GEOPHYSICS, VOL. 79, NO. 1 (JANUARY-FEBRUARY 2014); P. R13–R25, 12 FIGS.
10.1190/GEO2013-0108.1

Anisotropic migration velocity analysis using reverse-time migration

Wiktor Waldemar Weibull[1] and Børge Arntsen[1]

ABSTRACT

Seismic anisotropy, if not accounted for, can cause significant mispositioning of the reflectors in depth-migrated images. Accounting for anisotropy in depth migration requires velocity analysis tools that can estimate the anisotropic background velocity field. We extended wave equation migration velocity analysis to deal with 2D tilted transverse isotropic media. The velocities were obtained automatically by nonlinear optimization of the focusing and stack power of common-image point gathers constructed using an extended imaging condition. We used the elastic two-way wave equation to reconstruct the wavefields needed for the image and gradient computations. This led to an anisotropic migration velocity analysis algorithm based on reverse-time migration. We illustrated the method with synthetic and field data examples based on marine surface seismic acquisition. The results showed that the method significantly improves the quality of the depth-migrated image. However, as is common in the case of velocity analysis using surface seismic data, the estimation of anisotropic parameters seems to be strongly nonunique.

INTRODUCTION

Wave equation migration velocity analysis (WEMVA) can be described as a nonlinear least-squares inversion of prestack seismic reflection data in the image domain (Sava and Biondi, 2004). The procedure consists in setting up an objective function that can measure the misfit in the image domain due to a prestack depth migration with a nonoptimal velocity model and then minimizing this function with respect to the velocity parameters. The objective function for WEMVA is typically based on the focusing of common-image point gathers (CIGs) (Symes and Kern, 1994; Shen et al., 2003), stacking-power (Toldi, 1989; Chavent and Jacewitz, 1995), or a combination of both (Mulder, 2008; Shen and Symes, 2008). The procedure can be made fully automatic, and it is, to a certain extent, robust against poor initial guesses of the velocity field (Shen and Symes, 2008). But due to the assumptions of prestack depth migration, WEMVA is restricted to kinematic inversion of single scattering reflection data (Mulder and van Leeuwen, 2008).

A vertically transverse isotropic (VTI) model can be a good approximation to some horizontally or nearly horizontally layered sequences (Levin, 1979; Banik, 1984; Sayers, 1994). But for tectonically deformed geologic settings, such as fold thrust belts or the flanks of salt diapirs, a tilted transverse isotropic (TTI) model is a better approximation (Isaac and Lawton, 1999). A general 2D TTI medium can be described by five spatially varying parameters: (1) the P-wave velocity along the symmetry axis V_{P0}, (2) the S-wave velocity along the symmetry axis V_{S0}, Thomsen (1986) parameters (3) ε and (4) δ, (5) and the tilt θ of the symmetry axis with respect to the vertical. However, some assumptions can be used to reduce the number of parameters needed to describe the kinematics of the medium. In the context of P-wave velocity analysis, Tsvankin and Thomsen (1994) and Alkhalifah and Larner (1994) demonstrate that V_{S0} can be arbitrarily chosen. Also, a popular assumption, often referred to as *structural transverse isotropy* (STI), further reduces the number of parameters by assuming that θ is always perpendicular to the structure of the reflectors (Audebert et al., 2006).

Most implementations of WEMVA are based on the acoustic isotropic approximation (Sava and Vlad, 2008). However, in cases where the velocity field is anisotropic, velocity analysis under an isotropic assumption will ultimately lead to images that are well focused but mispositioned in space (Isaac and Lawton, 1999). In an attempt to overcome this problem, some research implement WEMVA under anisotropic assumptions. Li and Biondi (2011) propose a method based on the depth-oriented extension of the differential semblance objective function (Shen et al., 2003) and one-way wave equation migration in a VTI medium. Li et al. (2012) also present a method based on a similar objective function, but using the pseudoacoustic two-way wave equation for a VTI medium (Alkhalifah, 1998). Weibull et al. (2012) present a method where

Manuscript received by the Editor 13 March 2013; revised manuscript received 29 May 2013; published online 4 December 2013.
[1]Norwegian University of Science and Technology, Institute for Petroleum Technology and Applied Geophysics, Trondheim, Norway. E-mail: wiktor.weibull@ntnu.no; borge.arntsen@ntnu.no

they use an objective function based on the depth-oriented extension of differential semblance combined with stacking-power maximization (Toldi, 1989) and elastic reverse-time migration (RTM) to estimate anisotropic parameters over a VTI medium.

In this paper, we extend WEMVA to deal with a 2D TTI model of the subsurface and test it on synthetic and field surface seismic data. To account for anisotropy in the kinematics of wave propagation, we use a density-normalized elastic wave equation that is stable and accurately propagates waves at all angles, which is important for the estimation of anisotropic parameters. We use WEMVA to simultaneously estimate V_{P0}, ε, and δ. The parameter V_{S0} is assumed to have a negligible influence on P-wave propagation and is chosen arbitrarily. In addition, θ is assumed to conform to the geology and is estimated from the structure of the reflectors in the migrated image.

A major difficulty in the estimation of anisotropic velocities from the kinematics of surface reflection data is the inherent nonuniqueness related to the positioning of the reflectors in the subsurface and/or to the trade-off between the different parameters (Vestrum et al., 1999; Grechka et al., 2002). The trade-off between heterogeneity and anisotropy can, in principle, be reduced by considering the images in vertical time instead of depth, as suggested by Alkhalifah et al. (2001) or by considering a stretched depth axis as in Plessix and Rynja (2010). In this work, we use regularization to constrain the models to a physical set, and a coarse bicubic B-spline grid to confine the models to a sparse solution space. These measures help to obtain a convergent WEMVA algorithm, but they are not sufficient to obtain a unique geologic model of the subsurface. In practice, substantial additional information in the form of well logs and check-shot surveys are required to narrow down the range of possible solutions to the problem (Yan et al., 2004; Bakulin et al., 2010).

This paper starts by explaining the method and showing what can be expected in the ideal condition. The method is then tested on 2D synthetic and field data sets. Next, we discuss the main results and suggest potential ways forward. Finally, we present our conclusions.

2D TTI RTM

The basis for WEMVA is prestack depth migration. To build CIGs for velocity analysis, we use RTM with an extended imaging condition (Rickett and Sava, 2002):

$$\mathcal{R}(\mathbf{x},\mathbf{h}) = \int \mathrm{d}s \int \mathrm{d}t W^s(\mathbf{x}-\mathbf{h},t,s) W^r(\mathbf{x}+\mathbf{h},T-t,s), \tag{1}$$

where W^s are the forward-modeled source wavefields; W^r are the reverse time modeled receiver wavefields; $\mathbf{x} = (x,z)$ are the spatial coordinates, with z being the depth axis; $\mathbf{h} = (h_x,0)$ is the subsurface horizontal half-offset; t is the time; and s is the source index.

The computation of the W^s and W^r wavefields depends on the choice of the wave equation. To take anisotropy into account, an anisotropic wave equation must be used in the reconstruction of the wavefields. We model the wave propagation in a 2D TTI medium using a density-normalized elastic wave equation (Ikelle and Amundsen, 2005):

$$\frac{\partial^2 u_i}{\partial t^2}(\mathbf{x},t) - \frac{\partial}{\partial x_j}\left[a_{ijkl}(\mathbf{x})\frac{\partial u_l}{\partial x_k}(\mathbf{x},t)\right] = F_i(\mathbf{x},t), \tag{2}$$

where u_i is the displacement field, a_{ijkl} is the density-normalized elasticity tensor, F_i is a source term, and $i,j,k,l = x,z$ are indexes under the Einstein summation convention. The elastic wave equation and the density-normalized elastic parameters a_{ijkl} are described in more detail in Appendix A.

To obtain the W^s and W^r wavefields using equation 2, we first model the displacement vector wavefields, u_i^s and u_i^r, according to the following equations:

$$\frac{\partial^2 u_i^s}{\partial t^2}(\mathbf{x},t) - \frac{\partial}{\partial x_j}\left[a_{ijkl}(\mathbf{x})\frac{\partial u_l^s}{\partial x_k}(\mathbf{x},t)\right] = \frac{\partial S}{\partial x_i}(x_s,t,s), \tag{3}$$

and

$$\frac{\partial^2 u_i^r}{\partial t^2}(\mathbf{x},t) - \frac{\partial}{\partial x_j}\left[a_{ijkl}(\mathbf{x})\frac{\partial u_l^r}{\partial x_k}(\mathbf{x},t)\right] = \frac{\partial P}{\partial x_i}(x_r,T-t,s), \tag{4}$$

for the source and receiver displacements, respectively. In equation 3, S is the source time function for source s, at location $\mathbf{x_s}$. Whereas in equation 4, P is the time-reversed recorded seismic reflection data for source s at receiver positions given by $\mathbf{x_r}$. Also, note that equation 3 is to be solved forward in time, while equation 4 is to be solved in reverse time.

We then extract a scalar wavefield from the source and receiver displacement wavefields by taking the divergence of the displacement vector scaled by the density-normalized bulk modulus:

$$W^s(\mathbf{x},t,s) = V_{P0}^2(\mathbf{x})\frac{\partial u_i^s}{\partial x_i}(\mathbf{x},t,s), \tag{5}$$

and

$$W^r(\mathbf{x},t,s) = V_{P0}^2(\mathbf{x})\frac{\partial u_i^r}{\partial x_i}(\mathbf{x},t,s). \tag{6}$$

Here, the divergence operator acts to attenuate the S-wave modes present in the elastic wavefields and thus produces approximate P-wave scalar wavefields. While this operation perfectly separates P- and S-waves in 2D isotropic media, in anisotropic media, because the polarizations of P- and S-wave modes are nonorthogonal, some S-wave modes leak into the resulting scalar wavefields (Dellinger and Etgen, 1990). However, the residual S-waves are generally of small amplitude. In addition, due to the large difference in phase velocity between the P- and S-waves, the latter tend to "stack out" during imaging, where the phase velocities of the P-waves are used. For these reasons, the divergence operator is a cheap and effective operator for PP imaging. A better separation can be achieved by designing spatially variant operators based on the solution of the Christofel equation, but at a much higher computational cost (Yan and Sava, 2011).

