

# Long-wavelength FWI updates in the presence of cycle skipping

Jaime Ramos-Martínez<sup>1</sup>, Lingyun Qiu<sup>1</sup>, Alejandro A. Valenciano<sup>1</sup>, Xiaoyan Jiang<sup>1</sup>, and Nizar Chemingui<sup>1</sup>

<https://doi.org/10.1190/tle38030193.1>

## Abstract

Full-waveform inversion (FWI) has become the tool of choice for building high-resolution velocity models. Its success depends on producing seamless updates of the short- and long-wavelength model features while avoiding cycle skipping. Classic FWI implementations use the L2 norm to measure the data misfit in combination with a gradient computed by a crosscorrelation imaging condition of the source and residual wavefields. The algorithm risks converging to an inaccurate result if the data lack low frequencies and/or the initial model is far from the true one. Additionally, the model updates may display a reflectivity imprint before the long-wavelength features of the model are fully recovered. We propose a new solution to this fundamental challenge by combining the quadratic form of the Wasserstein distance (W2 norm) for measuring the data misfit with a robust implementation of a velocity gradient. The W2 norm reduces the risk of cycle skipping, whereas the velocity gradient effectively eliminates the reflectivity imprint and emphasizes the long-wavelength model updates. We illustrate the performance of the new solution on a field survey acquired offshore Brazil. We demonstrate how FWI successfully updates the earth model and resolves a high-velocity carbonate section missing from the initial velocity model.

## Introduction

Classic full-waveform inversion (FWI) (Tarantola, 1984) can produce velocity updates with reflectivity imprint before the long-wavelength components of the model are constructed. This is because the high wavenumbers provided by the reflections often dominate the inversion over the low-wavenumber updates (Mora, 1989). To mitigate the problem, practitioners follow different data selection strategies to separate diving waves from reflections. However, the separation in the data space can be challenging. This, and the fact that for deep targets only reflections are available for the inversion, has motivated the development of FWI gradients that isolate the different wavenumber ranges in the velocity updates (e.g., Xu et al., 2012; Zhou et al., 2015; Ramos-Martínez et al., 2016).

Moreover, the misfit function based on the L2 norm measures the difference between the recorded and modeled oscillatory signals on a point-by-point basis. This constrains FWI to use initial models that allow the simulation of waveforms within half the period of the recorded waveforms. If the initial model does not satisfy this condition, the inversion may suffer from cycle skipping and converge to the wrong velocity model. In practice, this limitation can be overcome by applying a laborious data selection strategy. The events with the nearest offsets and the lowest frequencies are inverted first. In subsequent stages, increased offset ranges and broader frequency bandwidths are considered. However, in many cases the acquired seismic data do not have

enough low frequencies to comply with the half-period condition. Moreover, in complex geologic settings such as salt environments, a small error in the location of the reflectors would lead to large kinematic errors. Hence, there is an incentive to use a different metric than the L2 norm for quantifying the data misfit (e.g., Engquist et al., 2016; Métivier et al., 2016; Qiu et al., 2017).

Here, we have combined the quadratic form of the Wasserstein distance (W2 norm) to measure the data misfit with a robust implementation of the velocity gradient. Our numerical implementation introduces dynamic weights (Ramos-Martínez et al., 2016) in the velocity sensitivity kernel derived from the impedance and velocity parameterization of the objective function. It effectively separates the migration isochrones produced by the specular reflectivity from the components produced by transmitted arrivals. Our implementation (Qiu et al., 2017) uses an encoding scheme based on a logistic function that assures the positiveness and mass conservation conditions required by the optimal-transport theory.

## Theory

FWI is formulated as a nonlinear inverse problem matching modeled data to recorded field data (Tarantola, 1984). Generally, a least-squares objective function is used for measuring the data misfit. Here, we measure the data difference using the W2 norm:

$$J = \sum_s \sum_r W_2^2(\tilde{u}, \tilde{d}), \quad (1)$$

where  $\tilde{u}(t)$  and  $\tilde{d}(t)$  are encoded versions of the modeled and field data. The W2 norm and the resulting Frechet derivative are explained in Qiu et al. (2017). We use the logistic function to encode both the field and modeled data.

