Challenges and solutions for performing 3D time-domain elastic full-waveform inversion

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

Abstract

The full-waveform inversion (FWI) method relies on an effective numerical solution of the wave equation. The wave equation must be solved numerous times during an inversion run. In the past, to be able to use FWI in practice, it was necessary to assume that the earth's subsurface was a 2D acoustic medium. Recent increases in computational power have made it possible to include more real-world physics in the FWI method, such that the computational subsurface can mimic the real-world subsurface as closely as possible. Going from 2D to 3D is challenging, primarily due to the numerical methods involved in solving the wave equation. Including elastic effects is not straightforward due to the increase in possible models that can explain the data, more complicated wave phenomena involved in the wave propagation, as well as trade-off between the subsurface elastic parameters during the inversion. We discuss some of the challenges and solution strategies for using the FWI method in the time domain using a 3D elastic computational domain.

Introduction

Seismic waves that have propagated through the earth and are recorded at the earth's surface contain information about the physical properties of the subsurface. The physical properties of the subsurface affect the traveltime as well as the amplitude of the seismic signal and are of great interest for a proper characterization of the subsurface. One of the goals of exploration geophysics is to use the seismic signal to obtain a better understanding of the subsurface that can improve the success rate in the search for hydrocarbons. Among other seismic imaging methods, the full-waveform inversion (FWI) method has proved in recent decades to be an effective tool for finding high-resolution models of the physical properties of the subsurface.

In the first investigations of the earth's interior, a few seismograms from different earthquakes around the globe were used (Oldham, 1906). Since then, much effort has been put into developing systems for acquiring better seismological data. Nowadays, accurate sensors such as hydrophones and geophones are placed around the globe to continuously monitor waves related to earthquakes and other natural disasters. In exploration geophysics, on the other hand, the development has been driven toward acquisition systems that can collect seismic data using active sources in the water layer, at the sea bottom, on land, and inside well bores. Today, it is possible to acquire dense, multicomponent seismic data with better quality than ever before. To account for the full potential of high-quality seismic data, it is important to use advanced methods like FWI that incorporate as much real-world physics as possible.

The first methods for estimating subsurface parameters were based on using traveltime information of the waves recorded at the earth's surface (Oldham, 1906). Claerbout (1971) introduced migration of seismic data based on the wave equation, which subsequently has developed into present-day advanced imaging methods such as reverse time migration (RTM). Common for modern methods is the requirement for a kinematically correct velocity model of the subsurface that is used to propagate the seismic signal recorded at the surface back to the origin of the event (typically an interface) in the subsurface.

Tarantola (1984) and Mora (1987) introduced the concept of FWI. The FWI method gives estimates of the physical properties for the medium the waves have traveled through and is formulated as an inverse problem in which one tries to estimate the subsurface parameters directly by minimizing the difference between recorded and synthetic seismic data. This is done in a nonlinear optimization procedure. The key requirement for the method is the numerical solution of the wave equation, which is used both in generating the synthetic data and in forming the model gradient used to update the physical property model in the optimization method.

Imaging methods are commonly divided into two main categories. The first category consists of migration methods like RTM that utilize the high-frequency features of the subsurface, and thus a detailed image of the subsurface is the output from these methods. The second category, normally called seismic inversion, consists of methods that estimate the physical properties of the subsurface through an inversion process. FWI belongs to the second category. As the migration methods require a velocity model, common practice is to use a two-step approach in a conventional imaging workflow: A velocity model is created using a combination of seismic tomography and FWI. This inverted velocity model is then used as input in the migration method. By using this approach, the advantages from both categories are combined to improve understanding of the subsurface.

Because FWI requires numerical solutions of the wave equation, it is computer intensive. As a consequence, the development and applications of the method have somewhat followed the development of computer technology. The first applications of FWI, therefore, were restricted to the assumption that the subsurface is a 2D acoustic medium (Pratt, 1999). This assumption reduces the computational cost dramatically, as the 2D acoustic wave equation requires at least an order of magnitude less computer resources to solve than the 3D elastic counterpart. The restriction to 2D acoustic wave propagation is an oversimplification, and hence a large effort has been invested to extend FWI to a real 3D world.

