Bayesian inversion of non-sonic well log data for the determination of P-wave velocity: application to data from Jequitinhonha Basin, Brazil

Odette Caroline Aquino Aragão* (IGEO/UFBA), Amin Bassrei[1] (CPGG/IGEO/UFBA & INCT-GP), Geraldo Girão Nery (Hydrolog)

Summary

A new stochastic approach is presented to invert non-sonic well logs (density, neutronic porosity and lithology) in order to obtain P-wave velocity. The probabilistic analysis combines rock physics relationships between the model and data parameters. The prior, theoretical and posterior joint probability density functions (PDF) for each depth, where the well logs were recorded are estimated. Finally, from the posterior marginal model PDF, the posterior model is determined. This approach was applied to well log data from three boreholes in the Jequitinhonha Basin, state of Bahia, Brazil. The results presented very good agreement to the actual sonic logs from the same well.

Introduction

Borehole geophysics plays a key role in the exploration of hydrocarbons and minerals (Morozov and Ma, 2009). In situ information is acquired by various kinds of probes with different vertical resolutions and different depths of investigation from the vicinity of the borehole wells. The collected data can be related to various physical rock properties located at the top of the Earth's crust (Ellis and Singer, 2007). One of the main geophysical well logs is the sonic one. It measures the travel time of an ultrasonic wave propagating through the rock layers (Bassiouni, 1994). Despite the fact that it is an indispensable tool in geophysical surveys, many wells do not have this data, due to its high cost. Thus, it is necessary to use other data for the estimation of the sonic log.

The Bayesian inversion, represented by Bayes' theorem (Bayes, 1763), provides a robust arrangement that takes into account the data uncertainties and associates, in an inverse problem formulated stochastically, the model and the data from physical and mathematical relationships (Tarantola, 1987). In this structure, the solution of the inverse problem is represented by the posterior probability density function (PDF) over the model parameters (Tarantola, 2005). The incorporation of prior information can significantly speed up the convergence of the inversion process toward the most likely solution in model space (Scales and Tenorio, 2001).

In this work we will use Bayesian inversion in order to estimate the values of the P-wave velocities in the vicinity of three wells from Jequitinhonha Basin, state of Bahia, Brazil. The PDFs will be built using the data from the density, neutron porosity and lithology profiles. The actual sonic logs were also available, but not used in the inversion procedure. They were used for the evaluation of the proposed method.

Theory and Method

A probability density function (PDF) $f(\boldsymbol{x})$ satisfies the following conditions (Ulrych et al., 2001):

$$f(\boldsymbol{x}) \geq 1, \forall\, \boldsymbol{x};\ \int_{-\infty}^{+\infty} f(\boldsymbol{x})\, d\boldsymbol{x} = 1. \qquad (1)$$

Considering a Cartesian product between two spaces, $\chi = \mathcal{M} \times \mathcal{D}$, given a joint PDF $f(\boldsymbol{m}, \boldsymbol{d})$, the marginal PDF related to the model and data are given, respectively, by (Charlin and Louis, 1996):

$$f_{\boldsymbol{m}}(\boldsymbol{m}) = \int_{\mathcal{D}} f(\boldsymbol{m}, \boldsymbol{d})\, d\boldsymbol{d}, \qquad (2)$$

and

$$f_{\boldsymbol{d}}(\boldsymbol{d}) = \int_{\mathcal{M}} f(\boldsymbol{m}, \boldsymbol{d})\, d\boldsymbol{m}. \qquad (3)$$

According to Bayes' theorem, for independent variables, a joint PDF can be expressed by (Mosegaard and Tarantola, 2002):

$$f(\boldsymbol{m}, \boldsymbol{d}) = f_{\boldsymbol{m}}(\boldsymbol{m}) f_{\boldsymbol{d}}(\boldsymbol{d}). \qquad (4)$$

In this work, we assume that all PDF conform to a Gaussian probability distribution. A Gaussian function $f(\boldsymbol{x})$ is given by:

$$f(\boldsymbol{x}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\frac{(\boldsymbol{x} - \bar{\boldsymbol{x}})^2}{\sigma^2}\right), \qquad (5)$$

where $\bar{\boldsymbol{x}}$ and σ, respectively, the mean and the standard deviation of $\boldsymbol{x}$.