Note that, different from the pseudoacoustic approximation of Alkhalifah (1998), equation 2 requires V_{S0} to be provided. However, if only the kinematics of P-wave propagation are considered, the S-wave velocities are of minor importance (Alkhalifah and

Tsvankin, 1995). In this work, the S-wave velocities are heuristically chosen to be 0.9 km/s and are spatially invariant.

WEMVA

We quantify a misfit in the prestack depth-migrated image using the same objective function that Shen and Symes (2008) describe. The objective function consists of a combination of the depth-oriented extension of differential semblance optimization (Shen et al., 2003) with stack-power maximization (Toldi, 1989; Chavent and Jacewitz, 1995; Zhou et al., 2009). The objective function can be written as

$$\begin{aligned}\mathcal{J} &= \frac{1}{2}\left\|\mathbf{h}\frac{\partial \mathcal{R}}{\partial z}(\mathbf{x},\mathbf{h})\right\|^2 - \frac{\gamma}{2}\left\|\frac{\partial \mathcal{R}}{\partial z}(\mathbf{x},0)\right\|^2 \\ &= \frac{1}{2}\int \mathbf{dx}\int \mathbf{dh}\hat{\mathbf{h}}\left[\frac{\partial \mathcal{R}}{\partial z}(\mathbf{x},\mathbf{h})\right]^2, \end{aligned} \tag{7}$$

where $\hat{\mathbf{h}} = \mathbf{h}^2 - \gamma\delta(\mathbf{h})$, with δ being the Kronecker delta and γ being a constant weight that balances the contribution of differential semblance and stack power to the total value of the objective function. The spatial derivative operator ∂_z attenuates the low vertical wavenumber components in the RTM image (Guitton et al., 2007). This helps to improve the stability and convergence properties of WEMVA using the two-way wave equation (Mulder, 2008; Weibull and Arntsen, 2011). We use the vertical derivative operator for its simplicity and robustness. There are more sophisticated noise-reducing imaging conditions available, which also attenuate horizontally oriented noise, as well as better preserve vertical reflectors (Douma et al., 2010; Whitmore and Crawley, 2012).

In addition to quantifying the misfit in the CIGs, we use regularization to constrain the parameters to a feasible set and also to prevent excessive roughness in the solution (Tikhonov and Arsenin, 1977). The regularization is implemented by adding the following term to the objective function:

$$\begin{aligned}\mathcal{J}_R = &\frac{\alpha_1}{2}\left\|\frac{\partial V_{\mathrm{P0}}}{\partial x_i}(\mathbf{x}) - \frac{\partial V^0_{\mathrm{P0}}}{\partial x_i}(\mathbf{x})\right\|^2 + \frac{\beta_1(\mathbf{x})}{2}\|V_{\mathrm{P0}}(\mathbf{x}) - V^0_{\mathrm{P0}}(\mathbf{x})\|^2 \\ &+ \frac{\alpha_2}{2}\left\|\frac{\partial \varepsilon}{\partial x_i}(\mathbf{x}) - \frac{\partial \varepsilon^0}{\partial x_i}(\mathbf{x})\right\|^2 + \frac{\beta_2(\mathbf{x})}{2}\|\varepsilon(\mathbf{x}) - \varepsilon^0(\mathbf{x})\|^2 \\ &+ \frac{\alpha_3}{2}\left\|\frac{\partial \delta}{\partial x_i}(\mathbf{x}) - \frac{\partial \delta^0}{\partial x_i}(\mathbf{x})\right\|^2 + \frac{\beta_3(\mathbf{x})}{2}\|\delta(\mathbf{x}) - \delta^0(\mathbf{x})\|^2, \end{aligned} \tag{8}$$

where V^0_{P0}, ε^0, and δ^0 represent initial values of the target parameters and α_n and β_n, with $n = 1, 2, 3$, being constant weights, one for each spatially varying parameter.

One limitation of the objective function given by equation 7 is that it is strictly valid under the single scattering assumption. Because multiples and primaries will focus at different velocities in the CIGs, the presence of multiples will introduce local minima in the objective function. One of the simplest solutions to this problem is to include multiple attenuation as a part of the preprocessing of the data used for velocity analysis and use absorbing boundary conditions at all sides for the source and receiver wavefield reconstructions. In case multiple attenuation fails, there are other approaches that might be useful. Mulder and van Leeuwen (2008) propose to reduce the influence of the low-velocity free-surface multiples by modifying the objective function with an asymmetric weighting function. Another proposed method consists in modeling and subtracting the multiples as a part of the velocity analysis (van Leeuwen and Mulder, 2008).

To check the sensitivity of the objective function to the parameters V_{P0}, ε, and δ, we evaluate the objective function in a simple 1D three-layer model, as shown in Figure 1a. In a first stage, we used this acoustic model to simulate surface seismic data using a split-spread geometry with a maximum offset of 1.4 km. We used a monopole point source, consisting of a Ricker pulse with dominant frequency of 15 Hz. No free surface was used in the modeling. The preprocessing of the data was limited to muting the direct wave and wide angle reflections. The resulting data are shown in

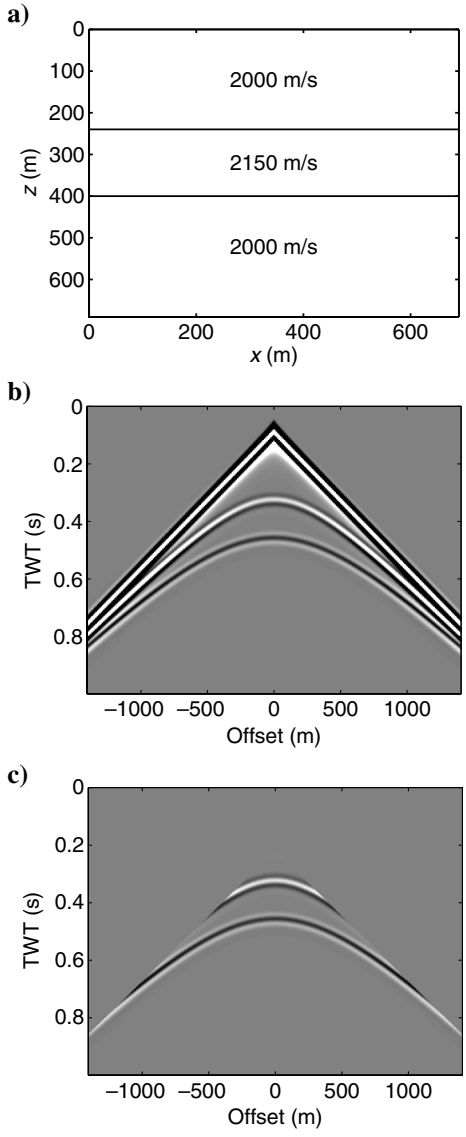

Figure 1. Model and data used to generate the sensitivity plots in Figure 2; (a) 1D velocity model. (b) Synthetic shot gather modeled using the model in (a). (c) Shot gather in (b) after mute to remove direct wave and postcritical reflections.

Figure 1b and 1c. The data are virtually single scattering and therefore ideal within the assumptions of the method. Next, we perturbed the magnitude of the true parameters in the second layer for different values of V_{P0}, ε, and δ and migrated the data. In this procedure, only one parameter is perturbed at each time, leaving the other parameters set at their true value. The tests were conducted two times, one time with θ fixed to 0°, and a second time with θ fixed at 45°. We then used the resulting images to compute the objective function values for three different combinations of the misfit functions: For stack power alone, for differential semblance alone, and for the combination of stack power and differential semblance.

The results in Figure 2 show a comparison of the variation of the values of the different objective functions for each parameter. Figure 2a–2c shows the results with θ fixed at 0°, while Figure 2d–2f shows the results with θ fixed at 45°. What this idealized experiment shows is that the misfit functions are quasiconvex for a wide range of model perturbations, and they can therefore be suitable for gradient-based optimization. Another fact shown is that, in this ideal case, if all but one parameter are known precisely, the unknown parameter can be uniquely determined. Within the Thomsen's parameters, the objective functions are more sensitive to errors in ε than they are to errors in δ. The effect of having the wrong θ model is largest for δ than it is for ε. And because the test model is isotropic, the sensitivity to the choice of θ is zero for the V_{P0} tests.

Gradient computation

To minimize the objective function, we use an L-BFGS method (Byrd et al., 1995; Nocedal and Wright, 2000). This method requires the evaluation of the objective function and its gradient with respect to the parameters at each iteration or line search step. We compute the gradient of equation 7 using the adjoint state method (Lions and Magenes, 1972; Chavent and Lemonnier, 1974; Plessix, 2006). This method gives the following equations for the gradients with respect to $V_{P0}(\mathbf{x})$, $\varepsilon(\mathbf{x})$, and $\delta(\mathbf{x})$:

$$\begin{aligned}\frac{\partial \mathcal{J}}{\partial V_{P0}}(\mathbf{x}) = &\int ds \int dt \frac{\partial a_{ijkl}}{\partial V_{P0}}(\mathbf{x}) \frac{\partial u_l^s}{\partial x_k}(\mathbf{x},t,s) \frac{\partial \tilde{u}_i^s}{\partial x_j}(\mathbf{x},T-t,s)\\ &+ \int ds \int dt \frac{\partial a_{ijkl}}{\partial V_{P0}}(\mathbf{x}) \frac{\partial u_l^r}{\partial x_k}(\mathbf{x},T-t,s) \frac{\partial \tilde{u}_i^r}{\partial x_j}(\mathbf{x},t,s)\\ &+ \int ds \int dt 2V_{P0}(\mathbf{x}) \frac{\partial u_i^s}{\partial x_i}(\mathbf{x},t,s)\\ &\times \int dh \frac{\partial^2 \mathcal{R}}{\partial z^2}(\mathbf{x}+\mathbf{h},\mathbf{h}) W^r(\mathbf{x}+2\mathbf{h},T-t,s)\\ &+ \int ds \int dt 2V_{P0}(\mathbf{x}) \frac{\partial u_i^r}{\partial x_i}(\mathbf{x},T-t,s)\\ &\times \int dh \frac{\partial^2 \mathcal{R}}{\partial z^2}(\mathbf{x}-\mathbf{h},\mathbf{h}) W^s(\mathbf{x}-2\mathbf{h},t,s), \end{aligned} \tag{9}$$

