Bayesian algorithm to jointly estimate wavelet, seismic noise level and correlation

Leandro Passos de Figueiredo, UFSC, Mauro Roisenberg, UFSC, Marcio Santos UFSC, Guenther Schwedersky Neto, PETROBRAS S/A – CENPES/PDGP/CMR

This paper was prepared for presentation at the 13th International Congress of the Brazilian Geophysical Society, held in Rio de Janeiro, Brazil, August 26-29, 2013.

Abstract

This paper describes how a Bayesian framework can be modeled and applied on seismic data to estimate the wavelet. The method works on post-stack and pre-stack data, in both, the convolutional forward model is considered, but it differ in how the reflectivity is calculated from the well log. We expanded the method to estimate the seismic noise correlation range jointly with the wavelet, the seismic noise level and uncertainties. The method is applied in synthetic and real post-stack seismic data. The Gaussian assumption for the likelihood models enables to obtain the analytical expressions for the conditioned distributions, which allows sampling from the posterior distribution via Gibbs Sampling Algorithm.

Introduction

Inversion of seismic data plays a vital processing step in reservoir modeling and characterization. It helps to improve exploration and management success, once that estimates the elastic properties from the seismic data, which has a great correlation with many petrophysical properties.

The Bayesian formulation for the inverse problem has been demonstrated an efficient and robust technique to estimate uncertainties and obtain multiples realizations of properties, as can be seen in Bosch et al. (2010), Rimstad et al. (2012), Buland and Omre (2003*b*) and Buland and Omre (2003*a*).

In this work, we adapted the Bayesian wavelet estimation method proposed by Buland and Omre (2003*a*), which basically estimates the wavelet, its uncertainties and the seismic noise level. We expanded the method to jointly estimate the seismic noise correlation range that is related with its frequencies, which are important knowledge about the seismic. We also present explicitly how some variables of the stochastic model were defined in a geophysical interpretation, and discuss how the conditional distributions of the Gibbs algorithm is obtained.

The information about the noise correlation can help the seismic inversion algorithms, once it is associated with the covariance matrix of the misfit function, which has a great importance in the optimization process.

The Gibbs algorithm is a Markov Chain Monte Carlo (MCMC) method similar to Metropolis, which has been applied in many multidimensional problems, obtaining a random walk in the parameter space, which approximately, sample the desired posterior distribution (Geman and Geman, 1984) (Gelman et al., 2004).

Methodology

In the stochastic model, the distributions are considered multivariate normal distributions denoted by $N_n(\boldsymbol{\mu},\boldsymbol{\Sigma})$, where $\boldsymbol{\mu}$ is the mean vector, $\boldsymbol{\Sigma}$ is the covariance matrix and n is the dimension of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. Its explicit form is expressed by the equation below.

$$p(\boldsymbol{x}) = \frac{1}{(2\pi)^{\frac{n}{2}} \mid \boldsymbol{\Sigma} \mid^{\frac{1}{2}}} \mathbf{exp}\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right) \quad (1)$$

where $\boldsymbol{x}$ is a random field that satisfies the multivariate distribution (Anderson, 1984).

Seismic Model

The forward seismic model is considered the convolutional model in a discrete setting (Sen, 2006), as shown in equation 2,

$$\boldsymbol{d}_o = \boldsymbol{R}\boldsymbol{s} + \boldsymbol{e}_d \quad (2)$$

where $\boldsymbol{d}_o$ is the seismic data, $\boldsymbol{s}$ is the wavelet, $\boldsymbol{e}_d$ is a noise, $\boldsymbol{R}$ is the convolutional matrix formed by the reflectivity $\boldsymbol{r}$ that depends of impedance z by the relation below.

$$\boldsymbol{r} = \frac{1}{2}\frac{\partial}{\partial t} ln(z(t)) \quad (3)$$

Stochastic Model

Assuming that seismic noise $\boldsymbol{e}_d$ is a Gaussian noise, the Gaussian likelihood model for $\boldsymbol{d}_o$ with expectation $\boldsymbol{\mu}_d = \boldsymbol{R}\boldsymbol{s}$ and covariance matrix $\boldsymbol{\Sigma}_d$ is proposed on equation 4.

$$p(\boldsymbol{d}_o|\boldsymbol{\mu}_d,\boldsymbol{\Sigma}_d) = N_{n_d}(\boldsymbol{\mu}_d,\boldsymbol{\Sigma}_d) \quad (4)$$

