

Evaluating the potential of reflection-based waveform inversion

Khalid Almuteri, Yu Geng and Kris Innanen

ABSTRACT

Full waveform inversion (FWI) is a powerful tool to build high-resolution velocity models, from recorded seismic data. However, a major issue with FWI is that it fails at reconstructing the low-wavenumber components in the absence of low-frequency information in the data. Generally, for a limited-offset acquisition geometry, deep targets are only sampled by reflected waves with narrow scattering angles, which makes such failure inevitable. In this paper, we point out the limitation of conventional FWI when applied to reflection data, and review an alternative approach to overcome this limitation. The new waveform inversion formalism relies on decomposing the subsurface model into a background part that we seek to resolve, and a reflectivity part that we assume to be known. We show that separating the decoupled velocity model into long-wavelength and short-wavelength components permit us to extract the contribution of the reflected data to the background part of the velocity model.

INTRODUCTION

Full waveform inversion (FWI) is an ill-posed data-fitting procedure that aims at reconstructing the earth's physical parameters by iteratively minimizing the least-squares norm of the difference between the predicted and observed data (Lailly, 1983; Tarantola, 1984; Virieux and Operto, 2009). The rise of FWI, as a tool for velocity model building, is due to the high-resolution velocity models that it provides (Pan and Innanen, 2015).

The resolution of the FWI reconstruction is related to the diffraction tomography principle (Devaney, 1982; Miller et al., 1987; Wu and Toksöz, 1987; Brossier et al., 2015); which relates the recoverable wavenumber \mathbf{k} , sampled at a point diffractor (Figure 1), to the local wavelength and aperture angle θ according to equation (1).

$$\mathbf{k} = \frac{4\pi}{\lambda_0} \cos\left(\frac{\theta}{2}\right) \mathbf{n} \quad (1)$$

where λ_0 is the local wavelength, and $\mathbf{n} = \frac{\mathbf{q}_s + \mathbf{q}_r}{\|\mathbf{q}_s + \mathbf{q}_r\|}$ (Figure 1).

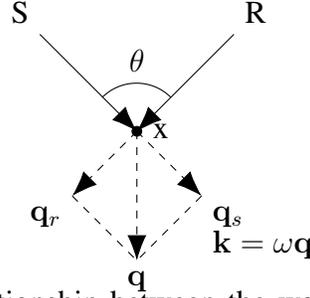


FIG. 1. Illustration of the relationship between the wavenumber k and the acquisition geometry for a point diffractor. S denotes source, R denotes receiver, θ is the aperture angle, x is a point diffractor, \mathbf{q}_s and \mathbf{q}_r are the slowness vectors, $\mathbf{q} = \mathbf{q}_s + \mathbf{q}_r$, ω is the angular frequency, and k is the recoverable wavenumber.

In a narrow-aperture acquisition geometry, where the depth of investigation is larger than the offset range, the shallow area of the subsurface is sampled by reflections, refractions, and direct waves; while only reflections sample the deep section. According to equation (1), in the deep part of the model, due to the small range of θ in narrow-aperture acquisition with smooth background velocity, only high-wavenumbers will be reconstructed, while in the shallow part both low and high wavenumbers will be reconstructed (Brossier et al., 2015); hence, a successful FWI requires the presence of long-offset data, low frequency data, and an accurate starting model.

An alternative approach to the conventional FWI, which aims at retrieving the low-wavenumber components of the velocity in areas sampled by reflected data only, is proposed by Xu et al. (2012), which is known as reflection-based waveform inversion (RWI). In this new approach, the velocity model is decomposed into a background/transmission model that we seek to resolve and a reflectivity model that is assumed to be known, allowing the emphasis on the transmission wavepaths of the reflected data in the inversion process. There are three main differences between conventional FWI and RWI. First, the primary goal of RWI is to invert for the background model, and not to obtain a high-resolution model. Second, FWI uses the full wavefield in the inversion process, including direct waves, refracted and reflected waves, while RWI uses only reflected waves. Third, RWI relies on a migration/demigration process (Zhou et al., 2012).

The outline of this paper is as follows. First, we cover the derivation of the necessary equations for and the theory behind RWI. Second, we demonstrate the effectiveness of RWI by comparing its results to conventional FWI. Finally, we briefly discuss the process of migration and demigration in Appendix A. Although we perform all the derivations in the frequency domain, we carry out the inversion in time domain.

THEORY

Derivation

We start by define two equations in an acoustic medium

$$[\nabla^2 + \omega^2 \mathbf{m}_0] G_0(\mathbf{r}_g, \mathbf{r}_s, \omega) = \delta(\mathbf{r} - \mathbf{r}_s) \quad (2)$$

$$[\nabla^2 + \omega^2 \mathbf{m}] G(\mathbf{r}_g, \mathbf{r}_s, \omega) = \delta(\mathbf{r} - \mathbf{r}_s) \quad (3)$$

where $\mathbf{m} = \mathbf{m}_0 + \delta\mathbf{m}$ is the actual subsurface model, \mathbf{m}_0 is the background model, $\delta\mathbf{m}$ is the reflectivity model, $G_0(\mathbf{r}_g, \mathbf{r}_s, \omega)$ is the predicted data sampled at \mathbf{r}_g due to a source at \mathbf{r}_s , and $G(\mathbf{r}_g, \mathbf{r}_s, \omega)$ is the observed data sampled at \mathbf{r}_g due to a source at \mathbf{r}_s . It should be noted that the wave equation is parametrized in terms of the slowness squared.

We define the misfit function as a least-squares norm given by

$$E(\mathbf{m}_0) = \frac{1}{2} \|\delta P(\mathbf{r}_g, \mathbf{r}_s, \omega)\|^2, \quad (4)$$

where $\delta P(\mathbf{r}_g, \mathbf{r}_s, \omega)$ is the data residual (equation 5).

$$\delta P(\mathbf{r}_g, \mathbf{r}_s, \omega) = G(\mathbf{r}_g, \mathbf{r}_s, \omega) - G_0(\mathbf{r}_g, \mathbf{r}_s, \omega). \quad (5)$$

Taking the derivative of equation (4) in the vicinity of the model parameter \mathbf{m}_0 gives

$$\frac{\partial E(\mathbf{m}_0)}{\partial \mathbf{m}_0(\mathbf{r})} = - \sum_{\mathbf{r}_s, \mathbf{r}_g} \int d\omega \Re \left\{ \frac{\partial G_0(\mathbf{r}_g, \mathbf{r}_s, \omega)}{\partial \mathbf{m}_0(\mathbf{r})} \delta P^*(\mathbf{r}_g, \mathbf{r}_s, \omega) \right\}, \quad (6)$$

where

$$g(\mathbf{r}) = \frac{\partial E(\mathbf{m}_0)}{\partial \mathbf{m}_0(\mathbf{r})} \quad (7)$$

is the gradient, and $\frac{\partial G_0(\mathbf{r}_g, \mathbf{r}_s, \omega)}{\partial \mathbf{m}_0(\mathbf{r})}$ is the sensitivity or Fréchet derivative. The sensitivity describes the changes in the wavefield due to changes in the model parameters.

