# Tomographic full waveform inversion (TFWI) by combining full waveform inversion with wave-equation migration velocity analysis

*Biondo Biondi and Ali Almomin*

## ABSTRACT

By extending the velocity-model domain to subsurface offsets we solve the local-minima problem of data-fitting waveform inversion. We then regularize the extended-model data-fitting inversion with the addition of an image-focusing term to the objective function, therefore achieving robust global convergence of the waveform inversion problem. The method shares with full waveform inversion the advantage of simultaneously solving for all the wavelengths of the model, but it also has the global convergence characteristics of wave-equation migration velocity analysis. Numerical implementation of the proposed inversion method requires the solution of an extended wave-equation where velocity is a convolutional, instead of scalar, operator. The resulting method is therefore computationally intensive, but it can be easily tested in 2D. A simple example with a Gaussian velocity anomaly illustrates how the reflections from the anomaly recorded in the low-frequency components of the data increase the spatial resolution of the final inversion results. Numerical tests performed on synthetic data from a modified Marmousi model demonstrate the global convergence as well the high-resolution potential of the method.

## INTRODUCTION

Conventional seismic imaging relies on a separation of scales between migration velocity model (long-wavelength components) and reflectivity (short-wavelength components). The migration velocity model is estimated first, and then it is used as input to migration for imaging reflectivity. Even when wave-equation operators are employed to estimate the velocity, such as in wave-equation migration velocity analysis (WEMVA) methods, reflectivity is used only indirectly to measure the focusing power of the velocity model (Biondi and Sava, 1999; Shen and Symes, 2008). The only important exceptions in current practice occur when migrated volumes are used: 1) to interpret boundaries of geobodies (e.g. salt bodies), whose interior are assigned predefined velocities, and 2) to estimate predominant dips in the geologic layering that are then used to constraint a tomographic velocity updating (Clapp et al., 2004).

As the industry strives to widen the data frequency band at both the low and high end, the advantages of overcoming the limitations of conventional imaging, and of

exploiting reflectivity information for velocity estimation, are becoming more relevant to important imaging problems. One of the main attractions of full waveform inversion (FWI) (Tarantola, 1987; Pratt, 1999) is to overcome the limitations imposed by the conventional approach that may limit the quality of the imaging results by finding a suboptimal solution. However, FWI suffers from well-known convergence problems when the starting model is far from the correct one and low frequencies are missing from the data.

We discuss an inversion framework that overcomes FWI difficulties by supplementing an FWI-like data-fitting objective function with a WEMVA-like term that measures the reflection-focusing power of the velocity model. The method fits the recorded data with data modeled using a generalized version of the acoustic wave equation; the domain of velocity model is extended to include subsurface offsets. The extension of the reflectivity along the subsurface offset axes (or reflection angles) is a well-established technique for migration, linearized waveform inversion, and WEMVA (Biondi, 2006). In a data-fitting inversion, extending reflectivity to the prestack domain has the critical advantage that the kinematics of the modeled data will not be too distant from the ones of the recorded data, no matter the magnitude of the background velocity error.

Symes (2008) introduced the idea of using a wave equation with an extended velocity. By extending velocity the convergence difficulties of conventional FWI are overcome and all scales can be solved simultaneously. In the same paper, he also introduced the waveform-inversion formulation used in this paper, and described its application to the solution of 1D inversion problem in presence of multiple reflections.

The main goal of our research is not to tackle the problem of multiples, but to perform simultaneous inversion for all scales of the velocity model. Therefore, we apply the theory and numerically solve the extended wave equation in 2D. We also derive an effective scheme to linearize the extended wave equation and to compute the gradient of the objective-function by an adjoint-state method. In 3D the proposed method would be extremely expensive. In a companion report, (Almomin and Biondi, 2012) we present an approximation to the method presented here that drastically reduces the computational cost, but still retains the capability of simultaneously solving for all the wavelengths of the velocity model.

Another potential problem with strict coupling of velocity with reflectivity arises when the assumption of constant density cannot be made, as is the case in most of field data problems. In this case density variations may create reflections that do not correspond to velocity contrasts. However, we still would like to avoid the addition of density to the problem parameters for computational and convergence reasons. The approximate method proposed in Almomin and Biondi (2012) has the potential of being more flexible in accommodating these discrepancies.

# TOMOGRAPHIC FULL WAVEFORM INVERSION (TFWI)

The conventional FWI objective function $J_{\text{FWI}}$ can be written as:

$$J_{\text{FWI}}(\mathbf{v}) = \|\mathbf{d}(\mathbf{v}) - \mathbf{d}_{\text{obs}}\|_2^2, \tag{1}$$

where $\mathbf{v}$ is the velocity model, $\mathbf{d}(\mathbf{v})$ is the modeled data, and $\mathbf{d}_{\text{obs}}$ is the observed data.

