GEOPHYSICS, VOL. 82, NO. 5 (SEPTEMBER-OCTOBER 2017); P. S391–S401, 9 FIGS.
10.1190/GEO2016-0318.1

Check for updates

Downloaded 10/27/17 to 138.67.129.100. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

Cascaded wavefield tomography and inversion using extended common-image-point gathers: A case study

Esteban Díaz[1] and Paul Sava[2]

ABSTRACT

Image-domain wavefield tomography (WT) exploits focusing characteristics of extended images for updating the velocity field. To make good use of this information, one must understand how such images behave if the migration velocity is accurate. This is not trivial because focusing depends on not only the model error, but also on the acquisition setup, the data bandwidth, and illumination variation caused by the overburden. We address this problem by constructing penalty functions based on the point spread functions of the imaging operator that characterize focusing in extended images. Moreover, instead of sampling the extended images at preset distances along the surface, we sample the image by constructing common-image-point gathers, which are more economical from a computational point of view and also measure the spatial and temporal focusing of the wavefield. Coupled with image residuals exploiting illumination-based penalty functions, we construct robust wavefield tomographic updates from sparse locations of the image in areas of poor or uneven illumination. Models obtained with this type of methodology are a good starting point for more sensitive, but less robust, waveform inversion methods. Starting with the velocity model obtained by illumination-based tomography (where important low-wavenumber information is added to the model), we perform an inversion using waveform tomography in the data domain. We describe an application of the method to synthetic examples and to a marine 2D data set. We complement the low-resolution WT resolution with a high-resolution data-domain inversion, which has the potential to add details to the reconstructed model. We describe the tomographic implementation with a synthetic data set and a marine field data example.

INTRODUCTION

An accurate velocity model is the main requirement for successful imaging. To be consistent with the typical band-limited seismic data, one ought to use wavefield-based tomographic velocity estimation. The wavefield-based approach avoids shortcomings inherent in ray-based methods, such as limited model sensibility (a ray travels through an infinitesimally narrow path inside the model) and instability around sharp boundaries in the velocity model.

Velocity analysis methods based on wavefield extrapolation are commonly referred to as wavefield tomography (WT) (Tarantola, 1984; Woodward, 1992; Pratt, 1999; Biondi and Symes, 2004; Sava and Biondi, 2004a, 2004b; Shen and Symes, 2008; Symes, 2008); such tomographic approaches can be formulated either in the image domain, in which one tries to improve image quality, or in the data domain, in which one seeks consistency between modeled and observed data.

Image-domain WT can be formulated in many ways. A common approach aims to improve the flatness of angle gathers, or equivalently one can improve the focusing of space-lag gathers (Sava and Biondi, 2004a; Shen and Calandra, 2005). Space-lag gathers (Rickett and Sava, 2002; Sava and Fomel, 2006) measure the spatial similarity between source and receiver wavefields. Hence, during tomography, one seeks to increase the similarity of the spatial correlation for a collection of seismic experiments (Shen and Symes, 2008; Shan and Wang, 2013; Weibull and Arntsen, 2013; Yang et al., 2013; Biondi and Almomin, 2014). The concept of focusing in extended images can be applied to space-lag gathers, to time-lag gathers, or to a combination of space-lag and time-lag gathers (Yang

Manuscript received by the Editor 15 June 2016; revised manuscript received 10 March 2017; published ahead of production 16 June 2017; published online 16 August 2017.
[1]Formerly Center for Wave Phenomena, Houston, Texas, USA; presently BP America Inc., Houston, Texas, USA. E-mail: esteban.diaz@bp.com.
[2]Center for Wave Phenomena, Colorado School of Mines, Houston, Texas, USA. E-mail: psava@mines.edu.
© 2017 Society of Exploration Geophysicists. All rights reserved.

Downloaded 10/27/17 to 138.67.129.100. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

and Sava, 2015). To define the tomographic objective, one must first define an operator whose purpose is to highlight the defocused events in the gather that one needs to focus trough iterative tomography. In general, this operator is defined by a weight proportional to the lag axis, as introduced by Shen and Symes (2008). In contrast, Yang et al. (2013) advocate for using a more complicated penalty function to account for illumination of the seismic data, which can improve the convergence rate of the problem. By incorporating illumination information into the penalty function, the tomography goal is more realistic because it has the objective of improving the focus given the bandwidth and acquisition parameterization of the data. In contrast, conventional penalties seek to obtain extended images whose focus approaches a delta function, which is an unrealistic goal given the limitations of seismic data.

Data-domain WT is generally formulated by improving the consistency between modeled and observed data. Originally, Tarantola (1984) introduces the data difference as a similarity estimate in the time domain. Alternatively, the problem can be solved in the frequency domain (Pratt, 1999). Contrary to the image-domain formulation, data-domain WT is highly nonlinear; i.e., the objective function has many local minima. To overcome the nonlinearity, a multiscale separation approach is needed (Bunks et al., 1995). Within each scale (frequency or frequency band), the problem can be more linear if the initial model is closer to the one corresponding to the global minimum. Another loop of multiscale can be added by introducing time damping, a method commonly referred to as Laplace-Fourier waveform inversion (Sirgue and Pratt, 2004; Shin and Ho Cha, 2009). The purpose of the time damping is to fit earlier arrivals first and then to fit later arrivals progressively.

Data-domain and image-domain tomographic methods share similarities in their processes: Both use the same extrapolation engine (the two-way wave equation) and build the gradient of the corresponding objective function through the adjoint state method (ASM) framework (Tarantola, 1984; Plessix, 2006; Symes, 2008). In this paper, we interleave a time-domain application for image-domain tomography using common-image-point (CIP) gathers under illumination constraints with a frequency-domain application of data-domain tomography. We further analyze the tomographic term in the image-domain inversion, using different adjustments to the objective function, such as illumination and maximization to include the zero-lag gather contribution directly into the inversion (Zhang and Shan, 2013). We complement the low WT resolution with a high-resolution data-domain inversion that has the potential to add resolution to the reconstructed model. We illustrate the tomographic implementation with a synthetic data set and with a marine field data example.

IMAGE-DOMAIN TOMOGRAPHY

Extended-image gathers (Rickett and Sava, 2002; Sava and Fomel, 2006; Sava and Vasconcelos, 2011) highlight the spatial and temporal consistencies between wavefields by exploring the focusing information in the image domain characterized with common-image gathers. The moveout in the gather is sensitive to velocity perturbations and hence can be optimized. A general extended image can be defined as follows:

$$R(\mathbf{x},\boldsymbol{\lambda},\tau)=\sum_{e}\sum_{t}u_s(e,\mathbf{x}-\boldsymbol{\lambda},t-\tau)u_r(e,\mathbf{x}+\boldsymbol{\lambda},t+\tau), \quad (1)$$

where $\boldsymbol{\lambda}$ is the space-lag vector, τ is the time-lag vector, $\mathbf{x}$ is the image location, e is the experiment index, u_s is the source wavefield, and u_r is the receiver wavefield. Given the increase in dimensions for extensions in space and time, one can take advantage of a sparse sampling of the extended images in a subset of image locations $\mathbf{x}_c$ instead of the full image space $\mathbf{x}$. One efficient way to decide the locations $\mathbf{x}_c$ is by placing the observation points at the reflector locations (Cullison and Sava, 2011). Note that the process for computing the extended image is linear with respect to one of the wavefields. Hence, one can define the extended image in a matrix-vector form as

$$\mathbf{R}=\mathbf{I}_s\mathbf{u}_r=\mathbf{I}_r\mathbf{u}_s, \quad (2)$$

where $\mathbf{R}$ is the vector representation of the extended image, and $\mathbf{I}_s$ and $\mathbf{I}_r$ are the imaging operators for the source and receiver wavefields, respectively.

