ASEG-PESA-AIG 2016
25TH GEOPHYSICAL
CONFERENCE & EXHIBITION
Interpreting the Past, Discovering the Future

Obtaining low frequencies for Full Waveform Inversion by using augmented physics

Eldad Haber*
The University of British Columbia
Dept of EOAS, UBC, Vancouver BC
ehaber@eos.ubc.ca

Eran Treister
The University of British Columbia
Dept of EOAS, UBC, Vancouver BC
erantreister@gmail.com

Elliot Holtham
Kyma Geosciences / Computational Geosciences Inc.
2386 East Mall, Vancouver, BC
elliot@compgeoinc.com

SUMMARY

Full waveform inversion (FWI) is a process in which seismic data in the frequency or time domain are being fitted by changing the velocity of the media under investigation. The problem is nonlinear, and therefore, optimization techniques have been used to attempt and find a geological solution to the problem. The main problem in fitting the data is the lack of low special frequencies. This deficiency often leads to a local minimum and to non-geologic solutions. In this work we discuss how to obtain low frequency information that can augment FWI. We explore two techniques, the first, travel time tomography and the second, controlled-source electromagnetics. We then discuss a framework for joint inversion and show that by considering these problems jointly we are able to steer the direction of FWI towards the global minimum.

Key words: FWI, tomography, CSEM

INTRODUCTION

Full waveform inversion (FWI) is a process in which the velocity of the earth is estimated by using measured wave-field data. The estimation is done by fitting the field data to computed data, that are a simulation of the Helmholtz equation (in frequency), or the wave equation (in time). The data fitting requires some optimization algorithm, typically some descent algorithm that gradually reduces the misfit. To keep the velocity model geologic, regularization is typically added to the process. However, while the method has been investigated over 30 years ago by Tarantola (1987) with little success, it has regained new traction when computing power and algorithms have been investigated and we can now handle bigger volume of data using sophisticated algorithms (see for example Pratt (1998, 1999), Krebs et-al (2009), Epanomeritakis et-al (2008), and Biondi and Almomin (2014)).

Nonetheless, solving the FWI problem is difficult in practice and one typically converges to a local minimum that is not geologically feasible. There are two main problems to consider when discussing the convergence behaviour of FWI algorithms. First, since the problem contains many local minima, standard optimization techniques converge only to a local minimizer of the data regardless to presence of low frequencies in the data. Second, if the data do not contain low frequencies, then there may not be information in the data to lead it to a minimizer that geologically makes sense. If low frequencies exist in the data then it is possible to use a process of frequency continuation in order to obtain the global minimizer. This process is known to be both efficient and stable in producing FWI results. However, most acquisition still does not acquire sufficiently low frequencies in order to use the process of frequency continuation.

A number of authors have attempted to address the problem of the lack of low frequencies by changing the misfit function from least squares to some other misfit. However, there is a fundamental problem with any such approach when low frequency data are missing. To explain, consider that the forward problem is $F(m)$, where m is the earth's model and F is the forward modelling operator. The lack of sensitivity to low special frequencies implies that $F(m) \cong F(m+s)$ where s is some smooth perturbation to the model. It implies that the sensitivity matrix $J = \nabla F$ has an approximate null space that can be characterized by low frequencies. Now, consider any objective function $S(F(m),d)$. The common one is the least square one $S(F(m),d) = ||F(m)-d||_2$ but others have been suggested in the past (see Virieux *et al.* (2009) and Warner *et al.* (2013) for a review). The gradient of the objective function is $g = J^* \nabla S$ and since J does not contain low frequencies the gradient will also not contain low frequencies. Since a descent direction is based on the gradient, it is impossible to recover low frequencies in the model. This fundamental observation suggests that in order to find a geologically feasible minimizer to the data, the physics of the problem needs to be augmented with low frequency information. This low frequency information can then steer FWI in the direction of geologically reasonable models.

