

Image registration guided wavefield tomography for shear-wave velocity model building

Di Yang¹, Xuefeng Shang², Alison Malcolm¹, Michael Fehler¹, and Hyongsu Baek³

ABSTRACT

Multicomponent acquisitions offer the opportunity to form elastic migration images and to estimate elastic parameters of the subsurface. Compared with better constrained P-wave velocity inversions, it is more difficult to estimate the S-wave velocity due to strong nonlinearities introduced by converted S-waves. We have developed an iterative S-wave velocity inversion method guided by image registration. Given an accurate P-wave velocity and a simple initial S-wave model, we form P-P and P-S images using elastic reverse time migration. We use image registration to find the shifts between the P-S and P-P images. The S-wave velocity model could be updated iteratively by minimizing the differences between the original and the fractionally warped P-S images. A simple layered model and a modified Marmousi model are used to demonstrate the viability of the new method. In both examples, high-quality P-S images, as well as smooth S-wave velocity models, are inverted efficiently with a homogeneous S-wave initial model.

INTRODUCTION

Multicomponent seismographs have been deployed in global seismology for decades to estimate the earth's elastic properties, such as P- and S-wave velocities. In hydrocarbon exploration, the S-wave velocity, coupled with the P-wave velocity, provides crucial information for lithology estimation and reservoir characterization (Stewart et al., 2002).

Multicomponent imaging methods have been proposed in the time and depth domains. Herrenschmidt et al. (2001) compare several converted-wave imaging approaches using real data, and show

that prestack time migration provides interpretable results when lateral velocity variations are not significant. Kuo and Dai (1984) propose Kirchhoff elastic-wave depth migration based on Kirchhoff-Helmholtz-type integrals. Hokstad (2000) presents multicomponent Kirchhoff migration using the survey-sinking concept. These ray-based methods are likely to fail when ray theory breaks down in complex media as does acoustic Kirchhoff migration (Gray et al., 2001). One-way migration methods can also be extended to elastic applications. Wapenaar and Haime (1990) propose separating wave modes on the surface before one-way migration in isotropic media. Yan and Sava (2008) advocate an alternative procedure to separate the wave modes during the vector wavefield propagation and reconstruction in elastic reverse time migration (ERTM). Shang et al. (2012) and Sava (2011) extend the ERTM concept to passive source applications, e.g., in teleseismic or microseismic studies.

Accurate velocity models are necessary for elastic depth migration (Alkhalifah, 2003). P-wave model building is relatively mature and robust. On the other hand, S-wave velocity or V_P/V_S ratio estimation is less well understood. A converted-wave migration velocity model is often obtained in the time domain by tuning the V_P/V_S ratio (Fomel et al., 2005; Hale, 2013). For example, S-wave velocities can be estimated by registering corresponding P-P and P-S reflections in time-migrated sections. Assuming that the P-wave velocity is correct, the time shifts between P-P and P-S events can be transformed into V_P/V_S ratio corrections. However, this method suffers from the limitations of time migration in handling lateral inhomogeneity, and it does not provide an accurate interval S-wave velocity model with which to depth migrate the data. Du et al. (2012a) propose a joint migration velocity analysis in the angle domain for P-P and P-S depth images. However, they use a Kirchhoff-based migration, which is likely to break down in complex structures. Yan and Sava (2010) present a wave-equation migration velocity analysis (WEMVA) method that finds the S-wave velocities and P-S depth-migration images simultaneously, but the computation

Manuscript received by the Editor 6 August 2014; revised manuscript received 24 November 2014; published online 27 April 2015.

¹Earth Resource Lab, Cambridge, Massachusetts, USA. E-mail: diyang.mit@gmail.com; amalcolm@mun.ca; fehler@mit.edu.

²Formerly Earth Resource Lab, Cambridge, Massachusetts; presently Shell International Exploration and Production, Houston, Texas, USA. E-mail: xfshang@mit.edu.

³Aramco Research Center, Houston, Texas, USA. E-mail: hbaek@math.mit.edu.

© 2015 Society of Exploration Geophysicists. All rights reserved.

of elastic angle gathers is expensive. Moreover, the method does not impose any constraint on the proximity of P-S to P-P images.

Yang et al. (2014a) propose an image-domain wavefield tomography (IDWT) method for time-lapse velocity inversion assuming that the reflector location does not change between surveys. A similar principle can be used for S-wave velocity inversion. We assume that the reflectors in P-S depth-migrated images should be at the same depths as those in P-P images. If the S-wave velocities are incorrect, we can measure and minimize the depth shifts between P-S and P-P images to update the S-wave velocity model. The calculation of depth shifts can be achieved by image registration. Fomel and Backus (2003) introduce a least-squares optimization method for multicomponent data registration, but this method requires a good initial guess. The local similarity attribute is used for registering time-lapse images in Fomel and Jin (2009). Hale (2013) computes time shifts in a robust and efficient manner with an improvement of a dynamic warping method originally developed for speech recognition, which he applies to the registration of P-P and P-S time-migration images. Baek et al. (2014) present a robust time/amplitude warping method with piecewise polynomials and low-frequency augmented signals, and they successfully use their method in waveform inversion to mitigate cycle skipping. All of these methods can potentially be applied for the registration of depth-migrated images.

In this study, we propose a new methodology to invert for S-wave velocities by combining IDWT (Yang et al., 2014a), dynamic image warping (DIW) (Hale, 2013), and the registration-guided least-squares (RGLS) method (Baek et al., 2014). The paper is organized as follows: We first briefly describe the ERTM and DIW algorithms that are used to form and register P-S images, respectively. We then introduce the theory of elastic IDWT, and we modify it under the RGLS framework for S-wave model updating. A simple and intuitive three-layer model is used to illustrate the method. A modified elastic Marmousi model demonstrates the robustness and efficiency of our new method for complex structures. In the ‘‘Discussion’’ section, we cover the advantages, limitations, and practical issues of this method.

THEORY

Elastic reverse time migration

Acoustic reverse time migration (RTM) (Baysal et al., 1983; McMechan, 1983) is robust for imaging in complex geology, and is becoming the imaging workhorse in exploration seismics. ERTM is, however, more complicated because the wavefields are vectors rather than scalars. To form P-P and P-S images separately, the wave modes should be separated during migration (Yan and Sava, 2008). Dellinger and Etgen (1990) propose separating the extrapolated wavefield into P and S potentials; we follow their approach in isotropic media. Any vector field $\mathbf{u}(\mathbf{x}, t)$ can be written as

$$\mathbf{u} = \nabla\Phi + \nabla \times \Psi, \quad (1)$$

where Φ is the scalar potential, Ψ is the vector potential, and $\nabla \cdot \Psi = 0$ (Aki and Richards, 2002). The potentials can be obtained separately, but indirectly, by applying the divergence and curl operators to the field $\mathbf{u}(\mathbf{x}, t)$:

$$P = \nabla \cdot \mathbf{u} = \nabla^2\Phi, \quad (2)$$

$$\mathbf{S} = \nabla \times \mathbf{u} = -\nabla^2\Psi. \quad (3)$$

For isotropic elastic media, the P mode is the compressional component of the wavefield propagating at speed V_p , and the \mathbf{S} mode is the transverse component propagating at speed V_s . It is worth noting that the theory is valid when the medium is homogeneous. However, in practice, a smoothed medium can be considered to be piecewise homogeneous. If a smooth velocity model as required in migration algorithms is provided, the decomposition is numerically stable.

In ERTM, we forward propagate the source to simulate a source wavefield \mathbf{u}_s , and back propagate the multicomponent data to form a receiver wavefield \mathbf{u}_r . During the propagations, we use the Helmholtz decomposition to separate \mathbf{u}_s into compressional component P_s and transverse component \mathbf{S}_s , and we separate \mathbf{u}_r into P_r and \mathbf{S}_r . By permuting the separated wave modes, we can form P_sP_r , $P_s\mathbf{S}_s$, $P_s\mathbf{S}_r$, \mathbf{S}_sP_r , and $\mathbf{S}_s\mathbf{S}_r$ images (Yan and Sava, 2008). Without hurting the generality, here we choose P_sP_r and $P_s\mathbf{S}_r$ images for the following analysis. To simplify notations, we use P-P for P_sP_r , and P-S for $P_s\mathbf{S}_r$.

Because the \mathbf{S}_r mode is a vector field, the imaging conditions for P-S images vary in how the vectors are treated. For example, a P-S image can be obtained by applying

$$I_{PS} = \int_0^T (\nabla P_s) \cdot (\nabla \times \mathbf{S}_r) dt, \quad (4)$$

or

$$I_{PS} = \int_0^T P_s (\mathbf{I}_3 \cdot \mathbf{S}_r) dt, \quad (5)$$

where $\mathbf{I}_3 = [1 \ 1 \ 1]$, and T is the seismic record length. Other imaging conditions (e.g., crosscorrelating component by component, Yan and Sava [2008] and crosscorrelation between spatial gradients and temporal derivatives of different modes, Shang et al., 2012) can also be applied. The choice of imaging condition does not influence the validity of our framework.

Given a well-estimated P-wave velocity model and a starting S-wave velocity model, we can use ERTM to form P-P and P-S images. The structures of the subsurface reflectors can be identified from the P-P images. Due to the inaccurate S-wave velocity model, the reflectors in the P-S images deviate from those in the P-P images in position. This deviation contains the information of the S-wave velocity errors, and it can be represented by a shift function obtained with image registration methods such as DIW.