To produce long-wavelength velocity updates, we adapted the equations for the velocity gradient (Ramos-Martínez et al., 2016) to work with the W2 misfit function. The dynamic weight implementation of the velocity kernel was translated to the equivalent expressions such that the first-order time derivatives of the source and residual wavefields are computed before the adjoint source backward propagation. The resulting velocity kernel has the form.

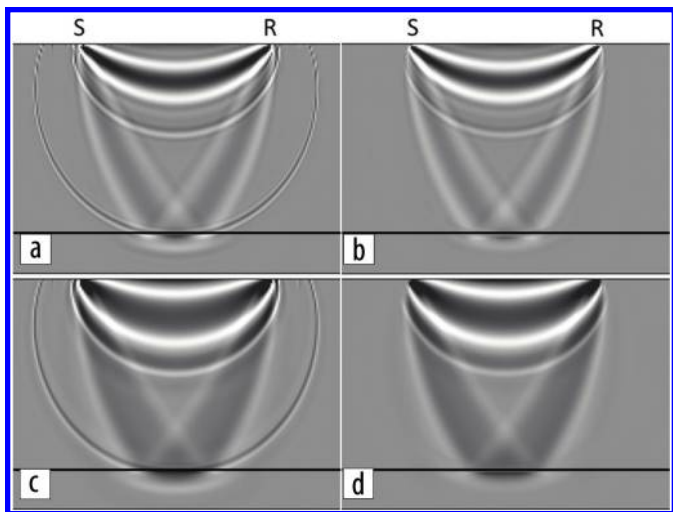
$$G_v(\mathbf{x}) = \frac{1}{A(\mathbf{x})} \int \left[ \begin{aligned} &W_1(\mathbf{x}, t) \frac{1}{V(\mathbf{x})^2} S(\mathbf{x}, t) R(\mathbf{x}, T-t) \\ &+ W_2(\mathbf{x}, t) \nabla S(\mathbf{x}, t) \nabla \hat{R}(\mathbf{x}, T-t) \end{aligned} \right] dt, \quad (2)$$

where  $S(\mathbf{x}, t)$  is the source wavefield, and  $R(\mathbf{x}, t)$  is the wavefield computed from the adjoint-state equation using the time reversal of a preconditioned adjoint source.  $\hat{R}(\mathbf{x}, T-t)$  results from applying

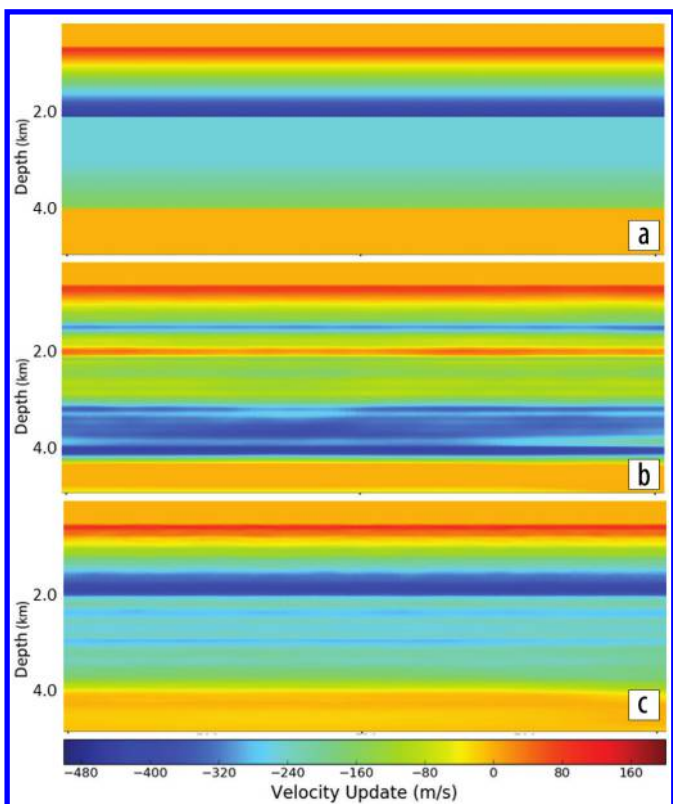
<sup>1</sup>PGS, Houston, Texas, USA. E-mail: [jaime.ramos@pgs.com](mailto:jaime.ramos@pgs.com); [lingyun.qiu@pgs.com](mailto:lingyun.qiu@pgs.com); [alejandro.valenciano.mavilio@pgs.com](mailto:alejandro.valenciano.mavilio@pgs.com); [xiaoyan.jiang@pgs.com](mailto:xiaoyan.jiang@pgs.com); [nizar.chemingui@pgs.com](mailto:nizar.chemingui@pgs.com).