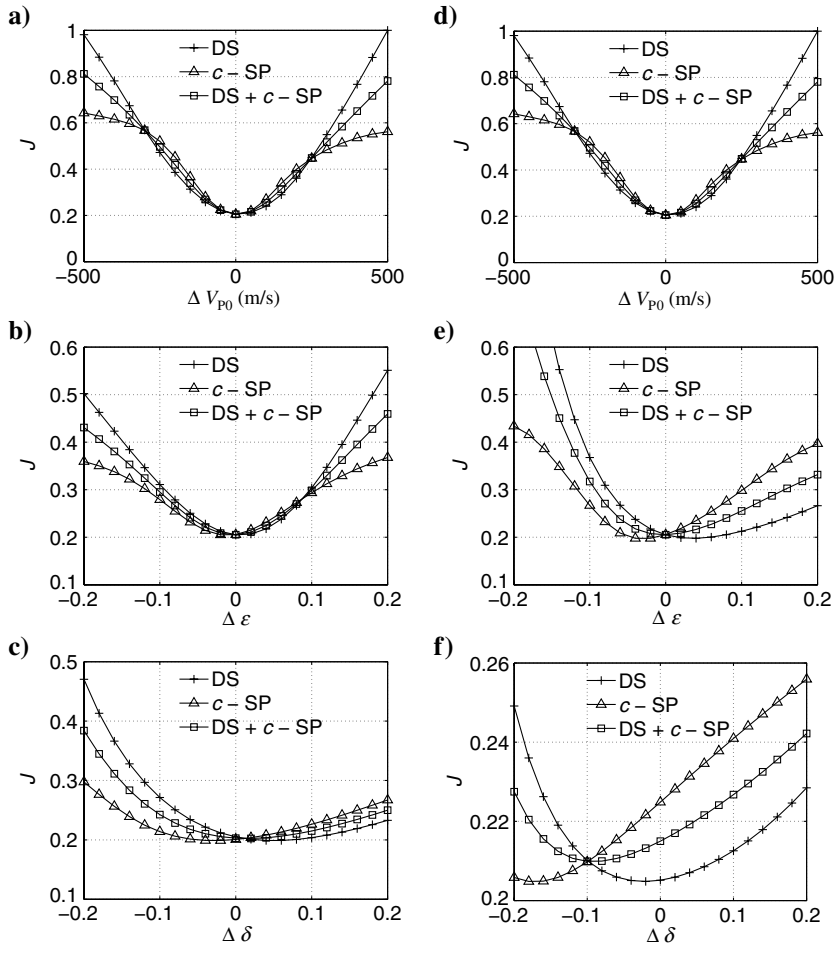

Figure 2. Objective function value computed from migrating the data in Figure 1c using the model in Figure 1a modified by perturbations in (a) V_{P0}, (b) ε, (c) δ. Panels (d-f) are similar to (a-c) but with θ set to 45°. The DS corresponds to the differential semblance misfit function, while the SP corresponds to stack power, with c being a constant added so that the minimum of −SP is equal to the minimum of DS.

$$\frac{\partial \mathcal{J}}{\partial \varepsilon}(\mathbf{x}) = \int \mathrm{d}s \int \mathrm{d}t \frac{\partial a_{ijkl}}{\partial \varepsilon}(\mathbf{x}) \frac{\partial u_l^s}{\partial x_k}(\mathbf{x}, t, s) \frac{\partial \tilde{u}_i^s}{\partial x_j}(\mathbf{x}, T-t, s) + \int \mathrm{d}s \int \mathrm{d}t \frac{\partial a_{ijkl}}{\partial \varepsilon}(\mathbf{x}) \frac{\partial u_l^r}{\partial x_k}(\mathbf{x}, T-t, s) \frac{\partial \tilde{u}_i^r}{\partial x_j}(\mathbf{x}, t, s), \tag{10}$$

and

$$\frac{\partial \mathcal{J}}{\partial \delta}(\mathbf{x}) = \int \mathrm{d}s \int \mathrm{d}t \frac{\partial a_{ijkl}}{\partial \delta}(\mathbf{x}) \frac{\partial u_l^s}{\partial x_k}(\mathbf{x}, t, s) \frac{\partial \tilde{u}_i^s}{\partial x_j}(\mathbf{x}, T-t, s) + \int \mathrm{d}s \int \mathrm{d}t \frac{\partial a_{ijkl}}{\partial \delta}(\mathbf{x}) \frac{\partial u_l^r}{\partial x_k}(\mathbf{x}, T-t, s) \frac{\partial \tilde{u}_i^r}{\partial x_j}(\mathbf{x}, t, s). \tag{11}$$

Note that because we scale the divergence of displacement vector by the density-normalized bulk modulus V_{P0}^2, the gradient with respect to V_{P0} (equation 9) has two additional terms in the right-hand side, when compared to the formula for the other gradients (equations 10 and 11).

The adjoint state wavefields $\tilde{u}_i^s$ and $\tilde{u}_i^r$ can be computed by the following adjoint modelings:

$$\frac{\partial^2 \tilde{u}_i^s}{\partial t^2}(\mathbf{x}, t) - \frac{\partial}{\partial x_j}\left[a_{ijkl}(\mathbf{x}) \frac{\partial \tilde{u}_l^s}{\partial x_k}(\mathbf{x}, t)\right] = \frac{\partial A^s}{\partial x_i}(\mathbf{x}, T-t, s) \tag{12}$$

and

$$\frac{\partial^2 \tilde{u}_i^r}{\partial t^2}(\mathbf{x}, t) - \frac{\partial}{\partial x_j}\left[a_{ijkl}(\mathbf{x}) \frac{\partial \tilde{u}_l^r}{\partial x_k}(\mathbf{x}, t)\right] = \frac{\partial A^r}{\partial x_i}(\mathbf{x}, t, s), \tag{13}$$

where A^s and A^r are given by

$$A^s(\mathbf{x}, t, s) = V_{\mathrm{P0}}^2(\mathbf{x}) \int \mathrm{d}h \hat{h} \frac{\partial^2 \mathcal{R}}{\partial z^2}(\mathbf{x}+\mathbf{h}, \mathbf{h}) W^r(\mathbf{x}+2\mathbf{h}, T-t, s), \tag{14}$$

and

$$A^r(\mathbf{x}, t, s) = V_{\mathrm{P0}}^2(\mathbf{x}) \int \mathrm{d}h \hat{h} \frac{\partial^2 \mathcal{R}}{\partial z^2}(\mathbf{x}-\mathbf{h}, \mathbf{h}) W^s(\mathbf{x}-2\mathbf{h}, t, s). \tag{15}$$

The adjoint sources $\partial A^s/\partial x_i$ and $\partial A^r/\partial x_i$ represent, respectively, the source- and receiver-side displacement residuals. These residuals are obtained by taking the kernel of the derivatives of the objective function with respect to the displacement wavefields u_i^s and u_i^r. Because the displacement wavefields u_i^s and u_i^r are originally shifted by, respectively, $+\mathbf{h}$ and $-\mathbf{h}$, a shift with opposite sign must be applied to the objective function prior to taking the respective derivatives. This explains why A^s have dependencies in $\mathbf{x}+\mathbf{h}$ and $\mathbf{x}+2\mathbf{h}$ and A^r have dependencies in $\mathbf{x}-\mathbf{h}$ and $\mathbf{x}-2\mathbf{h}$. For example, in the case of the source side, we have $\int \mathrm{d}x \int \mathrm{d}h \frac{\partial^2 \mathcal{R}}{\partial z^2}(\mathbf{x}, \mathbf{h}) W^s(\mathbf{x}-\mathbf{h}) W^r(x+h) = \int \mathrm{d}x \int \mathrm{d}h \frac{\partial^2 \mathcal{R}}{\partial z^2}(\mathbf{x}+\mathbf{h}, \mathbf{h}) W^s(\mathbf{x}) W^r(\mathbf{x}+2\mathbf{h})$. The same argument can be used for the receiver side.

Finally, note that equation 12 is to be solved in reverse time and equation 13 is to be solved forward in time.

If regularization is applied, the gradients of equation 7 with respect to the velocity parameters must be augmented with the gradients of equation 8, which are given by

$$\frac{\partial \mathcal{J}_R}{\partial V_{\mathrm{P0}}}(\mathbf{x}) = \beta_1(\mathbf{x})(V_{\mathrm{P0}}(\mathbf{x}) - V_{\mathrm{P0}}^0(\mathbf{x})) - \alpha_1\left(\frac{\partial^2 V_{\mathrm{P0}}}{\partial x_i^2}(\mathbf{x}) - \frac{\partial^2 V_{\mathrm{P0}}^0}{\partial x_i^2}(\mathbf{x})\right), \tag{16}$$

$$\frac{\partial \mathcal{J}_R}{\partial \varepsilon}(\mathbf{x}) = \beta_2(\mathbf{x})(\varepsilon(\mathbf{x}) - \varepsilon^0(\mathbf{x})) - \alpha_2\left(\frac{\partial^2 \varepsilon}{\partial x_i^2}(\mathbf{x}) - \frac{\partial^2 \varepsilon^0}{\partial x_i^2}(\mathbf{x})\right), \tag{17}$$

$$\frac{\partial \mathcal{J}_R}{\partial \delta}(\mathbf{x}) = \beta_3(\mathbf{x})(\delta(\mathbf{x}) - \delta^0(\mathbf{x})) - \alpha_3\left(\frac{\partial^2 \delta}{\partial x_i^2}(\mathbf{x}) - \frac{\partial^2 \delta^0}{\partial x_i^2}(\mathbf{x})\right). \tag{18}$$

Velocity preconditioning

To speed up convergence, and to restrict the space of possible solutions, we precondition the velocity parameters using bicubic B-splines (Dierckx, 1993)

$$V_{\mathrm{P0}}(\mathbf{x}) = \sum_m b_{m1} B_{m1}(\mathbf{x}) + V_{\mathrm{P0}}^0(\mathbf{x}), \tag{19}$$

$$\varepsilon(\mathbf{x}) = \sum_m b_{m2} B_{m2}(\mathbf{x}) + \varepsilon^0(\mathbf{x}), \tag{20}$$

$$\delta(\mathbf{x}) = \sum_m b_{m3} B_{m3}(\mathbf{x}) + \delta^0(\mathbf{x}), \tag{21}$$

where B_{mn} are cubic splines defined at predetermined points $\mathbf{m} = (m_x, m_z)$ in a spline grid and b_{mn} are coefficients to be determined by the velocity analysis. The spacing of the spline grid controls the sparseness of the solution and can be chosen differently for V_{P0}, ε, and δ, hence the $n = 1, 2, 3$ index. We only use B-splines to represent the velocity updates. This avoids having to fit the initial model to a B-spline basis, which would otherwise result in unnecessary smoothing.

