



Dynamic Relative Entropy Regularization for Full Waveform Inversion

Danilo Santos Cruz^{1*}, Joao M. de Araújo² and Carlos C. Nascimento da Silva³

¹ Petróleo Brasileiro S.A. ² DFTE - CISCOPET /Universidade Federal do Rio Grande do Norte

³ CCET - DGEF /Universidade Federal do Rio Grande do Norte

Copyright 2019, SBGf - Sociedade Brasileira de Geofísica.

This paper was prepared for presentation at the 16th International Congress of the Brazilian Geophysical Society, held in Rio de Janeiro, Brazil, 19-22 August, 2019.

Contents of this paper were reviewed by the Technical Committee of the 16th International Congress of the Brazilian Geophysical Society and do not necessarily represent any position of the SBGf, its officers or members. Electronic reproduction or storage of any part of this paper for commercial purposes without the written consent of the Brazilian Geophysical Society is prohibited.

Abstract

Full-Waveform Inversion is an ill-posed inverse problems so, eventually, regularization schemes are needed to constrain the solution. In most geophysical methods, prior information is available in wells, and these can be used to increase the reliability of the inversion result.

The regularization term with a priori information can significantly reduce the sensitivity of the inversion to incorrect initial conditions since the information a priori will contribute mainly in the regions of low illumination. The weighting of the term of the relative entropy was done in a dynamically decreasing way. Once the problem of cycle skipping has been solved, the importance of a priori information is gradually reduced within the objective function and is hence only directed by seismic data.

For this, we propose the introduction of the relative entropy term in the definition of the FWI objective function, that is, the misfit of the data plus the relative entropy. This will allow us the introduction of the previous information in the FWI workflow. In the synthetic example of Marmousi, a significant improvement was obtained by using the entropy term from the previous model for noise-free and noisy synthetic data.

Introduction

Regularization techniques, also known as preconditioning, can help to address the issue of non-uniqueness in ill-posed problems. As an example of these techniques, we can mention the proposal of (Tikonov and Arsenin, 1977) wherein the optimization stage it was sought to find a smooth model that explained the data.

The preconditioning technique that acts as a smooth operator in updating the model can add (Operto et al., 2006) strong previous features of the expected structure through the directive Laplacian smooth, for example (Guitton et al., 2010).

For the specific problem of full wave inversion (FWI), regularization schemes were developed that preserve edges and contrast through the L1 norm penalty of the model (Peters and Herrmann, 2017). All these regularization techniques allow the inversion process to be stabilized assuming a particular representation or structure

of the speed model (smoothness, sparsity, etc.).

During an exploratory process, no prior model information is normally used in full waveform inversion schemes since generally no information is available for application. However, for the purposes of monitoring reservoirs where many different types of data have already been collected, such as sonic profiles, stratigraphic data, etc., one may wish to use such information in fwi schemes, as may be done for other construction techniques of velocity models.

Based on this (Asnaashari et al., 2013) proposed a regularized FWI scheme based on a two-term model penalty in the misfit definition: The Tikonov term to ensure smoothness and a generalized Tikhonov term to attract the inversion towards a prior model. In addition, they investigated the influence of a simple decreasing dynamic weighting of the previous model term.

In this work, we will use the dynamic smoothing scheme for the FWI proposed by (Asnaashari et al., 2013) to study the contribution of the entropy of the model. The application of the method will be made in the synthetic model of the Marmousi aiming at the recovery not only of the shallower region but also of the deep part of the model

Methodology

The Full Waveform Inversion is aimed at minimizing the residue between observed (real data) and modelled wave fields. This problem is usually solved through an iterative local optimization being developed with linearization of least squares problems.

Even if the problem is linearized, it remains ill-posed. For this reason, additional information is usually required, and these are introduced to the problem through regularization.

Our objective function is based on the classical formulation of the FWI (Tarantola, 1987), plus a penalty term associated with the certainty measure of random variables of the model parameters (Skilling and S.F.Gull, 1985), in other words, a term referring to the relative entropy of the model parameters.

