



Uncertainty quantification in linear inverse problems with dimension reduction

Leandro Passos de Figueiredo, LTrace and UFSC; Dario Grana, UW; Leonardo Azevedo, IST; Mauro Roisenberg, UFSC; Bruno B. Rodrigues PETROBRAS S/A – CENPES/PDGP/CMR

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, August 19-22, 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

Dimension reduction is a process used to simplify high dimensional inverse problems in geophysics. In this work, we analyze the impact of the model and data reduction on the uncertainty quantification of the inversion results. In the Bayesian approach for inverse problems, the solution is given by the posterior distribution and the uncertainty is interpreted as the posterior variance. Under the Gaussian assumption for the noise and prior distribution and if the forward model and reduction operator are linear, we analytically obtain the posterior distribution. We use the analytical result to compute the solution of the acoustic inversion problem with different reduction levels of both data and model parameters. The method allows us to analyze the impact of the linear reduction on the uncertainty quantification of linear inverse problems.

Introduction

Several geophysical applications can be faced as an inverse problem, such as seismic inversion (Sen, 2006), tomography (Nolet, 1987) and inversion of CSEM data (Gunning et al., 2010). In all applications, we aim to estimate the subsurface properties (model parameters) based on geophysical measurements (observed data) and theoretical relations (forward model) between the observed and the unobserved parameters (Tarantola, 2005).

The common techniques to solve an inverse problem include stochastic optimization methods, where the best subsurface model is chased by minimizing the error/misfit function (Sen, 2006; Azevedo et al., 2015; Bordignon et al., 2017); and Bayesian sampling algorithms, where the solution of the inverse problem is the posterior distribution of the subsurface properties (Buland and Omre, 2003; Tarantola, 2005; Bosch et al., 2010). In the Bayesian approach, the uncertainty is interpreted as the posterior variance, which is a inherent part of the solution (Gelman et al., 2004; Tarantola, 2005; Doyen, 2007). Although the Bayesian algorithms might be inefficient for general inverse problems, it has been successfully applied to geophysical problems during the last decades (Gunning and Glinsky, 2004; Bosch et al., 2010; Grana and Della Rossa, 2010; Grana et al., 2017; de Figueiredo et al., 2018b). In particular cases where the forward model is linear

and under the Gaussian assumption for the prior and noises, the posterior distribution is analytically obtained, leading to very fast algorithms (Buland and Omre, 2003; de Figueiredo et al., 2014).

Dimension reduction, or compression, is a mathematical process used to reduce the number of variables of a problem (Roweis and Saul, 2000; Tompkins et al., 2011). In inverse problem, the reduction can be applied to both model parameters or observed data. In methods based on ensemble Kalman filter, the reduction of the data can improve the efficiency of algorithm. On the other hand, the efficiency of stochastic optimization methods depends on the dimension of the model parameters (Marzouk and Najm, 2009; Lieberman et al., 2010).

With the objective of analyzing the impact of the data/model reduction on the uncertainty of the inversion results, we discuss a general linear inverse problem assuming a linear operator for dimension reduction. With these assumptions, the posterior can be analytically treated by applying Multivariate Statistics results (Anderson, 1984). In the application section, we discuss the methodology applied to the acoustic seismic inversion, where the forward model can be linearized.

Methodology

Bayesian linear inversion

Assuming a general linear inverse problem, the objective is estimating the model parameters \mathbf{m} based on the linear forward operator \mathbf{G} and the observed data \mathbf{d} :

$$\mathbf{d} = \mathbf{G}\mathbf{m} + \mathbf{e}_d \quad (1)$$

where \mathbf{e}_d represents the seismic noise and modeling errors.

In the Bayesian approach for inverse problems, the solution is given by the posterior distribution

$$p(\mathbf{m}|\mathbf{d}) \propto p(\mathbf{d}|\mathbf{m})p(\mathbf{m}), \quad (2)$$

where $p(\mathbf{d}|\mathbf{m})$ is the likelihood distribution (relating the model parameters to the observed data) and $p(\mathbf{m})$ is the prior distribution (relating the model parameters to prior information).