Full-waveform inversion can be performed in the time domain or in the frequency domain. In 3D, the most common implementation for FWI is the time domain due to the scaling of the computational burden for the numerical methods. The time-domain implementation has proven worthy of tackling a large range of applications in both controlled-source exploration geophysics and

---

Special Section: Full-waveform inversion Part II

earthquake seismology. In the frequency domain, the wave equation is reduced to a boundary value problem that is formulated as a system of linear equations. This system is solved using iterative methods or direct Gaussian elimination techniques. For the 3D problem, the linear algebra techniques have been considered intractable due to the memory demands in solving the linear problem (Operto et al., 2015). Because of this fact, we have implemented our elastic FWI method in the time domain.

In this paper, we discuss the importance of fully accounting for real-world physics in FWI. We mention some of the challenges and solution strategies in performing 3D elastic FWI in the time domain. In the end, we show examples of applications using 3D elastic time-domain FWI on synthetic and real data.

Why 3D elastic FWI?

The real world is 3D in its property variations, and the recorded wavefield is a result of 3D wave propagation. One important difference between 2D and 3D wave propagation is the geometrical spreading factor that controls the amplitude of the signal. In a 2D setting, the amplitude of the wavefield (from a point source in a homogeneous medium) decay is $1/\sqrt{r}$ where r is the distance the wave has traveled. In 3D, on the other hand, the decay is $1/r$. Moreover, there is a phase difference between 2D and 3D wave propagation. In some cases, the phase differences may be large. Thus, when the 2D wave equation is used in FWI, the amplitudes of the synthetic data do not match those of the real data, and there is a challenge in getting the phase information correct. The consequence is that real data must be modified prior to the inversion. This modification has proved to be problematic and may give artifacts in the inverted subsurface models (Auer et al., 2013).

The acoustic assumption is problematic, in general, simply because the real world is not an acoustic medium. A better approximation is to assume that the subsurface is an elastic medium. Wave phenomena like shear waves and converted waves are not included in the acoustic assumption as the medium is assumed to be a fluid (i.e., no shear forces are allowed in the medium). However, the acoustic wave equation models the kinematics of the pressure waves in an approximately correct fashion. This fact is taken to be one of the main reasons FWI has been applied with success using the acoustic assumption. On the other hand, the wavefield dynamics are not modeled correctly such that the reflection and transmission coefficients are incorrectly estimated. For the synthetic data, this may have a large impact on the amplitudes, with the result being that the synthetic-data amplitudes are not correctly estimated compared to the real-data amplitudes.

To illustrate the difference between acoustic and elastic FWI, we have constructed a synthetic elastic 3D model consisting of several layers in depth, each with different physical properties. We simulate a streamer survey with a cable length of 1700 m and use the elastic wave equation to create the synthetic data. Since this is short-offset data, we use FWI to invert for the P-wave velocity (V_P) and link the other parameters using simple empirical relationships. In Figure 1, we show vertical profiles of the density (ρ) and V_P for acoustic and elastic FWI. We observe that down to approximately 400 m depth the acoustic FWI fails completely in approximating both the density and V_P model. This is a direct consequence of the incorrect reflection and transmission coefficients in the acoustic assumption. The elastic FWI, on the other hand, does not have the same challenges in approximating the

parameters in the same interval. At interfaces, there are oscillations in both the acoustic and elastic FWI results. This is a consequence of the low frequencies used in the inversion. In addition, the layer interfaces are slightly shifted downward for the acoustic FWI compared to the elastic FWI. This example shows that it is important to include elastic effects in FWI, particularly when we use short-offset data.

The benefit of performing 3D elastic FWI is that the synthetic data is approximated more correctly since the decay of the amplitudes and the phase are correct. Thus, no data modification is needed prior to inversion. Moreover, out-of-plane events become part of the synthetic data, which may improve the imaging results. In the end, as more real-world physics is included in the method, accurately estimated subsurface elastic models may yield better images and thus increase understanding of the subsurface. The understanding also may be improved simply because the physical models and thus the seismic images are 3D.

Challenges in performing 3D elastic FWI

The challenges in going from the 2D acoustic to the 3D elastic computational setting can be divided, in general terms, into two categories: (1) computational challenges and (2) theoretical challenges.