The wavelet $\boldsymbol{s}$ is also modeled by a Gaussian likelihood (equation 5), with wavelet expectation $\boldsymbol{\mu}_s$ defined by a simple null vector with n_s components.

$$p(\boldsymbol{s}|\boldsymbol{\mu}_s,\boldsymbol{\Sigma}_s) = N_{n_s}(\boldsymbol{\mu}_s,\boldsymbol{\Sigma}_s) \quad (5)$$

Based on the convolutional model and these likelihoods models, a stochastic model is proposed and shown in

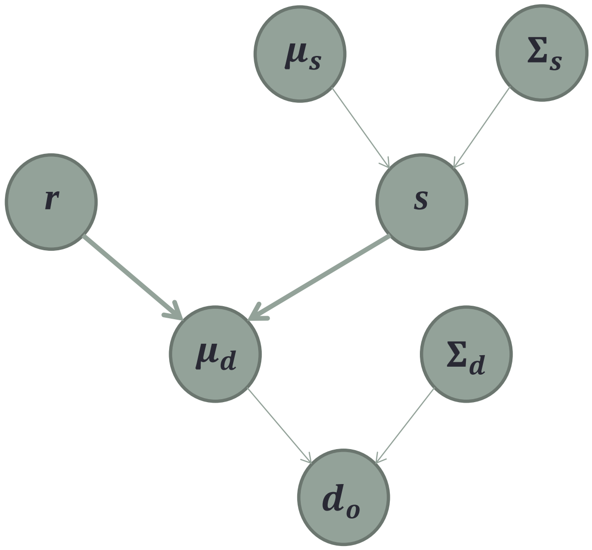

Figure 1: Directed acyclic graph representing the stochastic model.

the directed acyclic graph (DAG) on figure 1 (Buland and Omre, 2003*b*).

Considering that the wavelet expectation $\boldsymbol{\mu}_s$ and the reflectivity $\boldsymbol{r}$ are completely known, its associated prior distributions are delta distributions. The covariance matrices $\boldsymbol{\Sigma}_s$ and $\boldsymbol{\Sigma}_d$ are assumed to be known up to an unknown multiplicative variance factor σ_d^2 and σ_s^2 (equation 6), which its uncertainties are modeled by a constant prior distribution, because no significant difference was observed when using other distribution.

$$\forall\, \boldsymbol{v} \in \{\boldsymbol{s},\boldsymbol{d}\}: \quad \boldsymbol{\Sigma}_v = \sigma_v^2 \boldsymbol{\Sigma}_{0v} \tag{6}$$

The covariance matrices structures $\boldsymbol{\Sigma}_{0v}$ are defined using the prior knowledge about the variables of interest.

The wavelet covariance matrix $\boldsymbol{\Sigma}_{0s}$ is defined by equation 7, which impose smoothness to the wavelet by the second-order exponential correlation function on equation 8, and that its components approaching zero at the ends defining the variance of each component by equation 9.

$$\boldsymbol{\Sigma}_{0s_{t,t'}} = \delta_t \delta_{t'} \nu_{t,t'}, \tag{7}$$

$$\nu_{t,t'} = exp\left(-\frac{(t-t')^2}{L_s^2}\right) \tag{8}$$

$$\delta_t = exp\left(-\frac{1}{0.02 n_s^2}\left(t-(n_s+1)/2\right)^2\right), \tag{9}$$

The range parameter L_s is defined observing the range of the seismic vertical variogram, which is assumed to be approximately equal to wavelet variogram range.

The covariance matrix for the seismic data $\boldsymbol{\Sigma}_{0d}$ is defined by the sum of an exponential second-order correlation function with range L_d and a first-order correlation function with range 32 ms (equation 10). The first-order correlation function is added only to avoid the matrix singularity associated with the second-order function.

$$\boldsymbol{\Sigma}_{0d_{t,t'}} = exp\left(-\frac{(t-t')^2}{L_d^2}\right) + 10^{-3} exp\left(-\frac{|t-t'|}{32}\right), \tag{10}$$

The range parameter L_d is an unknown variable which has to be estimated in the process, this parameter is related with the seismic noise correlation and its frequencies, which are important knowledge about the seismic.