Model and Data Parameters. The model parameter $\boldsymbol{m}$ contained in the $\mathcal{M}$ space is the P- wave velocity ($\boldsymbol{v_p}$), and the data parameters $\boldsymbol{d}$contained in the $\mathcal{D}$ space is represented by the density ($\boldsymbol{RHOB}$) and neutronic porosity ($\boldsymbol{NPHI}$) available in the well log file.

Prior Joint Probability Density Function. The prior joint PDF $\rho(\boldsymbol{m}, \boldsymbol{d})$ is determined using the Bayes' rule as follows:

[1] Currently visiting professor at Stanford University

$$\rho(\boldsymbol{m},\boldsymbol{d}) = \rho_{\boldsymbol{m}}(\boldsymbol{v_p})\rho_d(\boldsymbol{RHOB},\boldsymbol{NPHI}). \quad (6)$$

The posterior probabilities are strongly influenced by the prior probabilities. Therefore any error caused by misreading the prior probability will be amplified in the posterior probability. The marginal distribution $\rho_{\boldsymbol{m}}(\boldsymbol{m})$ was established from the data presented in the lithology profile. Table 1 shows the ranges of $\boldsymbol{v_p}$for all lithologies present in both wells, according to data from Bourbie et al. (1987) and also the code that identify each lithology.

The values of $\boldsymbol{RHOB}, \boldsymbol{NPHI}$ and $\boldsymbol{v_p}$ were discretized in 30 intervals, from the minimum to the maximum values from the well logs. Then, for each depth of investigation, 30 different values of $\boldsymbol{v_p}$ were stochastically combined with 30 different values of $\boldsymbol{RHOB}$ and $\boldsymbol{NPHI}$.

Lithology	Code	Range of v_p (*km/s*)
Calcilutite	6	2.5 – 5.0
Calcarenite	8	2.5 – 5.0
Dolomite	30	3.5 – 8.0
Conglomerate	42	2.0 – 4.0
Sandstone	49	2.0 – 3.5
Clay	57	1.0 – 2.5
Marl	58	2.0 – 3.0
Shale	73	2.0 – 5.0
Anhydrite	82	4.0 – 7.0

Table 1: Ranges of v_p values and codes for all lithologies present in the three well logs.

For each depth of investigation the averages of $\boldsymbol{RHOB}$ and $\boldsymbol{NPHI}$ correspond to their actual values related to that depth. Also, the average of $\boldsymbol{v_p}$is the value estimated from the lithology log. Finally, the standard deviation of each parameter is related to the standard deviation of the caliper log, because this log provides the quality control over the entire data (Nery, 2013), considering its average value as the well drill size.

Theoretical Joint Probability Density Function. In order to calculate the theoretical joint PDF, the $\boldsymbol{NPHI}$ vector of the prior PDF was used. On the other hand, $\boldsymbol{v_p}$ and $\boldsymbol{RHOB}$ are built by relating them to $\boldsymbol{NPHI}$ by the following relations:

$$\frac{1}{\boldsymbol{v_p}} = \frac{1-\boldsymbol{NPHI}}{v_{pm}} + \frac{\boldsymbol{NPHI}}{v_{pf}}, \quad (7)$$

$$\boldsymbol{RHOB} = (1-\boldsymbol{NPHI})\rho_m + \boldsymbol{NPHI}\rho_f. \quad (8)$$

The first relation is known as Wyllie's equation (Wyllie et al, 1956). v_{pm} and ρ_f are, respectively, the rock matrix P-wave velocity and density. It was considered a matrix composed of calcite with $v_{pm} = 6.4$ km/s and $\rho_m = 2.71$ g/cm³. v_{pf} and ρ_f are, respectively, the fluid (in the pore space) P-wave velocity and density. The considered fluid was fresh water with $v_{pf} = 1.5$ km/s and $\rho_f = 1.0$ g/cm³.Then, the theoretical joint PFD $\vartheta(\boldsymbol{m},\boldsymbol{d})$ is expressed as

$$\vartheta(\boldsymbol{m},\boldsymbol{d}) = \vartheta_{\boldsymbol{m}}(\boldsymbol{v_p})\vartheta_d(\boldsymbol{RHOB}). \quad (9)$$

Because $\vartheta(\boldsymbol{m},\boldsymbol{d})$ is also represented by a Gaussian function, it requires a standard deviation, which was established as half of the value of the prior joint PDF standard deviation. Thus, the posterior joint PDF will depend more on the theoretical model than on the prior model.