the inverse of the preconditioning operator to the adjoint wavefield.  $A(\mathbf{x})$  is the illumination term. The dynamic weights  $W_i(\mathbf{x}, t)$  are designed to optimally suppress the unwanted specular reflectivity (migration isochrones).

Figure 1 shows the kernels for different combinations of the L2 and W2 norms and FWI gradients. They were computed for a source-receiver pair in a layer where velocity increases with depth. Notice that the W2 velocity kernel accentuates the long-wavelength components compared to the L2 velocity kernel.



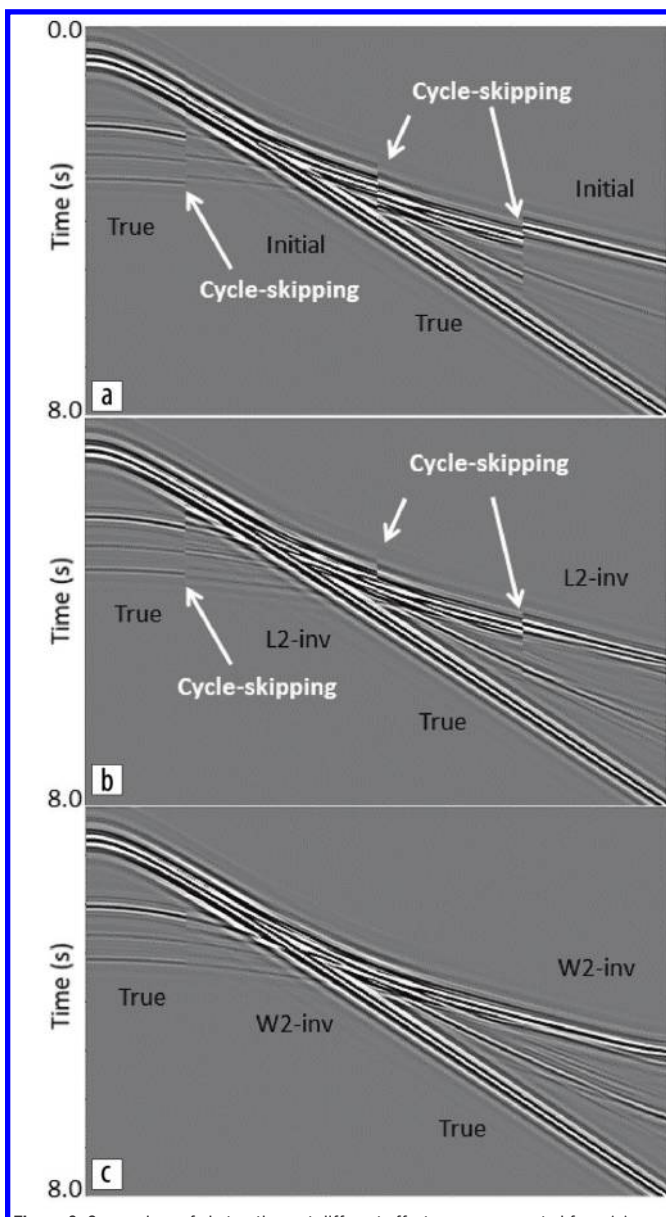
**Figure 1.** Sensitivity kernels of a source-receiver pair in a model with a  $v(z)$  layer overlaying a half-space for (a) the L2 norm and crosscorrelation FWI gradient, (b) the L2 norm and FWI velocity gradient, (c) the W2 norm and crosscorrelation gradient, and (d) the W2 norm and velocity gradient.



