

Combining forward-scattered and back-scattered wavefields in velocity analysis

Ali Almomin, Biondo Biondi, and Robert Clapp

ABSTRACT

Wave-Equation Migration Velocity Analysis (WEMVA) is a family of techniques that aims to improve the subsurface velocity model by minimizing the residual in the image space. This process is performed iteratively by linearizing the imaging operator in order to relate image perturbations to model updates. However, WEMVA techniques only utilize the kinematics in the forward-scattered wavefield. Ignoring the back-scattered wavefield results in model updates that have low vertical resolution. I present a method that combines the forward-scattered wavefield information from WEMVA with the back-scattered wavefield information from full-waveform inversion (FWI). This can be done by first decomposing the FWI gradient into forward- and back-scattered gradients, and then applying the proper weights to combine the back-scattered FWI gradient with the WEMVA gradient. These weights aim to enhance the components of the FWI gradients that overlap with the WEMVA gradient. Preliminary results show that the combined gradient converges faster and to a better solution.

INTRODUCTION

Seismic velocity-analysis methods can be divided into two major groups. First, there are techniques that aim at minimizing the misfit in the data domain, such as full-waveform inversion (Tarantola, 1984; Luo and Schuster, 1991). Second, there are other techniques that aim at improving the quality in the image domain, such as migration velocity analysis (MVA) (Symes and Carazzone, 1991; Biondi and Sava, 1999; Shen, 2004; Biondi, 2010). These techniques try to measure the quality of the image in several ways and then invert the estimated image perturbation using a linearized wave-equation operator.

There are several advantages to minimizing the residual in the image-space, such as increasing the signal-to-noise ratio and decreasing the complexity of the data (Tang et al., 2008). However, a common drawback in doing velocity analysis in the image domain is that only forward-scattered wavefields are used. This results in a loss of vertical resolution in the estimated model updates.

On the other hand, full-waveform inversion (FWI) does not have that problem, since it utilizes the information from both the forward-scattered and back-scattered

wavefields, i.e. both the kinematics and the dynamics. This results in higher resolution in the model estimates. However, FWI has the disadvantage of being highly nonlinear, which requires the starting model to be very close to the true model to avoid converging to local minima.

A straightforward solution is to first invert for the velocity model using MVA techniques and then use the output as the initial model for FWI. However, this practice might not work if the results of MVA are not accurate enough for FWI to start. This could be a result of the larger null space that forward-scattered wavefields do not constrain. Moreover, the convergence rate of the MVA techniques is going to be sub-optimal, since they do not use all of the information in the data.

In this paper, I try to improve the MVA results and convergence rate by supplying the back-scattered information to the gradient. First, I decompose the FWI gradient into two components: the forward-scattering gradient and the back-scattering gradient. The decomposition is done in Fourier domain based on the direction of wave propagation. Second, I develop a new gradient that combines the MVA gradients with the back-scattered component of the FWI gradient using the proper weighting function. This weighting function aims to emphasize the components of the back scattered FWI gradients that overlap with the MVA gradient. Finally, I apply the combined gradient in a synthetic example.

METHOD

When linearized over slowness, the FWI gradient can be written as follows (Tarantola, 1984):

$$g(\mathbf{x}) = \sum_{\omega, \mathbf{x}_s, \mathbf{x}_r} -2\omega^2 s_0(\mathbf{x}) G^*(\mathbf{x}, \mathbf{x}_s, \omega) G^*(\mathbf{x}, \mathbf{x}_r, \omega) \Delta d(\mathbf{x}_r, \mathbf{x}_s, \omega), \quad (1)$$

where G is the Green's function, s_0 is the background slowness, Δd is the data residual, \mathbf{x}_s and \mathbf{x}_r are the source and receiver coordinates, \mathbf{x} is the Green's functions' coordinate, and ω is frequency. FWI updates contain two components: a tomographic component and a migration component. To better understand these components, we will look at the migration and tomographic operators.