If the noise \mathbf{e}_d of Equation 1 is normally distributed with zero mean, we obtain the following Gaussian multivariate likelihood distribution:

$$p(\mathbf{d}|\mathbf{m}) = \frac{1}{(2\pi)^{\frac{n_d n_d}{2}} |\boldsymbol{\Sigma}_d|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\mathbf{d} - \mathbf{G}\mathbf{m})^T \boldsymbol{\Sigma}_d^{-1} (\mathbf{d} - \mathbf{G}\mathbf{m})\right) \equiv N(\mathbf{d}; \mathbf{G}\mathbf{m}, \boldsymbol{\Sigma}_d). \quad (3)$$

where $\boldsymbol{\Sigma}_d$ is the covariance matrix of the seismic noise (Buland and Omre, 2003).

Assuming a prior knowledge about the model parameters, such as a background model, we can also propose a Gaussian prior distribution:

$$p(\mathbf{m}) = N(\mathbf{m}; \boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m), \quad (4)$$

where $\boldsymbol{\mu}_m$ is the background model and $\boldsymbol{\Sigma}_m$ is the prior covariance matrix that can include spatial correlations (Buland and Omre, 2003; de Figueiredo et al., 2018a).

Based on the statistical model presented in Equations 3 and 4 and applying standard results of multivariate statistical analysis (Anderson, 1984), we can analytically calculate the posterior distribution (Buland and Omre, 2003) of the Bayesian linear inversion (BLI):

$$p(\mathbf{m}|\mathbf{d}) = N(\boldsymbol{\mu}_{m|d}, \boldsymbol{\Sigma}_{m|d}), \quad (5)$$

where the expectation value $\boldsymbol{\mu}_{m|d}$ and the covariance matrix $\boldsymbol{\Sigma}_{m|d}$ are:

$$\boldsymbol{\mu}_{m|d} = \boldsymbol{\mu}_m + \boldsymbol{\Sigma}_m \mathbf{G}^T (\mathbf{G} \boldsymbol{\Sigma}_m \mathbf{G}^T + \boldsymbol{\Sigma}_d)^{-1} (\mathbf{d} - \mathbf{G} \boldsymbol{\mu}_m) \quad (6)$$

are

$$\boldsymbol{\Sigma}_{m|d} = \boldsymbol{\Sigma}_m - \boldsymbol{\Sigma}_m \mathbf{G}^T (\mathbf{G} \boldsymbol{\Sigma}_m \mathbf{G}^T + \boldsymbol{\Sigma}_d)^{-1} \mathbf{G} \boldsymbol{\Sigma}_m. \quad (7)$$

Linear dimension reduction

Principal component analysis (PCA) is a mathematical method that allows obtaining a set of uncorrelated variables from observed correlated variables, which maintains the full observed variance (Pearson, 1901). PCA can be applied to the covariance matrix to reduce the dimension of the problem by selecting only the first components that capture most of the variance.

For example, from a set of observations of the $(n \times 1)$ -vector \mathbf{x} with a mean $\boldsymbol{\mu}$ and a $(n \times n)$ -covariance matrix $\boldsymbol{\Sigma}$, we can apply the Eigen Decomposition theorem to $\boldsymbol{\Sigma}$:

$$\boldsymbol{\Sigma} = \mathbf{V} \boldsymbol{\Lambda} \mathbf{V}^T, \quad (8)$$

where \mathbf{V} is the matrix of orthogonal eigenvectors (principal components) and $\boldsymbol{\Lambda}$ is the diagonal matrix with the respective eigenvalues (variances).