Thus, the cost function can be written using the L_2 norm as:

$$C(m) = \underbrace{\|\mathbf{u}_{obs} - \mathbf{u}(m)_{calc}\|^2}_F + \mu\gamma \underbrace{\left\| \sum_{k=1}^N m_k \ln \left(\frac{m_k}{m_k^r} \right) - (m_k - m_k^r) \right\|^2}_S \quad (1)$$

where \mathbf{u}_{obs} are the observed data and \mathbf{u}_{calc} are simulated data generated by a wave equation which depends on the

unknown model parameters $\mathbf{m} = m_i; i = 1, N_m$. These model parameters should be determined through the inverse procedure. The trade-off parameter μ balances the data fitting and the minimum negative entropy regularization, m_k is the model parameter at position k , m_k^r is the prior information. γ is a normalization factor expressed by the ratio between the misfit term gradient trace and the entropy term gradient trace, ie:

$$\gamma = \frac{tr(\nabla F)}{tr(\nabla S)} \quad (2)$$

The objective of this parameter is to maximize the importance of the information a priori in the first iterations and consequently to guide the inversion. As the model parameters are iterated, this term progressively decreases the weight of the regularization term, thus avoiding that the information a priori leaves footprints in the inverted final model (Asnaashari et al., 2013).

Minimizing the objective function (equation 1) leads to the normal equation system which can be written as:

$$H_m \Delta \mathbf{m} = -g_m \quad (3)$$

where g_m and H_m are the gradient and the Hessian of the objective function, respectively. The gradient expression can be written with two terms as:

$$g_m = -\mathbf{J}'(\mathbf{u}_{obs} - \mathbf{u}(\mathbf{m})_{calc}) + \mu \gamma S \sum_{k=1}^N \left[\ln \left(\frac{m_k}{m_k^r} \right) \right] \quad (4)$$

The sensitivity matrix $\mathbf{J} = \frac{\partial \mathbf{u}(\mathbf{m})}{\partial \mathbf{m}}$ is composed by the Fréchet derivatives of the synthetic data with respect to the model parameters. For the data contribution, the gradient is efficiently computed with an adjoint formulation (Plessix, 2006) without an explicit computation of \mathbf{J} . We solve the normal equation system by a quasi-Newton method using *LBF GS - B* scheme (Nocedal and Wright, 2006). We have performed the inversion in the time domain although the approach is also valid in the frequency domain.

Results

Our study was carried out in the Marmousi P wave velocity model as can see in Figure 1a. The density, in this case, was considered homogeneous.

The acquisition geometry contains 17 explosive sources, located along a horizontal line 48m, each 540m. The configuration is the same for all the shots, one fixed horizontal receiver line at 24m depth in all surface. A Ricker wavelet source with a central frequency of 8Hz is used for all shots.

The grid is regular, with the grid size equal to 24m, and it is consistent for modelling and inversion. The time seismograms are generated using finite-difference modelling in the time-domain with a fourth-order stencil in space and a second-order integration in time. Perfectly-matching-layer (CPML) absorbing boundary conditions (Komatitsch and Martin, 2007; Pasalic and McGarry, 2010) are used for nonreflecting boundaries.

For the initial model, we did a smoothing in the real model (Figure 1a), as can be seen in the figure in Figure 1b. For prior information, the true velocity values inside the two wells are considered to be known, as sonic profile are often used in such case. From the well profiles, a linear interpolation was performed for the whole model as can be seen in Figure 1c.

The first (Figure 2a) is performed with a "classical" FWI method, that is, without considering any prior model (equivalent $\mu = 0$ in eq.(1)). Although the shallow part of the model has been satisfactorily recovered, there are differences in the deep part. Probably this optimization is trapped in a local minimum due to cycle skipping problem, principally in the regions indicated by the arrows. In these positions, there is an error in the position and inclination of the events.

In the second test, we try to answer the following question: If this a priori information is so good as to contribute to the inversion, why not use it as the initial model. Therefore, we use prior information as the initial model.