In the center of every FWI algorithm lies the numerical solution of the wave equation. The numerical solution is used both to create the synthetic data and in the computation of the model gradient. A seismic wave propagating in an elastic medium is governed by the following partial differential equation (Aki and Richards, 2002)

$$\rho(\mathbf{x})\partial_t^2 u_i(\mathbf{x},t) = \partial_j \sigma_{ij}(\mathbf{x},t) + f_i(\mathbf{x},t), \qquad (1)$$

and

$$\sigma_{ij}(\mathbf{x},t) = c_{ijkl}(\mathbf{x})\varepsilon_{kl}(\mathbf{x},t), \qquad (2)$$

where $\rho(\mathbf{x})$ is the density, $u_i(\mathbf{x},t)$ is the particle displacement in direction i, $\sigma_{ij}(\mathbf{x},t)$ is the stress tensor, $f_i(\mathbf{x},t)$ is a body force, $c_{ijkl}(\mathbf{x},t)$ is the elasticity tensor, and $\varepsilon_{kl}(\mathbf{x},t)$ is the strain tensor given as

$$\varepsilon_{kl}(\mathbf{x},t) = \frac{1}{2}\left(\frac{\partial u_k(\mathbf{x},t)}{\partial x_l} + \frac{\partial u_l(\mathbf{x},t)}{\partial x_k}\right). \qquad (3)$$
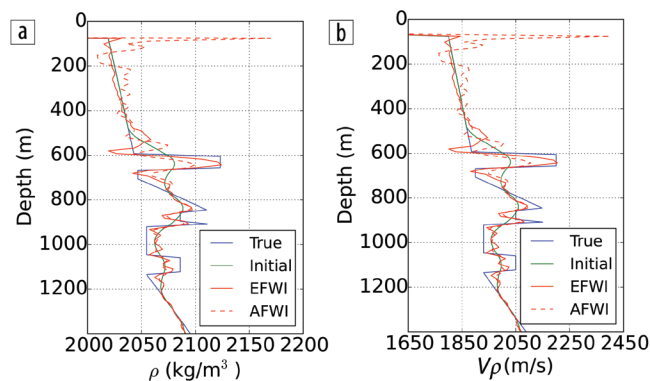

Figure 1. Vertical profiles of the models in the synthetic test for the difference between acoustic FWI (AFWI) and elastic FWI (EFWI). (a) The density (ρ) and (b) the P-wave velocity (V_P) (modified from Raknes et al., 2015).

In the above equations, the Einstein's summation convention is used and $i, j, k, l \in \{x, y, z\}$.

The above equation can be solved by a wide range of different numerical methods (Igel, 2016). In the following we use the finite-difference approach. From a computational point of view, the most expensive part of solving the wave equation is the computation of the derivatives in both space and time. This is done by discretizing the equation and computing the derivatives using a finite number (typically below 20) of local points. The wave equation is hyperbolic by nature (Evans, 2010), and thus the wave equation is normally solved by an explicit method that requires small time steps to be numerically stable. If we let n be the number of points on each spatial axis and n_t be the number of time steps, the computational cost for solving the equation scales as $\mathcal{O}(n^3 \cdot n_t)$. In 3D, the discrete model often consists of several millions of points — a number which is large compared to a standard 2D discrete model. More-over, the number of derivatives to compute is larger in 3D than in 2D simply because we have one more dimension to consider. Hence, solving the 3D elastic wave equation for a realistic-sized problem is a computationally demanding procedure.

To solve equations 1 and 2 efficiently, it is important to keep all variables and physical parameters in memory on the computer as there is a large bottleneck for file input and output operations. In the past, there was a limitation on the size of the problems possible to solve due to limited memory. On today's modern computer systems, there are still memory limitations, but in practice they are nonexistent as the time spent to solve the wave equation is impractically large if the whole memory is used for one single simulation.

The amount of data in a conventional seismic survey is large — typically several terabytes. The number of shots is typically several hundred thousands. In FWI, a synthetic data set and a model gradient for each shot must be computed in every iteration of the optimization method. Hence, to be able to perform a successful inversion run, the method must be parallelized for use on large high-performance computing (HPC) facilities. Further-more, FWI requires a large amount of temporal data storage during the computations. Consequently, FWI puts large require-ments on the computational system to be used.