Even though the commutation of the forward mappings (equation 2) produces the same result ($\mathbf{R}$), the adjoint mapping satisfies different equations. The operator $\mathbf{I}_s^{\mathsf{T}}$ implements

$$\tilde{u}_r(e,\mathbf{x},t)=\sum_{\tau,\boldsymbol{\lambda}}R(\mathbf{x}-\boldsymbol{\lambda},\boldsymbol{\lambda},\tau)u_s(e,\mathbf{x}-2\boldsymbol{\lambda},t-2\tau), \quad (3)$$

whereas $\mathbf{I}_r^{\mathsf{T}}$ satisfies

$$\tilde{u}_s(e,\mathbf{x},t)=\sum_{\tau,\boldsymbol{\lambda}}R(\mathbf{x}+\boldsymbol{\lambda},\boldsymbol{\lambda},\tau)u_r(e,\mathbf{x}+2\boldsymbol{\lambda},t+2\tau). \quad (4)$$

The symbol $+=$ implies that the adjoint wavefield is accumulated in the output after summing all possible lags $\boldsymbol{\lambda},\tau$. The output wavefields $\tilde{u}_r$ and $\tilde{u}_s$ are obtained after the application of the adjoint imaging operators to the input wavefields u_s and u_r, respectively.

The source and receiver wavefields follow the wave equation, which in this case is the acoustic wave equation:

$$m(\mathbf{x})\frac{\partial^2 u(e,\mathbf{x},t)}{\partial t^2}-\nabla^2 u(e,\mathbf{x},t)=f(e,\mathbf{x},t), \quad (5)$$

where $m=s^2(\mathbf{x})$ is the squared slowness. The source wavefield $u_s(e,\mathbf{x},t)$ is computed by forward propagation of the source function $f_s(e,\mathbf{x},t)$ for experiment index e, whereas the receiver wavefield $u_r(e,\mathbf{x},t)$ is created by backward (adjoint) propagation of the recorded data $f_r(e,\mathbf{x},t)$ as summarized by the following system of equations:

$$\begin{bmatrix}\mathcal{L}(m,t) & 0\\ 0 & \mathcal{L}^{\mathsf{T}}(m,t)\end{bmatrix}\begin{bmatrix}u_s\\ u_r\end{bmatrix}=\begin{bmatrix}f_s\\ f_r\end{bmatrix}, \quad (6)$$

where $\mathcal{L}$ and $\mathcal{L}^{\mathsf{T}}$ are the matrix representation of the forward and adjoint (reverse) propagators, respectively.

A model with good kinematic agreement with the true background velocity exhibits a maximum correlation between source and receiver wavefields around $(\boldsymbol{\lambda},\tau)=(0,0)$. For an inaccurate velocity, extended gathers exhibit moveout indicative of model errors (Yang and Sava, 2010). Hence, the energy in the gathers can be used to formulate an inversion process to update the velocity model. A typical objective function (Shen and Symes, 2008; Weibull and Arntsen, 2013; Yang and Sava, 2015) based on extended images is

$$J_1(m)=\|P(\boldsymbol{\lambda},\tau)R(\mathbf{x},\boldsymbol{\lambda},\tau)\|^2, \quad (7)$$

where P is a penalty operator whose purpose is to highlight the events in the image that show velocity inaccuracies. The penalty operator is sometimes defined using the lags:

Downloaded 10/27/17 to 138.67.129.100. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

$$P(\boldsymbol{\lambda},\tau)=\sqrt{\boldsymbol{\lambda}^2+(v\tau)^2}, \tag{8}$$

where v is the local velocity (Yang and Sava, 2015). This penalty function is highly idealized because it produces a zero residual if the extended image is a perfect spike (i.e., $R(\mathbf{x},\boldsymbol{\lambda},\tau)=\delta(\mathbf{x},\boldsymbol{\lambda},\tau)$). To relax such requirements, Yang et al. (2013) propose a penalty function based on illumination that depends on the acquisition setup, data bandwidth, and the velocity model itself. To obtain the penalty function, they perform a demigration on the conventional image to obtain data as if they were propagated with the current velocity model; then, they remigrate these data to obtain extended images that are perfectly focused given the acquisition conditions:

$$R_e(\mathbf{x},\boldsymbol{\lambda})=\mathbf{M}\mathbf{M}_{\mathbf{z}}^{\top}R(\mathbf{x}). \tag{9}$$

In their notation, the operator $\mathbf{M}_{\mathbf{z}}^{\top}$ performs a demigration on the conventional image $R(\mathbf{x})$ to obtain synthetic data without kinematic errors, whereas $\mathbf{M}$ performs extended imaging. Once the reference image R_e is obtained, the penalty is constructed by measuring the inverse of the energy of the image:

$$P_I(\mathbf{x},\boldsymbol{\lambda})=\frac{1}{|R_e(\mathbf{x},\boldsymbol{\lambda})|+\varepsilon^2}. \tag{10}$$

The illumination-based penalty operator is minimum around the reflector position and increases its value as it moves away from the reflector. Because the data are obtained by demigration of the conventional image, its spectrum is narrower than the original data (because it contains twice the effect of the source wavelet and the data spectrum). Hence, the illumination effect due to the bandwidth of the data might be exaggerated. Here, instead of computing extended image gathers at continuous locations, we sample the extended image at the reflector position in a sparse fashion. We use this sparse sampling to our advantage to compute. The sparseness allows us to obtain the point spread functions (PSF) of the extended imaging operator directly at the reflector positions (Valenciano et al., 2009; Fletcher et al., 2012).

We obtain the effect of the illumination (blurring operator), by cascading a demigration/migration process applied to spikes placed at the extended image locations $\mathbf{x}_c$:

$$\tilde{R}(\mathbf{x}_c,\boldsymbol{\lambda},\tau)=\mathbf{M}\mathbf{M}^{\top}\sum_i\delta(\mathbf{x}-\mathbf{x}_c^i,\boldsymbol{\lambda},\tau), \tag{11}$$

with $\mathbf{M}^{\top}$ being an extended demigration operator and $\mathbf{M}$ the extended imaging process. The index i corresponds to the CIP location $\mathbf{x}_i$. The synthetic image $\tilde{R}$ contains the action of the source function, the acquisition, and the model at the image locations $\mathbf{x}_c$.

Instead of minimizing the energy outside zero lag, one can maximize the energy around zero lag. Conceptually, both approaches are reciprocal; however, the second approach (Zhang and Shan, 2013) has the advantage of including the zero-lag information in the residual. Thus, it avoids calculating the contribution of the image power terms used by Shen and Symes (2008) and Weibull and Arntsen (2013). To keep the tomography as a minimization problem, we use the following objective function:

$$J_2(m)=-\|H(\boldsymbol{\lambda},\tau)R(\mathbf{x},\boldsymbol{\lambda},\tau)\|^2, \tag{12}$$

where H is the operator that highlights the information at around zero lag. Our illumination-based highlighting operator is computed as

$$\tilde{H}(\mathbf{x}_c,\boldsymbol{\lambda},\tau)=\langle\mathrm{Env}(\tilde{R}(\mathbf{x}_c,\boldsymbol{\lambda},\tau))\rangle, \tag{13}$$

where Env is an envelope function that captures the energy of the synthetic image $\tilde{R}$ and $\langle\rangle$ denotes the smoothing in the $\boldsymbol{\lambda}$-τ plane. The smoothing is done with the objective of preserving a quasi-monotonic behavior of the highlighting operator H. Once the operator H and objective function J_2 are defined, we proceed to update the model iteratively using the gradient ∇J_2. We show how to obtain the gradient in Appendix A.