The B-spline representation is attractive because it allows for local velocity variations, while, at the same time, it ensures continuous second-order spatial derivatives. These properties help to obtain a numerically stable and well-posed velocity analysis algorithm.

In practice, we compute the gradients in Cartesian coordinates and subsequently transform them to the spline basis

$$\frac{\partial \mathcal{J}_{m1}}{\partial b_{m1}} = \int d\mathbf{x} B_{m1}(\mathbf{x}) \frac{\partial \mathcal{J}}{\partial V_{\mathrm{P0}}}(\mathbf{x}), \tag{22}$$

$$\frac{\partial \mathcal{J}_{m2}}{\partial b_{m2}} = \int d\mathbf{x} B_{m2}(\mathbf{x}) \frac{\partial \mathcal{J}}{\partial \varepsilon}(\mathbf{x}), \tag{23}$$

$$\frac{\partial \mathcal{J}_{m3}}{\partial b_{m3}} = \int d\mathbf{x} B_{m3}(\mathbf{x}) \frac{\partial \mathcal{J}}{\partial \delta}(\mathbf{x}). \tag{24}$$

Diagonal scaling

One problem associated with the estimation of more than one parameter at the same time is that the magnitudes of the gradients with respect to the different parameters can be very different. The different sensitivities are normally compensated for in a full-Newton optimization method, through the scaling given by the inverse Hessian matrix (Nocedal and Wright, 2000). However, for quasi-Newton methods, this is not the case. The poor relative scaling causes the optimization to be dominated by the parameters with the largest gradient magnitudes. To mitigate this problem, we apply so-called *diagonal scaling* (Nocedal and Wright, 2000), where the optimization variables are related to the bicubic spline coefficients by a linear coordinate transformation:

$$b'_{\mathrm{mn}} = \frac{b_{\mathrm{mn}}}{k_n}. \tag{25}$$

By proper choice of the constants k_n for each spline coefficient array b_{mn}, we can rescale the gradients because

$$\frac{\partial \mathcal{J}_{\mathrm{mn}}}{\partial b'_{\mathrm{mn}}} = k_n \frac{\partial \mathcal{J}_{\mathrm{mn}}}{\partial b_{\mathrm{mn}}}. \tag{26}$$

This helps to equalize the contribution of each parameter to the descent direction and hence simultaneous estimation of all parameters. One drawback of this approach, compared to a full-Newton method, is that the chosen scaling must be fixed at the beginning of the optimization, and it can only be changed by restarting the optimization as a steepest descent.

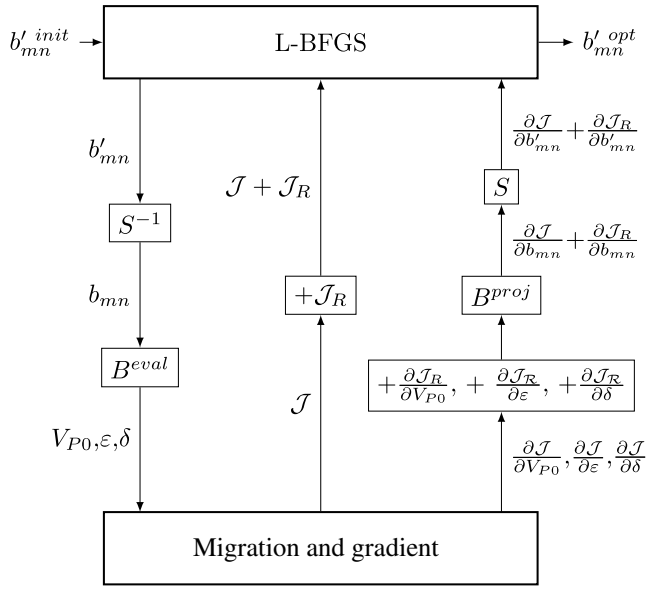

Figure 3. Organization of the 2D TTI WEMVA algorithm: B^{eval} and B^{proj} correspond to, respectively, B-spline evaluation and projection; S and S^{-1} refer to, respectively, the forward and inverse diagonal scaling; $b'^{\mathrm{init}}_{\mathrm{mn}}$ are the initial diagonally scaled B-spline coefficients; and $b'^{\mathrm{opt}}_{\mathrm{mn}}$ are the optimized diagonally scaled B-spline coefficients.

Numerical optimization

The organization of the numerical optimization scheme in our 2D TTI WEMVA implementation is shown in Figure 3. In a first stage, the L-BFGS algorithm is fed with an initial model (as shown in the left side of Figure 3), which corresponds to the initial diagonally scaled B-spline coefficient $b'^{\mathrm{init}}_{\mathrm{mn}}$. These are typically zero because we are not fitting the initial models to B-spline coefficients, and the initial updates are zero. At each iteration or line search step, the objective function and gradient need to be evaluated. The modeling variables V_{P0}, ε, and δ necessary for the migration and gradient computation are obtained from the optimization variables b'_{mn}, through a two-stage process, as shown in the left of Figure 3. First, the diagonal scaling is removed by solving equation 25 for the B-spline coefficients b_{mn}. Then, equations 19–21 are used to evaluate the B-splines. A step-by-step procedure for computing the objective function and gradient can be described as follows:

1) Construct $\mathcal{R}$ and evaluate the objective function, and in this procedure store the displacements u_i^s and u_i^r for each shot.
2) Perform separately the two adjoint modelings for each shot, according to equations 12 and 13 to compute, respectively, the adjoint states $\tilde{u}_i^s$ and $\tilde{u}_i^r$, and at each time step use equations 9–11 to build the gradients.
3) Stack each gradient over all shots to obtain the full gradient.

The computed objective function $\mathcal{J}$ must be augmented by the regularization term $\mathcal{J}_R$, as shown in the center of Figure 3. And the gradients of $\mathcal{J}$ with respect to the parameters V_{P0}, ε, and δ are augmented with the respective regularization gradients, as shown in the right side of Figure 3. Finally, these gradients are projected into a B-spline basis using equations 22–24 and diagonally scaled using equation 26.

The loop is repeated until some convergence criteria is met or a predetermined number of iterations has been run. At which point, the optimized parameters $b'^{\mathrm{opt}}_{\mathrm{mn}}$ are output, as shown in the right of Figure 3.

SYNTHETIC 2D EXAMPLE

The first example of TTI WEMVA is based on the synthetic velocity model shown in Figure 4. The model is a 2D synthetic cross section of a North Sea offshore reservoir. The anisotropic model simulates a 2D TTI medium. This model was used to generate synthetic seismic data using a finite-difference solution to the elastic wave equation (Lisitsa and Vishnevskiy, 2010). The geometry of the data consists of a line with minimum offset of 0.15 km and maximum offset of 5 km. Absorbing boundary conditions were used to ensure that the data are free from surface-related multiples. However, interbed multiples and converted waves are still present in the data.

In this example, as in the next example, we simultaneously estimate V_{P0}, ε, and δ. All parameters are optimized over bicubic B-spline grids with 0.8 km spacing in the lateral direction and 0.2 km in the vertical direction. We assume an initial θ model and keep it constant over the course of the minimization. In this case, we use the true θ model, shown in Figure 4d.

Regularization consisted in constraining the anisotropic parameters ε and δ to be zero at the water layer and positive in the sediments. The constant α_1, controlling the derivative regularization of the V_{P0} model, was set so that the value of the regularization was 1% of the total initial objective function value, while α_2 and α_3 were set to values 10,000 times larger than that of α_1. A taper is applied to mute the gradient in the water layer, simulating a situation in which the velocity of the water is known. The maximum frequency of the data was filtered down to 30 Hz, so that a coarse grid of 0.02 by 0.02 km could be used for modeling and migration.

The starting point for the velocity analysis is an isotropic 1D velocity model, shown in Figure 5a. The model is constructed from a single smoothed trace of the true velocity model. The result of optimization on the parameters after 27 iterations is shown in Figure 5b–5d. From this figure, we can see that the updated V_{P0} model is able to partially capture the main background features of the true V_{P0} model. On the other hand, the estimated anisotropic parameters show a strong imprint of the P-wave velocity and a general lack of structure. This reveals a strong dependency between the different parameters, which in this case appears to be a major contributor to the nonuniqueness of the result. Of all parameters estimated, δ seems to be the most poorly constrained. We note that there is a tendency for δ to be overestimated in the shallow parts of the model and underestimated in the deeper parts of the model. The reason for this behavior is twofold. The first reason is the poor separation between the effects of V_{P0} and δ on the kinematics of the

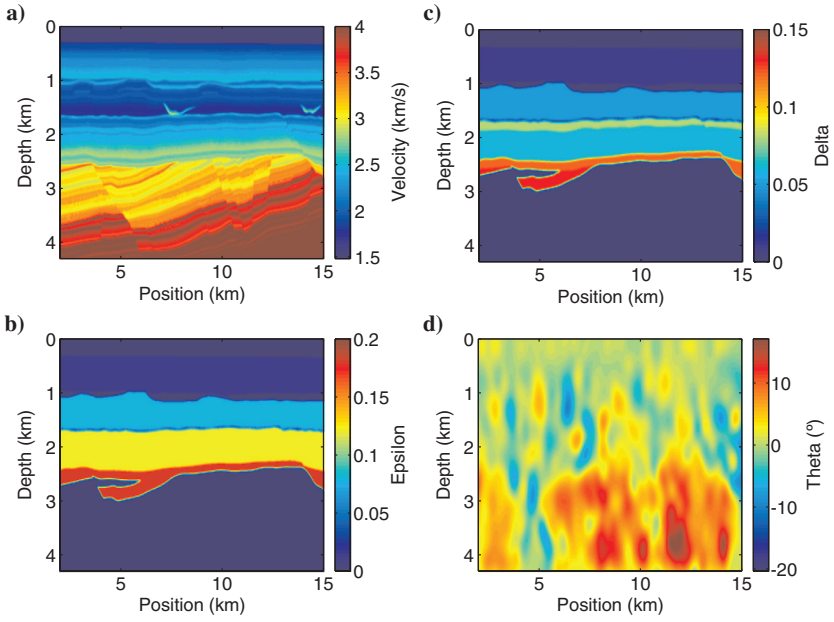

Figure 4. 2D TTI synthetic model of a North Sea reservoir: (a) V_{P0}, (b) ε, (c) δ, and (d) θ.

