Multi-objective full waveform inversion in the absence of low frequencies

Pawan Bharadwaj, Delft University of Technology, Wim A. Mulder, Delft University of Technology and Shell Global Solutions International B.V., and Guy Drijkoningen, Delft University of Technology*

SUMMARY

Least-squares inversion of seismic reflection waveforms can reconstruct remarkably detailed models of the Earth's subsurface. However, the cycle-skipping associated with the high-frequency waveforms are responsible for spurious local minima in its objective function. Therefore, it is often difficult for descent methods to converge to the true model without starting from an accurate large-scale velocity estimate. To partially overcome this difficulty, we propose to use multiple objective functions for inversion. An additional constraint based on cross-correlation is added to the conventional least-squares (LS) inversion. Observations suggest this will result in a model with an accurate background velocity and reflectivity that corresponds to the global minimum of the least-squares objective function. Optimization of a cross-correlation based function (CC) in the data domain appears to pull the trapped solution out of the local minima associated with the least-squares objective function, and vice versa. Some 2-D numerical tests confirm the validity of the approach in the absence of low temporal data frequencies, starting from a constant initial velocity model.

INTRODUCTION

Conventionally, full waveform inversion (FWI) minimizes the least-squares data misfit to estimate the Earth's model parameters that explain the observed data (Tarantola, 1984; Virieux and Operto, 2009). Due to the computational cost of wave-equation modeling, this non-linear optimization can only be performed using gradient-based techniques, which are known to get stuck in the nearest local minimum of the objective function if the initial model is too far from the true one. The effect of this local minima problem becomes severe when:

1. Reliable low frequencies are absent in the observed data. Low frequencies which aren't cycle skipped can be efficiently inverted to provide long-scale updates (necessary to update the background velocity) to the model. Global convergence requires both the background velocity and reflectivity to be updated. Bunks et al. (1995) suggested the inversion of low frequencies prior to full bandwidth inversion since they help in the reconstruction of a kinematically correct background velocity model. Inversion of higher frequencies after such a reconstruction improves the resolution of the model.
2. No prior information about the Earth model is available. In that case, conventional velocity analysis methods can be used to construct a reasonable initial velocity model, but such methods are not completely automatic and involve manual picking of velocities.

Alternative functionals can be formulated that give more weight to the traveltime content in the seismic waveforms. They are known to have a smooth topography and their associated optimization is called wave-equation traveltime tomography (Luo and Schuster, 1991). The traveltime shift between the observed and the calculated waveforms can be quantified using cross-correlation (van Leeuwen and Mulder, 2010) rather than manual picking. These inversion methods are proven to be efficient compared to existing ray-based traveltime tomography, which involve a high-frequency approximation, when inverting transmitted waves in seismic data. However, the application to reflection data suffers from the fact that the reflectivity and background velocity of the model have to be estimated simultaneously. Zhang et al. (2011) proposed converting the reflection experiment into two virtual transmission experiments, but this requires manual picking of reflectors. Migration-based traveltime tomography (Clément et al., 2001) attempts to ameliorate the cycle-skipping problem by comparing the data after migration/demigration and to original observed data. This leaves the zero-offset data in place, even in the wrong velocity model, and then should correct for move-out differences at larger offsets. Van Leeuwen (2010) considered a correlation-based generalization of the method by maximizing the correlation at zero lag. These methods depend, however, on constructive interference after stacking, otherwise, no reflections will appear in the migrated/demigrated data. The required amplitude scaling is not trivial and may require iterative migration. The computation of the gradient of the objective function with respect to the model parameters is more complicated than for classic FWI. In order to overcome these difficulties, we propose a multi-objective inversion scheme that can potentially output a velocity model with accurate rough and smooth components. The multi-objective inversion is performed using two different functions: 1) the classic least-squares objective function (LS) and 2) a cross-correlation based misfit functional (CC) in the data domain. The CC objective function acts as a necessary constraint to the least-squares inversion and improves convergence. If the solution is trapped in a local minimum of one of these objective functions, it can be pulled out by the other, allowing the inversion to converge to an acceptable output model.

