

FWI without tears: a forward modeling free gradient

*Marcelo Guarido, Laurence R. Lines and Robert Ferguson
University of Calgary*

Summary

Full waveform inversion (FWI) is a machine learning algorithm with the goal to find the Earth's model parameters that minimize the difference of acquired and synthetic shots. On this work, we are introducing a new interpretation of the gradient as a residual impedance inversion of the acquired data. Its estimation is forward modeling and wavelet free, reducing its costs drastically, as the inverted model was obtained on a personal laptop without parallel processing. The new method was successfully applied on the acoustic Marmousi simulation. The inverted model, when using the same starting point, is comparable to the results when using the migrated residuals. This approximation also made possible to invert the order of migration and the stack steps to use a post-stack depth migration to estimate the gradient with promising outputs. In the end, we are proposing a new FWI approximation that is cheap and stable and could be used on real data in the same processing center that has enough computer power to run a PSDM or even just a post-stack depth migration.

Introduction

Seismic inversion techniques are the ones that use intrinsic information contained in the data to determine rock properties by matching a model that "explains" the data. Some examples are the variation of amplitude per offset, or AVO (Shuey, 1985; Fatti et al., 1994), the traveltimes differences between traces, named traveltimes tomography (Langan et al., 1984; Bishop and Spongberg, 1984; Cutler et al., 1984), or even by matching synthetic data to the observed data, as it is done in full waveform inversion (Tarantola, 1984; Virieux and Operto, 2009; Margrave et al., 2010; Pratt et al., 1998), among others. These inversions can compute rock parameters as P and S waves velocities, density, viscosity and others. On this paper we are focusing in the inversion of the P wave velocity (acoustic).

Full waveform inversion (FWI) is a machine learning based method, which objective is to estimate the model parameters that minimizes the difference between observed (acquired) data and synthetic shots (Margrave et al., 2011). This is accomplished by iteratively updating the starting model with a new scaled gradient and computing new synthetic shots.

The full waveform inversion was proposed in the early 80's (Pratt et al., 1998) but the technique was considered too expensive in computational terms. Lailly (1983) and Tarantola (1984) simplified the methodology by using the steepest-descent method (or gradient method) in the time domain to minimize the objective function without calculate, explicitly, the partial derivatives. They compute the gradient by a reverse-time migration (RTM) of the residuals. Pratt et al. (1998) develop a matrix formulation for the full waveform inversion in the frequency domain and present efficient ways to compute the gradient and the inverse of the Hessian matrix (the step length for convergence in the FWI) the Gauss-Newton or the Newton approximations. The FWI is shown to be more efficient if applied in a multi-scale method, where lower frequencies are inverted first and is increased as more iterations are done (Pratt et al., 1998; Virieux and Operto, 2009; Margrave et al., 2010). An overview of the FWI theory and studies are compiled by Virieux and Operto (2009). Lindseth (1979) showed that an impedance inversion from seismic data is not effective due to the lack of low frequencies during the acquisition but could be compensated by the match with a sonic-log profile. Margrave et al. (2010) used a gradient method and matched it with sonic logs profiles to compensate the absence of the low frequency and to calibrate the model update by computing the step length and a phase rotation (avoiding cycle skipping). They also proposed the use of a PSPI

(phase-shift-plus-interpolation) migration (Ferguson and Margrave, 2005) instead of the RTM, so the iterations are done in time domain but only selected frequency bands are migrated, using a deconvolution imaging condition (Margrave et al., 2011; Wenyong et al., 2013) as a better reflectivity estimation, as used by Guarido et al. (2015a;2015b). Guarido et al. (2016) show the need of the application of an impedance inversion step in the gradient and use a band-limited impedance inversion (BLIMP) method using the algorithm implemented by Ferguson and Margrave (1996). Warner and Guasch (2014) use the deviation of the Weiner filters of the real and estimated data as the object function with great results.

We are proposing a new approximation for the FWI, where we treat the gradient as a residual impedance of the current model and the impedance inversion of the acquired data. On each iteration, the data is PSPI migrated (Ferguson and Margrave, 2005), with a deconvolution imaging condition, using the current model and applying a BLIMP inversion on the stacked data. A conjugate gradient is also used to improve the quality of the gradient and to reduce the number of iterations (Zhou et al., 1995; Vigh and Starr, 2008). The step length is computed by a least-square minimization (Pica et al., 1990) and is being estimated for individual frequencies. To compute the residuals on the classic method, a finite difference forward modelling algorithm is used to create the synthetic shots. The results of the new approximation are comparable with the classic method (steepest descent). We went further and inverted the migration and stack processing steps order computed the gradient using a zero-offset PSPI migration (post-stack). This test is preliminary but the results are really promising.

