FWI: The role of the imaging frequencies and the number of iterations in the reduction of the sensibility to noise

Franciane C. Peters (PEC/COPPE/UFRJ), Webe João Mansur (PEC/COPPE/UFRJ), Djalma Manoel Soares Filho (CENPES/PETROBRAS), Cid da Silva G. Monteiro (PEC/COPPE/UFRJ), Wellington Luís A. Pereira (PEC/COPPE/UFRJ)

Abstract

The full waveform inversion is a nonlinear inverse problem in which the values of physical properties recovered from seismic data are sensitive with respect to noise in data. The problem becomes more critical when the number of parameters increases or different kind of physical parameters are searched. The multi-scale approach is one of the most used strategies to make the solution process more stable. Tikhonov and other kinds of regularization are also used. However, these strategies have some drawbacks. The regularization operators and parameters influence the solution and the optimal ones are expensive to estimate. The multi-scale approach demands the choice of imaging frequencies and number of iterations for each frequency or other stop criteria. The knowledge of the misfit level to stop minimizing is a good stop criterion to avoid fit noise in data. The main objective of this work is to present some numerical experiments of acoustic inversion in the context of multi-scale approach considering the same total number of iterations but with a different number of imaging frequencies for the same bandwidth. The experiments show that the results can be less sensitive to noise in data if the number of iterations for each frequency is reduced and the number of frequencies is increased. The results are compared with images recovered from noise free data in order to show the effect of different settings.

Introduction

The full waveform inversion is a technique to recover an image of the subsurface using data from a seismic survey. It is computationally expensive because it is based on two-way wave equation. Although the first concepts about the technique were stated in the 80's (Tarantola, 1984), the progressive improvement of computational resources have made the technique viable. Nowadays, a lot of studies have contributed to this important and promising tool for exploration geophysics.

It is well known that many aspects are involved in solving such inverse, non-linear, and ill-posed problem. The solution process is usually sensitive to noise in data and the instability increases with the large number of parameters, different kinds of parameters, and limitations of acquired data. By this way, the development of stabilization strategies is very useful to generate robust FWI methodologies.

The multiscale approach is a well-known stabilization strategy (Bunks et al., 1995). In this approach, the inverse problem is solved in frequency domain using some previously selected imaging frequencies. The data is transformed to frequency domain and an image is obtained iteratively by means of minimization of residual considering low frequency data. Then, such image is used as input to the same procedure, now with a higher frequency. This is repeated up to the highest imaging frequency and the final image is obtained.

However, the multiscale approach is not enough and many other regularization tools are applied in order to avoid noise fitting by means of including the prior knowledge in the inversion procedure. Methods based on Tikhonov regularization (Aster et al., 2005) are the most often used. They change the basic functional based on the data residual adding a penalty factor that depends on estimated the parameters. Total variation regularization (Aster et al., 2005) and multiplicative regularization (Van den Berg et al., 2001 and Hu et al., 2011) are also used. During the minimization of the new functional, depending on the regularization operator, the contribution of noise in the recovered image is attenuated.

Although regularization methods are effective, they have some drawbacks. The first is that Tikhonov-type regularization leads to smoothed images. In order to avoid the defocusing on transitions between layers the regularized methods should be adapted (Zhang et al., 2013). The second problem is that regularization may limit the size of optimization step, reducing the convergence velocity and increasing the number of wave propagation modeling. Another problem is that it is difficult to determine the value of regularization parameters in advance. In fact, the minimization of a penalized functional means solving a new problem, not the original one. Then, the solution to the new problem may not be the best solution to the original problem.

Among other regularization methods, Fichtner (2011) discusses the “regularization through finite number of iterations”. It is an aspect that deserves special attention because it has an important role in nonlinear inverse problems. The author highlights that, in practice, the suitable number of iterations is difficult to determine.

Besides the regularization, the functional to be minimized can have other components. The correct choice of the norm of the difference between survey data and computational data is important because it makes the

functional more robust. The same can be achieved by the use of a suitable weighting operator of data. The fact is that different functionals lead to different solutions and we want to avoid this ambiguity. Here, we study the influence of the number of iterations of the minimization algorithm and the choice of frequencies used in the multiscale approach on the recovered image. The objective is to recover a better image, less sensitive to noise, considering no knowledge of the noise level. We can show that the correct choice of the number of frequencies combined with the number of iterations can be useful to obtain better images, more robust with respect to noisy data.