By arranging the eigenvalues in decreasing order, we can choose a subset of \tilde{n} eigen vectors (with $\tilde{n} \leq n$) that represents a fraction f_v of the total variance:

$$\frac{\sum_{i=1}^{\tilde{n}} \Lambda_{i,i}}{\sum_{j=1}^n \Lambda_{j,j}} \geq f_v, \quad (9)$$

Then, any observation \mathbf{x} can be represented a by a linear combination of the \tilde{n} eigenvectors:

$$\mathbf{x} = \boldsymbol{\mu} + \tilde{\mathbf{V}} \tilde{\mathbf{x}}, \quad (10)$$

where $\tilde{\mathbf{V}}$ is a $(n \times \tilde{n})$ -matrix with the subset of \tilde{n} principal eigenvectors (Tompkins et al., 2011), and $\tilde{\mathbf{x}}$ is a $(\tilde{n} \times 1)$ -vector that represents the sample \mathbf{x} in a reduced linear space.

To find the coefficients $\tilde{\mathbf{x}}$ that represents a given sample \mathbf{x} in the reduced space, we can use the generalize inversion of matrices:

$$\tilde{\mathbf{x}} = \tilde{\mathbf{V}}^{-1} (\mathbf{x} - \boldsymbol{\mu}). \quad (11)$$

where $\tilde{\mathbf{V}}^{-1} = \tilde{\mathbf{V}}^T (\tilde{\mathbf{V}} \tilde{\mathbf{V}}^T)^{-1}$ is the generalized inverse of the non-square operator $\tilde{\mathbf{V}}$.

Equations 10 and 11 define the linear operators to transform and back-transform from/to the reduced space. The scheme in Figure 1 illustrates the discussed framework for linear dimension reduction.

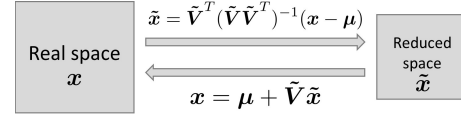


Figure 1: Framework for linear dimension reduction based on PCA.

Dimension reduction in inverse problems

- Data compression

Using the presented framework for data reduction, we obtain

$$\tilde{\mathbf{d}} = \tilde{\mathbf{V}}_d^{-1} (\mathbf{d} - \boldsymbol{\mu}_d), \quad (12)$$

where $\tilde{\mathbf{V}}_d$ is the matrix of the principal eigenvectors for a given fraction of total variance f_v , which is defined according to the Equations 8 and 9.

By applying Equation 12 to Equation 1 of a linear inverse problem, we obtain

$$\tilde{\mathbf{d}} - \tilde{\mathbf{V}}_d^{-1} \boldsymbol{\mu}_d = \mathbf{H} \mathbf{m} + \mathbf{e}_{\tilde{d}} \quad (13)$$

where $\mathbf{H} = \tilde{\mathbf{V}}_d^{-1} \mathbf{G}$ and $\mathbf{e}_{\tilde{d}} = \tilde{\mathbf{V}}_d^{-1} \mathbf{e}_d$ is the reduced noise vector with covariance $\boldsymbol{\Sigma}_{\tilde{d}} = \tilde{\mathbf{V}}_d \boldsymbol{\Sigma}_d \tilde{\mathbf{V}}_d^T$.

Because all the operators are linear, we can derive the multivariate Gaussian posterior distribution by using the same results of multivariate statistical analysis that are cited in the section Bayesian linear inversion (Equations 5-7):

$$p(\mathbf{m}|\tilde{\mathbf{d}}) = N(\mathbf{m}; \boldsymbol{\mu}_{m|\tilde{d}}, \boldsymbol{\Sigma}_{m|\tilde{d}}), \quad (14)$$

where the expectation value $\boldsymbol{\mu}_{m|\tilde{d}}$ and the covariance matrix $\boldsymbol{\Sigma}_{m|\tilde{d}}$ are:

$$\boldsymbol{\mu}_{m|\tilde{d}} = \boldsymbol{\mu}_m + \boldsymbol{\Sigma}_m \mathbf{H}^T (\mathbf{H} \boldsymbol{\Sigma}_m \mathbf{H}^T + \boldsymbol{\Sigma}_{\tilde{d}})^{-1} (\tilde{\mathbf{d}} - \tilde{\mathbf{V}}_d^{-1} \boldsymbol{\mu}_d - \mathbf{H} \boldsymbol{\mu}_m) \quad (15)$$

and

$$\boldsymbol{\Sigma}_{m|\tilde{d}} = \boldsymbol{\Sigma}_m - \boldsymbol{\Sigma}_m \mathbf{H}^T (\mathbf{H} \boldsymbol{\Sigma}_m \mathbf{H}^T + \boldsymbol{\Sigma}_{\tilde{d}})^{-1} \mathbf{H} \boldsymbol{\Sigma}_m. \quad (16)$$

Equation 14 is the solution of the inverse problem for the model parameter \mathbf{m} given a dimension reduction of the observed data \mathbf{d} .

