Fast Spatially Coupled Bayesian Linearized Acoustic Seismic Inversion in Time Domain

Fernando Bordignon[1], Leandro Figueiredo[2], Mauro Roisenberg[1], Bruno B. Rodrigues [3]

[1]Instituto de Informática e Estatística
Universidade Federal de Santa Catarina (UFSC)

[2]Departamento de Física
Universidade Federal de Santa Catarina (UFSC)

[3]CENPES/PDGP/CMR
Petróleo Brasileiro SA - PETROBRAS

{bordi,mauro}@inf.ufsc.br

Abstract. *Bayesian methods for seismic inversion assume multi-Gaussian distributions for the variables involved and can apply a linearization to the forward model to offer a mathematically tractable solution and uncertainty analysis. Results of the maximum a posteriori are robust, produced fast and have a high synthetic to real seismic correlation. The drawbacks of this inversion method are related to the lack of spatial continuity patterns, when the solution is computed trace-by-trace. This work proposes a methodology to tackle the scalability problem of the Bayesian linearized inversion when adding spatial continuity in the time domain. A sliding window and a manipulation of the equations are used to solve the problem locally and avoid boundary effects.*

1. Introduction

Acoustic seismic inversion aims to infer the acoustic impedance property of the Earth's subsurface via recorded seismic reflection data. It plays a key role in hydrocarbons reservoir modeling and characterization, and it is an ill-posed, nonlinear inverse problem [Tarantola 2005]. Despite its nonlinearity and ambiguousness, methodologies that linearizes the problem were proposed in the last decades [Buland and Omre 2003, Figueiredo et al. 2014]. The methodology results that are comparable to commercial software and is broadly used in the industry nowadays [Figueiredo et al. 2014, Grana and Della Rossa 2010].

The linearized Bayesian acoustic inversion uses a framework based on the Bayes rule and assumes log-normal distributions for the acoustic impedance and normal distribution for the error term of the seismic data. The solution is computed fast in the time domain if done trace-by-trace, because one can reuse calculations when the area of interest have stationary covariances. Performing the inversion trace-by-trace limits the spatial coupling of the results because no horizontal correlation is imposed, inheriting the continuity solely from the seismic data [Buland et al. 2003].

To impose lateral continuity into the solution, it is necessary to define a model covariance matrix that is $n \times n$ being n the number of cells in the area of interest. This leads to an exponential complexity on n, i.e. $O(n^3)$, which makes the process impractical due to the amount of memory needed to process big volumes of seismic data. To overcome this issue, [Buland et al. 2003] proposes an inversion in the frequency domain, lowering the complexity to $O(n \log n)$.

This paper proposes a moving window technique to account only for the neighboring seismic traces while imposing lateral continuity, overcoming the border effects by using a sliding window that inverts only the central trace at a time. The algorithm execution time is highly dependent on the window size but it is of linear complexity on the number of inversion cells, i.e. $O(n)$. Two applications examples are shown comparing the trace-by-trace inversion with the proposed method. The run time of the proposal is acceptable, exhibiting up to a tenfold increase in run time compared to the trace-by-trace method for the studied cases, in addition to the algorithm being easy to implement.

2. Bayesian Linearized Inversion

In the discrete domain, the seismic data $\boldsymbol{d}$ is given from its relation to the reflectivity $\boldsymbol{r}$ by:

$$\boldsymbol{d} = \boldsymbol{S}\boldsymbol{r} + \boldsymbol{e} \tag{1}$$

where $\boldsymbol{S}$ is the convolutional matrix constructed with a known wavelet. The relation between the reflectivity and the impedance z is given by:

$$r(t) = \frac{z(t+\delta t) - z(t)}{z(t+\delta t) + z(t)} \tag{2}$$

With the reflectivity smaller than 0.3, the relation above can be approximated by [Stolt and Weglein 1985]:

$$r(t) = \frac{1}{2}\Delta \ln(z(t)) \tag{3}$$

If we consider the model vector to be log-normal, e.g. $\boldsymbol{m} = \ln(\boldsymbol{z})$, and the differential matrix $\boldsymbol{D}$, the relationship between the seismic and the acoustic impedance is given by the linear operator $\boldsymbol{G} = \frac{\mathbf{1}}{\mathbf{2}}\boldsymbol{S}\boldsymbol{D}$ plus a white noise error term $\boldsymbol{e}$ as follows:

$$\boldsymbol{d} = \boldsymbol{G}\boldsymbol{m} + \boldsymbol{e} \tag{4}$$

The Gaussian likelihoods between the seismic data and the model parameters are written as in the following equations:

$$p(\boldsymbol{d}|\boldsymbol{\mu_d}, \boldsymbol{\Sigma_d}) = N(\boldsymbol{\mu_d}, \boldsymbol{\Sigma_d}), \tag{5}$$

$$p(\boldsymbol{m}|\boldsymbol{\mu_m}, \boldsymbol{\Sigma_m}) = N(\boldsymbol{\mu_m}, \boldsymbol{\Sigma_m}), \tag{6}$$

where $\boldsymbol{\mu_d} = \boldsymbol{G}\boldsymbol{m}$, $\boldsymbol{\Sigma_d}$ the covariance matrix of the seismic, $\boldsymbol{\mu_m}$ the low frequency model (LFM) and $\boldsymbol{\Sigma_m}$ the covariance matrix for the model parameters, e.g. defined by an normal decaying correlation neighborhood as in [Figueiredo et al. 2014]:

$$\boldsymbol{\nu}_{t,t'} = \sigma_m^2 exp\left(-\frac{(t-t')^2)}{L^2}\right), \tag{7}$$

where L is the desired correlation distance and $\boldsymbol{\nu}_{t,t'}$ is then multiplied by the variance to obtain the covariance.

Making use of this Bayesian linearized framework, the posterior distribution is written according to:

$$p(m|d, s, \mu_m, \sigma_d^2, \sigma_m^2) \propto p(d|s, m, \sigma_d^2)p(m|\mu_m, \sigma_m^2) \tag{8}$$

where s is the wavelet and σ_d^2 and σ_m^2 are seismic and model variances respectively.

The posterior mean and covariance matrix are given by [Figueiredo et al. 2014]:

$$\boldsymbol{\mu}_{m|} = \boldsymbol{\mu}_m + \boldsymbol{\Sigma}_m\boldsymbol{G}^T(\boldsymbol{G}\boldsymbol{\Sigma}_m\boldsymbol{G}^T + \boldsymbol{\Sigma}_d)^{-1}\left(\boldsymbol{d}_o - \boldsymbol{G}\boldsymbol{\mu}_m\right), \tag{9}$$

$$\boldsymbol{\Sigma}_{m|} = \boldsymbol{\Sigma}_m - \boldsymbol{\Sigma}_m\boldsymbol{G}^T(\boldsymbol{G}\boldsymbol{\Sigma}_m\boldsymbol{G}^T + \boldsymbol{\Sigma}_d)^{-1}\boldsymbol{G}\boldsymbol{\Sigma}_m. \tag{10}$$

where the mean is also referred to as maximum a posteriori (MAP), as the mean of a multivariate normal distribution is the most likely values of the vector, or the vector values with the highest likelihood.

The methodology presented is to invert a 3D seismic cube in a trace-by-trace manner. In this case, nothing but vertical correlations are imposed to the results, allowing horizontal continuity to be derived from seismic data alone. When performing the trace-by-trace inversion, one can reuse the computation of the matrix inversion in Equation (9) for an area with similar covariances, reducing the computational cost to a matrix inversion, sized according to the number of vertical samples, and matrix-vector products for each trace.

It is possible to invert all the 3D region of interest at once, defining a model covariance matrix $\boldsymbol{\Sigma}_m$ that imposes correlations in every desired direction. A drawback of this approach is the quadratic growth of the covariance matrices, as the number of cells to be inverted grows, which then propagates its size to the matrix inversion in Equations (9) and (10) yielding an exponential computational complexity, i.e. $O(n^3)$.

Results of the trace-by-trace method are reported to be robust and can be considered a smooth representation of the subsurface properties of interest. The procedure is easy to implement and faster compared to commercial available software whereas producing similar impedance models [Figueiredo et al. 2014]. Although the methodology does not have the practical ability to model complex spatial continuities, it is suitable for an expeditious inversion used for interpretation or simpler purposes.

