Porosity Bayesian inference from multiple well-log data

Luiz Lucchesi Loures

ABSTRACT

This paper reports on an inversion procedure for porosity estimation and uncertainty analysis in well locations from a series of well-logs consisting of neutron, sonic (compression and shear wave) and density logs. The inversion procedure is based on the Bayesian methodology of inference. The inversion formulation considers log data uncertainties and information from rock physics, which include the effects of clay content and pressure. The main goal of this methodology is to reduce the uncertainties associated with each type of well-log, within the estimated porosity model.

The Bayesian formulation developed assumes that all uncertainties can be described by Gaussian pdfs and we consider the variances of the well-log datasets to be unknown parameters. The posterior pdf is marginalized for the variances and the final posterior pdf is one for the interval porosity. The methodology is implemented using a moving window, which computes one posterior of the interval porosity for each interval of the discrete well.

Examples with synthetic data are produced to show how this methodology works. Tests performed with different well-log combinations have provided a way to analyze the contribution of each type of log data towards increased reliability of the estimated porosity. Tests with a real dataset are presented. The results of these tests are compared with a porosity model derived from laboratory experiments on a core sample.

INTRODUCTION

Determination of porosity distribution in a reservoir is usually preceded by porosity determination in well locations. In this procedure the interpreter relied on both well-log data and core measurements. The interval porosity along the well is derived, after suitable corrections, from well-log data, such as a neutron log, sonic logs (compression and shear waves), density log or a combination of those. The main problem with these approaches is that they fail to properly treat the uncertainties associated with each type of log. For example, the neutron porosity log is derived from a nuclear tool, which measures the amount of hydrogen contained in the formation. The amount of hydrogen is not constrained by pore volume alone. The presence of clay and the type of fluid strongly influence the hydrogen count and, as a result, a bias in the porosity values is introduced. In order to avoid the bias, knowledge of rock matrix and pore fluid types is required, to allow for proper calibration or correction of porosity values. A similar effect is found when the porosity is obtained from other types of well-log. For example, in the case of porosity determination from sonic or density logs the main problem is the effect derived from the mud cake. Calibrations are made to compensate this effect and the uncertainty associated with the calibration is stronger in the case of mud cake heterogeneous width.

In this paper, porosity determination along the well is addressed as an inference problem and independent well-log datasets are integrated in a full formulation. The

solution follows the Bayesian methodology of inference, as presented in Jeffreys (1939). The Bayesian approach focuses on obtaining a probability density function for the parameter under investigation, named posterior distribution, assimilating two kinds of information: information on data-fit and prior information on models. Once this goal is achieved, all inferences can be obtained from the posterior by computing statistics relative to individual parameters (e.g., marginal distributions). Previous geophysical applications of Bayesian inference are usually based on multidimensional Gaussian distributions as, for example, Duijndam (1988) and Gouveia and Scales (1998). These works reduce to finding the posterior mean and covariance, both of which can be computed by using standard least-squares methods.

The final result is a marginal distribution for porosity at each well interval based on a series of logs (porosity, sonic for compression wave, sonic for shear wave, and density) and petrophysical models. Following the Bayesian methodology of inference, the available information (from datasets) is treated as probability density functions (pdf). Since the final result is represented as pdfs, the proposed methodology provides not only the porosity estimates, but also the uncertainty associated with the estimates. The contribution of each piece of information towards increasing the estimates' reliability can be measured. Another advantage of this approach is that it provides a unified framework for reservoir characterization by including inversion of surface seismic data with assessment of the uncertainties associated with the well-log data and its effect on a porosity model obtained within an inter well space.

The well-logs are the effect of several petrophysical properties, mainly porosity, mineral composition, fluid properties, saturation, pressure and temperature. It is impractical to account for all such influences in a petrophysical inference problem. One solution is to consider some of these properties as prior knowledge. Another one is to consider some of these properties as constant and to use a calibrated rock physics model that does not account for them. These practical approaches incorporate uncertainty in the final result. Synthetic and real examples are presented. These tests show that such uncertainties should be strongly reduced by the integration of information regarding porosity from independent well-logs by this proposed methodology.

A brief introduction of the Bayes Theorem is given in the next section and the development of the Bayesian formulation to access the posterior distribution for this specific inference problem is described.

BAYESIAN FORMULATION

Bayes Theorem

The problem consists of making inferences about interval porosity ϕ from the set of well-log data **d**, mathematical formulas derived from experimental rock physics studies and prior information I, which is any additional information obtained independently from the data. Following the Bayesian approach of inference the solution is given by the posterior distribution $p(\phi \mid \mathbf{d}, I)$. This function is obtained by applying the Bayes Theorem; which can be expressed by the following equation

$$p(\phi|\mathbf{d}) = \frac{l(\mathbf{d}|\phi)\ r(\phi|I)}{h(\mathbf{d}|I)}, \tag{1}$$

where $r(\phi|I)$ is the prior pdf, $l(\mathbf{d}|\phi)$ is the joint pdf for the data, also known as the likelihood function, and $h(\mathbf{d}|I)$ is a normalizing pdf that ensures posterior distribution as a pdf. The posterior pdf should be expressed as the normalized product of prior distribution and likelihood function.

To consider the posterior pdf as the solution of an inverse problem, the likelihood must be defined (i.e. the relationship between the data **d** and the parameter ϕ exists and is known), and there are compatibilities between the prior understanding of the model and the final results, i.e., $l(\mathbf{d}|\phi)>0$ for some ϕ where $r(\phi|I) > 0$. Now it is necessary to define the likelihood function and the prior distribution to access the posterior distribution.

Likelihood Function

This work follows standard steps to construct the likelihood function, which is summarized by: i) selecting the datasets which carry information about ϕ ii) finding mathematical expressions relating each type of porosity; and iii) defining statistical models (pdfs) for data distributions, based on data uncertainty.

