A numerical study of 3D elastic time-lapse full-waveform inversion using multicomponent seismic data

Espen Birger Raknes[1] and Børge Arntsen[1]

ABSTRACT

A common assumption in wave-propagation problems is that the subsurface is approximately an acoustic medium. Under this assumption, important wave phenomena such as S-waves are not included. Due to the increase in computational power in recent years, the acoustic assumption may be left behind and replaced by the more physically correct elastic assumption. Time-lapse seismic data contain information about changes in the subsurface due to the production of hydrocarbons or injection of CO_2. Full-waveform inversion (FWI) is an inverse method that can be used to quantify these time-lapse changes in the subsurface. Using a 3D isotropic elastic implementation of the FWI method, we studied two strategies for performing time-lapse FWI. We used synthetic ocean-bottom multicomponent seismic time-lapse data to estimate changes in the P- and S-wave velocity models. A sensitivity analysis in which the sensitivities with respect to the magnitude and physical size of the time-lapse anomalies and the noise level in the data was performed. The strategy focusing on explaining the data differences between the baseline and monitor data sets provided fewer artifacts in the inverted elastic models than the strategy that tried to explain the full monitor data set, and it was therefore preferable. The data-difference strategy depends on good repeatability in the time-lapse data sets and sufficient convergence of the inversion of the baseline data set.

INTRODUCTION

A common assumption in wave propagation problems is that the subsurface is approximately an acoustic medium. Under this assumption, it is sufficient to use the acoustic wave equation (Aki and Richards, 2002) to model waves propagating through the medium of interest. Important wave phenomena such as S-waves are not included in the acoustic wave equation. In parameter estimation problems, the acoustic assumption may in the worst case lead to wrongly estimated models (Barnes and Charara, 2009; Raknes and Arntsen, 2014a). Due to the increase in computational power in recent years, we now can use the elastic wave equation (or other wave equations that involve more accurate physics) to model wave propagation in a medium. This equation is more complicated to solve than the acoustic wave equation, resulting in higher computational cost. More importantly, this leads to better agreement between simulated data and the real world wavefields, which may improve the results of methods relying on waveform fitting. Furthermore, it is now possible to numerically solve the elastic wave equations in a full 3D setup with acceptable computing times.

Time-lapse seismic data contain information about changes in the subsurface and have proven to be an effective tool in reservoir imaging and for monitoring of injected CO_2 in the subsurface (Biondi et al., 1996; Lumley et al., 2003). Conventional methods for quantifying time-lapse effects assume that the monitor model can be approximated by a linear perturbation of the baseline model (Greaves and Fulp, 1987; Landrø et al., 1999). This assumption is easily violated in cases in which there are strong changes in the rock properties between the baseline and monitor parameter models. Therefore, more robust time-lapse analysis methods are required.

The full-waveform inversion (FWI) method is a technique for estimating parameters affecting wave propagation (Tarantola, 1984; Mora, 1987). The method has been applied on synthetic and real data sets with success (Virieux and Operto, 2009). Because FWI in general is computer intensive and therefore time consuming, the method has mainly been applied using a 2D computational setup. Under these conditions, several modifications must be made to the real data sets to minimize 3D effects. The modifications may introduce artifacts in the inverted parameter models (Auer et al., 2013). Therefore, a 3D setup is preferable.

Increases in computational power over the last decade have made it possible to perform FWI using full 3D computational grids.

Sirgue et al. (2008) use a frequency-domain implementation to demonstrate FWI in 3D using a synthetic data set, whereas Vigh and Starr (2008) use a time-domain plane wave implementation and synthetic and real data sets in 3D. Plessix (2009) uses a frequency-domain implementation on synthetic and real ocean-bottom cable (OBC) data sets. Sirgue et al. (2009) use 3D FWI on a wide-azimuth real OBC data set from the Valhall area. Abubakar et al. (2011) use the contrast-source method in 3D to invert synthetic data sets. The previous examples assumed purely acoustic media, such that the acoustic wave equation was used to approximate the waves in the media. Recent developments have included more complicated wave phenomena in the modeling. Warner et al. (2013) apply a full 3D anisotropic pseudoacoustic implementation of FWI on a multicomponent ocean-bottom survey over the Tommeliten field in the North Sea. Butzer et al. (2013) apply a 3D isotropic elastic FWI on a small-scale cross-well acquisition geometry using synthetic and real data sets. Vigh et al. (2014) apply the same 3D wave equation using synthetic and real OBC data sets.

The reconstruction of time-lapse anomalies using FWI is a two-step procedure. The first step is the estimation of the baseline model, and the second step is the estimation of the monitor model. In the literature, at least three different strategies for performing time-lapse FWI (TLFWI) have been suggested (Zheng et al., 2011). The three strategies vary in how the second step is performed. The first strategy is formed by inverting independently for the baseline and monitor models starting from the same initial model. In cases in which the time-lapse changes are local, this is not a good idea from a computational point of view because FWI will use many iterations to explain already known events that are included in the inverted baseline model. The second strategy is therefore formed by using the inverted baseline model as the initial model for the monitor inversion. This strategy is called the *sequential strategy*. The third strategy, called the *data-difference strategy*, is formed by modifying the monitor data set such that the inversion tries to explain the true time-lapse data differences (Watanabe et al., 2004). The 2D implementations of the three strategies have been applied on the synthetic and real data sets with success (Zheng et al., 2011; Routh et al., 2012; Queißer and Singh, 2013; Zhang et al., 2013; Raknes and Arntsen, 2014a).

Even though TLFWI has been applied on the synthetic and real data sets, the advantages and disadvantages for the different time-lapse strategies are not well understood, particularly not in 3D. Because the number of iterations should be held to a minimum in 3D, we study the sequential and data-difference time-lapse strategies using a 3D isotropic elastic implementation of FWI. We use synthetic ocean-bottom multicomponent seismic time-lapse data to estimate time-lapse changes in the P- and S-wave velocity models. We perform a sensitivity analysis in which the sensitivities with respect to the magnitude and physical size of the time-lapse anomalies and the noise level in the data are studied. We find that both strategies are able to reveal the time-lapse anomalies, but the sequential strategy introduces more artifacts than the data-difference strategy in all examples. Hence, inverting for the data differences is preferred as long as the inverted baseline model is sufficiently close to the true baseline model, and the repeatability is good in the data sets.

THEORY AND METHODOLOGY

The theory that underlies FWI has been derived several times using different formulations. Here, we present only the results that are required to understand how our implementation works, and which assumptions we made during the implementation of the method. We refer to Pratt (1999), Fichtner et al. (2006), and Virieux and Operto (2009) for more details about the method.

Full-waveform inversion

The overall goal for FWI is to find a parameter model $\mathbf{m}$ that produces modeled data $\mathbf{q}$ that is close to some measured data $\mathbf{d}$. The foundation for the method is the assumption that synthetic data $\mathbf{q}$ can be generated using a numerical wave equation. Let $\mathcal{L}$ be the numerical wave operator that maps $\mathbf{m}$ from the model space into the data space. Then, the synthetic data can be generated as

$$\mathcal{L}(\mathbf{m}) = \mathbf{q}. \tag{1}$$

If the inverse operator of $\mathcal{L}$, that is, the mapping from the data space into the model space, exists, then the solution to the problem is given as

$$\mathbf{m} = \mathcal{L}^{-1}(\mathbf{d}), \tag{2}$$

where $\mathcal{L}^{-1}$ is the inverse operator. In practice, however, it is not possible to find an explicit expression for the inverse operator.

The standard way of solving the inverse problem is to define a measure, denoted as $\Psi(\mathbf{m})$, between $\mathbf{q}$ and $\mathbf{d}$. This measure is often called the *objective (or misfit) functional*. We require that the solution of the problem, that is, the point where $\mathbf{q}$ and $\mathbf{d}$ are equal, is an extreme point for $\Psi(\mathbf{m})$. Hence, the solution to the problem can be expressed as

$$\mathbf{m}' = \arg\min_{\mathbf{m}} \Psi(\mathbf{m}), \tag{3}$$

where $\mathbf{m}'$ is the model we are searching for. The inverse problem in equation 3 is nonlinear and ill posed.

The search for the extreme points of $\Psi(\mathbf{m})$ is done using an iterative optimization algorithm, written as

$$\mathbf{m}_{k+1} = \mathbf{m}_k - \alpha_k \mathbf{H}_k^{-1} \mathbf{g}_k, \tag{4}$$

where $\alpha_k > 0$ is the step length, $\mathbf{H}_k^{-1}$ is the inverse Hessian matrix, and $\mathbf{g}_k$ is the gradient of $\Psi(\mathbf{m})$ with respect to $\mathbf{m}$ at step k. To start the algorithm, an initial model $\mathbf{m}_0$ is required, and the algorithm is run until some convergence criteria are fulfilled.

In practice, the inverse Hessian matrix in equation 4 is complicated to compute because it involves second-order derivatives of the objective functional. To overcome this problem, we use the limited-memory Broyden, Fletcher, Goldfarb, and Shanno (L-BFGS) algorithm (Nocedal and Wright, 2006), which is a quasi-Newton method that tries to estimate the inverse Hessian matrix using a predefined number of gradients from previous iterations.

Time-lapse full-waveform inversion

The purpose of TLFWI is to reveal directly changes over time in the parameter models. Thus, at least two inversions must be performed, and the time-lapse images are obtained by subtracting the inverted parameter models obtained from each of the individual inversions. We present two time-lapse strategies that are different in the way the monitor inversion is performed.