Following the Bayesian rules, the DAG on the figure 1 leads to the posterior distribution:

$$p(\boldsymbol{s},\sigma_s^2,\sigma_d^2,L_d,|\boldsymbol{d}_o,\boldsymbol{r},\boldsymbol{\mu}_s) \propto \tag{11}$$
$$p(\boldsymbol{d}_o|\boldsymbol{s},\boldsymbol{r},\sigma_d^2,L_d)\,p(\boldsymbol{s}|\boldsymbol{\mu}_s,\sigma_s^2)$$

which is impossible to obtain in an analytical way, for this reason, a MCMC algorithm is necessary. More specifically, the Gibbs algorithm is used to perform the samples, which basically consists in, for each iteration, draw the unknown variables given all the others, and then, calculate the mean and the uncertainties of the variables of interest.

Algorithm

To develop the Gibbs algorithm, the conditioned distributions are necessary and need to be calculate from the posterior. Basically the distributions are obtained using the equation 11, considering the given variables constants jointly with the seismic model and the theorem of conditional distribution of multivariate Gaussian distributions, and the distribution of a non-singular linear transformation theorem, presented on Anderson (1984).

The conditional distribution for $\boldsymbol{s}$ is

$$p(\boldsymbol{s}|\boldsymbol{\mu}_s,\boldsymbol{d}_o,\boldsymbol{m},\boldsymbol{\Sigma}_s,\boldsymbol{\Sigma}_d) = N(\boldsymbol{\mu}_{s|},\boldsymbol{\Sigma}_{s|}), \tag{12}$$

where the mean and the covariance matrix are

$$\boldsymbol{\mu}_{s|} = \boldsymbol{\mu}_s + \boldsymbol{\Sigma}_s\boldsymbol{R}^T\left(\boldsymbol{R}\boldsymbol{\Sigma}_s\boldsymbol{R}^T+\boldsymbol{\Sigma}_d\right)^{-1}\left(\boldsymbol{d}_o - \boldsymbol{R}\boldsymbol{\mu}_s\right), \tag{13}$$

$$\boldsymbol{\Sigma}_{s|} = \boldsymbol{\Sigma}_s - \boldsymbol{\Sigma}_s\boldsymbol{R}^T\left(\boldsymbol{R}\boldsymbol{\Sigma}_s\boldsymbol{R}^T+\boldsymbol{\Sigma}_d\right)^{-1}\boldsymbol{R}\boldsymbol{\Sigma}_s, \tag{14}$$

The conditional distribution for the multiplicative variance factor σ_v^2 for all $\boldsymbol{v} \in \{\boldsymbol{s},\boldsymbol{d}\}$ is an inverse gamma distribution,

$$\sigma_v^2 \sim IG(\gamma_v,\lambda_v), \tag{15}$$

with shape parameter γ_v and scale parameter λ_v given by:

$$\gamma_v = \frac{n_v}{2}, \lambda_v = \frac{(\boldsymbol{v}-\boldsymbol{\mu}_v)^T \boldsymbol{\Sigma}_{0v}^{-1}(\boldsymbol{v}-\boldsymbol{\mu}_v)}{2}. \tag{16}$$

Finally, the conditional distribution for the range parameter L_d can not be derived analytically, however it is computationally obtained calculating the posterior probability on equation 11, for different values of the parameter L_d, considering that all the other variables as fixed parameters. In general the shape of calculated distribution is clearly normal, so after the calculations, the conditioned distribution is approximated by a normal gaussian univariate distribution, enabling the conditioned samples of L_d.

For each iteration, the algorithm consists on draw the unknown variables from these distributions, as it is demonstrated in Algorithm 1, summarizing the Gibbs sampling.

Algorithm 1 Gibbs sampling algorithm for stochastic wavelet estimation.

Define initial values for σ_d^2, σ_s^2 and L_d
for i=1, .. , k + n **do**
  Draw $\boldsymbol{s}(i)$ of $N(\boldsymbol{\mu}_{s|},\boldsymbol{\Sigma}_{s|})$
  Draw $\sigma_v^2(i)$ of $IG(\gamma_v,\lambda_v)\ \forall\ \boldsymbol{v} \in \{\boldsymbol{s},\boldsymbol{d}\}$
  $\boldsymbol{\Sigma}_v = \sigma_v^2(i)\boldsymbol{\Sigma}_{0v}\ \forall\ v \in \{\boldsymbol{s},\boldsymbol{d}\}$
  Compute $p(L_d|\boldsymbol{d}_o,\boldsymbol{s},\boldsymbol{r},\sigma_d^2) \approx N(\mu_L,\sigma_L^2)$
  Draw L_d of $N(\mu_L,\sigma_L)$
  Uptade $\boldsymbol{\Sigma}_d$
end for

Examples

To evaluate the algorithm efficiency, the method described is applied in real and synthetic post-stack seismic data.