Let $\mathbf{d}_\phi$, $\mathbf{d}_\alpha$, $\mathbf{d}_\beta$ and $\mathbf{d}_\rho$ be a set of vectors representing well-log data respectively composed of neutron porosity, sonic compression-velocity, sonic shear-velocity and density logs. Each of these data vectors should be represented as a sum of a porosity function, which is a deterministic variable, and an error component that is a probabilistic variable. The error components incorporate the uncertainties, which are associated with the process of data acquisition and with the theoretical knowledge. Using numbered subscripts to simplify the notions; these datasets can be expressed as the following equations:

$$\begin{aligned}
\mathbf{d}_\phi &\rightarrow \mathbf{d}_1 = \mathbf{f}_1(\phi) + \mathbf{e}_1, \\
\mathbf{d}\alpha &\rightarrow \mathbf{d}_2 = \mathbf{f}_2(\phi) + \mathbf{e}_2, \\
\mathbf{d}_\beta &\rightarrow \mathbf{d}_3 = \mathbf{f}_3(\phi) + \mathbf{e}_3, \\
\mathbf{d}_\rho &\rightarrow \mathbf{d}_4 = \mathbf{f}_4(\phi) + \mathbf{e}_4,
\end{aligned} \tag{2}$$

where f_i for i =1,…,4, represents the functions that relate the datasets and porosity and $\mathbf{e}_i$ for i=1,…,4, represents the errors in the data.

The next step is to define the relationships between data vectors and porosity, represented by the f_i, i=1,…,4 functions. As the value provided by the neutron well-log is a porosity estimate, function f_1 is the identity function. The expressions relating α and β to porosity employed in this report, respectively f_2 and f_3, are derived from Eberhart-Phillips et al. (1989). That article presents rock physics models derived from a multivariate analysis of laboratory measures on water-saturated samples of 64 different

sandstones. Such models consider the influence of effective pressure *Pe*, porosity ϕ, and clay content γ on velocities α and β. They are given by

$$\alpha = -5.77 - 6.94\phi - 1.73\sqrt{\gamma} + 0.446(Pe - e^{-16.7Pe}); \quad (3)$$

$$\beta = -3.70 - 4.94\phi - 1.57\sqrt{\gamma} + 0.361(Pe - e^{-16.7Pe}). \quad (4)$$

These rock-physics models are derived empirically from a set of sandstone samples and thus they apply only to the set of rocks studied. The regression coefficient should be recalibrated from core samples and logs at the site being studied.

The f_4 function is described by the equation

$$\rho = (1-\phi)\rho_{\mathrm{m}} + \phi\,\rho_{\mathrm{f}}, \quad (5)$$

where ρ_{m} is the matrix density and ρ_{f} is the fluid density.

Considering that these well-log data are independent, the likelihood function should be rewritten as the product of four independent data pdfs as follows

$$l(\mathbf{d}\,|\,\phi) = l_1(\mathbf{d}_\phi\,|\,\phi)\, l_2(\mathbf{d}_\alpha\,|\,\phi)\, l_3(\mathbf{d}_\beta\,|\,\phi)\, l_4(\mathbf{d}_\rho\,|\,\phi), \quad (6)$$

and consider $\mathbf{d} = [\mathbf{d}_\phi \quad \mathbf{d}_\alpha \quad \mathbf{d}_\beta \quad \mathbf{d}_\rho\ ]^{\mathrm{T}}$.

Next, it is necessary to establish the criteria to select the probability density models for l_{i}, i = 1,…,4 data pdfs. We chose to use the principle of maximum entropy to assign probabilities and assume that the first- and second-order momentum information was sufficient to describe the errors. This choice criterion to construct a likelihood pdf will be used as a standard criterion elsewhere in this work. According to such choices, the data pdfs are normal and should be expressed as

$$l_{\mathrm{i}}(\mathbf{d}_{\mathrm{i}}\,|\,\phi,\mathrm{I}) = \frac{1}{2\pi\sigma_{\mathrm{i}}^{\mathrm{N}}}\, exp\left\{-\frac{1}{2\sigma_{\mathrm{i}}^{2}}\left[\mathbf{d}_{\mathrm{i}} - \mathrm{f}_{\mathrm{i}}(\phi)\right]^{T}\left[\mathbf{d}_{\mathrm{i}} - \mathrm{f}_{\mathrm{i}}(\phi)\right]\right\}, \quad (7)$$

where σ_i is the standard deviation of the ith data vector. Consider its square, the data error variance σ_{i}^2. Realistically, these scaling parameters of the data functions are not known a priori. They must be treated as unknown parameters to be estimated from the data. In this case, the posterior distribution must be rewritten as

$$p^{+}(\phi,\sigma_\phi,\sigma_{\mathrm{VS}},\sigma_{\mathrm{VP}},\sigma_\rho\,|\,\mathbf{d}_\phi,\mathbf{d}_\alpha,\mathbf{d}_\beta,\mathbf{d}_\rho,\mathrm{I}). \quad (8)$$

Data error variances are not, however, important estimates to consider as a final result of the inference procedure. Interval porosity and associated uncertainties are all that matters in this problem. This question will be addressed later.

Prior distribution

A prior distribution is used to restrict the parameter space, which will provide an improvement of the ill-posed inverse problem. Following Jeffreys' point of view (Jeffreys, 1936) there are two types of prior distribution: a data base prior distribution (DBP) and a non-data base prior distribution (NDBP). The DBP is derived from datasets available and analyzed prior to the dataset used to construct the likelihood function of the current work. NDBP is derived from theoretical knowledge of the physical medium and from the investigator's background experiences. It is important to note that the form of two final prior distributions may differ from each other if two investigators use the same prior information to construct them.

Consider the posterior pdf p^+ described by Equation (8), where ϕ and σ_i, i=1,…,4 are unknown parameters. The prior distribution must describe the prior knowledge of those parameters. The porosity and random noise of datasets are statistically independent. Following Zellner (1996) this type of prior pdf should be expressed as a product of independent functions as

$$r(\phi,\sigma_1,\sigma_2,\sigma_3,\sigma_4 \mid I) \propto r_\phi(\phi \mid I)\, r_1(\sigma_1 \mid I)\, r_2(\sigma_2 \mid I)\, r_3(\sigma_3 \mid I)\, r_4(\sigma_4 \mid I). \quad (9)$$

This report follows the Bayesians' most conservative practice to define the prior pdf. It considers that with all previous pieces of information available, the only statement is that porosity should fall between a 0 and 1 interval. The use of a boxcar function is consistent in expressing this prior information (for more details please refer to Zellner, 1996).