In this work we focus on two such techniques. First, we use travel time tomography to replace the low frequency component in the data. We show that by jointly inverting the travel time data and the FWI data it is possible to obtain a model that fits both data sets and therefore, obtain low special frequencies. Second, we show that it is possible to augment the inversion using Controlled Source Electromagnetic (CSEM) data that also yields the low frequency components in the data. Since CSEM images conductivity, we require an objective function that either translates conductivity to velocity or uses the structure in the conductivity model to yield a similar structure in the velocity model.

In the following we present a methodology and a workflow that allow us to use augmented data in order to fit FWI data, starting from a very uninformative starting point.

METHOD

We consider the forward problem that is given by the Helmholtz equation for a constant density media

$$\nabla^2 u + m\omega^2 u = \delta(x - x_s).$$

Here, u is the wavefield, m is the model for the squared slowness, and ω is the frequency. The equation is given with absorbing boundary conditions to mimic the propagation of the wave in an unbounded domain.

Next, we consider the data, d, as an inner product of the form

$$d(\omega, x_r, x_s) = (p_r, u(\omega, x_s)) + \varepsilon$$

where p_r is a sampling operator that measures the field u, at some location x_r and $(\cdot, \cdot)$ is an inner product. The data is typically noisy and we assume that the noise, ε, is iid, Gaussian and with variance σ. Given data that is collected in a number of receiver locations and a number of frequencies we aim to estimate the model, m. This is done by solving the following regularized least squares problem:

$$\begin{aligned} \min_{m_L \le m \le m_H} \mathscr{J}(m) \quad &= \quad \frac{1}{2} \sum_{i,j,k} (d(\omega_i, x_{r_j}, x_{s_k}) - (p_{r_j}, u_{ik}))^2 + \alpha R(m) \\ \text{s.t} \qquad & \nabla^2 u_{ik} + m\omega_i^2 u_{ik} = \delta(x - x_{s_k}) \end{aligned}$$

Here, $R(m)$ is a regularization term (we use either smoothness or total variation, while $m_L < 0$ and $m_H > 0$ are bounds that keep the model physical). To solve the optimization problem a variety of methods are typically used. First order methods such as nonlinear conjugate gradient and limited memory BFGS (Nocedal and Wright, 1999) have the advantage of low memory but converge slowly. Our method of choice is the Gauss Newton method (Pratt *et al*., 1998), which incorporates curvature information and converges faster, especially if the noise level is low.

Solving the problem for all frequencies at once typically yields local minima. Therefore, it is common to solve the problem by frequency continuation, solving first the lowest frequency obtaining a model $m(\omega_1)$ and then, solving again the problem for 2 frequencies, starting from $m(\omega_1)$, and continuing forward by adding more frequencies, each time starting from the previous solution. This yields a stable process that converges to the global minima of the objective function. Nonetheless, in the absent of low frequencies this process cannot be used, and convergence to local minima is often observed. To alleviate this problem we propose a different process. It is fairly well known, that assuming that the wave field has a solution of the form $u = a(x, \omega)\, exp(i\omega\, T(x))$, where T is travel time, and substituting it into the Helmholtz equation we obtain the Eikonal equation for high frequencies

$$|\nabla T|^2 - m = 0, \qquad T(x_s) = 0.$$

This equation models the first arrivals of the waves. Since the travel time is an integral of the model over the ray path, its Jacobian with respect to the model contains mainly low frequencies. Thus, a way to overcome the lack of low frequencies in the data, d, is to extract it from the travel time, T. Travel time tomography has been considered by Benaichouche *et al.,* (2015) and by Li *et al.,* (2013). Here we use the Fast Marching method to solve the forward problem and compute the sensitivities directly.