The synthetic data is obtained by the convolution of the reflectivity calculated from the real acoustic impedance well data with a known wavelet. The aim of this application is to evaluate the estimation quality when using different kinds and levels of additional seismic noise comparing the estimated wavelet with the original wavelet, and the estimated noise parameter with the parameters used to generate the additional noise.

The noises are generated by sampling from the normal multivariate distribution $N_n(\boldsymbol{0},\boldsymbol{\Sigma}_d)$, given a correlation range L_d and noise variance factor σ_d^2 on equation 10.

The figure 2 shows the mentioned synthetic data, but with only four examples of noises (figure 2D) with same level but different correlation range.

The application on real data consists on consider the same well data to calculate the reflectivity, but the seismic is considered the real seismic measurement obtained on the well location.

On Gibbs algorithm, the results are calculated after 50 iterations to ensure the convergence to the posterior distribution in equation 11.

Firstly, three synthetic seismic were generated by the same wavelet and same noise correlation range, but with different noise levels with variance $\sigma_d^2 = 2.17\ 10^3$, $8.69\ 10^3$ and $34.8\ 10^3$, which corresponds to a signal to noise ratio (S/N) equals to 20, 10 and 5 respectively.

Figure 3 shows the wavelet mean calculated with the uncertainty for each point for the three different noise levels, where can be observed that the uncertainty is directly related with the noise level, and the wavelet mean has a good convergence to the original wavelet.

The histogram of the seismic noise variance σ_d^2 sampled during the same application (figure 4), which illustrates the posterior distribution of σ_d^2, shows that the distribution converges to the value that generates the seismic noise, as can be seen on table 1, which shows the mean of distribution compared with the value of the noise variance σ_d^2 used to generate the noise.

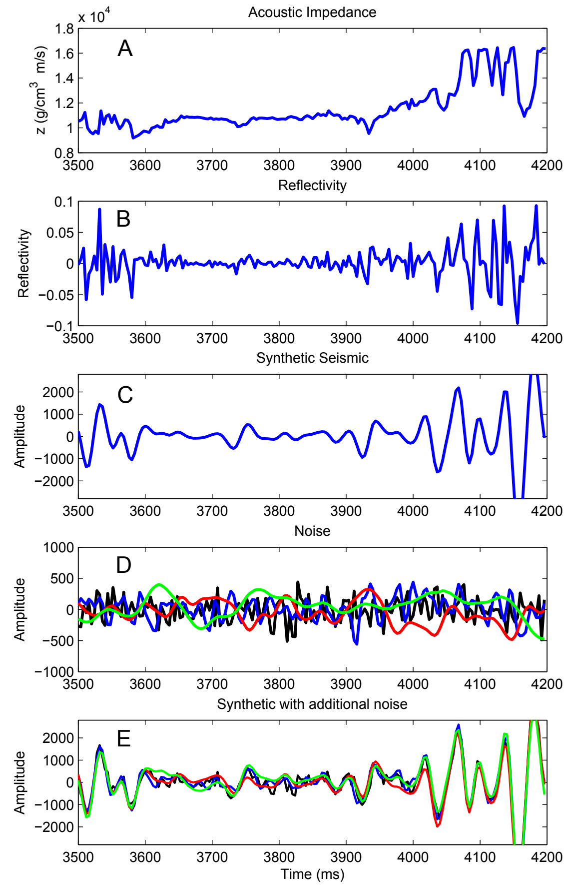

Figure 2: Synthetic data: acoustic impedance from well log (A); reflectivity (B); synthetic seismic (C); seismic noises with S/N = 5 sampled from $N_n(\boldsymbol{0},\boldsymbol{\Sigma}_d)$ with different correlation ranges L_d (D), white noise from $L_d \approx 0$ in black (noise 1), noise from $L_d = 8ms$ in blue (noise 2), noise from $L_d = 20ms$ in red (noise 3), noise from $L_d = 32ms$ in green (noise 4); and seismic with noise in respective color (E).

S/R	$\langle\sigma_d\rangle$ (10^3)	Noise variance (10^3)	Error(%)
20	2.21	2.17	1.7
10	9.11	8.69	4.8
5	35.7	34.8	2.8

Table 1: Seismic noise level estimation.