This paper is organized into three sections. The first section describes the method provides and the functions used for optimization. In the second section, we test the validity of the method by some 2-D numerical experiments. The last section summarizes the paper.

METHOD

Let $p^{\text{obs}}(x_s, x_g;\, t)$ denote the recorded data in the time domain at the receiver location, x_g, due to a source at x_s. Similarly, $p^{\text{cal}}(x_s, x_g;\, t)$ are data modeled with the constant-density

DOI http://dx.doi.org/10.1190/segam2013-1011.1

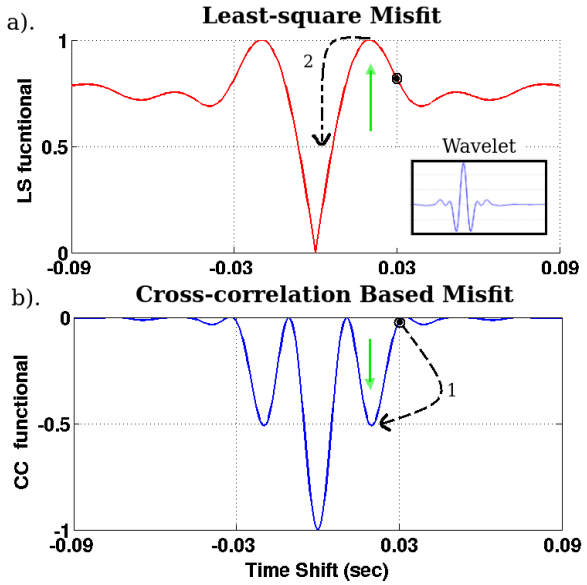

Figure 1: Plots of the LS and CC objective functions for a single reflected trace as a function of the time shift between observed and calculated traces. The source wavelet is a 10–40 Hz trapeziodal wavelet (inset). It turns out that the local extrema of the LS objective function correspond to local minima of the CC objective function.

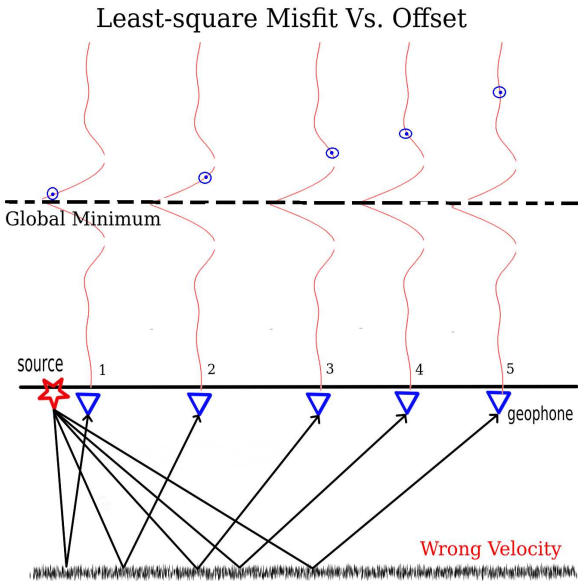

Figure 2: The least-squares error (blue dots) for traces at different offsets marked on the graph of Figure 1a. It is not possible that all blue dots lie simultaneously in a local minimum for a given interval velocity error.

acoustic wave equation. The LS functional can be written as (Tarantola, 1984):

$$J_{\rm ls} = \frac{1}{2}\int_{x_s}\int_{x_g} \chi_{\rm ls}(x_s,x_g)\, dx_g dx_s,$$

$$\chi_{\rm ls}(x_s,x_g) = \int_t \left(p^{\rm cal}(t) - p^{\rm obs}(t)\right)^2\, dt,$$

the velocity model corresponding to its global minimum completely explains the observed data. Following van Leeuwen and Mulder (2010), the CC objective function is

$$J_{\rm cc} = -\frac{1}{2}\int_{x_s}\int_{x_g} \chi_{\rm cc}(x_s,x_g)\, dx_g dx_s,$$

$$\chi_{\rm cc}(x_s,x_g) = \int_\tau \left(\int_t p^{\rm cal}(t) p^{\rm obs}(t+\tau)\, dt\right)^2 \underbrace{\left(1-e^{-\alpha\tau^2}\right)}_{\text{window at } \tau=0} d\tau.$$