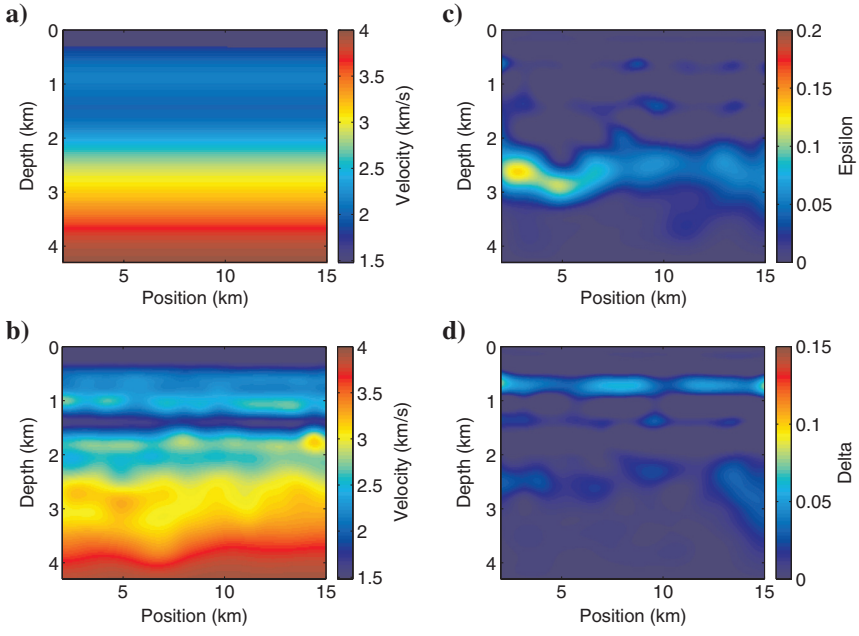

Figure 5. Initial and updated models used for the 2D synthetic data example: (a) initial V_{P0} model, (b) updated V_{P0} model after 27 iterations, (c) updated ε model after 27 iterations, (c) updated δ model after 27 iterations. The θ model used in this example is shown in Figure 4d.

image. The second reason is related to the poor scaling of the optimization, which causes the objective function to be dominated by the sensitivity to δ in the shallow parts of the image and by V_{P0} in the deeper parts. This means that in the shallow parts of the model, δ is mainly compensating for the kinematic errors introduced by an underestimated V_{P0} and vice versa for the deeper parts of the model.

Figure 6 shows a comparison of the RTM images produced with the initial 1D, WEMVA, and true model parameters. The initial image has large mispositionings and is poorly focused due to the inaccurate initial background velocities. These issues are partially fixed in the optimized migrated image, which is better focused. But there are some mispositionings (up to more than 0.05 km) in the optimized image, in particular below 3 km depth. To help better compare the spatial positioning of the reflectors in the images, we have drawn three arrows in fixed positions in the images.

A selection of subsurface offset CIGs constructed using the initial model, the optimized model, and the true model are shown in Figure 7a–7c. The figure shows that the energy in the updated CIGs is now better focused at the zero lag, when compared to the initial CIGs. A comparison of the true model CIGs with the updated CIGs reveals that the updated CIGs are slightly mispositioned (>0.05 km) in depth. The fact that the updated and true models can produce CIGs that are focused at zero lag, yet with different spatial positioning, reveals one source of nonuniqueness. This seems to repeat the generally known fact that, in presence of anisotropy, focusing of prestack depth-migrated images does not guarantee a unique positioning of the subsurface reflectors (Isaac and Lawton, 1999).

Now, we repeat the velocity analysis using the same parameters as before, but this time we use a fixed V_{P0} model obtained by smoothing the true model. The V_{P0} model is shown in Figure 8a, and the results of velocity analysis for ε, and δ are shown, respectively, in Figure 8a and 8b. The results show that by adding additional information, in this case V_{P0}, we can obtain better estimates of the anisotropic parameters.

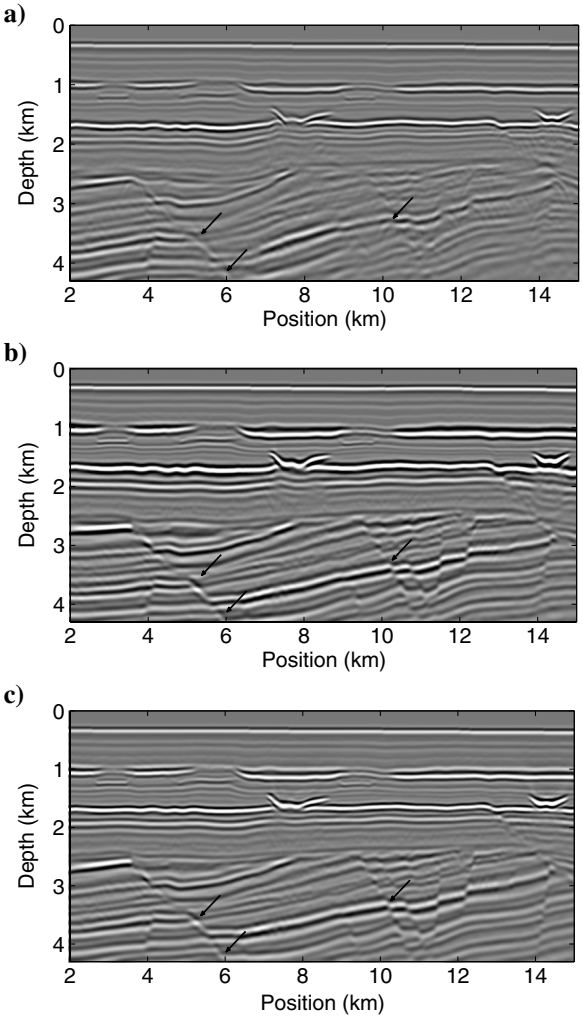

Figure 6. Images constructed using the (a) initial model, (b) optimized model, and (c) true model. Arrows are drawn to help visualize the changes in the spatial positioning of the reflectors.

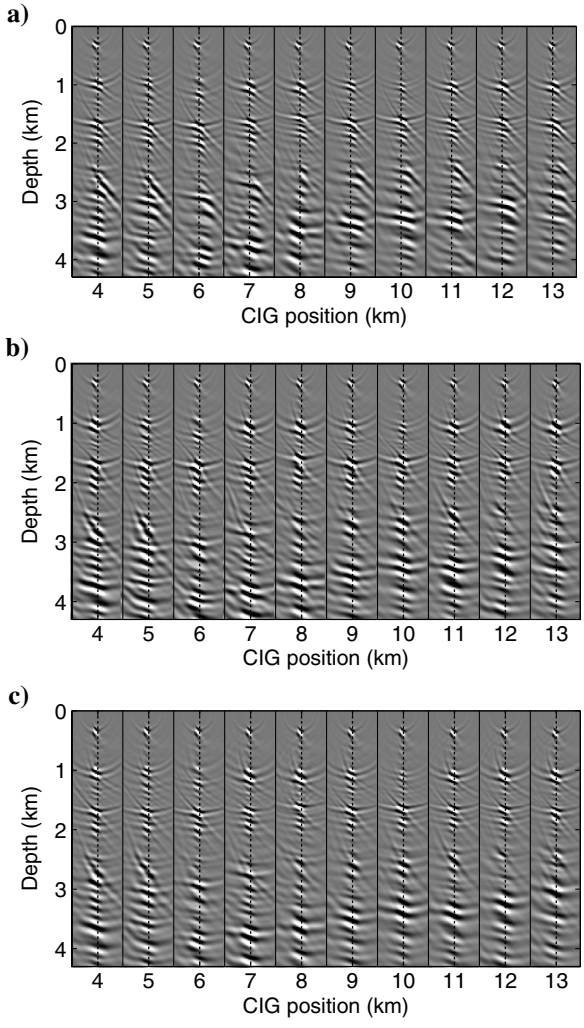

Figure 7. CIGs constructed using the (a) initial model, (b) optimized model, and (c) true model. The black dotted lines mark the position of the zero subsurface offset. The offsets range between −0.5 and 0.5 km.

FIELD 2D EXAMPLE

In the next example, we apply the method on a real data set taken off the North Sea, offshore Norway. The data are originally a 3D marine data set, from which we extracted a 2D line. The geometry of the data consists of a line with minimum offset of 0.15 km and maximum offset of 5 km. The data processing included multiple attenuation and muting of direct wave, wide-angle reflections and refractions. The maximum frequency of the data was filtered down to 30 Hz, so that a coarse grid of 0.02 by 0.02 km could be used for modeling and migration.