3. Fast Bayesian Linearized Inversion with Spatial Coupling

The Bayesian inversion methodology presented is the solution to a linear problem in the form $\boldsymbol{x} = \boldsymbol{A}\boldsymbol{y} + \boldsymbol{b}$, where $\boldsymbol{x}$ is the acoustic impedance, $\boldsymbol{A} = \boldsymbol{\Sigma}_m\boldsymbol{G}^T(\boldsymbol{G}\boldsymbol{\Sigma}_m\boldsymbol{G}^T + \boldsymbol{\Sigma}_d)^{-1}$, $\boldsymbol{y} = (\boldsymbol{d}_o - \boldsymbol{G}\boldsymbol{\mu}_m)$ and $\boldsymbol{b} = \boldsymbol{\mu}_m$.

Adding spatial coupling, as the problem is linearized, does not change the form of the solution. The spatial coupling is added via a linear function of the $\boldsymbol{y}$ vector. In other words, $\boldsymbol{A}$ gets bigger and more complicated but it is still a linear operator over $\boldsymbol{y}$. As the area of interest grows, $\boldsymbol{A}$ becomes even more sparse. This effect is due to the correlation length being limited to the nearest neighbours, i.e. the solution of the problem only depends on the samples that are in a close horizontal vicinity, having very little relation to distant samples.

Using the Equation 7 to define horizontal correlation, the sparsity of the matrix A is notable for $L = 2$. At the fourth horizontal index away from the diagonal, i.e. $t - t' = 4$, the covariance is already 50 times smaller than the diagonal. Therefore, it is reasonable to truncate the horizontal covariance at two times the L distance in this case.

To implement a linear complexity inversion methodology we explore the linearity of the solution and use a truncation of the lateral continuity model. Assume the goal is to invert a 3D grid with coordinates $G \in \mathbb{Z}^3$ where $|G| = n = ijk$, being i the number of cells in x direction, j the number of cells in y direction and k the number of cells in time t. Assume a contiguous window $W \subset G$ where $|W| = w^2k$ being w the size of the window both in x and y directions with $w < i$ and $w < j$.

Following, the problem is fully defined as we would invert only the traces inside the 3D window W, i.e. we define the covariance matrices and stack the seismic traces and LFM traces in its respective vectors $\boldsymbol{d}_o$ and $\boldsymbol{\mu}_m$. Accordingly, the covariance matrices will have size $w^2k \times w^2k$. Next, a new covariance matrix $\boldsymbol{\Sigma}'_m$ is defined using only the lines of $\boldsymbol{\Sigma}_m$ correspondent to the central trace of W, i.e. only the lines relative to the coordinates $(\lceil \frac{w}{2} \rceil, \lceil \frac{w}{2} \rceil, t) \forall t \in k$ are selected from $\boldsymbol{\Sigma}_m$ to compose $\boldsymbol{\Sigma}'_m$.

Finally a operator is computed with a modified Equation 9 as follows:

$$\boldsymbol{O} = \boldsymbol{\Sigma}'_m \boldsymbol{G}^T (\boldsymbol{G} \boldsymbol{\Sigma}_m \boldsymbol{G}^T + \boldsymbol{\Sigma}_d)^{-1} \tag{11}$$

The matrix $\boldsymbol{O}$ has k lines and w^2k columns, which when multiplied by $(\boldsymbol{d}_o - \boldsymbol{G}\boldsymbol{\mu}_m)$, defined with all samples inside the window W, gives the impedance solution for the central trace of the window. The window is then slided to the next x or y coordinate of the 3D grid by 1 index, overlapping most of the samples with the previous window. The $\boldsymbol{O}$ operator is applied again to compute the next central impedance trace of the window, wich is 1 index far from the previous central trace. The steps are repeated for every possible window in coordinates of the 3D grid G. The central traces are saved to compose the final solution to the problem. Figure 1 shows the graphic representation of the sliding window inside the 3D grid with its central trace at a (x, y) plane.

The truncation of $\boldsymbol{O}$ can also be applied at the vertical direction so that theoretically the algorithm will have $O(n)$ complexity. But, for most cases, the size k of the 3D grid is limited to a maximum of 1000 samples. Therefore, the procedure described above is of complexity $O(n)$ if k is considered a constant, i.e. the growth of n is mainly due to the growth of i and j.

4. Application Examples

In this section two datasets were used to compare the results of the inversion with omnidirectional correlation, also called spatially coupled inversion or coupled inversion for short, against the trace-by-trace inversion.