From a theoretical point of view, the inverse problem is ill-posed (in the Hadamard sense [Evans, 2010]) and nonlinear, demanding the use of complicated solution techniques to solve the problem. In addition, including elastic parameters in FWI is challenging because the solution space increases compared to a similar acoustic inverse problem, since more parameters are in-cluded in the model space. Hence, more parameter models can explain the data, which means the number of solutions to the inverse problem is larger in the elastic case.

A major challenge with including elastic parameters is that there is a coupling effect (so-called trade-off) between the elastic parameters affecting wave propagation and, thus, the seismic signal recorded at the receiver positions. Consequently, the inver-sion results may be degraded if these effects are not taken into account in the elastic FWI (Wang et al., 2016).

The standard solution method for the optimization problem is to use a local gradient-based optimization method. The major challenge with these methods is the convergence into so-called local minima that potentially can be far away from the global minimum sought for. To use such methods requires that the initial model is close to the solution.

Solution strategies

The increase in available computer power in recent years means we can solve bigger and more complex problems than ever before. However, the computational burden for performing 3D elastic FWI is so high that we still are not able to solve problems as large as we desire within a reasonable time. One way to circumvent the compute time is to use graphic cards to perform the computations (Venstad, 2016). With this strategy, it has been shown that the compute time can be reduced by a factor close to 30 (Weiss and Shragge, 2013).

Another way to efficiently solve the wave equation is to de-compose the computational domain into smaller parts and dis-tribute the different parts between each CPU on the HPC. Each CPU computes the wave equation in its respective part of the domain. This approach requires a large amount of communication between the compute nodes, as the wavefields at the boundaries of each part must be sent back and forth between the nodes.

One solution to reduce the storage needs necessary to perform the inversion is to use wavefield reconstruction methods (Nguyen and McMechan, 2015; Raknes and Weibull, 2016a; Yang et al., 2016). By using the Kirchhoff integral, the wavefields inside the computational domain can be reconstructed in the computation of the model gradient by using the wavefields recorded on the domain boundaries. For normal-sized seismic surveys, this will reduce the required storage needs from several terabytes down to tens of gigabytes. The extra cost to do this is the requirement of an extra numerical solution of the wave equation. In Figure 2, we show an example of a snapshot of the true and reconstructed wavefield using a synthetic anisotropic model. We observe that artifacts are part of the reconstructed wavefields. The amplitudes of the artifacts are smaller than the real events. For a single model-ing, the artifacts may be problematic, especially for long simulation times. When used in FWI, the artifacts are less problematic since the number of shots covering an area normally is large; thus the artifacts are reduced (stacked out) compared to real events in the computation of the updated FWI parameter model.

It is also important to reduce the number of shots used in the inversion to minimize the total computational time. The full data set therefore is decimated such that only a small portion of the shots in the area is used in FWI. The distance between each shot is dependent on the desired resolution and the frequencies used in the data (to avoid aliasing). To reduce the computational burden even more, the shots in the decimated data set are divided into different groups, which is a technique called shot subsampling. The shots from each group are used in different iterations of the optimization method.

There are two main strategies to avoid convergence into local minima. The first is to use optimization methods other than the local gradient-based methods. There is a large class of such methods ranging from methods that search the whole solution space (i.e., Monte Carlo methods) to pseudoglobal methods like the simulated annealing method. All these methods still are challenging to use as the number of simulations required is relatively large compared

to the local methods. Furthermore, constraints must be included in these methods to ensure that the solution found represents the physical properties of the underground correctly.

The second strategy is to improve the starting model using methods other than FWI. Examples of such methods are first-arrival traveltime tomography (Leung and Qian, 2006), reflection-based tomography (Prieux et al., 2013), or wave equation migration velocity analysis (Raknes and Weibull, 2016b). It is worth mentioning that this approach is not directly solving the challenge with convergence into a local minimum. It only makes a model that (hopefully) is closer to the solution sought for, and hence there are better changes for convergence towards to this solution.