We assume that we have two multicomponent data sets $\mathbf{d}_{\text{base}}$ and $\mathbf{d}_{\text{mon}}$, where $\mathbf{d}_{\text{base}}$ and $\mathbf{d}_{\text{mon}}$ are the data sets from the baseline and monitor surveys, respectively. The data sets are assumed to have been acquired using identical source-receiver geometries and source functions. In this sense, the repeatability of the data sets should be good.

The starting point in TLFWI is the inversion for the baseline data set. For the baseline inversion, we use the least-squares objective functional

$$\Psi(\mathbf{m}) = \frac{1}{2} \sum_{(s,r) \in \mathbf{S}} \left\| \mathbf{W} \left[\begin{pmatrix} \mathbf{q}_p \\ \mathbf{q}_v \end{pmatrix} - \begin{pmatrix} \mathbf{d}_p \\ \mathbf{d}_v \end{pmatrix} \right] \right\|^2, \qquad (5)$$

where $\mathbf{S}$ is the discrete set of all receiver and source enumerations, $\mathbf{W}$ is a data weighting matrix, $\mathbf{q}_p = \mathbf{q}_p(\mathbf{m}; \mathbf{x}_r, \mathbf{x}_s)$ is the modeled pressure data for a source at position $\mathbf{x}_s$ and receiver at $\mathbf{x}_r$ for source s and receiver r, $\mathbf{q}_v = \mathbf{q}_v(\mathbf{m}; \mathbf{x}_r, \mathbf{x}_s)$ is the modeled particle velocity data at the same position, $\mathbf{d}_p = \mathbf{d}_p(\mathbf{m}; \mathbf{x}_r, \mathbf{x}_s)$ is the measured pressure data, and $\mathbf{d}_v = \mathbf{d}_v(\mathbf{m}; \mathbf{x}_r, \mathbf{x}_s)$ is the measured particle velocity data at the same position.

It is important to let the pressure and particle velocity data contribute equally to the objective functional. By considering the units of the data components, the scale matrix in equation 5 is given as

$$\mathbf{W} = \begin{pmatrix} \mathbf{I}_1 & 0 \\ 0 & \rho V_P \mathbf{I}_3 \end{pmatrix}, \qquad (6)$$

where $\mathbf{I}_1$ is a 1×1 identity matrix, $\mathbf{I}_3$ is a 3×3 identity matrix, and ρ and V_P are the density and P-wave values at the receiver positions, respectively.

If we assume that the time-lapse changes are local and that the inverted baseline model is close enough to the true baseline model, it is best to use the inverted baseline model as an initial model for the monitor inversion, instead of using the same initial model for both. Thus, the first time-lapse strategy is formed by following this assumption, in which the objective functional in equation 5 is used in the monitor inversion. In the following, this scheme is called the *sequential strategy* (Figure 1a).

The time-lapse changes sought for may be small, and in some cases, the total error from the modeling and inversion is bigger than the actual time-lapse changes. By considering the error sources between the time-lapse data sets, some of the error sources can be avoided. Let $\mathbf{d}$ be one of the components from the recorded data set. By applying $\mathcal{L}$ in equation 1 on a given model $\mathbf{m}$, we may assume that

$$\mathbf{d} = \mathcal{L}(\mathbf{m}) + \varepsilon_n(\mathbf{m}) + \varepsilon_s + \varepsilon_r, \qquad (7)$$

where $\varepsilon_n(\mathbf{m})$ is the numerical error, ε_s is a static error, and ε_r is the random error. The numerical error is due to an imperfect forward operator in addition to an imperfect model. In real data sets, the static error term is a time-shift error that is constant in the data sets and easy to remove. Therefore, we assume that ε_s is equal for all time-lapse data. The random error corresponds to observational errors such as noise from other vessels in the acquisition area, noise from the water surface waves, and differences in the acquisition equipment, among others. The magnitude of the different error sources may vary for different source and receiver geometries.

By modifying all components in the measured monitor data set as follows (Zheng et al., 2011):

$$\hat{\mathbf{d}}_{\text{mon}} = \mathbf{q}_{\text{base}} + (\mathbf{d}_{\text{mon}} - \mathbf{d}_{\text{base}}), \qquad (8)$$

we observe, by using the assumption in equation 7 for $\mathbf{d}_{\text{mon}}$ and $\mathbf{d}_{\text{base}}$, that

$$\begin{aligned} \hat{\mathbf{d}}_{\text{mon}} &= \mathcal{L}(\mathbf{m}_{\text{mon}}) + (\varepsilon_n(\mathbf{m}_{\text{mon}}) - \varepsilon_n(\mathbf{m}_{\text{base}})) \\ &\quad + (\varepsilon_{r,\text{mon}} - \varepsilon_{r,\text{base}}) \\ &= \mathcal{L}(\mathbf{m}_{\text{mon}}) + \Delta\varepsilon_n + \Delta\varepsilon_r. \end{aligned} \qquad (9)$$

Thus, by comparing with the assumption in equation 7, ε_s is removed from this new data set and we are left with the differences in the numerical and random error terms. The consequence of the modification is that the monitor inversion does not try to explain the error from an imperfect inverted baseline model because this error is removed in the new data set by the $\Delta\varepsilon_n$ term. If we assume that $\Delta\varepsilon_n$ and $\Delta\varepsilon_r$ are small compared with the time-lapse data differences, the most important consequence of the modification is that the monitor inversion focuses on explaining the true data differences. The objective functional for the monitor inversion is

$$\Psi(\mathbf{m}) = \frac{1}{2} \sum_{(s,r) \in \mathbf{S}} \left\| \mathbf{W} \left[\begin{pmatrix} \mathbf{q}_p \\ \mathbf{q}_v \end{pmatrix} - \begin{pmatrix} \hat{\mathbf{d}}_p \\ \hat{\mathbf{d}}_v \end{pmatrix} \right] \right\|^2, \qquad (10)$$

with $\mathbf{W}$ as given in equation 6. The initial model for the monitor inversion is, as above, the inverted baseline model. In the following, this scheme is called the data-difference strategy (Figure 1b).

Implementation

Modeling

We assume that the subsurface is an isotropic elastic medium, so that $\mathcal{L}$ in equation 1 is the 3D isotropic elastic wave equation (Aki and Richards, 2002). Numerically, we solve the equation using a 3D implementation of the time-domain staggered-grid finite difference method described in Virieux (1986), in which we use high-order spatial differential operators (Holberg, 1987). We use regular cubic grids in which the source function is located at one single point in the grid. To simulate an unbounded half-space (nonreflecting boundaries), we use a 3D implementation of the perfectly matched layer
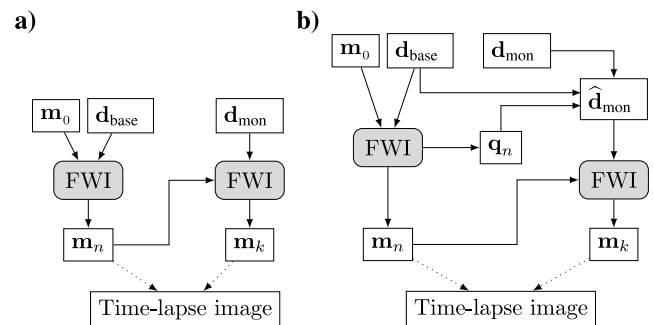

Figure 1. Schematic workflow of TLFWI. (a) The sequential strategy and (b) the data-difference strategy.

absorbing boundary conditions (Berenger, 1994; Zhen et al., 2009). The reflecting sea surface at the top of the model is approximated using the method described in Mittet (2002).

Gradient computation

The crucial step in FWI is the computation of the gradient in equation 4. Using the adjoint state method (Tarantola, 1984; Mora, 1987), the gradients for the P-wave V_P and S-wave V_S model for a single shot are given as (Mora, 1987)

$$g_p = -2\rho V_P \int_T (\nabla \cdot \mathbf{u}^f)(\nabla \cdot \mathbf{u}^b) dt, \qquad (11)$$

$$g_s = -2\rho V_S \int_T [\nabla \mathbf{u}^f : (\nabla \mathbf{u}^b + (\nabla \mathbf{u}^b)^T) - 2(\nabla \cdot \mathbf{u}^f)(\nabla \cdot \mathbf{u}^b)] dt, \qquad (12)$$

where $\mathbf{u}^f = \mathbf{u}^f(\mathbf{x}, t)$ is the particle displacement vector for the forward field propagating from the source position, $\mathbf{u}^b = \mathbf{u}^b(\mathbf{x}, t)$ is the particle displacement vector for the backward field propagating from the receiver positions, $\rho = \rho(\mathbf{x})$ is the density, $V_P = V_P(\mathbf{x})$ is the P-wave velocity, and $V_S = V_S(\mathbf{x})$ is the S-wave velocity. In the above equations, $\nabla \mathbf{u}^i$ is the Jacobian matrix and ":" is the Frobenius matrix inner product operator.

From equations 11 and 12, we observe that the V_P gradient depends only on normal stresses, whereas the V_S gradient depends on normal and shear stresses. From a computational point of view, we observe that to compute the gradients, snapshots of the wavefields at each time step are required for the forward and backward propagated wavefields. In two dimensions, a normal procedure is to store the required wavefields during the forward modeling. In three dimensions, however, this is not possible due to the extreme amounts of disk space needed when working on normal-sized models.