Computational aspects

The method proposed here is on the same line of those introduced by Shen and Symes (2008), Weibull and Arntsen (2013), Yang et al. (2013), and Shan and Wang (2013). It takes advantage of extended images to obtain a measurement of the focusing inside the medium. The key difference is that we use CIP gathers (Yang et al., 2013) instead of space-lag gathers. A common problem with space-lag gathers is its computational and storage cost. One way to alleviate the cost of obtaining the gathers is to extract them with a coarse x, y grid. As explained in Yang et al. (2013), this approach can be inefficient because it subsamples the x- and y-axes and oversamples the z-axis. Instead, CIP gathers sample the image space more efficiently because the coordinates are selected at the reflector position; thus, it avoids sampling of areas of the image that are not of interest (i.e., the water layer), whereas it focuses the sampling in areas of the image with dense reflectivity. The image sampling allows us to also extend the image along the time axis, which would be prohibitively costly for common-image gathers. Hence, the cost of computing the extended images is proportional to the selected number of points, which in general is much smaller than the number of points used in common-image gathers.

In terms of tomography, the conventional method requires the reconstruction of four wavefields for each experiment. Here, we also need to consider the cost of obtaining the illumination-based operators. The operator H depends on the image locations, which must be recomputed at every iteration. Because we need to demigrate and migrate the impulse responses at the image locations, we need to reconstruct two more wavefields, which means that at each iteration, one needs to retrieve a total of six wavefields per experiment.

The illumination operator H could also be computed only at the beginning for the initial model. Then, when the image locations are updated, one could interpolate the operator to the new locations. One disadvantage of this approach is that the illumination would not depend on the velocity model, which means that the operator should be recomputed every few iterations. In our implementation, we compute the operator H at each iteration.

DATA-DOMAIN WT

The construction of the tomography problem in the data domain amounts to measuring the error (or residual) at the receiver locations. For data-domain WT, one normally uses the data difference for the residual:

Downloaded 10/27/17 to 138.67.129.100. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

$$J_3 = \frac{1}{2}\|u_s(\mathbf{x}_r,\Omega) - f_r(\mathbf{x}_r,\Omega)\|^2 = \frac{1}{2}\|\Delta d\|^2, \tag{14}$$

where $\mathbf{x}_r$ are the receiver locations and Ω is the complex-valued frequency of which purpose we will explain later. Note that $f_r(\mathbf{x}_r,\Omega) = u_r(\mathbf{x}_r,\Omega)$. Because for building the residual, we only need to forward propagate the source function:

$$\mathcal{L}(m,\Omega)u_s = f_s(\mathbf{x}_s,\Omega), \tag{15}$$

and we have to compute one adjoint wavefield $a_s(\mathbf{x},\Omega)$. For data-domain WT, computing a_s involves backpropagating the data residual:

$$\mathcal{L}^\top(m,\Omega)a_s = (\Delta d)^*. \tag{16}$$

Here, $(\Delta d)^*$ is the complex conjugate residual and $\mathcal{L}(m,\Omega)$ is the acoustic wave equation in the frequency domain, defined as follows:

$$\mathcal{L}(m) = -\nabla\cdot\frac{1}{\rho(\mathbf{x})}\nabla - \frac{m(\mathbf{x})}{\rho(\mathbf{x})}\Omega^2, \tag{17}$$

where $\rho(\mathbf{x})$ is the density of the medium. In this paper, we do not invert for $\rho(\mathbf{x})$; instead, we parameterize it as a function of the velocity following Gardner et al. (1974).

Once we obtain $u_s(\mathbf{x},\Omega)$ and $a_s(\mathbf{x},\Omega)$, we can proceed to compute the gradient:

$$\nabla J_3(\mathbf{x}) = \mathbb{R}\left\{\sum_{\Omega,e}(\Omega^2 u_s(e,\mathbf{x},\Omega)a_s^*(e,\mathbf{x},\Omega))\right\}, \tag{18}$$

where * denotes the complex conjugate and $\mathbb{R}\{\}$ denotes the real part.

The data-domain WT objective function, equation 14, is highly nonlinear if we operate with signals in the normal frequency band. Hence, to increase the chances of convergence to the global minimum, it is customary to implement data-domain WT in a multiscale fashion. Bunks et al. (1995) propose to first invert lower frequencies and then move gradually to higher frequencies. The idea is that, within each scale, the problem looks more linear than when one inverts all the bandwidth at once.

An additional outer loop in the inversion is the time damping, which leads to the so-called Laplace-Fourier domain FWI (Sirgue and Pratt, 2004; Shin and Ho Cha, 2009). The purpose of this outer loop is to first fit earlier arrivals and then fit later arrivals. By first fitting early arrivals (with shorter traveltime), one reduces the risk of large phase differences between observed and modeled data, which can cause cycle skipping. Once the traveltime differences are solved for early arrivals, we can progressively increase ς. Introducing time damping requires the following transformation: $\Omega = \omega + i/\varsigma$, with ς being the time damping (Kamei et al., 2013). Thus, this transformation turns the real-valued angular frequency ω into a complex-valued angular frequency Ω. In order for the observed data to be consistent with the damped modeled data, one must also scale it as $f_r(\mathbf{x}_r,t) = d_{\text{obs}}(\mathbf{x}_r,t)e^{-t/\varsigma}$ before the transformation to the frequency domain.

The low frequencies of the data are sensitive to the long wavelength (smooth) components of the earth model. However, if the data do not have such frequencies, data-domain WT is unable to update such components. In contrast, focusing on extended images is mostly sensitive to the smooth components of the model. By implementing a joint workflow using image-domain WT for updating the smooth components of the model and later using data-domain WT for the high-resolution features of the model, we can obtain a more complete spectrum in the model. The first pass of using image-domain WT has the ability to mitigate the cycle-skipping problems in data-domain WT.

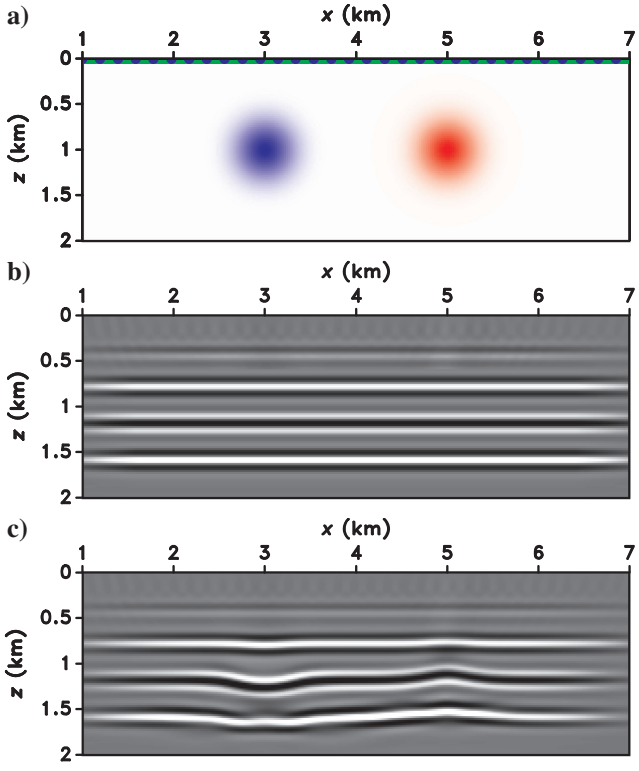


Figure 1. Synthetic examples: (a) true velocity model, (b) corresponding image, and (c) image after starting velocity model (constant background).

SYNTHETIC DATA EXAMPLE

In this section, we use a simple synthetic example to demonstrate the image-domain tomography process. Figure 1a shows a background velocity model of 3 km/s with two Gaussian anomalies of −0.5 and +0.5 km/s from left to right, respectively. The four reflectors, shown in the image computed with the true velocity model (Figure 1b), are produced by contrasts in the density model. The starting model for tomography is set constant at 3 km/s. The purpose of the tomographic step is to recover the effect of the Gaussian anomalies in the model. Figure 1c depicts the seismic image computed from the starting model.