Theory

The objective of the FWI methodology is to minimize an objective function. Here we minimize the residuals $\Delta d(m)$, that is the difference between observed data \mathbf{d}_0 and synthetic data $\mathbf{d}(m)$, when the model m (here P wave velocity) is changed:

$$C(m) = \|\mathbf{d}_0 - \mathbf{d}(m)\|^2 = \|\Delta d(m)\|^2 \quad (1)$$

Minimizing the objective function $C(m)$ in respect to the model m , we can to the steepest-descent formula (Pratt et al., 1998):

$$\mathbf{m}_{n+1} = \mathbf{m}_n - \alpha_n \mathbf{g}_n \quad (2)$$

where α is the step length, \mathbf{g} is the gradient and n is the n -th iteration. This equation shows that a model update can be obtained by adding a scaled gradient to the current model. This routine is kept until stoping criteria is reached. The gradient is, according to theory, computed by a reverse time migration of the residuals (Tarantola, 1984; Pratt et al., 1998; Virieux and Operto, 2009), but we decided to use the phase-shift-plus-interpolation (PSPI) migration. Later, the BLIMP algorithm uses the initial model as pilot to apply an impedance inversion of the gradient. The first iterations use only the low frequency on the data and higher frequencies are included later.

Understanding all the gradient estimation steps as seismic processing tools, equation 2 can be rewritten in terms of the migration \mathbf{M} , stacking \mathbf{S} and impedance inversion \mathbf{I} operators:

$$\mathbf{m}_{n+1} = \mathbf{m}_n - \alpha_n \mathbf{I} \{ \mathbf{S} [\mathbf{M} (\mathbf{d}_0 - \mathbf{d}_n)] \} \quad (3)$$

where \mathbf{d}_n is the synthetic shot. Guarido et al. (2016) assume all three operators are linear (true for migration and stack and approximate for impedance inversion) and the gradient can be interpreted as a residual difference of the processed acquired data and the processed synthetic data (both migrated using the current model). Second one, on a perfect case, is the current model itself (the migrated, stacked and impedance inversion of the synthetic data). This explanation is better visualized by looking to equation 4:

$$\begin{aligned} \mathbf{m}_{n+1} &= \mathbf{m}_n - \alpha_n (\mathbf{I} \{ \mathbf{S} [\mathbf{M} (\mathbf{d}_0)] \} - \underbrace{\mathbf{I} \{ \mathbf{S} [\mathbf{M} (\mathbf{d}_n)] \}}_{\text{Current model}}) \\ &= \mathbf{m}_n + \alpha_n (\mathbf{I} \{ \mathbf{S} [\mathbf{M} (\mathbf{d}_0)] \} - \mathbf{m}_n) \end{aligned} \quad (4)$$

Interpreting the gradient as the residual difference of the processed acquired data and the current model saves us to actually compute a synthetic data at each shot position. Source estimation is also not required. Two forward modeling are still required on the step length determination. But it is only an amplitude match and the correct wavelet is not needed.

We can make the method even cheaper if we invert the order of the migration and stacking operators on equation 4. This would result on using a stacked session as input and a post-stack migration at each iteration.

Examples

Simulations are done on the Marmousi velocity model (figure 1a). Synthetic acquired data are generated by a 2D acoustic finite difference code and a Ricker wavelet with 5Hz of dominant frequency on 104 different positions. Starting model (figure 1b) is a smoothed version of the real Marmousi. For the classic FWI method (figure 1c), forward modeling is done using current model and same wavelet as the acquired data. First iteration start from a frequency band of 4 to 6Hz, repeated, and the maximum frequency is increased by 2Hz when convergence is reached. BLIMP is applied on PSPI migrated residuals to convert the reflection coefficients to velocity using the initial model to fullfill the 1-3Hz gap on the data. Figure 1d is the inverted model based on the forward modeling free gradient method and the only difference is the interpretation of the gradient. For both methods, the step length is estimated as proposed by Pica et al. (1990). Both inverted models are comparable and with a good resolution. The advantage of the forward modeling free method is the computing requirement and processing time, which is reduced by about 70%.