**Figure 2.** (a) Velocity difference between the true and starting models. (b) Velocity updates using the L2 norm. (c) Velocity updates using the W2 norm.

### Synthetic example

We illustrate the advantages of using the W2 norm on a synthetic model consisting of four horizontal layers overlaying a half-space. The first layer is a water column of 0.5 km depth. The other three layers have thicknesses of 1.5, 1.0, and 1.0 km, respectively, with different depth-varying velocities. The inversion utilizes both transmitted and reflected events in a frequency band between 3 and 13 Hz. The differences between the true and starting velocity models (Figure 2a) are large enough to produce cycle skipping for the reflections from the second, third, and fourth interfaces. The cycle skipping is observed when comparing the modeled shot gathers using the true and initial velocity models (Figure 3a) and clearly noticeable for the long-offset transmitted events. FWI velocity updates using the different norms are shown in Figures 2b and 2c. The inversion with the L2 norm (Figure 2b) provides correct updates in the shallow



**Figure 3.** Comparison of shot gathers at different offset ranges computed from (a) the true and initial model, (b) the true and the inverted model using the L2 norm, and (c) the true and the inverted model using the W2 norm. Maximum offset is 12 km.

part of the model, but the solution diverges in the deeper part. In contrast, the W2 norm (Figure 2c) produces a good solution throughout the whole model. Figures 3b and 3c show comparison of the seismograms. Overall, the W2 norm provides the best result with the correct phase of the events.

### Field data example

We applied the new FWI algorithm to a field survey acquired in the Ceará Basin, offshore Fortaleza, Brazil. The acquisition comprises 14 multisensor streamers separated by 100 m, with a maximum inline offset of 8 km. The data show good signal-to-noise ratio at frequencies as low as 2.5 Hz. The maximum frequency used in the inversion was 8 Hz in a window containing transmitted and reflected events.

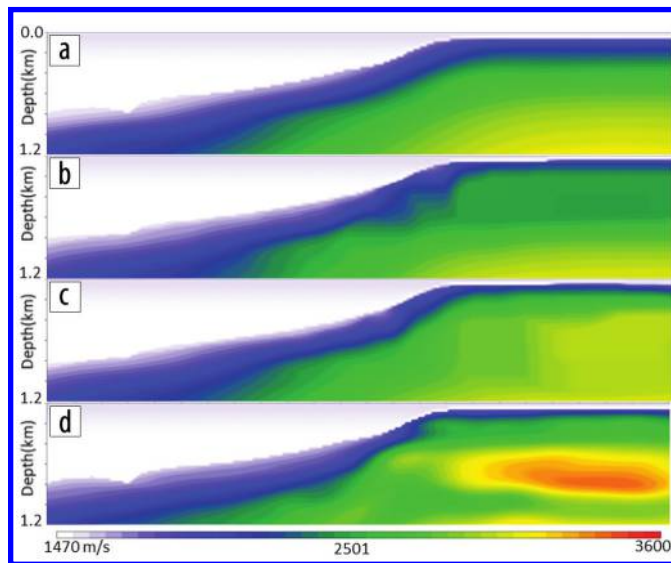
The legacy velocity model (Figure 4a) poorly resolved a package of near-seafloor carbonates in the shallow. The absence of the carbonates caused deeper distortions of the seismic image (Figure 6a) as well as cycle skipping (Figure 5a). The shallow water depth (about 50 m) limited the use of reflection tomography for updating the shallow layers. Meanwhile, the high-velocity carbonate layer (around 3300 m/s from nearby well log) limits the penetration of the refracted energy to a depth of approximately 1.2 km.