Inverse problem

In this work we treat the acoustic inversion of seismic data. Although more complex inversions based on elastic modeling of wave propagation have been developed, the acoustic inversion is still very much used because it is computationally efficient. Moreover, many works discuss the validity of this modeling that ignores elastic effects (Marelli et al., 2012), so the field of its application and its limitation have been well understood.

In acoustic inversion, the recovered image represents the p-wave velocities (v_p) at nodes of the grid. In order to constrain the parameters of inversion $\boldsymbol{\chi}$ we use the second transformation of velocities described by Habashy et al. (2004) and rewritten as follows:

$$v_p(\chi) = \frac{v_p^{max} e^{\chi} + v_p^{min} e^{-\chi}}{e^{\chi} + e^{-\chi}}. \qquad (1)$$

So, the velocities are in a range between $v_p^{min} = 100m/s$ and $v_p^{max} = 6900m/s$ and the inversion parameters are unconstrained, $\chi \in (-\infty, +\infty)$. In this paper we consider that the density of the model is uniform and known.

The misfit functional we use is the simplest possible, without weighting or regularization, the least-absolute-values norm ($l1$-norm) of the difference between observed data d^O and computational data d^χ:

$$C_{l1} = \sum_{i=1}^{M} |d_i^O - d_i^\chi|; \qquad (2)$$

where M is the number of traces. It is well known that, in general, $l1$-norm leads to more robust functionals with respect to noise in data than $l2$-norm (Fichtner, 2011 e Brossier et al., 2010). Furthermore, we want a robust but simpler functional, so we avoid the hybrid ones like the Huber norm (Guitton, 2003). Adopting the $l1$-norm, the gradient of the functional with respect to the model parameters is given by:

$$\boldsymbol{G}_{l1} = \Re\{\mathbf{J}^{\mathrm{t}}\mathbf{r}\}; \qquad (3)$$

with

$$r_i = \frac{(d_i^O - d_i^\chi)^*}{|d_i^O - d_i^\chi|}; \qquad (4)$$

where $\mathbf{r}$ is the data misfit vector, $\Re$ denotes the real part of the complex number, * denotes the conjugate of a complex number, and t denotes the transpose operator. Note that the gradient is computed using the adjoint formulation (Plessix et al., 2006). This avoids the explicit forming of the derivative matrix $\mathbf{J}$.

In order to minimize the misfit functional C_{l1}, we use a quasi-Newton method, the limited memory version of BFGS algorithm (Liu et al., 1989). The multiscale approach is adopted.

Forward problem

The forward problem is modeled by the second-order wave propagation PDE in frequency domain:

$$\frac{\omega^2}{\kappa(\mathbf{x})} P(\mathbf{x},\omega) + \nabla \cdot \left(\frac{1}{\rho}\nabla P(\mathbf{x},\omega)\right) = S(\mathbf{x},\omega); \qquad (5)$$

where $\mathbf{x}$ represents the spatial coordinates, $P(\mathbf{x},\omega)$ is the component of pressure field, ω is the frequency, κ is the Bulk Modulus, ρ is the density, and $S(\mathbf{x},\omega)$ is a source term. The upper border of the model is the free surface and the other borders are nonreflective boundaries. Furthermore, perfect matched layers are implemented combined with the nonreflective boundaries.

For each new velocity model, the numerical solution of Equation (5) is carried out by a staggered grid finite differences scheme using five points. In order to obtain better accuracy and avoid the "inverse crime", the synthetic acoustic data was generated by a staggered grid finite differences scheme using thirteen points. More details about these schemes and the absorbing boundaries used can be found in Hustedt et al. (2004).

The finite difference scheme used to solve Equation (5) leads to a system of linear equations that is solved by an implementation of LU factorization known as super-LU (Li et al., 2003). In this case, the LU factorization is suitable because we have to solve as many linear systems as shots, but the matrices of the systems are the same.

Numerical Experiments

In order to verify if the solution is sensitive to the number of imaging frequencies and iterations for each frequency, some numerical experiments of inversion of synthetic seismic data were performed.