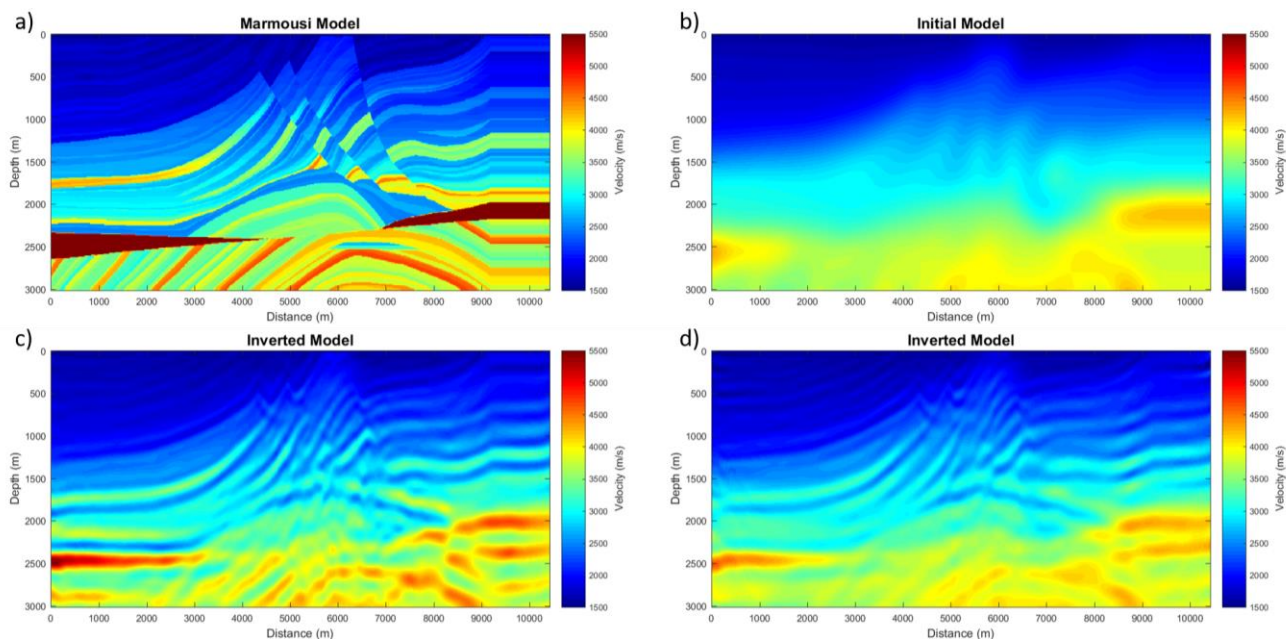


Figure 1: a) true Marmousi model, b) initial model for all runs, c) inverted model with classic FWI and d) inverted model with the forward modeling free gradient method.

Figure 2a is the stacked session used as input data for the post-stack FWI method and resulted model is shown on figure 2b. There is a loss of resolution if compared to the previous results but a gain in processing time and computer needs. For the classic method, it was used parallel processing in MatLab with 24 clusters and 48 hours of run time. To run the forward modeling free gradient method, it was used a personal gaming laptop, no parallel processing and 8 hours of run time in Octave. The post stack method ran on a tablet with dual core processor and 1 hour of run time. The resolution decreases as the cheaper the method is. The choice of the method is just a matter of cost and benefit. The best response will require the highest investment. However, we show that a reasonable result, with just a small loss of resolution, can be achieved by a drastically reduction of costs.