Posterior Joint Probability Density Function. Finally, the posterior joint PDF is determined as

$$\sigma(\boldsymbol{m},\boldsymbol{d}) = \frac{1}{v}\frac{\rho(\boldsymbol{m},\boldsymbol{d})\vartheta(\boldsymbol{m},\boldsymbol{d})}{\mu(\boldsymbol{m},\boldsymbol{d})}, \quad (10)$$

where $\mu(\boldsymbol{m},\boldsymbol{d})$ is the homogeneous probability density distribution and

$$v = \int_{\mathcal{D}\times\mathcal{M}} \frac{\rho(\boldsymbol{m},\boldsymbol{d})\vartheta(\boldsymbol{m},\boldsymbol{d})}{\mu(\boldsymbol{m},\boldsymbol{d})}\, d\boldsymbol{m}\, d\boldsymbol{d}. \quad (11)$$

The posterior marginal PDF related to the model is expressed by

$$\sigma_{\boldsymbol{m}} = \int_{\mathcal{D}} \sigma(\boldsymbol{m},\boldsymbol{d})\, d\boldsymbol{d}. \quad (12)$$

For each depth, the value related to the maximum probability in $\sigma_{\boldsymbol{m}}$will consist in the posterior model $\boldsymbol{v_{pinv}}$, that is, the inversion result.

Relative Error and Root Mean Square Deviation. The relative error (*RE*) and root mean square deviation (*RMSE*) between the values of $\boldsymbol{v_p}$ presented in the log file ($\boldsymbol{v_{pactual}}$) and $\boldsymbol{v_{pinv}}$ were calculated, respectively, by:

$$RE = \frac{\|\boldsymbol{v_{pactual}} - \boldsymbol{v_{pinv}}\|_2}{\|\boldsymbol{v_{pactual}}\|_2}, \quad (13)$$

and

$$RMSE = \sqrt{\frac{1}{n_{vp}}\sum_{i=1}^{n_{vp}} (\boldsymbol{v_{pactual}} - \boldsymbol{v_{pinv}})^2}, \quad (14)$$

where n_{vp} is the number of investigation depths in each well.

Discussion of Results

For the first well, 1-BAS-0037, seven lithologies are found. The chosen velocities for each lithology used for the estimation of the prior model are shown in Table 2. As the porosity log in this well indicates medium porosities, the velocity values are close to the mean of the velocity range presented in Table 1.

Figures 1 and 2 exhibit, respectively, $\rho_{\boldsymbol{m}}(\boldsymbol{v_p})$ and $\sigma_{\boldsymbol{m}}(\boldsymbol{v_p})$ for the whole well. It is noticeable that, for any

depth, the most probable values of $\sigma_m(v_p)$ and $\rho_m(v_p)$ are almost in the same range. However, $\sigma_m(v_p)$ always presents the highest probabilities.

Lithology	v_p (*km/s*)
Calcilutite	4.0
Calcarenite	4.0
Dolomite	7.0
Conglomerate	4.0
Sandstone	3.0
Clay	2.5
Shale	4.5
Anhydrite	6.0

Table 2: v_p values used to estimate the prior model for well 1-BAS-0037.

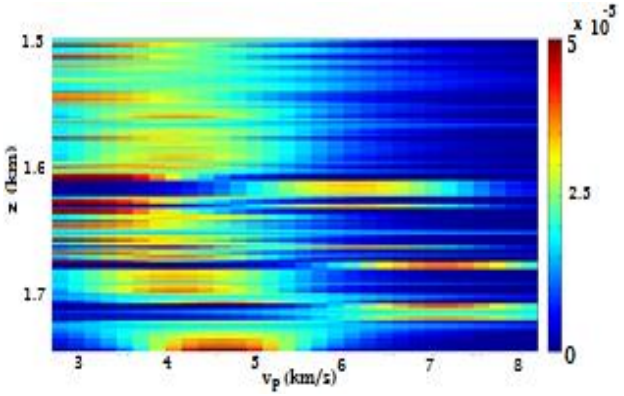

Figure 1: Prior marginal PDF related to the model for well 1-BAS-0037.