We perform two inversions using CIP gathers: (1) with the conventional differential semblance optimization (DSO) penalty function (equation 8) and (2) with the illumination-based highlighting operator H (from equation 13). Figure 2a shows the image corresponding to the inverted model in Figure 2b with the conventional DSO penalty function. Figure 2c depicts the image corresponding to the inverted model in Figure 2d using the illumination-based maximization approach. Both inverted models introduce velocity anomalies with

the expected sign. However, the model retrieved with the illumination-based framework better reconstructs the anomalies compared with the model from the DSO penalty functions. The reverse time migration (RTM) image corresponding to the illumination-based model (Figure 2c) is flatter than the corresponding DSO-based penalty model shown in Figure 2a.

To understand the fundamental difference between the two penalty methods, one can take a look at an ideal CIP image obtained from the correct model shown in Figure 3a and compare it with a CIP image to a −10% velocity anomaly (Figure 3b). The conventional DSO-based penalty function depicted in Figure 3c is ideal for an unrealistic CIP image described by a delta function:

$$R(\mathbf{x}, \boldsymbol{\lambda}, \tau) = \delta(\boldsymbol{\lambda})\delta(\tau). \tag{19}$$

In contrast, Figure 3d is computed through a migration/demigration process that mimics the ideal focusing for the given acquisition parameters. One can observe the consistency between the illumination-based penalty function, Figure 3d, and the ideal CIP image, Figure 3a.

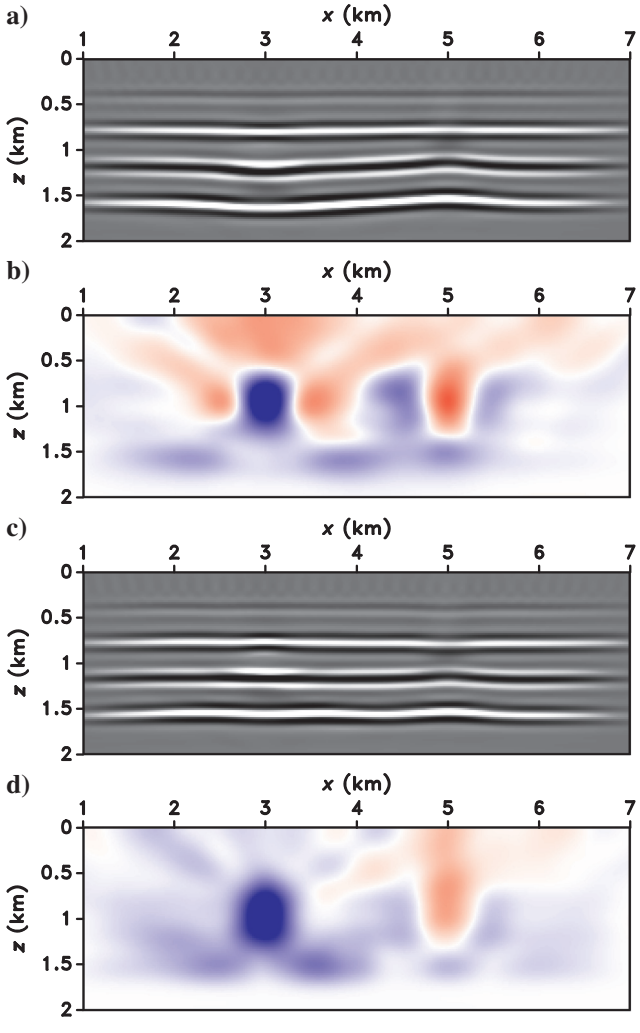


Figure 2. Synthetic example: (a) updated image, (b) model after DSO-based tomography, (c) updated image, and (d) model after illumination-based tomography. Note how the structure is better recovered with the illumination-based approach.

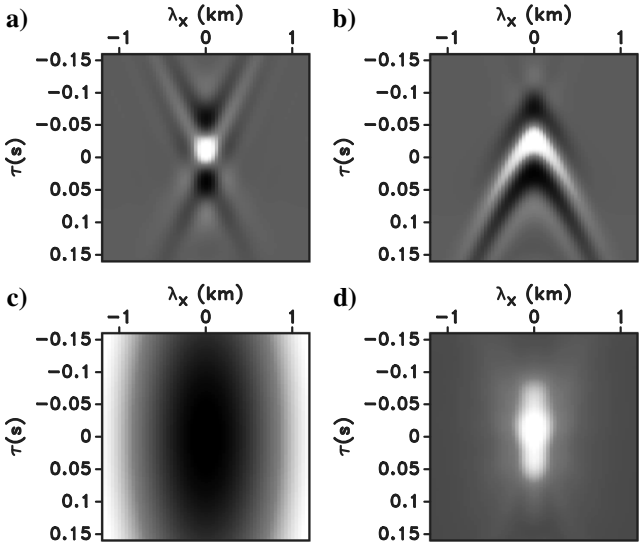


Figure 3. (a) An example of a CIP gather from the correct model, (b) CIP gather with 10% velocity error, (c) conventional penalty function, and (d) illumination-based operator H. Note the consistency between the penalty functions in (d) and the correct gather in (a).

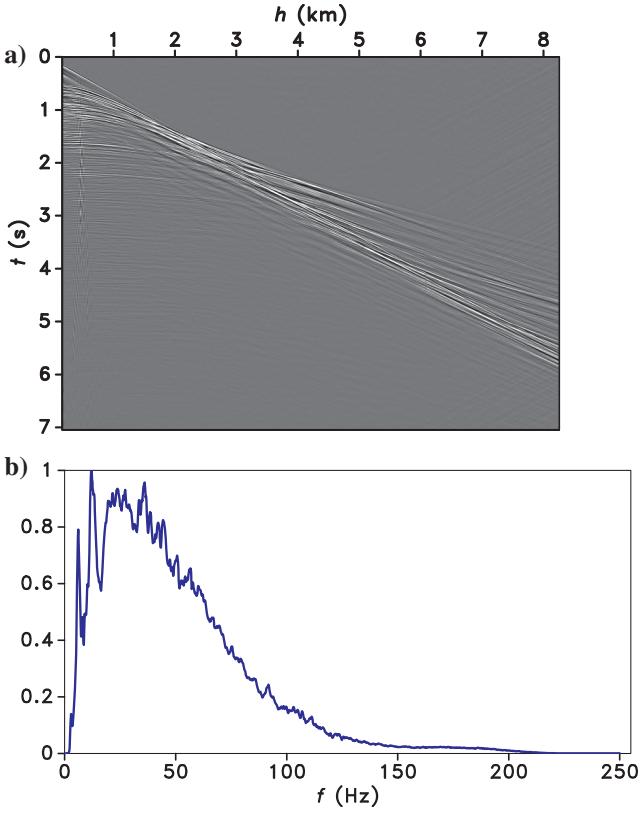


Figure 4. (a) Shot gather at $x = 14$ km and (b) its average amplitude spectrum. Note the low-frequency response of the data with a good signal-to-noise ratio starting at approximately $f = 2.5$ Hz.

Downloaded 10/27/17 to 138.67.129.100. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

Downloaded 10/27/17 to 138.67.129.100. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

FIELD DATA EXAMPLE

In this section, we apply the cascaded workflow of image-domain WT followed by data-domain WT. We use image-domain WT to correct for most kinematics errors and then data-domain WT to refine the model and add details. We obtain three models: (1) the initial model, (2) the model obtained by image-domain WT, and (3) the data-domain WT model.