Figure 2b shows the velocity model reconstructed after inversion. In this case, we also observed a good recovery in the shallow part of the model. However, it is clear the presence of footprints coming from the interpolated initial model in the final result. Figure 2c shows the result of the inversion using the systematic of the dynamic power for the entropy. The initial model used in this process is that shown in Figure 1b and with the prior information of Figure 1c. The lambda parameter was set to 1. The analysis of this result shows a gain in the quality of the inverted result, mainly with respect to the reliability of the structural positioning of the model.

Figure 3 shows the contribution of the relative entropy during the FWI process. Note that the entropy term helps to decrease the misfit of the data.

In the third step, we studied the effect of noise on the FWI. We keep the same acquisition configuration, whereas an artificial Gaussian noise in the range of 0 – 30Hz, the bandwidth of the source, is added to the true noise-free data. The signal-to-noise ratio (S/N) is around 6dB.

Tests were also performed using the initial model shown in figure 1b with and without regularization. We also used the model shown in Figure 1c. The results with noisy data remain consistent with the noise-free tests: The relative entropy term still drives the inversion toward a more realistic and accurate final model, although it is noisier.

We also perform the inversion using a linear velocity gradient as the initial model. Even though we have a priori information available, when we use an initial model that is very far from the true model, this proposed scheme is not capable of providing a satisfactory inverted model.

Conclusion

We have proposed a regularized FWI scheme based on relative entropy model in the misfit definition. Entropy model penalty allows the non-uniqueness issue of FWI to be reduced, particularly for time lapsed applications where prior such as sonic log can be available. We also show that the use of information a priori contributes to the decrease of the misfit data.

Acknowledgments

The authors acknowledge the support of Petrobras.

References

- Asnaashari, A., R. Brossier, S. Garambois, F. Audebert, P. Thore, and J. Virieux, 2013, Regularized seismic full waveform inversion with prior model information: *Geophysics*, **78**, 25–36.
- Guitton, A., G. Ayeni, and G. Gonzales, 2010, A preconditioning scheme for full waveform inversion: SEG Technical Program Expanded Abstracts, Expanded Abstracts, 1008–1012.
- Komatitsch, D., and R. Martin, 2007, An unsplit convolutional perfectly matched layer improved at grazing incidence for the seismic wave equation: *Geophysics*, **72**, SM155–SM167.
- Nocedal, J., and S. J. Wright, 2006, *Numerical optimization*, second ed.: Springer.
- Operto, S., J. Virieux, J. X. Dessa, and G. Pascal, 2006, Crustal imaging from multifold ocean bottom seismometers by frequency-domain full-waveform tomography: application to the eastern nankai trough: *Journal of Geophysical Research*, **111**.
- Pasalic, D., and R. McGarry, 2010, Convolutional perfectly matched layer for isotropic and anisotropic acoustic wave equations: Presented at the SEG Technical Program Expanded Abstracts 2010, Society of Exploration Geophysicists.
- Peters, B., and F. J. Herrmann, 2017, Constraints versus penalties for edge-preserving full-waveform inversion: *The Leading Edge*, **36**.
- 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.
- Skilling, J., and S.F.Gull, 1985, *Algorithms and applications in maximum-entropy and bayesian methods in inverse problems*: SPRINGER.
- Tarantola, A., 1987, *Inverse problem theory: Methods for data fitting and model parameter estimation*, first ed.: ELSEVIER.
- Tikonov, A., and V. Arsenin, 1977, *Solution of ill-posed*

problems: Winston, Washington, DC.