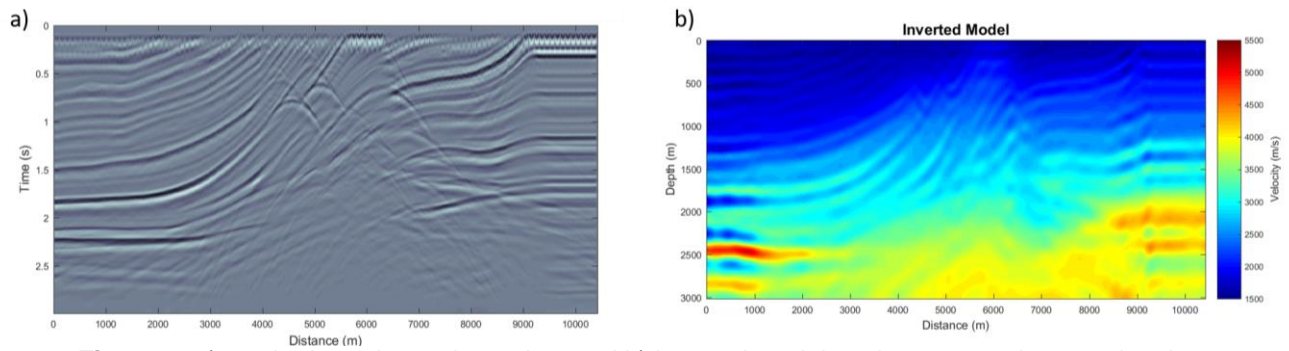


Figure 2: a) stacked session as input data and b) inverted model on the post-stack approximation.

Figure 3 compares the shots and models errors of the 3 methods. All the methods show to be stable and reach some convergence. The “break” of the couves are when the inversion is around the dominant frequency of the acquired data (12Hz).

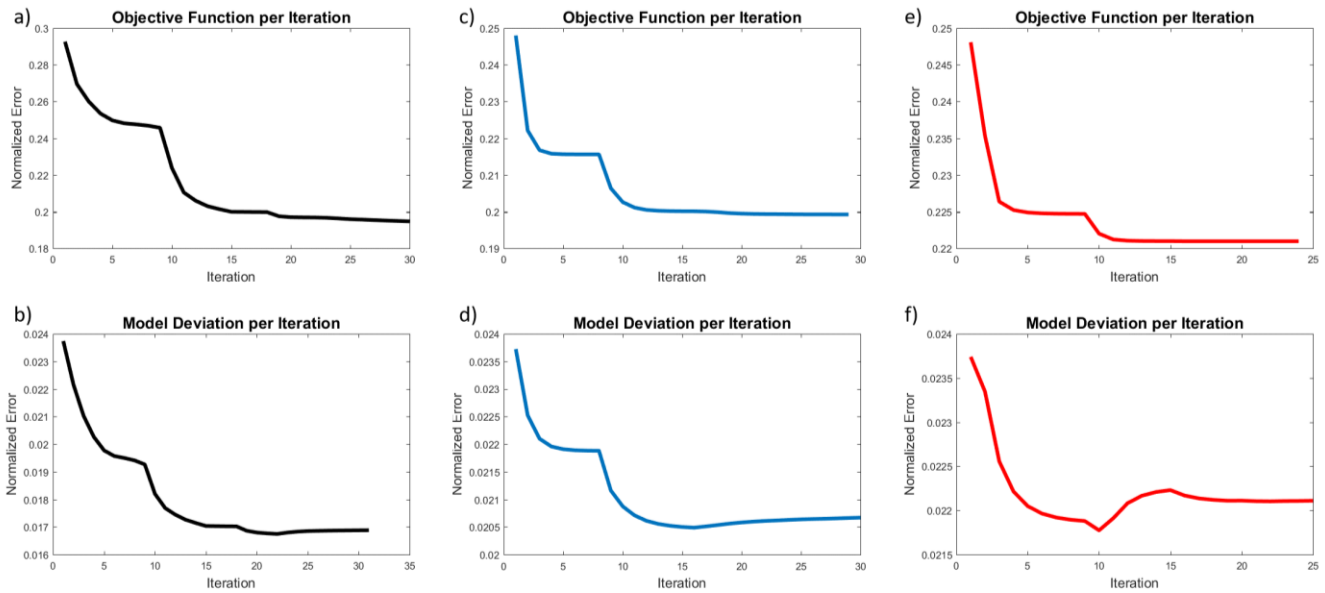


Figure 3: Respectively, shot and model errors of a) and b) classic FWI, c) and d) forward modeling free gradient method and e) and f) post-stack approximation.

Conclusions

It was presented a new FWI method based on interpreting the gradient as a residual difference of the impedance inversion of acquired data and the current inverted model, removing the need to compute one forward modeling per shot location on every iteration. Comparing with the classic FWI, the results are comparable with some loss of resolution as costs go cheaper but the cost-benefit trade looks to worth it.

A post stack method with preliminary results was also presented reducing even more the costs for a FWI run but also losing resolution. However, we are confident that this is a safe strategy to follow with the goal of applying the FWI method on large datasets with reduced computer requirements. In the end, the choice of which method to use will depend on the investment power of the user.