The modeled data is computed as:

$$d(\mathbf{x}_s, \mathbf{x}_r, \omega; \mathbf{v}) = f(\mathbf{x}_s, \omega) G(\mathbf{x}_s, \mathbf{x}, \omega; \mathbf{v}) \delta(\mathbf{x}_r - \mathbf{x}), \tag{2}$$

where $f(\mathbf{x}_s, \omega)$ is the source function, $\omega$ is frequency, $\mathbf{x}_s$ and $\mathbf{x}_r$ are the source and receiver coordinates, and $\mathbf{x}$ is the model coordinate. In the acoustic, constant-density case the Green's function $G(\mathbf{x}_s, \mathbf{x}, \omega; \mathbf{v})$ satisfies:

$$\left(\frac{\omega^2}{v^2(\mathbf{x})} + \nabla^2\right) G(\mathbf{x}_s, \mathbf{x}, \omega; \mathbf{v}) = \delta(\mathbf{x}_s - \mathbf{x}). \tag{3}$$

For the sake of compact notation, in the rest of the paper we present the expressions for computing the data and the gradient of objective functions in the frequency domain. However, we perform the computation in the time domain.

We can extend the velocity in the subsurface-offset dimension $\mathbf{h}$ which changes the wave equation into the following form

$$\left(v^2(\mathbf{x}, \mathbf{h}) *^{-1} \omega^2 + \nabla^2\right) G(\mathbf{x}_s, \mathbf{x}, \omega; \mathbf{v}) = \delta(\mathbf{x}_s - \mathbf{x}), \tag{4}$$

where with $*^{-1}$ we indicate deconvolution. Notice that the division by velocity in equation 4 becomes a deconvolution over the offset axis. Once we define the Green's function, the data could be computed similarly to equation 2. We now write the new objective function as follows:

$$J_{\text{EFWI}}\left(\mathbf{v}(\mathbf{h})\right) = \|\mathbf{d}\left(\mathbf{v}(\mathbf{h})\right) - \mathbf{d}_{\text{obs}}\|_2^2 . \tag{5}$$

The long-wavelength components of the solution of the optimization problem defined by equation 5 are not likely to be substantially different from the long-wavelength components of the initial model. The extension of the model, and in particular of its reflectivity component, to non-zero subsurface offset causes the kinematics of the modeled data to match the kinematics of the recorded data independently from the accuracy of the long-wavelength components.

Another term must be added to the objective function to drive the solution towards a model that focuses the image. Symes (2008) suggests the addition of a differential semblance penalty function (DSO); that is,

$$J_{\text{DSO}}(\mathbf{v}(\mathbf{h})) = \||\mathbf{h}|\mathbf{v}(\mathbf{h})\|_2^2. \tag{6}$$

We use this focusing term in the numerical experiments described in this paper. Another valid choice would be the maximization of the normalized power of the stack over reflection angles as a function of a residual moveout parameter $\rho$, as suggested by Zhang and Biondi (2012).

The important characteristic of the second term is that its gradient "imposes" on the current model only a phase shift, and not a bulk vertical shift. This assures that the corresponding perturbations on the modeled data are mere phase shifts, and not bulk time shifts. The absence of bulk time shifts in the modeled data avoids large discrepancies between the kinematics of modeled data and recorded data. These large discrepancies are at the root of the convergence problems in conventional FWI.

Another practically important consideration is that in the proposed formulation the computation of the gradient of a term like the one presented in Zhang and Biondi (2012) is straightforward because it does not require back-projection of image perturbations. This is in contrast with WEMVA-like methods, where the computation of the gradient must take into account the constraint that the image is the result of migrating the recorded data. Therefore, at least in principle, it would be equally easy to add to the objective function other terms that reward focusing of the model along the midpoint spatial axis, in addition to the subsurface offset or reflection angle (Biondi, 2010).