The data set is a marine 2D line acquired with a variable depth cable. The towed streamer contains increasing depths as a function of offset, which enhances the frequency content of the data by producing a mixed-notch response. Hence, the increased cable depths improve the low-frequency content at intermediate and far offsets, which can be very helpful for data-domain WT. The cable contains offsets ranging from 0.169 to 8.256 km with a receiver spacing of 0.0125 km. Figure 4a shows a shot gather from $x = 14$ km, and

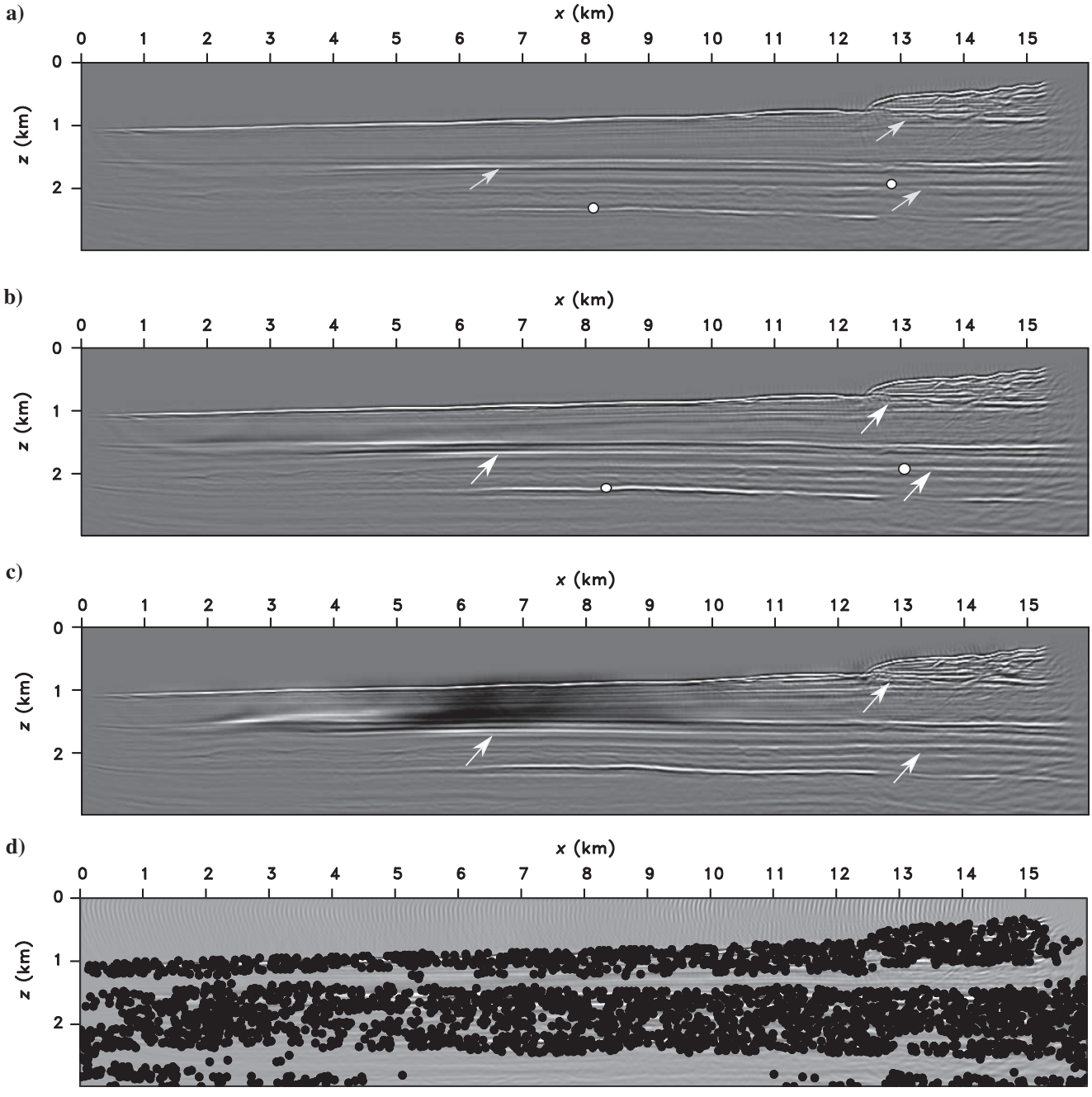


Figure 5. The RTM images from the (a) initial model, (b) illumination-based tomographic model, and (c) data-domain inversion model. Note how the focusing (image strength) increases from (a) to (b). Also, one can see the shallow part of the image improved due to the data-domain tomography at approximately $z = 1.1$ km. (d) An example of the CIP locations selected during tomography. The overlaid dots on (a and b) show the locations used for the selected CIP locations in the next figure. The arrows indicate locations of sequential improvement after each step of the process.

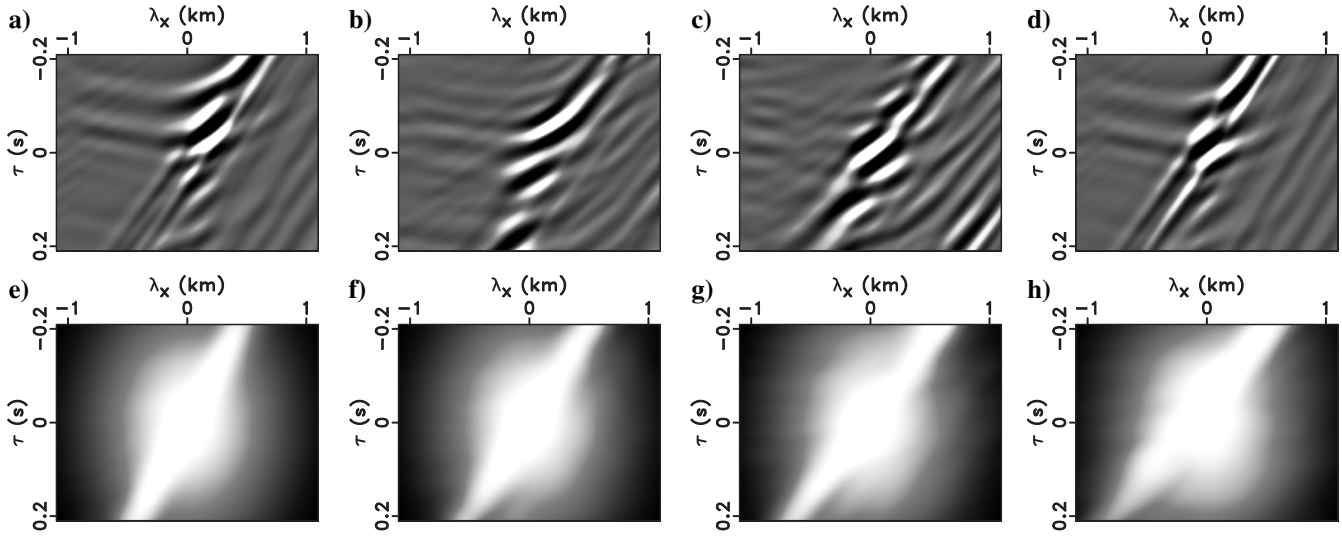


Figure 6. Selected CIP gathers and illumination operator H for the locations highlighted in Figure 5a. Panels (a and b) depict the CIP gathers obtained from the initial model and (c and d) show the CIP gathers after tomography. Note how the peak of the correlation is shifted toward zero lag after the velocity update. The bottom panels show the operator H, which contains small variations after the tomographic update. Also, the peak of the operator follows the one-sided illumination pattern of the CIP gathers introduced by the single-sided towed streamer acquisition.