The trade-offs between the different elastic parameters is dependent on the source and receiver geometry. One approach to reduce the trade-off is by data filtering, using different data components or applying different problem parametrizations (Köhn et al., 2012). Another approach is to use the Hessian matrix in the optimization method. The Hessian matrix holds information about the curvature of the misfit functional with respect to the model parameters. In addition, it is considered as an operator that accounts for the limited illumination at greater depth. Furthermore, the inverse Hessian is expected to correct the model gradients from the trade-offs between the elastic parameters (Wang et al., 2016). The major challenge in elastic FWI is that the Hessian and its inverse are more expensive to compute than the gradients, as second-order derivatives of the misfit functional with respect to the model parameters are needed. With the increase in computer power, we currently see an increasing interest in including the Hessian in the FWI method (Wang et al., 2016).

Figure 2. (a) The true wavefield and (b) the reconstructed wavefield using wavefield recordings back-propagated from the domain boundaries using the Kirchhoff integral. In this example, the wavefields are recorded and back-propagated at the position of the dashed line (from Raknes and Weibull, 2016a).

Figure 3. Seismic images and overlay plots of the seismic image and the corresponding V_P model for the shallow part in the Sleipner area. (a) The initial seismic image created using the tomography model, (b) the seismic image created using the FWI model, (c) overlay plot using the initial model, and (d) overlay plot using the FWI model (modified from Raknes et al., 2015).

Examples

The first example we show is taken from the Sleipner area, offshore Norway. The real data set used is a conventional short-offset streamer data set. Therefore, the inversion is restricted to the V_P model, whereas the other parameter models are computed using simple empirical relationships, as in the synthetic example above. An 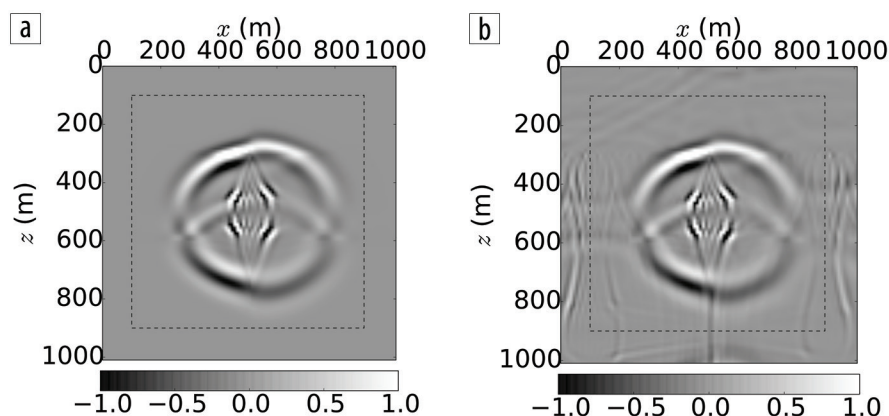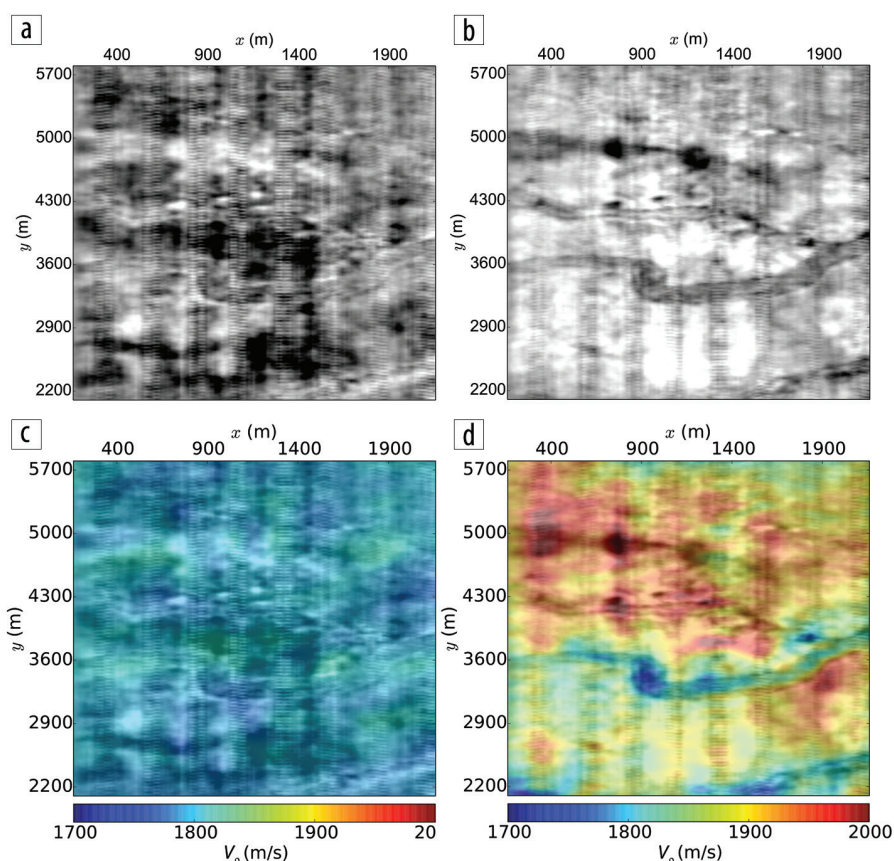informative view of the quality of the FWI model is made by an overlay plot of the seismic image and the corresponding velocity model. In Figure 3, we have included the seismic images and the corresponding overlay plots for the shallow part in the Sleipner area. The repeated vertical lines in the plots are imprints of the acquisition lines. The initial tomography model is smooth without large velocity changes. The corresponding seismic image (Figure 3a) is noisy, and details like the channel are not well