- Model compression

Applying Equation 10 to the model parameter, we obtain

$$\mathbf{m} = \boldsymbol{\mu}_m + \tilde{\mathbf{V}}_m \tilde{\mathbf{m}}, \quad (17)$$

where $\tilde{\mathbf{V}}_m$ is the matrix of the principal eigenvectors for a given fraction of total variance f_v , which is defined according to the Equations 8 and 9.

From Equations 1 and 17, we can write the forward model as

$$\mathbf{d} - \mathbf{G}\boldsymbol{\mu}_m = \mathbf{F}\tilde{\mathbf{m}} + \mathbf{e}_d, \quad (18)$$

where $\mathbf{F} = \mathbf{G}\tilde{\mathbf{V}}_m$, $\tilde{\mathbf{m}}$ is the reduced model parameter with Gaussian prior distribution

$$p(\tilde{\mathbf{m}}) = N(\tilde{\mathbf{m}}; \mathbf{0}, \boldsymbol{\Sigma}_{\tilde{\mathbf{m}}}), \quad (19)$$

and the covariance $\boldsymbol{\Sigma}_{\tilde{\mathbf{m}}}$ is a diagonal matrix with the eigenvalues of the prior covariance $\boldsymbol{\Sigma}_m$.

Given Equations 18 and 19, we can derive the posterior distribution of the reduced model parameter $\tilde{\mathbf{m}}$ by using the standard results of multivariate statistical analysis (Equations 5-7).

$$p(\tilde{\mathbf{m}}|\mathbf{d}) = N(\tilde{\mathbf{m}}; \boldsymbol{\mu}_{\tilde{\mathbf{m}}|\mathbf{d}}, \boldsymbol{\Sigma}_{\tilde{\mathbf{m}}|\mathbf{d}}), \quad (20)$$

where the expectation value $\boldsymbol{\mu}_{\tilde{\mathbf{m}}|\mathbf{d}}$ and the covariance matrix $\boldsymbol{\Sigma}_{\tilde{\mathbf{m}}|\mathbf{d}}$ are

$$\boldsymbol{\mu}_{\tilde{\mathbf{m}}|\mathbf{d}} = \boldsymbol{\Sigma}_{\tilde{\mathbf{m}}}\mathbf{F}^T (\mathbf{H}\boldsymbol{\Sigma}_{\tilde{\mathbf{m}}}\mathbf{F}^T + \boldsymbol{\Sigma}_d)^{-1} (\mathbf{d} - \mathbf{G}\boldsymbol{\mu}_m) \quad (21)$$

and

$$\boldsymbol{\Sigma}_{\tilde{\mathbf{m}}|\mathbf{d}} = \boldsymbol{\Sigma}_{\tilde{\mathbf{m}}} - \boldsymbol{\Sigma}_{\tilde{\mathbf{m}}}\mathbf{F}^T (\mathbf{H}\boldsymbol{\Sigma}_{\tilde{\mathbf{m}}}\mathbf{F}^T + \boldsymbol{\Sigma}_d)^{-1} \mathbf{F}\boldsymbol{\Sigma}_{\tilde{\mathbf{m}}}. \quad (22)$$