Dynamic image warping for elastic images

As described in Hale (2013) and Yang et al. (2014a), a migrated image I made with an incorrect velocity can be considered a warped version of the true image \hat{I} , if I and \hat{I} are close. In equation 6, $\mathbf{w}(\mathbf{x})$ is a vector warping function that specifies how much the image point at \mathbf{x} in $\hat{I}(\mathbf{x})$ is shifted from the same image point in $I(\mathbf{x})$:

$$I(\mathbf{x}) = \hat{I}(\mathbf{x} + \mathbf{w}(\mathbf{x})). \quad (6)$$

Given $I(\mathbf{x})$ and $\hat{I}(\mathbf{x})$, we can pose the optimization problem to solve for $\mathbf{w}(\mathbf{x})$ as

$$\mathbf{w}(\mathbf{x}) = \arg \min_{\mathbf{l}(\mathbf{x})} D(\mathbf{l}(\mathbf{x})), \quad (7)$$

where

$$D(\mathbf{l}(\mathbf{x})) = \int_{\mathbf{x}} |I(\mathbf{x}) - \hat{I}(\mathbf{x} + \mathbf{l}(\mathbf{x}))|^2 d\mathbf{x}. \quad (8)$$

Recent developments in Hale (2013) and Baek et al. (2014) provide efficient algorithms to solve similar problems for time warping with smooth constraints. The application of these methods for image warping is straightforward. To be consistent with our previous work in Yang et al. (2014a), we choose to use the dynamic time/image warping (Hale, 2013) to find the depth shifts between P-S and P-P images.

It should be noted that if the S-wave velocity is incorrect, the differences between P-P and P-S images are more than just misalignments. Amplitudes, phases, and wavelengths of the reflectors may vary due to different radiation patterns of P- and S-waves, and misfocusing of P-S images, not to mention coherent or incoherent noise in the images. Hale (2013) proves the robustness of the DIW algorithm with a field data example, in which P-P and P-S events are well aligned by registration despite such differences.

One issue that does require discussion, however, is that in P-S RTM images, the events with flipped polarity in the P-S image will be misregistered with the P-P events. To mitigate this, we modify equation 8 into

$$D(\mathbf{l}(\mathbf{x})) = \int_{\mathbf{x}} |\mathcal{E}[I_{PP}(\mathbf{x})] - \mathcal{E}[I_{PS}(\mathbf{x} + \mathbf{l}(\mathbf{x}))]|^2 d\mathbf{x}, \quad (9)$$

where \mathcal{E} is an operator that fixes the polarity issues of the P-S images. There are several ways to correct the polarities in P-S images. One efficient way is to use Poynting vectors (Du et al., 2012b; Shang et al., 2012). In this case, \mathcal{E} is a mask that reverses all the flipped polarities in a P-S image, and has no effect on the P-P image. One can also choose \mathcal{E} as the Hilbert transform or absolute value function to register the envelope or magnitude of the image, respectively. In 3D, the operator \mathcal{E} is complicated. In our synthetic examples, we use Poynting vectors to correct the polarities because this is also necessary to visualize the final imaging results. Although the Poynting vector method is difficult to stabilize, it is easy to implement in the time domain where our methodology is derived.

With DIW, we extract the S-wave velocity error information from the P-S images using the P-P images as references. The warping function $\mathbf{w}(\mathbf{x})$ is a quantitative measure of the P-S image distortion. It can be minimized to update the velocities as in Yang et al. (2014a), or it can be used to correct the P-S image directly by shifting the image points as in the following section.

Elastic image-domain wavefield tomography

Data-domain inversion methods such as full-waveform inversion (FWI) (Tarantola, 1984; Virieux and Operto, 2009), are designed to estimate model parameters by fitting observed data with simulated

data. If we assume that an observed image $I_{\text{obs}}(\mathbf{x})$ is available, a similar least-squares fitting cost function can be written as

$$J(m) = \frac{1}{2} \sum_{x_s} \int_{\mathbf{x}} |I(\mathbf{x}, x_s, m) - I_{\text{obs}}(\mathbf{x}, x_s)|^2 d\mathbf{x}, \quad (10)$$

where I is the image we want to construct, \mathbf{x} is the spatial vector, x_s is the source index, and m is the velocity model to be recovered. Such methods are not commonly used for initial model building (e.g., WEMVA) because there are no observed images. Instead, velocity errors are characterized by the features of the events in image gathers, for example, misfocusing in time-lag gathers (Sava and Fomel, 2006; Yang and Sava, 2011) or flatness in angle gathers (Sava and Fomel, 2003; Yan and Sava, 2008).

In time-lapse situations, we use the baseline image as the observed image. The time-lapse velocity changes are estimated by fitting the baseline images with the monitor images (Yang et al., 2014b). In the context of S-wave velocity model building, we can use the P-P image as the observed image, given an accurate P-wave velocity model. However, the P-P image cannot be directly compared with the P-S image in equation 10 for several reasons. First, the P-P and P-S images generally have different wavelengths and amplitudes. Second, the P-S image has zero values at normal incidence locations. Third, the P-S and P-P images, from the same shot profile, often have different areas of illumination.

We thus require a way of estimating a modeled P-S image from the P-P image. This can be achieved by the reflector locations in the P-P image to synthesize a target P-S image \hat{I}_{PS} . The key step is to extract the location information from the P-P image. Without needing to identify individual reflectors or define a wavelet, DIW as discussed above is an efficient tool to estimate the warping function $\mathbf{w}(\mathbf{x})$ that describes how much the reflectors in the P-S image are shifted away from those in the P-P image. With $\mathbf{w}(\mathbf{x})$, we are able to match the reflectors in the two images, so that we know which reflectors we need to synthesize the response from. The target P-S image can then be obtained efficiently by shifting the P-S reflectors in I_{PS} to the locations of the P-P reflectors in I_{PP} using $\mathbf{w}(\mathbf{x})$. The cost function for S-wave velocity inversion can then be written as

$$J(\beta) = \frac{1}{2} \sum_{x_s} \int_{\mathbf{x}} |I_{PS}(\mathbf{x}, x_s, \beta) - \hat{I}_{PS}(\mathbf{x}, x_s)|^2 d\mathbf{x}, \quad (11)$$

where $\hat{I}_{PS}(\mathbf{x}, x_s) = I_{PS}(\mathbf{x} + \mathbf{w}(\mathbf{x}), x_s)$, and β is the S-wave velocity model. Because we minimize the image differences in equation 11, the P-S images will have the same reflector locations as the P-P image, and the S-wave velocity will be recovered.

We call this least-squares optimization problem in equation 11 *elastic image-domain wavefield tomography* (EIDWT). Similar to derivations for FWI and acoustic IDWT (Yang et al., 2014a), the formula for the gradient of $J(\beta)$ can be derived using the adjoint-state method (Plessix, 2006). The gradient of $J(\beta)$ with respect to the Lamé parameters λ and μ can be written as a correlation between wavefields:

$$\frac{\partial J}{\partial \lambda} = - \int_0^T (\nabla \cdot \nu_s)(\nabla \cdot \mathbf{u}_s) + (\nabla \cdot \nu_r)(\nabla \cdot \mathbf{u}_r) dt, \quad (12)$$

$$\begin{aligned} \frac{\partial J}{\partial \mu} = & -\frac{1}{2} \int_0^T \{ [\nabla \nu_s + (\nabla \nu_s)^T] : [\nabla \mathbf{u}_s + (\nabla \mathbf{u}_s)^T] \\ & + [\nabla \nu_r + (\nabla \nu_r)^T] : [\nabla \mathbf{u}_r + (\nabla \mathbf{u}_r)^T] \} dt, \end{aligned} \quad (13)$$

where \cdot is the dot product and $:$ is the Frobenius inner product. We denote by \mathbf{u}_s and \mathbf{u}_r the forward-propagated source wavefield and back-propagated receiver wavefield, respectively; ν_s and ν_r are the associated adjoint wavefields. These are obtained by solving the elastic-wave equation

$$\rho \ddot{\mathbf{v}} = \nabla \cdot (\mathbf{c} : \nabla \nu) + \mathcal{A}, \quad (14)$$

where ρ is the density; the elasticity tensor \mathbf{c} can be noted by $c_{ijkl} = \lambda \delta_{jk} \delta_{lm} + \mu (\delta_{ji} \delta_{km} + \delta_{jm} \delta_{kl})$ for an isotropic elastic medium; δ is the Kronecker delta; and \mathcal{A} is the adjoint source, which varies with different elastic imaging conditions. To make the following derivation concise and consistent with our numerical examples, we consider an ERTM algorithm that uses the scalar imaging condition from equation 5:

$$I_{\text{PS}} = \int_0^T (\nabla \cdot \mathbf{u}_s(t)) \mathbf{I}_3 \cdot (\nabla \times \mathbf{u}_r(T-t)) dt, \quad (15)$$

where $\mathbf{I}_3 = [1 \ 1 \ 1]$. In this case, the adjoint sources are

$$\mathcal{A}_s = -\nabla \cdot (\phi_I \mathbf{I}_3 \cdot (\nabla \times \mathbf{u}_r)), \quad (16)$$

and

$$\mathcal{A}_r = \nabla \times (\phi_I \mathbf{I}_3 \cdot (\nabla \cdot \mathbf{u}_s)), \quad (17)$$

for ν_s and ν_r , respectively, where $\phi_I = I_{\text{PS}} - \hat{I}_{\text{PS}}$.

Based on the relationship between the S-wave velocity β and the Lamé parameters (Mora, 1987), we have

$$\frac{\partial J}{\partial \beta} = 2\rho\beta \frac{\partial J}{\partial \mu} - 4\rho\beta \frac{\partial J}{\partial \lambda}. \quad (18)$$

Because we assume the P-wave velocity is known, and the source side wavefield $\nabla \cdot \mathbf{u}_s$ used to form I_{PS} is controlled by only the P-wave velocity, the actual contribution to $\partial J / \partial \beta$ from the source side is negligible. Therefore, a more economical formula for the gradient can be formed by combining equations 12, 13, and 18 and dropping all the source side terms, resulting in

$$\begin{aligned} \frac{\partial J}{\partial \beta} = & -\rho\beta \int_0^T \{ [\nabla \nu_r + (\nabla \nu_r)^T] : [\nabla \mathbf{u}_r + (\nabla \mathbf{u}_r)^T] \} dt \\ & + 4\rho\beta \int_0^T (\nabla \cdot \nu_r) (\nabla \cdot \mathbf{u}_r) dt. \end{aligned} \quad (19)$$

From equation 19, we see that only ν_r needs to be calculated with \mathcal{A}_r (equation 17). More details on EIDWT with different imaging conditions are presented in Appendix A.

