Elastic full-waveform inversion for a synthetic VTI model from Valhall field

Nishant Kamath, Ilya Tsvankin & Esteban Díaz

ABSTRACT

One of the main challenges for full-waveform inversion (FWI) is taking into account both anisotropy and elasticity. Here, we perform FWI for a synthetic 2D elastic VTI (transversely isotropic with a vertical symmetry axis) model based on the geologic section at Valhall field in the North Sea. Multicomponent data are generated by a finite-difference code for surface acquisition. The model is parameterized in terms of the P- and S-wave vertical velocities ($V_{P0}$ and $V_{S0}$) and the P-wave NMO and horizontal velocities ($V_{nmo}$ and $V_{hor}$). We apply FWI in the time domain using a multiscale approach with three frequency bands. An approximate Hessian matrix, computed using the L-BFGS algorithm, is employed to scale the gradients of the objective function and improve the convergence. In the absence of significant diving-wave energy, FWI allows us to update the model primarily using reflection data. The largest and most accurate updates are obtained for the velocities $V_{P0}$ and $V_{hor}$. The vertical velocities $V_{P0}$ and $V_{S0}$ are updated both in the shallow and deeper parts of the model, whereas $V_{nmo}$ and $V_{hor}$ change primarily in the near-surface horizons. Further improvement is expected after extending the frequency range for the inversion.

1 INTRODUCTION

Full-waveform inversion (FWI), as originally proposed by Tarantola (1984), is designed to estimate an earth model that minimizes the difference between the observed and simulated seismic wavefields. The main advantage of this method is the possibility of achieving high resolution by employing the phase and amplitude information contained in the recorded waveforms.

FWI typically operates with diving-wave energy to update the long-wavelength (or smoothly varying) component of the velocity model. Once an accurate background model has been obtained using low frequencies, the FWI gradient becomes more sensitive to reflection events, which can help update the reflectivity and improve the resolution (Bunks et al., 1995).

In most implementations of FWI, only the shallow part of the model probed by diving waves can be updated to obtain an accurate background velocity field. In the deeper horizons, where the background velocity is incorrect, the reflectors are imaged at the wrong depths, which hampers model updates using reflected waves. Hence, reflection data are insufficient to update the model without an accurate background velocity field.

Most existing FWI algorithms assume the earth model to be isotropic and acoustic (Pratt, 1999; Sirgue and Pratt, 2004; Shin and Cha, 2008; Plessix et al., 2010; Baeten et al., 2013; Choi and Alkhalifah, 2013). Recent advances include methods for mitigating the nonlinearity of the problem (Choi and Alkhalifah, 2013; Alkhalifah, 2015) and using the entire bandwidth of the data at the same time (Biondi and Almomin, 2014). When FWI is performed for anisotropic media, it is typically done in the acoustic approximation. Plessix and Cao (2011), Gholami et al. (2013), and Alkhalifah and Plessix (2014) discuss the sensitivity of the objective function to different parameters of acoustic VTI media. They apply FWI for various combinations of the P-wave vertical ($V_{P0}$), NMO ($V_{nmo} = V_{P0} \sqrt{1 + 2\delta}$), and horizontal ($V_{hor} = V_{P0} \sqrt{1 + 2\delta}$) velocities, and the anisotropy coefficients $\epsilon$, $\delta$, and $\eta = (\epsilon - \delta)/(1 + 2\delta)$. Optimal parameter combinations for FWI depend on acquisition geometry, the data used for the inversion (diving waves or reflection events), and type of inversion (simultaneous or hierarchic). Although the acoustic approximation is computationally efficient, it cannot properly model reflection amplitudes and handle shear- and mode-converted waves.

Barnes et al. (2008) conduct a feasibility study for elastic FWI of cross-well seismic data from VTI media. For noise-free data, although $V_{P0}$, $V_{S0}$, and density are well-constrained, the anisotropy parameters $\epsilon$ and $\delta$ are resolved only in the laterally homogeneous parts of the model. Lee et al. (2010) perform FWI of synthetic surface data for 2D elastic VTI media. Parameterizing the model in terms of the stiffness coefficients, however, leads to ambiguity in their results.