## GRADIENT COMPUTATION

To compute the gradient of the objective function expressed in equation 5, we need to linearize the extended wave equation 4. Usually equation 3 is linearized over slowness or velocity. However, the extended wave equation 4 includes a deconvolution over the offset axis that is not easy to implement. Hence, we will first rearrange equation 4 to a form that facilitates the computation of the gradient, and that is actually solved numerically in the propagation. This can be achieved by convolving both sides of the equation by the square of velocity then rearranging the terms as follows:

$$\omega^2 G(\mathbf{x}_s, \mathbf{x}, \omega, \mathbf{v}) = v^2(\mathbf{x}, \mathbf{h}) * \left(\delta(\mathbf{x}_s - \mathbf{x}) - \nabla^2 G(\mathbf{x}_s, \mathbf{x}, \omega, \mathbf{v})\right).$$

We now can linearize the relationship between the Green's function and the model by perturbing the model around a background value as follows:

$$v^2(\mathbf{x}, \mathbf{h}) = v_0^2(\mathbf{x}, \mathbf{h}) + \Delta v^2(\mathbf{x}, \mathbf{h}),$$

where $v_0(\mathbf{x}, \mathbf{h})$ is the background component and $\Delta v(\mathbf{x}, \mathbf{h})$ is the perturbation component, i.e. the model update. After this separation, the first-order Born approximation can be used to define the gradient as follows:

$$\begin{aligned} g_d(\mathbf{x}, \mathbf{h}) = & \\ & \sum_{\mathbf{x}_s, \mathbf{x}_r, \omega} \left[\nabla^2 f(\mathbf{x}_s, \omega) G\left(\mathbf{x}_s, \mathbf{x} - \mathbf{h}, \omega; \mathbf{v}_0(\mathbf{h})\right)\right. \\ & \left. G\left(\mathbf{x}_r, \mathbf{x} + \mathbf{h}, \omega; \mathbf{v}_0(\mathbf{h})\right)\ \Delta d^*(\mathbf{x}_s, \mathbf{x}_r, \omega; \mathbf{v}_0(\mathbf{h}))\right], \end{aligned}$$

where $\Delta d$ are the data residuals and $^*$ indicates the complex conjugate. Unlike the usual linearization of the wave equation, the scattering term includes a Laplacian operator instead of the second derivative in time. The Laplacian operator does not add to the computational cost since it is already computed in the propagation of the background wavefield.

# SYNTHETIC DATA EXAMPLES

We tested the proposed inversion method with two different synthetic data sets. The first one was generated assuming a simple model where the goal was to estimate a Gaussian positive velocity anomaly. The second data set was generated assuming a slightly modified Marmousi model.

## Gaussian Anomaly

To illustrate the interplay between different scales of the velocity during the inversion we applied the proposed method to a two-layer model with a Gaussian anomaly in the middle of the first layer, as shown in Figure 1. The velocity of the top layer is 3 km/s and of the bottom layer is 3.5 km/s. The velocity at the center of the Gaussian anomaly is 3.5 km/s. The interface between the layers is at 2 km depth and the Gaussian anomaly is centered at 1 km depth.

A Ricker wavelet with a fundamental frequency of 15 Hz and temporal sampling of 1.5 ms was used as a source function to model the data. The wavefields were generated by 31 sources with a spacing of 100 m and recorded by 151 fixed receivers with a spacing of 20 m. The maximum offset is 1.5 km. The initial model is a constant model of 3 km/s velocity.

Figure 2 shows the difference between the velocity model estimated after one iteration and the starting model. As expected, the reflector is well imaged, though not perfectly focused under the velocity anomaly. Figure 3 shows a rescaled window of Figure 2 around the anomaly. It shows how the reflections from the anomaly measured from the low-frequency components of the data start outlining the contour of the anomaly. Figure 4 shows the difference between the initial and the inverted model after 2000 iterations. Figure 5 shows a rescaled window of Figure 4 around the anomaly. The anomaly is now fairly well focused. However, the maximum amplitude of the estimated anomaly is still a fraction of the amplitude of the true anomaly, but better recovered than if it had been estimated using conventional velocity estimation methods based only on the transmission effects (Almomin and Biondi, 2012).

Figure 6 shows three sections of the difference cube obtained by subtracting the starting model from the model obtained after one iteration. These sections are taken at fixed horizontal coordinates and are functions of depth and subsurface offset. These images are analogous to migration subsurface-offset common image gathers; they

show the lack of focusing of the model along the subsurface-offset axis. Figure 7 shows the gathers taken at the same locations as the ones shown in Figure 6, but after 2000 iterations. The gathers are now well focused around the zero-offset axis, indicating that the final model well explains the kinematic effects of the reflections that propagated through the velocity anomaly. Indeed, the gathers shown in Figure 7 appear artificially focused around zero subsurface offset. This appearance is caused by the DSO term in the objective function forcing the focusing of the image even beyond what would have been the focusing with the true model. In the data domain this "extra" focusing causes extrapolation of events beyond the recorded offsets (Clapp, 2005).

Figure 8 shows the normalized data residual as a function of iterations. The residual has not completely not flatten out and converged to zero; further iterations would improve the results.