Figures 4b and 4c show the inverted models obtained in a first stage using the L2 and W2 norms. In both cases, we used the velocity gradient to minimize the high-wavenumber artifacts produced by the multiples. The L2-norm inversion produced an update in the wrong direction due to cycle skipping. In contrast, the W2 norm yielded an increase in velocity where the carbonates are expected.

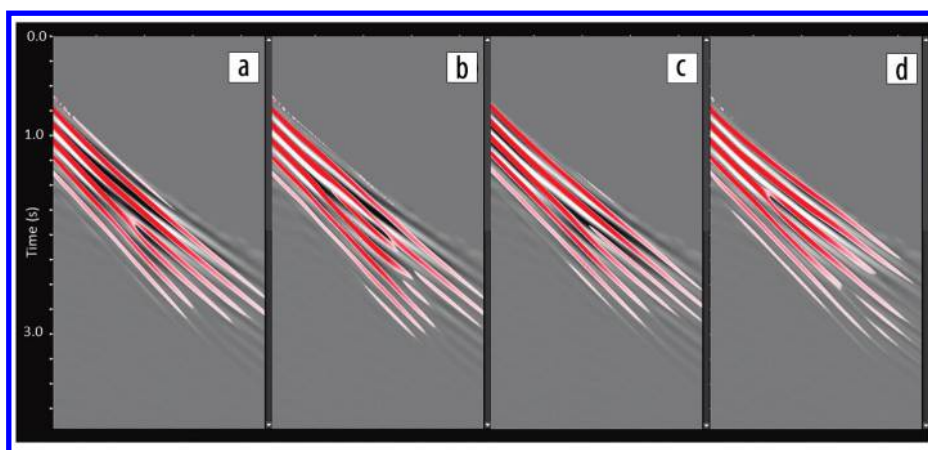
We validated the FWI results by comparing synthetic and recorded shot gathers (Figure 5). The positive amplitudes of the synthetic traces (displayed in red) should match the corresponding positive amplitudes of the recorded traces (displayed in black) if the model is accurate. Figure 5a shows a recorded shot gather overlaid by the synthetic using the initial model and clearly shows the cycle skipping at intermediate and long offsets. Figure 5b shows the comparison using the L2-norm inverted model. The match of near offsets improves, but the intermediate and far offsets suffer from cycle skipping. In contrast, the W2-norm synthetics (Figure 5c) show better match at all offset ranges.

After resolving the cycle-skipping problem we continued the inversion using the L2-norm FWI (Figure 4d). The high-velocity carbonates are now resolved after eliminating the cycle skipping, and the match between field and modeled traces is good for all offsets (Figure 5d).

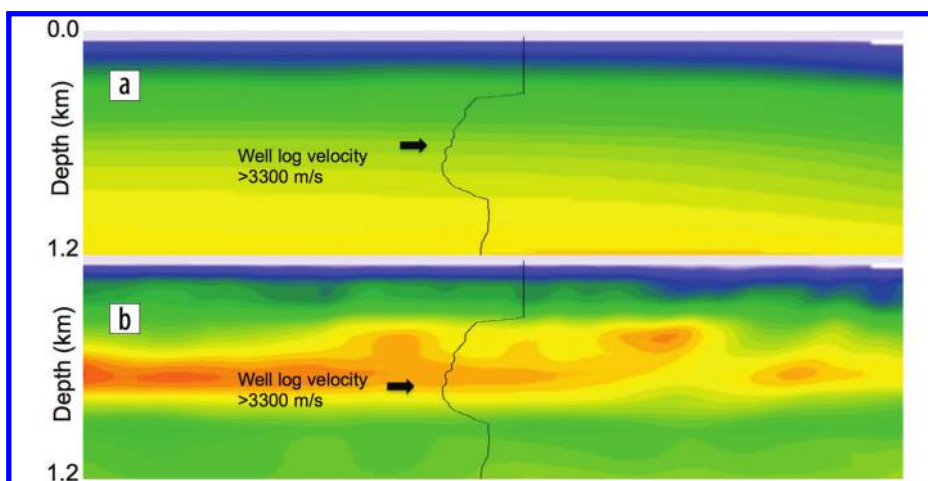
The velocity increase was corroborated with the well log data. Figures 6a and 6b show the starting and final FWI models for a line in the proximity of the well. The final FWI model matches