Kamath and Tsvankin (2013) perform elastic FWI of multicomponent data from a horizontally layered VTI model. They show that it is possible to estimate $V_{P0}$, $V_{S0}$, $\epsilon$, and $\delta$ even for offset-to-depth ratios close to unity, primarily because the algorithm accurately models reflection coefficients. The
gradients of the FWI objective function for elastic anisotropic media are derived by Kamath and Tsvankin (2014a). They also develop a methodology for FWI in 2D elastic VTI media for different parameterizations and test it on models with Gaussian anomalies in the Thomsen parameters. Kamath and Tsvankin (2014c) present general expressions for “radiation” (i.e., sensitivity) patterns of model parameters for elastic anisotropic FWI and use this formalism to explain the FWI results for transmission data in 2D VTI media.

Here, we invert multicomponent surface data generated for a 2D elastic VTI model based on the geologic section at Valhall field. First, we review the inversion methodology based on the algorithm of Kamath and Tsvankin (2014a). Then we describe the model used to generate the “observed” wavefield and the processing applied to the simulated data. A multiscale approach is employed to perform FWI in three frequency bands. Finally, we discuss the preliminary results obtained for each range of frequencies.

2 METHODOLOGY

2.1 Inversion methodology

We carry out FWI in the time domain by minimizing the following objective function:

\[
\mathcal{F} = \frac{1}{2} \sum_{r=1}^{N} \int_0^T \| \mathbf{u}(x_r, t) - \mathbf{d}(x_r, t) \|^2 dt,
\]

where \( N \) is the number of receivers, \( T \) is the trace length, \( \mathbf{u}(x_r, t) \) is the displacement computed for a trial model at receiver location \( x_r \), and \( \mathbf{d}(x_r, t) \) is the recorded displacement. The gradient of \( \mathcal{F} \) with respect to the stiffness coefficients \( c_{ijkl} \) is obtained using the adjoint-state method (Kamath and Tsvankin, 2014a):

\[
\frac{\partial \mathcal{F}}{\partial c_{ijkl}} = - \int_0^T \frac{\partial \mathbf{u}_k}{\partial x_j} \frac{\partial \psi_h}{\partial x_l} dt,
\]

where \( \mathbf{u} \) and \( \psi \) are the forward and adjoint displacement fields, respectively. The gradient for a chosen model parameter \( m_n \) is obtained as:

\[
\frac{\partial \mathcal{F}}{\partial m_n} = \sum_{ijkl} \frac{\partial \mathcal{F}}{\partial c_{ijkl}} \frac{\partial c_{ijkl}}{\partial m_n}.
\]

Here, we define the model in terms of the velocities \( V_{P0}, V_{S0}, V_{S0n}, \) and \( V_{S0r} \) (parameterization I in Kamath and Tsvankin, 2013). The gradients of the objective function \( \mathcal{F} \) with respect to the velocities are given in Appendix A.

The 2D elastic finite-difference modeling code from the Madagascar package is used to generate the data. To obtain the model update at each iteration, the gradient is scaled by a factor which can be chosen in different ways. In the widely used steepest-descent method, the gradient is scaled by the step length. This approach, however, does not account for the energy loss due to geometric spreading. In the absence of a Hessian term (or its approximation), the gradient has too high amplitudes near the sources and receivers, which complicates velocity estimation. Therefore, the gradients at the source and receiver locations have to be masked to ensure meaningful model updates. In addition, the steepest-descent technique could result in slow convergence if the objective function has a long and narrow valley.

The problems associated with the steepest-descent algorithm can be circumvented by applying the Gauss-Newton method and computing the model update \( \Delta m \) from

\[
\Delta m = [\mathcal{H}]^{-1} G,
\]

where \( \mathcal{H} \) is the Hessian operator and \( G \) is the gradient of the objective function. Computation of the Hessian, however, is extremely expensive. The Broyden-Fletcher-Goldfarb-Shannon (BFGS) algorithm or its low-memory equivalent (L-BFGS) is often used to obtain an approximate Hessian. We use the L-BFGS package of Nocedal (1980) to compute the operator \( \mathcal{H} \) and update the model from equation 4.

2.2 Synthetic model and data processing

The synthetic model used here is based on the parameters estimated for Valhall field in the North Sea (Munns, 1985). The fields of \( V_{P0}, V_{S0}, \) and \( \delta \) were provided to us by Romain Brossiere (Université Joseph Fourier, Grenoble) and Olav Barkved (BP, Norway). The \( V_{P0} \) field (Figure 1(a)) includes low-velocity gas layers above the reservoir, which is located at a depth of 2.5 km. The original model is sampled every 3.125 m in the horizontal and vertical directions. However, to reduce computational time in the finite-difference modeling, the grid spacing is increased to 20 m. Although the original Valhall model has a water column on top, we make the entire section elastic and clip the shear-wave velocity \( V_{S0} \) so that its minimum value near the surface is 700 m/s.