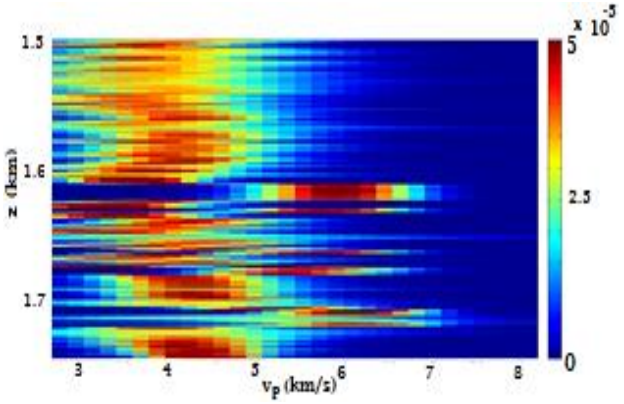

Figure 2: Posterior marginal PDF related to the model for well 1-BAS-0037.

Figure 3 shows the result of the inversion, v_{pinv}, and other logs such as $v_{pactual}$. The larger differences between v_{pinv} and $v_{pactual}$ are associated with the lithologies dolomite, clay and shale, represented respectively by codes 30, 57 and 73. The dolomite presented around 1700m is located in a region where there was collapse, according to the analysis of the caliper profile. For this well, $ER = 7.95\%$ and $RMSE = 0.50$ km/s.