**Figure 4.** Vertical velocity for the field survey: (a) initial model, (b) FWI model using the L2 norm, (c) FWI model using the W2 norm, and (d) FWI model produced by cascading the W2 norm and L2 norm.



**Figure 5.** Comparison of field records (grayscale) and synthetic shot gathers computed with (a) initial model, (b) FWI model using the L2 norm, (c) FWI model using the W2 norm, and (d) the W2 norm followed by the L2 norm.



**Figure 6.** Comparison at the proximity of the well: (a) starting model and (b) final FWI model.

the well trend capturing the spatial variability of the carbonates.

Finally, reflection tomography was applied to update the deeper part of the model. Figures 7a–7c show the Kirchhoff-migrated images overlaying their corresponding velocity models. Note the improvement in the continuity of the deeper reflectors from the initial to final models.

## Conclusions

We combine a robust implementation of velocity gradient and optimal-transport norm (W2) to overcome the FWI cycle-skipping problem and retrieve the long-wavelength velocity updates. The proposed solution expands the use of FWI for velocity model building since it reduces its dependency on accurate starting velocity models and therefore relaxes the requirements for long offsets and low-frequency data. We illustrated the advantages of the solution on a synthetic example and a field survey where the algorithm resolved high-velocity carbonates that were missing from the initial model. The final velocity model resulted in an improved image of the survey including both shallow and deep structures. **11E**

## Acknowledgments

We thank Mikhail Orlovich for his assistance in the field data example, and Yunan Yang, Dan Whitmore, and Faqi Liu for helpful discussions. We also thank PGS for permission to publish this work.