A second way to obtain low frequencies in the data is by using electromagnetic imaging, and in particular Controlled Source Electromagnetics (CSEM) (Haber, 2014). In this problem the forward problem can be modelled by solving Maxwell's equations

$$\nabla \times \mu^{-1} \nabla \times \vec{E} + i\omega\sigma\vec{E} = -i\omega\vec{s}$$

Here, σ is the conductivity that is inverted for by measuring the electric field, E. In order to use CSEM as a second modality, one requires using a relation that connects the conductivity to seismic velocity. In this work we have used Gassmann's equation (Berryman 2009) to obtain such a relation. Since electromagnetic fields represent a decaying wavefield, the sensitivity with respect to the data contains mainly smooth components and it is "blind" to high special frequency variations in the model. Thus, the CSEM method supplements seismic data that is not sensitive to smooth perturbations in the model. In many cases, CSEM is preferred to travel time data as it is sensitive to deeper parts of the model and is not affected by low velocity zones.

The three techniques are put together through the following optimization problem:

$$
\begin{aligned}
\min_{m_L \le m \le m_H} \mathscr{J}(m) \quad = \quad & \frac{w_{\mathrm{fwi}}}{2} \sum_{i,j,k} (d(\omega_i, x_{r_j}, x_{s_k}) - (p_{r_j}, u_{ik}))^2 + \frac{w_{\mathrm{eik}}}{2} \sum_{j,k} (d_{\mathrm{T}}(x_{r_j}, x_{s_k}) - (p_{r_j}, T_k))^2 + \\
& \frac{w_{\mathrm{em}}}{2} \sum_{j,k} (d_{\mathrm{EM}}(x_{r_j}, x_{s_k}) - (p_{r_j}, \vec{E}_k))^2 + \alpha R(m) \\
\text{s.t} \qquad & \nabla^2 u_{ik} + m\omega_i^2 u_{ik} = \delta(x - x_{s_k}) \\
& |\nabla T_k|^2 - m = 0, \qquad T_k(x_{s_k}) = 0. \\
& \quad \text{or/and} \\
& \nabla \times \mu^{-1} \nabla \times \vec{E} + i\omega\sigma(m)\vec{E} = -i\omega\vec{s}
\end{aligned}
$$

The optimization problem is then solved using the Gauss-Newton method where we start by fitting the Eikonal and/or the CSEM data and then augment the FWI data once low frequency components of the model have been recovered.

RESULTS

To experiment with our model we use the SEG salt model, plotted in the left panel of Figure 1 as a test case. The model is discretized using 512 x 140 cells with dimensions of 15.6m x 15.6m. There are 30 equally spaced sources and 140 receivers placed on the surface and the data is collected at different frequencies ranging from a very low frequency of 0.75 Hz to 14 Hz. The low frequencies are not used for all experiments and we assume to have it so we can compare inversion results with and without it. We start all inversions from a gradient velocity model shown in the right panel of Figure 1. The data is contaminated with 1% noise that is Gaussian.

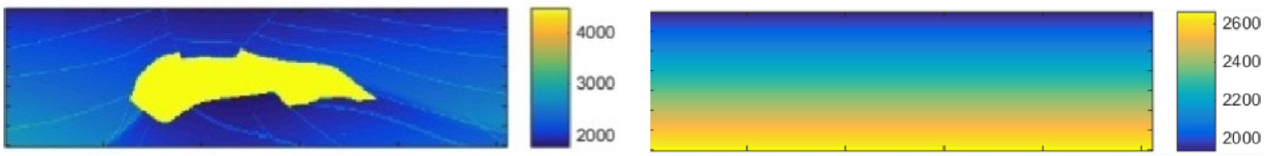

Figure 1: SEG velocity salt model (left panel) and velocity starting model (right panel).

The results of the 3 different experiments are presented in Figure 2.