Here, τ denotes the cross-correlation lag. The parameter α controls the width of the Gaussian window around zero lag. In the work of van Leeuwen and Mulder (2010), it was chosen fairly small to improve the convexity of the objective function. Here, we choose α deliberately large, so that $\chi_{\rm cc}(x_s,x_g) \sim (\langle p^{\rm cal}, p^{\rm obs}\rangle)^2$, where $\langle\cdot,\cdot\rangle$ denotes the inner product. If the normalized observed data perfectly match the normalized calculated data, then $\chi_{\rm cc}$ should reach its maximum.

Local minima of $J_{\rm cc}$ correspond to the local extrema of $J_{\rm ls}$. As a very simple example, consider two mono-frequency sinusoids representing observed and calculated waveforms. The least-squares functional will have an extremum when the phase difference between the two is $n\pi$, with n integer. It is obvious that their inner product (since α is large) and hence the absolute value of $J_{\rm cc}$ will be maximal at those phase shifts where $J_{\rm ls}$ has extrema. Figure 1 plots, in red and blue, both these functionals obtained by shifting a trapezoidal wavelet (10–40 Hz) instead of a sinusoid. Green vertical arrows in the figure mark the position of a local maximum of $J_{\rm ls}$ that corresponds to a local minimum of the $J_{\rm cc}$ plot. Minimizing either of these functionals individually may not be effective since both of them suffer from the cycle-skipping problem, as is obvious from the local minima in Figure 1a and Figure 1b. Mora (1989) noted that $J_{\rm ls}$ optimization can effectively provide low-wavenumber updates to the velocity model, but in practice, these long wavelengths are masked by the large amplitudes of the reflectors. Optimization of $J_{\rm cc}$ will help the solution to climb the local maxima associated with $J_{\rm ls}$, indicated as step 1 in Figure 1. This will help the least-squares inversion to reach the correct global minima, as indicated by step 2 in Figure 1. Since the data have multiple offsets, it is unlikely that we reach a local minimum of both these objective functions simultaneously, as illustrated in Figure 2. When $J_{\rm ls}$ is optimized starting from a homogeneous initial velocity model, the interval velocity is not updated and the long-offset traces are cycle skipped, as in Figure 2.

Optimization method. Now that we have two functionals with complementary behavior, we adopt the following strategy to perform the optimizations. The LS and CC functions are alternatively minimized. The minimization of the $J_{\rm ls}$ should be carried out first, because it produces the reflectors in the model, which are needed for the $J_{\rm cc}$ functional. The velocity model output of the $J_{\rm ls}$ inversion always acts as an input to the $J_{\rm cc}$ inversion and vice-versa. The number of iterations performed