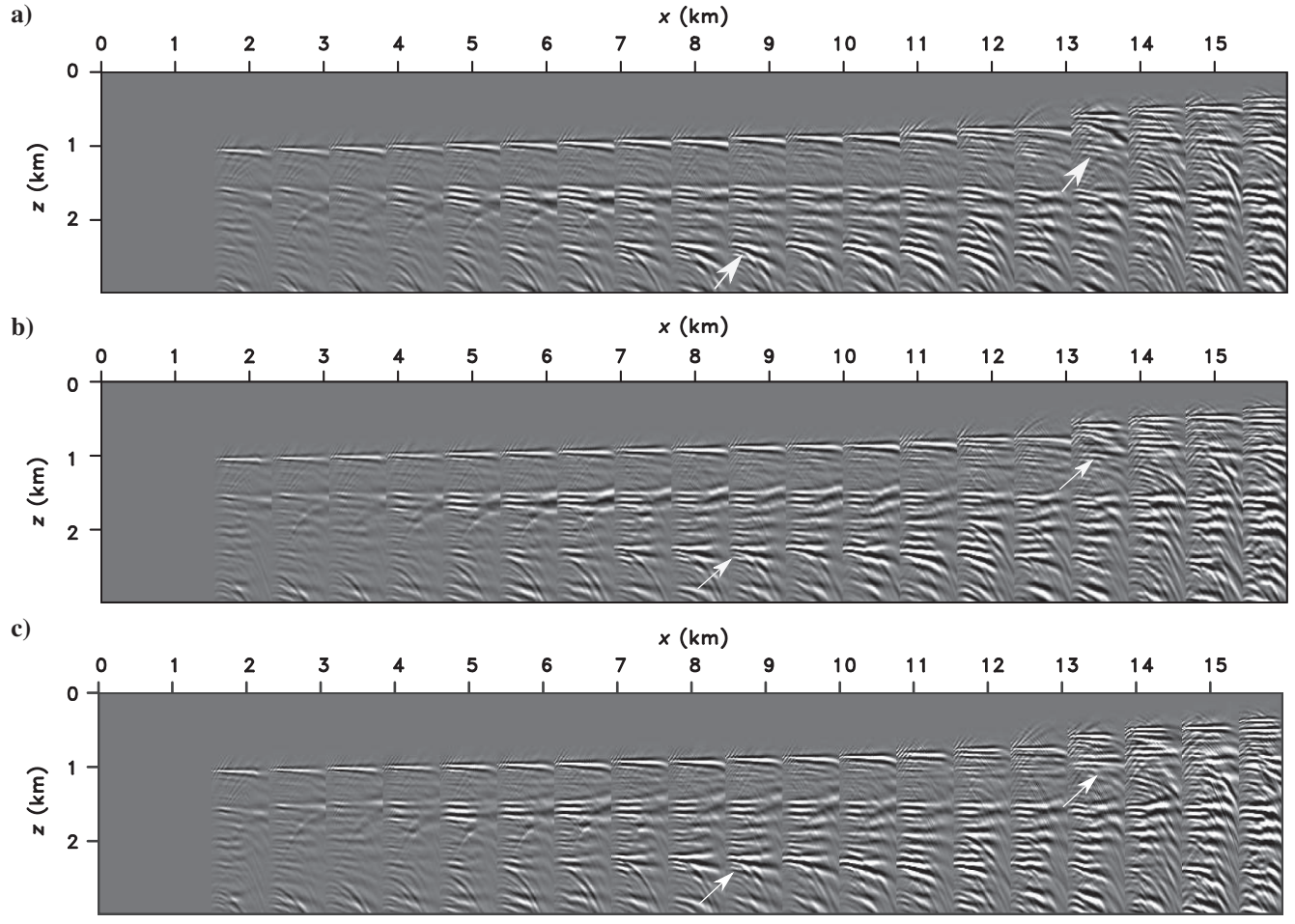


Figure 7. Angle gathers retrieved at sparse surface locations from the (a) initial model, (b) illumination-based tomographic model, and (c) data-domain inversion model. Observe how the flatness of the gather increases from (a) to (b). The improvement in (c) is especially evident at the shallow part $z < 1.2$ km between $x = 10$ and 15 km. The arrows indicate the locations of sequential improvement after each step of the process.

Downloaded 10/27/17 to 138.67.129.100. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

Figure 4b depicts the average amplitude spectrum for the same gather. For both inversions, we use 250 shot gathers with an interval of 0.0625 km. Figure 5a–5c shows the RTM images corresponding to the initial image-domain tomography and data-domain tomography models, respectively. Note how the image improves in each subsequent step of the inversion. Figure 5d shows an example of CIP locations at a given tomographic iteration. The dots in Figure 5a display the locations for the selected CIP gathers shown in Figure 6. Figure 7a–7c shows angle gathers corresponding each velocity model. The gathers can be seen as an independent tool for confirming the improvement in image focusing.

We build the initial model, Figure 8a, by performing time-domain normal moveout analysis followed by smoothing, root-mean-square (stacking) conversion to interval velocity (Dix, 1955), and time-to-depth conversion. The velocity grid used for tomography contains 1100 points along the x-axis spaced at $dx = 0.030$ km and 100 points along the z-axis spaced at $dz = 0.030$. Figure 5a shows the RTM image produced by the model in Figure 8a; one can see poor

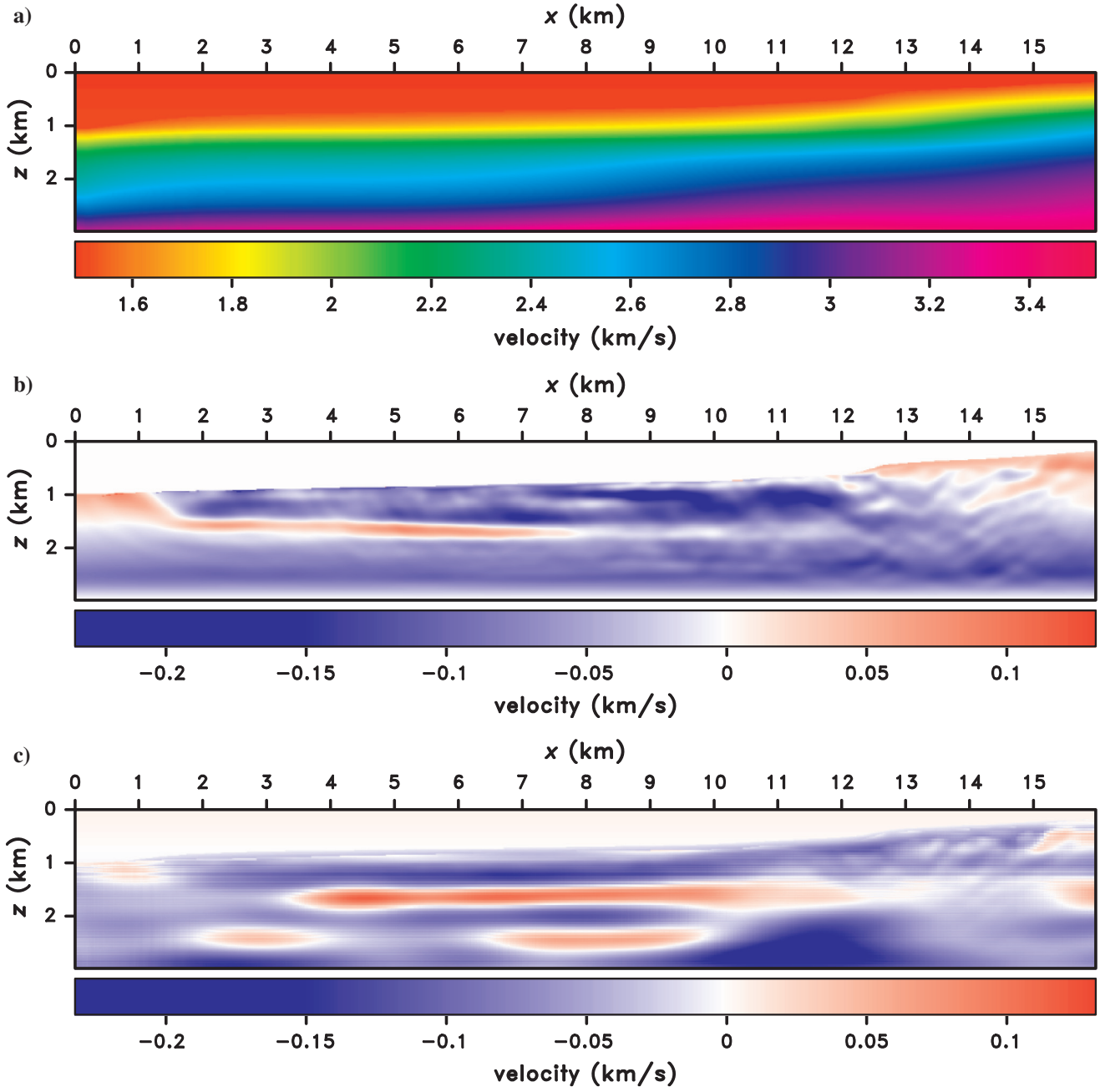


Figure 8. Velocity profiles from the (a) initial model, (b) illumination-based tomographic model, and (c) data-domain inversion model. Panels (b and c) show the update with respect to the initial model in (a). Note how the velocity decreases from (a) to (b) and how further details are added in the model in (c).