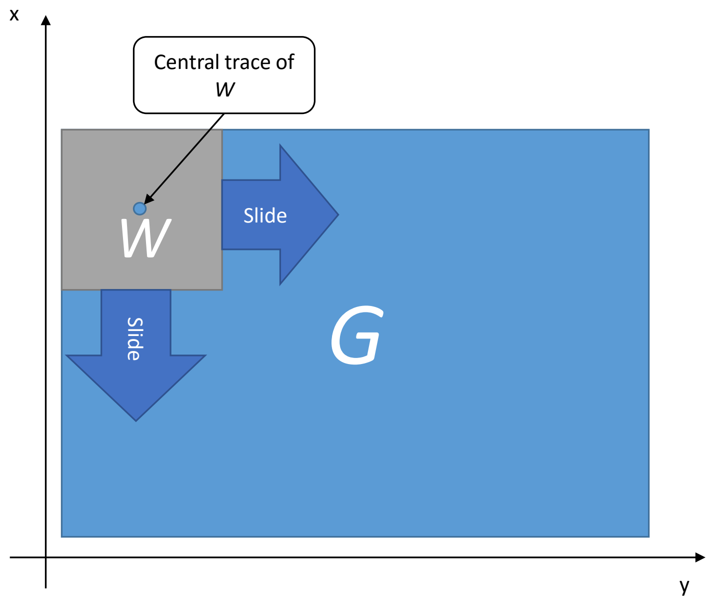


Figure 1. Graphic representation of the sliding window at the 3D grid

4.1. Synthetic Dataset

The first experiment was conducted on a synthetic dataset with $101 \times 101 \times 90$ cells. The seismic data has no noise, hence a trace-by-trace inversion of the original data is used as a gold standard for the inversion. A white noise with amplitude of 10% of the seismic data RMS value was added to the seismic data, therefore it is possible to compare the trace-by-trace inversion with the coupled inversion, evidentiating the lateral continuity imposed. Figure 2 depicts a vertical section of the trace-by-trace inversion with the original noise free seismic data.

For this case, the correlation distance L was set to 1.3 in all directions. A constant low frequency model was used with $13000g/cm^3m/s$ for all cell nodes. A previously extracted wavelet was provided by geophysicists. The model variance σ_m^2 from Equation 7 was set to 0.0077 which was extracted from the logarithm of the high-pass filtered impedance well logs (cutoff frequency of $8Hz$). The seismic variance σ_d^2 was set to 10% of the seismic data RMS value. The window W was defined as being of size $5 \times 5 \times 90$ for the inversion with horizontal correlation.

The trace-by-trace inversion took 2 seconds to run and yielded a synthetic to real seismic correlation coefficient of 0.99. Figure 3 shows a vertical section of the trace-by-trace inversion of the noisy seismic data. Notice the vertical stripes of noise that appear due to the lack of lateral coupling of the results.

The coupled inversion took 18 seconds to run for the entire grid, yielding a synthetic to real seismic correlation coefficient of 0.99. The impedance results are shown at Figure 4. As a result of spatial coupling, some noise is filtered and it is possible to see the improvement at the definition of the deepest layer at around the time index 80, which is now smoother and has more contrast compared to the trace-by-trace inversion.

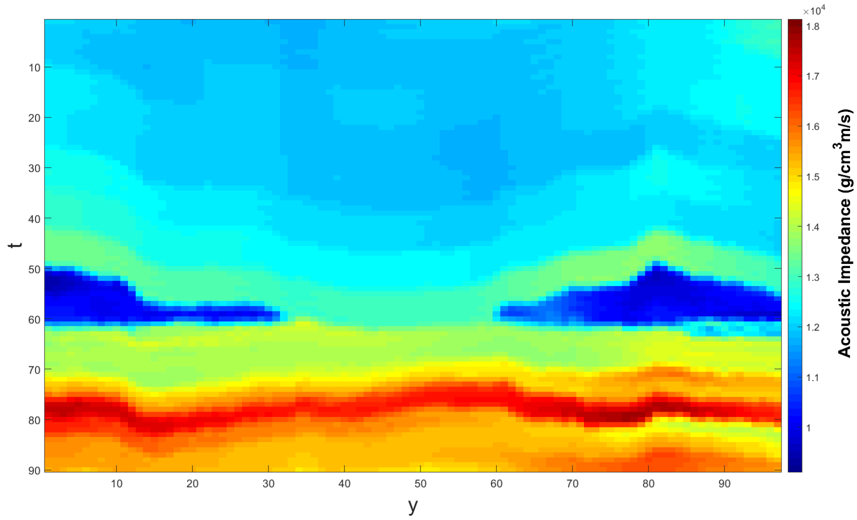


Figure 2. Impedance result from trace-by-trace inversion without noise for the synthetic dataset