To overcome this problem, we need to reconstruct the forward wavefields in equations 11 and 12 when they are needed. During the forward modeling, we store the wavefields at the boundaries of the grid, i.e., all six sides of the computational cube. In addition, we save the last snapshot of all the wavefields in the complete model at the last time step, such that waves inside the model are not lost when the reconstruction starts. When we compute the backward propagating wavefields, we reconstruct the forward propagated wavefields using the saved wavefields at the boundaries (McMechan, 1983). The crosscorrelation in equations 11 and 12 is computed during the time stepping of the two equations. Thus, the gradient computation is twice as costly as the forward modeling itself. To conclude, to compute the modeled data and the gradients, we need to numerically solve the elastic wave equation three times.

Algorithm complexity

Let n_i be the number of grid points on the spatial axis i, and assume that $n_i = n$ for an arbitrary number n. The computational cost of performing the forward modeling and the gradient computation for a single shot is $\mathcal{O}(n^3 \cdot n_t)$, where n_t is the number of time steps. Because we only store the boundaries for the reconstruction of the forward wavefield, the storage needed for this operation is $\mathcal{O}(n^2 \cdot n_t)$ for a single shot. As a result, if n_s is the number of shots in the data sets, the total computational cost of our implementation is $\mathcal{O}(n^3 \cdot n_t \cdot n_s)$, and the necessary storage is $\mathcal{O}(n^2 \cdot n_t \cdot n_s)$. Thus, as the model size increases, the computational cost and the required storage increase severely.
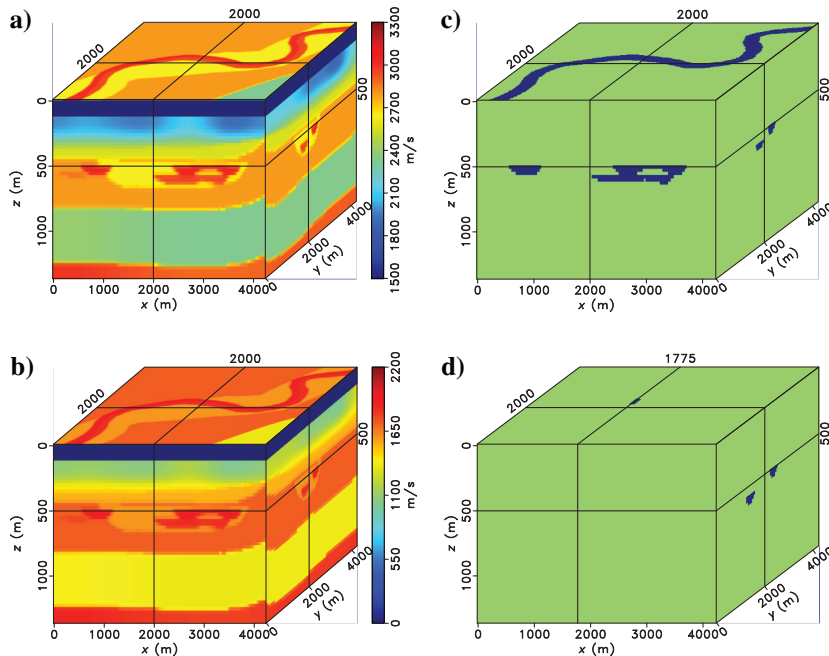
RESULTS

We study the behavior of the two time-lapse strategies using a synthetic model adapted from the SEG/EAGE overthrust model (Figure 2). Using this model, the sensitivity with respect to noise, the magnitude of time-lapse anomaly, and the physical size of the time-lapse anomaly are studied. For the first two test cases, the target zone is a channel system at 500–600-m depth, in which a large time-lapse effect in the whole channel system is introduced (Figure 2c). For the third test case, a small and local time-lapse effect consisting of two anomalies in the channel system is used (Figure 2d). For all tests, the density assumes the constant value of 1000 kg/m³. The size of each grid cell for the model is $25 \times 25 \times 25$ m³, resulting in a computational grid of $65 \times 190 \times 190$ grid points. Thus, the total number of grid points is 2,346,500 for each parameter model. To avoid numerical aliasing, the source signature is a Ricker wavelet with center frequency of 6.0 Hz. We assume that the source signature is known, such that the same signature is used both to generate the data and in the inversion. The initial model (Figure 3) is

Figure 2. The true models: (a) V_P, (b) V_S, (c) large time-lapse effect, and (d) small time-lapse effect. The position of the slices is indicated in the plots and is projected to the boundaries of the cubes.

made by smoothing the true models using a triangle smoothing operator.

The relative time-lapse changes in the parameter models in a real setting may be as small as 1%. Such small changes may be difficult to observe in seismic time-lapse data, and thus difficult to invert for. To study the sensitivity of TLFWI with respect to the magnitude of the time-lapse effect, we use monitor models in which the V_P and V_S velocities in the channel system are changed with 1%, 5%, and 10%, respectively. The changes in magnitude for the three cases are as follows:

1) **1%:** 0–30 m/s for V_P and 0–20 m/s for V_S,
2) **5%:** 0–152 m/s for V_P and 0–96 m/s for V_S,
3) **10%:** 0–304 m/s for V_P and 0–191 m/s for V_S.

We simulate a realistic OBC setup, in which 16 multicomponent cables with a length of 4 km are put on the seafloor. The crossline distance between each cable is 250 m, and the inline distance between each receiver is 25 m. The total number of receivers in the grid is 2560. We perform 441 shots on a square grid with a shot sampling of 125 m in both horizontal directions. The source depth is 25 m.

Data without noise

For monitoring purposes, it is important to obtain a good estimate of the baseline model because the model is used as the starting model for the monitor inversion, and thus for the reconstruction of the time-lapse anomalies. We invert simultaneously for V_P and V_S, and because the method is implemented in the time domain, all frequencies are included in the inversion. In addition, the full offset range is used in the inversion. Hence, one inversion run is performed to estimate the baseline V_P and V_S models.

The inverted baseline models for V_P and V_S are given in Figure 4a and 4b, respectively. We observe that the inversion is able to reveal the channel system and in general sharpen the initial model. The V_S model is sharper than the V_P model, which is expected due to the lower wave velocities for the S-waves. The inversion is not able to estimate sharp edges, and in these areas oscillations are clearly visible. This behavior can be explained by the relatively low frequencies used in the inversion. The quality of the inverted baseline models can be visualized using the normalized velocity error, given as

$$V^{err} = (V^{true} - V^{inv})/V^{true}. \qquad (13)$$

Figure 4c and 4d shows the normalized error for V_P and V_S, respectively. We observe that the error is smaller in the center of the model than at the edges of the model and that the error is at the largest in the bottom part of the model. This is expected because information about these areas is not present in the data.

The inverted time-lapse models for V_P and V_S for the two time-lapse strategies are given in Fig-

ures 5 and 6, respectively. The results show that the amount of artifacts is higher in the sequential strategy than in the data-difference strategy. For the 1% change, it is difficult to distinguish the time-lapse effects and the artifacts, particularly for the V_S model, for the sequential strategy. The artifacts are similar for the 5% and 10% changes, but because the magnitude of the time-lapse changes is higher, the artifacts are not dominant to the same extent. Because no noise is added in the data, the artifacts are due to the fact that the monitor inversion is trying to update the baseline model in the areas in which the time-lapse anomaly is placed, in addition to further update the areas of the model that were not well resolved in the baseline inversion. For the data-difference strategy, however, the error from the imperfect inverted baseline model is included in the new monitor data set (see equation 9), and therefore the monitor inversion is not trying to explain the nonperfect inverted baseline model. In Figure 7, the difference between the baseline and monitor data sets is shown for the 1% case. From this figure, we see that the time-lapse difference for the sequential strategy is far away from the
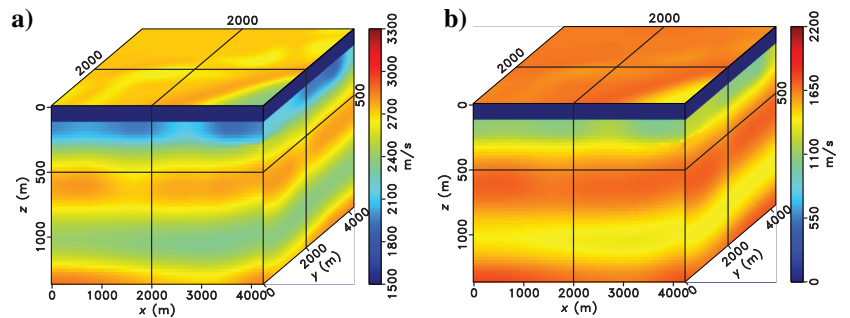