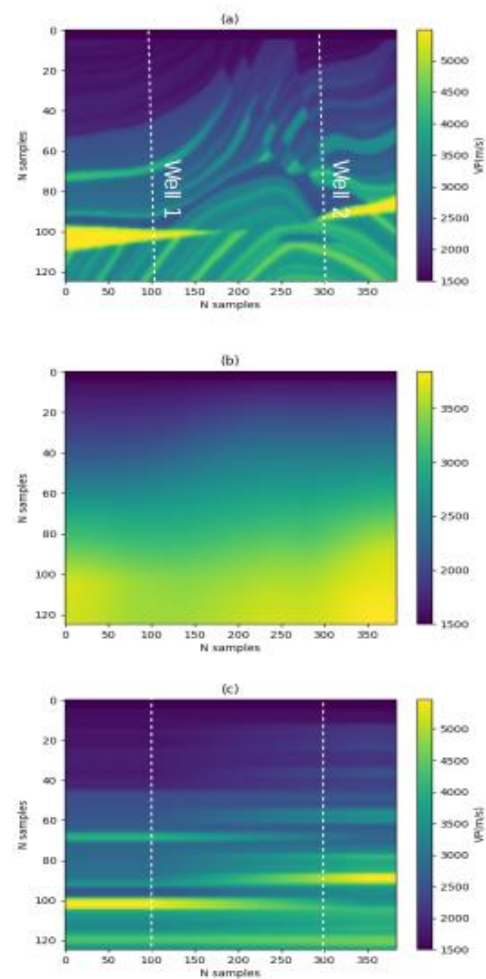


Figure 1: (a) True model with wells that were used to extract the sonic log; (b) The smoothed model; (c) the prior model built by linear interpolation of the two well velocity profiles.

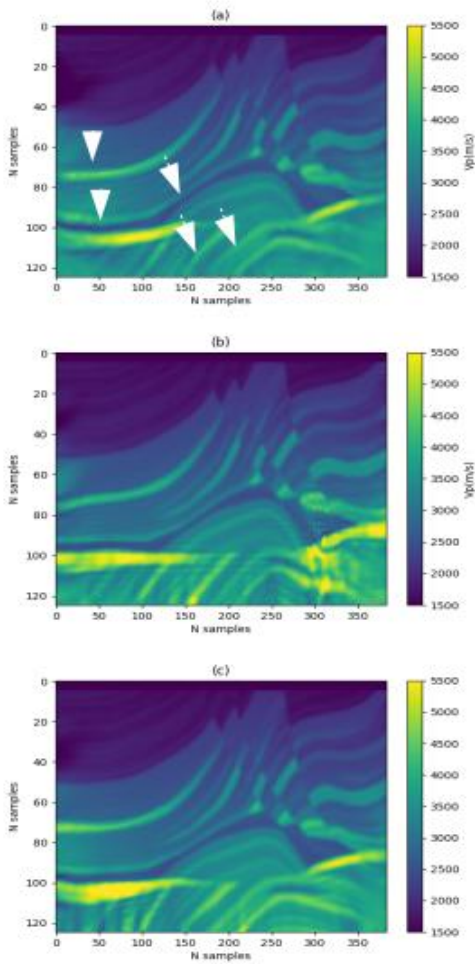


Figure 2: The inverted V_p models by FWI, (a) starting from the smooth initial model with "classical" FWI (b) starting from the interpolated initial model with "classical" FWI (c) as case (a) using a relative entropy

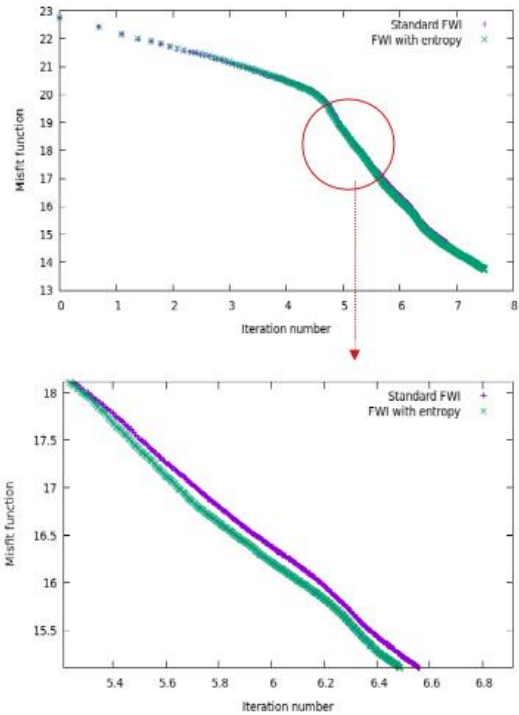


Figure 3: Comparison of the misfit function curves in the case of using the standard FWI and relative entropy. On the left for all iterations and on the right a detail.