DOI http://dx.doi.org/10.1190/segam2013-1011.1

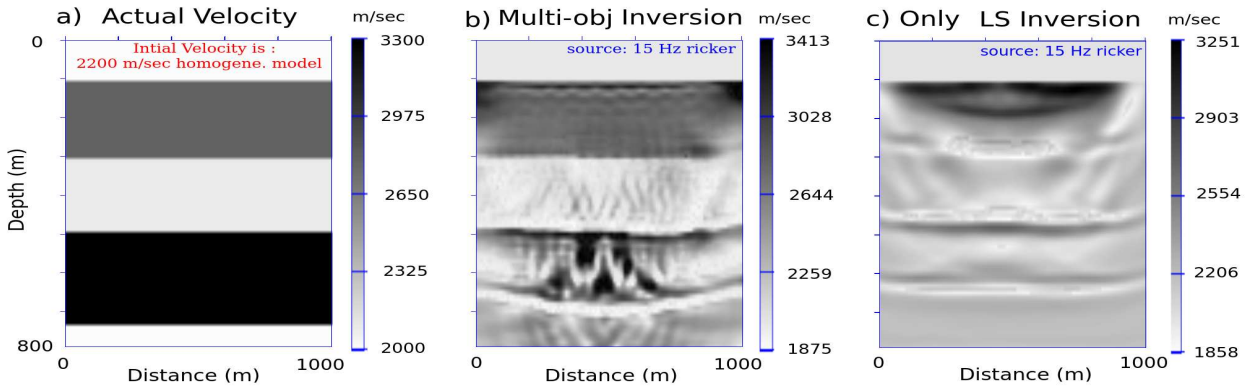

Figure 3: Layered model inversion results for a 15-Hz Ricker source wavelet. a) Actual velocity model used to generate the synthetic data. b) Full-bandwidth multi-objective inversion result. c) Conventional least-squares waveform inversion output where the optimization gets stuck in a local minimum.

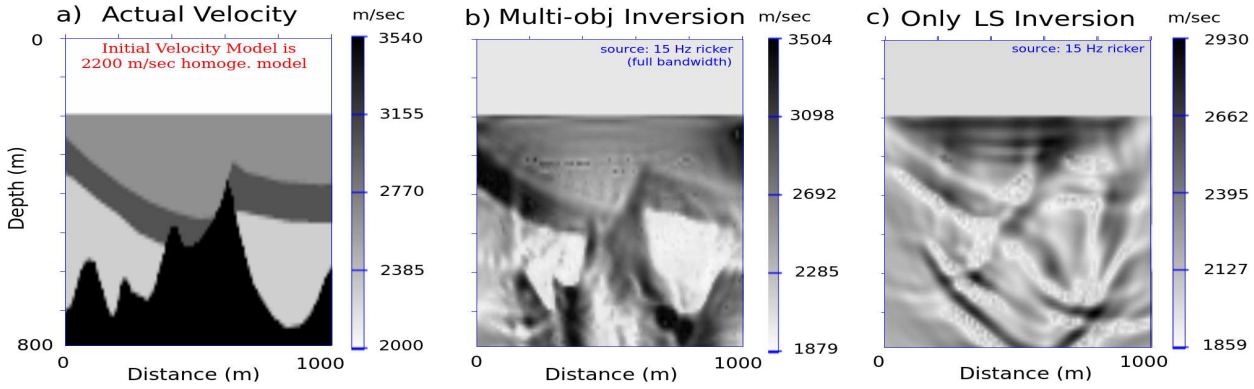

Figure 4: Fault model inversion results using a 15-Hz Ricker source. a) Actual velocity model used to generate the synthetic data. b) Full bandwidth multi-objective inversion result. c) Conventional least-squares waveform inversion result.