Figure 3. The initial model for the baseline inversion: (a) V_P and (b) V_S.
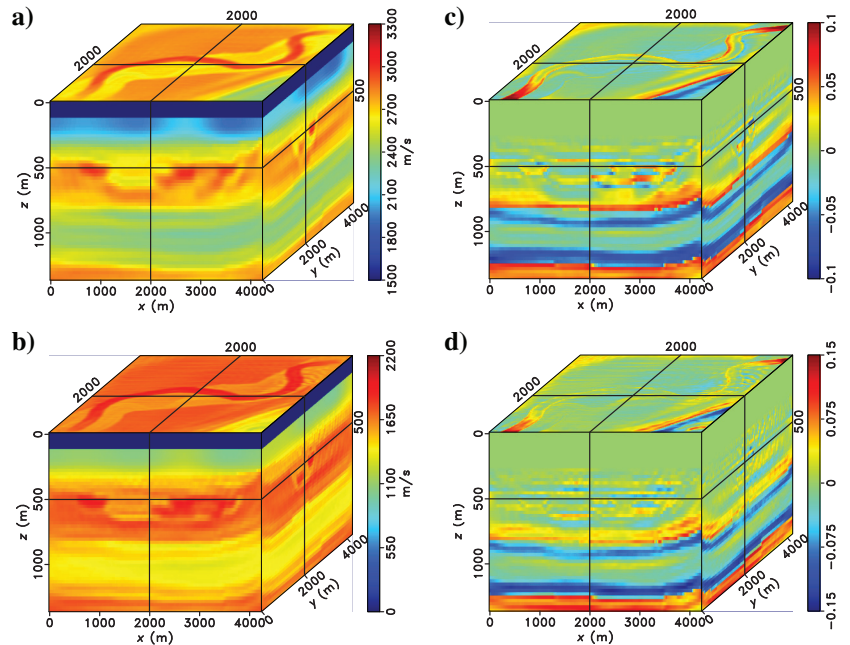

Figure 4. (a) Inverted V_P baseline model using data without noise, (b) inverted V_S baseline model using data without noise, (c) normalized V_P error, and (d) normalized V_S error (see equation 13).

true time-lapse difference, whereas the data-difference strategy is close to the true difference. By comparing the objective functionals (Figure 8) for the two strategies, we observe that the data-difference strategy reduces the objective functional more than the sequential counterpart, which is a result of the above-mentioned error present in the latter strategy.

Data with noise

We define the signal-to-noise ratio (S/N) as $S/N = rms_d/rms_n$, in which rms_d and rms_n are the root-mean-square values of the noise-free data and the noise, respectively. To simulate real noise, we create a data set with pure (pseudo-) random noise. A Gaussian distribution that uses a seed as input is used to generate the noise. To ensure that the noise is not repeated in the shots, different seeds are used for each shot in the baseline and monitor data sets. To color the noise, the noise data set is filtered using a bandpass filter with a window of 2–10 Hz before it is added to the data.

We use two noise levels to study the sensitivity with respect to the noise. The weak noise data sets are made by using $S/N = 5.0$, and the strong noise data sets are made by using $S/N = 2.0$. Figure 9 shows the different S/Ns for a pressure recording from the baseline data set, whereas Figure 10 shows the time-lapse data differences for the two S/Ns. We observe that with $S/N = 2.0$, the time-lapse data difference is more or less completely distorted (Figure 10b), whereas with $S/N = 5.0$, it is possible to see some parts of the time-lapse data difference, particularly between 1.0 and 1.5 s (Figure 10c). Because the 1% time-lapse model introduced a lot of artifacts for the sequential strategy for noise-free data, we use the 5% time-lapse model in the tests that follow.

The inverted time-lapse images for V_P are given in Figure 11, and those for V_S are given in Figure 12. We observe that as the noise level is increased the amount of artifacts are increased as well, which is expected. As in the noise-free examples, there are fewer artifacts with the data-difference strategy than with the sequential strategy The reason for the success of TLFWI for $S/N = 2.0$ is due to the fact that when data from all the shots are summed together in the stacking of the global gradient, the time-lapse information is repeated and thus dominates compared with the random noise.

Small-sized time-lapse anomaly

The spatial size of the time-lapse anomaly studied so far has been large. To test if TLFWI is able to resolve small spatial changes, we now use a model in which the time-lapse anomaly is two relatively small anomalies located in the channel system (Figures 2d and 13). The distance between the two anomalies is 500 m, and the sizes of the anomalies are approximately 150 m in all three directions. To simulate a real-world time-lapse anomaly, data with $S/N = 2.0$ are used, and the magnitude of the time-lapse anomaly is 5%.

From the final inverted models for the two time-lapse strategies in Figures 14 and 15, we observe that both strategies are able to detect both time-lapse anomalies. As in the previous examples, the sequential strategy introduces several artifacts that can be interpreted as time-lapse anomalies. The magnitudes of the artifacts are of the same order as the time-lapse anomalies sought for. The data-difference strategy, however, manages to reveal the two anomalies without distorting noise. The estimated magnitudes of the V_P anomalies are lower than the true anomalies. This can be explained by the low frequencies, and thus the wavelengths in the data. The low frequencies result in that the anomalies are smeared out in all directions. This behavior is also visible in the estimated V_S anomalies, but here, due to the shorter wavelengths, the images are sharper with magnitudes relatively closer to the true anomalies.

Computational aspects

The nodes on the computer cluster used in this study have 16 (two eight-core processors) cores (Intel Xeon E5-2670, 2.6 GHz processor speed)
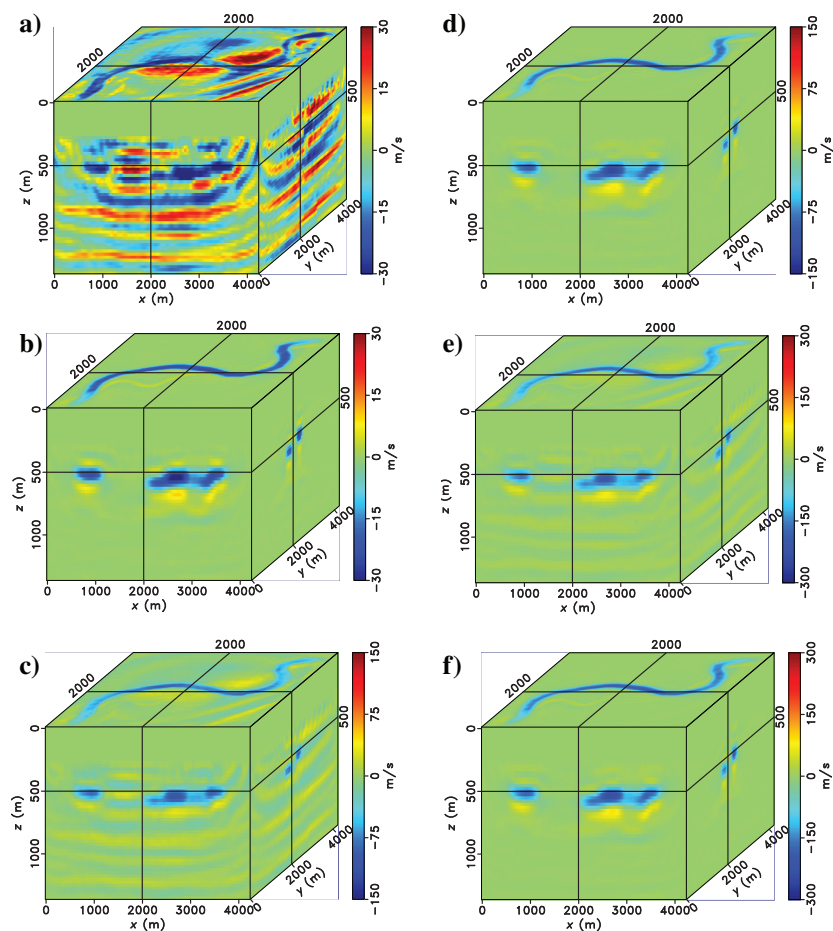