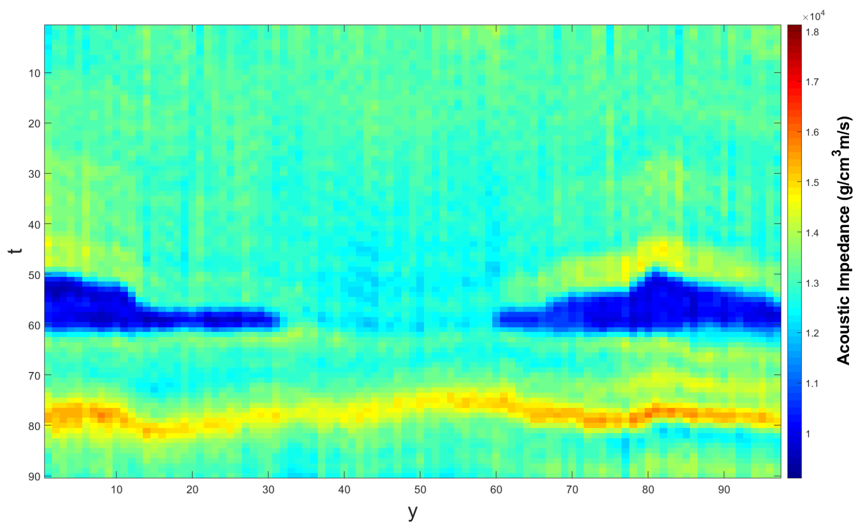


Figure 3. Impedance result from trace-by-trace inversion with noise for the synthetic dataset

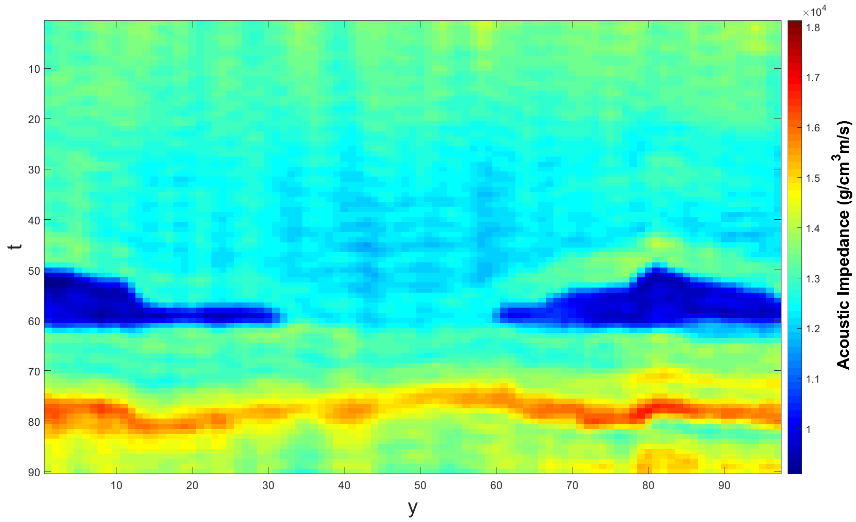


Figure 4. Impedance result from the coupled inversion for the synthetic dataset

4.2. Real case application

For this case, a real dataset was used that has a relatively high signal to noise ratio. A 2D arbitrary line that passes through the 4 wells was selected for the inversion. The section has 707 traces with 250 time samples each. The window W has size 5×250 because the inversion is 2D. The trace-by-trace inversion took 0.6 seconds to execute, while the coupled inversion 2.2 seconds. The acoustic impedance result of the trace-by-trace inversion is shown at Figure 5.