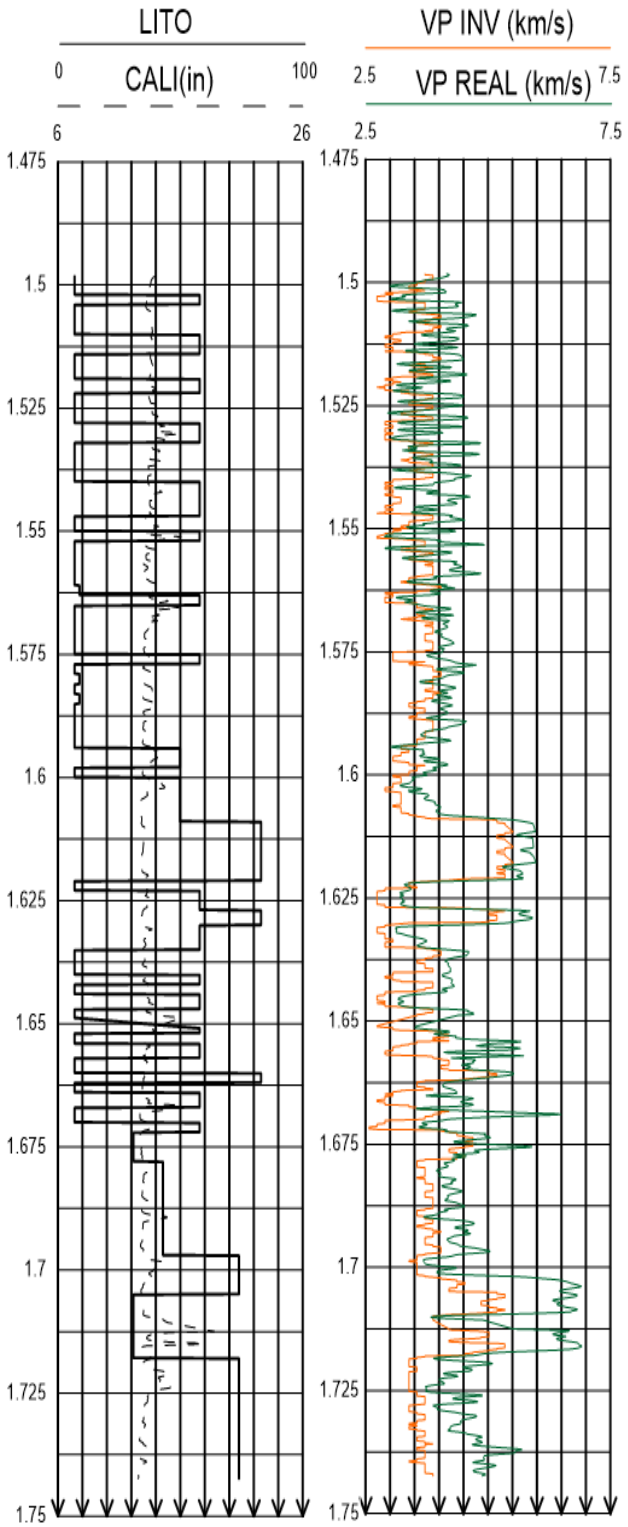

Figure 3: Results for well 1-BAS-0037. On the left, the continuous and dashed black lines represent, respectively, the lithology and the caliper logs. On the right, the orange line represents the actual sonic log and the green line represents inversion result.

For well 1-BAS-0068, according to the lithology log, calcilutite and calcarenite are both present. The velocities chosen for these lithologies for the prior model estimation were the maximum values in Table 1 (5.0 km/s for both lithologies) because the porosity log in this well indicates very low porosities. Figures 4 and 5 represent, respectively, $\rho_m(v_p)$ and $\sigma_m(v_p)$ for the whole well. For any depth, the values of $\sigma_m(v_p)$ are always greater than $\rho_m(v_p)$.

The inversion result, v_{pinv}, showed in Figure 6 is presented with other well logs, including $v_{pactual}$. It is possible to see that the largest errors are associated with landslides, visible from the caliper profile, from the top of the profile up to about 4975 m. For this well, $RE = 2.92\ \%$ and *RMSE* $= 0.23$ km/s.

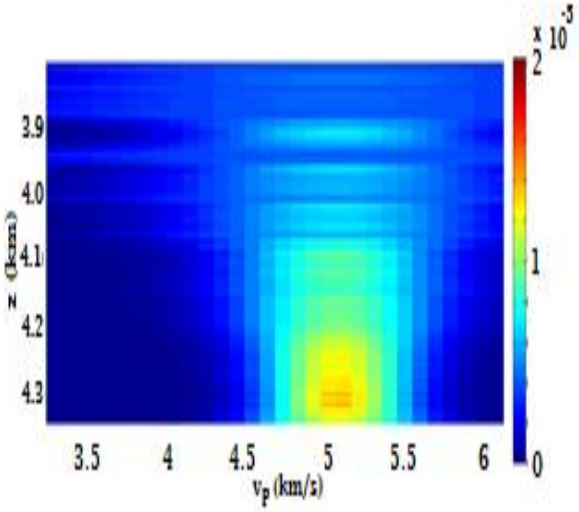

Figure 4: Prior marginal PDF related to the model for well 1-BAS-0068.

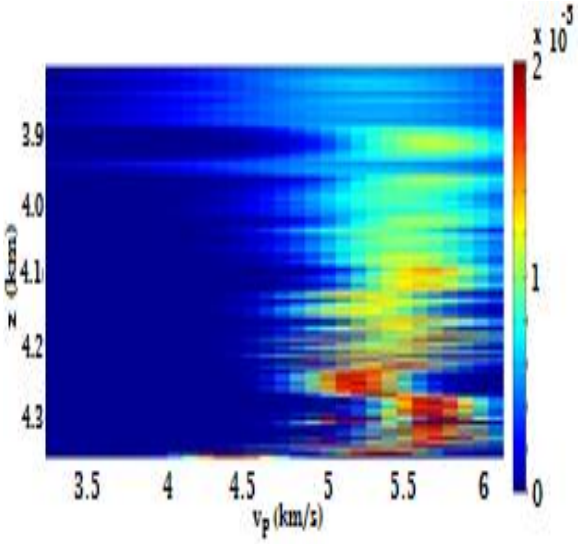

Figure 5: Posterior marginal PDF related to the model for well 1-BAS-0068.