Multilevel optimization

As with the traditional FWI cost function, if the reflector shift is larger than a half-wavelength, measured normal to a reflector, cycle skipping makes the (minus) gradient of the cost function point in a direction that is not necessarily toward the global minimum. Starting from an empirical V_p/V_s ratio may help. However, convergence is still not guaranteed.

Back et al. (2014) propose an RGLS method to mitigate the cycle-skipping effects in FWI. In the cost function, fractionally warped synthetic data are used as the new observed data to make sure the phase differences between the synthetic waveforms and the new observed data are less than a half-wavelength. Here, we borrow this idea, and substitute a fractionally warped image I_{frac} for the fully warped image \hat{I}_{PS} in equation 11 to avoid cycle-skipping. The fractionally warped image is defined as

$$I_{\text{frac}} = I_{\text{PS}}(\mathbf{x} + \alpha \mathbf{w}(\mathbf{x})), \quad (20)$$

where $0 < \alpha \ll 1$, and $\mathbf{w}(\mathbf{x})$ is the original warping function. A sufficiently small α should be chosen to ensure that I_{frac} is close enough to I_{PS} .

The registration-guided EIDWT (RG-EIDWT) method is summarized as the following multilevel optimization problem: Given the S-wave velocity model β_k after iteration k :

- 1) We use DIW to solve equation 7 for the warping function $\mathbf{w}(\mathbf{x})$ that registers the current I_{PS} to I_{PP} .
- 2) We fractionally warp I_{PS} to I_{frac} with $\alpha \mathbf{w}(\mathbf{x})$, and use EIDWT to minimize

$$J(\beta_{k+1}) = \frac{1}{2} \sum_{x_s} \int_{\mathbf{x}} |I_{\text{PS}}(\mathbf{x}, x_s, \beta_{k+1}) - I_{\text{frac}}(\mathbf{x}, x_s)|^2 d\mathbf{x}, \quad (21)$$

to get β_{k+1} iteratively.

- 3) We go back to step (1), adjust α , and repeat step (2).

In the process above, the overall image shift between the original I_{PS} and I_{PP} is minimized fraction by fraction. The parameter α determines the size of each fraction; this choice could also be optimized to expedite the overall convergence. To avoid cycle skipping, the safest choice for α should satisfy $\max(\alpha \mathbf{w}(\mathbf{x})) = d/2$, where d is the normal spatial wavelength of the reflectors in I_{PS} . Because d varies with the S-wave velocity, further optimization can be achieved by an $\alpha(\mathbf{x})$ that is a function of space. In this paper, we use a single-valued α to simplify the process.

EXAMPLES USING SYNTHETIC DATA

We demonstrate our new method with two synthetic examples. A simple three-layer model is used to illustrate how RG-IDWT recovers the interval S-wave velocities when the image resolution is high with respect to the layer thickness. A modified Marmousi model is used to show the robustness of RG-IDWT when subsurface structures are complex. It also shows the potential of the resulting S-wave velocity model as a starting model for FWI. The DIW is only executed in the depth direction, which is sufficient in our tests with reflection data. For near-vertical reflectors, DIW in the horizontal directions will add more constraints. All of our examples are in 2D;

the extension to 3D is straightforward except for the handling of polarity in 3D.

Three-layer model

Figure 1 shows the true P- and S-wave velocity models used in this test. Eight sources are placed on the surface, evenly spaced at an interval of 300 m. We use 300 receivers with a 10-m spacing, also on the surface, to cover the entire model. Synthetic data sets are generated by an elastic finite-difference solver. Perfectly matched layers are used on all four sides of the model. Therefore, the synthetic data are free of surface-related multiples. The source wavelet is a standard Ricker wavelet centered at 15 Hz. The x - and z -components of the waves are collected.

We assume that through velocity model building, a smooth version of the P-wave velocity model is available, sufficient for migration. Instead of assigning an empirical V_P/V_S ratio, we use a constant 1900 m/s S-wave velocity model as the starting model. We perform ERTM beginning from the smooth P and constant S velocity models. In Figure 2, we piece the P-P and P-S images together to show the depth mismatch of the reflectors. The left half is the P-P image, and the right half is the P-S image. Both are formed by stacking the RTM images of all the shots. Because the constant S-wave velocity is higher than the true velocity in the first layer (1767 m/s), the first reflector in the P-S image is shifted downward by about a half-wavelength. The S-wave velocity of the second layer is higher than 1900 m/s, so the second reflector is shifted a little less.

Figure 3a shows the P-S image produced with one shot gather. The black star marks the location of the shot. To register with the

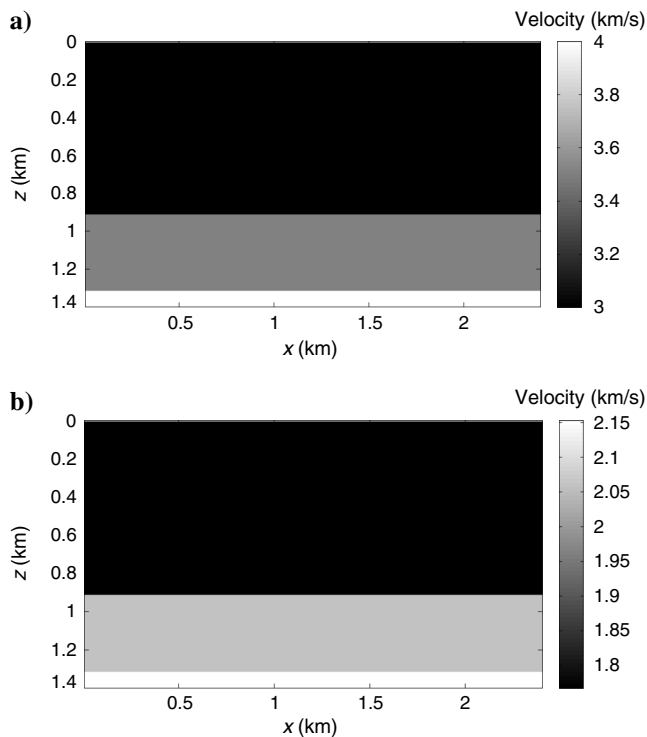


Figure 1. (a) True P-wave velocity model and (b) true S-wave velocity model. The S-wave velocities in three layers are 1767, 2060, and 2150 m/s from top to bottom.

P-P image, the polarities of the reflectors to the left of the shot are corrected to improve the consistency. The DIW calculates the warping function in Figure 3b that shows the maximum depth shift of the P-S image to be 70 m. To form I_{frac} in equation 20, we need to use the original P-S image (without polarity correction) as shown in Figure 4a instead of the polarity-corrected image. We multiply the warping function by $\alpha = 0.5$, and use it to warp Figure 4a toward the reflectors in the P-P image. Figure 4b shows a zoom of the reflectors (dashed line in Figure 4a). The blue wiggles, which are

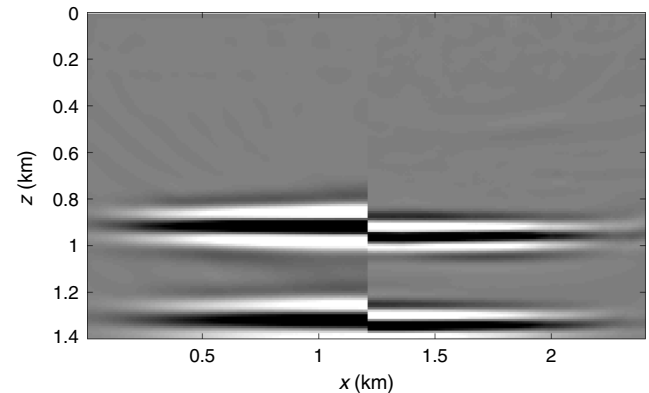


Figure 2. Comparison between the P-P (left half) and P-S (right half) images. Both images are formed with all eight sources. The top reflector in the P-S image is shifted downward by about a half-wavelength.

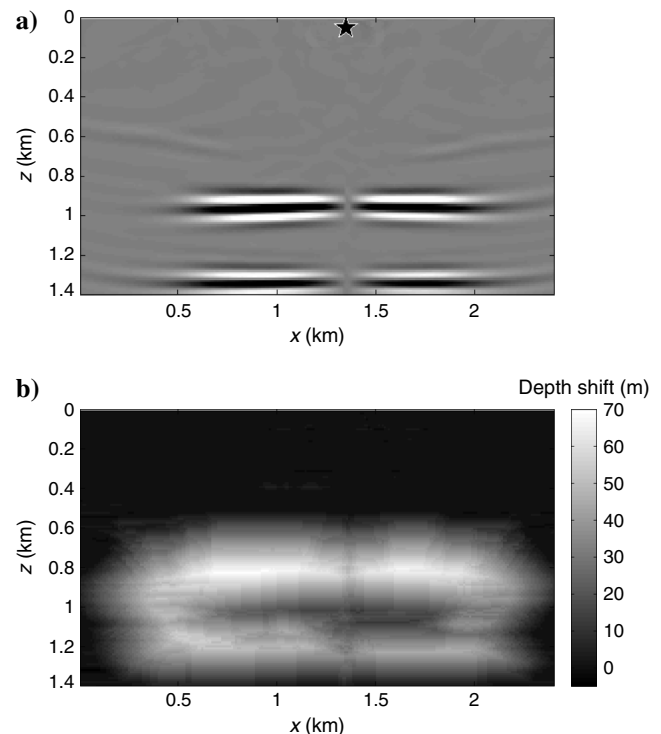


Figure 3. (a) Polarity-corrected P-S image formed with one shot gather. The black star marks the location of the source. (b) The warping function is calculated by DIW. It describes how much the depth shift is for each image point in the P-S image in panel (a).

the warped reflectors, are shifted upward from the red ones (original image). The subtraction between the red and blue images generates the ϕ_l in equations 16 and 17 to form the adjoint sources.