Acknowledgements

The authors 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. We thank Soane Mota dos Santos for the suggestions, tips and productive discussions.

References

- Bishop, T. N., and Spongberg, M. E., 1984, Seismic tomography: A case study: SEG Technical Program Expanded Abstracts, 313, 712–713.
- Cutler, R. T., Bishop, T. N., Wyld, H. W., Shuey, R. T., Kroeger, R. A., Jones, R. C., and Rathbun, M. L., 1984, Seismic tomography: Formulation and methodology: SEG Technical Program Expanded Abstracts, 312, 711–712.
- Fatti, J. L., Smith, G. C., Vail, P. J., Strauss, P. J., and Levitt, P. R., 1994, Detection of gas in sandstone reservoirs using avo analysis: A 3-d seismic case history using the geostack technique: *Geophysics*, 59, No. 9, 1362–1376.
- Ferguson, R., and Margrave, G., 2005, Planned seismic imaging using explicit one-way operators: *Geophysics*, 70, No. 5, S101–S109.
- Guarido, M., Lines, L., and Ferguson, R., 2015b, Full waveform inversion: a synthetic test using pspi migration: SEG Technical Program Expanded Abstract, 279, 1456–1460.
- Guarido, M., Lines, L., and Ferguson, R., 2015a, Convergence of a FWI scheme based on PSPI migration: GeoConvention Technical Program Expanded Abstract.
- Guarido, M., Lines, L., and Ferguson, R., 2016, FWI without tears: a forward modeling free gradient: CREWES Research Report, 28, 26.1–26.20.
- Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migrations: Conference on inverse scattering, theory and application: Society of Industrial and Applied Mathematics, Expanded Abstracts, 206–220.
- Langan, R. T., Lerche, I., Cutler, R. T., Bishop, T. N., and Spera, N. J., 1984, Seismic tomography: The accurate and efficient tracing of rays through heterogeneous media: SEG Technical Program Expanded Abstracts, 314, 713–715.
- Lindseth, R. O., 1979, Synthetic sonic logs—a process for stratigraphic interpretation: *Geophysics*, 44, No. 1, 3–26.
- Ma, Y., Hale, D., Meng, Z. J., and Gong, B., 2010, Full waveform inversion with image-guided gradient: SEG Technical Program Expanded Abstracts, 198, 1003–1007.
- Margrave, G., Ferguson, R., and Hogan, C., 2010, Full waveform inversion with wave equation migration and well control: CREWES Research Report, 22, 63.1–63.20.
- Margrave, G., Yedlin, M., and Innanen, K., 2011, Full waveform inversion and the inverse hessian: CREWES Research Report, 23, 77.1–77.13.
- Pica, A., Diet, J. P., and Tarantola, A., 1990, Nonlinear inversion of seismic reflection data in a laterally invariant medium: *Geophysics*, 55, No. 2, R59–R80.
- Pratt, R. G., Shin, C., and Hick, G. J., 1998, Gauss-newton and full newton methods in frequency-space seismic waveform inversion: *Geophysical Journal International*, 133, No. 2, 341–362.
- Shuey, R. T., 1985, A simplification of the Zoeppritz equations: *Geophysics*, 50, No. 4, 609–614.
- Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: *Geophysics*, 49, No. 8, 1259–1266.
- Treitel, S., Lines, L., and Ruckgaber, G., 1995, Seismic impedance estimation: *Geophysical Inversion and Applications - Memorial University of Newfoundland*, 1, 6–11.
- Vigh, D., and Starr, E.W., 2008, 3d prestack plane-wave, full-waveform inversion: *GEOPHYSICS*, 73, No. 5, VE135–VE144.
- Virieux, J., and Operto, S., 2009, An overview of full-waveform inversion in exploration geophysics: *Geophysics*, 74, No. 6, WCC1–WCC26.
- Warner, M., and Guasch, L., 2014, Adaptive waveform inversion: Theory: SEG Technical Program Expanded Abstracts, 207, 1089–1093.
- Wenyong, P., Margrave, G., and Innanen, K., 2013, On the role of the deconvolution imaging condition in full waveform inversion: CREWES Research Report, 25, 72.1–72.19.
- Zhou, C., Cai, W., Luo, Y., Schuster, G. T., and Hassanzadeh, S., 1995, Acoustic wave-equation traveltimes and waveform inversion of crosshole seismic data: *GEOPHYSICS*, 60, No. 3, 765–773.