First of all, synthetic seismic data were generated and used as observed data. The 13 grid points scheme of finite difference was used to solve the acoustic wave propagation over the "exact model" (Figure 2) for hundreds of frequencies. After that, the Inverse Fourier Transform was used to generate seismograms in time domain. Such seismograms compose the first set of observed data. In this work we call this set of "exact data". Random noise with Gaussian distribution was added to "exact data" in order to generate the "noisy data". Both

sets of data were inverted to recover an image of p-wave velocity model. The seismic data simulates a split-spread seismic survey with 382 traces for each one of the 188 shots. A Ricker wavelet of 8Hz of pick frequency was used as a source to generate synthetic data and during the inversion.

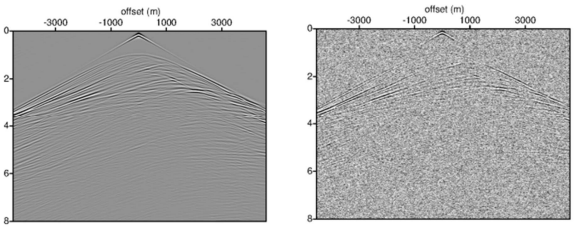


Figure 1: One of the seismograms of the exact data set (left) and noisy data set (right).

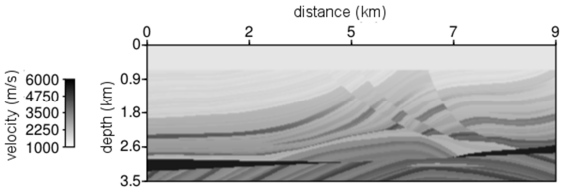


Figure 2: The exact p-wave velocity model used to generate exact seismic data. It is the acoustic Marmousi/SEG model with a layer of water.

Eighteen experiments of seismic data inversion were performed. Some of them were done using exact data and others were done using noisy data. Two different initial models were given as initial guess for the minimization process. Both of them were obtained by smoothing of the exact model in different degrees. The second model is smoother than the first one and they can be seen in Figure 4(a) and Figure 5(a). A different number of imaging frequencies and total number of iterations were applied in order to verify their influence over the final image. Note that the first and last imaging frequency is the same for all results, 6Hz and 15Hz, respectively. For each imaging frequency, the only stop criterion is the number of iterations to have a fair comparison between results. The inversion experiments can be summarized as follows:

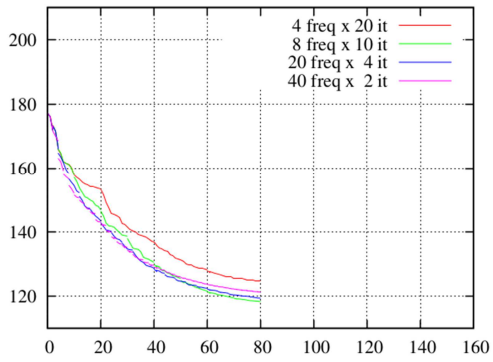
1. Inversion of the exact seismic data using the first initial model and the total number of iterations is 80, equally divided between the 4, 8, 20, or 40 imaging frequencies;
2. Inversion of the noisy seismic data using the first initial model and the total number of iterations is 80, equally divided between the 4, 8, 20, or 40 imaging frequencies;
3. Inversion of the noisy seismic data using the second initial model and the total number of iterations is 120, equally divided between the 4, 8, 20, 40 or 60 imaging frequencies;
4. Inversion of the noisy seismic data using the second initial model and the total number of iterations is 160, equally divided between the 4, 8, 20, 40 or 80 imaging frequencies.

In all of the inversion experiments, the velocity at each grid node of the finite differences scheme is considered an unknown parameter of the inverse problem, except the layer of water and the PML. So, for all the frequencies, the grid is the same and has 767x293 nodes in x and z direction respectively. The number of parameters is 767x243.