Substituting $\mathbf{m} = \mathbf{m}_0 + \delta\mathbf{m}$ into equation (3) yields

$$[\nabla^2 + \omega^2 \mathbf{m}_0] G(\mathbf{r}_g, \mathbf{r}_s, \omega) = \delta(\mathbf{r}_g - \mathbf{r}_s) - \omega^2 \delta\mathbf{m}(\mathbf{r}_g) G(\mathbf{r}_g, \mathbf{r}_s, \omega), \quad (8)$$

where its integral form solution is given by

$$\begin{aligned} G(\mathbf{r}_g, \mathbf{r}_s, \omega) &= G_0(\mathbf{r}_g, \mathbf{r}_s, \omega) - \omega^2 \int d\mathbf{r}' G_0(\mathbf{r}_g, \mathbf{r}', \omega) \delta\mathbf{m}(\mathbf{r}') G(\mathbf{r}', \mathbf{r}_s, \omega) \\ \Rightarrow \delta G(\mathbf{r}_g, \mathbf{r}_s, \omega) &= -\omega^2 \int d\mathbf{r}' G_0(\mathbf{r}_g, \mathbf{r}', \omega) \delta\mathbf{m}(\mathbf{r}') G(\mathbf{r}', \mathbf{r}_s, \omega). \end{aligned} \quad (9)$$

Using born series to eliminate $G(\mathbf{r}', \mathbf{r}_s, \omega)$ yields

$$\begin{aligned} \delta G(\mathbf{r}_g, \mathbf{r}_s, \omega) &= -\omega^2 \int d\mathbf{r}' G_0(\mathbf{r}_g, \mathbf{r}', \omega) \delta\mathbf{m}(\mathbf{r}') G_0(\mathbf{r}', \mathbf{r}_s, \omega) \\ &+ \omega^4 \int d\mathbf{r}' G_0(\mathbf{r}_g, \mathbf{r}', \omega) \delta\mathbf{m}(\mathbf{r}') \int d\mathbf{r}'' G_0(\mathbf{r}', \mathbf{r}'', \omega) \delta\mathbf{m}(\mathbf{r}'') G_0(\mathbf{r}'', \mathbf{r}_s, \omega) \\ &+ \dots \end{aligned} \quad (10)$$

We next localize the perturbation model $\delta\mathbf{m}(\mathbf{r})$ by introducing the delta function so that

$$\delta\mathbf{m}(\mathbf{r}') = \delta m \delta(\mathbf{r}' - \mathbf{r}), \quad (11)$$

which allows us to evaluate the integrals easily. By substituting equation (11) into equation (10) we obtain

$$\begin{aligned} \delta G(\mathbf{r}_g, \mathbf{r}_s, \omega) &= -\omega^2 G_0(\mathbf{r}_g, \mathbf{r}, \omega) \delta m G_0(\mathbf{r}, \mathbf{r}_s, \omega) \\ &+ \omega^4 G_0(\mathbf{r}_g, \mathbf{r}, \omega) \delta m G_0(\mathbf{r}, \mathbf{r}, \omega) \delta m G_0(\mathbf{r}, \mathbf{r}_s, \omega) + \dots \\ &= -\omega^2 G_0(\mathbf{r}_g, \mathbf{r}, \omega) \delta m G_0(\mathbf{r}, \mathbf{r}_s, \omega) [1 - \omega^2 \delta m G_0(\mathbf{r}, \mathbf{r}, \omega) + \dots]. \end{aligned} \quad (12)$$

Noting that the series expansion of $\frac{1}{1+x}$ is $(1 - x + x^2 - \dots)$, equation (12) can be re-written as

$$\begin{aligned} \delta G(\mathbf{r}_g, \mathbf{r}_s, \omega) &= -\frac{\omega^2 G_0(\mathbf{r}_g, \mathbf{r}, \omega) \delta m G_0(\mathbf{r}, \mathbf{r}_s, \omega)}{1 + \omega^2 \delta m G_0(\mathbf{r}, \mathbf{r}, \omega)} \\ \Rightarrow \frac{\delta G(\mathbf{r}_g, \mathbf{r}_s, \omega)}{\delta m} &= -\frac{\omega^2 G_0(\mathbf{r}_g, \mathbf{r}, \omega) G_0(\mathbf{r}, \mathbf{r}_s, \omega)}{1 + \omega^2 \delta m G_0(\mathbf{r}, \mathbf{r}, \omega)}. \end{aligned} \quad (13)$$

Taking the limit so that δm vanishes gives the conventional FWI kernel

$$\frac{\partial G(\mathbf{r}_g, \mathbf{r}_s, \omega)}{\partial \mathbf{m}_0} = \lim_{\delta m \rightarrow 0} \frac{\delta G(\mathbf{r}_g, \mathbf{r}_s, \omega)}{\delta m} = -\omega^2 G_0(\mathbf{r}_g, \mathbf{r}, \omega) G_0(\mathbf{r}, \mathbf{r}_s, \omega). \quad (14)$$

In conventional FWI, we tend to resolve and update a velocity model that consists of a background part and a reflectivity part. In other words, to find an updated model \mathbf{m}_{n+1} , we need to update a current model \mathbf{m}_n , which is composed of a background part and perturbation part, by solving for a model update $\delta \mathbf{m}_n$. However, in RWI we are only interested in updating the background model; hence, we need to take the derivative of equation (9) with respect to the background model \mathbf{m}_0 to obtain

$$\begin{aligned} \frac{\partial \delta G(\mathbf{r}_g, \mathbf{r}_s, \omega)}{\partial \mathbf{m}_0} &= -\omega^2 \int d\mathbf{r}' \left(\frac{\partial G_0(\mathbf{r}_g, \mathbf{r}', \omega)}{\partial \mathbf{m}_0} G(\mathbf{r}', \mathbf{r}_s, \omega) \right. \\ &\quad \left. + G_0(\mathbf{r}_g, \mathbf{r}', \omega) \frac{\partial G(\mathbf{r}', \mathbf{r}_s, \omega)}{\partial \mathbf{m}_0} \right) \delta \mathbf{m}(\mathbf{r}'), \end{aligned} \quad (15)$$

where

$$\frac{\partial G_0(\mathbf{r}_g, \mathbf{r}', \omega)}{\partial \mathbf{m}_0} = -\omega^2 G_0(\mathbf{r}_g, \mathbf{r}'', \omega) G_0(\mathbf{r}'', \mathbf{r}', \omega) \quad (16)$$

and

$$\frac{\partial G(\mathbf{r}', \mathbf{r}_s, \omega)}{\partial \mathbf{m}_0} = -\omega^2 G(\mathbf{r}', \mathbf{r}'', \omega) G(\mathbf{r}'', \mathbf{r}_s, \omega). \quad (17)$$