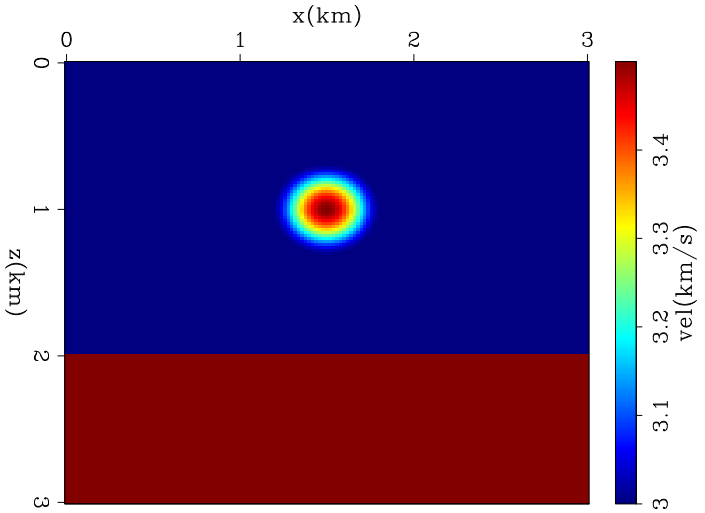


Figure 1: The true velocity model. [**ER**]

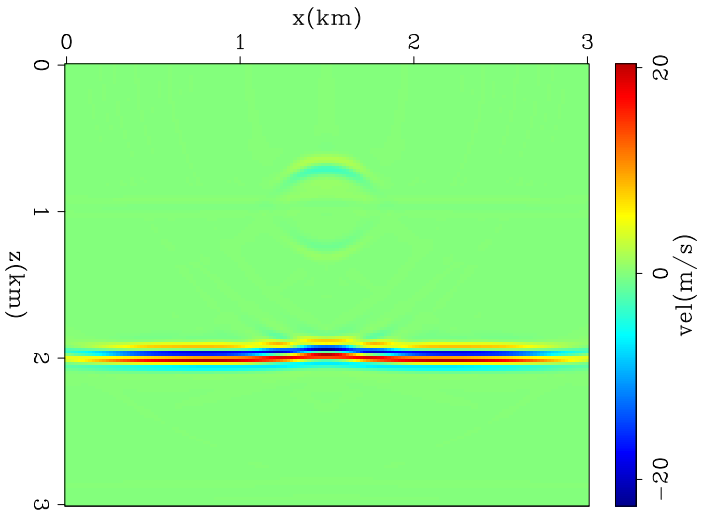


Figure 2: The difference between inverted and initial model after one iteration. [**CR**]

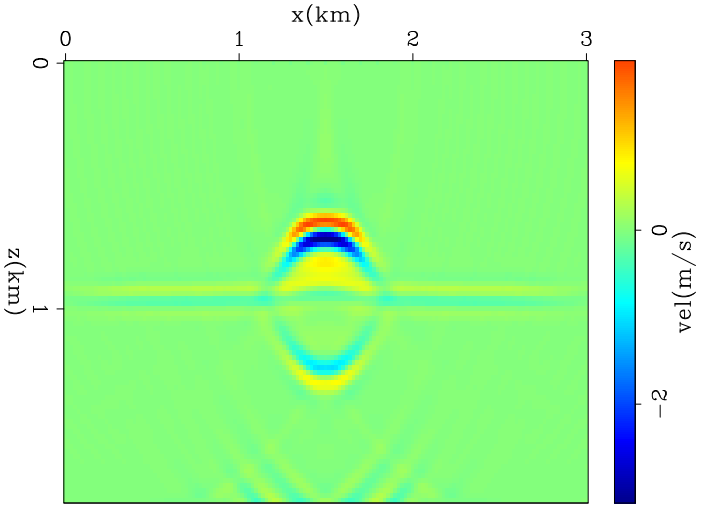


Figure 3: Zoom of the difference between inverted and initial model after one iteration. Notice the slight vertical stretch with respect to Figure 2. [**CR**]

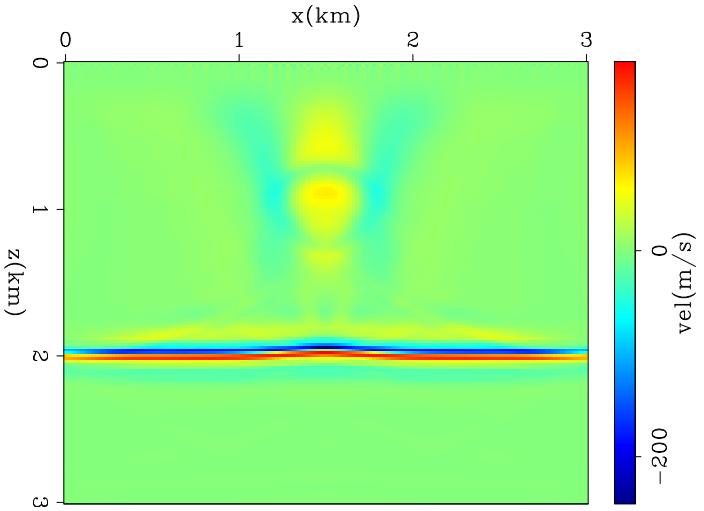


Figure 4: The difference between inverted and initial model after 2000 iterations. [**CR**]