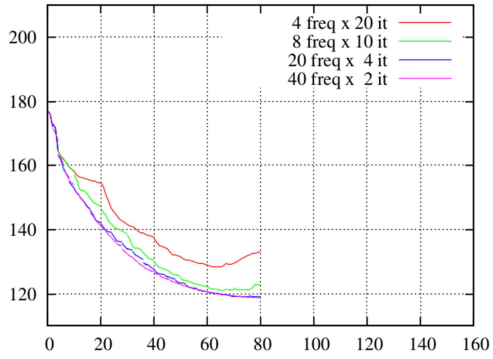
In order to compare the results, the mean absolute error between the recovered image and the exact image was computed for each updated model during the inversion process. Figure 3 shows the evolution of the error. The evolution of the misfit is not presented here because the comparison is difficult since different frequencies have different levels of misfit and are not comparable. Furthermore, for each image frequency, the value of the misfit is reduced (or remains constant) by the optimization algorithm. However, the reduction of the misfit value does not ensure that the image error is reduced because of the noise fitting. So we have opted to show the evolution of the image error instead of the misfit. In each graphic, the legend shows the number of imaging frequencies used and the number of iterations for each frequency.

Figure 4 shows the recovered images using the noisy data after a total of 80 iterations considering the same initial model, Figure 4(a). Note that the recovered image using 4 imaging frequencies in the range 6Hz-15Hz is rough. Using 8 frequencies, such artefacts are reduced and using the greatest number of frequencies tested, such defects almost disappear. The recovered images after 80 iterations using noisy data are not presented because the images are similar to each other and they do not suffer of noise fitting.

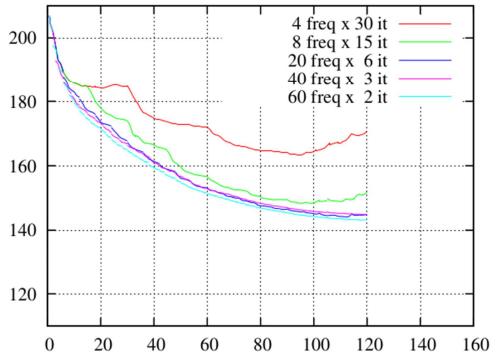
Figure 5 and Figure 6 present the recovered images after 120 iterations and 160 iterations. Note that, in both cases, the used initial model was the second one. So, since it is smoother than the first one, consequently far from the exact model, a greater number of iterations were applied. Comparing Figure 5 with Figure 6 (b) and (c), respectively, the increment in the number of iterations increases the image error. Such comparison is easier to observe comparing red and green curves of Figure 3(c) and (d). On the other hand, increasing the number of imaging frequencies in the same range, the greater number of iterations allows obtaining better images (smaller errors).



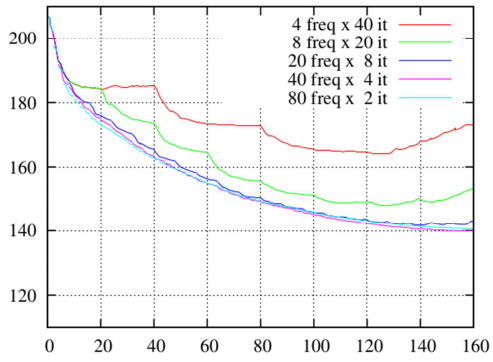
(a) 80 iterations, exact data



(b) 80 iterations, noisy data

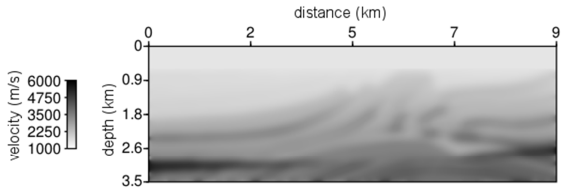


(c) 120 iterations, noisy data

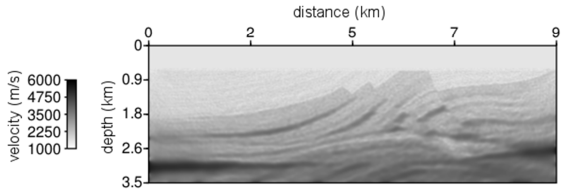


(d) 160 iterations, noisy data

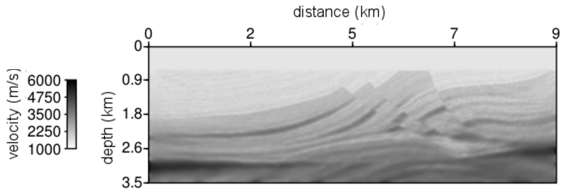
Figure 3: The evolution of the image error (m/s) during the process of inversion.