Figure 5a shows the gradient calculated with the adjoint sources. The dominant energy is on the wavepaths from the adjoint sources to the receivers. This is because only the shear wavefield from the receivers is used to form the P-S image. As a result, the perturbation of the P-S image is only sensitive to the S-wave velocities along the S-wave propagation paths in the receiver field. Because the P-wave velocity is correct, there is no energy on the source-side wavepath. By summing the gradients from all of the shots, we obtain the total gradient in Figure 4b used to update the velocity model. The positive values in the gradient indicate that the current velocity is generally too high.

After 20 iterations, we obtain the final S-wave velocity model as shown in Figure 6a. Both the low velocity (1757 m/s) in the first layer and the high velocity (2060 m/s) in the second layer are recovered. The third layer still has the starting velocity because there is no reflection from below. On the edges of the model, the recovery is poorer particularly for deeper parts, due to illumination limits. We compare the final P-S image with the P-P image in Figure 6b. It is obvious that the reflectors in the P-S image (right half) are aligned with those in the P-P image (left half). The alignment is also not as

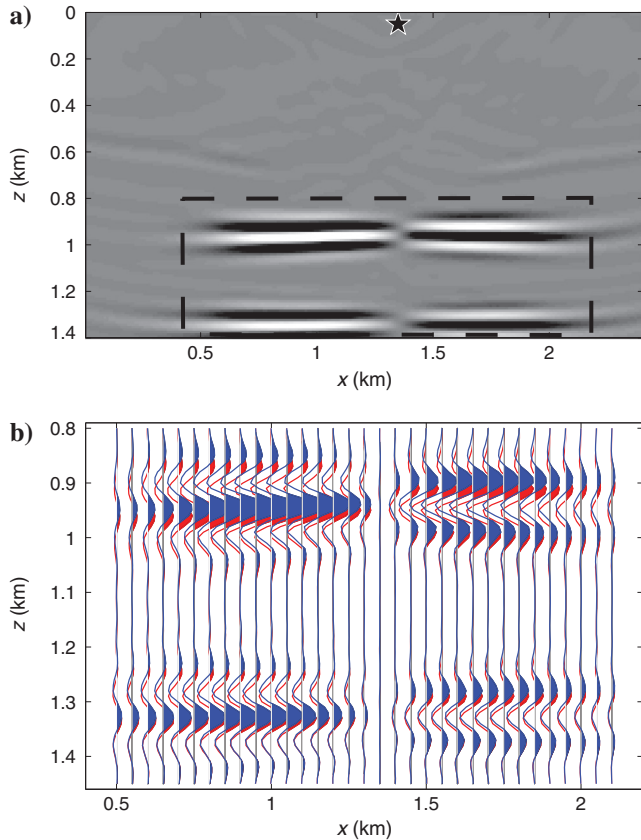


Figure 4. (a) Original P-S image without polarity correction. The same source is used as in Figure 3a. (b) Magnified view of the reflectors marked by the black dashed line in panel (a). The original image (red wiggles) and the fractionally warped image (blue wiggles) are shown together. The differences between them are used to generate the adjoint sources.

good on the edges due to the same illumination limits. We expect the recovery to be improved with a wider acquisition.

Modified Marmousi

The three-layer model example is a good showcase for interval velocity recovery when the imaging resolution is much higher than the layer thickness. However, the layering of the real subsurface can be very detailed and complicated, and the resolutions of the seismic images are limited by the frequency content of the data. In this example, we use a modified Marmousi model to show the behavior of RG-EIDWT in realistic geology, when the layer thickness is lower than, or comparable with, the image resolution.

Figure 7 shows the true P- and S-wave velocity models. To facilitate numerical experiments, we compress both velocity ranges, as compared with standard Marmousi, to allow for a larger grid size and time step in the finite-difference modeling. We use 18 sources, evenly distributed with 480-m spacing on the surface, and 750 receivers are also placed on the surface, covering the entire model, with a spacing of 12 m. We simulate a land acquisition, so the top layer of the model is solid, but perfectly matched layers are used to avoid a free surface condition. We collect the x - and z -components data. The source wavelet is a standard Ricker centered at 20 Hz.

We assume that a smooth P-wave velocity model, as shown in Figure 8a, is obtained with initial model building tools. The S-wave velocity is taken as unknown, so we use a constant velocity model ($V_S = 2500$ m/s) as the starting model for ERTM. Figure 8b is the P-P image produced with all 18 sources using the model in

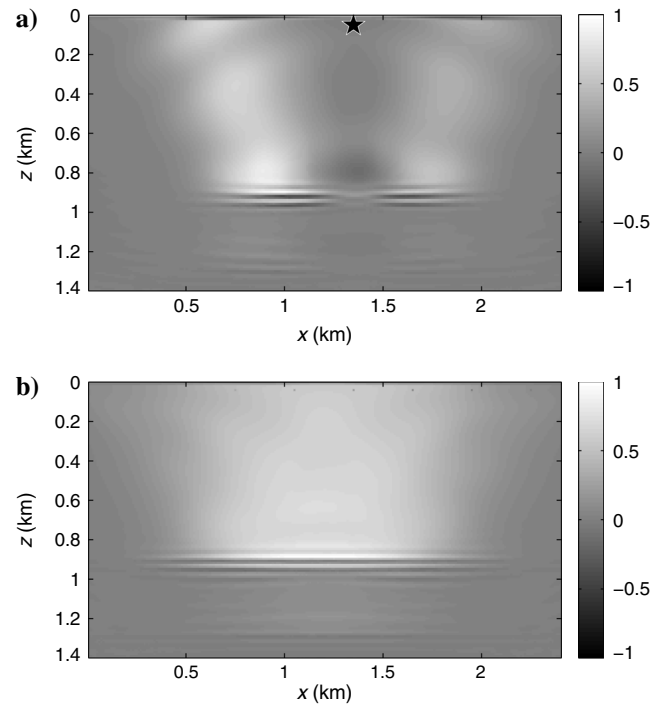


Figure 5. (a) The gradient for the source marked by the black star. The dominant energy is along the receiver-side wavepath because the source-side depends only on the P-wave velocity. (b) Total gradient formed by stacking the partial gradients from each source. Positive values indicate that the current velocity is too high.

Figure 8a. In ERTM, the wave modes are separated into P and S potentials, and the propagation of the P mode is dependent only on the P-wave velocities. As a result, even without the correct

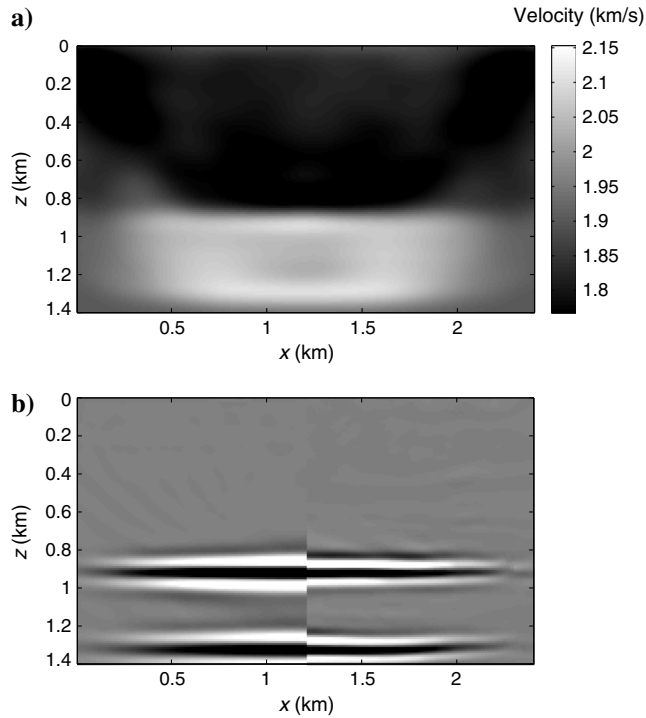


Figure 6. (a) The recovered S-wave velocity model after 20 iterations. Both the low- and high-velocity layers are resolved. The recovery is limited by the illumination of the survey. (b) The P-S image (right) formed with the recovered S-wave velocity model in panel (a) is compared with the P-P image (left). Both reflectors are aligned. The alignment is poor on the edges due to the same illumination limits.

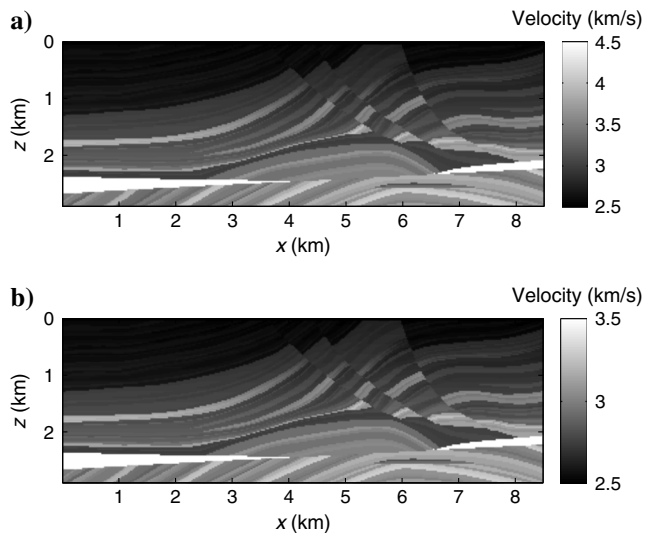


Figure 7. The modified elastic Marmousi model: (a) True P-wave velocity model and (b) true S-wave velocity model. Velocity ranges are modified to be compressed compared with the original Marmousi model to allow larger grid sizes and time steps in finite difference. The top layer is solid instead of water.

S-wave velocities, the P-P image still shows the reflectors at the correct locations. This image is used as the reference image for registering the P-S images.