Multicomponent synthetic data are generated by a horizontal array of sources placed with a 80 m increment at a depth of 20 m. The receivers are located at every grid point at the same depth (20 m). The wavefield is generated by a vertical point source; the source signal is a Ricker wavelet with a peak frequency of 5 Hz.

The actual parameter fields are smoothed to generate the starting model. The smoothing is applied directly to the velocities \( V_{P0}, V_{S0}, V_{S0n}, \) and \( V_{S0r} \) (Figure 2). The L-BFGS implementation requires putting bounds on the model parameters, so we restrict each velocity to lie between the maximum and minimum values for the actual model. The density is fixed at the correct value and not updated during the inversion. FWI is carried out using a multiscale approach, with 20 iterations for each of the following frequency bands: \( 0 - 1.5 \text{ Hz}, 0 - 3 \text{ Hz}, \) and \( 0 - 5 \text{ Hz} \).

3 PRELIMINARY RESULTS

Because of the lack of significant diving-wave energy in the observed (modeled) data, the updates are based primarily on reflections. Comparison of the difference between the actual and starting velocity fields (Figure 3) with the velocity updates
Elastic FWI for 2D VTI media

Figure 1. Parameters (a) $V_p^0$, (b) $V_s^0$, (c) $\epsilon$, and (d) $\delta$ of a synthetic VTI model based on sections from the Valhall field. The velocities have units of km/s.

Figure 2. Starting model for FWI: (a) $V_p^0$, (b) $V_s^0$, (c) $V_{nmo}$, and (d) $V_{hor}$.

Figure 3. Difference between the actual and starting models: (a) $V_p^0$, (b) $V_s^0$, (c) $V_{nmo}$, and (d) $V_{hor}$. The units for the velocity differences updat es here and in the plots below are m/s.

Figure 4. Model updates (with respect to the starting model in Figure 2) for the frequency range 0−1.5 Hz: (a) $V_p^0$, (b) $V_s^0$, (c) $V_{nmo}$, and (d) $V_{hor}$.
for the first frequency band (Figure 4) indicates the magnitude of the improvements. The \( V_{P0} \)- and \( V_{S0} \)-fields are updated throughout the section, although the most pronounced changes are above 2.5 km. The updates in the parameters \( V_{nmo} \) and \( V_{hor} \), however, are restricted to the shallow part of the model (down to 1.5 km). Although the largest updates amount to only about 10% of the required change, they have the correct sign (e.g., compare figures 3(a) and 4(a)).

The radiation patterns (Kamath and Tsvankin, 2014b) of the model parameters help explain the obtained results. A \( V_{P0} \)-anomaly on a horizontal reflector scatters P-wave energy mostly along the vertical symmetry axis. In the case of an anomaly in \( V_{hor} \), most of the energy is scattered near the isotropy plane, which explains why the updates in the horizontal velocity are concentrated in the shallow part of the model. The maximum energy scattered by a \( V_{nmo} \)-anomaly is only about 25% of that for \( V_{P0} \), and is focused at angles near 45°. Therefore, the algorithm cannot update \( V_{nmo} \) at depths below 2 km. A scatterer in \( V_{S0} \) has a pattern similar to that in \( V_{nmo} \), but the maximum energy is higher (comparable to that of \( V_{P0} \), assuming \( V_{P0}/V_{S0} = 2 \)). Hence, there is a trade-off between the P- and S-wave vertical velocities.

The inversion result for the first frequency band is used as the starting model for the second one (0–3 Hz). Including higher frequencies improves model resolution, with the low-velocity updates in Figure 5 helping identify the gas clouds (Figure 1). Also, the magnitude of the updates is substantially increased (compare Figures 4 and 5). On the whole, the inversion updates the \( V_{P0} \)- and \( V_{hor} \)-fields more significantly than those of \( V_{S0} \) and \( V_{nmo} \).

The horizontal and vertical components of the data misfit computed for the starting and final models demonstrate significant improvement in waveform matching (Figure 6). The decrease in data misfit is particularly visible at the far offsets on the horizontal component (Figures 6(a) and 6(b)) and the near offsets on the vertical component (Figures 6(c) and 6(d)).