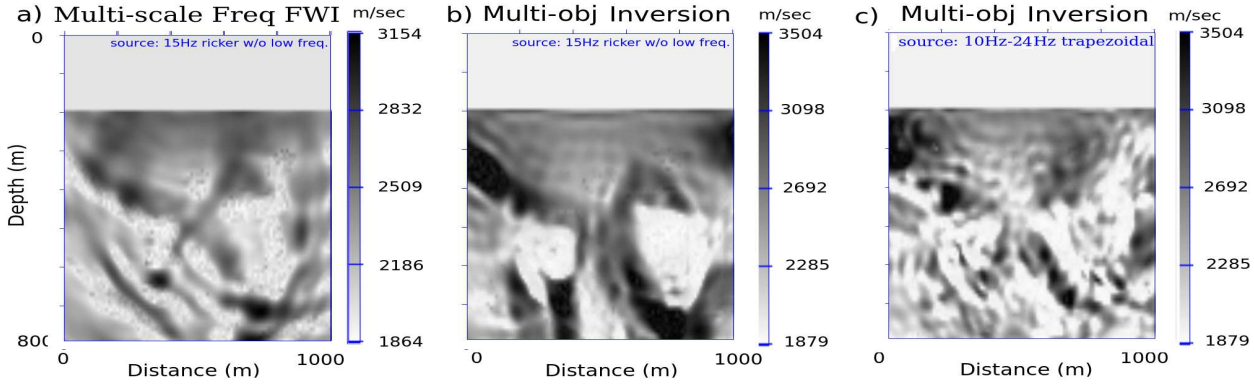

Figure 5: Fault model inversion results for a Ricker wavelet after a 6-Hz low-cut filter. a) Time domain multi-scale frequency FWI reconstruction result, using 4 bands between 6 and 24 Hz. b) Multi-objective inversion result for low-cut Ricker source. c) Multi-objective inversion result for a 10–24 Hz trapezoidal source wavelet.

DOI http://dx.doi.org/10.1190/segam2013-1011.1

during minimization of the functions is variable. The iterations stop when convergence stagnates, indicating that a local minimum has been reached. The method then switches to the other functional, using the latest inversion result as a starting model.

2-D NUMERICAL EXAMPLES

We tested the multi-objective method using two small synthetic examples involving different acoustic velocity models both of 1000-m width and 800-m depth. We placed 100 evenly spaced receivers at a depth of 40 m. Synthetic data were generated for 10 sources, 100 m apart, for the layered model shown in Figure 3a and for 50 sources, 20 m apart, for the fault model in Figure 4a. In both cases, we assumed the velocity of the top layer, 2000 m/s, to be known so that the direct arrivals can be muted from the 'observed' data.

We use a time-domain staggered-grid finite-difference code for the constant-density acoustic equation to model the data and to perform the adjoint wavefield computations required for calculating the gradients (Plessix, 2006; Fichtner, 2010). The optimization was performed by the L-BFGS quasi-Newton method (Byrd et al., 1995). In all experiments, we started from the same initial homogeneous model with a velocity of 2200 m/s below the top layer. And verified that cycle-skipping is avoided if J_{ls} is used with frequencies between 1 and 6 Hz in a multi-scale approach (Bunks et al., 1995).

Inversion of full-bandwidth using a 15-Hz Ricker wavelet. A one-time J_{ls} inversion of the full-bandwidth data suffers from the cycle-skipping problem. Figures 3c and 4c show the output of conventional full waveform inversion corresponding to the layered and the fault model, respectively. The full bandwidth multi-objective inversion produces much better results, shown in Figures 3b and 4b and provides a reconstruction of the long-scale features in the model.

Inversion without frequencies below 6 Hz for the fault model. Next, we used a Ricker source time function as a source wavelet after applying a 6-Hz low-cut filter to serve as the 'observed' data of the fault model. We tested the performance of the multi-scale FWI by first inverting the data between 6 and 10 Hz, followed by the three subsequent 4-Hz bands up to 24 Hz. The result is plotted in Figure 5a, which is not better than the conventional least-squares FWI output in Figure 4c. The multi-objective inversion produces a similar result as before, shown in the Figure 5b. Finally, we tried to break the method by selecting a ringy trapeziodal wavelet (10–24 Hz) as source time function. The result in Figure 5c shows the failure of our approach in that case.

Convergence behavior of multi-objective inversion. Figure 6 displays the convergence behavior of the J_{ls} and J_{cc} objective functions applied to the Ricker-source fault model inversion of Figure 4b. We see that whenever J_{cc} is minimized, the solution reaches a local maximim of J_{ls}. The minimization of J_{ls} in the next phase prevents the optimizer to get struck in the local minimum. If many such multi-objective inversion steps are performed, then both these objective functions eventually reach their respective global minimum in this case.