Figure 9a is the P-S image generated with one shot gather with the polarities corrected. The shot location is marked by the black star. Due to the low starting S-wave velocities, the entire image is shifted upward from the P-P image. DIW is used to calculate the warping function as shown in Figure 9b, which describes the depth shift of each image point in the P-S image from the corresponding image point in the P-P image. As described in Hale (2013), DIW is robust to the differences between P-S and P-P images. Close scrutiny finds that the amplitudes of the reflectors in the P-P and P-S images are different. The P-S image also has some

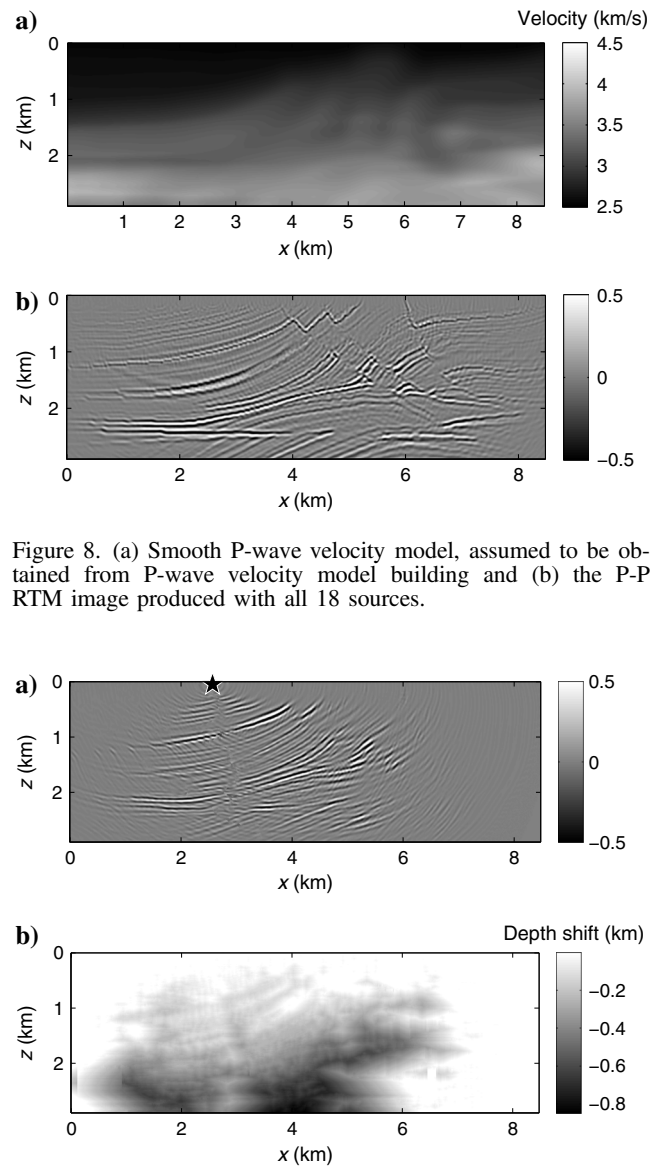


Figure 8. (a) Smooth P-wave velocity model, assumed to be obtained from P-wave velocity model building and (b) the P-P RTM image produced with all 18 sources.

Figure 9. (a) A P-S image produced with one shot gather. The polarities are corrected. The black star marks the location of the source. (b) The warping function that registers the P-S image in panel (a) to the P-P image in Figure 8b. The shifts are all negative (upward) because the starting S-wave velocity is lower than all velocities in the true S-wave velocity model.

imaging artifacts that do not exist in the P-P image. The limited acquisition aperture could cause the spurious patterns in the P-S image due to the edge effect. The incorrect S-wave velocities also would likely introduce wave-mode leakage, and generate artificial reflectors in the P-S image. The wave-mode separation is also not perfect because of the variations in the velocity model, which clearly violates the homogeneous assumption in the Helmholtz decomposition (Aki and Richards, 2002). Nonetheless, DIW is robust to these artifacts because it looks for a global solution for the entire image volume. The real P-S events are more coherent with those in the P-P image and have relatively higher amplitudes than the artifacts. Misregistration might happen to individual events, but DIW mitigates these errors by forcing the smoothness of the warping function along the events. For locations that are outside of the illumination, for example, the right end in Figure 9a, DIW assigns zeros because no registration can be achieved in these areas, shown with white in Figure 9b. The success of image registration is the foundation of the RG-EIDWT process.

We warp the original P-S image (without polarity correction) by a fraction of the warping function, and form the adjoint sources to calculate the gradient as shown in Figure 10a. The adjoint sources are distributed sources in the entire image volume. By comparing Figure 10a and Figure 9a, we see that the amplitudes in the gradient are proportional to the amplitudes in the image because the magnitude of the adjoint sources is scaled by the magnitude of the image differences (equations 16 and 17). By stacking the gradients from all 18 sources, we obtain the total gradient shown in Figure 10b. It only has negative values, indicating that the S-wave velocities for the current iteration are too low.

After 50 iterations, we obtain the final S-wave velocity model in Figure 11. The macrovelocity distribution is consistent with the true S-wave velocity model. However, the interval velocities of individual layers are not resolved because the image resolution is not high enough. Here, the maximum image resolution is 62.5 m, whereas the typical layer thickness is 60 m. The salt bodies on the sides are also not recovered because of illumination limitations. The struc-

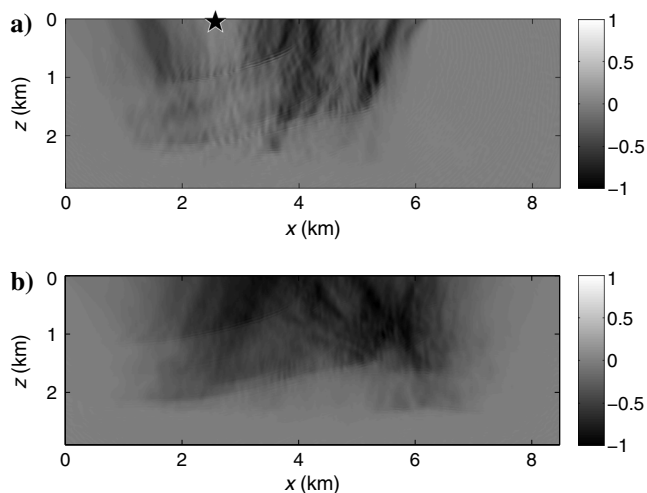


Figure 10. (a) The normalized gradient calculated with the same source. The black star marks the source location. (b) The normalized total gradient by summing the gradients from all 18 sources. Negative values indicate that the current velocities need to be increased.

tures beneath the salt layers are not recovered because the converted S-waves are outside the acquisition aperture due to the reflector dips. We expect this region to improve with a wider acquisition.

Although the model in Figure 11 has low resolution, it contains the kinematics to place the P-S events at the correct locations. Figure 12 shows the P-S image before and after the inversion. We divide the entire image into seven sections, and the P-P and P-S image sections are shown in an alternating pattern (i.e., the odd sections are P-P images, and the even sections are P-S images). In Figure 12a, the mismatch between the images is obvious, and it gets larger deeper in the model because the kinematic errors accumulate with depth. In particular, for the reservoir location, outlined by the dashed circle, the image is severely shifted from the correct location. The coherency of the reflectors is poor, and the polarity of the strong reflector is wrong. Interpretations would likely be unreliable based on this image. Figure 12b shows the same alternating P-P/P-S

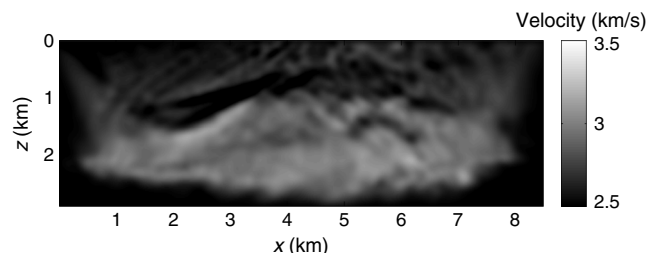


Figure 11. The final S-wave velocity model after 50 iterations. The macrovelocity distribution is recovered. The bottom part of the model is poorly resolved because the converted S-waves are outside of the acquisition surface due to the dipping reflectors.

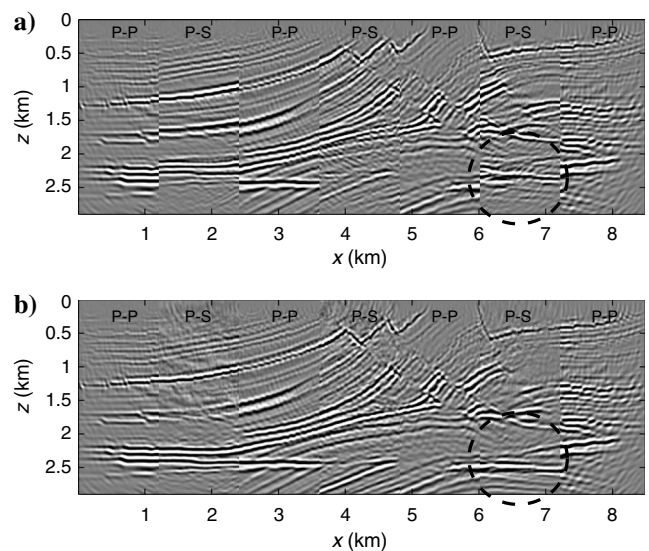


Figure 12. (a) Comparison between P-P and P-S images before the inversion. The image is divided into seven sections in the horizontal direction. Odd sections are P-P images, and even sections are P-S images. The mismatch is clear at the section interfaces. In the circled area, the reflectors are not well imaged due to the incorrect S-wave velocity and (b) the P-S image based on the inverted S-wave velocity model is compared with the P-P image in the same setup as in panel (a). The coherency at the section interfaces is markedly improved. The reflectors in the circled area are all well-resolved and aligned with those in the P-P image.

image generated with the inverted S-wave velocity model in Figure 11. The P-P and P-S events are well aligned, even for the deep reflectors that are shifted by more than one wavelength in Figure 12a. For the reservoir area (black dashed circle), the strong reflector is at the correct location, and the polarity is also corrected. The interfaces that are blurry before the inversion are coherent and well-imaged.