Substituting equations (16) and (17) into equation (15) results in

$$\begin{aligned} \frac{\partial \delta G(\mathbf{r}_g, \mathbf{r}_s, \omega)}{\partial \mathbf{m}_0} &= \omega^4 \int d\mathbf{r}' \left(G_0(\mathbf{r}_g, \mathbf{r}'', \omega) G_0(\mathbf{r}'', \mathbf{r}', \omega) G(\mathbf{r}', \mathbf{r}_s, \omega) \right. \\ &\quad \left. + G_0(\mathbf{r}_g, \mathbf{r}', \omega) G(\mathbf{r}', \mathbf{r}'', \omega) G(\mathbf{r}'', \mathbf{r}_s, \omega) \right) \delta \mathbf{m}(\mathbf{r}'). \end{aligned} \quad (18)$$

By re-writing equation (18) into

$$\begin{aligned} \frac{\partial \delta G(\mathbf{r}_g, \mathbf{r}_s, \omega)}{\partial \mathbf{m}_0} = & -\omega^2 G_0(\mathbf{r}_g, \mathbf{r}'', \omega) \left[-\omega^2 \int d\mathbf{r}' G_0(\mathbf{r}'', \mathbf{r}', \omega) \delta \mathbf{m}(\mathbf{r}') G(\mathbf{r}', \mathbf{r}_s, \omega) \right] \\ & - \omega^2 G(\mathbf{r}'', \mathbf{r}_s, \omega) \left[-\omega^2 \int d\mathbf{r}' G_0(\mathbf{r}_g, \mathbf{r}', \omega) \delta \mathbf{m}(\mathbf{r}') G(\mathbf{r}', \mathbf{r}'', \omega) \right], \end{aligned} \quad (19)$$

we note that the expressions within $[\cdot]$ can be replaced by the perturbed wavefield (equation 9) to obtain

$$\frac{\partial \delta G(\mathbf{r}_g, \mathbf{r}_s, \omega)}{\partial \mathbf{m}_0} = -\omega^2 \left(G_0(\mathbf{r}_g, \mathbf{r}'', \omega) \delta G(\mathbf{r}'', \mathbf{r}_s, \omega) + G(\mathbf{r}'', \mathbf{r}_s, \omega) \delta G(\mathbf{r}_g, \mathbf{r}'', \omega) \right), \quad (20)$$

where $\frac{\delta G(\mathbf{r}_g, \mathbf{r}_s, \omega | \mathbf{m}_0)}{\partial \mathbf{m}_0}$ is the RWI kernel. In equation (20), $G(\mathbf{r}'', \mathbf{r}_s, \omega)$ corresponds to the actual source wavefield, however, since obtaining the actual source wavefield is not possible, it is replaced with the modeled source wavefield $G_0(\mathbf{r}'', \mathbf{r}_s, \omega)$. The gradient of the misfit function with respect to the background model \mathbf{m}_0 is obtained by substituting equation (20) into equation (6) to give

$$\begin{aligned} g(\mathbf{r}) = \sum_{\mathbf{r}_s, \mathbf{r}_g} \int d\omega \omega^2 \left(\left[\delta G(\mathbf{r}, \mathbf{r}_s, \omega) \right] \times \left[G_0(\mathbf{r}_g, \mathbf{r}, \omega) \delta P^*(\mathbf{r}_s, \mathbf{r}_g, \omega) \right] \right. \\ \left. + \left[G_0(\mathbf{r}, \mathbf{r}_s, \omega) \right] \times \left[\delta G(\mathbf{r}_g, \mathbf{r}, \omega) \delta P^*(\mathbf{r}_s, \mathbf{r}_g, \omega) \right] \right), \end{aligned} \quad (21)$$

where $\delta G(\mathbf{r}, \mathbf{r}_s, \omega)$ is the demigrated source wavefield (Appendix A), $G_0(\mathbf{r}_g, \mathbf{r}, \omega)$ is the receiver wavefield, $G_0(\mathbf{r}, \mathbf{r}_s, \omega)$ is the source wavefield, $\delta G(\mathbf{r}_g, \mathbf{r}, \omega)$ is the demigrated receiver wavefield, and $\delta P^*(\mathbf{r}_s, \mathbf{r}_g, \omega)$ is the conjugated data residual. Equation (21) depends on both the background model and the perturbation model (through $\delta G(\mathbf{r}, \mathbf{r}_s, \omega)$ and $\delta G(\mathbf{r}_g, \mathbf{r}, \omega)$).

The corresponding equation in time domain to equation (21) is given by

$$\begin{aligned} g(\mathbf{r}) = \sum_{\mathbf{r}_s, \mathbf{r}_g} \int_0^T dt \left(\left[\delta \ddot{G}(\mathbf{r}, \mathbf{r}_s, t) \right] \left[G_0(\mathbf{r}_g, \mathbf{r}, T-t) * \delta G(\mathbf{r}_g, \mathbf{r}_s, t) \right] \right. \\ \left. + \left[\ddot{G}_0(\mathbf{r}, \mathbf{r}_s, t) \right] \left[\delta G(\mathbf{r}_g, \mathbf{r}, T-t) * \delta G(\mathbf{r}_g, \mathbf{r}_s, t) \right] \right), \end{aligned} \quad (22)$$

where $[G_0(\mathbf{r}_g, \mathbf{r}, T-t) * \delta G(\mathbf{r}_g, \mathbf{r}_s, t)]$ is the back-propagated data residual wavefield, $[\delta G(\mathbf{r}_g, \mathbf{r}, T-t) * \delta G(\mathbf{r}_g, \mathbf{r}_s, t)]$ is the demigrated receiver wavefield, \ddot{G} corresponds to the second time derivative of G , and $*$ represents convolution. The parametrization of the gradient (equations 21 and 22) is in terms of slowness squared $\left(m(\mathbf{r}) = \frac{1}{v(\mathbf{r})^2} \right)$.