The initial model for the optimization is shown in Figure 9a. It consists of an isotropic 1D model created by smoothing a well log of the vertical slowness. Here, we numerically estimate θ from the initial image, and we keep it constant during optimization. We try to approximate an STI model of the subsurface, where the symmetry axis is perpendicular to the dip of the beddings. To estimate the tilt angle, we first estimate the smallest positive angle between the spatial gradient of the image and an unitary positive vertical vector. This can be done using the equation

$$\phi(\mathbf{x}) = \cos^{-1}\left[\frac{\left|\frac{\partial \mathcal{R}}{\partial z}(\mathbf{x},0)\right|}{\left(\frac{\partial \mathcal{R}^2}{\partial x}(\mathbf{x},0)+\frac{\partial \mathcal{R}^2}{\partial z}(\mathbf{x},0)\right)^{1/2}}\right]. \tag{27}$$

To find the sign of the angle, we use the following convention:

$$\theta(\mathbf{x}) = \begin{cases} -\phi(\mathbf{x}) & \text{if } \frac{\partial \mathcal{R}}{\partial x}(\mathbf{x},0) \le 0, \\ \phi(\mathbf{x}) & \text{if } \frac{\partial \mathcal{R}}{\partial x}(\mathbf{x},0) > 0. \end{cases} \tag{28}$$

To avoid excessive roughness in the estimate of θ, we low-pass filter the image gradient before computing ϕ and subsequently low-pass filter θ. The estimated tilt angles are shown in Figure 9b. We can see that the estimated angles partially capture the general background trend of the structure in the image. However, some artifacts are apparent, which make the estimate look nongeologic. These appear in particular at areas where there is a lack of reflectivity or due to the smearing caused by the smoothing. Because these artifacts are of small angle magnitude ($|\theta| < 5°$), we deem that they should have a minor or negligible effect on wave propagation.

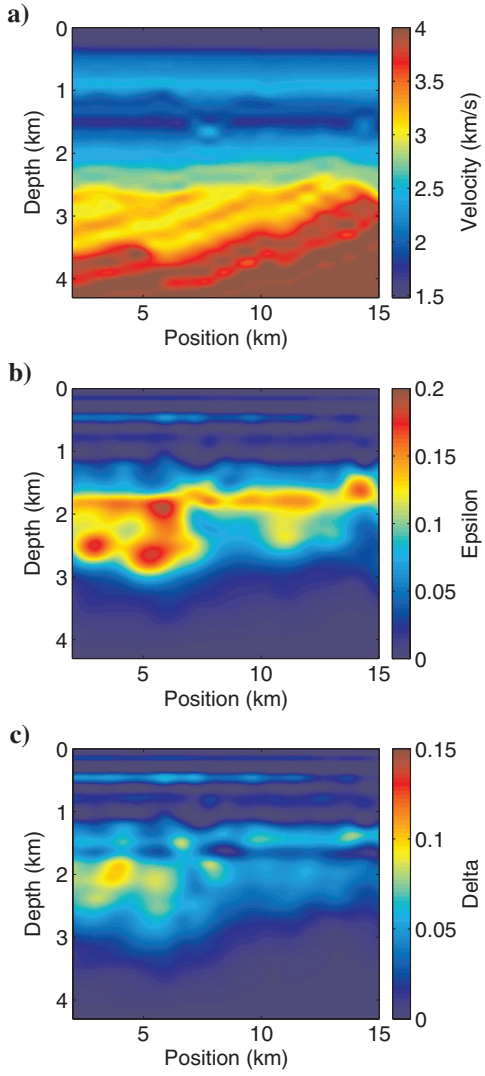

Figure 8. Results of 2D TTI WEMVA with a fixed V_{P0} model obtained by smoothing the true model: (a) smoothed V_{P0}, (b) ε, and (c) δ.

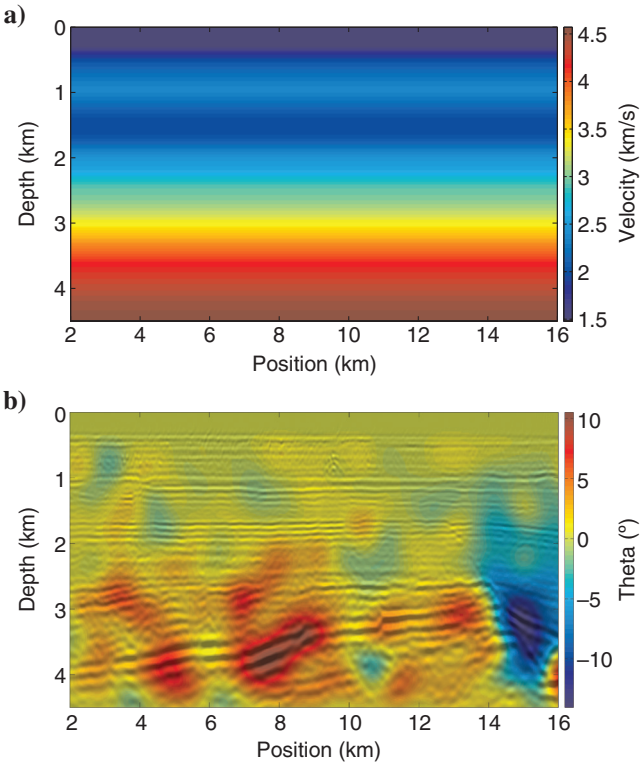

Figure 9. Initial models used for the 2D field data example: (a) initial V_{P0} model and (b) θ model overlaid by the initial image. The tilt angles in (b) are estimated from the reflector dips of the initial image.

As in the previous example, the parameters are optimized over a bicubic B-spline grid with 0.8 km spacing in the lateral direction and 0.2 km in the vertical direction. Regularization followed the same guidelines as in the previous example.

The resultant estimated parameters after 47 iterations of 2D TTI WEMVA are shown in Figure 10. The estimated V_{P0} model reveal a nearly 1D plane-layered overburden down to 2 km. Below 2 km, as expected due to the normal faulted structures, the optimized V_{P0}, ε, and δ models show more lateral variation.

The migrated images computed with the initial and updated models are shown in Figure 11, while Figure 12 shows a comparison on selected CIGs. We can see that optimization locally improves the focusing of the RTM image. These improvements are marked by arrows in Figure 11. Also, the energy in the updated CIGs is more focused at zero offset, as can be seen in Figure 12. However, as in the synthetic data example, there is some uncertainty about the positioning of the reflectors in the final image.

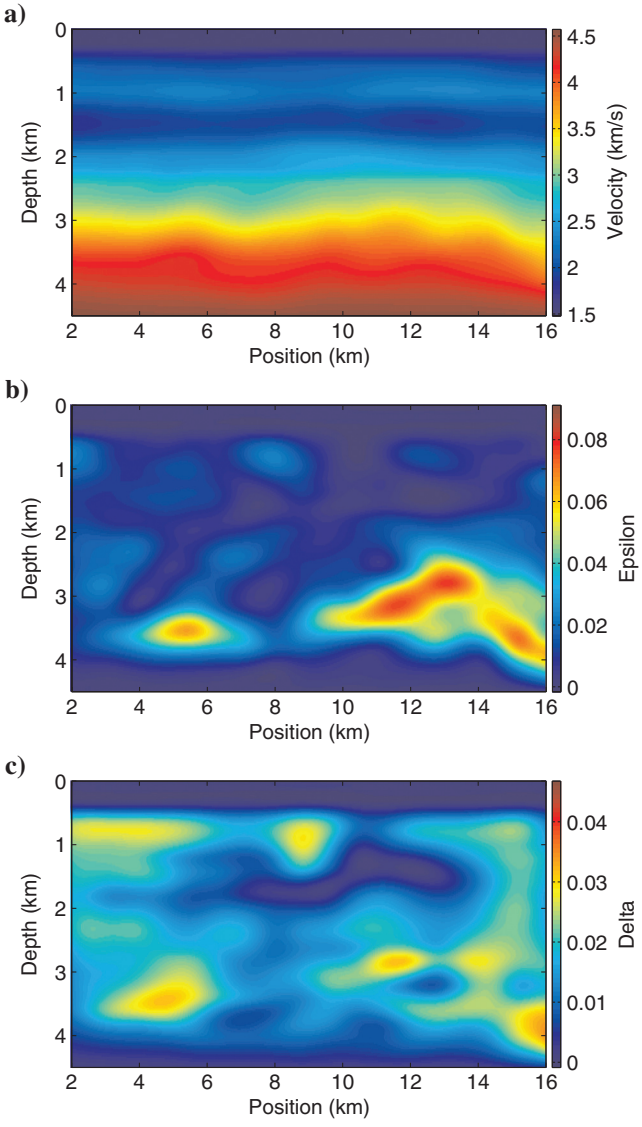

Figure 10. Optimized models after 47 iterations of 2D TTI WEMVA: (a) V_{P0}, (b) ε, and (c) δ.

DISCUSSION

Prestack depth migration of P-wave surface seismic data in 2D TTI media using a density-normalized elastic wave equation requires an estimate of the P-wave velocity along the symmetry axis V_{P0}, the S-wave velocity along the symmetry axis V_{S0}, Thomsen's parameters ε and δ, and the tilt θ of the symmetry axis with respect to the vertical. To reduce the number of parameters, we need to estimate; we assume that the kinematics of P-waves are independent of the choice of V_{S0} and that θ is structurally conforming, which reduces the velocity estimation problem to three parameters: V_{P0}, ε, and δ. Here, we use 2D TTI WEMVA to simultaneously estimate these three parameters from surface seismic data. The results show that the method is able to improve the focusing of the depth-migrated image. However, the synthetic data example clearly shows that a unique set of anisotropic parameters is not constrained by the method. There are several contributors to this issue. One of them is that, over a transverse isotropic medium, focusing of the prestack depth images does not constrain a unique spatial positioning of the reflectors in a depth-migrated image (Isaac and Lawton, 1999; Bakulin et al., 2010). Another reason is the existence of a strong interdependency between the parameters V_{P0}, ε, and δ (Alkhalifah and Tsvankin, 1995; Jones et al., 2003). In fact, Alkhalifah et al. (2001) show that, under certain conditions, at most two parameters can be recovered from P-wave seismic reflection data. Therefore, if the goal is to obtain a geologic model of the subsurface from surface seismic data, a combination of different methods and additional information such as well data, must be used to constrain the updated models to a narrower space of possible solutions (Alkhalifah and Tsvankin, 1995; Alkhalifah et al., 2001; Yan et al., 2004; Bakulin et al., 2010).