Downloaded 10/27/17 to 138.67.129.100. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

Downloaded 10/27/17 to 138.67.129.100. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

focusing through the section and a relatively flat geology below $z = 1.5$ km. Figure 7a shows angle gathers extracted at sparse locations in the model. Note that we do not use the angle gathers for inversion; instead, we use the gathers as an independent quality control tool. The transformation from space-lag gathers $R(\mathbf{x}, \boldsymbol{\lambda})$ to angle domain $R(\mathbf{x}, \theta)$ follows the method of Sava and Fomel (2003). The angles vary from 0° to 45° for all the gathers shown in this paper. The moveout in the gathers confirms that the velocity is too fast. Some of the events in the gathers, however, correspond to migrated surface-related multiples, and their moveout is not indicative of velocity error.

Figure 8b shows the update added to the initial model by the CIP tomography approach with illumination constraints. The update is mostly negative, which means that the tomography process constructed an overall slower model; however, a high-velocity layer is introduced into the velocity model at approximately $z = 1.5$ km in the left part of the model. The CIP gathers contain a maximum space and time lag of $\boldsymbol{\lambda} = 1.5$ km and $\tau = 1$ s, respectively. We fix the number of CIP observation points to 10,000 per iteration and perform 11 iterations. Figure 5d shows a subset of the selected CIP locations for the initial image. The automatic picker (Cullison and Sava, 2011) works by placing CIP locations at areas of the image with high coherency, planarity, and amplitude strength. One can also restrict the locations in such a way that the picks are not clustered together. The idea of this tomographic step is to correct for the bulk of the kinematic errors in the model. As an example of the focusing of CIP gathers, we extract two panels at the highlighted locations in Figure 5a for the initial model and in Figure 5b for the updated model. Figure 6a and 6b shows how the CIP gathers for the initial model are poorly focused because the maximum of the correlation deviates from zero lag. Note how the gathers are illuminated from one side because of the towed streamer acquisition. After updating the model, the CIP gathers in Figure 6c and 6d show an improved focusing in which now the energy is concentrated close to zero lag. The illumination operators in Figure 6e and 6f for the initial model and in Figure 6g and 6h for the updated model show correlation with the CIP gathers. The maximum of the operator is at zero lag and is spread along the line with the slope given by the local velocity. Figure 5b shows the RTM image corresponding to the updated model, in which the focusing improves significantly throughout the section. This observation is confirmed in Figure 7b, in which the gathers are flatter along the section. To construct the illumination-based highlighting operators, we perform a demigration/migration step at each iteration. Assuming that the illumination does not vary significantly within iterations, one could update the operators every few iterations, which would help to diminish the overall cost of the tomography.

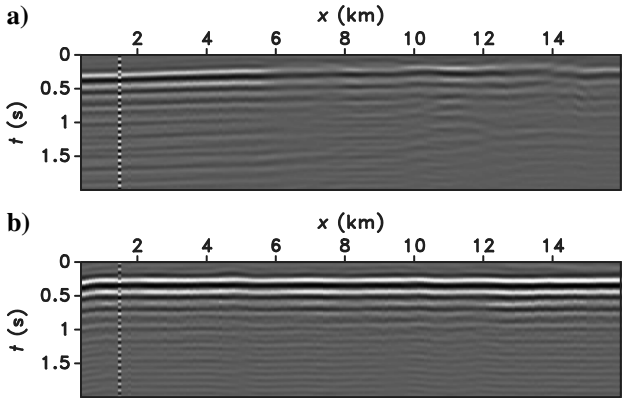


Figure 9. Estimated source signatures as a function of shot coordinate: (a) before data-domain tomography and (b) after data-domain tomography. After data-domain tomography, the source signatures the lateral consistency improves. For inversion, we use the average of the individual signatures.

Finally, we update the image-domain wavefield tomographic model with the data-domain wavefield tomographic approach. Figure 8c depicts the difference with the initial reference model, which changes considerably in the interval $x = 6$–12 km. The velocity slows down in the shallow area, and also a higher velocity layer is introduced at approximately $z = 1.5$ km. Figures 5c and 7c are the corresponding RTM image and angle gathers, respectively. The velocity contrast introduces enough scattering in the model to produce RTM artifacts that are due to the synchronization of source and receiver wavefields at the imaging condition. Note how further details are added into the velocity model, which changes the structure of the image and flattens the reflector at $z = 2.2$ km. For data-domain WT, we use seven frequency blocks with five frequencies each. The center frequency for each block ranges from $f = 2.6$–8.9 Hz. For the time damping constant, we use $\varsigma = 1.6$ s, it is important to highlight that the time damping is carried from $t = 0$ s. The first step in data-domain WT involves estimating the source function $f_s(\Omega)$; later, we compare the inverted source functions for each model. Figure 9a and 9b shows the estimated source signatures as a function of shot position before and after data-domain tomography, respectively. One can observe how the consistency of the source signature considerably improves after the model update.

CONCLUSION

The combination between image-domain and data-domain WT seeks to exploit the features of each method. The image-domain WT methods are sensitive to the smooth components of the model due to the definition of the inverse problem. Once we obtain a smooth model that improves focusing in the extended images, we can proceed to further refine the model using data-domain WT. We demonstrate the cascaded workflow using a real 2D marine data set. Our image-domain WT is based on illumination constraints imposed in the inversion through the highlighting operator. This operator allows one to emphasize aspects of the gather that one wants to improve. The criteria are based on what the gather should look like if the velocity model were correct. This knowledge is estimated by a demigration/migration process at each iteration. The illumination-based criterion is more realistic than the conventional DSO-based penalty functions because it incorporates information acquisition, data bandwidth, and velocity information. Our tests with the 2D data set show how the velocity model improves with the illumination-based tomography followed by the data-domain tomography. The second adds details in the shallower part of the model, thus increasing the consistency between modeled and observed early arrivals (not constrained by the image-domain tomography).

ACKNOWLEDGMENTS

We thank the sponsor companies of the Center for Wave Phenomena. The seismic data shown in this paper are proprietary to and provided courtesy of CGG. We thank B. VerWest for the support

with the data set. E. Díaz would like to thank G. Pratt for his feedback on the data-domain tomography strategies and for kindly providing his software, FULLWV. Also, we would like to thank the following people for constructive feedback: associate editor A. Cheng, reviewers F. Liu and L. Guasch, and an anonymous reviewer. The reproducible numeric examples in this paper use the Madagascar open-source software package (Fomel et al., 2013) freely available from http://www.ahay.org.