focused. FWI introduces large changes to the seismic image (Figure 3b) and the velocity model (Figure 3d). The most pronounced change between the seismic images is the focusing of the channel. In the FWI velocity model, the channel is estimated as a low-velocity zone. In general, there is a good correlation between the FWI model and the seismic image, which is not the case for the initial model and the seismic image.

In Figure 4, vertical slices through the migrated cubes are shown using the parameter model before and after FWI. In general we observe that the image after FWI is more focused with more continuous reflectors than the image before FWI. The improvements are most pronounced in the upper part of the cube, which is due to the short offsets in the data set resulting in lower resolution at greater depths in the model.

The second example illustrates the wavefield reconstruction technique and its influence on the FWI results. To create this example, we have used a submodel of the SEG/EAGE Overthrust model. We have created a synthetic multicomponent ocean-bottom data set that we use in the inversion for V_P and V_S. In Figure 5, vertical slices through the 3D cubes of V_P and V_S models are shown. We observe that there are small differences between the inversion results with and without the wavefield reconstruction method. Elastic FWI is able to resolve the shallow part of the V_P and V_S models. The deeper parts in the models are partly updated, but the details are not well resolved. This is a result of the limited frequency content used in the inversion, poor illumination at these depths, as well as the relatively few iterations performed to estimate the models.

Figure 4. Vertical profile through the seismic image cube using (a) the initial tomography model and (b) the final FWI model. The arrows on the seismic images indicate positions where there are improvements in the image between the initial and final images (from Raknes et al., 2015).

Figure 5. Vertical slice through the V_P and V_S synthetic model for the wavefield reconstruction FWI example. (a) The true V_P model, (b) the true V_S model, (c) the inverted V_P model not using wavefield reconstruction, (d) the inverted V_S model not using wavefield reconstruction, (e) the inverted V_P model using wavefield reconstruction, (f) the inverted V_P model using wavefield reconstruction (from Raknes and Weibull, 2016a).

The two examples show that it is possible to run elastic FWI in 3D within reasonable time and cost and with improvements in both the parameter models as well as the seismic images. Thus, the extra effort invested in performing elastic FWI may improve the understanding of the subsurface.