For the second well, 1-BAS-0080, the present lithologies presented calcilutite, shale and marl and the adopted velocities for the prior model estimation for each lithology are, respectively, 2.5 km/s, 1.0 km/s and 2.0 km/s. These values are close to the minimum values in Table 1, because the porosity log in this well shows very high porosities. Figures 7 and 8 respectively represent $\rho_m(v_p)$ and $\sigma_m(v_p)$ for the whole well. The most probable values of $\rho_m(v_p)$ are below 1.5 km/s while for $\sigma_m(v_p)$ the highest probabilities are between 2.0 and 2.5 km/s.

Figure 9 shows the result of the inversion, v_{pinv}, and other logs. It is noteworthy that calcilutite, represented in the *LITHO* log by code 6, always causes velocity peaks. These peaks are also present in the inverted profile. Even considering the fluctuating values of the caliper and despite the presence of lithologies that have clay (shale and marl), the inversion was successful with $RE = 4.61\%$ and $RMSE = 0.15$km/s.

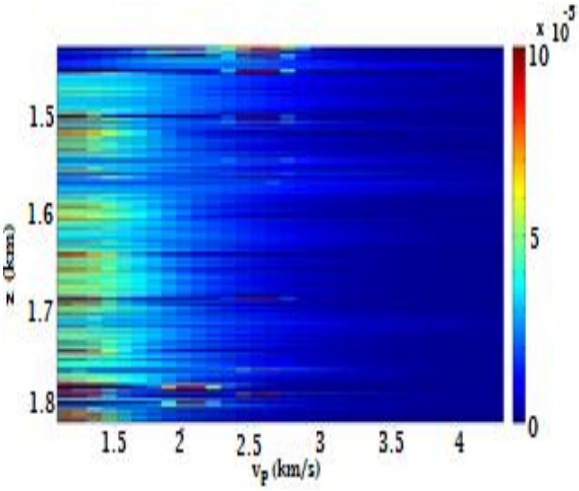

Figure 7: Prior marginal PDF related to the model for well 1-BAS-0080.

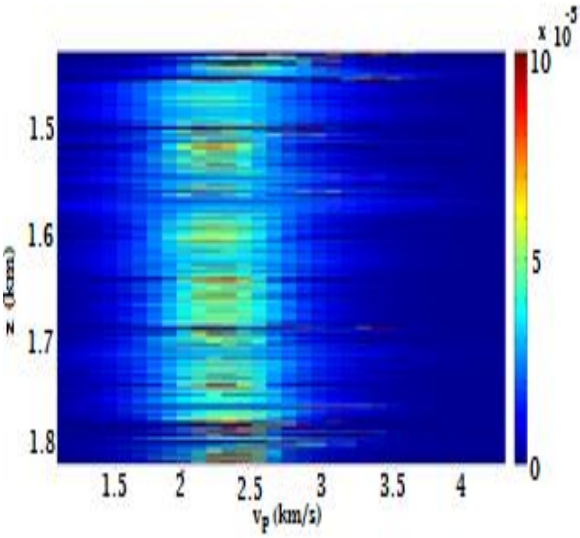

Figure 8: Posterior marginal PDF related to the model for well 1-BAS-0080.