The operator relating the slowness model and the image is nonlinear. Therefore, the first step in evaluating a tomographic operator is to linearize the image \mathbf{I} around the background slowness \mathbf{s}_0 , as follows:

$$\mathbf{I} = \mathbf{I}_0 + \left. \frac{\partial \mathbf{I}}{\partial \mathbf{s}} \right|_{s_0} (\mathbf{s} - \mathbf{s}_0) + \dots, \quad (2)$$

where \mathbf{I}_0 is the background image. Under the Born approximation, the background image can be obtained as follows:

$$\mathbf{I}_0 = \mathbf{L}^* \mathbf{d}, \quad (3)$$

where \mathbf{L} is the Born modeling operator, and its adjoint \mathbf{L}^* is the migration operator. By neglecting the higher-order terms in the image series, we can define the tomographic operator as follows:

$$\Delta \mathbf{I} = \frac{\partial \mathbf{I}}{\partial \mathbf{s}}|_{s_0} \Delta \mathbf{s} = \mathbf{T} \Delta \mathbf{s}, \quad (4)$$

where \mathbf{T} is the tomographic operator. Now, we use the conventional imaging condition as follows:

$$I(\mathbf{x}, \mathbf{h}) = \sum_{\omega, \mathbf{x}_s, \mathbf{x}_r} G^*(\mathbf{x} - \mathbf{h}, \mathbf{x}_s, \omega) G^*(\mathbf{x} + \mathbf{h}, \mathbf{x}_r, \omega) d(\mathbf{x}_r, \mathbf{x}_s, \omega), \quad (5)$$

where I is the image, d is the surface data, and \mathbf{h} is the subsurface offset. The background image I_0 can be written as follows:

$$I_0(\mathbf{x}, \mathbf{h}) = \sum_{\omega, \mathbf{x}_s, \mathbf{x}_r} G_0^*(\mathbf{x} - \mathbf{h}, \mathbf{x}_s, \omega) G_0^*(\mathbf{x} + \mathbf{h}, \mathbf{x}_r, \omega) d(\mathbf{x}_r, \mathbf{x}_s, \omega), \quad (6)$$

where G_0 is the background Green's function. After derivation, as shown in Almomin and Tang (2010), we obtain the forward tomographic operator as follows:

$$\begin{aligned} \Delta I(\mathbf{x}, \mathbf{h}) &= \sum_{\mathbf{y}} \frac{\partial I(\mathbf{x}, \mathbf{h})}{\partial s(\mathbf{y})} \Delta s(\mathbf{y}) \\ &= \sum_{\omega, \mathbf{x}_s, \mathbf{x}_r, \mathbf{y}} \left\{ -2\omega^2 s_0(\mathbf{y}) G_0^*(\mathbf{y}, \mathbf{x}_s, \omega) G_0^*(\mathbf{x} - \mathbf{h}, \mathbf{y}, \omega) \right\} G_0^*(\mathbf{x} + \mathbf{h}, \mathbf{x}_r, \omega) d(\mathbf{x}_r, \mathbf{x}_s, \omega) \Delta s(\mathbf{y}) \\ &+ \sum_{\omega, \mathbf{x}_s, \mathbf{x}_r, \mathbf{y}} \left\{ -2\omega^2 s_0(\mathbf{y}) G_0^*(\mathbf{x} - \mathbf{h}, \mathbf{x}_s, \omega) G_0^*(\mathbf{x} + \mathbf{h}, \mathbf{y}, \omega) \right\} G_0^*(\mathbf{y}, \mathbf{x}_r, \omega) d(\mathbf{x}_r, \mathbf{x}_s, \omega) \Delta s(\mathbf{y}), \end{aligned} \quad (7)$$

where \mathbf{y} is the slowness coordinate. Notice that all the Green's functions in the migration and tomographic operators are background Green's functions, which means that they do not contain any perturbations. Thus, the migration operator correlates two background wavefields, whereas the tomographic operator correlates a background and a perturbed wavefield. Therefore, a migration component can be generalized as the result of correlating wavefields moving in the opposite direction, whereas a tomographic component is the result of correlating wavefields moving in the same direction. In FWI, the background and perturbed wavefields are summed together in one wavefield. Nonetheless, these components are correlated together to compute the gradient.