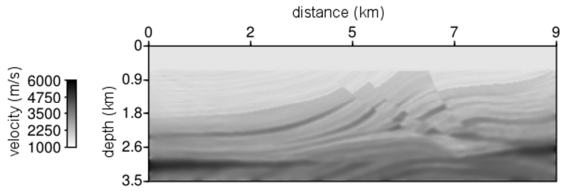
(a) Initial Model 1



(b) 4 frequencies

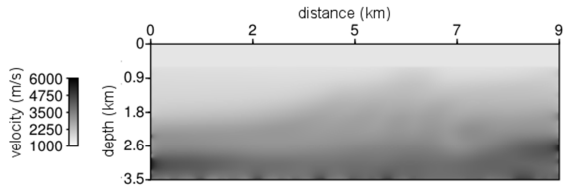


(c) 8 frequencies

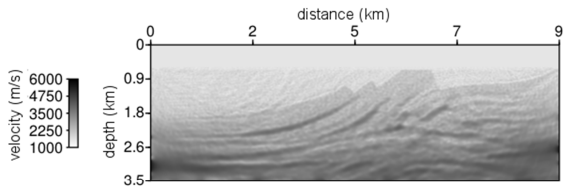


(d) 40 frequencies

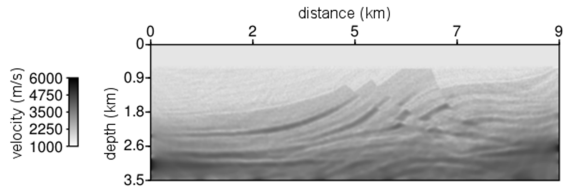
Figure 4: The model used as initial guess to the minimization process (a) and recovered images after 80 iterations using 4, 8, and 40 frequencies.



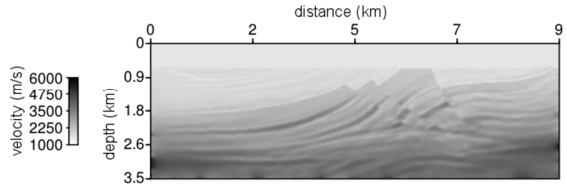
(a) Initial Model 2



(b) 4 frequencies

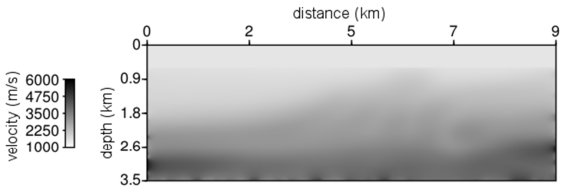


(c) 8 frequencies

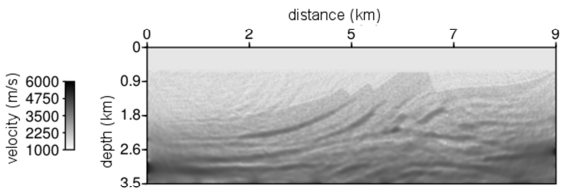


(d) 60 frequencies

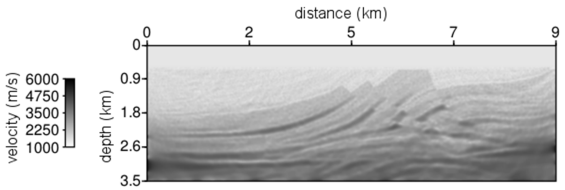
Figure 5: The model used as initial guess to the minimization process (a) and recovered images after 120 iterations using 4, 8, and 60 frequencies.



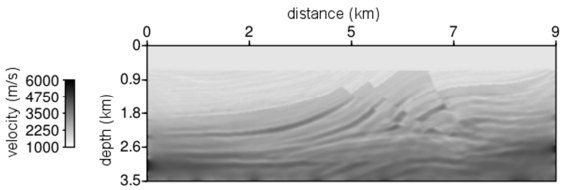
(a) Initial Model 2



(b) 4 frequencies



(c) 8 frequencies



(d) 80 frequencies

Figure 6: The model used as initial guess to the minimization process (a) and recovered images after 160 iterations using 4, 8, and 80 frequencies.

Conclusions

It is well-known that the bandwidth of imaging frequencies is very important because it defines the size of structures that can be recovered from seismic data. However, theoretical studies show that it is possible to estimate the imaging frequencies and a relatively small number of frequencies is enough to obtain good images (Sirgue et al., 2004).