With the corrected background velocities, we can apply FWI to further improve the resolution of the models. We use the same data but only the reflection part with a maximum offset of 3.6 km to perform FWI starting from the models in Figure 8a (for P) and Figure 11 (for S). Figure 13 shows the P- and S-wave reflectivity models after the inversion. The image resolution is remarkably improved, and the reflectivities are true physical parameters rather than image points. Except for the areas with poor illumination, the S-wave reflectivity model is of similar quality to the P-wave reflectivity model. Without the S-wave velocity inversion with RG-EIDWT, it is difficult to obtain the image in Figure 13b.

DISCUSSION

Effect of image quality

We have shown that DIW and RG-EIDWT are robust to imaging artifacts. However because the method is based on single shot-migration images, random noise in the image might be a cause for concern. To mitigate this, we could stack images from adjacent shots, relying on the robustness of RTM and DIW to random noise to reduce the artifacts. RG-EIDWT could use the stacked image to do registration and form adjoint sources.

Multiples in the data can also cause imaging artifacts. Because of the velocity variations, the artificial reflectors generated by the P-wave multiples and those generated by the S-wave multiples are not necessarily consistent in depth, even if the correct velocity model is

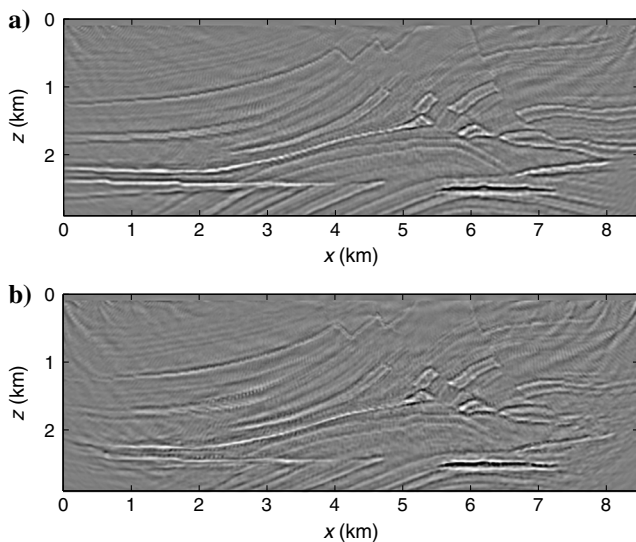


Figure 13. (a) The P-wave reflectivity model resolved by FWI starting from the smooth model in Figure 8a and (b) the S-wave reflectivity model resolved by FWI starting from the smooth model in Figure 11. Except for poorly illuminated areas, the recovery of the S-wave model is of similar quality to that of the P-wave model in panel (a). This demonstrates the success of the S-wave velocity model building with RG-EIDWT.

provided. When multiples are strong, the P-P image might provide the inaccurate guidance for the P-S image. Therefore, demultiple processing in the data domain is preferred before applying our method.

Our method depends on the consistency between the P-P and P-S reflectors. However, missing P-P or P-S events are common in certain geologic settings such as under the gas cloud or special lithology. The registration and minimization will not work properly when the P-S image cannot find a target reflector in the P-P image. As an initial model building tool, RG-EIDWT does not require every P-S reflector to have a counterpart in the P-P image. As long as there are sufficient common strong events in both images, this technique will make use of that information and try to update the velocity model. However, under some circumstances in which there is no P-P image at all (e.g., under the gas cloud), it is impossible to update that part of the model.

Potential applications

The two synthetic examples we use to explain the process of RG-EIDWT also showcase the two different field data applications that we have in mind. The Marmousi example is meant to highlight the problem encountered in seismic exploration, that the image resolution is often not high enough to discriminate between thin layers in the subsurface. RG-EIDWT can be used to build a smooth background velocity model for S-waves, without requiring long-offset, low-frequency acquisitions. FWI or least-squares migration can improve the resolution of the model based on the RG-EIDWT results. With the layered model, we mimic the global seismology problem, in which reflections are often sparse in depth. RG-EIDWT can resolve the interval S-wave velocities by registering the P-P and P-S images of deep earth discontinuities and provide an average estimate of Poisson's ratio for the layers in the crust and mantle.

In the derivation and numerical tests in this paper, we show only the S-wave velocity inversion with registration between the P-P and P-S images. However, within the same framework, RG-EIDWT can be modified to invert for P-wave velocities in situations in which the S-wave velocity model is known. In addition, S-P and S-S images can also be integrated into the method if the source has strong shear components. It is also a natural extension to apply a similar methodology to 4D inversion, in which the baseline image serves as the target image and the time-lapse P- and S-wave velocity changes are inverted by matching monitor images with the baseline images.

Extension to anisotropy

Isotropic elasticity is assumed in this work. However, S-wave anisotropy is common in reality. Anisotropy would not violate the principle of our methodology; however, it would raise complications in practice. Isotropic modeling and wave-mode separation tools will not handle the anisotropy in the data correctly. Dellinger and Etgen (1990) propose an anisotropic wavefield separation method using projection of the polarization vectors. Yan and Sava (2009) improve this method with spatially varying polarization projections, and further improve the separation efficiency combining wavenumber domain separation and space domain interpolation (Yan and Sava, 2011). Given the success of the wave-mode separation in anisotropic media, we can form the P-P and P-S images, calculate the shifts, and use that information to update the S-wave velocity and its anisotropy parameter. Nonetheless, these separation

methods need to construct operators based on accurate models. The sensitivity of the separation to velocity or anisotropy parameter errors has not been studied to the best of our knowledge. If severe mode leakage appears with inaccurate anisotropy, its effect on the imaging quality and model building is not well understood. It is outside the scope of this paper, but it is an important topic for future research.

CONCLUSION

We have proposed an image registration guided wavefield tomography method in the image domain for S-wave velocity model building with the knowledge of P-wave velocities. The shifts of the P-S images with respect to the P-P image are minimized fraction by fraction to recover the S-wave velocities iteratively. The method is wave-equation-based and has no assumptions about the smoothness of the subsurface. It works well with high-frequency reflection data, and it can start with an arbitrary constant S-wave velocity model. It is computationally efficient without the calculation of angle gathers or extended images. The resulting model is smooth and serves as a good starting model for FWI or least-squares migration.

ACKNOWLEDGMENTS

This work was supported by the MIT Earth Resources Laboratory Founding Members Consortium. We especially thank the associate editor T. Alkhalifah, J. Cheng, S. Gray, and an anonymous reviewer, for their insightful comments and suggestions which helped to improve the quality of this work.

APPENDIX A

ADJOINT-STATE METHOD FOR ELASTIC IMAGE-DOMAIN WAVEFIELD TOMOGRAPHY

Here, we present the mathematic derivation of the adjoint wavefields and the gradient for elastic IDWT using the associate Lagrangian in the time domain. Following the approach of Plessix (2006), the steps of the derivation are as follows for model parameter m and cost function $J(m)$:

- 1) List all the state equations $F_i = 0$.
- 2) Build the augmented functional L by associating the independent adjoint state variables ν_i with the state equations F_i .
- 3) Define the adjoint-state equations by $\partial L / \partial u_i = 0$.
- 4) Compute the gradient by $\partial L / \partial m = \partial J / \partial m$.

To simplify the process, we assume a single shot. A more general derivation can be easily achieved by summing over all shots. The least-squares functional is

$$J(\mathbf{c}) = \frac{1}{2} \int_{\mathbf{x}} |I_{\text{PS}}(\mathbf{x}) - \hat{I}_{\text{PS}}(\mathbf{x})|^2 d\mathbf{x}, \quad (\text{A-1})$$

where the model parameter is \mathbf{c} , which is the elasticity tensor. $\hat{I}_{\text{PS}}(\mathbf{x})$ is the target P-S image, and $I_{\text{PS}}(\mathbf{x})$ is the P-S image produced with the current elasticity tensor \mathbf{c} , with the imaging condition

$$I_{\text{PS}} = \int_0^T (\nabla(\nabla \cdot \mathbf{u}_s(t))) \cdot (\nabla \times (\nabla \times \mathbf{u}_r(T-t))) dt, \quad (\text{A-2})$$

where \mathbf{u}_s is the source wavefield obtained by solving the following wave equations:

$$\begin{cases} \rho \ddot{\mathbf{u}}_s = \mathbf{f} + \nabla \cdot (\mathbf{c} : \nabla \mathbf{u}_s), \\ \mathbf{u}_s|_{t=0} = 0, \\ \dot{\mathbf{u}}_s|_{t=0} = 0, \end{cases} \quad (\text{A-3})$$

where \mathbf{f} is the source term.

The value \mathbf{u}_r is the receiver wavefield obtained by solving the following equations:

$$\begin{cases} \rho \ddot{\mathbf{u}}_r = \mathbf{d}(T-t) + \nabla \cdot (\mathbf{c} : \nabla \mathbf{u}_r), \\ \mathbf{u}_r|_{t=0} = 0, \\ \dot{\mathbf{u}}_r|_{t=0} = 0, \end{cases} \quad (\text{A-4})$$

where \mathbf{d} is the multicomponent data that are collected in the seismic survey.

For simplicity, density is assumed to be constant, and the spatial boundary conditions are left unspecified because any condition that guarantees a unique solution is acceptable. In our numerical examples, we use absorbing boundary conditions.