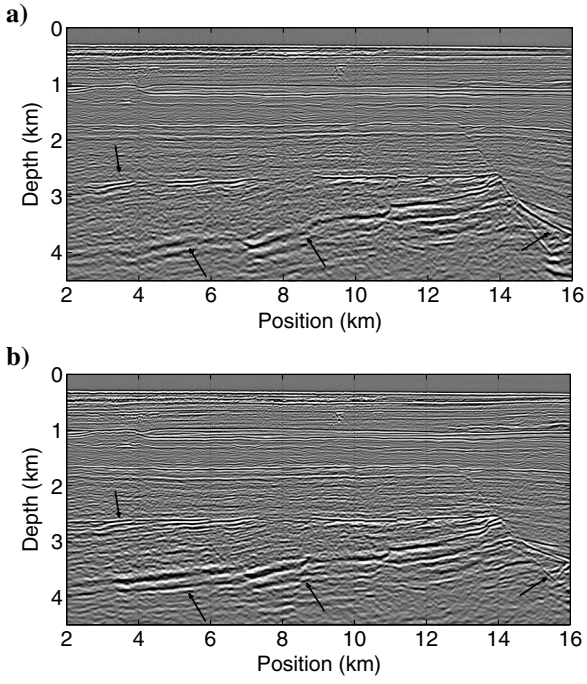

Figure 11. Images constructed using the (a) initial model and (b) optimized model. Arrows are drawn at fixed positions in the images to mark changes in the spatial positioning and focusing.

The choice of parameterization can be important for the results of velocity analysis because it changes the scaling of the optimization problem (Nocedal and Wright, 2000). Here, we use WEMVA to estimate V_{P0}, ε, and δ. To estimate more than one parameter simultaneously using an L-BFGS optimization algorithm, we apply diagonal scaling. The idea is to approximately equalize the sensitivity of the objective function to the different parameters. This procedure can be avoided if the parameters have the same units and vary approximately over the same ranges. It is thus possible that the parameterization given by V_{P0}, $V_h = V_{P0}\sqrt{1+2\varepsilon}$ and $V_n = V_{P0}\sqrt{1+2\delta}$, which was suggested by Alkhalifah and Tsvankin (1995), provides a better scaling for the problem. Although we develop the method using V_{P0}, ε, and δ, it is trivial to change the parameterization to V_{P0}, V_h, and V_n by substitution of the quantities on the density-normalized stiffness tensor.

The main difference between isotropic and anisotropic WEMVA is in the reconstruction of the source and receiver wavefields. Here, we use a density-normalized elastic wave equation to model the source and receiver displacement wavefields over 2D TTI media. We then extract quasi-P wavefields by taking the divergence of the displacements scaled by the density-normalized bulk modulus. These wavefields are input to an extended crosscorrelation imaging condition to create PP CIGs of the subsurface. However, the elastic wave equation also naturally models S-waves. This opens up the possibility to apply WEMVA to multicomponent data, which would allow us to create PP, PS, and SS images of the subsurface (Hokstad et al., 1998). The 2D TTI WEMVA algorithm we present can be extended to use all these images, allowing to simultaneously estimate V_{P0} and V_{S0}, as well as to improve the accuracy in the estimation of the anisotropic parameters ε and δ (Tsvankin and Thomsen, 1995; Grechka et al., 2002).

The WEMVA method we present is restricted to single-scattering reflection data. This is a limitation shared by all methods based on the image domain, and it prevents us from using valuable information contained in refractions and multiple reflections. Data domain methods based on full-waveform inversion (FWI) can, in theory, use all information contained in the data (Tarantola, 1984; Virieux and Operto, 2009; Plessix and Rynja, 2010). On the other hand, for FWI to converge to the solution, it requires either the presence of a useful signal at very low frequencies (<4 Hz) or an initial model that is kinematically close to the solution (Virieux and Operto, 2009). Combining FWI and WEMVA is nontrivial, because in practice the two methods have very different restrictions when it comes to the input data and initial models. However, one possible strategy is to use WEMVA to create initial models that are kinematically close enough to the solution and then further refine them using FWI.

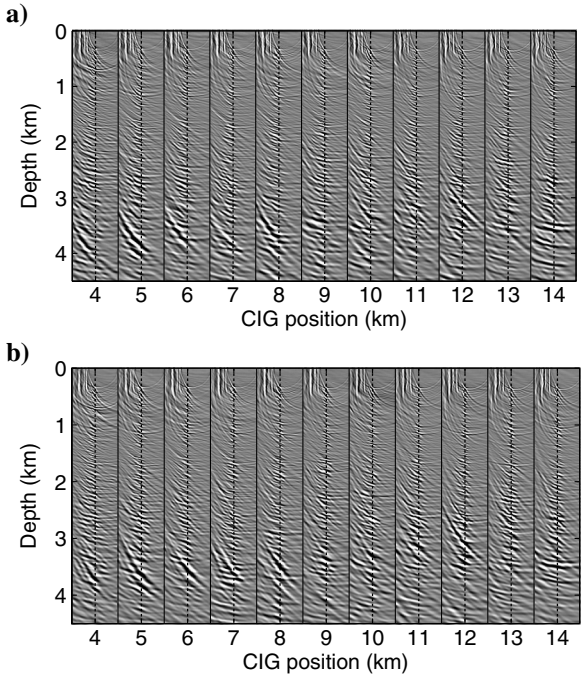

Figure 12. CIGs constructed using the (a) initial model and (b) optimized model. The black dotted lines mark the position of the zero subsurface offset. The offsets range between −0.5 and 0.5 km.

CONCLUSION

We present a WEMVA method based on anisotropic 2D TTI RTM. To reconstruct the wavefields needed for the imaging condition, we use a density-normalized elastic wave equation. The WEMVA method can be used to simultaneously estimate the parameters V_{P0}, ε, and δ from surface seismic reflection data, using an objective function consisting of a combination of depth-oriented differential semblance and stack-power maximization. In this procedure, the tilt of the symmetry axis with respect to the vertical is assumed to conform to the initial reflectivity geometry and not updated, and the S-wave velocities are heuristically chosen and also kept constant.

The method is tested on synthetic and field data. The tests show that the method is well posed and converges to a model that produces well-focused images. On the other hand, the results also show that 2D TTI WEMVA of surface seismic data alone is not sufficient to constrain a unique anisotropic model of the subsurface. This means that to obtain a correct positioning of the reflectors in the subsurface, the method must be complemented with additional information.

Because we employ the elastic wave equation for wavefield reconstruction, the method can be easily generalized to any anisotropic medium. Furthermore, the quantities in the stiffness tensor can easily be transformed to honor other parameterizations than the one used in this paper.

ACKNOWLEDGMENTS

W. Weibull would like to thank M. Houbiers, J. Mispel, and A. Kritski for interesting discussions about the subject. This author would also like to thank Statoil Petroleum AS for financial support of this work. We acknowledge the partners in the Snorre license, Statoil Petroleum AS, Petoro AS, ExxonMobil Exploration and Production Norway AS, Idemitsu Petroleum Norge AS, RWE Dea Norge AS, and Core Energy AS for permission to publish the results. We would also like to acknowledge the associate editor I. Vasconcelos and the reviewers, J. Shragge, T. Alkhalifah, and F. Perrone, for their constructive comments and suggestions, which greatly helped to improve this manuscript.

APPENDIX A

2D ELASTIC TTI MEDIUM

We consider the following system describing 2D elastic wave propagation:

$$\partial_t u_x = \partial_x \tau_{xx} + \partial_z \tau_{xz}, \tag{A-1}$$

$$\partial_t u_z = \partial_x \tau_{xz} + \partial_z \tau_{zz}, \tag{A-2}$$

$$\begin{pmatrix} \partial_t \tau_{xx} \\ \partial_t \tau_{zz} \\ \partial_t \tau_{xz} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{13} & a_{15} \\ a_{13} & a_{33} & a_{35} \\ a_{15} & a_{35} & a_{55} \end{pmatrix} \begin{pmatrix} \partial_x u_x \\ \partial_z u_z \\ \partial_x u_z + \partial_z u_x \end{pmatrix} + \begin{pmatrix} S_{xx} \\ S_{zz} \\ S_{xz} \end{pmatrix}, \tag{A-3}$$

where $\partial_t = \frac{\partial}{\partial t}, \partial_x = \frac{\partial}{\partial x}$, and $\partial_z \frac{\partial}{\partial z}$; u_i is the particle displacement vector; τ_{ij} is a stress tensor; S_{ij} is an external source tensor; and $A = (a_{ij})$ is the matrix of density-normalized elastic coefficients. This matrix is obtained from the fourth-order tensor a_{ijkl} using the Voigt notation: $xx \to 1$, $zz \to 3$, and $xz \to 5$ (Winterstein, 1990).

In a 2D TTI medium, A is given by the following elastic coefficients:

$$a_{11} = a'_{11}\cos^4\theta + a'_{33}\sin^4\theta + 2(a'_{13} + 2a'_{55})\sin^2\theta\cos^2\theta, \tag{A-4}$$

$$a_{13} = (a'_{11} + a'_{33} - 4a'_{55})\sin^2\theta\cos^2\theta + a'_{13}(\sin^4\theta + \cos^4\theta), \tag{A-5}$$

$$\begin{aligned} a_{15} &= (a'_{13} - a'_{11} + 2a'_{55})\sin\theta\cos^3\theta, \\ &\quad + (a'_{33} - a'_{13} - 2a'_{55})\sin^3\theta\cos\theta, \end{aligned} \tag{A-6}$$

$$\begin{aligned} a_{33} &= a'_{11}\sin^4\theta + a'_{33}\cos^4\theta + 2(a'_{13} + 2a'_{55})\sin^2\theta\cos^2\theta, \\ a_{35} &= (a'_{13} + 2a'_{55} - a'_{11})\sin^3\theta\cos\theta, \end{aligned} \tag{A-7}$$

$$\begin{aligned} &+ (a'_{33} - a'_{13} - 2a'_{55})\sin\theta\cos^3\theta, \\ a_{55} &= (a'_{11} + a'_{33} - 2a'_{13} - 2a'_{55})\sin^2\theta\cos^2\theta \\ &+ a'_{55}(\sin^4\theta + \cos^4\theta), \end{aligned} \tag{A-8}$$

where θ is the tilt angle of the symmetry axis with respect to the vertical and $A' = (a'_{ij})$ are quantities given by

$$a'_{11} = V_{P0}^2(1 + 2\varepsilon), \tag{A-9}$$

$$a'_{13} = [2\delta V_{P0}^2(V_{P0}^2 - V_{S0}^2) + (V_{P0}^2 - V_{S0}^2)^2]^{1/2} - V_{S0}^2, \tag{A-10}$$

$$a'_{33} = V_{P0}^2, \tag{A-11}$$

$$a'_{55} = V_{S0}^2, \tag{A-12}$$

where V_{P0} is the P-wave velocity along the symmetry axis, V_{S0} is the S-wave velocity along the symmetry axis, and ε and δ are Thomsen's anisotropic parameters.