I now compute the operators' responses in a surface acquisition. The receiver spacing is 20 m, the source spacing is 80 m, and the temporal sampling is 2 ms. Figure 1(a) shows the tomographic operator response by applying its forward and adjoint on a spike. This operator response represents one column of the Hessian matrix of the tomographic operator. Figure 1(b) shows the amplitude spectrum of the tomographic operator response, which shows that the tomographic updates illuminate only the very small vertical wavenumbers. Similarly, figure 2(a) shows the

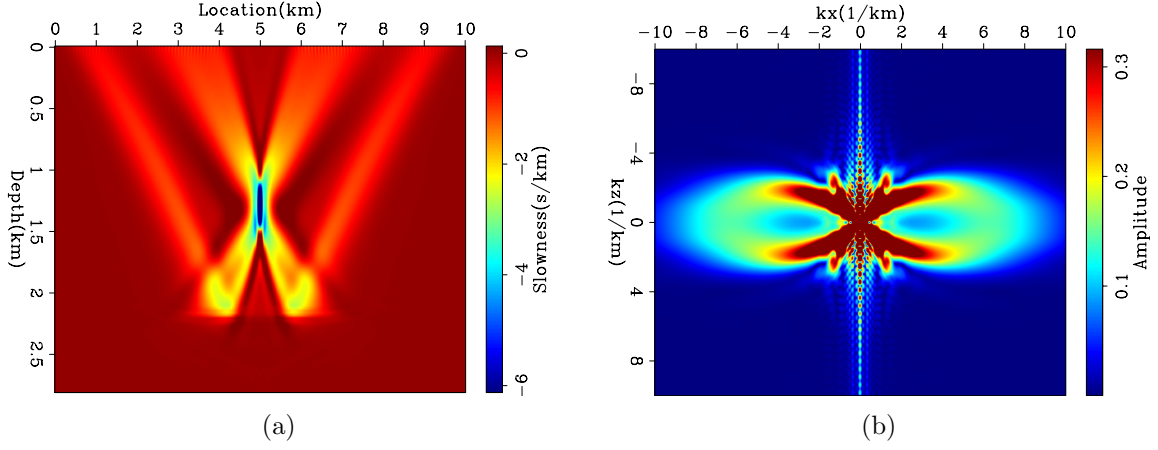


Figure 1: The tomographic operator response (a) and its amplitude spectrum (b). [CR]

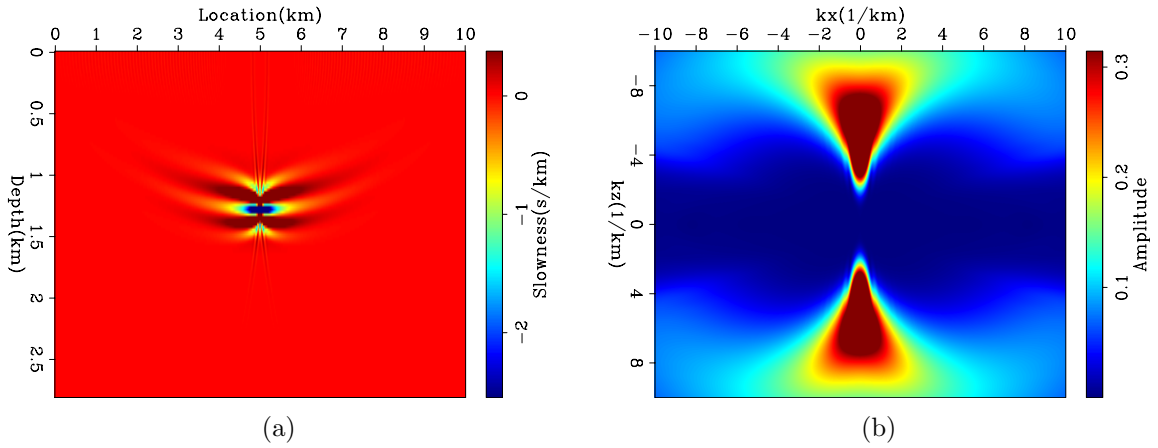


Figure 2: The migration operator response (a) and its amplitude spectrum (b). [CR]

migration operator response of a spike, and figure 2(b) shows its amplitude spectrum. Unlike the tomographic component, the migration component illuminates larger vertical wavenumbers. The illuminated wavenumbers in both operators depend on the reflection angles as well as the frequency content of the data. In this case, the data was created with a Ricker wavelet that has a dominant frequency of 15 Hz.