From the posterior distribution of the reduced model and the Equation 17, we obtain the posterior distribution of the model parameter given the linear transformation between the variables, then

$$p(\mathbf{m}|\mathbf{d}, \tilde{\mathbf{m}}) = N(\mathbf{m}; \boldsymbol{\mu}_{\mathbf{m}|\mathbf{d}, \tilde{\mathbf{m}}}, \boldsymbol{\Sigma}_{\mathbf{m}|\mathbf{d}, \tilde{\mathbf{m}}}), \quad (23)$$

where

$$\boldsymbol{\mu}_{\mathbf{m}|\mathbf{d}, \tilde{\mathbf{m}}} = \tilde{\mathbf{V}}_m \boldsymbol{\mu}_{\tilde{\mathbf{m}}|\mathbf{d}} + \boldsymbol{\mu}_m \quad (24)$$

and

$$\boldsymbol{\Sigma}_{\mathbf{m}|\mathbf{d}, \tilde{\mathbf{m}}} = \tilde{\mathbf{V}}_m \boldsymbol{\Sigma}_m \tilde{\mathbf{V}}_m^T. \quad (25)$$

Application

To analyze the theoretical results of the methodology section, we apply the linear Bayesian inversion and linear dimension reduction to the acoustic seismic inversion problem. In this case, the observed data \mathbf{d} is the full stack seismic data that can be modeled by the convolutional model based on the acoustic impedance values (model parameter \mathbf{m}) of the subsurface (Buland and Omre, 2003; de Figueiredo et al., 2014, 2018a). The convolutional forward model can be linearized assuming reflectivities lower than 0.3. Assuming that the seismic noise and the prior distribution of the model parameters are both Gaussian, we can apply the methodology presented in this paper.

To validate the inversion results, we apply the methods to a synthetic case of acoustic impedance with a signal to noise of 10. The reference model is shown in black curves in the Figure 2, where the reference seismic trace is on the left and the log P-impedance on the right. In the same figure, we show the traditional BLI (without dimension reduction) in blue, and we observe a good match between the reference model and the estimated one.

Figures 3 and 4 show the results of the BLI with dimension reduction with fractions of the total variance of 0.9 and 0.75, respectively. In both results we can observe a smoothing

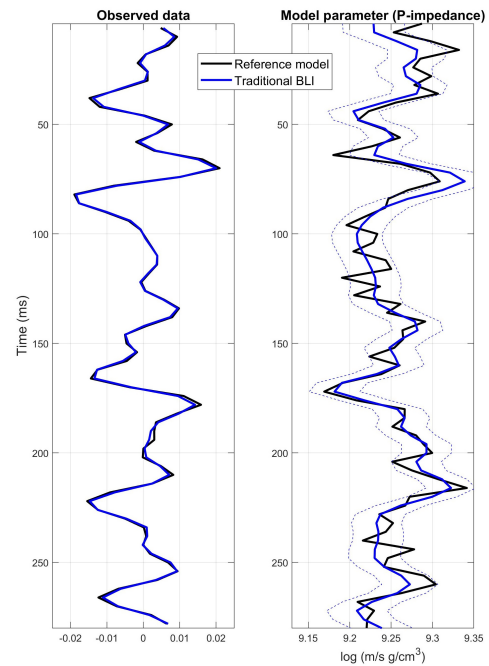


Figure 2: Synthetic case of acoustic inversion, seismic data (left plot) and log P-impedance (right plot). The reference data is in black and the inversion results of the traditional BLI is in blue with the confidence interval of 68% (dashed lines).

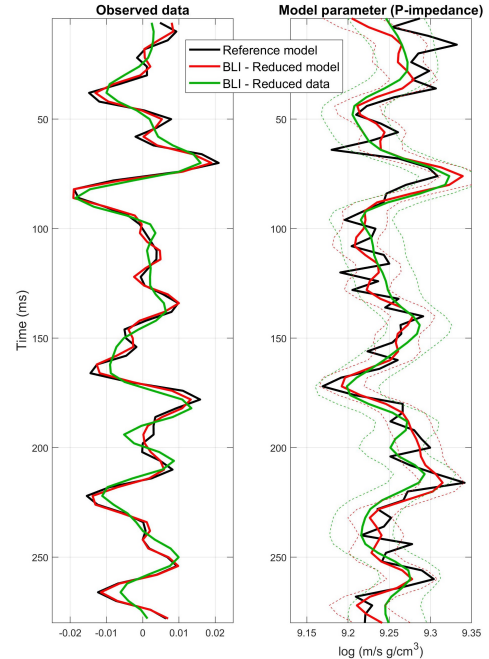


Figure 3: BLI results with dimension reduction with a fraction 0.9 of the total variance. Model reduction results in red and data reduction in green. Dashed lines represent the confidence interval of 68%.