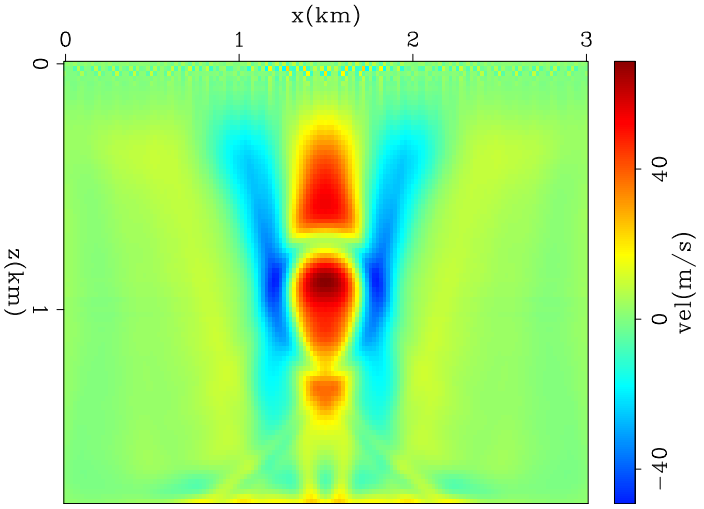


Figure 5: Zoom of the difference between inverted and initial model after 2000 iterations. Notice the slight vertical stretch with respect to Figure 4. [**CR**]

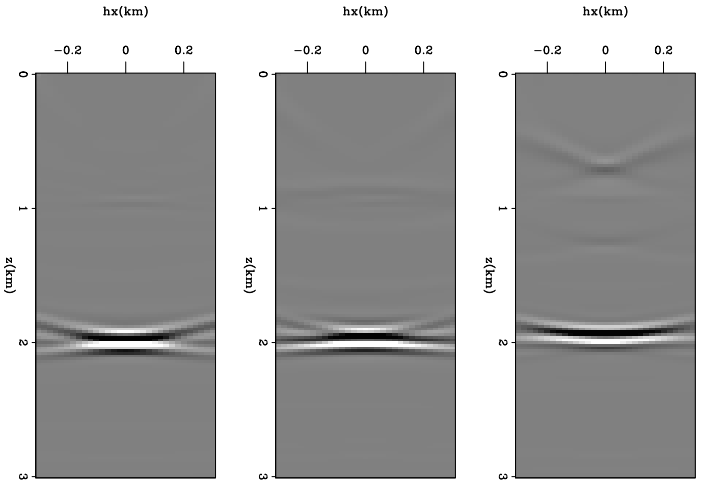


Figure 6: The difference between inverted and initial model after one iteration. These sections were taken at x=.09, 1.2, and 1.5 km. [**CR**]

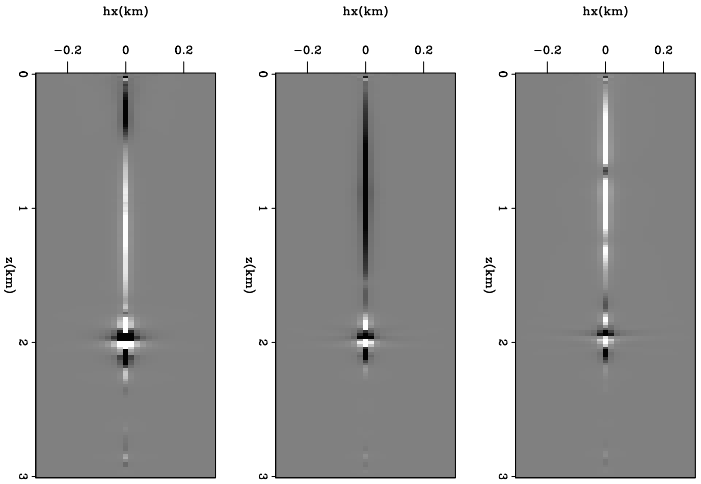


Figure 7: The difference between inverted and initial model after 2000 iterations. These sections were taken at x=.09, 1.2, and 1.5 km. Compare with Figure 6. [**CR**]