APPENDIX A

TOMOGRAPHY SETUP

We are interested in the gradient of equation 12 to update the background velocity model. The inverse problem is set up with constraints on the wave equation:

$$\begin{aligned} &\underset{m}{\text{minimize}} \quad J_2(m) \\ &\text{subject to} \quad \mathcal{L}u_s = f_s, \quad \mathcal{L}^\top u_r = f_r. \end{aligned} \tag{A-1}$$

To compute the gradient, we follow the methodology described by Plessix (2006) for objective functions constrained by partial differential equations. First, we define the augmented functional

$$\begin{aligned} A(m, u_s, u_r, a_s, a_r) = &-J_2(m) + \langle a_s, \mathcal{L}u_s - f_s\rangle \\ &+ \langle a_r, \mathcal{L}^\top u_r - f_r\rangle, \end{aligned} \tag{A-2}$$

where $\langle .,. \rangle$; defines a dot product between vectors in the $(\mathbf{x}, t)$ space, a_s is the adjoint source wavefield, and a_r is the adjoint receiver wavefield. The augmented functional depends on the model m, the forward state variables u_s and u_r, and the adjoint state variables a_s and a_r. From the first optimality conditions, $\partial A/\partial a_s = 0$ and $\partial A/\partial a_r = 0$, we obtain the system in equation 6.

To find the adjoint source wavefield a_s, we rewrite equation A-2 as

$$\begin{aligned} A(m, u_s, u_r, a_s, a_r) = &-\langle HI_r u_s, HI_r u_s\rangle \\ &+ \langle a_s, \mathcal{L}u_s - f_s\rangle + \langle a_r, \mathcal{L}^\top u_r - f_r\rangle. \end{aligned} \tag{A-3}$$

By setting $\partial A/\partial u_s = 0$, we obtain the adjoint state variable a_s:

$$\mathcal{L}^\top a_s = I_r^\top H^\top H I_r u_s = I_r^\top H^\top H R. \tag{A-4}$$

The operator H is diagonal and real, which means $H^\top = H$. This equation tells us to use the receiver wavefield u_r as an imaging operator together with the image residual $H^\top H R$. The wavefield a_s travels from the image toward the source.

To find the adjoint receiver wavefield a_r, we redefine equation A-2 to the equivalent form:

$$\begin{aligned} A(m, u_s, u_r, a_s, a_r) = &-\langle HI_s u_r, HI_s u_r\rangle \\ &+ \langle a_s, \mathcal{L}u_s - f_s\rangle + \langle a_r, \mathcal{L}^\top u_r - f_r\rangle. \end{aligned} \tag{A-5}$$

By setting $\partial A/\partial u_r = 0$, we obtain the adjoint variable a_r:

$$\mathcal{L}a_r = I_s^\top H^\top H R. \tag{A-6}$$

From the last optimality condition, we obtain the actual gradient

$$\frac{\partial J_2}{\partial m} = -\sum_{t,e} \frac{\partial^2 u_s}{\partial t^2} a_s + \frac{\partial^2 u_r}{\partial t^2} a_r. \tag{A-7}$$

REFERENCES

Biondi, B., and A. Almomin, 2014, Simultaneous inversion of full data bandwidth by tomographic full-waveform inversion: Geophysics, **79**, no. 3, WA129–WA140, doi: 10.1190/geo2013-0340.1.

Biondi, B., and W. Symes, 2004, Angle-domain common-image gathers for migration velocity analysis by wavefield-continuation imaging: Geophysics, **69**, 1283–1298, doi: 10.1190/1.1801945.

Bunks, C., F. Saleck, S. Zaleski, and G. Chavent, 1995, Multiscale seismic waveform inversion: Geophysics, **60**, 1457–1473, doi: 10.1190/1.1443880.

Cullison, T., and P. Sava, 2011, An image-guided method for automatically picking common-image-point gathers: 81st Annual International Meeting, SEG, Expanded Abstracts, 3908–3912.

Dix, C., 1955, Seismic velocities from surface measurements: Geophysics, **20**, 68–86, doi: 10.1190/1.1438126.

Fletcher, R. P., S. Archer, D. Nichols, and W. Mao, 2012, Inversion after depth imaging: 82nd Annual International Meeting, SEG, Expanded Abstracts, doi: 10.1190/segam2012-0427.1.

Fomel, S., P. Sava, I. Vlad, Y. Liu, and V. Bashkardin, 2013, Madagascar: Open-source software project for multidimensional data analysis and reproducible computational experiments: Journal of Open Research Software, **1**, e8, doi: 10.5334/jors.ag.

Gardner, G., L. Gardner, and A. Gregory, 1974, Formation velocity and density the diagnostic basics for stratigraphic traps: Geophysics, **39**, 770–780, doi: 10.1190/1.1440465.

Kamei, R., R. G. Pratt, and T. Tsuji, 2013, On acoustic waveform tomography of wide-angle OBS data — Strategies for pre-conditioning and inversion: Geophysical Journal International, **194**, 1250–1280, doi: 10.1093/gji/ggt165.

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.

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

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, 2004a, Wave-equation migration velocity analysis — I: Theory: Geophysical Prospecting, **52**, 593–606, doi: 10.1111/j.1365-2478.2004.00447.x.

Sava, P., and B. Biondi, 2004b, Wave-equation migration velocity analysis — II: Subsalt imaging examples: Geophysical Prospecting, **52**, 607–623, doi: 10.1111/j.1365-2478.2004.00448.x.

Sava, P., and S. Fomel, 2006, Time-shift imaging condition in seismic migration: Geophysics, **71**, no. 6, S209–S217, doi: 10.1190/1.2338824.

Sava, P., and I. Vasconcelos, 2011, Extended imaging conditions for wave-equation migration: Geophysical Prospecting, **59**, 35–55, doi: 10.1111/j.1365-2478.2010.00888.x.

Sava, P. C., and S. Fomel, 2003, Angle-domain common-image gathers by wavefield continuation methods: Geophysics, **68**, 1065–1074, doi: 10.1190/1.1581078.

Shan, G., and Y. Wang, 2013, RTM based wave equation migration velocity analysis: 83rd Annual International Meeting, SEG, Expanded Abstracts, 4726–4731.

Shen, P., and H. Calandra, 2005, One-way waveform inversion within the framework of adjoint state differential migration: 75th Annual International Meeting, SEG, Expanded Abstracts, 1709–1712.

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.

Shin, C., and Y. Ho Cha, 2009, Waveform inversion in the Laplace-Fourier domain: Geophysical Journal International, **177**, 1067–1079, doi: 10.1111/j.1365-246X.2009.04102.x.

Sirgue, L., and R. Pratt, 2004, Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies: Geophysics, **69**, 231–248, doi: 10.1190/1.1649391.

Symes, W. W., 2008, Migration velocity analysis and waveform inversion: Geophysical Prospecting, **56**, 765–790, doi: 10.1111/j.1365-2478.2008.00698.x.

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

Downloaded 10/27/17 to 138.67.129.100. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

Valenciano, A. A., B. L. Biondi, and R. G. Clapp, 2009, Imaging by target-oriented wave-equation inversion: Geophysics, **74**, no. 6, WCA109–WCA120, doi: 10.1190/1.3250267.

Weibull, W., and B. Arntsen, 2013, Automatic velocity analysis with reverse-time migration: Geophysics, **78**, no. 4, S179–S192, doi: 10.1190/geo2012-0064.1.

Woodward, M., 1992, Wave-equation tomography: Geophysics, **57**, 15–26, doi: 10.1190/1.1443179.

Yang, T., and P. Sava, 2010, Moveout analysis of wave-equation extended images: Geophysics, **75**, no. 4, S151–S161, doi: 10.1190/1.3460296.

Yang, T., and P. Sava, 2015, Image-domain wavefield tomography with extended common-image-point gathers: Geophysical Prospecting, **63**, 1086–1096, doi: 10.1111/1365-2478.12204.

Yang, T., J. Shragge, and P. Sava, 2013, Illumination compensation for image-domain wavefield tomography: Geophysics, **78**, no. 5, U65–U76, doi: 10.1190/geo2012-0278.1.

Zhang, Y., and G. Shan, 2013, Wave-equation migration velocity analysis using partial stack-power maximization: 83rd Annual International Meeting, SEG, Expanded Abstracts, 4847–4852.

Downloaded 10/27/17 to 138.67.129.100. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/