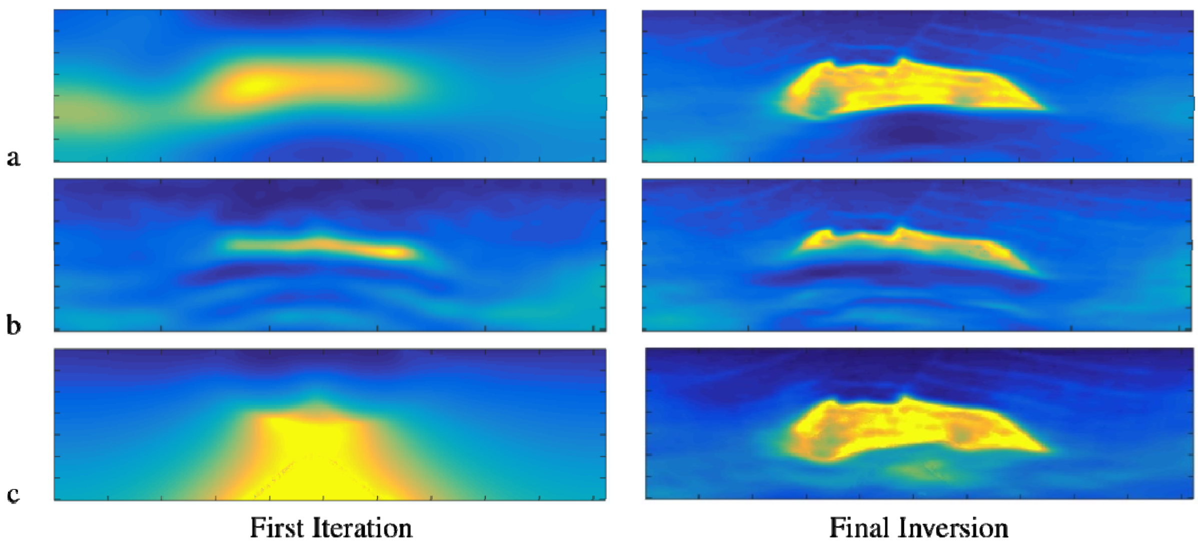

Figure 2: First iteration (left panel) and final inversion (right panel) of a) all frequencies (low and high), b) with first two frequencies missing and c) with the Eikonal equation replacing the first two frequencies.

The left column of the figure shows the inversion result after the first iteration (using the lowest frequency or the Eikonal). The right column shows the final result of the inversion. When low frequency is available, (first row) a blurred representation of the model is obtained in the first iteration. This blurred version is then sharpened when more frequencies are added. When the low frequencies are missing, the lowest frequency yields high frequency features in the initial model. Since high frequencies do not contain low frequency content, the inversion does not recover and the results are far from the true model. Finally, when the Eikonal is used to replace the low frequency data, a low frequency representation of the model is also obtained. Even though this representation is incorrect at depth, the additional frequencies manage to overcome this and the final result is equivalent to the result obtained by using the low frequency data.

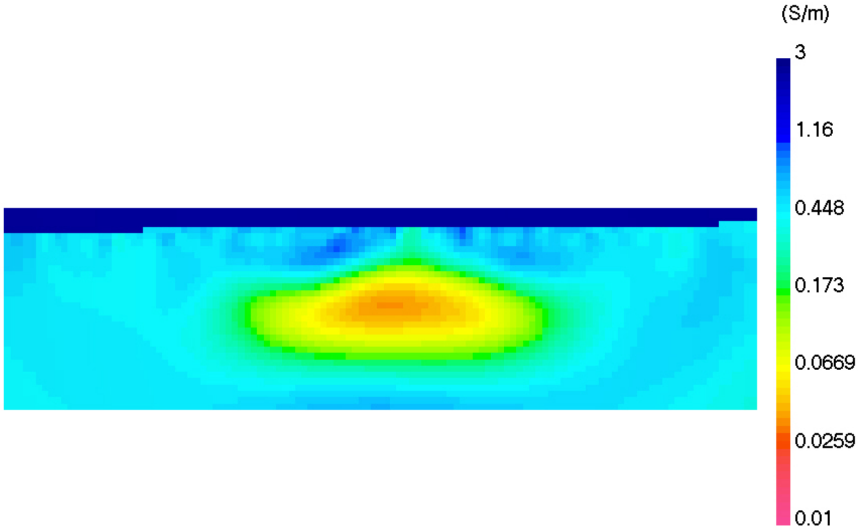

Figure 3: Initial conductivity model obtained by using CSEM data. Note the identification of the bottom of the salt body.