effect in the impedance models and synthetic seismic data. The uncertainty of the estimates is represented by the dashed lines in the plots, which show the confidence interval of 68%. In both figures, we can visually conclude that the posterior variance is lower when reducing the model.

In fact, by calculating Equation 16 and 22 for several values of fraction of total variance used for dimension reduction, we obtain the result of Figure 5. Posterior variances of inverse problem solutions with data and model reduction converge to the traditional BLI. However, when reducing the observed data, the posterior variance is higher than the traditional BLI, whereas it is lower when reducing the model parameter. This effect can be explained by the fact that the inverse problem becomes well posed when we reduce the model parameters, and ill posed when we reduce the observed data.

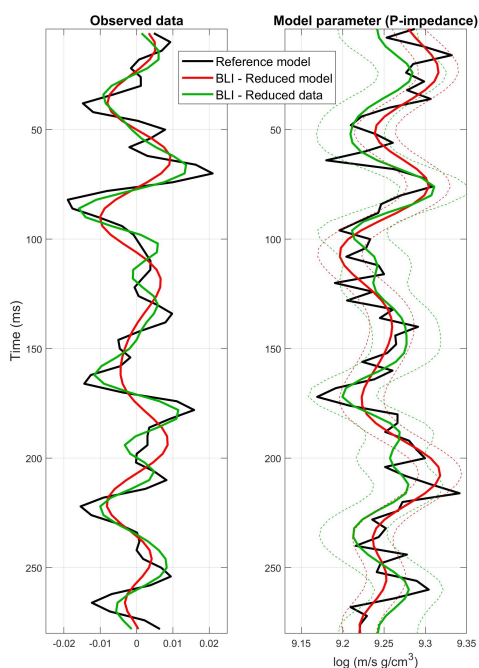


Figure 4: BLI results with dimension reduction with a fraction 0.75 of the total variance. Model reduction results in red and data reduction in green, Dashed lines represent the confidence interval of 68%.

Conclusion

Assuming that the forward model and the dimension reduction operator are both linear, and under the Gaussian assumption for the observed noise and prior distribution of the model parameters, we analytically compute the solution of the linear inversion problem. The preliminary application to the acoustic inversion shows that the effects on uncertainty quantification are different for data and model reduction. The results indicate that the uncertainty is underestimated when the reduction is applied to the model parameters, whereas the reduction of the observed data causes higher variances due to the lack of observations.

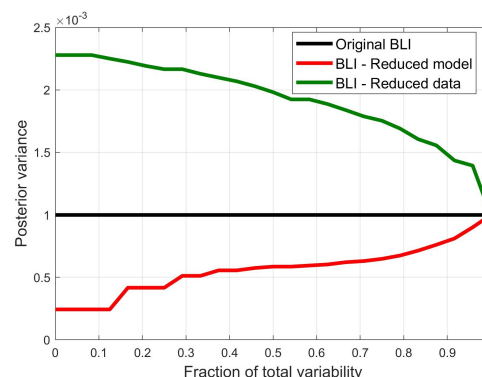


Figure 5: Posterior variance versus the fraction of the total variance used for reduction. Model reduction results in red and data reduction in green.