Now that I have defined the components in FWI, we can take a closer look at their error sensitivity. The migration operator is linear with respect to model updates. Hence, the migration component in FWI also behaves in a linear sense. This can be better explained by observing that wavefields moving in opposite directions are always going to result in the correct sign of update, as long as the source wavelet is correct. Background velocity changes, or errors, will mostly affect the position at which the wavefields correlate. On the other hand, tomographic components correlate wavefields that are moving in the same direction. For the correlation to have the correct direction, the two wavefields should be within the wavelength of one another. This strict requirement causes the nonlinear behavior of FWI. Since WEMVA techniques do not have such a requirement, they tend to have more stable behavior when the velocity error is large. Our goal is to combine the robust components of FWI and WEMVA.

One way to decompose the FWI gradient into its tomographic and migration components is to separate the wavefields based on their direction of propagation, e.g. up or down, which can be done in Fourier domain (Hu and McMechan, 1987; Liu et al., 2007; Taweesintananon, 2011). The decomposed FWI gradient can be written as follows:

$$\begin{aligned}
 g(\mathbf{x}) = & - \sum_{\omega, \mathbf{x}_s, \mathbf{x}_r} 2\omega^2 s_0(\mathbf{x}) G^{*+}(\mathbf{x}, \mathbf{x}_s, \omega) G^{*-}(\mathbf{x}, \mathbf{x}_r, \omega) \Delta d(\mathbf{x}_r, \mathbf{x}_s, \omega) \\
 & - \sum_{\omega, \mathbf{x}_s, \mathbf{x}_r} 2\omega^2 s_0(\mathbf{x}) G^{*-}(\mathbf{x}, \mathbf{x}_s, \omega) G^{*+}(\mathbf{x}, \mathbf{x}_r, \omega) \Delta d(\mathbf{x}_r, \mathbf{x}_s, \omega) \\
 & - \sum_{\omega, \mathbf{x}_s, \mathbf{x}_r} 2\omega^2 s_0(\mathbf{x}) G^{*+}(\mathbf{x}, \mathbf{x}_s, \omega) G^{*+}(\mathbf{x}, \mathbf{x}_r, \omega) \Delta d(\mathbf{x}_r, \mathbf{x}_s, \omega) \\
 & - \sum_{\omega, \mathbf{x}_s, \mathbf{x}_r} 2\omega^2 s_0(\mathbf{x}) G^{*-}(\mathbf{x}, \mathbf{x}_s, \omega) G^{*-}(\mathbf{x}, \mathbf{x}_r, \omega) \Delta d(\mathbf{x}_r, \mathbf{x}_s, \omega), \quad (8)
 \end{aligned}$$

where the positive and negative superscript signs of G indicate the direction of propagation of the wavefield.

In order for the migration component of FWI gradient to improve the convergence and results of WEMVA gradient, the two need to have some overlap of their illuminated wavenumbers. Figure 4(a) shows the multiplication of the amplitude spectra in 1(b) and 2(b). The overlap seems to be only at very low horizontal wavenumbers, which have some truncation artifacts. However, the operator's response is a function of frequency. Therefore, by taking lower frequencies when computing the migration component, we may increase the overlap region of the two gradients.