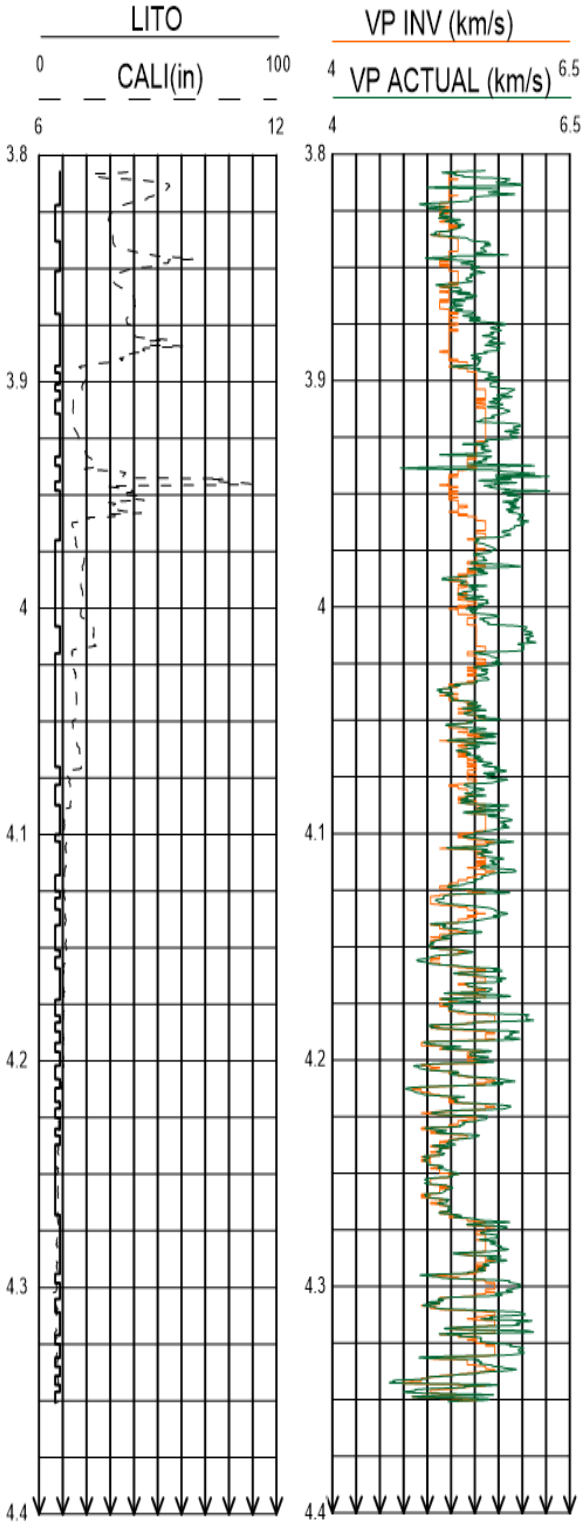

Figure 6: Results for well 1-BAS-0068. On the left, the continuous and dashed black lines represent, respectively, the lithology and the caliper logs. On the right, the orange line represents the actual sonic log and the green line represents inversion result.