The second test with synthetic data is with three synthetic seismic generated by the same wavelet, but in this time, the noise levels are the same and the correlation range varies to $L_d = 8ms$, $20ms$ and $32ms$ (noises on figure 2D).

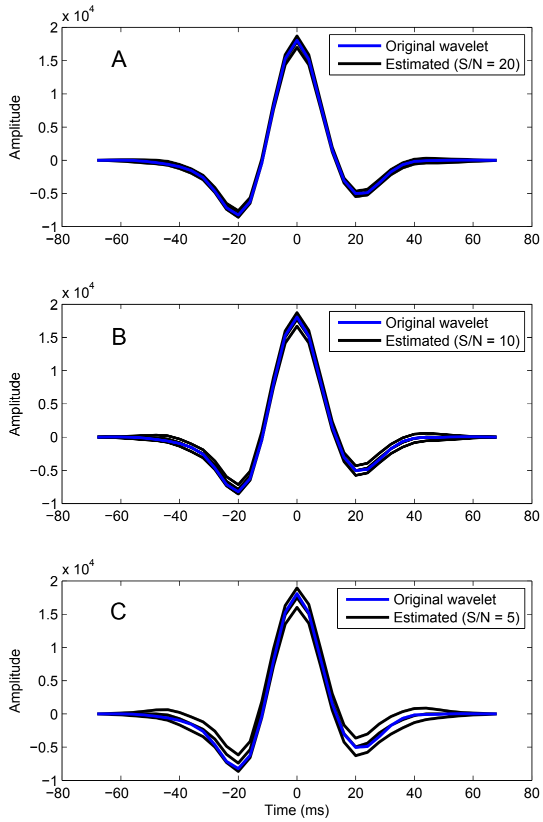

Figure 3: Original wavelet (blue) and estimated wavelet with uncertainties (black) for three different noise levels, S/N = 20 in A, S/N = 10 in B and S/N = 5 in C

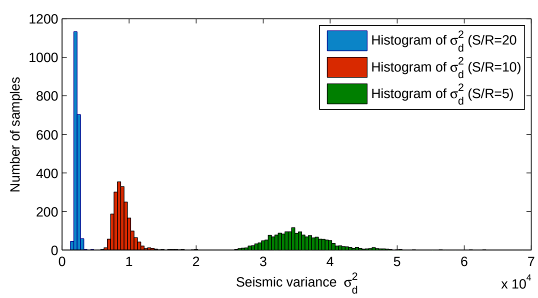

Figure 4: Histogram of the seismic noise variance σ_d^2 sampled during the algorithm.

The wavelet mean converges to the original wavelet. And as can be seen on figure 5, the posterior distribution for L_d also converges to the value that generates the seismic noises of figure 2.

The last application of the algorithm is with the same real acoustic impedance, but with the real seismic data on the location well (figure 6).

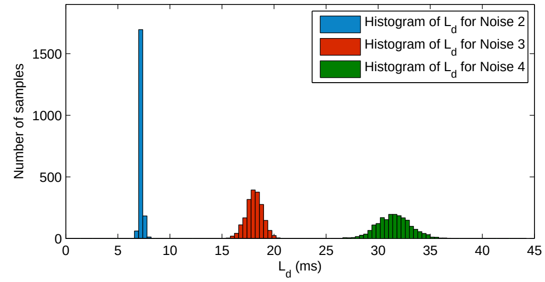

Figure 5: Histogram of the seismic noise correlation range L_d sampled during the algorithm for different noises.

The estimated wavelet mean calculated with the uncertainty for each point is shown in figure 7, which can be observed a reasonable uncertainty. The wavelet generated in the process, models a seismic that has a good fit with the experimental seismic data as can be seem on figure 6.

The histogram of the seismic noise variance σ_d^2 sampled during the execution (figure 8) showed that the posterior distribution converges to a higher value than those used in the tests with synthetic data. And the histogram of L_d (figure 9) showed that the posterior distribution have a good convergence to a mean value (10 ms).

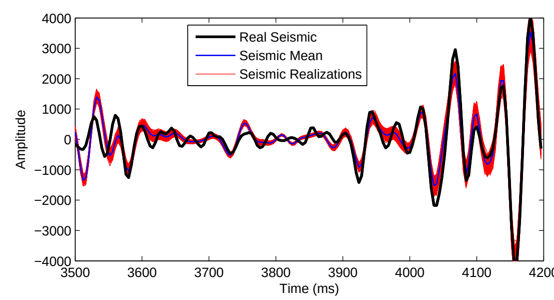

Figure 6: Real post-stack seismic data in black, the synthetic seismics generated by the wavelet samples in red and its mean in blue.