Using the Lagrangian formulation, we associate the adjoint states $\tilde{\mu}_s^0$, $\tilde{\mu}_s^1$, $\tilde{\mu}_r^0$, and $\tilde{\mu}_r^1$ with the initial conditions in equations A-3 and A-4, respectively. Adjoint states $\tilde{\nu}_s$ and $\tilde{\nu}_r$ are associated with the wave equations in equations A-3 and A-4. Adjoint state $\tilde{\phi}_I$ is associated with equation A-2. With the operations above, the augmented functional is defined by

$$\begin{aligned} L(\tilde{\phi}_I, \tilde{\nu}_s, \tilde{\nu}_r, \tilde{\mu}_s^0, \tilde{\mu}_s^1, \tilde{\mu}_r^0, \tilde{\mu}_r^1, \tilde{\mathbf{u}}_s, \tilde{\mathbf{u}}_r, \tilde{I}_{\text{PS}}, \mathbf{c}) \\ = \int_{\mathbf{x}} \frac{1}{2} |\tilde{I}_{\text{PS}}(\mathbf{x}) - \hat{I}_{\text{PS}}(\mathbf{x})|^2 d\mathbf{x} \\ - \int_0^T \langle \tilde{\nu}_s, \ddot{\mathbf{u}}_s - \mathbf{f} - \nabla \cdot (\mathbf{c} : \nabla \tilde{\mathbf{u}}_s) \rangle_{\mathbf{x}} dt \\ - \int_0^T \{ \langle \tilde{\mu}_s^0, \tilde{\mathbf{u}}_s \rangle_{\mathbf{x}} + \langle \tilde{\mu}_s^1, \dot{\tilde{\mathbf{u}}}_s \rangle_{\mathbf{x}} \} \delta(t) dt \\ - \int_0^T \langle \tilde{\nu}_r, \ddot{\mathbf{u}}_r - \mathbf{d}(T-t) - \nabla \cdot (\mathbf{c} : \nabla \tilde{\mathbf{u}}_r) \rangle_{\mathbf{x}} dt \\ - \int_0^T \{ \langle \tilde{\mu}_r^0, \tilde{\mathbf{u}}_r \rangle_{\mathbf{x}} + \langle \tilde{\mu}_r^1, \dot{\tilde{\mathbf{u}}}_r \rangle_{\mathbf{x}} \} \delta(t) dt \\ - \left\langle \tilde{\phi}_I, \tilde{I}_{\text{PS}}(\mathbf{x}) - \int_0^T (\nabla(\nabla \cdot \tilde{\mathbf{u}}_s(t))) \right. \\ \left. \cdot (\nabla \times (\nabla \times \tilde{\mathbf{u}}_r(T-t))) dt \right\rangle_{\mathbf{x}}, \end{aligned} \quad (\text{A-5})$$

with $\langle \mathbf{a}, \mathbf{b} \rangle_{\mathbf{x}} = \int_x \int_y \int_z \mathbf{a} \cdot \mathbf{b} dx dy dz$ the inner product of vector-valued functions in space.

By integration by parts, it is easy to prove that the second-order time derivative operator and the second-order spatial derivative operator are both self-adjoint, so we have

$$\begin{aligned} \int_0^T \langle \tilde{\nu}_s, \ddot{\tilde{\mathbf{u}}}_s \rangle_x dt &= \int_0^T \langle \ddot{\tilde{\nu}}_s, \tilde{\mathbf{u}}_s \rangle_x dt \\ &+ \int_0^T \{ \langle \tilde{\nu}_s, \dot{\tilde{\mathbf{u}}}_s \rangle_x \delta(t-T) - \langle \tilde{\nu}_s, \dot{\tilde{\mathbf{u}}}_s \rangle_x \delta(t) \\ &- \langle \dot{\tilde{\nu}}_s, \tilde{\mathbf{u}}_s \rangle_x \delta(t-T) + \langle \dot{\tilde{\nu}}_s, \tilde{\mathbf{u}}_s \rangle_x \delta(t) \} dt \end{aligned} \quad (\text{A-6})$$

and

$$\begin{aligned} \langle \tilde{\nu}_s, \nabla \cdot (\mathbf{c} : \nabla \tilde{\mathbf{u}}_s) \rangle_x &= \langle \nabla \cdot (\mathbf{c} : \nabla \tilde{\nu}_s), \tilde{\mathbf{u}}_s \rangle_x \\ &+ \{ \tilde{\nu}_s \cdot (\mathbf{c} : \nabla \tilde{\mathbf{u}}_s) - (\mathbf{c} : \nabla \tilde{\nu}_s) \cdot \tilde{\mathbf{u}}_s \} \Big|_{x=\pm\infty}. \end{aligned} \quad (\text{A-7})$$

For the imaging condition term, we can similarly derive

$$\begin{aligned} \langle \tilde{\phi}_I, (\nabla(\nabla \cdot \tilde{\mathbf{u}}_s)) \cdot (\nabla \times (\nabla \times \tilde{\mathbf{u}}_r)) \rangle_x &= \langle \tilde{\mathbf{u}}_s, \nabla((\nabla \tilde{\phi}_I) \cdot (\nabla \times (\nabla \times \tilde{\mathbf{u}}_r))) \rangle_x \\ &+ \{ (\nabla \cdot \tilde{\mathbf{u}}_s) \cdot (\tilde{\phi}_I \nabla \times (\nabla \times \tilde{\mathbf{u}}_r)) \\ &- (\tilde{\mathbf{u}}_s \cdot \mathbf{I}_3)(\nabla \tilde{\phi}_I) \cdot (\nabla \times (\nabla \times \tilde{\mathbf{u}}_r)) \} \Big|_{x=\pm\infty} \\ &= \langle \tilde{\mathbf{u}}_r, \nabla \times ((\nabla \tilde{\phi}_I) \times (\nabla(\nabla \cdot \tilde{\mathbf{u}}_s))) \rangle_x \\ &+ \{ \nabla(\nabla \cdot \tilde{\mathbf{u}}_s) \cdot (\tilde{\phi}_I \mathbf{I}_3 \times (\nabla \times \tilde{\mathbf{u}}_r)) \\ &+ (\mathbf{I}_3 \times \tilde{\mathbf{u}}_r) \cdot (\nabla \tilde{\phi}_I \times \nabla(\nabla \cdot \tilde{\mathbf{u}}_s)) \} \Big|_{x=\pm\infty}, \end{aligned} \quad (\text{A-8})$$

where $\mathbf{I}_3 = [1 \ 1 \ 1]$.

With equations A-5–A-8, we can compute the derivatives with respect to the adjoint states, and we evaluate them at $(\nu_s, \nu_r, \mathbf{u}_s, \mathbf{u}_r, \phi_I, I_{PS})$ to obtain the adjoint-state equations.

With respect to $\tilde{\mathbf{u}}_s$, we have equations

$$\begin{cases} \ddot{\nu}_s = \nabla \cdot (\mathbf{c} : \nabla \nu_s) + \mathcal{A}_s, \\ \nu_s|_{t=T} = 0, \\ \dot{\nu}_s|_{t=T} = 0, \end{cases} \quad (\text{A-9})$$

in which

$$\mathcal{A}_s = -\nabla((\nabla \phi_I) \cdot (\nabla \times (\nabla \times \mathbf{u}_r(T-t))). \quad (\text{A-10})$$

With respect to $\tilde{\mathbf{u}}_r$, we have equations

$$\begin{cases} \ddot{\nu}_r = \nabla \cdot (\mathbf{c} : \nabla \nu_r) + \mathcal{A}_r, \\ \nu_r|_{t=T} = 0, \\ \dot{\nu}_r|_{t=T} = 0, \end{cases} \quad (\text{A-11})$$

in which

$$\mathcal{A}_r = -\nabla \times ((\nabla \phi_I) \times (\nabla(\nabla \cdot \mathbf{u}_s(T-t))). \quad (\text{A-12})$$

With respect to \tilde{I}_{PS} , we have equation

$$\phi_I = I_{PS} - \hat{I}_{PS}. \quad (\text{A-13})$$

By taking the derivative of L with respect to the model parameter \mathbf{c} , we have the gradient of the cost function:

$$\frac{\partial L}{\partial \mathbf{c}} = \frac{\partial J}{\partial \mathbf{c}} = - \int_0^T (\nabla \nu_s \nabla \mathbf{u}_s + \nabla \nu_r \nabla \mathbf{u}_r) dt, \quad (\text{A-14})$$

where $\partial J / \partial \mathbf{c}$ is a fourth-order tensor. If we consider an elastic isotropic medium, the elasticity tensor \mathbf{c} can be noted by $c_{ijklm} = \lambda \delta_{jk} \delta_{il} + \mu (\delta_{jl} \delta_{km} + \delta_{jm} \delta_{kl})$. Thus, we have

$$\frac{\partial J}{\partial \lambda} = - \int_0^T (\nabla \cdot \nu_s)(\nabla \cdot \mathbf{u}_s) + (\nabla \cdot \nu_r)(\nabla \cdot \mathbf{u}_r) dt \quad (\text{A-15})$$

and

$$\begin{aligned} \frac{\partial J}{\partial \mu} &= - \frac{1}{2} \int_0^T \{ [\nabla \nu_s + (\nabla \nu_s)^T] : [\nabla \mathbf{u}_s + (\nabla \mathbf{u}_s)^T] \\ &+ [\nabla \nu_r + (\nabla \nu_r)^T] : [\nabla \mathbf{u}_r + (\nabla \mathbf{u}_r)^T] \} dt, \end{aligned} \quad (\text{A-16})$$

where $:$ is the Frobenius inner product. Based on the relationship between the S-wave velocity β and the Lamé parameters, we have

$$\frac{\partial J}{\partial \beta} = 2\rho\beta \frac{\partial J}{\partial \mu} - 4\rho\beta \frac{\partial J}{\partial \lambda}. \quad (\text{A-17})$$

The derivation varies subtly with different kinds of imaging conditions. Suppose we use a more naive imaging condition such as

$$I_{PS} = \int_0^T (\nabla \cdot \mathbf{u}_s(t)) \mathbf{I}_3 \cdot (\nabla \times \mathbf{u}_r(T-t)) dt. \quad (\text{A-18})$$

Accordingly, for the divergence and curl operators, we can derive

$$\begin{aligned} \langle \tilde{\phi}_I, ((\nabla \cdot \tilde{\mathbf{u}}_s)) \mathbf{I}_3 \cdot ((\nabla \times \tilde{\mathbf{u}}_r)) \rangle_x &= -\langle \tilde{\mathbf{u}}_s, \nabla(\tilde{\phi}_I \mathbf{I}_3 \cdot (\nabla \times \tilde{\mathbf{u}}_r)) \rangle_x + \{ \tilde{\mathbf{u}}_s \cdot \mathbf{I}_3 (\tilde{\phi}_I \nabla \times \tilde{\mathbf{u}}_r) \} \Big|_{x=\pm\infty} \\ &= \langle \tilde{\mathbf{u}}_r, \nabla \times (\tilde{\phi}_I \mathbf{I}_3 (\nabla \cdot \tilde{\mathbf{u}}_s)) \rangle_x + \{ \tilde{\mathbf{u}}_r \cdot (\tilde{\phi}_I \mathbf{I}_3 \nabla \cdot \tilde{\mathbf{u}}_s) \} \Big|_{x=\pm\infty}. \end{aligned} \quad (\text{A-19})$$