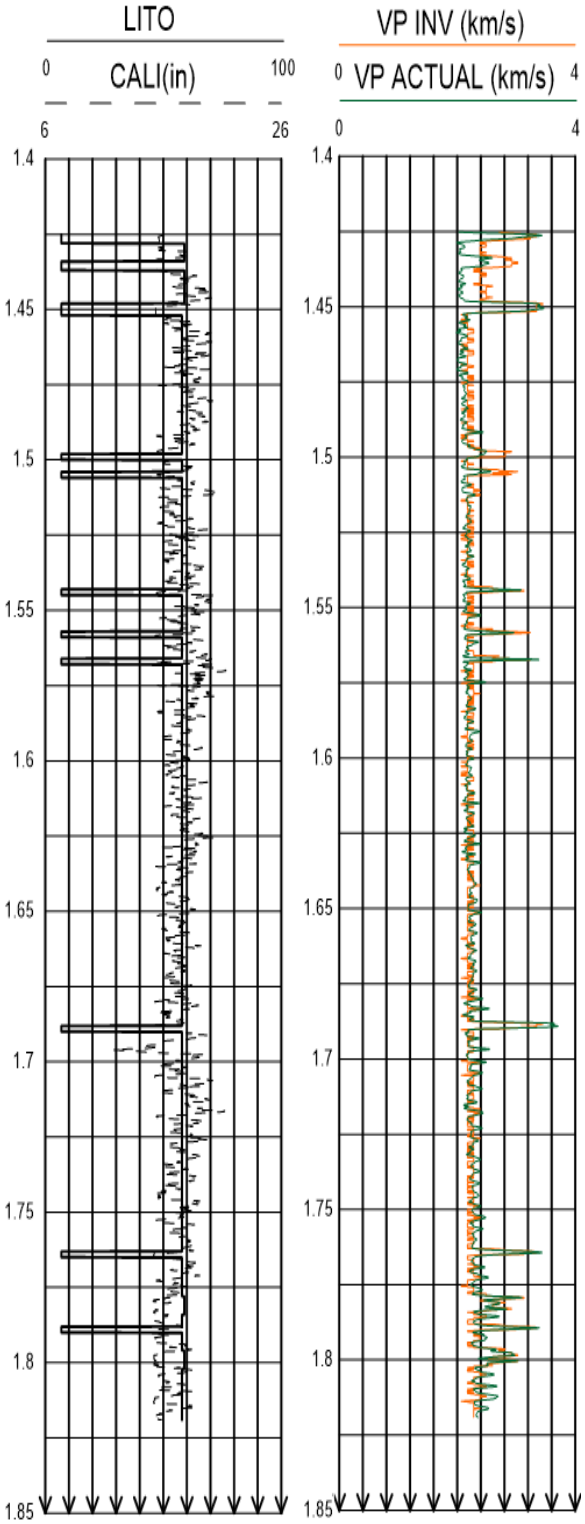

Figure 9: Results for well 1-BAS-0080. On the left, the continuous and dashed black lines represent, respectively, the lithology and the caliper logs. On the right, the orange line represents the actual sonic log and the green line represents inversion result.

Conclusion

Well 1-BAS-0068 presented the best result with the lowest RE, which can be related to the presence of only two lithologies that usually do not interfere in the quality of the all log recordings. Furthermore, calcarenite and calcilutite have a matrix formed in part by calcite, that makes the value of v_{pm} and ρ_m suitable. On the other hand, well 1-BAS-0037, which contains seven different lithologies had the largest relative error (7.95%).

As a rule, for the three wells, the larger differences between $\boldsymbol{v_{pinv}}$ and $\boldsymbol{v_{pactual}}$ are associated with landslides areas, indicated by the caliper profile, and the presence of clay minerals, as in shales and marls. When the well diameter is greater than the drill size, what is an indicative of collapse, the inversion quality greatly decreases. It is necessary to emphasize the importance of determining an appropriate prior model because it plays a key role in the process of inversion. The proposed method produced very good results, even without performing any correction in the well log data, and using the same values of v_{pm}, v_{pf}, ρ_m and ρ_f for the entire three wells.

Acknowledgements

Odette C. A. Aragão thanks CAPES for the M.Sc. scholarship. A. Bassrei thanks both CNPq and Stanford University for sponsoring his one year visiting professorship stay. Both authors thank CNPq and PETROBRAS for sponsoring the project INCT-GP, and FINEP for sponsoring the project Rede 01. Both authors also thank PETROBRAS for the well log data provided to the project Rede 01.

References

Bassiouni, Z., 1994, Theory, Measurement, and Interpretation of Well Logs: SPE Textbook Series.

Bayes, T. and M. Price, 1763, An essay towards solving a problem in the doctrine of chances: Philosophical Transactions, **53**, 370–418.

Carlin, B. P. and T. A. Louis, 1996, Bayes and Empirical Bayes Methods for Data Analysis: Chapman and Hall.

Ellis, D. V. e J. M. Singer, 2007, Well Logging for Earth Scientists: Springer.

Morozov, I., and J. Ma, 2008, Accurate poststack acoustic-impedance inversion by well-log calibration: Geophysics, **74**, 59–67.

Mosegaard, K. and A. Tarantola, 2002, Probabilistic approach to inverse problems, *in* International Handbook of Earthquake and Engineering Seismology, 237–265.

Nery, G. G., 2013, Perfilagem Geofísica em Poço Aberto: Fundamentos Básicos com Ênfase em Petróleo: Sociedade Brasileira de Geofísica.

Scales, J. A. and L. Tenorio, 2001, Prior information and uncertainty in inverse problems: Geophysics, **66**, 389–397.

Tarantola, A., 1987, Inverse Problem Theory: Science Publ. Co.

Tarantola, A., 2005, Inverse Problem Theory and Methods for Model Parameter Estimation: SIAM.

Ulrych, T. J., M. D. Sacchi, and A. Woodbury, 2001, A Bayes tour of inversion: tutorial: Geophysics, **66**, 55–69.

Wyllie, M. R. J., A. R. Gregory and L. W. Gardner, 1956, Elastic wave velocity in heterogeneous and porous media: Geophysics, **21**, 41–70.