Figure 5. Inverted time-lapse images for V_P using data without noise. (a) 1% time-lapse effect using the sequential strategy, (b) 1% time-lapse effect using the data-difference strategy, (c) 5% time-lapse effect using the sequential strategy, (d) 5% time-lapse effect using the data-difference strategy, (e) 10% time-lapse effect using the sequential strategy, (f) 10% time-lapse effect using the data-difference strategy.
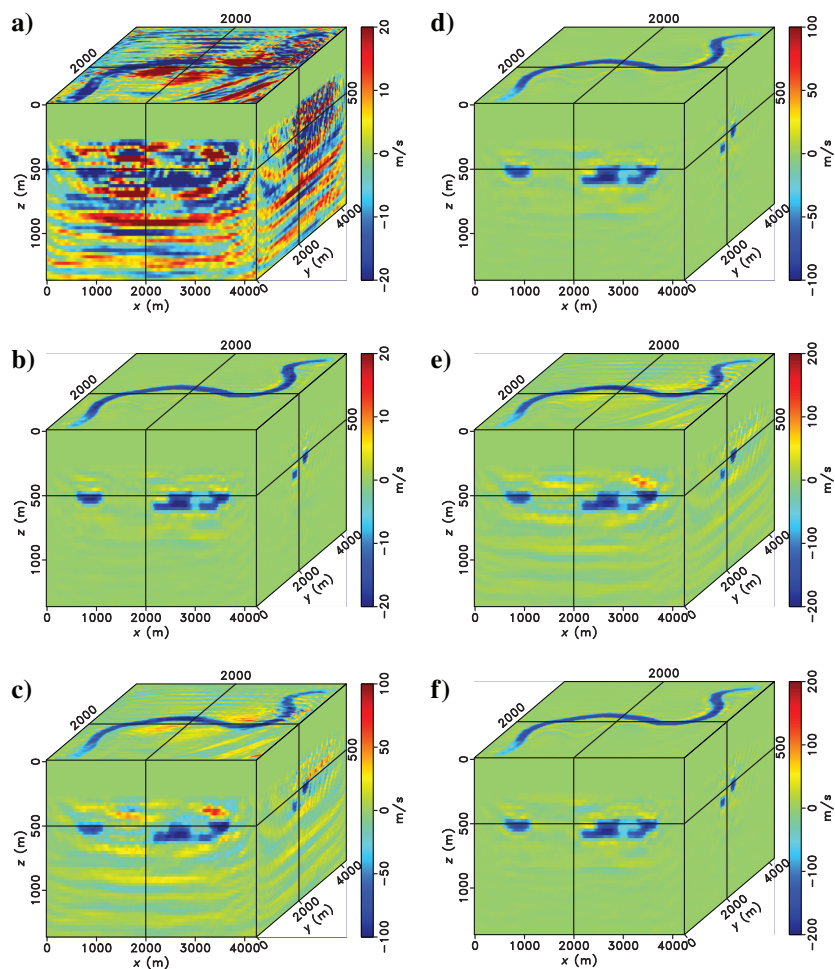
Figure 6. Inverted time-lapse images for V_S using data without noise. (a) 1% time-lapse effect using the sequential strategy, (b) 1% time-lapse effect using the data-difference strategy, (c) 5% time-lapse effect using the sequential strategy, (d) 5% time-lapse effect using the data-difference strategy, (e) 10% time-lapse effect using the sequential strategy, (f) 10% time-lapse effect using the data-difference strategy.
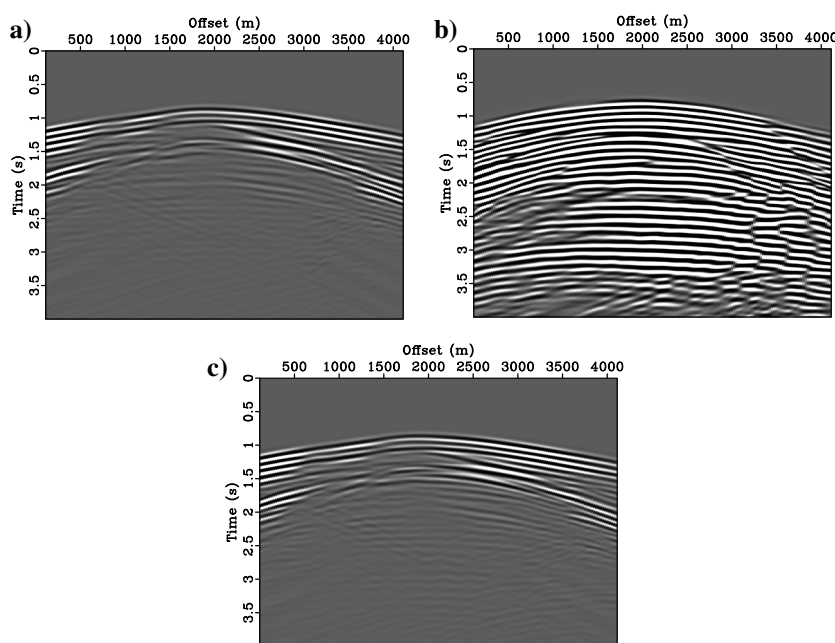
Figure 7. The time-lapse data difference for a pressure recording from a single cable for the 1% time-lapse change in V_P and V_S: (a) true, (b) sequential strategy, and (c) data-difference strategy. The same clip value is used in all plots.

and a total memory of 32 GB. With optimal spreading of the work to the nodes, such that each core computed one modeling and one gradient, one iteration took approximately 1.5 h. The size of the temporary data generated during one iteration was approximately 1500 GB. In total, 99% of the computing time was spent on the modeling and gradient calculations, and the remaining time was used on stacking the local gradients, calculating the objective functional, and updating the model.

DISCUSSION

Conventional time-lapse methodology requires careful matching of baseline and monitor data sets before and throughout the entire processing sequence (Kristiansen et al., 2000), including poststack matching. The resulting 4D data volumes are usually interpreted qualitatively in terms of amplitude information (Landrø et al., 1999). Measurements of time delays are to a certain degree used to obtain quantitative velocity information (Landrø, 2002). The TLFWI offers a completely new way of time-lapse processing by aiming at computing differences in important parameters such
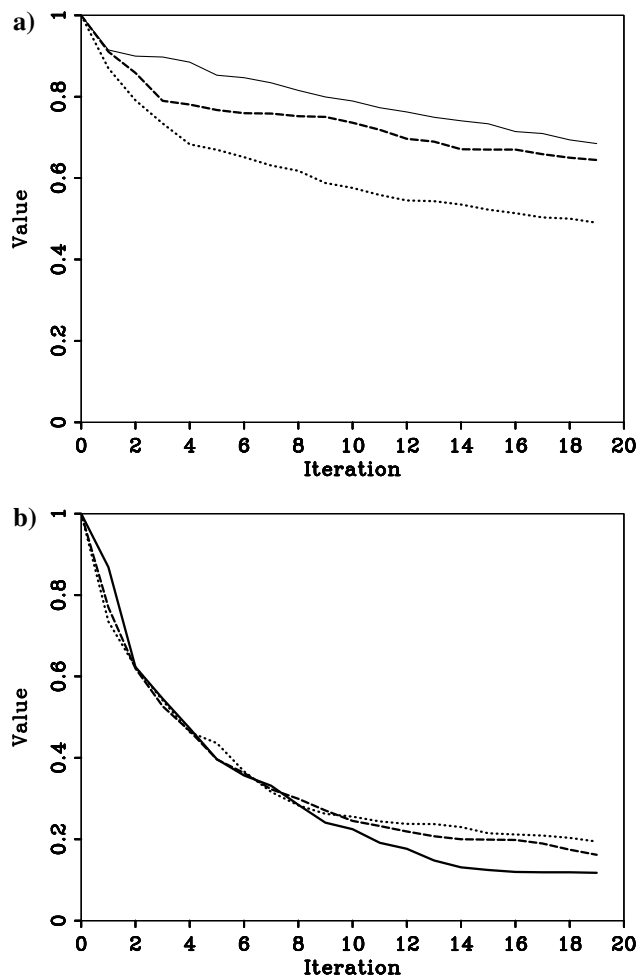

Figure 8. Normalized objective functionals for the monitor inversions without noise in the data. (a) Sequential strategy and (b) data-difference strategy (solid line: 1% change, dashed line: 5% change, dotted line: 10% change).

as wave velocity throughout the entire data volume as a function of depth.

The noise-free examples given above show that, in principle, the FWI approach can resolve velocity changes of the order of 1%. Even though it is possible, in theory, to resolve even smaller changes, there are fundamental limitations to the resolving power even for perfect acquisition repeatability. The spatial resolution of velocity changes can in principle be calculated using the Hessian matrix (Fichtner and Trampert, 2011), which very roughly depends on the frequency content of the source signature, the focusing power of the acquisition geometry, and the overburden velocity model. Due to the large computational cost of FWI, we are (at the moment) forced to use the low-frequency end of the seismic spectrum, thus limiting the accuracy. Conventional time-lapse processing can use a much larger frequency spectrum and should in principle achieve better resolution, but will also be limited by the acquisition geometry in the same way as FWI. However, both methodologies depend on the accuracy of the overburden velocity model that to a very large extent determines the resolving power. There is now considerable evidence that FWI has the ability to estimate accurate velocity models (Sirgue et al., 2009), thus counteracting the low-frequency problem.

Conventional processing aims at estimating reflection coefficients and smooth velocity models, whereas FWI sets a more ambitious goal of trying to add detailed velocity information to the smooth models obtained from tomography. Parametrization of the forward problem is important because sensitivity and coupling between different types of parameters have a large effect on the final solution. Multicomponent data using V_P and V_S seem to be a natural choice because the pressure and the vertical components of the particle velocity contain velocity information through P-wave traveltime and the horizontal component of the particle velocity contains S-wave information through converted (PS) waves. Estimating V_P and V_S simultaneously ensures that the final velocity models are consistent, which is usually a problem in conventional processing of OBC data (Szydlik et al., 2007; Mathewson et al., 2013). The disadvantage is that density is difficult to estimate due to the strong coupling to V_P and V_S at small and intermediate offsets.

From a computational point of view, it is favorable to invert simultaneously for V_P and V_S because the number of required iterations is kept at a minimum. However, because the model space is increased in a simultaneous inversion compared to, for instance, a sequential inversion, it is more likely that the inversion gets stuck in local minima, particularly with real data and more complex media. In our preliminary test runs, we find that the weighting matrix (equation 6) has a major impact on the inversion results. This can be explained by the fact that there is a different sensitivity of V_P and V_S with respect to the data. Another subject of discussion is the parametrization of the inversion which also has impact on the results (Mora, 1987; Prieux et al., 2013a). Thus, multiparameter inversion is a challenging task and is still a topic of research (Operto et al., 2013; Prieux et al., 2013b; Raknes and Arntsen, 2014b).

The quality of FWI results depend on the initial model, but conventional processing suffers from the same problem, and this cannot be viewed as a major drawback. The examples shown in the preceding section show that by using initial velocity models accurate enough to avoid cycle skipping, the final solution is reasonably accurate. An important observation is the presence of strong surface waves in the data sets due to the high contrast between the water

layer and the sea bottom. The surface waves, particularly, the low-frequency Scholte waves, were complicated to estimate using FWI. As a result, the initial model had to be close to the true model in the parts close to the sea floor, such that the surface waves were approximated correctly. Wrong estimates of the shallow parts lead to convergence into a local minimum far from the true solution.