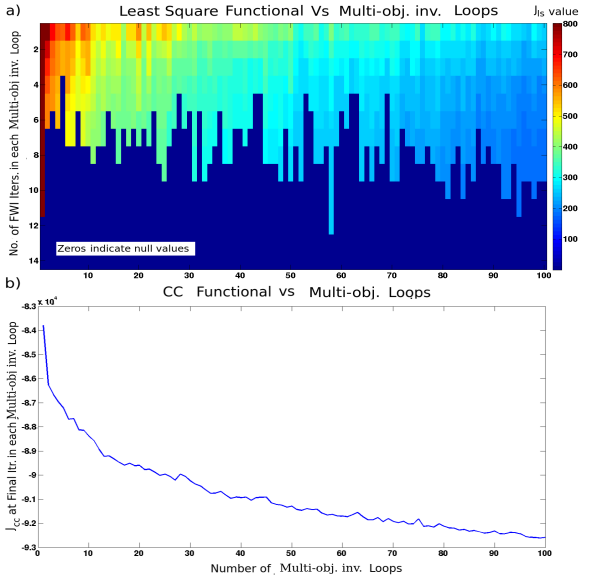

Figure 6: a) Least-squares misfit, plotted as a function of the multi-objective iteration count. The color scale displays the value of the least-squares misfit function. At each multi-objective iteration, several least-squares iterations are performed until convergence halts. The number of these iterations is indicated on the vertical axis and the change in color from red to blue shows its decrease, both in each column and overall, after restarting from the CC result. b) Value of the CC functional at the final iteration of each multi-objective inversion step.

CONCLUSIONS

We presented a multi-objective inversion scheme where the conventional least-squares misfit functional is coupled with a correlation-based cost functional. Synthetic results demonstrate that this can be effective even in the absence of low temporal frequencies in the observed data. The correlation-based cost functional appears to be capable of pulling the inversion result of least-squares objective function out of a local minimum, and the other way around. By performing a number of iterations with each functional in an alternating way, we were able to construct an acceptable inversion result that could not be reached with each of the methods independently. Future work is necessary to better understand the behavior of these two functionals and their potential to reconstruct more complex models.

ACKNOWLEDGEMENTS

This work was carried out as part of the NeTTUN project, funded from the European Commissions Seventh Framework Programme for Research, Technological Development and Demonstration (FP7 2007-2013) under Grant Agreement 280712. The first author thanks Jan Thorbecke (TU Delft) for helpful discussions on numerical wave-equation modeling.

DOI http://dx.doi.org/10.1190/segam2013-1011.1

http://dx.doi.org/10.1190/segam2013-1011.1

EDITED REFERENCES

Note: This reference list is a copy-edited version of the reference list submitted by the author. Reference lists for the 2013 SEG Technical Program Expanded Abstracts have been copy edited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES

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

Byrd, R. H., P. Lu, J. Nocedal, and C. Y. Zhu, 1995, A limited memory algorithm for bound constrained optimization: SIAM Journal on Scientific Computing, **16**, 1190–1208, http://dx.doi.org/10.1137/0916069.

Clement, F., G. Chavent, and S. Gomez, 2001, Migration-based traveltime waveform inversion of 2D simple structures: A synthetic example: Geophysics, **66**, 845–860, http://dx.doi.org/10.1190/1.1444974.

Fichtner, A., 2010, Full seismic waveform modeling and inversion: Springer.

Luo, Y., and G. Schuster, 1991, Wave-equation traveltime inversion: Geophysics, **56**, 645–653. http://dx.doi.org/10.1190/1.1443081.

Mora, P., 1989, Inversion = migration + tomography: Geophysics, **54**, 1575–1586, http://dx.doi.org/10.1190/1.1442625.

Plessix, R.-E., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications: Geophysical Journal International, **167**, 495–503, http://dx.doi.org/10.1111/j.1365-246X.2006.02978.x.

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

van Leeuwen, T., 2010: Ph.D. thesis, Delft University of Technology.

van Leeuwen, T., and W. A. Mulder, 2010, A correlation-based misfit criterion for wave-equation traveltime tomography: Geophysical Journal International, **182**, 1383–1394, http://dx.doi.org/10.1111/j.1365-246X.2010.04681.x.

Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, no. 6, WCC1–WCC26, http://dx.doi.org/10.1190/1.3238367.

Zhang, S., G. Schuster, and Y. Luo, 2011, Wave-equation reflection traveltime inversion: Presented at the 81st Annual International Meeting, SEG.