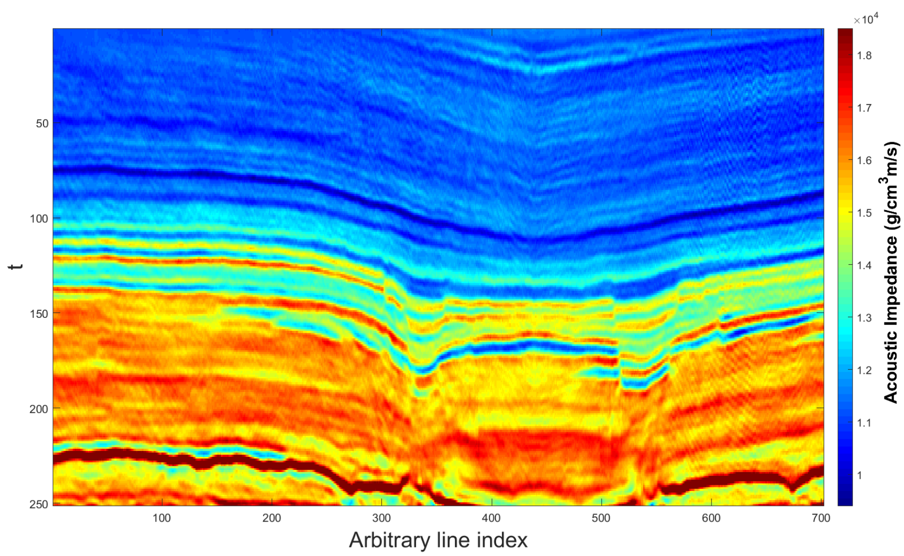


Figure 5. Impedance result from trace-by-trace inversion with noise for the real dataset

In this case it is possible to notice some noise throughout the section, specially at the upper right part. Note that the vertical continuity is present, while at the noisy areas there are abrupt discontinuities in the horizontal direction, indicating noise still left on the data.

The coupled inversion result is depicted at Figure 6. In this case the horizontal continuity is present, as it is possible to see mainly at the upper right area of the section.

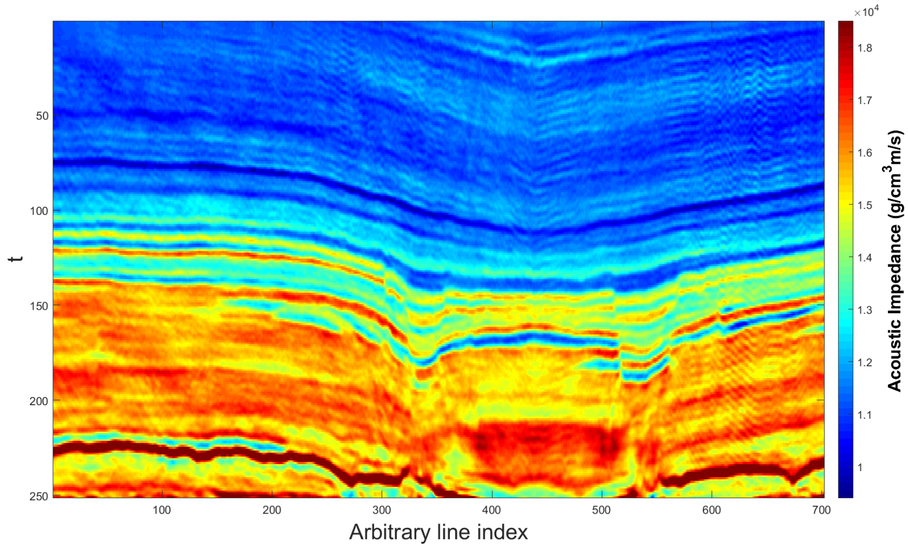


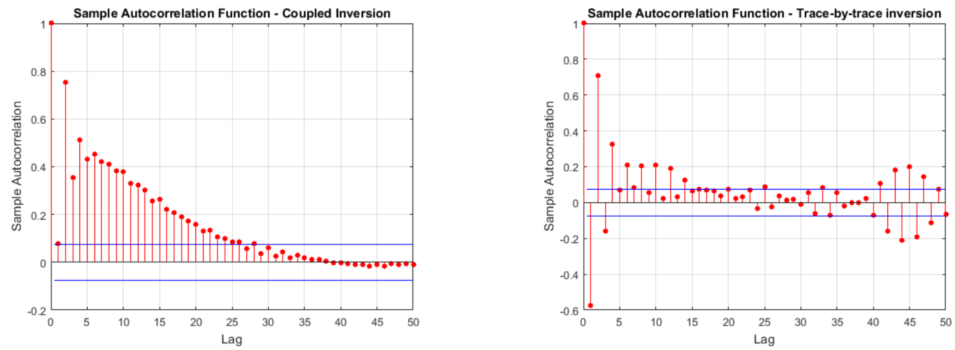
Figure 6. Impedance result from the coupled inversion for the real dataset