Sensitivity Kernels

To understand the advantage of model decomposition we look at the contribution of direct waves and reflected waves to the FWI and the RWI sensitivity kernels. The sensitivity kernel describes the perturbation in the data domain due to a perturbation in the model parameter (Chi et al., 2015). Equation (14) is the conventional FWI sensitivity kernel in frequency domain. The corresponding FWI kernel in time domain, \mathbf{K}_{FWI} , is given by the cross-correlation of the second time derivative of the source wavefield with the time-reversed back-propagated receiver wavefield (equation 23).

$$\mathbf{K}_{FWI} = \ddot{G}(\mathbf{r}, \mathbf{r}_s, t) \otimes G(\mathbf{r}_g, \mathbf{r}, -t), \quad (23)$$

where \otimes represents cross-correlation, and $G(\mathbf{r}_g, \mathbf{r}, t)$ satisfies

$$\begin{cases} \left[\nabla^2 + \mathbf{m} \frac{\partial^2}{\partial t^2} \right] G(\mathbf{r}, t) = 0 \\ G(\mathbf{r}, \mathbf{r}_g, t) = G(\mathbf{r}_g, \mathbf{r}_s, t), \end{cases} \quad (24)$$

and $G(\mathbf{r}, \mathbf{r}_s, t)$ satisfies

$$\left[\nabla^2 + \mathbf{m} \frac{\partial^2}{\partial t^2} \right] G(\mathbf{r}, \mathbf{r}_s, t) = \delta(\mathbf{r} - \mathbf{r}_s), \quad (25)$$

where $G(\mathbf{r}_g, \mathbf{r}_s, t)$ is the recorded data. The sensitivity kernel of RWI in frequency domain is given by equation (20), and the corresponding kernel in time domain, \mathbf{K}_{RWI} , is given by

$$\mathbf{K}_{RWI} = \delta \ddot{G}(\mathbf{r}, \mathbf{r}_s, t | \mathbf{m}_0) \otimes G(\mathbf{r}_g, \mathbf{r}, -t | \mathbf{m}_0) + \ddot{G}(\mathbf{r}, \mathbf{r}_s, t | \mathbf{m}_0) \otimes \delta G(\mathbf{r}_g, \mathbf{r}, -t | \mathbf{m}_0), \quad (26)$$

where $\delta \ddot{G}(\mathbf{r}, \mathbf{r}_s, t | \mathbf{m}_0)$ is the second time derivative of the demigrated source wavefield, $G(\mathbf{r}_g, \mathbf{r}, -t | \mathbf{m}_0)$ is the time reversed receiver wavefield, $\ddot{G}(\mathbf{r}, \mathbf{r}_s, t | \mathbf{m}_0)$ is the second time derivative of the source wavefield, and $\delta G(\mathbf{r}_g, \mathbf{r}, -t | \mathbf{m}_0)$ is the demigrated receiver wavefield.

To construct the FWI and RWI sensitivity kernels, we use a two-layer model, where the velocity of the first layer is 2000 m/s and the velocity of the second layer is 3000 m/s , and the interface is at 0.6 km depth. In this model, a source is placed at (0.33, 0.33) km and a receiver is placed at (0.66, 0.33) km , (Figure 2). In constructing the sensitivity kernels, we use the full bandwidth by utilizing the time domain formalism of the kernels.

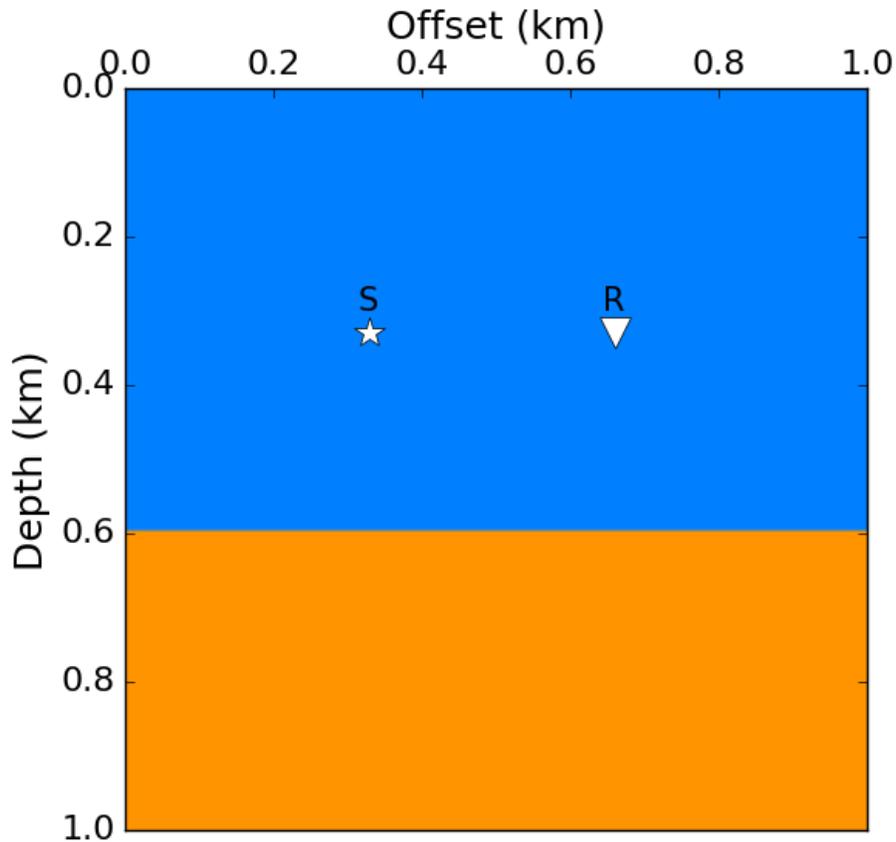


FIG. 2. Velocity model used to generated sensitivity kernels. The velocity of the first layer is 2000 m/s and the velocity of the second layer is 3000 m/s . S, ☆, denotes the source and R, ▽, denotes the receiver.

Analyzing the conventional FWI sensitivity sub-kernels (Figures 3a, 3b, and 3c) indicate that with a smooth initial model, early iterations will update the low-wavenumber components in the shallow part of the model (Figure 3a), and will update the high-wavenumber components in the deep part of the model, (Figure 3b), (Chi et al., 2015). Figure 3b indicates that the high-wavenumber update, in the deep section, will exhibit a migration-like reconstruction (Brossier et al., 2015).