The prior knowledge of data variances is that these scaling parameters may vary between 0 and ∞. Following Jeffreys (1939), the use of an improper distribution is consistent with expressing complete unawareness of such parameters (i.e. any value between 0 and ∞ is equiprobable). A logarithm pdf should be deemed as uniformly distributed and the use of a $r(\sigma_i) \propto 1/\sigma_i$ function is considered a reasonable choice in describing this information. This pdf is invariant under power transformation. Thus, prior distribution should be expressed as

$$r(\phi,\sigma_1,\sigma_2,\sigma_3,\sigma_4 \mid I) \propto \frac{1}{\sigma_1\sigma_2\sigma_3\sigma_4},\ 0 \le \phi \le 1. \quad (10)$$

Note that the prior pdf is improper, i.e. it is not normalized.

Posterior distribution

With the Bayes Theorem applied, the posterior distribution should be

$$p^+(\phi,\sigma_\phi,\sigma_{VS},\sigma_{VP},\sigma_\rho \mid \mathbf{d}_\phi,\mathbf{d}_\alpha,\mathbf{d}_\beta,\mathbf{d}_\rho,I) \propto$$
$$\frac{1}{\sigma_1\sigma_2\sigma_3\sigma_4}\prod_{i=1}^{4}\sigma_i^{-N_i}\, exp\left\{-\frac{1}{2\sigma_i^2}[\mathbf{d}_i - f_i(\phi)]^T[\mathbf{d}_i - f_i(\phi)]\right\},\ 0 \le \phi \le 1. \quad (11)$$

where N_i is the number of data points in vector $\mathbf{d}_i$

Data error variances are not important estimates to be considered as the final result of the inference procedure. Interval porosity and associated uncertainty is all that matters in this problem. The Bayesian methodology provides a standard treatment for such cases of uninteresting parameters, which are referred to as nuisance parameters. The standard treatment consists of eliminating the nuisance parameters by integration of the joint posterior. This procedure, which is known in statistics as marginalization of the joint distribution, is applied and the marginal posterior pdf for porosity is obtained:

$$p(\phi \mid \mathbf{d}_\phi, \mathbf{d}_\alpha, \mathbf{d}_\beta, \mathbf{d}_\rho, \mathrm{I}) \propto \int_0^\infty p^+(\phi, \sigma_\phi, \sigma_{\mathrm{VS}}, \sigma_{\mathrm{VP}}, \sigma_\rho \mid \mathbf{d}_\phi, \mathbf{d}_\alpha, \mathbf{d}_\beta, \mathbf{d}_\rho, \mathrm{I}) \mathrm{d}\sigma_\phi \mathrm{d}\sigma_{\mathrm{VS}} \mathrm{d}\sigma_{\mathrm{VP}} \mathrm{d}\sigma_\rho$$

$$= \prod_{i=1}^{4} \left\{ [\mathbf{d}_i - \mathrm{f}_i(\phi)]^T [\mathbf{d}_i - \mathrm{f}_i(\phi)] \right\}^{-\frac{N_i - 1}{2}}, 0 \le \phi \le 1. \qquad (12)$$

Any inference question can be addressed to the posterior. For example, one can use the mean, median, or mode as estimates for the interval porosity and the standard deviation or confidence intervals as measure of uncertainty.

PRACTICAL IMPLEMENTATION

The procedure for the previous formulation to be applied to a set of well-log data is as follows. The one-dimensional space representing the well position is discretized at *M* regular intervals. The goal is a collection of posterior pdfs representing one posterior for each interval. A moving window, which has the length of *N* well-log observations, is applied to compute those posteriors. Every such posterior provides a porosity estimate, by its mode, and a measure of the associated uncertainty, by its spread.

Such posterior pdfs are presented by an image with a colour scale. The vertical axis of this image represents the depth, the horizontal axis represents porosity values and the colour scale represents the probability amplitudes of the posteriors. This procedure is schematically shown in Figure 1. This image provides an idea of the porosity variation along the well, given by the maximum probability density for the posteriors and the associated uncertainty given by the form of the posteriors. A value for the estimates should be easily numerically computed from the posteriors. A confident interval of the posterior pdf, which is easily computed, should be considerd a measure of the associated uncertainty.

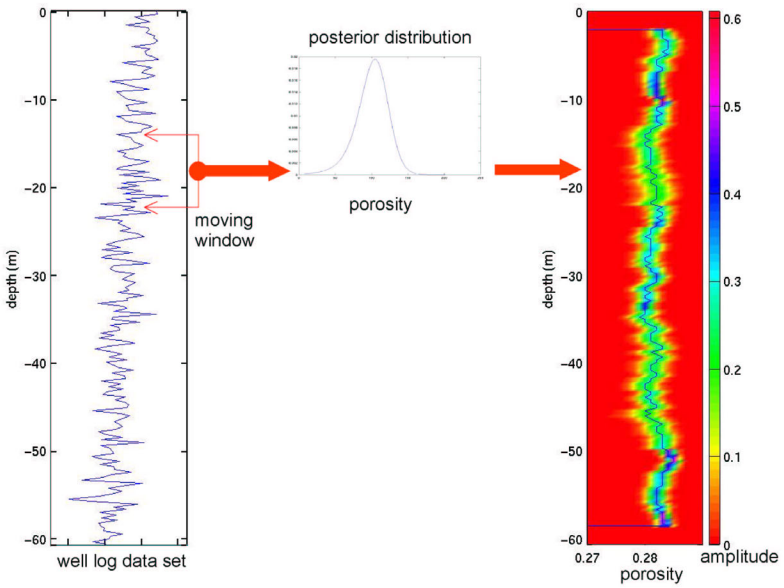

FIG 1: This figure shows how the formulation is implemented with a moving window. The vertical graph on the left represents a set of well-logs data used. A moving window (red line) runs along the well and for each window position, a posterior pdf is computed. The central graph represents a posterior pdf for one window position. In the end a collection of posterior pdfs is computed; one pdf for each interval of the discrete well and an image with a colour scale representing the amplitude of the posteriors is constructed. The image on the right shows an example. Visual analysis of this image provides an estimated model for porosity along the well, given by a curve that marks the high probability and provides a measure of uncertainty associated with this estimated model, given by the spread of the image around the maximum, which represents the spread of the posteriors around the mode.

This method has been tested on both synthetic and real data. In the next section we give a synthetic example to show how this methodology works.

SYNTHETIC DATA EXAMPLE

This section describes a reservoir model created to simulate synthetic well-log data of porosity, sonics (compression and shear waves), and density logs. The synthetic data are used to test the proposed methodology, after being corrupted with pseudo-random and coherent noises.

To evaluate the importance of each type of well-log data in increasing the reliability of porosity estimates, several inversions are performed considering both each log type alone and different combinations of logs.