Figure 3(a) shows the migration operator response of a spike with a 5 Hz Ricker

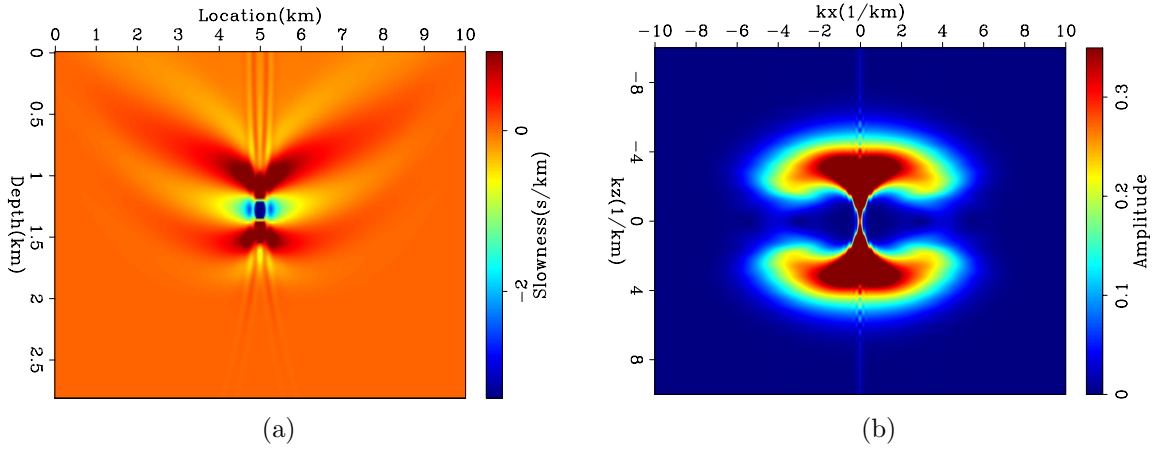


Figure 3: The low-frequency migration operator response (a) and its amplitude spectrum (b). [CR]

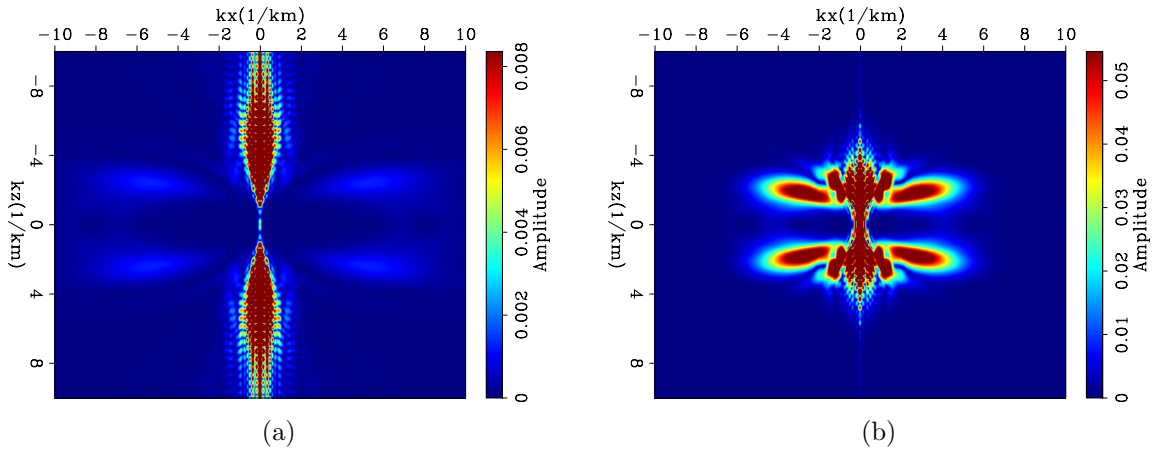


Figure 4: The spectra product of the tomographic operator with (a) the migration operator and (b) the low-frequency migration operator. [CR]

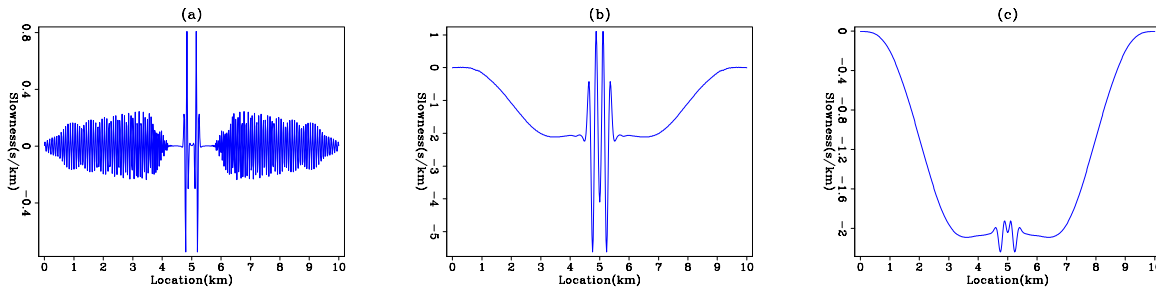


Figure 5: The vertical stack of (a) migration operator response, (b) low-frequency migration operator response and (c) laterally smooth low-frequency migration operator response. [CR]

wavelet, and figure 3(b) shows its amplitude spectrum. With lower frequencies, the migration operator response shifts toward smaller vertical wavenumbers. Figure 4(b) shows the multiplication of the amplitude spectra in 1(b) and 3(b), which shows a much larger and stronger overlap compared to figure 4(a). Figure 5 shows the vertical average migration response at 15 Hz and 5 Hz. It is clear that at low frequencies, the migration update starts to have small vertical-wavenumber contribution.