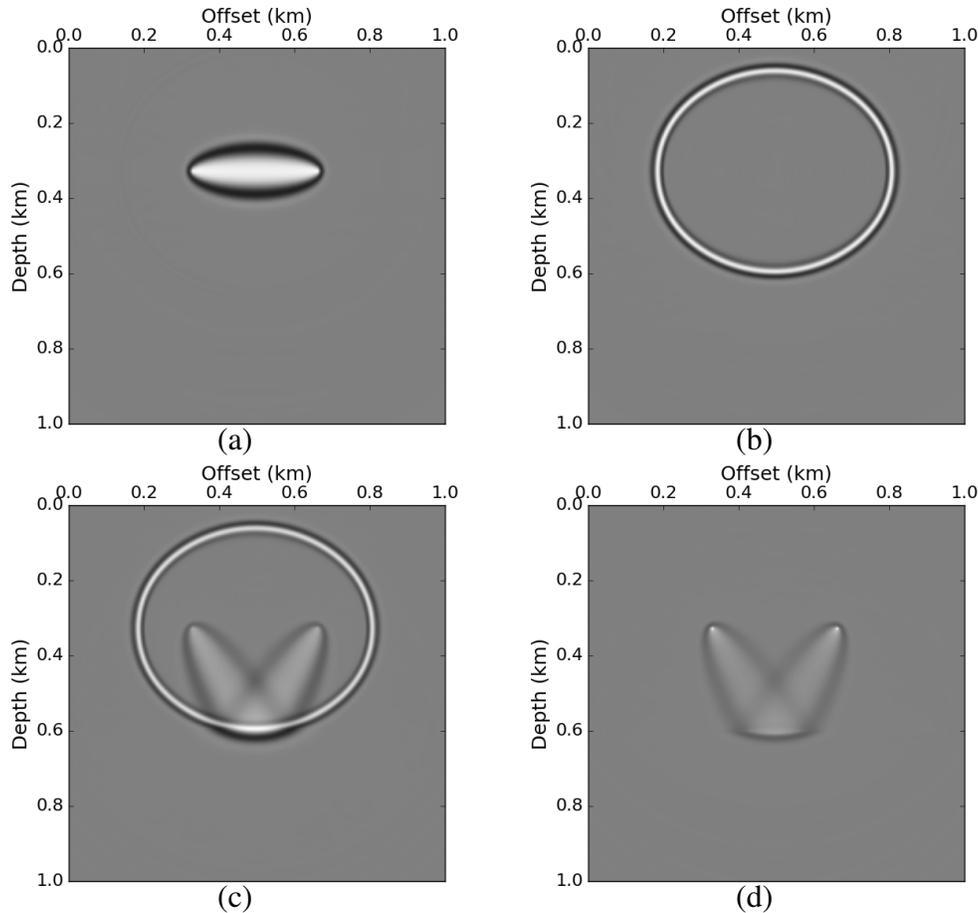


FIG. 3. FWI and RWI sensitivity kernels. (a) FWI direct wave sub-kernel. (b) Migration ellipse. This is the FWI reflected wave sub-kernel when model is smooth and does not contain reflectivity information. (c) FWI reflected wave sub-kernel. (d) RWI sensitivity kernel. We use a Ricker wavelet as a source function and the full bandwidth in generating the sensitivity kernels.

Such behavior in the inversion process is a result of the absence of reflected waves in the predicted data in early iterations. With increasing the number of iterations, the low-wavenumber components in the deep part will start updating (Figure 3c); however, the contribution of the high-wavenumber components is stronger; hence, the inversion process will try to match the predicted data and observed data by updating the high-wavenumber components (Chi et al., 2015). On the other hand, the RWI sensitivity kernel indicates that the reflected waves contribute to updating the background model (Figure 3d). Moreover, the predicted data in RWI will have reflected waves, as modeled data is generated through a migration/demigration process (Zhou et al., 2012).

NUMERICAL EXAMPLE

To examine RWI, we use a three-layer model of dimension $5000\text{ m} \times 3000\text{ m}$ (Figure 4). The model is composed of a background velocity of 2500 m/s with a low-velocity Gaussian anomaly, where the center of the lens has a velocity of 2200 m/s . The first re-

flector is at 3000 m in depth, with a velocity of 2750 m/s . The second reflector has a velocity of 3000 m/s , with a dipping angle. Fifty shots were used in this test, with the first shot at $x = 75\text{ m}$, the last shot at $x = 7425\text{ m}$, and a shot spacing of 150 m . There are 499 receivers in this test, with the first receiver at $x = 5\text{ m}$, the last receiver at $x = 7500\text{ m}$, with receiver spacing of 5 m . The recording time is 3.5 s and the time interval is 3.0 m.s. A Ricker wavelet with 10 Hz dominant frequency is used to generate the data. A constant velocity model is used as an initial model with a velocity of 2500 m/s (Figure 5). RWI is formulated such that the observed data consists of only reflected waves; hence, we mute direct waves in our observed data. We compare RWI results with FWI results for the same model; however, in FWI we utilize the full recorded information.

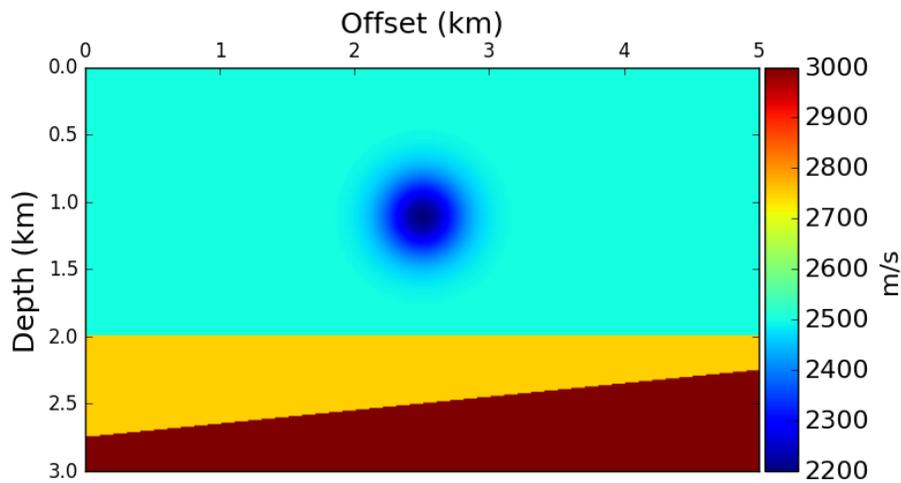


FIG. 4. True velocity model.