Industrial time-lapse data are not noise free and contain (often serious) errors due to the differences in acquisition geometries and source characteristics between baseline and monitor surveys. Much of the effort in conventional time-lapse processing is spent trying to compensate and rectify these errors. When it comes to us-

ing the FWI method, little experience has been gained, and the knowledge of the sensitivity of the method to repeatability problems is not widespread. However, as with conventional processing, acquisition differences can be attacked prior to inversion by data interpolation. Extensive synthetic testing can to a certain degree give some insight, but would, due to the computational effort involved, be prohibitively expensive. As an inexpensive alternative, we have shown results when colored noise is added to the data, and it is clear that time-lapse effects can be estimated even at fairly low S/Ns. It is worth mentioning that the Gaussian noise used here is not a realistic type of noise. A more realistic noise type could have been achieved
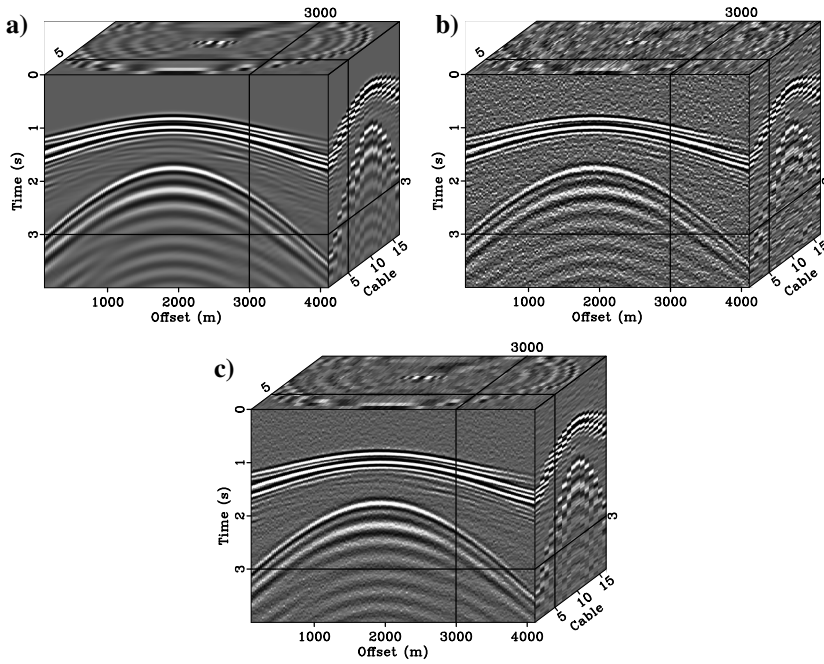
Figure 9. (a) Data without noise, (b) data with S/N = 2.0, and (c) data with S/N = 5.0. The same clip value is used in all plots.

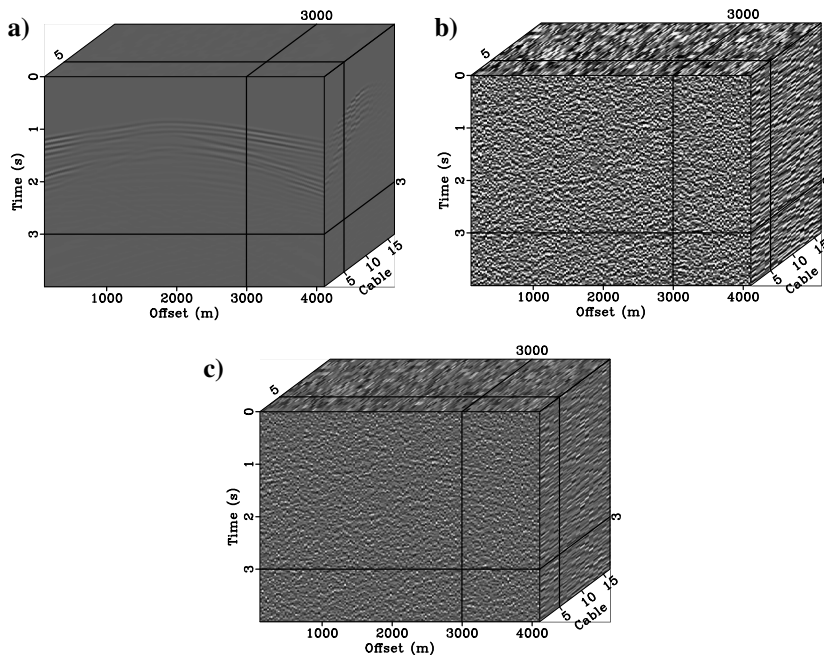Figure 10. (a) True time-lapse data difference without noise, (b) time-lapse data difference with S/N = 2.0, and (c) time-lapse data difference with S/N = 5.0. The same clip values are used in all plots.
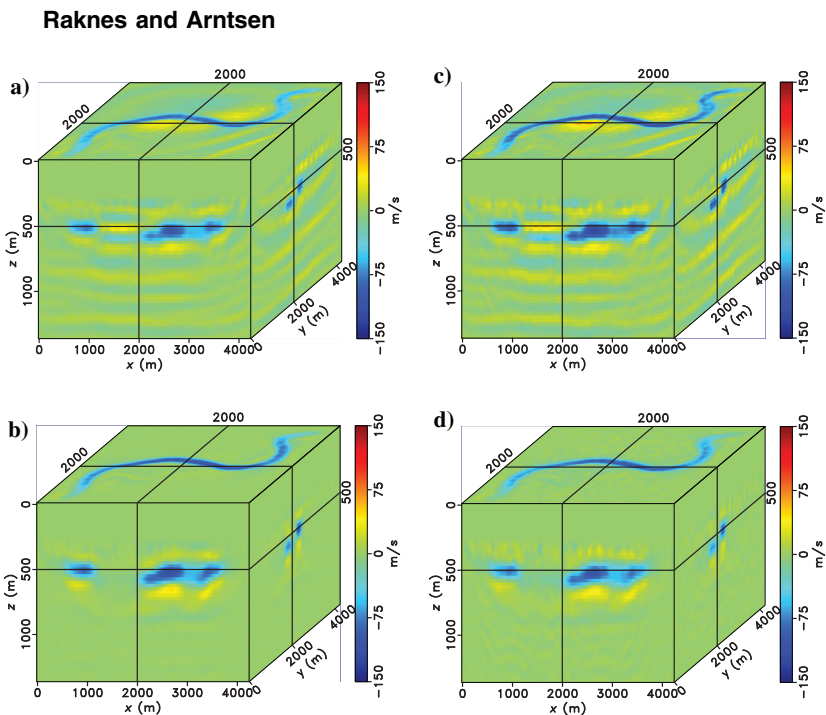
Raknes and Arntsen

Figure 11. Inverted time-lapse images for V_P for data with noise (top row: sequential strategy, bottom row: data-difference strategy): (a) S/N = 5.0, (b) S/N = 5.0, (c) S/N = 2.0, and (d) S/N = 2.0.

Figure 12. Inverted time-lapse images for V_S for data with noise (top row: sequential strategy, bottom row: data-difference strategy): (a) S/N = 5.0, (b) S/N = 5.0, (c) S/N = 2.0, and (d) S/N = 2.0.
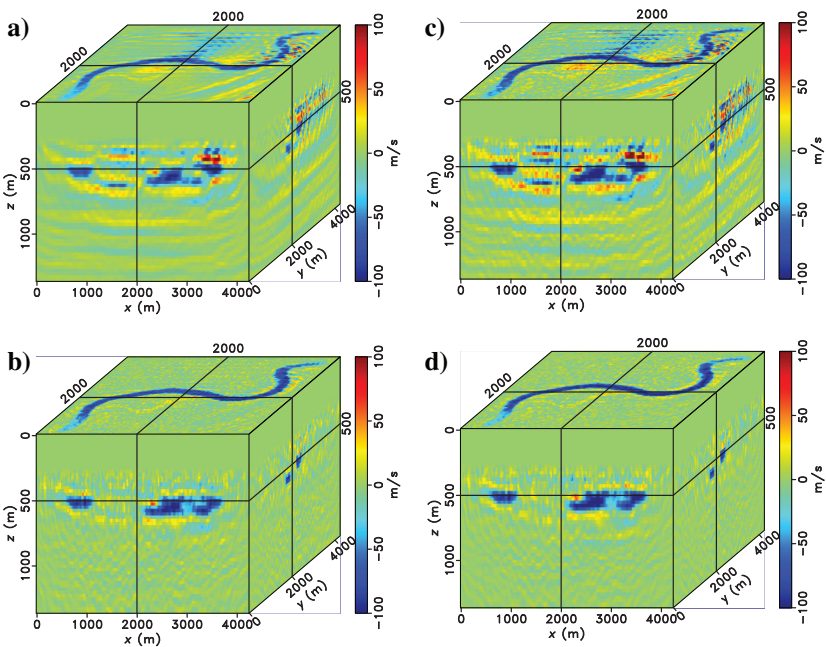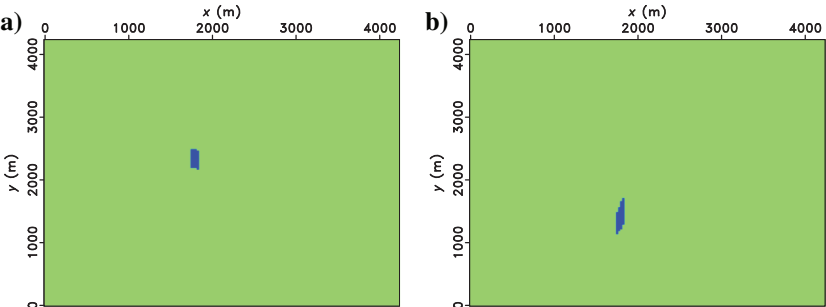
Figure 13. Horizontal slice through the small-size time-lapse model at (a) $z = 500$ m and (b) $z = 575$ m. The magnitude for the V_P anomaly is $0 - 152$ and $0 - 96$ m/s for the V_S anomaly.

by introducing effects related to the coupling of the receivers to the seafloor and differences in the rotation of the particle velocity sensors (geophones). In general, it is difficult to simulate real-world noise because it is caused by a large variety error sources.