The main goal of this example is to evaluate the methodology when the porosity log has an unknown shift due, for example, to a calibration error. The model consists of a 60 m thick sandstone layer, 0.26 porosity and 0.4 kbar/cm^3 effective pressure. The clay content is set to vary linearly from 0.10, at the top of the layer, to 0.20, at the bottom. Using the correct values for porosity, clay content, and pressure, a series of sonics (compression and shear sonic wave velocities) and density logs are generated from

Equations (3), (4) and (5) with an additional pseudo-random Gaussian noise and sample interval of 30 cm. For the porosity neutron log computation a true value of porosity is used, with additional pseudo-random Gaussian noise, and also affected by a constant 0.10 shift of the true porosity value to simulate a calibration error. All logs are shown in Figure 2.

The inversion procedure is applied with a 3 m-wide moving window. The correct clay content is assumed to be known and introduced in the likelihood function. As previously mentioned, different tests are performed to evaluate the contribution of each type of log to increasing the reliability of the estimated porosity model provided.

Figure 3 shows the inversion results. These colour images represent the results from various tests. In all images, the vertical blue line represents the true porosity of the model. The first four images in the top row of this figure, images (a), (b), (c) and (d) respectively, represent the images obtained from the use of a single log in the process of inversion; respectively porosity neutron, compression wave, shear wave, and density log: i.e. the posterior distributions $p(\phi|\mathbf{d}_\phi, I)$, $p(\phi|\mathbf{d}_\alpha, I)$, $p(\phi|\mathbf{d}_\beta, I)$, and $p(\phi|\mathbf{d}_\rho, I)$. The inversion of the neutron log is clearly affected by the calibration error. The results also demonstrate that lowest porosity resolution is obtained from the density log by the larger spread of the posterior pdfs and the best resolution is provided from the shear sonic-log.

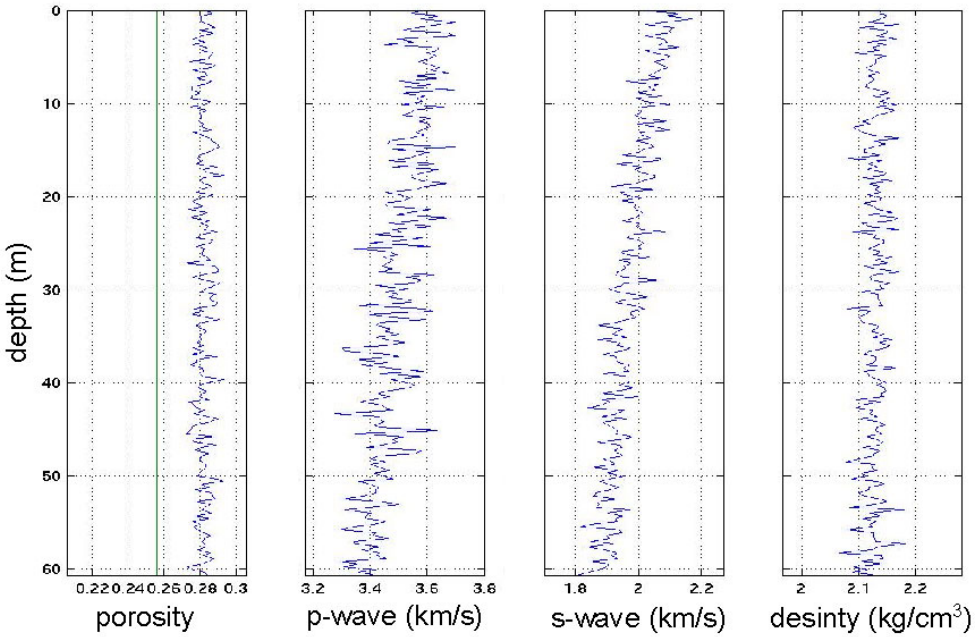

FIG 2: Well-log profiles generated from a sandstone model used for a synthetic test, which are from left to right, porosity neutron log, sonic logs (compressional and shear wave) and density log, respectively. These logs are simulated with the 30 cm sample interval and are corrupted with pseudo-random Gaussian noise. In addition to the Gaussian noise, the porosity log is also affected by a constant 0.10 shift of the true porosity value to simulate a calibration error. The vertical line on the porosity neutron log shows the actual porosity value of the model.

Tests with a pair of logs are performed to provide analysis of the improvement in the reduction of the shift bias in porosity estimate from a neutron log. Tests results are given

by the posterior pdfs $p(\phi|\mathbf{d}_\phi, \mathbf{d}_\alpha, I)$, $p(\phi|\mathbf{d}_\phi, \mathbf{d}_\beta, I)$, and $p(\phi|\mathbf{d}_\phi, \mathbf{d}_\rho, I)$ respectively, which are shown the three first images in the bottom row of this figure, images (e), (f), and (g) respectively. All results are still strongly biased by the calibration error. When using porosity with a shear sonic log the results are still strongly biased by the calibration error, but at some depths a second mode clearly marks the true porosity value. Two tests are performed using neutron porosity and sonic logs and all logs, resulting in the posterior pdfs $p(\phi|\mathbf{d}_\phi, \mathbf{d}_\alpha, \mathbf{d}_\beta, I)$ and $p(\phi|\mathbf{d}_\phi, \mathbf{d}_\alpha, \mathbf{d}_\beta, \mathbf{d}_\rho, I)$ respectively, shown in images (i) and (j). The bias is completely eliminated only when all logs are used in the inversion; image (j). The graph (e) in the top row of this figure shows the misfit between the true porosity and the estimated porosity model for this last test ($p(\phi|\mathbf{d}_\phi, \mathbf{d}_\alpha, \mathbf{d}_\beta, \mathbf{d}_\rho, I)$). Maximum misfit for porosity estimated from the mode of the posteriors showed in image (j) is 0.005.