Conclusion

We have discussed why it is important to use FWI in a full 3D elastic computational setting to account for as much real-world physics as possible. Solving the 3D elastic FWI problem is challenging both from a computational and a theoretical point of view.
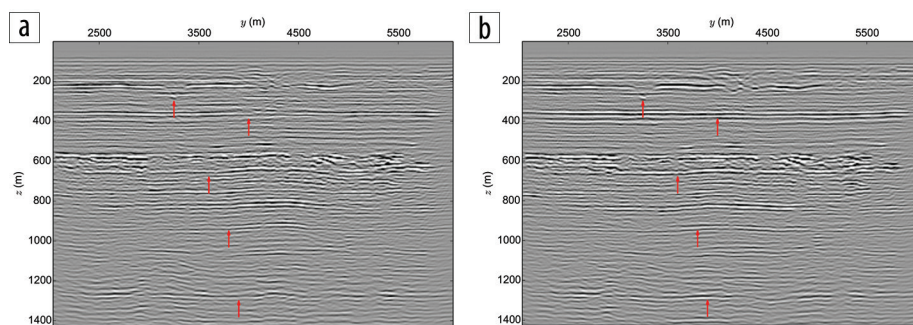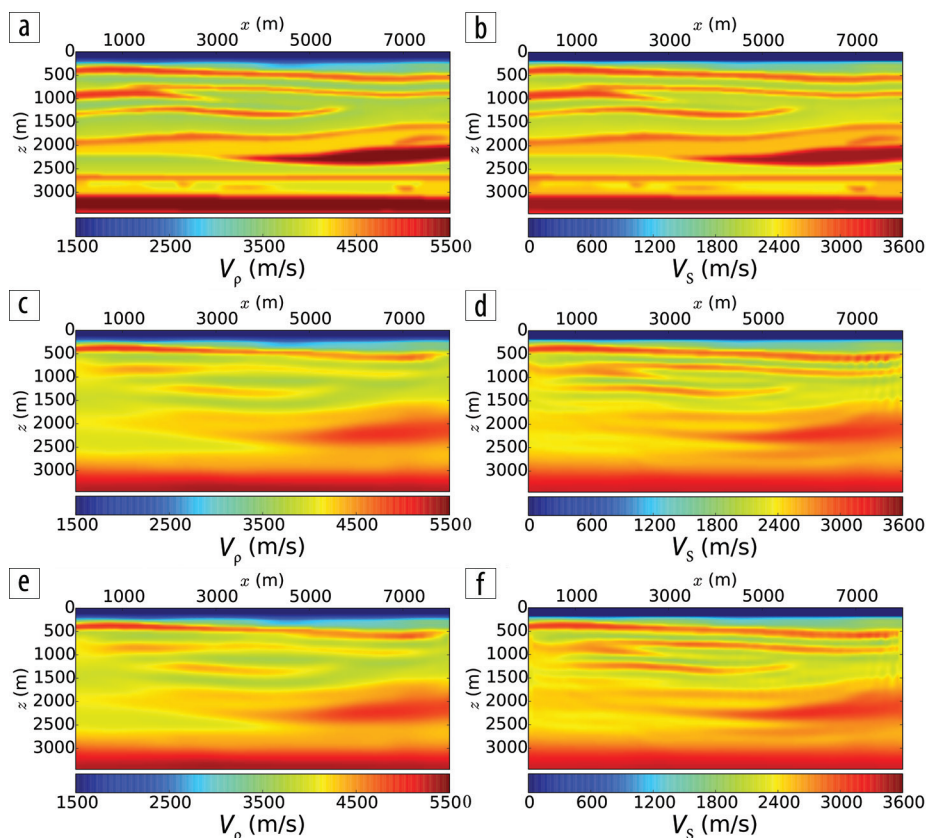
We have shown both real and synthetic examples that demonstrate the potential in using 3D elastic FWI in subsurface parameter estimation. Using an acoustic approach may result in wrongly estimated parameters due to the difference in the reflection and transmission coefficients. **TLE**

Corresponding author: espen.raknes@ntnu.no

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

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, http://dx.doi.org/10.1190/geo2012-0382.1.

Claerbout, J. F., 1971, Toward a unfied theory of reflector mapping: Geophysics, **36**, no. 3, 467–481, http://dx.doi.org/10.1190/1.1440185.

Evans, L. C., 2010, Partial Differential Equations, 2nd ed.: American Mathematical Society.

Igel, H., 2016, Computational Seismology: A Practical Introduction, 1st ed.: Oxford University Press.