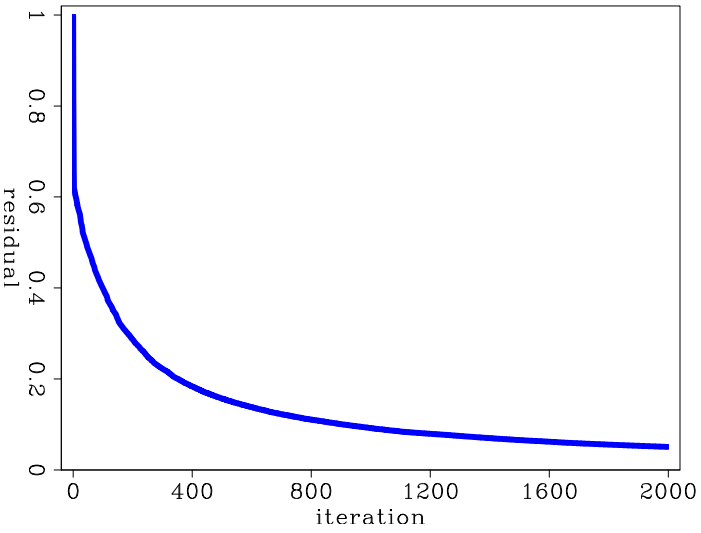


Figure 8: The normalized residual as a function of the number of iterations. [**CR**]

## Marmousi model

To test the convergence and resolution characteristics of the TFWI method, we applied it to the inversion of a synthetic data set generated assuming the modified Marmousi model shown in Figure 9. This model is the classical Marmousi with the addition of a 500 m water layer. The data were modeled assuming 93 shots with 100 m spacing and a Ricker wavelet with fundamental frequency of 15 Hz. The reflected wavefields were recorded by 461 receivers with fixed spread and 20 m spacing.

We compare the results of three inversion methods starting from the model shown in Figure 10, which was obtained by applying strong horizontal smoothing to the true model. Figure 11 shows the model obtained with conventional FWI; that is, by minimizing the objective function expressed in equation 1. As expected, conventional FWI fails because of its well-known difficulties with global convergence. The resulting model is almost identical to the starting one, with only a few shallow velocity discontinuities being imaged.

Figure 12 and 15 show the results of extended FWI (EFWI); that is, the minimization of the objective function expressed in equation 5. Figure 12 shows the zero subsurface offset section, displayed in color to facilitate the analysis of the long-wavelength components of the velocity. These components are similar in the final model and in the initial model. However, in contrast with the simple FWI, reflectivity is now imaged across the section, though misplaced and only partially focused because of the persistent errors in the long-wavelength components of the velocity. Figure 15 shows three sections of the difference cube obtained by subtracting the starting model from the final model. These sections are taken at fixed horizontal coordinates and are functions of depth and subsurface offset. These images are analogous to migration subsurface-offset common image gathers; they show the lack of focusing of the model along the subsurface-offset axis.

Figure 13 and 16 show the results of the proposed tomographic full waveform inversion (TFWI). To compute these results we minimized an objective function defined by the sum of equation 5 and equation 6. The zero subsurface-offset section (Figure 13) shows remarkable convergence towards the true model, in particular in the middle of the section where the model is well illuminated by the modeled data. The offset-domain common image gathers shown in Figure 16 confirm that the final model correctly focuses the reflected events. Similarly to the previous example (Gaussian anomaly), the gathers shown in Figure 16 appear artificially focused around zero subsurface offset. As discussed previously, this appearance is caused by the DSO term in the objective function forcing the focusing of the image even beyond what would have been the focusing with the true model.

Finally, Figure 14 shows normalized values of the norm of the data residuals as a function of iterations for all three methods we discussed: FWI in blue, EFWI (i.e. minimizing only objective function 5) in red and TFWI (i.e. minimizing sum of objective functions 5 and 6) in magenta. Notice that in the case of TFWI, even if the objective function had two terms, only the value of $J_{\mathrm{EFWI}}$ (i.e. the data fitting

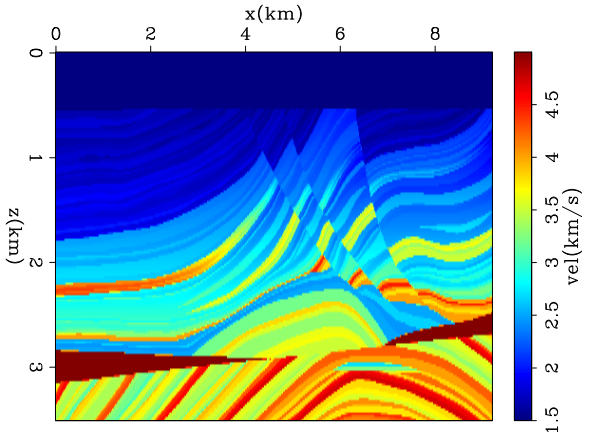


Figure 9: Modified Marmousi model used for numerical tests. [**ER**]