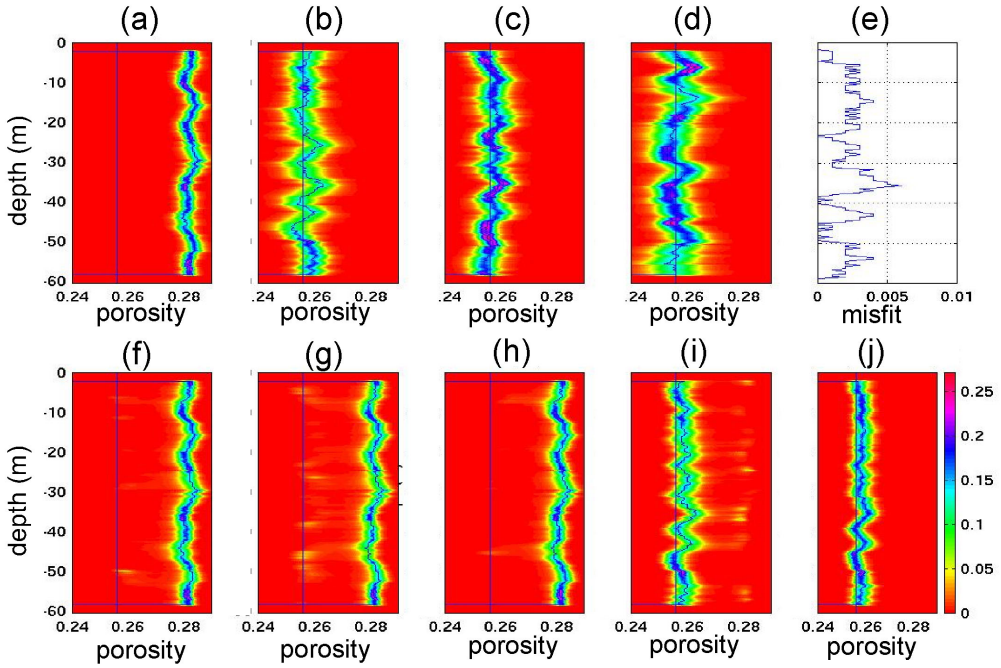

FIG 3: Images and a graphic showing the results of the tests with a synthetic model. The images represent the posterior distribution for porosity from different tests. A vertical line in all images represents the actual value of the porosity model. These colour images are as presented in Figure 1. The first four images in the top row, images (a), (b), (c), and (d), respectively, represent the results obtained from the use of a single log in the process of inversion; neutron ($p(\phi|\mathbf{d}_\phi, I)$), compression sonic wave ($p(\phi|\mathbf{d}_\alpha, I)$), shear sonic wave ($p(\phi|\ \mathbf{d}_\beta,\ I)$), and density $p(\phi|\mathbf{d}_\rho, I)$) logs respectively. The results of tests with pairs of logs, with the posteriors $p(\phi|\mathbf{d}_\phi, \mathbf{d}_\alpha, I)$, $p(\phi|\mathbf{d}_\phi, \mathbf{d}_\beta, I)$, and $p(\phi|\mathbf{d}_\phi, \mathbf{d}_\rho, I)$, are shown by images (f), (g), and (h), respectively. Tests using neutron and sonic logs ($p(\phi|\mathbf{d}_\phi, \mathbf{d}_\alpha, \mathbf{d}_\beta, I)$) and using all logs ($p(\phi|\mathbf{d}_\phi, \mathbf{d}_\alpha, \mathbf{d}_\beta, \mathbf{d}_\rho, I)$, are shown by images (i) and (j), respectively. Graphic (e) in the top row of this figure shows the misfit between the actual porosity and the estimated porosity model from the modes of $p(\phi|\mathbf{d}_\phi, \mathbf{d}_\alpha, \mathbf{d}_\beta, \mathbf{d}_\rho, I)$.

These synthetic tests have shown that the proposed methodology can integrate the information from independent datasets in a posterior pdf. The information on the medium that is consistently present in all datasets, such as porosity information, strongly influences the posterior. Otherwise, error components, which are present in some particular datasets, have their effect minimized on the posterior.

Next, a real test is presented to evaluate the practical applicability of this methodology.

REAL DATA EXAMPLE

The applicability of this methodology was investigated in a well-log dataset from lower Cretaceous sandstone of Glauconitic Formation at the Blackfoot Field, located 15 km southeast of Strathmore, Alberta, Canada. The target rocks are incised valley-fill sediments within the Glauconitic Formation. In the Blackfoot area the Glauconitic sands thickness varies from 0 to over 35 m. It is subdivided into three phases of valley incision which, however, is discontinuous. The lower and upper members are made of quartz sandstone with 0.18 average porosity, while the middle member consists of low porosity tight lithic sandstone.

Figure 4 shows the datasets from well 0808 used for this example, i.e. the porosity neutron log together with the porosity estimated from core laboratory analysis, sonic logs (compression and shear waves), density log, and the clay content estimated from the gamma-ray log. Note the vertical heterogeneity shown by the clay content plot. The 1683 m - 1692 m interval has a high clay volume and the porosity estimated by the porosity log at this interval is biased by a shift that results in an upper evaluation of the porosity when compared with the porosity estimated from core samples. This shift clearly represents the effect of the hydrogen from the clay content on the porosity neutron log.

The main goal of this example is an analysis of the improved resolution of the estimated porosity model from the posterior pdfs, which integrates information from independent well-log datasets.

One key step of this methodology is to define the rock physics models to construct the likelihood function. The relations described from equations (3) and (4) are empirical and they apply only to the set of rocks studied for that work (Eberhart-Phillips et al., 1989). Although these authors have shown that the results should extend in general to many consolidated sandstones in any case, if possible, the regression coefficients should be considered for the specific site being studied.

So as to improve the process of application of this methodology we conducted a petrophysical study using laboratory core samples and well-log data from the same reservoir, through from a different well. In the next section we present a pretophysical study using a collection of well-log dataset from Glauconitic producer sandstone formation in Blackfoot Area.

Rock physics regression-coefficients inference

Core samples and well-log data from wells 16-05 and 16-08 from the Blackfoot area provided the information on porosity (from core samples analysis), clay content (from gamma ray log), and sonic velocities (from compression and shear wave logs) for the following rock physics study.

The normal multiple-regression model analysis (Zellner, 1996) was adopted in a Bayesian inference for the regression coefficients of the Eberhart-Phillips models (equations (3) and (4)).

The effective pressure found from p-wave sonic log is 0.11 kbar in both wells. Note that as we do not have a pressure gradient it is impossible to compute the regression coefficient for effective pressure. In this work the fourth term of the Eberhart-Phillips model is incorporated in the first regression. Substituting the resulting rock physics models for f_2 and f_3 and writing the expressions in, a vector form is found

$$\mathbf{d}_i = \mathbf{X}\mathbf{m}_\mathbf{i} + \mathbf{e}_i, \tag{13}$$

where the vector of n observations $\mathbf{d}_i$, represents the dependent variable and $\mathbf{X}$ is a matrix with the known petrophysical properties porosity ϕ and clay content γ along the wells, which are the independent variables. $\mathbf{m}_\mathrm{i}$ is a vector of the regression coefficients a_i, b_i, and c_i. $\mathbf{e}_i$ is a vector of the disturbances in the observations and incorporate the uncertainties related with the rock physics model too. For compression i=2 and i=3 for shear wave.