To better demonstrate the horizontal correlation imposed on the results, a high pass filter with cutoff frequency of $8Hz$ was applied vertically to the results and the autocorrelation function was calculated for a horizontal line in both cases. The filtering was applied to remove the low frequency provided by the low frequency model, leaving only the higher frequencies provided by the seismic data, containing noise and signal, to be examined by the autocorrelation function. Figure 7 shows both sample autocorrelation functions.

As expected, the coupled inversion has a greater correlation distance due to higher correlation at the near lags, reducing slowly until lag 25 while the trace-by-trace inversion has lower correlations and goes to near zero at lag number 15.

5. Discussion

The proposed methodology adds spatial coupling to the impedance results of the seismic inversion via a simple manipulation of the inversion operator. This manipulation was possible because of the linear nature in which the problem was casted and its solution. The assumptions are the same as in [Buland et al. 2003], which is a stationary prior model for all traces in the 3D model. The boundary effects cited by the same authors are now treated, as the sliding operator proposed in this paper explores the linearity and sparsity of the solution.



(a) Autocorrelation function for the coupled inversion

(b) Autocorrelation function for the trace-by-trace inversion

Figure 7. Autocorrelation of a horizontal 1D line at $t = 150$

The main reason why the proposal works is that, even when coupling all cell nodes at once, the correlations imposed comes from a linear operator on the seismic data. Hence, the result is mathematically independent from the neighbor impedance traces, in other words, the resulting spatial coupling is the result of a linear filter applied to the seismic data.

Another positive effect observed in this proposal is the coupling of samples at greater distances than the window size w, which is evidentiated by having correlations up until lag 25 with a window size of 5, shown at Figure 7.

6. Conclusion

In this paper we presented a simple technique for spatially coupled acoustic seismic inversion in the time domain, which has linear computation time, i.e. $O(n)$ on the number of cells to be inverted. The two test cases showed the addition of spatial coupling without any boundary effects. The proposed technique has a time overhead compared to the trace-by-trace inversion that is a constant which depends on the window size. The window size is determined based on the correlation distance L, in a similar way in which is defined the search radius of kriging algorithms [Caers 2011].

The proposal can be further optimized if the vertical sparsity of the solution is explored. In the cases studied, the vertical number of cells was not enough to justify the truncation of the operator in t. Considering that the vertical dependency of samples are larger vertically due to the effect of the wavelet.

7. Acknowledgement

The authors would like to thank CAPES - Brazilian Federal Agency for Support and Evaluation of Graduate Education - scholarship process no. 4041/15-1, PETROBRAS and CERENA/Instituto Superior Técnico for supporting this work. We also thank CGG for providing an academic license for the Jason software package, used in this work for data preparation and other manipulations.

References

Buland, A., Kolbjørnsen, O., and Omre, H. (2003). Rapid spatially coupled avo inversion in the fourier domain. *Geophysics*, 68(3):824–836.

Buland, A. and Omre, H. (2003). Bayesian linearized avo inversion. *Geophysics*, 68(1):185–198.

Caers, J. (2011). *Modeling Uncertainty in the Earth Sciences*. Wiley.

Figueiredo, L. P., Santos, M., Roisenberg, M., Neto, G., and Figueiredo, W. (2014). Bayesian framework to wavelet estimation and linearized acoustic inversion. *Geoscience and Remote Sensing Letters, IEEE*, 11(12):2130–2134.

Grana, D. and Della Rossa, E. (2010). Probabilistic petrophysical-properties estimation integrating statistical rock physics with seismic inversion. *Geophysics*, 75(3):O21–O37.

Stolt, R. H. and Weglein, A. B. (1985). Migration and inversion of seismic data. *Geophysics*, 50(12):2458–2472.

Tarantola, A. (2005). *Inverse Problem Theory and Methods for Model Parameter Estimation*. Society for Industrial and Applied Mathematics.