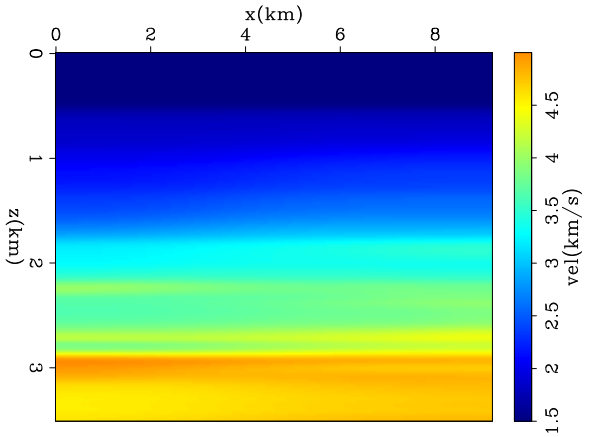


Figure 10: Starting model for inversions. [**ER**]

term) is plotted in the graph. The graphs show that FWI has not converged even after hundreds of iterations. The data residuals decrease more quickly for EFWI than for TFWI because EFWI does not need substantial changes in the long-wavelength components of the model to fit the data. However, the magnitude of the final residuals is comparable between the two methods.

## CONCLUSIONS

We presented and tested a method to invert for all wavelengths of the velocity model simultaneously from seismic reflection data. The method does not suffer from the global convergence limitations of conventional full waveform inversion. The method is based on the extension of the velocity model to subsurface offsets and on the addition of a reflection focusing objective function to the conventional FWI data-fitting objective function. The Marmousi model numerical tests indicate that the method can robustly converge to very accurate models when FWI fails. The numerical solution of the wave equation with the extended velocity is computationally expensive. The excellent results, and the computational expense, justify the search of an approximate solution that still enables the simultaneous solution for all wavelengths in the model. We present such approximation in a companion report (Almomin and Biondi,

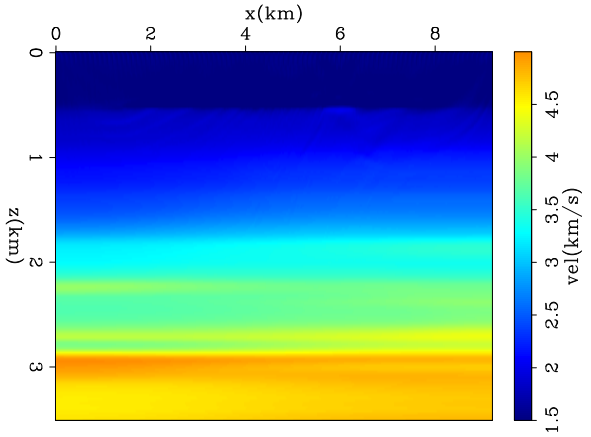


Figure 11: Final model when minimizing objective function 1 (FWI). [**CR**]

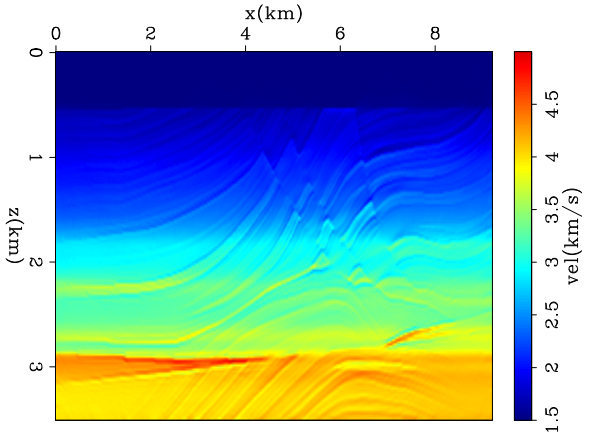


Figure 12: Final model when minimizing objective function 5 (EFWI). [**CR**]

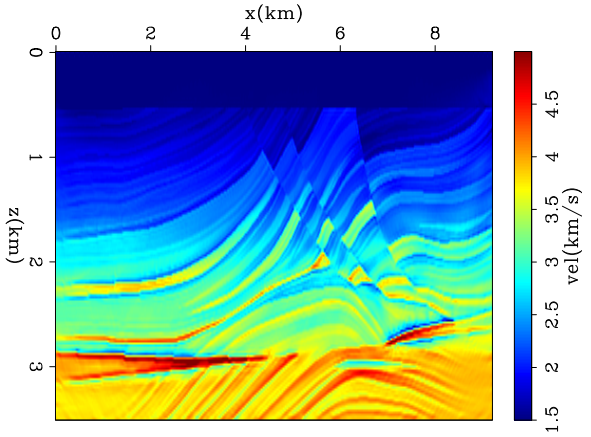


Figure 13: Final model when minimizing sum of objective functions 5 and 6 (TFWI). [**CR**]