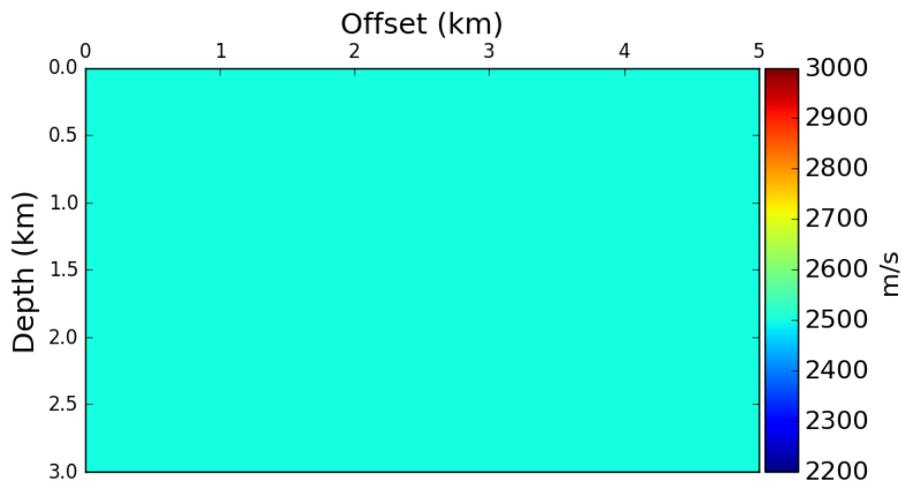


FIG. 5. Initial velocity model.

Figures 6 and 7 shows the conventional FWI gradient and the updated model, respectively, after one iteration. As expected, conventional FWI recovered the reflectivity model, but not the Gaussian anomaly that contributes to the background model. FWI failed in retrieving the low-wavenumber components of the velocity model due to the lack of low-frequency information in the data. As a consequence, the reflectivity model below the Gaussian anomaly is unfocused and mispositioned. Figures 8 and 9 shows the RWI gradient and the inverted velocity model after one iteration, respectively. We note that, unlike FWI, RWI recovered the general characteristics of the background model, and the misplaced reflectors in FWI are now correctly positioned.

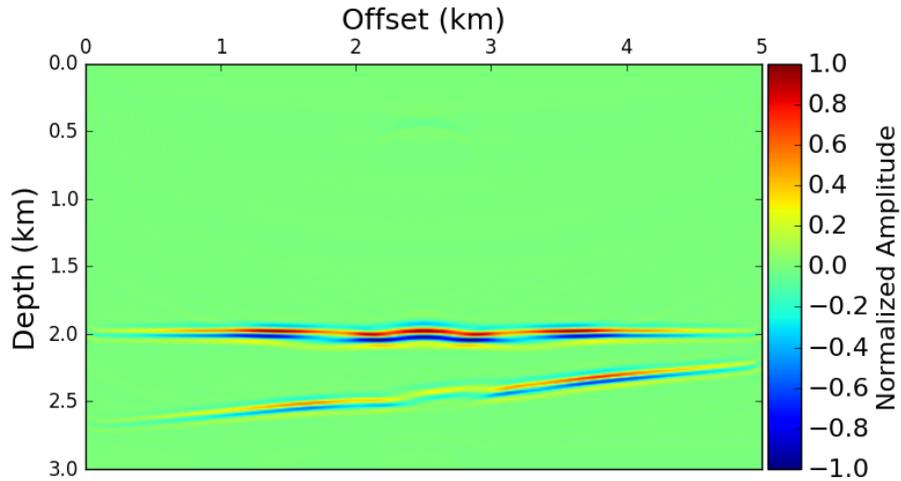


FIG. 6. FWI gradient.

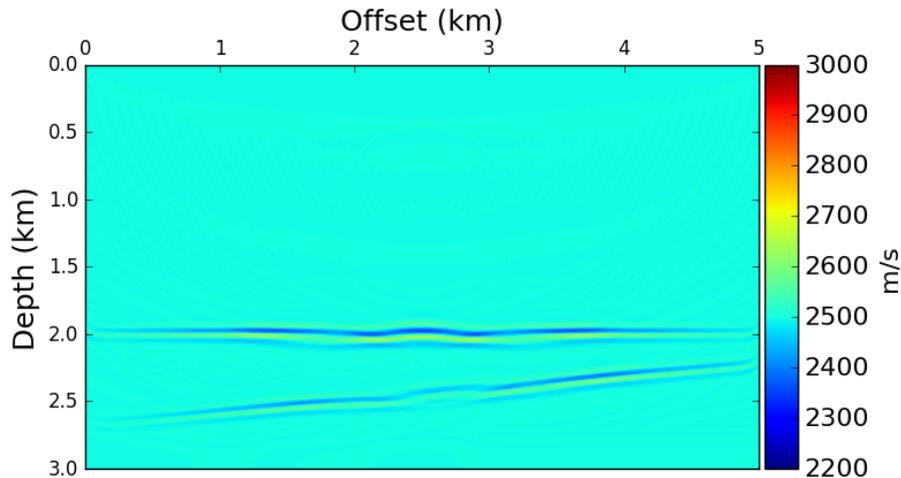


FIG. 7. Updated velocity model using FWI after one iteration.

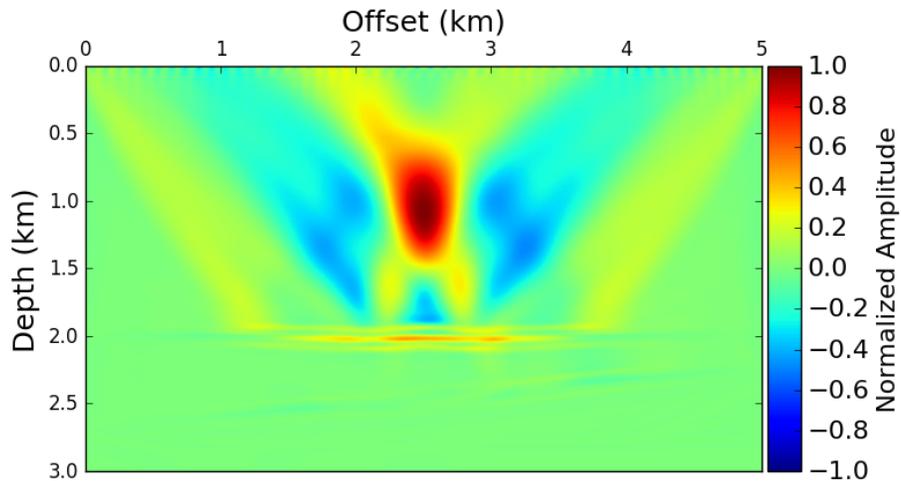


FIG. 8. RWI gradient.

