objective function is checked, and if it has decreased, the step is accepted and $\lambda$ is reduced. If not, the calculation is redone with a larger $\lambda$. When convergence is obtained, we let $\hat{m}_k = m^t$, and the covariance matrix is also updated:

$$P_k = (H_i^T R_k^{-1} H_i + P_{k-1}^{-1})^{-1}.$$  

(16)

To relate $h(m)$ to the seismic theory presented earlier, we must handle the fact that the model variables, $m$, for the Bayesian problem is defined in terms of log-velocity, whereas the model variable for the seismic forward modeling is the perturbation of the squared slowness defined in Eq. (5). The sensitivity required for the seismic signal with respect to the model variable $m$ can be obtained with the use of the chain rule.

**Numerical experiments**

**The model**

Our model is a two-dimensional vertical cross section. The size of the target is 1200 m vertically and 3000 m horizontally. We use gridblocks of size 20 m, thus the number of gridblocks is $60 \times 150 = 9000$. We assume that the velocities in the target are log-normally distributed. The logarithm of the velocities in Figure 1 was used as mean of the prior distribution. The velocities in this figure range between 2500 m/s at the top and 3200 m/s at the bottom. The covariance matrix for the prior distribution was constructed from an exponential variogram with a practical range of 50 gridblocks vertically and 300 horizontally. (The practical range for an exponential variogram is defined as the distance where 95% of the sill is reached, see Remy et al. (2009).) The true model is a realization from the prior distribution, see Figure 2a. The surroundings of the target are assumed to have a constant known velocity of 2600 m/s, and this velocity is used as background velocity ($c_0$ in equation (5)).

**Seismic forward modelling**

We use equation (12) for each frequency to make the observed data. A Ricker wavelet with center frequency 10 Hz is used as source. We use 60 sources and 150 receivers evenly distributed 20 m above the target. The reflected waves are sampled every 4 ms for 4 s. The sampling rate gives frequencies up
Figure 1 The logarithm of the velocities in this figure is the mean of the prior distribution, and is also used as starting model for the inversion. The velocities range from 2500 m/s at the top to 3200 m/s at the bottom.

Figure 2 True velocities, inverted velocities (MAP estimate) and error in the baseline survey.
to 125 Hz, but we will only use the frequencies from 4-18 Hz for the inversion, that is 15 frequencies in total. We add complex Gaussian white noise to the waveforms in the frequency domain with a signal to noise ratio 3. We use the following definition of signal to noise ratio $r$:

$$r = \frac{\|Y\|^2}{\|w\|^2},$$

where $Y$ is the data and $w$ is the noise. Let $Y$ be the data of the frequencies used for inversion and let $v_1$ and $v_2$ be vectors of the same length as $Y$ of normally distributed random numbers, then the white noise is constructed as $w = v_1 + iv_2$. We get the desired signal to noise ratio ($r$) by scaling it in the following way:

$$Y_{\text{noisy}} = Y + \frac{\|Y\|}{\sqrt{(r \cdot E(\|w\|^2))}} w,$$

where $E$ denotes the expectation. Since $\|w\|^2$ is sum of the squares of independent standard normal random variables it follows a chi-squared distribution with the number of degrees of freedom equal to the number of variables.

**Baseline survey**

The covariance matrix $R$ for the noise in the seismic waveforms is set diagonal with a constant value on the diagonal equal to $\|Y\|^2/(r \cdot E(\|w\|^2))$, see (17). The mean of the prior distribution (see Figure 1) was used as starting model for the inversion. We used equation (15) and (16) for the updating. We tried updating using one frequency and three frequencies at a time. We always started with the lowest frequency or frequency group. As stopping criteria we used an improvement in the objective function of less than 0.2% and a change in $m$ of less than 1%. We noticed that it was important to start with a large $\lambda$ in the first iteration for the first frequency to make the first step small. We started with an initial $\lambda = 10000$ and reduced it by a factor of 10 after each iteration if the step was accepted. For the next frequency we could start with a smaller $\lambda$.