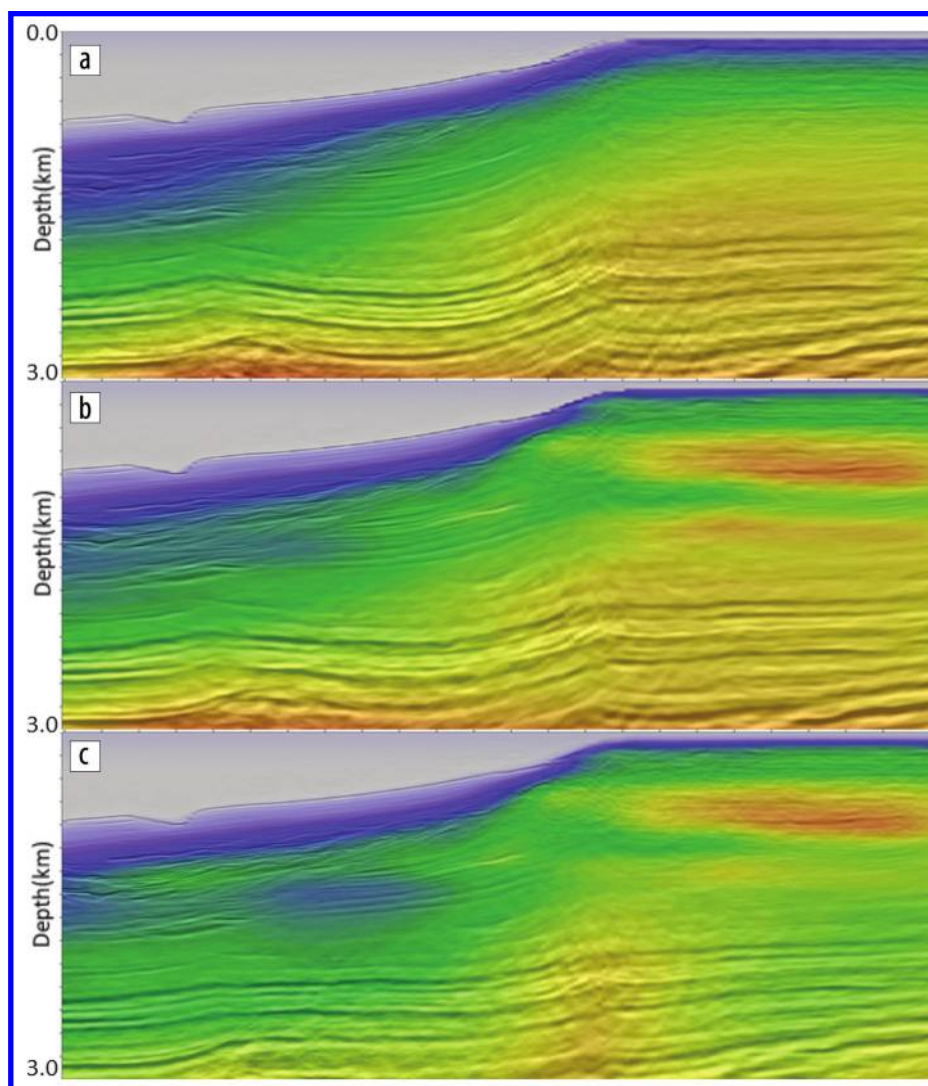
## Data and materials availability

Data associated with this research are confidential and cannot be released.

Corresponding author: nizar.chemingui@pgs.com

## References

- Engquist, B., B. D. Froese, and Y. Yang, 2016, Optimal transport for seismic full waveform inversion: *Communications in Mathematical Sciences*, **14**, no. 8, 2309–2330, <https://doi.org/10.4310/CMS.2016.v14.n8.a9>.
- Métivier, L., R. Brossier, Q. Mérigot, E. Oudet, and J. Vireux, 2016, Measuring the misfit between seismograms using an optimal transport distance: Application to full waveform inversion: *Geophysical Journal International*, **205**, no. 1, 345–377, <https://doi.org/10.1093/gji/ggv014>.
- Mora, P., 1989, Inversion = migration + tomography: *Geophysics*, **54**, no. 12, 1575–1586, <https://doi.org/10.1190/1.1442625>.



**Figure 7.** Kirchhoff-migrated images overlaying the velocity model: (a) starting model, (b) final FWI model, and (c) model after deeper tomography. Note the improvement in the continuity of the deeper reflectors from initial to final model.

- Qiu, L., J. Ramos-Martinez, A. Valenciano, Y. Yang, and B. Enquist, 2017, Full-waveform inversion with an exponentially encoded optimal-transport norm: 87<sup>th</sup> Annual International Meeting, SEG, Expanded Abstracts, 1286–1290, <https://doi.org/10.1190/segam2017-17681930.1>.
- Ramos-Martinez, J., S. Crawley, Z. Zou, A. A. Valenciano, L. Qiu, and N. Chemingui, 2016, A robust gradient for long wavelength FWI updates: 76<sup>th</sup> Conference and Exhibition, EAGE, Extended Abstracts, <https://doi.org/10.3997/2214-4609.201601536>.
- Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: *Geophysics*, **49**, no. 8, 1259–1266, <https://doi.org/10.1190/1.1441754>.
- Xu, S., D. Wang, F. Chen, Y. Zhang, and G. Lambaré, 2012, Full waveform inversion for reflected seismic data: 74<sup>th</sup> Conference and Exhibition, EAGE, Extended Abstracts, <https://doi.org/10.3997/2214-4609.20148725>.
- Zhou, W., R. Brossier, S. Operto, and J. Vireux, 2015, Full waveform inversion of diving waves for velocity model building with impedance inversion based on scale separation: *Geophysical Journal International*, **202**, no. 3, 1535–1554, <https://doi.org/10.1093/gji/ggv228>.