In principle, TLFWI should be able to compensate for some of the acquisition differences usually found between baseline and monitor surveys. If a full Gauss-Newton approach is used for the inversion, information contained in the Hessian matrix is able to compensate to a certain degree for the acquisition imprint (Virieux and Operto, 2009), which is supported with experience from least-squares migration (Nemeth et al., 1999). In our examples, we use the gradient-based L-BFGS method, which only partially exploits the Hessian matrix. If the repeatability in the baseline and monitor data sets is poor, then the data-difference strategy will in-
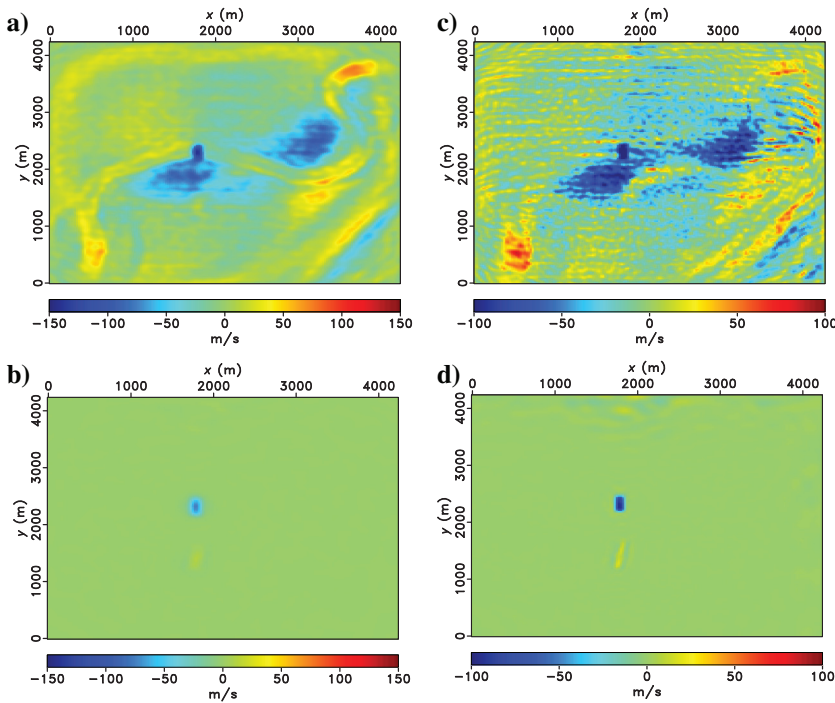

Figure 14. Horizontal time-lapse slices at $z = 500$ m for the small size time-lapse model: (a) V_P using the sequential strategy, (b) V_P using the data-difference strategy, (c) V_S using the sequential strategy, and (d) V_S using the data-difference strategy.
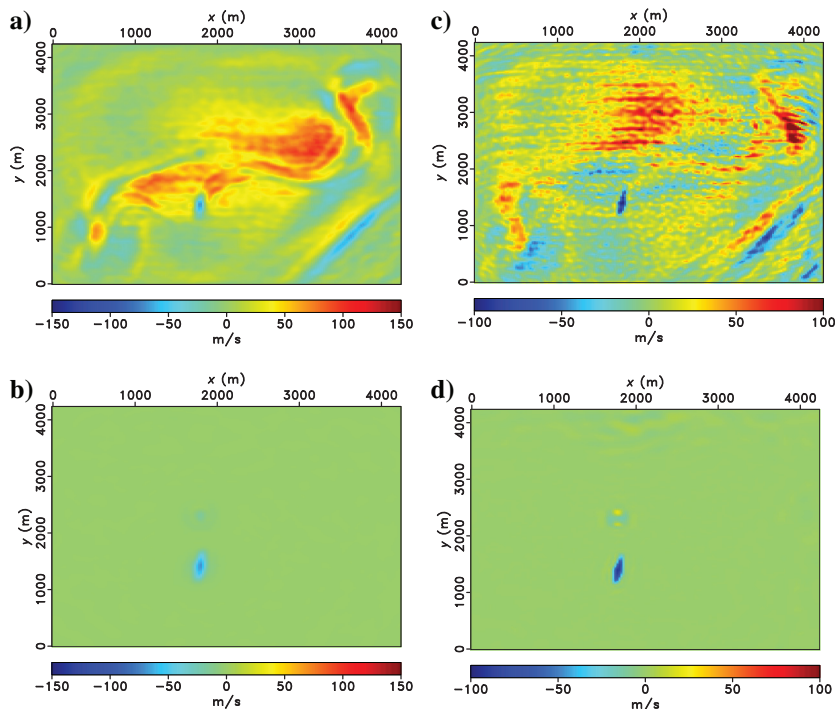

Figure 15. Horizontal time-lapse slices at $z = 575$ m for the small size time-lapse model: (a) V_P using the sequential strategy, (b) V_P using the data-difference strategy, (c) V_S using the sequential strategy, and (d) V_S using the data-difference strategy.

troduce an error in the modified data set that the Hessian matrix is not able to compensate for because this error is not due to simplifications in the method itself.

In standard monitoring setups, the location of the reservoir is often well known, and it may be used to constrain the inversion. This prior information may be incorporated either by introducing a model term in the objective functional (Asnaashari et al., 2013) or by constraining the gradient itself (Zhang and Huang, 2013). Another option is to use the local migration regularization (Raknes and Arntsen, 2014a), which migrates the data differences to constrain the model update. Using the above-mentioned methods, it is possible to reduce the time-lapse artifacts in the images.

The goal with this study is to compare the sequential and data-difference TLFWI strategies in 3D, and thus give insights into finding the best strategy for performing TLFWI. From our examples, it is clear that the data-difference strategy is favorable, particularly for cases in which the amplitudes of the time-lapse changes sought for are small. The reason for the success is that the monitor inversion is, as long as the convergence of the baseline inversion is sufficient close to the true solution, not trying to fit the nonperfect baseline model and the time-lapse anomalies at the same time. From this argument, it follows that having baseline and time-lapse variations as simultaneous unknowns in the sequential strategy yields a more difficult inverse problem to solve than inverting for the differences in the data-difference strategy.

The disadvantage with the data-difference strategy is that it has strong requirements on the repeatability of the data sets because the receivers must be positioned on identical points in space in both surveys to consider the data differences. In a permanent OBC setup, this is easy, whereas using conventional streamer geometries it is more difficult. In addition, it is difficult in practice to put the source on identical positions in a full-scale survey. Furthermore, factors related to the acquisition (i.e., different equipment, time of the year of the acquisitions, weather conditions) of the data sets may cause differences between the baseline and monitor data sets that may be interpreted as real time-lapse signals because $\Delta\varepsilon_r$ in equation 9 may become larger in magnitude than the time-lapse signals. Moreover, in some settings, $\Delta\varepsilon_n$ in equation 9 may become as large or even larger in magnitude than the time-lapse signal yielding false time-lapse changes in the elastic parameters. Therefore, the data difference based inversion strategy may not be used in cases in which the time-lapse data sets are not acquired using the same source-receiver geometries, or in which the repeatability in the data sets is not sufficiently good. In these cases, the sequential strategy is the best approach. It is also worth mentioning that once the data difference is computed it cannot be undone and all errors are locked in the resulting data set.

It remains a task for the future to determine the full characteristics of FWI when it comes to acquisition errors. It is clear from the discussion above that to release the full potential of TLFWI, data with a broad frequency range have to be used. Unfortunately, the amount of computer power needed scales as the frequency to the fifth power, implying that presently this cannot be done efficiently.

CONCLUSION

We have investigated two strategies for performing time-lapse FWI using synthetic ocean-bottom multicomponent seismic time-lapse data. A 3D isotropic elastic FWI implementation was used to investigate the sensitivity with respect to the magnitude and the physical size of the time-lapse anomalies in the P- and S-wave velocity models, and the noise level in the data. Both strategies were able to detect the time-lapse anomalies, but the strategy that focuses on the differences in the data is preferable due to fewer artifacts in the final time-lapse images. This strategy depends on compatible source and receiver geometries in the data sets, and sufficient convergence of the inversion of the baseline data set.

ACKNOWLEDGMENTS

REFERENCES

Abubakar, A., G. Pan, M. Li, L. Zhang, T. Habashy, and P. van den Berg, 2011, Three-dimensional seismic full-waveform inversion using the finite-difference contrast source inversion method: Geophysical Prospecting, **59**, 874–888, doi: 10.1111/j.1365-2478.2011.00953.x.

Aki, K., and P. G. Richards, 2002, Quantitative seismology 2nd ed.: University Science Books.

Asnaashari, A., R. Brossier, S. Garambois, F. Audebert, P. Thore, and J. Virieux, 2013, Regularized seismic full waveform inversion with prior model information: Geophysics, **78**, no. 2, R25–R36, doi: 10.1190/geo2012-0104.1.

Auer, L., A. Nuber, S. Greenhalgh, H. Maurer, and S. Marelli, 2013, A critical appraisal of asymptotic 3D-to-2D data transformation in full-waveform seismic crosshole tomography: Geophysics, **78**, no. 6, R235–R247, doi: 10.1190/geo2012-0382.1.