REFERENCES

Alkhalifah, T., 1998, Acoustic approximations for processing in transversely isotropic media: Geophysics, **63**, 623–631, doi: 10.1190/1.1444361.

Alkhalifah, T., S. Fomel, and B. Biondi, 2001, The space-time domain: Theory and modelling for anisotropic media: Geophysical Journal International, **144**, 105–113, doi: 10.1046/j.1365-246x.2001.00300.x.

Alkhalifah, T., and K. Larner, 1994, Migration error in transversely isotropic media: Geophysics, **59**, 1405–1418, doi: 10.1190/1.1443698.

Alkhalifah, T., and I. Tsvankin, 1995, Velocity analysis for transversely isotropic media: Geophysics, **60**, 1550–1566, doi: 10.1190/1.1443888.

Audebert, F., V. Dirks, and A. Pettenati, 2006, TTI anisotropic depth migration: What tilt estimate should we use?: 76th Annual International Meeting, SEG, Expanded Abstracts, 2382–2386.

Bakulin, A., M. Woodward, D. Nichols, K. Osypov, and O. Zdraveva, 2010, Building tilted transversely isotropic depth models using localized anisotropic tomography with well information: Geophysics, **75**, no. 4, D27–D36, doi: 10.1190/1.3453416.

Banik, N., 1984, Velocity anisotropy of shales and depth estimation in the North Sea basin: Geophysics, **49**, 1411–1419, doi: 10.1190/1.1441770.

Byrd, R. H., P. Lu, J. Nocedal, and C. Zhu, 1995, A limited memory algorithm for bound constrained optimization: SIAM Journal on Scientific Computing, **16**, 1190–1208, doi: 10.1137/0916069.

Chavent, G., and C. A. Jacewitz, 1995, Determination of background velocities by multiple migration fitting: Geophysics, **60**, 476–490, doi: 10.1190/1.1443785.

Chavent, G., and P. Lemonnier, 1974, Identification de la non-linéarité d'une équation parabolique quasilinéaire: Applied Mathematics and Optimization, **1**, 121–162, doi: 10.1007/BF01449027.

Dellinger, J., and J. Etgen, 1990, Waveafield separation in two-dimensional anisotropic media: Geophysics, **55**, 914–919, doi: 10.1190/1.1442906.

Dierckx, P., 1993, Curve and surface fitting with splines: Oxford University Press.

Douma, H., D. Yingst, I. Vasconcelos, and J. Tromp, 2010, On the connection between artifact filtering in reverse-time migration and adjoint tomography: Geophysics, **75**, no. 6, S219–S223, doi: 10.1190/1.3505124.

Grechka, V., A. Pech, and I. Tsvankin, 2002, P-wave stacking-velocity tomography for VTI media: Geophysical Prospecting, **50**, 151–168, doi: 10.1046/j.1365-2478.2002.00307.x.

Guitton, A., B. Kaelin, and B. Biondi, 2007, Least-squares attenuation of reverse-time-migration artifacts: Geophysics, **72**, no. 1, S19–S23, doi: 10.1190/1.2399367.

Hokstad, K., R. Mittet, and M. A. Landr, 1998, Elastic reverse time migration of marine walkaway vertical seismic profiling data: Geophysics, **63**, 1685–1695, doi: 10.1190/1.1444464.

Ikelle, L. T., and L. Amundsen, 2005, Introduction to petroleum seismology: SEG.

Isaac, J., and D. Lawton, 1999, Image mispositioning due to dipping TI media: A physical seismic modeling study: Geophysics, **64**, 1230–1238, doi: 10.1190/1.1444629.

Jones, I. F., M. L. Bridson, and N. Bernitsas, 2003, Anisotropic ambiguities in TI media: First Break, **21**, 15–21, doi: 10.3997/1365-2397.2003006.

Levin, F., 1979, Seismic velocities in transversely isotropic media: Geophysics, **44**, 918–936, doi: 10.1190/1.1440985.

Li, Y., and B. Biondi, 2011, Migration velocity analysis for anisotropic models: 81st Annual International Meeting, SEG, Expanded Abstracts, 201–206.

Li, Y. E., P. Shen, and C. Perkins, 2012, VTI migration velocity analysis using RTM: 82nd Annual International Meeting, SEG, Expanded Abstracts, doi: 10.1190/segam2012-0128.1.

Lions, J. L., and E. Magenes, 1972, Nonhomogeneous boundary value problems and applications: Springer Verlag.

Lisitsa, V., and D. Vishnevskiy, 2010, Lebedev scheme for the numerical simulation of wave propagation in 3D anisotropic elasticity: Geophysical Prospecting, **58**, 619–635, doi: 10.1111/j.1365-2478.2009.00862.x.

Mulder, W. A., 2008, Automatic velocity analysis with the two-way wave equation: 70th Annual International Conference and Exhibition, EAGE, Extended Abstracts, P165.

Mulder, W. A., and T. van Leeuwen, 2008, Automatic migration velocity analysis and multiples: 78th Annual International Meeting, SEG, Expanded Abstracts, 3128–3132.

Nocedal, J., and S. J. Wright, 2000, Numerical optimization: Springer.

Plessix, R., and H. Rynja, 2010, VTI full waveform inversion: A parametrization study with a narrow azimuth streamer data example: 80th Annual International Meeting, SEG, Expanded Abstracts, 962–966.
Plessix, R. E., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications: Geophysical Journal International, **167**, 495–503, doi: 10.1111/j.1365-246X.2006.02978.x.
Rickett, J. E., and P. C. Sava, 2002, Offset and angle-domain common image-point gathers for shot-profile migration: Geophysics, **67**, 883–889, doi: 10.1190/1.1484531.
Sava, P., and B. Biondi, 2004, Wave-equation migration velocity analysis. I. Theory: Geophysical Prospecting, **52**, 593–606, doi: 10.1111/j.1365-2478.2004.00447.x.
Sava, P., and L. Vlad, 2008, Numeric implementation of wave-equation migration velocity analysis operators: Geophysics, **73**, no. 5, VE145–VE159, doi: 10.1190/1.2953337.
Sayers, C. M., 1994, The elastic anisotrophy of shales: Journal of Geophysical Research: Solid Earth, **99**, 767–774, doi: 10.1029/93JB02579.
Shen, P., and W. W. Symes, 2008, Automatic velocity analysis via shot profile migration: Geophysics, **73**, no. 5, VE49–VE59, doi: 10.1190/1.2972021.
Shen, P., W. W. Symes, and C. C. Stolk, 2003, Differential semblance velocity analysis by wave-equation migration: 73rd Annual International Meeting, SEG, Expanded Abstracts, 2132–2135.
Symes, W. W., and M. Kern, 1994, Inversion of reflection seismograms by differential semblance analysis: Algorithm structure and synthetic examples: Geophysical Prospecting, **42**, 565–614, doi: 10.1111/j.1365-2478.1994.tb00231.x.
Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266, doi: 10.1190/1.1441754.
Thomsen, L., 1986, Weak elastic anisotropy: Geophysics, **51**, 1954–1966, doi: 10.1190/1.1442051.
Tikhonov, A. N., and V. Y. Arsenin, 1977, Solution of ill-posed problems: W. H. Winston and Sons.
Toldi, J. L., 1989, Velocity analysis without picking: Geophysics, **54**, 191–199, doi: 10.1190/1.1442643.
Tsvankin, I., and L. Thomsen, 1994, Nonhyperbolic reflection moveout in anisotropic media: Geophysics, **59**, 1290–1304, doi: 10.1190/1.1443686.
Tsvankin, I., and L. Thomsen, 1995, Inversion of reflection traveltimes for transverse isotropy: Geophysics, **60**, 1095–1107, doi: 10.1190/1.1443838.
van Leeuwen, T., and W. A. Mulder, 2008, Velocity analysis with multiples — NMO modeling for layered velocity structures: 78th Annual International Meeting, SEG, Expanded Abstracts, 1925–1929.
Vestrum, R., D. Lawton, and R. Schmid, 1999, Imaging structures below dipping TI media: Geophysics, **64**, 1239–1246, doi: 10.1190/1.1444630.
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.
Weibull, W., B. Arntsen, M. Houbiers, and J. Mispel, 2012, Automatic anisotropic migration velocity analysis for reverse-time migration: 82nd Annual International Meeting, SEG, Expanded Abstracts, doi: 0.1190/segam2012-0650.1.
Weibull, W. W., and B. Arntsen, 2011, Reverse time migration velocity analysis: 73rd Annual International Conference and Exhibition, EAGE, Extended Abstracts, B012.
Whitmore, N. D., and S. Crawley, 2012, Applications of RTM inverse scattering imaging conditions: 82nd Annual International Meeting, SEG, Expanded Abstracts, 10.1190/segam2012-0779.1.
Winterstein, D. F., 1990, Velocity anisotropy terminology for geophysicists: Geophysics, **55**, 1070–1088, doi: 10.1190/1.1442919.
Yan, J., and P. Sava, 2011, Improving the efficiency of elastic wave-mode separation for heterogeneous tilted transverse isotropic media: Geophysics, **76**, no. 4, T65–T78, doi: 10.1190/1.3581360.
Yan, L., L. Lines, and D. Lawton, 2004, Influence of seismic anisotropy on prestack depth migration: The Leading Edge, **23**, 30–36, doi: 10.1190/1.1645453.
Zhou, H., Q. Liao, and F. Ortigosa, 2009, Migration velocity inversion with semblance analysis (mVisa): 71st Annual International Conference and Exhibition, EAGE, Extended Abstracts, U022.