This inference aims at the posterior pdf for the regression coefficients m_i, given the data d_i and the prior information I represented by $p(m_i|d_i,I)$. The Bayes theorem is applied. We followed the same statements described in this report to access the likelihood pdf and the prior information. The posterior distribution mode is considered an estimate for the regression coefficients. The rock physics models with these estimated regression-coefficients are considered the calibrated Eberhart-Phillips models for the lithologies being studied.

The marginal joint posterior distributions for regression coefficients are shown in Figures 5 and 6, respectively for compression and shear wave. In both figures, the marginal joint posterior pdfs for a pair of regression coefficients, $p(a_i, b_i \mid \mathbf{d}_i)$, $p(a_i, c_i \mid \mathbf{d}_i)$, and $p(b_i, c_i \mid \mathbf{d}_i)$, are shown at the top, from left to right. The ellipse form showed in these images represents the correlation between the petrophysical properties imposed by the rock physics models. The marginal posterior pdfs for each of the regression-coefficients, $p(a_i, \mid \mathbf{d}_i)$, $p(b_i \mid \mathbf{d}_i)$, and $p(c_i \mid \mathbf{d}_i)$, respectively, are shown at the bottom of those figures, from left to right. The form of these functions represents the resolution of the estimated coefficients.

The final calibrated rock physics models are

$$\alpha = 4.45 - 4.12\phi + 0.38\sqrt{\gamma}, \tag{14}$$

$$\beta = 2.89 - 1.61\phi + 0.08\sqrt{\gamma}. \tag{15}$$

To evaluate such rock physics models, compression and shear velocities are computed using equations (14) and (15) for the well 0808 and using porosity estimates from core laboratory analysis and clay content estimates from gamma ray log. Figures 7 and 8 display the cross-plots for compression- and shear-waves, respectively between the computed velocities and the measured velocities from the well and cross-plots between velocities and porosity and clay content, with the observed and computed values.

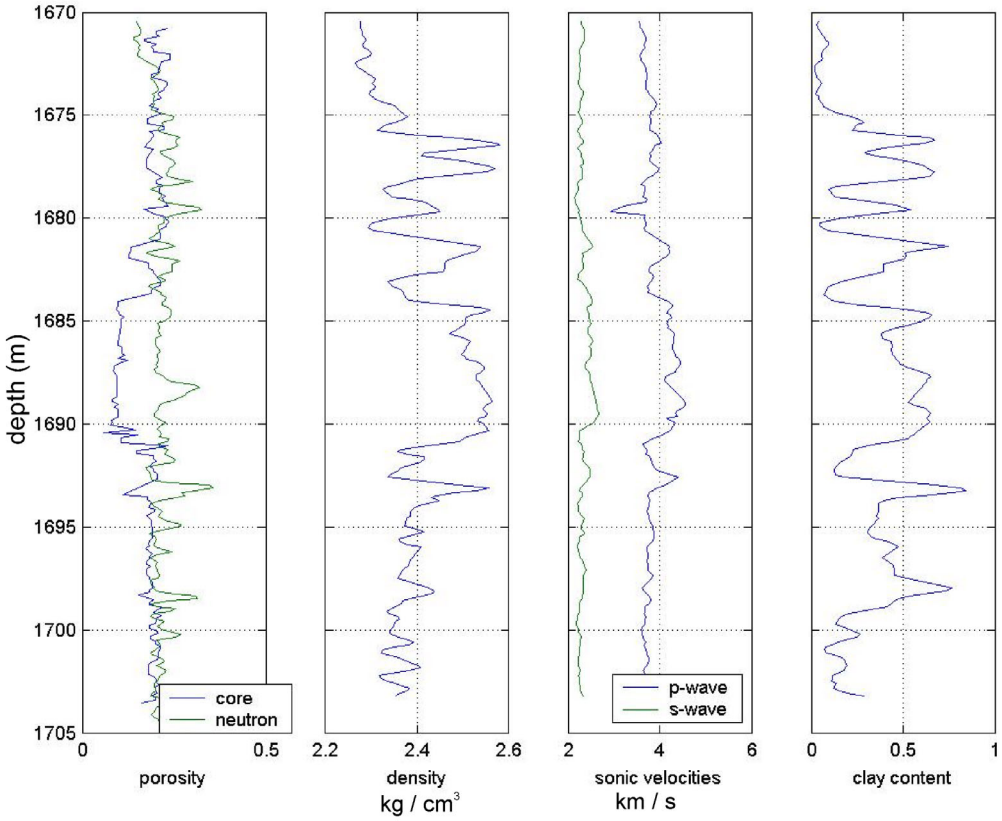

FIG 4: These plots represent the well-logs dataset used in the real example. In the first plot the porosity neutron log and the porosity estimated from core analysis. The clay content plot is derived from the gamma-ray log. The interval represents the Glauconitic sandstone from the Blackfoot area (well 08-08) that has 0.18 average porosity.