Barnes, C., and M. Charara, 2009, The domain of applicability of acoustic full-waveform inversion for marine seismic data: Geophysics, **74**, no. 6, WCC91–WCC103, doi: 10.1190/1.3250269.

Berenger, J.-P., 1994, A perfectly matched layer for the absorption of electromagnetic waves: Journal of Computational Physics, **114**, 185–200, doi: 10.1006/jcph.1994.1159.

Biondi, B., C. Deutsch, R. Gundesø, D. Lumley, G. Mavko, T. Mukerji, J. Rickett, and M. Thiele, 1996, Reservoir monitoring: A multi-disciplinary feasibility study: 66th Annual International Meeting, SEG, Expanded Abstracts, 1775–1778.

Butzer, S., A. Kurzmann, and T. Bohlen, 2013, 3D elastic full-waveform inversion of small-scale heterogeneities in transmission geometry: Geophysical Prospecting, **61**, 1238–1251, doi: 10.1111/j.1365-2478.12065.

Fichtner, A., H.-P. Bunge, and H. Igel, 2006, The adjoint method in seismology: I — Theory: Physics of the Earth and Planetary Interiors, **157**, 86–104, doi: 10.1016/j.pepi.2006.03.016.

Fichtner, A., and J. Trampert, 2011, Hessian kernels of seismic data functionals based upon adjoint techniques: Geophysical Journal International, **185**, 775–798, doi: 10.1111/j.1365-246X.2011.04966.x.

Greaves, R., and T. Fulp, 1987, Three-dimensional seismic monitoring of an enhanced oil recovery process: Geophysics, **52**, 1175–1187, doi: 10.1190/1.1442381.

Holberg, O., 1987, Computational aspects of the choice of operator and sampling interval for numerical differentiation in large-scale simulation of wave phenomena: Geophysical Prospecting, **35**, 629–655, doi: 10.1111/j.1365-2478.1987.tb00841.x.

Kristiansen, P., P. Christie, J. Bouska, A. O'Donovan, P. Westwater, and E. Thorogood, 2000, Foinaven 4D: Processing and analysis of two designer 4Ds: 70th Annual International Meeting, SEG, Expanded Abstracts, 1456–1459.

Landrø, M., 2002, Uncertainties in quantitative time-lapse seismic analysis: Geophysical Prospecting, **50**, 527–538, doi: 10.1046/j.1365-2478.2002.00330.x.

Landrø, M., O. A. Solheim, E. Hilde, B. O. Ekren, and L. K. Strønen, 1999, The Gullfaks 4D seismic study: Petroleum Geoscience, **5**, 213–226, doi: 10.1144/petgeo.5.3.213.

Lumley, D., D. C. Adams, M. Meadows, S. Cole, and R. Wright, 2003, 4D seismic data processing issues and examples: 73rd Annual International Meeting, SEG, Expanded Abstracts, 1394–1397.

Mathewson, J., M. J. Woodward, D. Nichols, C. Xu, and C. Leone, 2013, Joint PP/PS tomography with floating event constraints: 83rd Annual International Meeting, SEG, Expanded Abstracts, 1649–1653.

McMechan, G., 1983, Migration by extrapolation of time-dependent boundary values: Geophysical Prospecting, **31**, 413–420, doi: 10.1111/j.1365-2478.1983.tb01060.x.

Mittet, R., 2002, Free-surface boundary conditions for elastic staggered-grid modeling schemes: Geophysics, **67**, 1616–1623, doi: 10.1190/1.1512752.

Mora, P., 1987, Nonlinear two-dimensional elastic inversion of multioffset seismic data: Geophysics, **52**, 1211–1228, doi: 10.1190/1.1442384.

Nemeth, T., C. Wu, and G. Schuster, 1999, Least-squares migration of incomplete reflection data: Geophysics, **64**, 208–221, doi: 10.1190/1.1444517.

Nocedal, J., and S. J. Wright, 2006, Numerical optimization 2nd ed.: Springer Science + Business Media, LLC.

Operto, S., Y. Gholami, V. Prieux, A. Ribodetti, R. Brossier, L. Metivier, and J. Virieux, 2013, A guided tour of multiparameter full-waveform inversion with multicomponent data: From theory to practice: The Leading Edge, **32**, 1040–1054, doi: 10.1190/tle32091040.1.

Plessix, R.-E., 2009, Three-dimensional frequency-domain full-waveform inversion with an iterative solver: Geophysics, **74**, no. 6, WCC149–WCC157, doi: 10.1190/1.3211198.

Pratt, R. G., 1999, Seismic waveform inversion in the frequency domain, part 1: Theory and verification in a physical scale model: Geophysics, **64**, 888–901, doi: 10.1190/1.1444597.

Prieux, V., R. Brossier, S. Operto, and J. Virieux, 2013a, Multiparameter full waveform inversion of multicomponent ocean-bottom-cable data from the Valhall field. Part 1: Imaging compressional wave speed, density and attenuation: Geophysical Journal International, **194**, 1640–1664, doi: 10.1093/gji/ggt177.

Prieux, V., R. Brossier, S. Operto, and J. Virieux, 2013b, Multiparameter full waveform inversion of multicomponent ocean-bottom-cable data from the Valhall field. Part 2: Imaging compressive-wave and shear-wave velocities: Geophysical Journal International, **194**, 1665–1681, doi: 10.1093/gji/ggt178.

Queißer, M., and S. Singh, 2013, Localizing CO_2 at Sleipner a — Seismic images versus P-wave velocities from waveform inversion: Geophysics, **78**, no. 3, B131–B146, doi: 10.1190/geo2012-0216.1.

Raknes, E., and B. Arntsen, 2014a, Time-lapse full-waveform inversion of limited-offset seismic data using a local migration regularization: Geophysics, **79**, no. 3, WA117–WA128, doi: 10.1190/geo2013-0369.1.

Raknes, E. B., and B. Arntsen, 2014b, Strategies for elastic full waveform inversion: 84th Annual International Meeting, SEG, Expanded Abstracts, 1222–1226.

Routh, P., G. Palacharla, I. Chikichev, and S. Lazaratos, 2012, Full wavefield inversion of time-lapse data for improved imaging and reservoir charac-terization: 82nd Annual International Meeting, SEG, Expanded Abstracts, doi: 10.1190/segam2012-1043.1.

Sirgue, L., O. Barkved, J. V. Gestel, O. Askim, and J. Kommedal, 2009, 3D waveform inversion on Valhall wide-azimuth OBC: 71st Annual International Conference and Exhibition, EAGE, Extended Abstracts, U038.

Sirgue, L., J. Etgen, and U. Albertin, 2008, 3D frequency domain waveform inversion using time domain finite difference methods: 70th Annual International Conference and Exhibition, EAGE, Extended Abstracts, F022.

Szydlik, T., P. Smith, S. Way, L. Aamodt, and C. Friedrich, 2007, 3D PP/PS prestack depth migration on the volve field: First Break, **25**, 43–47.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266, doi: 10.1190/1.1441754.

Vigh, D., K. Jiao, D. Watts, and D. Sun, 2014, Elastic full-waveform inversion application using multicomponent measurements of seismic data collection: Geophysics, **79**, no. 2, R63–R77, doi: 10.1190/geo2013-0055.1.

Vigh, D., and E. Starr, 2008, 3D prestack plane-wave, full-waveform inversion: Geophysics, **73**, no. 5, VE135–VE144, doi: 10.1190/1.2952623.

Virieux, J., 1986, P-SV wave propagation in heterogeneous media: Velocity-stress finite-difference method: Geophysics, **51**, 889–901, doi: 10.1190/1.1442147.

Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, no. 6, WCC1–WCC26, doi: 10.1190/1.3238367.

Warner, M., A. Ratcliffe, T. Nangoo, J. Morgan, A. Umpleby, N. Shah, V. Vinje, I. Štekl, L. Guasch, C. Win, G. Conroy, and A. Bertrand, 2013, Anisotropic 3D full-waveform inversion: Geophysics, **78**, no. 2, R59–R80, doi: 10.1190/geo2012-0338.1.

Watanabe, T., S. Shimizu, E. Asakawa, and T. Matsuoka, 2004, Differential waveform tomography for time lapse crosswell seismic data with application to gas hydrate production monitoring: 74th Annual International Meeting, SEG, Expanded Abstracts, 2323–2326.

Zhang, F., C. Juhlin, M. Ivandic, and S. Lüth, 2013, Application of seismic full waveform inversion to monitor CO_2 injection: Modelling and a real data example from the Ketzin site, Germany: Geophysical Prospecting, **61**, 284–299, doi: 10.1111/1365-2478.12021.

Zhang, Z., and L. Huang, 2013, Double-difference elastic-waveform inversion with prior information for time-lapse monitoring: Geophysics, **78**, no. 6, R259–R273, doi: 10.1190/geo2012-0527.1.

Zhen, Q., L. Minghui, Z. Xiaodong, Y. Yao, Z. Cai, and S. Jianyong, 2009, The implementation of an improved NPML absorbing boundary condition in elastic wave modeling: Applied Geophysics, **6**, 113–121, doi: 10.1007/s11770-009-0012-3.

Zheng, Y., P. Barton, and S. Singh, 2011, Strategies for elastic full waveform inversion of time-lapse ocean bottom cable (OBC) seismic data: 81st Annual International Meeting, SEG, Expanded Abstracts, 4195–4200.