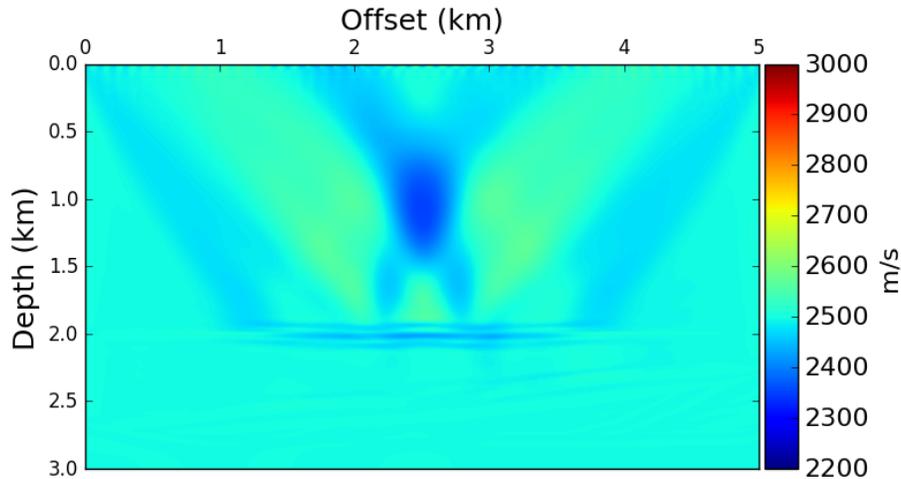


FIG. 9. Updated velocity model using RWI after one iteration.

Figures 10, 11, and 12 are the RTM images produced by using the initial velocity, the FWI inverted velocity, and the RWI inverted velocity, respectively, as the migration velocity. The migrated images using the initial velocity and the FWI velocity are both unfocused due to the inaccuracy of the models. On the other hand, the migrated image using the RWI inverted velocity shows better results and produces a well-focused image.

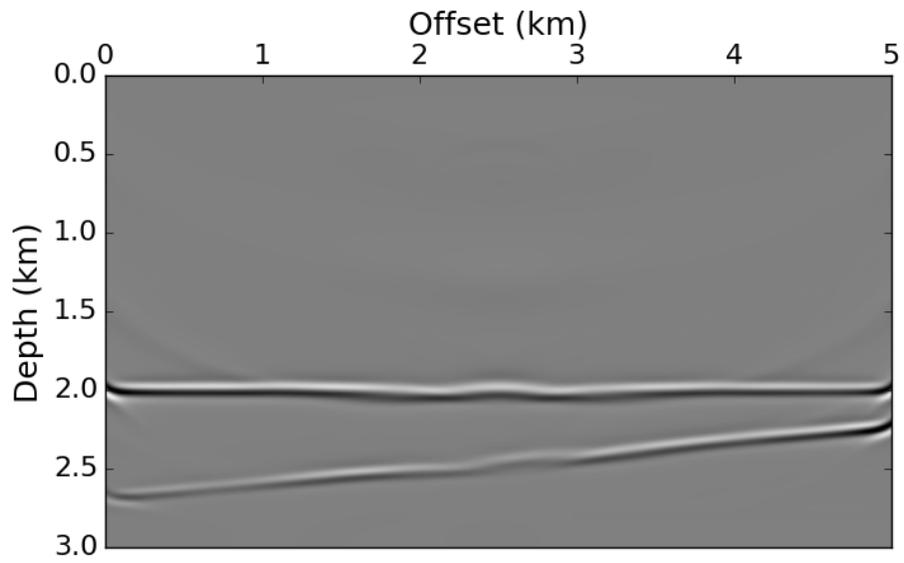


FIG. 10. RTM image migrated using the initial velocity.

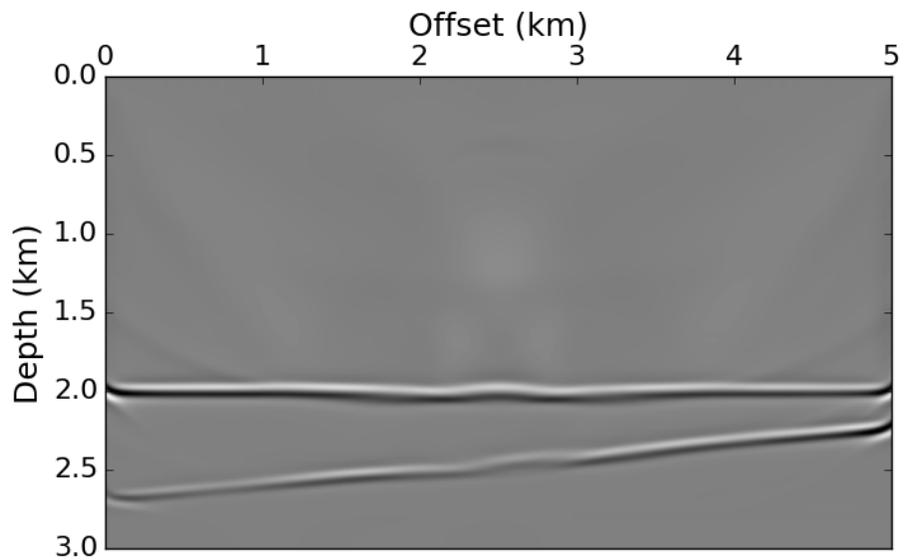


FIG. 11. RTM image migrated using the FWI inverted velocity.