In Figure 3 we show the inversion results obtained by simulating a CSEM experiment over the same domain as before to show that CSEM data can replace low frequency data. As can be observed, the salt body is placed at the right location and even its bottom is imaged with some accuracy, which gives the FWI a good starting model in order to converge to an improved result.

CONCLUSIONS

In this work we have explored a methodology that aids full waveform inversion to converge to the global minimum in the absence of low frequency data. The method is based on the extraction of travel time from high frequency data or collecting CSEM data and using either the Eikonal or Maxwell's equations in order to model the travel time or conductivity. We use these augmented physical experiments instead of low frequency data to jointly invert with the rest of the data using a frequency continuation process. Since both travel time and CSEM inversions are sensitive to low spatial frequencies modes in the model, they yield an initial model that enables a recovery that is a close approximation to the true model. Furthermore, since we jointly invert the full waveform and travel time/CSEM data, our final model is consistent for both physical models.

While our method seems to be robust in the presence of noise, it has two main limitations. First, long offset data must be recorded in order to have a meaningful first arrival inversion. Similarly, CSEM data needs to be collected over the same domain as the seismic one in order to image the low frequencies using Maxwell's equations. Second, our approach requires travel time picking and EM processing. This increases the cost of FWI and requires codes that can handle both EM and seismic inversion preferably within the same computational framework.

ACKNOWLEDGMENTS

This work was supported by Kyma Geosciences Inc. and MITACS.

The research leading to these results has received funding from the European Union's - Seventh Frame- work Programme (FP7/2007-2013) under grant agreement no 623212 - MC Multiscale Inversion.

REFERENCES

Benaichouche, A., Noble, M. and Gesret, A. [2015] First Arrival Traveltime Tomography Using the Fast Marching Method and the Adjoint State Technique. In: 77th EAGE Conference Proceedings.

Biondi, B. and Almomin, A. [2014] Simultaneous inversion of full data bandwidth by tomographic full-waveform inversion. Geophysics, 79(3), WA129–WA140.

Epanomeritakis, I., Akcelik, V., Ghattas, O. and Bielak, J. [2008] A Newton-CG method for large-scale three-dimensional elastic full-waveform seismic inversion. Inverse Problems.

Haber, E. 2015, Computational Methods in Geophysical Electromagnetics, SIAM Philadelphia.

Krebs, J.R., Anderson, J.E., Hinkley, D., Neelamani, R., Lee, S., Baumstein, A. and Lacasse, M.D. [2009] Fast full-wavefield seismic inversion using encoded sources. Geophysics, 74(6), WCC177– WCC188.

Li, S., Vladimirsky, A. and Fomel, S. [2013] First-break traveltime tomography with the double-square- root eikonal equation. Geophysics, 78(6), U89–U101.

Nocedal, J. and Wright, S. [1999] Numerical Optimization. Springer, New York.

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

Pratt, R.G., Shin, C. and Hick, G. [1998] Gauss–Newton and full Newton methods in frequency–space seismic waveform inversion. Geophysical Journal International, 133(2), 341–362.

Tarantola, A. [1987] Inverse problem theory. Elsevier, Amsterdam.

Virieux, J. and Operto, S. [2009] An overview of full-waveform inversion in exploration geophysics. Geophysics, 74(6), WCC1–WCC26.

Warner, M., Ratcliffe, A., Nangoo, T., Morgan, J., Umpleby, A., Shah, N., Vinje, V., S□tekl, I., Guasch, L., Win, C. et al. [2013] Anisotropic 3D full-waveform inversion. Geophysics, 78(2), R59–R80.