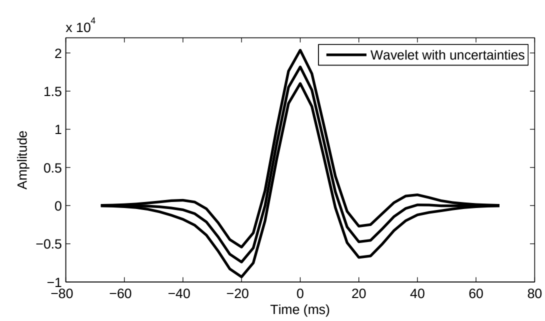

Figure 7: Estimated wavelet with uncertainties for the real seismic data.

The method was also applied in pre-stack seismic data, in which the reflectivity depends of density, p-velocity and s-

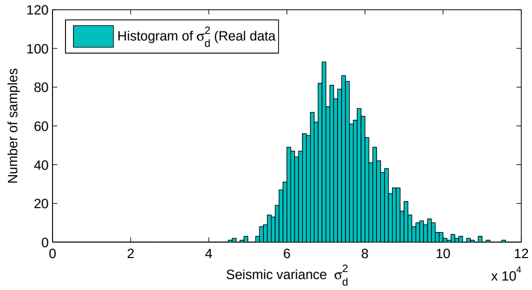

Figure 8: Histogram of the seismic noise variance σ_d^2 sampled during the algorithm (real data).

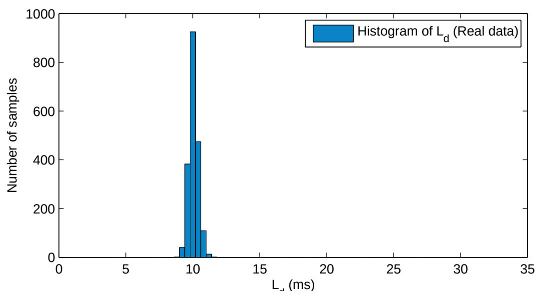

Figure 9: Histogram of the seismic noise correlation range L_d sampled during the algorithm (real data).

velocity from well data (Aki and Richards, 2002), and the algorithm also showed results that has a good fit with the real data.

Conclusion

In all the applications, the method presents good results without any assumption about the wavelet phase. In synthetic seismic data, not only the estimated wavelet has a good convergence to the original wavelet, but also the noise level and correlation range converge to the original values. In real seismic data, the synthetic seismic obtained during the execution has a good fit with the experimental seismic. All this results indicates that the method is viable and reliable. The information about the noise correlation can help the seismic inversion algorithms, once that is associated with the covariance matrix of the misfit function, which has a great importance in the optimization process.

References

Aki, K. and Richards, P. (2002), *Quantitative Seismology:*, Geology (University Science Books).: Seismology, Univ Science Books.

Anderson, T. (1984), *An introduction to multivariate statistical analysis*, Wiley series in probability and mathematical statistics: Probability and mathematical statistics, Wiley.

Bosch, M., Mukerji, T. and Gonzalez, E. (2010), 'Seismic inversion for reservoir properties combining statistical rock physics and geostatistics: A review', *Geophysics* **75**(5), 75A165.

Buland, A. and Omre, H. (2003*a*), 'Bayesian wavelet estimation from seismic and well data', *Geophysics* **68**(6), 2000–2009.

Buland, A. and Omre, H. (2003*b*), 'Joint avo inversion, wavelet estimation and noise-level estimation using a spatially coupled hierarchical bayesian model', *Geophysical Prospecting* **51**(6), 531–550.

Gelman, A., Carlin, J., Stern, H. and Rubin, D. (2004), *Bayesian Data Analysis*, Chapman & Hall/CRC.

Geman, S. and Geman, D. (1984), 'Stochastic relaxation, gibbs distributions, and the bayesian restoration of images', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6**(6), 721–741.

Rimstad, K., Avseth, P. and Omre, H. (2012), 'Hierarchical bayesian lithology/fluid prediction: A north sea case study', *Geophysics* **77**(2), B69–B85.

Sen, M. K. (2006), *Seismic Inversion*, Society of Petroleum Engineers, Richardson, TX, USA.

Acknowledgments

The authors wish to express their appreciation to Petrobras for the support and availability during the work.