The L-BFGS algorithm allows the user to specify the desired number of calculations of the gradients, but the actual number of model updates depends on the values of the objective function and the gradients. Hence, although we compute 50 and 70 gradients in the first and second frequency bands, respectively, the number of model updates is only 18 and 31 (Figure 7).

The gradients are smoothed before each model update, with the same smoothing operator applied to all four velocities. Because \( V_{S0} \) is much smaller than the three P-wave velocities, the updates in \( V_{S0} \) have larger wavenumbers. The artifacts in the inverted \( V_{S0} \)-field (Figures 4(b) and 5(b)) might be due to the high shear-wave energy near the source. The Hessian computed by the L-BFGS algorithm is approximate and may be unable to effectively suppress the energy at the source location.

4 CONCLUSIONS

We presented preliminary results of elastic FWI for a 2D synthetic VTI model fashioned after the geologic section at Valhall field in the North Sea. The model includes low-velocity lenses above the reservoir, which are supposed to represent gas clouds.

The inversion is carried out in the time domain using the
Figure 7. Change in the normalized objective function with iterations for the (a) first and (b) second frequency bands.

5 ACKNOWLEDGMENTS

We are grateful to the members of the A(nisotropy)- and i(maging)- teams at CWP and to Chunlei Chu and Phuong Vu (BP, Houston) for fruitful discussions. We would also like to thank Romain Brossiere (Université Joseph Fourier, Grenoble) and Olav Barkved (BP, Norway) for providing the synthetic model. This work was supported by the Consortium Project on Seismic Inverse Methods for Complex Structures at CWP and by the CIMMM Project of the Unconventional Natural Gas Institute at CSM. The reproducible numerical examples in this paper are generated with the Madagascar open-source software package freely available from http://www.ahay.org.

REFERENCES

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.
APPENDIX A: GRADIENTS FOR DIFFERENT PARAMETERIZATIONS

The gradient of the objective function with respect to the stiffness coefficient $c_{ijkl}$ is given in equation 2. Expressing the stiffnesses in terms of chosen model parameters $m_n$ makes it possible to obtain the gradients with respect to $m_n$ from equation 3. Kamath and Tsvankin (2014a) derive the following gradients for VTI media with respect to the velocities $V_P^0$, $V_S^0$, $V_{nmo}$, and $V_{hor}$:

$$\frac{\partial F}{\partial V_P^0} = -2\rho V_P^0 \int_0^T \left[ \frac{\partial \psi_3}{\partial x_3} \frac{\partial u_3}{\partial x_3} + \frac{1}{2} \frac{V_{nmo,P}^2 - V_{S0}^2}{V_P^2 - V_{S0}^2} \left( \frac{\partial \psi_1}{\partial x_1} \frac{\partial u_3}{\partial x_3} + \frac{\partial \psi_3}{\partial x_3} \frac{\partial u_1}{\partial x_1} \right) \right] dt,$$  \hspace{1cm} (A1)

$$\frac{\partial F}{\partial V_S^0} = -2\rho V_S^0 \int_0^T \left\{ \left[ \frac{2V_{S0}^2 - V_P^2 - V_{nmo,P}^2}{2(\sqrt{V_{nmo,P}^2 - V_{S0}^2})(V_P^2 - V_{S0}^2)} - 1 \right] \left( \frac{\partial \psi_3}{\partial x_1} \frac{\partial u_3}{\partial x_3} + \frac{\partial \psi_3}{\partial x_3} \frac{\partial u_1}{\partial x_1} \right) \right. \right.$

$$\left. + \left( \frac{\partial \psi_3}{\partial x_3} + \frac{\partial \psi_1}{\partial x_1} \right) \left( \frac{\partial u_1}{\partial x_3} + \frac{\partial u_3}{\partial x_1} \right) \right\} dt,$$ \hspace{1cm} (A2)

$$\frac{\partial F}{\partial V_{nmo}} = -\rho V_{nmo} \int_0^T \sqrt{\frac{V_P^2 - V_{S0}^2}{V_{nmo,P}^2 - V_{S0}^2}} \left( \frac{\partial \psi_1}{\partial x_1} \frac{\partial u_3}{\partial x_3} + \frac{\partial \psi_3}{\partial x_3} \frac{\partial u_1}{\partial x_1} \right) dt,$$ \hspace{1cm} (A3)

$$\frac{\partial F}{\partial V_{hor}} = -2\rho V_{hor} \int_0^T \frac{\partial \psi_1}{\partial x_1} \frac{\partial u_1}{\partial x_1} dt.$$ \hspace{1cm} (A4)