SYNTHETIC EXAMPLES

In this synthetic example, we will compare three inversion results after 20 iterations: low-frequency back-scattered FWI, correlation-based WEMVA (Almomin, 2011), and the combined gradient of WEMVA and the back-scattered FWI weighted by a scalar. The receiver spacing is 20 m, the source spacing is 80 m, and the temporal sampling is 2 ms. Figure 6(a) shows the true slowness anomaly we are inverting for. There is one reflector at the bottom of the slowness model. Figure 6(b) shows the results of low-frequency back-scattered FWI. Although it has the correct anomaly shape, the result has strong side-lobes and the amplitude is still very weak, because convergence is very slow.

Figure 7(a) shows the results of correlation-based WEMVA. As shown earlier, the gradient has low vertical resolution. Also, the anomaly has negative side-lobes. However, the amplitude of the anomaly is much stronger than with back-scattered FWI, and the kinematics of the reflector below converges much faster. Figure 7(b) shows the results of the combined gradient. The amplitude of the anomaly is even stronger and is closer to the true anomaly than is the WEMVA inversion. Furthermore, the side-lobes are much weaker, and the model has much better vertical resolution. Notice that figures 6(b), 7(a) and 7(b) have different color scales.

DISCUSSION AND CONCLUSIONS

I showed that the forward-scattered and back-scattered wavefields have overlap in their estimated slowness models. The synthetic example showed that supplying the back-scattered information can improve the tomographic inversion results in terms of convergence rate and accuracy. The overlap is strongest in the low-frequency part of the data. Therefore, the data must contain low frequencies for the combined gradient to give the best results. Moreover, I used a scalar weight that does not vary with iterations to calculate the combined gradient. Varying the weights with iterations could improve the results even more.

REFERENCES

- Almomin, A., 2011, Correlation-based wave-equation migration velocity analysis: SEP-Report, **143**, 33–42.