Köhn, D., D. De Nil, A. Kurzmann, A. Przebindowska, and T. Bohlen, 2012, On the influence of model parametrization in elastic full waveform tomography: Geophysical Journal International, **191**, no. 1, 325–345, http://dx.doi.org/10.1111/j.1365-246X.2012.05633.x.

Leung, S., and J. Qian, 2006, An adjoint state method for three-dimensional transmission traveltime tomography using first-arrivals: Communications in Mathematical Sciences, **4**, no. 1, 249–266, http://dx.doi.org/10.4310/CMS.2006.v4.n1.a10.

Mora, P., 1987, Nonlinear two-dimensional elastic inversion of multioffset seismic data: Geophysics, **52**, no. 9, 1211–1228, http://dx.doi.org/10.1190/1.1442384.

Nguyen, B., and G. McMechan, 2015, Five ways to avoid storing source wave field snapshots in 2D elastic prestack reverse time migration: Geophysics, **80**, no. 1, S1–S18, http://dx.doi.org/10.1190/geo2014-0014.1.

Oldham, R. D., 1906, The constitution of the interior of the earth, as revealed by earthquakes: Quarterly Journal of the Geological Society, **62**, 456–475, http://dx.doi.org/10.1144/GSL.JGS.1906.062.01-04.21.

Operto, S., A. Miniussi, R. Brossier, L. Combe, L. Metivier, V. Monteiller, A. Ribodetti, and J. Virieux, 2015, Efficient 3-d frequency-domain mono-parameter full-waveform inversion of ocean-bottom cable data: application to Valhall in the viscoacoustic vertical transverse isotropic approximation: Geophysical Journal International, **202**, no. 2, 1362–1391, http://dx.doi.org/10.1093/gji/ggv226.

Pratt, R. G., 1999, Seismic waveform inversion in the frequency domain, Part 1: Theory and verification in a physical scale model: Geophysics, **64**, no. 3, 888–901, http://dx.doi.org/10.1190/1.1444597.

Prieux, V., G. Lambare, S. Operto, and J. Virieux, 2013, Building starting models for full waveform inversion from wide-aperture data by stereotomography: Geophysical Prospecting, **61**, 109–137, http://dx.doi.org/10.1111/j.1365-2478.2012.01099.x.

Raknes, E. B., B. Arntsen, and W. Weibull, 2015, Three-dimensional elastic full waveform inversion using seismic data from the Sleipner area: Geophysical Journal International, **202**, no. 3, 1877–1894, http://dx.doi.org/10.1093/gji/ggv258.

Raknes, E. B., and W. Weibull, 2016a, Efficient 3d elastic full-waveform inversion using wavefield reconstruction methods: Geophysics, **81**, no. 2, R45–R55, http://dx.doi.org/10.1190/geo2015-0185.1.

Raknes, E. B., and W. Weibull, 2016b, Combining wave-equation migration velocity analysis and full-waveform inversion for improved 3D elastic parameter estimation: 86th Annual International Meeting, SEG, Expanded Abstracts, 1320–1324, http://dx.doi.org/10.1190/segam2016-13858670.1.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, no. 8, 1259–1266, http://dx.doi.org/10.1190/1.1441754.

Venstad, J. M., 2016, Industry-scale finite-difference elastic wave modeling on graphics processing units using the out-of-core technique: Geophysics, **81**, no. 2, T35–T43, http://dx.doi.org/10.1190/geo2015-0267.1.

Wang, Y., L. Dong, Y. Liu, and J. Yang, 2016, 2D frequency-domain elastic full-waveform inversion using the block-diagonal pseudo-Hessian approximation: Geophysics, **81**, no. 5, R247–R259, http://dx.doi.org/10.1190/geo2015-0678.1.

Weiss, R. M., and J. Shragge, 2013, Solving 3D anisotropic elastic wave equations on parallel GPU devices: Geophysics, **78**, no. 2, F7–F15, http://dx.doi.org/10.1190/geo2012-0063.1.

Yang, P., R. Brossier, L. Metivier, and J. Virieux, 2016, Wave field reconstruction in attenuating media: a checkpointing-assisted reverse-forward simulation method: Geophysics, **81**, no. 6, R349–R362, http://dx.doi.org/10.1190/geo2016-0082.1.