References

- Anderson, T., 1984, An introduction to multivariate statistical analysis: Wiley. Wiley series in probability and mathematical statistics: Probability and mathematical statistics.
- Azevedo, L., R. Nunes, A. Soares, E. C. Múndin, and G. S. Neto, 2015, Integration of well data into geostatistical seismic amplitude variation with angle inversion for facies estimation: *Geophysics*, **80**, M113–M128.
- Bordignon, F. L., L. P. de Figueiredo, L. Azevedo, A. Soares, M. Roisenberg, and G. S. Neto, 2017, Hybrid global stochastic and bayesian linearized acoustic seismic inversion methodology: *IEEE Transactions on Geoscience and Remote Sensing*, **55**, 4457–4464.
- Bosch, M., T. Mukerji, and E. Gonzalez, 2010, Seismic inversion for reservoir properties combining statistical rock physics and geostatistics: A review: *Geophysics*, **75**, 75A165–75A176.
- Buland, A., and H. Omre, 2003, Bayesian linearized AVO inversion: *Geophysics*, **68**, 185–198.
- de Figueiredo, L. P., D. Grana, F. L. Bordignon, M. Santos, M. Roisenberg, and B. B. Rodrigues, 2018a, Joint bayesian inversion based on rock-physics prior modeling for the estimation of spatially correlated reservoir properties: *Geophysics*, **83**, M49–M61.
- de Figueiredo, L. P., D. Grana, M. Roisenberg, and B. B. Rodrigues, 2018b, Gaussian mixture mcmc method for linear seismic inversion. Accepted with minor revisions.
- de Figueiredo, L. P., M. Santos, M. Roisenberg, G. Neto, and W. Figueiredo, 2014, Bayesian framework to wavelet estimation and linearized acoustic inversion: *Geoscience and Remote Sensing Letters, IEEE*, **11**, 2130–2134.
- Doyen, P., 2007, Seismic reservoir characterization: an earth modelling perspective: EAGE publications. Education tour series.
- Gelman, A., J. Carlin, H. Stern, and D. Rubin, 2004, Bayesian data analysis: Chapman & Hall/CRC.
- Grana, D., and E. Della Rossa, 2010, Probabilistic petrophysical-properties estimation integrating statistical rock physics with seismic inversion: *Geophysics*, **75**, O21–O37.
- Grana, D., T. Fjeldstad, and H. Omre, 2017, Bayesian

- gaussian mixture linear inversion for geophysical inverse problems: *Mathematical Geosciences*, **49**, 493–515.
- Gunning, J., and M. E. Glinsky, 2004, Delivery: an open-source model-based Bayesian seismic inversion program: *Computers & Geosciences*, **30**, 619–636.
- Gunning, J., M. E. Glinsky, and J. Hedditch, 2010, Resolution and uncertainty in 1d csem inversion: A bayesian approach and open-source implementation: *GEOPHYSICS*, **75**, F151–F171.
- Lieberman, C., K. Willcox, and O. Ghattas, 2010, Parameter and state model reduction for large-scale statistical inverse problems: *SIAM Journal on Scientific Computing*, **32**, 2523–2542.
- Marzouk, Y. M., and H. N. Najm, 2009, Dimensionality reduction and polynomial chaos acceleration of bayesian inference in inverse problems: *Journal of Computational Physics*, **228**, 1862 – 1902.
- Nolet, G., 1987, *in* Seismic wave propagation and seismic tomography: Springer Netherlands, 1–23.
- Pearson, F. R. S. K., 1901, Liii. on lines and planes of closest fit to systems of points in space: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, **2**, 559–572.
- Roweis, S. T., and L. K. Saul, 2000, Nonlinear dimensionality reduction by locally linear embedding: *Science*, **290**, 2323–2326.
- Sen, M. K., 2006, *Seismic Inversion*: Society of Petroleum Engineers.
- Tarantola, A., 2005, *Inverse problem theory and methods for model parameter estimation*: Society for Industrial and Applied Mathematics.
- Tompkins, M., J. Fernández Martínez, D. Alumbaugh, and T. Mukerji, 2011, Scalable uncertainty estimation for nonlinear inverse problems using parameter reduction, constraint mapping, and geometric sampling: *Marine controlled-source electromagnetic examples: Geophysics*, **76**, F263–F281.

Acknowledgments

The authors acknowledge Petrobras, School of Energy Resources of the University of Wyoming and Instituto Superior Tecnico for the support and availability during the work.