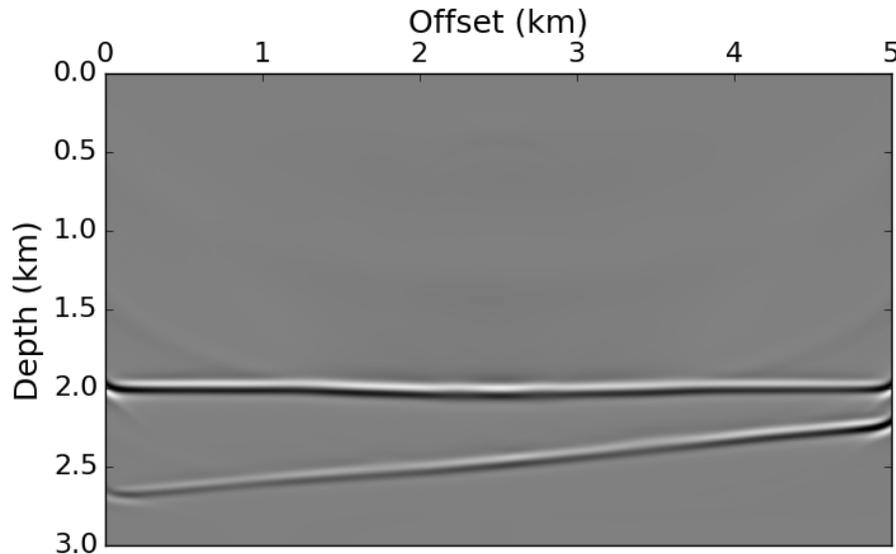


FIG. 12. RTM image migrated using the RWI inverted velocity.

CONCLUSION

In this paper, we showed that RWI is superior to FWI in retrieving long-wavelength components when building velocity models from reflection data. The background model does not generate backscattering; however, it controls the arrival times of reflected data. As a result, an accurate background model is necessary for seismic migration coherency.

APPENDIX A: SEISMIC MIGRATION AND DEMIGRATION

In FWI, the initial model is normally a smooth model that doesn't generate any reflections. As a result, for early iterations, the predicted data will only contain direct waves, which causes the FWI to fail in retrieving the low-wavenumber components of the model, which will result in cycle-skipping if the initial model is not accurate and close enough to the actual background model (Zhou et al., 2012). This failure is attributed to the fact that, at later iterations, FWI will try to minimize the data residual by updating the reflectivity model, constructed from migrating the reflection data residual, instead of the background model (Chi et al., 2015). The low-frequency and high-frequency components of the velocity model are decoupled, as the low-frequency components do not produce any reflections, and it only controls the arrival time of reflected data generated by the high-wavenumber components (Zhou et al., 2012). As a result, modeling by migration/demigration process produces reflected waves in the predicted data, even though the velocity model in use has only low-wavenumber components.

The first step in the migration/demigration process is to migrate the observed data, using a true-amplitude migration method, to generate the reflectivity model. The migrated image $I(\mathbf{r})$, (equation A-3), is formed by multiplying the source wavefield $G(\mathbf{r}, \mathbf{r}_s, \omega)$ that satisfy

$$[\nabla^2 + \omega^2 \mathbf{m}] G(\mathbf{r}, \mathbf{r}_s, \omega) = \delta(\mathbf{r} - \mathbf{r}_s), \quad (\text{A-1})$$

with the conjugated receiver wavefield $G^*(\mathbf{r}_g, \mathbf{r}, \omega)$ that satisfy

$$\begin{cases} [\nabla^2 + \omega^2 \mathbf{m}] G(\mathbf{r}, \omega) = 0 \\ G(\mathbf{r}, \mathbf{r}_g, \omega) = G(\mathbf{r}_g, \mathbf{r}_s, \omega), \end{cases} \quad (\text{A-2})$$

where $G(\mathbf{r}_g, \mathbf{r}_s, \omega)$ is the observed data at the receiver location \mathbf{r}_g due to a source at the location \mathbf{r}_s , which gives an RTM migrated image.

$$I(\mathbf{r}) = \int d\omega G(\mathbf{r}, \mathbf{r}_s, \omega) \times G^*(\mathbf{r}_g, \mathbf{r}, \omega). \quad (\text{A-3})$$

Then the demigrated data (equation A-4), is produced by using the migrated image as a source in depth in a concept similar to the exploding reflector model. The source term in the wave equation, used to generate the demigrated data, is created by correlating the source wavefield with the migrated image (equation A-5).

$$d_{cal}(\mathbf{r}_g, \mathbf{r}_s, \omega) = \delta G(\mathbf{r}_g, \mathbf{r}_s, \omega), \quad (\text{A-4})$$

where $\delta G(\mathbf{r}_g, \mathbf{r}_s, \omega)$ is given by

$$[\nabla^2 + \omega^2 \mathbf{m}] \delta G(\mathbf{r}, \mathbf{r}_s, \omega) = I(\mathbf{r}) \cdot G(\mathbf{r}, \mathbf{r}_s, \omega). \quad (\text{A-5})$$

ACKNOWLEDGEMENTS

We thank the sponsors of CREWES for continued support. This work was funded by CREWES industrial sponsors and NSERC (Natural Science and Engineering Research Council of Canada) through the grant CRDPJ 461179-13.

REFERENCES

- Brossier, R., Operto, S., and Virieux, J., 2015, Velocity model building from seismic reflection data by full-waveform inversion: *Geophysical Prospecting*, **63**, No. 2, 354–367.
- Chi, B., Dong, L., and Liu, Y., 2015, Correlation-based reflection full-waveform inversion: *Geophysics*, **80**, No. 4, R189–R202.
- Devaney, A. J., 1982, A filtered backpropagation algorithm for diffraction tomography: *Ultrasonic imaging*, **4**, No. 4, 336–350.
- Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migrations, *in* Conference on Inverse Scattering: Theory and Application, Society of Industrial and Applied Mathematics, 206–220.
- Miller, D., Oristaglio, M., and Beylkin, G., 1987, A new slant on seismic imaging: Migration and integral geometry: *Geophysics*, **52**, No. 7, 943–964.
- Pan, W., and Innanen, K., 2015, Full-waveform inversion in the frequency-ray parameter domain: CREWES Annual Report, **27**.
- Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: *Geophysics*, **49**, No. 8, 1259–1266.
- Virieux, J., and Operto, S., 2009, An overview of full-waveform inversion in exploration geophysics: *Geophysics*, **74**, No. 6, WCC127–WCC152.

- Wu, R.-S., and Toksöz, M. N., 1987, Diffraction tomography and multisource holography applied to seismic imaging: *Geophysics*, **52**, No. 1, 11–25.
- Xu, S., Wang, D., Chen, F., Zhang, Y., and Lambare, G., 2012, Full waveform inversion for reflected seismic data, *in* 74th EAGE Conference and Exhibition incorporating EUROPEC 2012.
- Zhou, H., Amundsen, L., and Zhang, G., 2012, Fundamental issues in full waveform inversion, *in* SEG Technical Program Expanded Abstracts 2012.