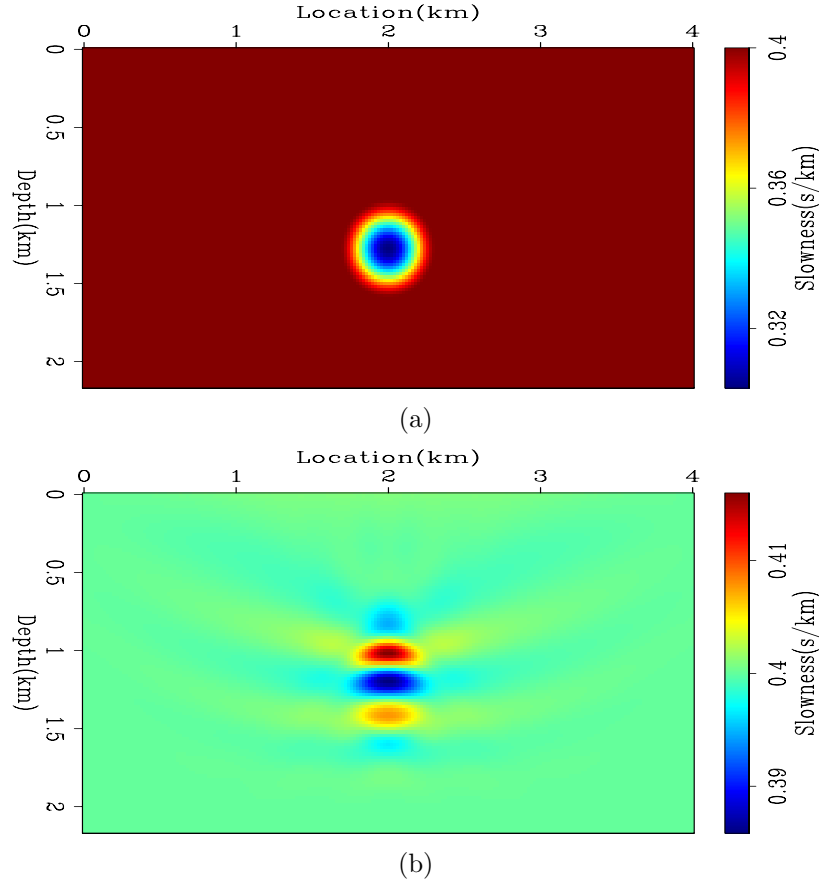


Figure 6: The true slowness model with a cosine perturbation (a) and the inversion results using low-frequency back-scattered FWI (b). [ER] [CR]

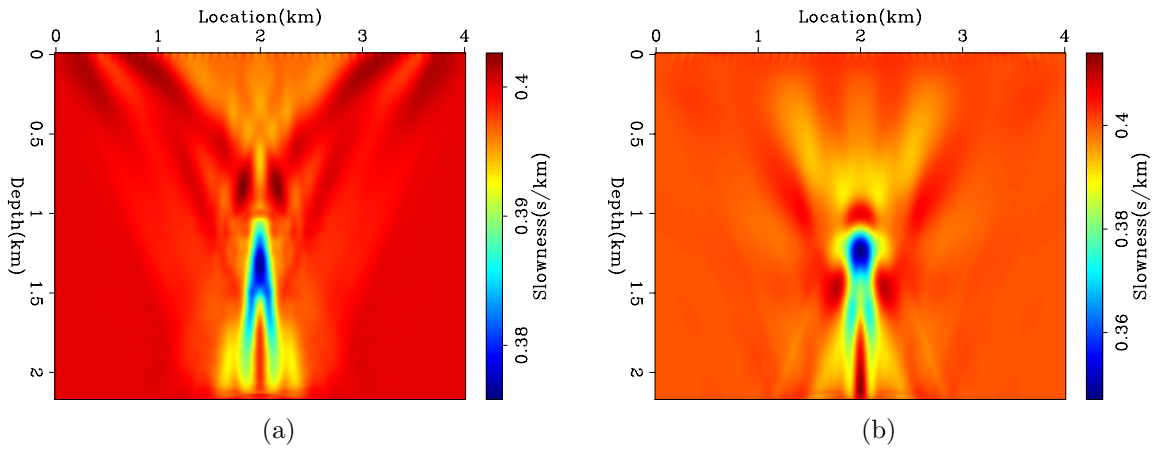


Figure 7: Slowness inversion results using (a) correlation-based WEMVA inversion and (b) the combined gradient. [CR]

- Almomin, A. and Y. Tang, 2010, Migration velocity analysis based on linearization of the two-way wave equation: SEP-Report, **142**, 13–24.
- Biondi, B., 2010, Wave-equation migration velocity analysis by residual moveout fitting: SEP-Report, **142**, 25–32.
- Biondi, B. and P. Sava, 1999, Wave-equation migration velocity analysis: SEG Technical Program Expanded Abstracts, **18**, 1723–1726.
- Hu, L.-Z. and G. A. McMechan, 1987, Wave-field transformations of vertical seismic profiles: Geophysics, **52**, 307–321.
- Liu, F., G. Zhang, S. A. Morton, and J. P. Leveille, 2007, Reverse-time migration using one-way wavefield imaging condition: SEG Technical Program Expanded Abstracts, **26**, 2170–2174.
- Luo, Y. and G. T. Schuster, 1991, Wave-equation travelttime inversion: Geophysics, **56**, 645–653.
- Shen, P., 2004, Wave equation migration velocity analysis by differential semblance optimization: PhD thesis, Rice University.
- Symes, W. W. and J. J. Carazzone, 1991, Velocity inversion by differential semblance optimization: Geophysics, **56**, 654–663.
- Tang, Y., C. Guerra, and B. Biondi, 2008, Image-space wave-equation tomography in the generalized source domain: SEP-Report, **136**, 1–22.
- Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266.
- Taweesintananon, K., 2011, Reverse-time migration using wavefield decomposition: SEP-Report, **143**, 171–190.