In this work we present some numerical experiments of inversion of synthetic seismic data with and without noise. The inversion results of exact data seem to be insensible to the number of imaging frequencies. As expected, using 4 up to 40 imaging frequencies the images are almost the same. On the other hand, the results obtained with noisy data and a fewer number of imaging frequencies (4 and 8) seems to suffer with noise fitting because the error increases while the misfit function is minimized. However, using a greater number of frequencies (20 and 40) the

results are not affected. Note that, using 40 imaging frequencies, the curve of the error is very similar to the curve of the best result obtained with exact data. It shows that the number of frequencies has a role of regularization.

Considering a smoother initial model for the p-wave velocity, we have to use a greater number of iteration to recover the image. However, the results show that for the same number of imaging frequencies, the increasing in the number of frequencies is not enough to obtain better images. This occurs because of noise fitting. The results presented here indicate that the increasing of the number of iterations should be done together with the increasing of the number of imaging frequencies.

It is important to observe that, for the performed experiments, the increasing of the number of frequencies does not affect the convergence of the inversion procedure. Even in cases where only two iterations for each imaging frequency were used, the rates of convergence of the image error are very similar to those that use up to 8 iterations for frequency. This is interesting because, in a context where there is no knowledge about the level of the noise and a budget (in the sense of total number of iterations or solutions of the forward problem) to recover the image, taking into account a greater number of frequencies in the bandwidth with a small number of iterations to each frequency can be a good decision to obtain images less sensitive to noise.

In order to conclude if such strategy is useful in the practice of geophysics, more complex inversion experiments should be performed. First, inversion of synthetic data with different kinds and levels of noise should be done. Furthermore, the reconstruction of more complex models including the density and s-wave velocity should be carried out. Moreover, such strategy should be tested to the inversion of real data.

Acknowledgments

The authors would like to thank PETROBRAS, FAPERJ, and CNPq for the financial support.

References

Aster, R. C., B. Borchers, and C. H. Thurber, 2005, Parameter estimation and inverse problems, 1. ed.: Elsevier Academic Press.

Brossier, R., S. Operto, and J. Virieux, 2010, Which data residual norm for robust elastic frequency-domain full waveform inversion?: Geophysics, 75, R37–46.

Bunks, C., F. M. Saleck, S. Zaleski, and G. Chavent, 1995, Multiscale seismic waveform inversion: Geophysics, 60, 1457–1473.

Fichtner, A., 2011, Full seismic waveform modelling and inversion: Springer.

Guitton, A., and W. W. Symes, 2003, Robust inversion of seismic data using the Huber norm: Geophysics, 68, 1310–1319.

Habashy, T. M., and A. Abubakar, 2004, A general framework for constraint minimization for the inversion of electromagnetic measurements: Progress In Electromagnetics Research, 46, 265–312.

Hu, W., A. Abubakar, T. M. Habashy, and J. Liu, 2011, Preconditioned non-linear conjugate gradient method for frequency domain full-waveform seismic inversion: Geophysical Prospecting, 59, 477–491.

Hustedt, B., S. Operto, and J. Virieux, 2004, Mixed-grid and staggered-grid finite-difference methods for frequency-domain acoustic wave modelling: Geophysical Journal International, 157, 1269–1296.

Li, X. S., and J. W. Demmel, 2003, Superlu dist: A scalable distributed-memory sparse direct solver for unsymmetric linear systems: ACM Trans. Mathematical Software, 29, 110–140.

Liu, D. C., and J. Nocedal, 1989, On the limited memory bfgs method for large scale optimization: Mathematical Programming, 45, no. 1, 503–528.

Marelli, S., H. Maurer, and E. Manukyan, 2012, Validity of the acoustic approximation in full-waveform seismic crosshole tomography: Geophysics, 77, R129–R139.

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.

Sirgue, L., and R. G. Pratt, 2004, Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies: Geophysics, 69, 231–248.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, 49, 1259–1266.

Van den Berg, P. M., and A. Abubakar, 2001, Contrast source inversion method: state of art: Progress In Electromagnetics Research, 34, 189–218.

Zhang, R., M. Sen, and S. Srinivasan, 2013, A prestack basis pursuit seismic inversion: GEOPHYSICS, 78, R1–R11.