The results were a bit better when using three frequencies at a time, so we will only show the figures for this case. We have not tested other frequency selection strategies. The inverted velocities (the MAP estimate) are shown in Figure 2b and the error in Figure 2c. Most of the features are captured in the inversion. The initial standard deviation in the prior was approximately 0.16. After the inversion the average standard deviation was 0.025; see Figure 3 for the standard deviation after the baseline survey and error in log-velocity.

**Monitor survey**

To illustrate a time-lapse effect we made a model which was equal to the first one except for a perturbation in a small area, see Figure 4a. The perturbation is shown in Figure 5a. The velocity is reduced, and is meant to illustrate the velocity reduction that can come from gas injection. We use the same frequencies for the second inversion and add noise in the same way (but different noise).

There are several ways to choose a prior for the second inversion. The inversion can be done independently from the baseline survey, or we can use some or all of the information from the first inversion.
**Figure 4** True velocities, inverted velocities (MAP estimate) and error in the monitor survey.

**Figure 5** True and estimated difference between the two seismic surveys, and error in the difference.
We investigated four ways of choosing the prior:

1. Using the same prior as for the baseline survey.
2. Using the posterior mean from the baseline survey as prior mean, but the same prior covariance matrix as was used for the baseline survey.
3. Using both the posterior mean and covariance matrix from the baseline survey as prior.
4. Using the posterior mean from the baseline survey, and a modified version of the posterior covariance with increased uncertainty in areas where we expect there can be velocity changes due to fluid flow.

Of these, number 4 gave the best results when evaluating the norm of the difference between the true and estimated change, and we will only show the results for this case. We modified the posterior covariance from the first inversion by adding a matrix that was made with the same variogram as the prior covariance matrix and scaled by multiplying with a diagonal matrix $D$ that is gradually reducing away from a center point where the injection is imagined to take place (i.e. $P = P_{\text{posterior}} + DBD$), where $B$ is similar to the prior covariance, but with 1 on the diagonal, see Figure 6. In future work we plan to use our inversion method as part of an ensemble based approach for history matching, and then a more natural way of adding extra uncertainty would be to estimate it from an ensemble of reservoir simulations as was done in Eikrem et al. (2016). In Eikrem et al. (2016) the mean of the ensemble after using the reservoir simulator was also used to update the posterior mean from the first inversion before using it as prior for the second inversion. Here we used the posterior mean without change.

The inverted velocities are shown in Figure 4b and the error in Figure 4c. The average posterior standard deviation after the monitor survey was 0.024, as opposed to 0.032 before the survey; see Figure 7 for the standard deviation and error in log-velocity. The true and estimated time-lapse change are shown in Figure 5. The velocity reduction is clearly visible in the figure.
Conclusion

We have developed a non-linear full waveform inversion method that also gives an estimate of the uncertainty in the inversion. We used a Levenberg-Marquardt version of the iterated extended Kalman filter to update the estimated log-velocities using three frequencies at a time. Our method is an extension of the method developed in Jakobsen and Ursin (2015). We performed two surveys, and investigated the change in velocity between the two surveys. The main features of the models were captured in the inversions, and the velocity change was clearly visible when taking the difference between the surveys. We compared different choices of prior for the monitor survey, and the best results were obtained when we used the posterior mean from the baseline survey as prior mean, and a modified version of the posterior covariance with increased uncertainty in areas where we expected there could be velocity changes.

In future work, we will test the inversion method when history matching a model using ensemble based methods with both production data and time-lapse seismic data, as was done in Eikrem et al. (2016).

Acknowledgements

The authors acknowledge the Research Council of Norway and the industry partners; ConocoPhillips Skandinavia AS, Aker BP ASA, Eini Norge AS, Maersk Oil Norway AS, DONG Energy A/S, Denmark, Statoil Petroleum AS, ENGIE E&P NORGE AS, Lundin Norway AS, Halliburton AS, Schlumberger Norge AS, Wintershall Norge AS of The National IOR Centre of Norway for support. The research work performed by Morten Jakobsen was partially associated with Petromaks II project 225387 (Modelling and Inversion of seismic waveform and electromagnetic data using integral equation methods) funded by the Norwegian Research Council.

References


Muhumuza, K. [2015] *Modelling and inversion of time-lapse seismic waveform data using scattering