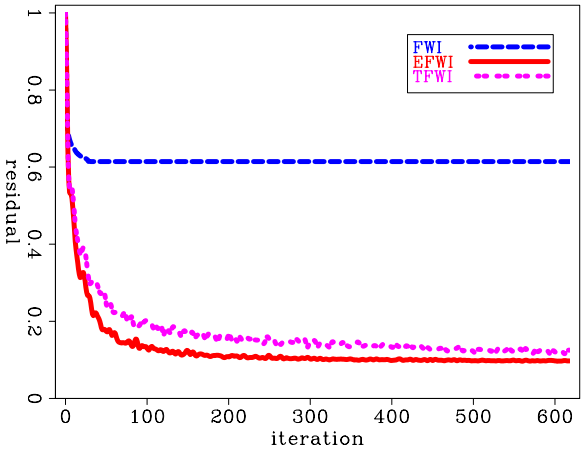


Figure 14: Graphs of data residuals as function of iterations for: FWI (blue), EFWI (red), and TFWI (magenta). [**CR**]

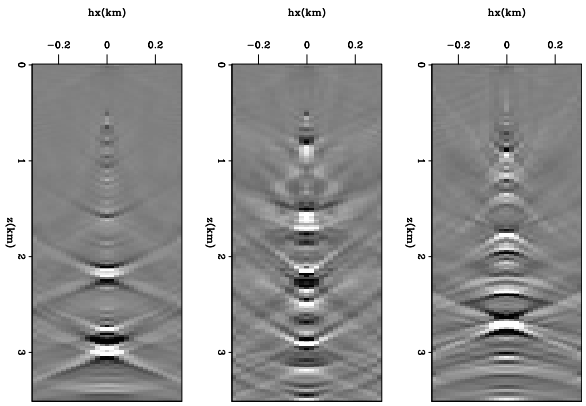


Figure 15: Final model subtracted from initial model taken at x=2.5, 5, and 7.5 km for EFWI. [**CR**]

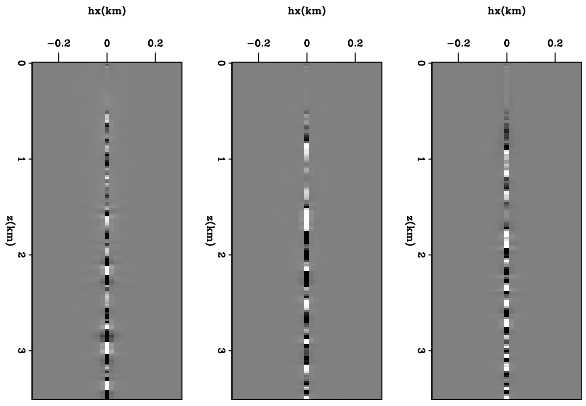


Figure 16: Final model subtracted from initial model taken at x=2.5, 5, and 7.5 km for TFWI. [**CR**]

2012).

## REFERENCES

Almomin, A. and B. Biondi, 2012, Tomographic full waveform inversion: Practical and computationally feasible approach: SEP-Report, **147**, 13–26.

Biondi, B., 2006, 3D Seismic Imaging: Society of Exploration Geophysicists.

———, 2010, Velocity estimation by image-focusing analysis: Geophysics, **75**, U49–U60.

Biondi, B. and P. Sava, 1999, Wave-equation migration velocity analysis: SEG Technical Program Expanded Abstracts, **18**, 1723–1726.

Clapp, M., 2005, Imaging under salt: illumination compensation by regularized inversion: PhD thesis, Stanford University.

Clapp, R. G., B. Biondi, and J. F. Claerbout, 2004, Geologically constrained migration velocity analysis: Geophysics, **69**, 533–546.

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

Shen, P. and W. W. Symes, 2008, Automatic velocity analysis via shot profile migration: Geophysics, **73**, VE49–VE59.

Symes, W. W., 2008, Migration velocity analysis and waveform inversion: Geophysical Prospecting, **56**, 765–790.

Tarantola, A., 1987, Inverse problem theory: Elsevier.

Zhang, Y. and B. Biondi, 2012, Residual moveout-based wave-equation migration velocity analysis in 3-D: SEP-Report, **147**, 27–38.