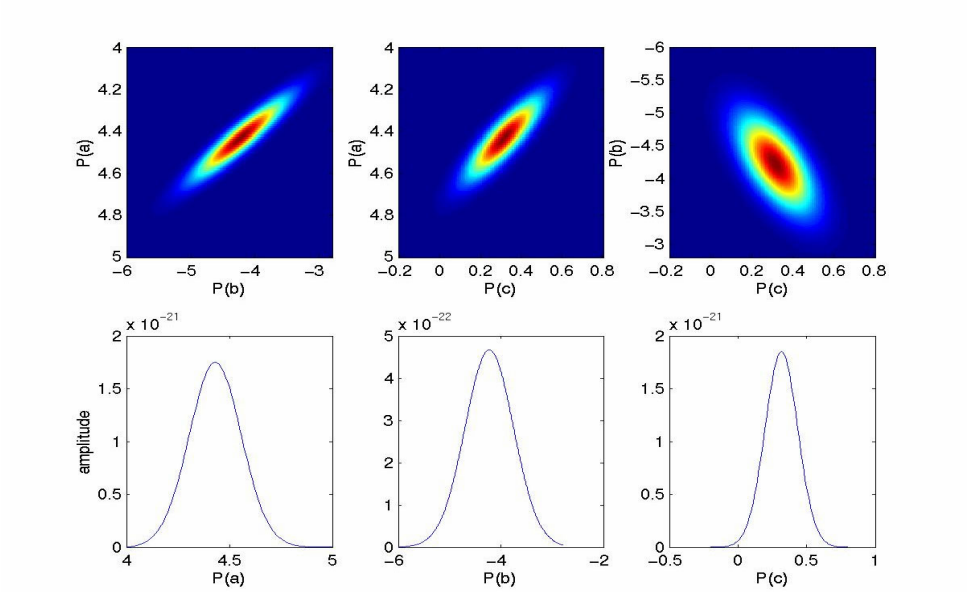

FIG 5: Resulting marginal posteriors for the regression coefficients of the compression-wave rock-physics model. The marginal joint posterior pdfs for a pair of regression coefficients, $p(a_i, b_i \mid \mathbf{d}_i)$, $p(a_i, c_i \mid \mathbf{d}_i)$, and $p(b_i, c_i \mid \mathbf{d}_i)$, are shown at the top, from left to right. The marginal posterior pdfs for each regression coefficients, $p(a_i, \mid \mathbf{d}_i)$, $p(b_i \mid \mathbf{d}_i)$, and $p(c_i \mid \mathbf{d}_i)$, respectively, are shown at the bottom, from left to right.

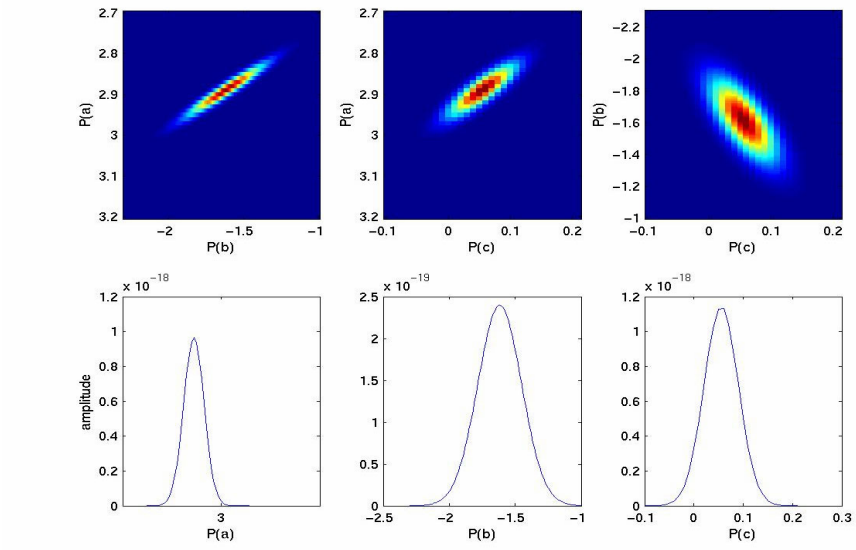

FIG 6: Resulting marginal posteriors for the regression coefficients of the shear-wave rock-physics model. Marginal joint posterior pdfs for a pair of regression coefficients, $p(a_i, b_i \mid \mathbf{d}_i)$, $p(a_i, c_i \mid \mathbf{d}_i)$, and $p(b_i, c_i \mid \mathbf{d}_i)$, are shown at the top, from left to right. Marginal posterior pdfs for each regression coefficients, $p(a_i, \mid \mathbf{d}_i)$, $p(b_i \mid \mathbf{d}_i)$, and $p(c_i \mid \mathbf{d}_i)$, respectively, are shown at the bottom, from left to right.

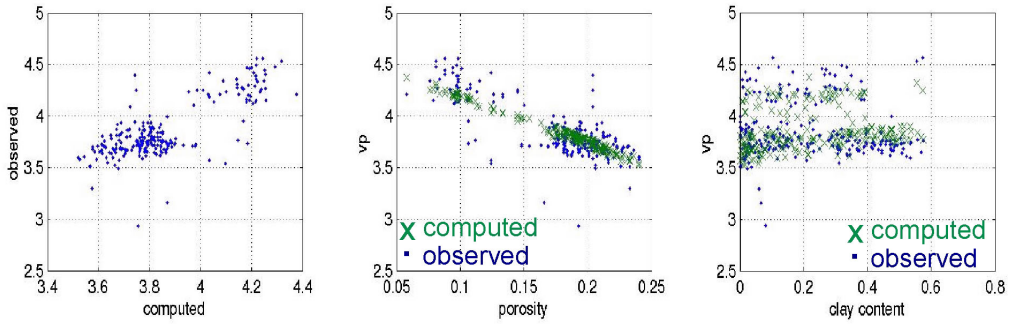

FIG 7: These cross-plots depict the computed compression-wave velocity from Equation (14) against the observed compression sonic-log (left), porosity against compression-wave velocity (centre); observed and computed from Equation (14) and clay content against compression-wave velocity (right); observed log and computed from Equation (14). These data are from well-log 0808. The porosity estimates used are derived from core samples, the clay content from gamma-ray log and the sonic velocity data from compression sonic-log.

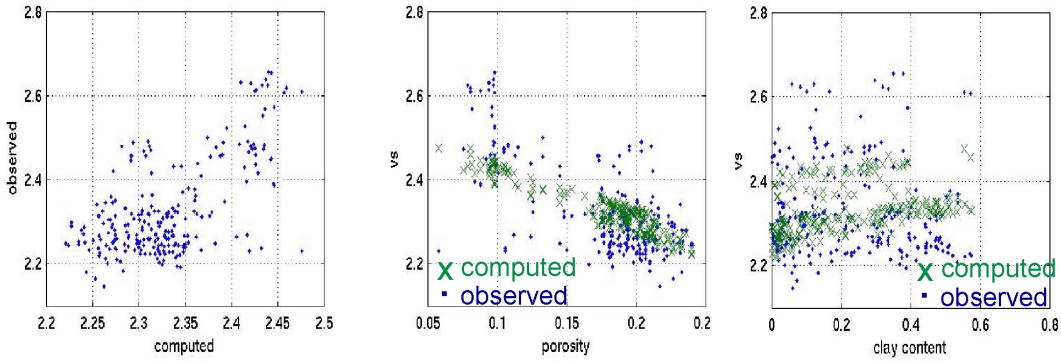

FIG 8: These cross-plots depict the computed shear-wave velocity from Equation (15) against the observed shear sonic-log (left), porosity against shear-wave velocity (centre); observed log and computed from Equation (15) and clay content against shear-wave velocity (right); observed log and computed from Equation (15). These data are from well-log 0808. The porosity estimates used are derived from core samples, the clay content estimates from gamma-ray log, and the sonic velocity data from shear sonic-log.