Thus, the adjoint sources are

$$\mathcal{A}_s = -\nabla(\phi_I \mathbf{I}_3 \cdot (\nabla \times \mathbf{u}_r)) \quad (\text{A-20})$$

and

$$\mathcal{A}_r = \nabla \times (\phi_I \mathbf{I}_3 (\nabla \cdot \mathbf{u}_s)). \quad (\text{A-21})$$

This simpler formula makes it easier when we reduce the problem to 2D because the curl term will only have one nonzero component:

$$\nabla \times \mathbf{u}_r = \left(0\mathbf{i}, \left(\frac{\partial u_{rx}}{\partial z} - \frac{\partial u_{rz}}{\partial x} \right) \mathbf{j}, 0\mathbf{k} \right). \quad (\text{A-22})$$

Then, the adjoint sources in 2D are

$$A_s = \left(-\frac{\left[\partial \phi_I \left(\frac{\partial u_{rx}}{\partial z} - \frac{\partial u_{rz}}{\partial x} \right) \right]}{\partial x} \mathbf{i}, -\frac{\left[\partial \phi_I \left(\frac{\partial u_{rx}}{\partial z} - \frac{\partial u_{rz}}{\partial x} \right) \right]}{\partial z} \mathbf{k} \right) \quad (\text{A-23})$$

and

$$A_r = \left(-\frac{\left[\partial \phi_I \left(\frac{\partial u_{sx}}{\partial x} + \frac{\partial u_{sz}}{\partial z} \right) \right]}{\partial z} \mathbf{i}, \frac{\left[\partial \phi_I \left(\frac{\partial u_{sx}}{\partial x} + \frac{\partial u_{sz}}{\partial z} \right) \right]}{\partial x} \mathbf{k} \right). \quad (\text{A-24})$$

REFERENCES

- Aki, K., and P. G. Richards, 2002, *Quantitative seismology*: University Science Books.
- Alkhalifah, T., 2003, Tau migration and velocity analysis: Theory and synthetic examples: *Geophysics*, **68**, 1331–1339, doi: [10.1190/1.1598126](https://doi.org/10.1190/1.1598126).
- Back, H., H. Calandra, and L. Demanet, 2014, Velocity estimation via registration-guided least-squares inversion: *Geophysics*, **79**, no. 2, R79–R89, doi: [10.1190/geo2013-0146.1](https://doi.org/10.1190/geo2013-0146.1).
- Baysal, E., D. D. Kosloff, and J. W. C. Sherwood, 1983, Reverse time migration: *Geophysics*, **48**, 1514–1524, doi: [10.1190/1.1441434](https://doi.org/10.1190/1.1441434).
- Dellinger, J., and J. Etgen, 1990, Wave-field separation in two-dimensional anisotropic media: *Geophysics*, **55**, 914–919, doi: [10.1190/1.1442906](https://doi.org/10.1190/1.1442906).
- Du, Q., F. Li, J. Ba, Y. Zhu, and B. Hou, 2012a, Multicomponent joint migration velocity analysis in the angle domain for PP-waves and PS-waves: *Geophysics*, **77**, no. 1, U1–U13, doi: [10.1190/geo2010-0423.1](https://doi.org/10.1190/geo2010-0423.1).
- Du, Q., Y. Zhu, and J. Ba, 2012b, Polarity reversal correction for elastic reverse time migration: *Geophysics*, **77**, no. 2, S31–S41, doi: [10.1190/geo2011-0348.1](https://doi.org/10.1190/geo2011-0348.1).
- Fomel, S., and M. Backus, 2003, Multicomponent seismic data registration by least squares: 83rd Annual International Meeting, SEG, Expanded Abstracts, 781–784.
- Fomel, S., M. Backus, K. Fouad, B. Hardage, and G. Winters, 2005, A multistep approach to multicomponent seismic image registration with application to a west Texas carbonate reservoir study: 85th Annual International Meeting, SEG, Expanded Abstracts, 1018–1021.
- Fomel, S., and L. Jin, 2009, Time-lapse image registration using the local similarity attribute: *Geophysics*, **74**, no. 2, A7–A11, doi: [10.1190/1.3054136](https://doi.org/10.1190/1.3054136).
- Gray, S., J. Etgen, J. Dellinger, and D. Whitmore, 2001, Seismic migration problems and solutions: *Geophysics*, **66**, 1622–1640, doi: [10.1190/1.1487107](https://doi.org/10.1190/1.1487107).
- Hale, D., 2013, Dynamic warping of seismic images: *Geophysics*, **78**, no. 2, S105–S115, doi: [10.1190/geo2012-0327.1](https://doi.org/10.1190/geo2012-0327.1).
- Herrenschmidt, A., P.-Y. Granger, F. Audebert, C. Gerea, G. Etienne, A. Stopin, M. Alerini, S. Lebeget, G. Lambar, P. Berthet, S. Nebieridze, and J.-L. Boelle, 2001, Comparison of different strategies for velocity model building and imaging of PP and PS real data: *The Leading Edge*, **20**, 984–995, doi: [10.1190/1.1487320](https://doi.org/10.1190/1.1487320).
- Hokstad, K., 2000, Multicomponent Kirchhoff migration: *Geophysics*, **65**, 861–873, doi: [10.1190/1.1444783](https://doi.org/10.1190/1.1444783).
- Kuo, J., and T. Dai, 1984, Kirchhoff elastic wave migration for the case of noncoincident source and receiver: *Geophysics*, **49**, 1223–1238, doi: [10.1190/1.1441751](https://doi.org/10.1190/1.1441751).
- McMechan, G. A., 1983, Migration by extrapolation of time-dependent boundary values: *Geophysical Prospecting*, **31**, 413–420, doi: [10.1111/j.1365-2478.1983.tb01060.x](https://doi.org/10.1111/j.1365-2478.1983.tb01060.x).
- Mora, P., 1987, Nonlinear two-dimensional elastic inversion of multioffset seismic data: *Geophysics*, **52**, 1211–1228, doi: [10.1190/1.1442384](https://doi.org/10.1190/1.1442384).
- 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](https://doi.org/10.1111/j.1365-246X.2006.02978.x).
- Sava, P., 2011, Micro-earthquake monitoring with sparsely sampled data: *Journal of Petroleum Exploration and Production Technology*, **1**, 43–49, doi: [10.1007/s13202-011-0005-7](https://doi.org/10.1007/s13202-011-0005-7).
- Sava, P., and S. Fomel, 2003, Angle-domain common-image gathers by wavefield continuation methods: *Geophysics*, **68**, 1065–1074, doi: [10.1190/1.1581078](https://doi.org/10.1190/1.1581078).
- Sava, P., and S. Fomel, 2006, Time-shift imaging condition in seismic migration: *Geophysics*, **71**, no. 6, S209–S217, doi: [10.1190/1.2338824](https://doi.org/10.1190/1.2338824).
- Shang, X., M. V. Hoop, and R. D. Hilst, 2012, Beyond receiver functions: Passive source reverse time migration and inverse scattering of converted waves: *Geophysical Research Letters*, **39**, doi: [10.1029/2012GL052289](https://doi.org/10.1029/2012GL052289).
- Stewart, R., J. Gaiser, R. Brown, and D. Lawton, 2002, Converted-wave seismic exploration: *Methods: Geophysics*, **67**, 1348–1363, doi: [10.1190/1.1512781](https://doi.org/10.1190/1.1512781).
- Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: *Geophysics*, **49**, 1259–1266, doi: [10.1190/1.1441754](https://doi.org/10.1190/1.1441754).
- 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](https://doi.org/10.1190/1.3238367).
- Wapenaar, C. P. A., and G. C. Haime, 1990, Elastic extrapolation of primary seismic P- and S-waves: *Geophysical Prospecting*, **38**, 23–60, doi: [10.1111/j.1365-2478.1990.tb01833.x](https://doi.org/10.1111/j.1365-2478.1990.tb01833.x).
- Yan, J., and P. Sava, 2008, Isotropic angle-domain elastic reverse-time migration: *Geophysics*, **73**, no. 6, S229–S239, doi: [10.1190/1.2981241](https://doi.org/10.1190/1.2981241).
- Yan, J., and P. Sava, 2009, Elastic wave-mode separation for VTI media: *Geophysics*, **74**, no. 5, WB19–WB32, doi: [10.1190/1.3184014](https://doi.org/10.1190/1.3184014).
- Yan, J., and P. Sava, 2010, Analysis of converted-wave extended images for migration velocity analysis: 80th Annual International Meeting, SEG, Expanded Abstracts, 1666–1671.
- 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](https://doi.org/10.1190/1.3581360).
- Yang, D., A. Malcolm, and M. Fehler, 2014a, Using image warping for time-lapse image domain wavefield tomography: *Geophysics*, **79**, no. 3, WA141–WA151, doi: [10.1190/geo2013-0424.1](https://doi.org/10.1190/geo2013-0424.1).
- Yang, D., A. Malcolm, M. Fehler, and L. Huang, 2014b, Time-lapse walk-away vertical seismic profile monitoring for CO₂ injection at the SAC-ROC enhanced oil recovery field: A case study: *Geophysics*, **79**, no. 2, B51–B61, doi: [10.1190/geo2013-0274.1](https://doi.org/10.1190/geo2013-0274.1).
- Yang, T., and P. Sava, 2011, Wave-equation migration velocity analysis with time-shift imaging: *Geophysical Prospecting*, **59**, 635–650, doi: [10.1111/j.1365-2478.2011.00954.x](https://doi.org/10.1111/j.1365-2478.2011.00954.x).