The next step is to apply a moving window operator, using the rock-physics models of equations (14) and (15) in the likelihood functions for a porosity inference. The results of the process of porosity inference are shown in the section below.

Porosity inference

As done in the synthetic example, several inversions are performed from the real dataset, considering each log type alone and different combinations of logs. Figure 9 shows the images representing the posterior pdfs for the use of single logs, porosity neutron, compression wave, shear wave, and density logs - images $p(\phi|\mathbf{d}_\phi,I)$, $p(\phi|\mathbf{d}_\alpha,I)$, $p(\phi|\mathbf{d}_\beta,I)$, and $p(\phi|\mathbf{d}_\rho,I)$ respectively. The dark dots in these images represent the porosity estimates from core samples laboratory analysis. Note in the first image on the left, representing the inversion results from a porosity neutron log ($p(\phi|\mathbf{d}_\phi,I)$) alone, the effect of high clay content in the 1683 m - 1692 m interval. The fourth image, which represents

the result from the inversion of the density log ($p(\phi|\mathbf{d}_\rho,I)$), shows a good porosity estimate within this interval but outside this interval, the porosity is under-evaluated. The other images show at some depths the modes of the posterior pdfs with a misfit regarding the porosity from core samples. The plot (e) on the right represents the modes from these four images' posterior pdfs. We can see that each log type provides a different porosity model after inversion. These different results are the effect of the independent uncertainties associated with each well-log method.

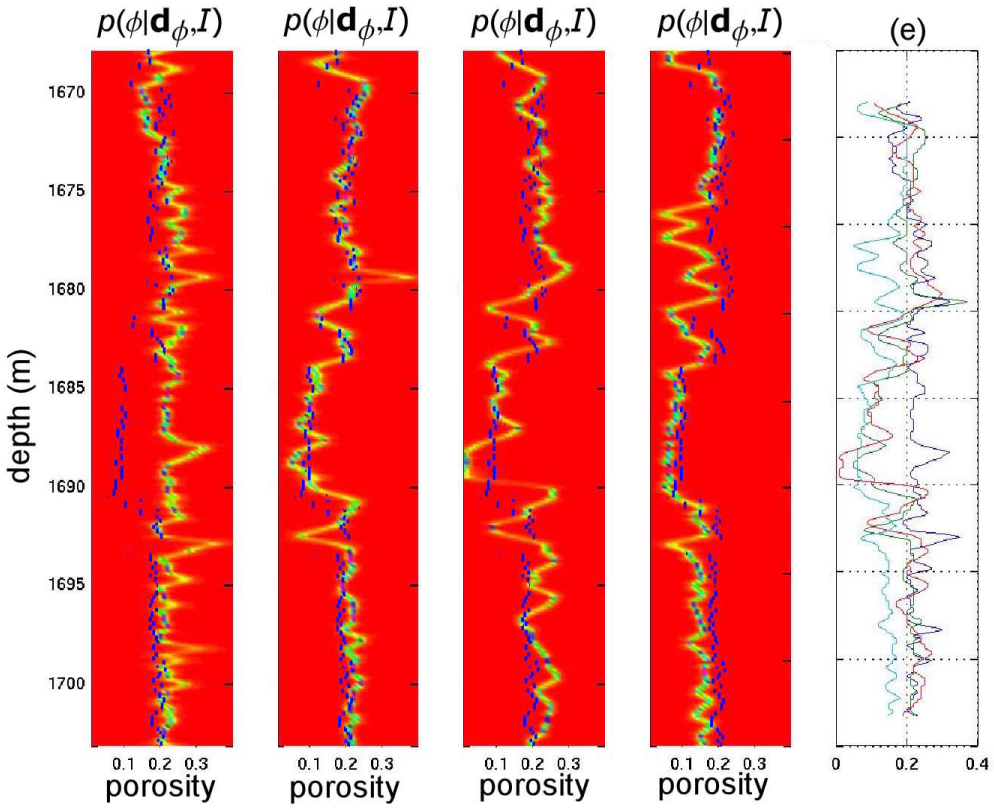

FIG 9: Images representing the posterior pdfs from tests using single logs, i.e. porosity neutron, compression wave, shear wave, and density logs - images $p(\phi|\mathbf{d}_\phi,I)$, $p(\phi|\mathbf{d}_\alpha,I)$, $p(\phi|\mathbf{d}_\beta,I)$, and $p(\phi|\mathbf{d}_\rho,I)$, respectively. The blue dots in these images represent the porosity estimates from core sample laboratory analysis. The plot (e) on the right represents the modes from these four images' posterior pdfs.

The results of the inversion from combinations of logs are shown in the images in Figure 10. The three images on the left represent the results from pairs of logs, porosity neutron with compression wave sonic logs ($p(\phi|\mathbf{d}_\phi,\mathbf{d}_\alpha,I)$), porosity neutron with shear wave sonic logs ($p(\phi|\mathbf{d}_\phi,\mathbf{d}_\beta,I)$), and porosity neutron with density logs ($p(\phi|\mathbf{d}_\phi,\mathbf{d}_\rho,I)$), respectively. As in the synthetic examples, these posterior pdfs show high influence of the porosity neutron log. The fourth image represents the result of the inversion from porosity neutron with sonic (compression and shear wave) logs ($p(\phi|\mathbf{d}_\phi,\mathbf{d}_\alpha,\mathbf{d}_\beta, I)$). This image shows a smaller influence of the neutron porosity log, but the modes of the posteriors pdfs and the core porosity estimates still have a considerable misfit. The last image on the right represents the results of the inversion from all log data put together

($p(\phi|\mathbf{d}_\phi,\mathbf{d}_\alpha,\mathbf{d}_\beta,\mathbf{d}_\rho,I)$). This image shows a considerably improved porosity estimated from the posterior pdf modes. Within the 1683 m - 1692 m interval the influence of the porosity neutron log vanished completely and out of this interval good porosity estimates could be found.

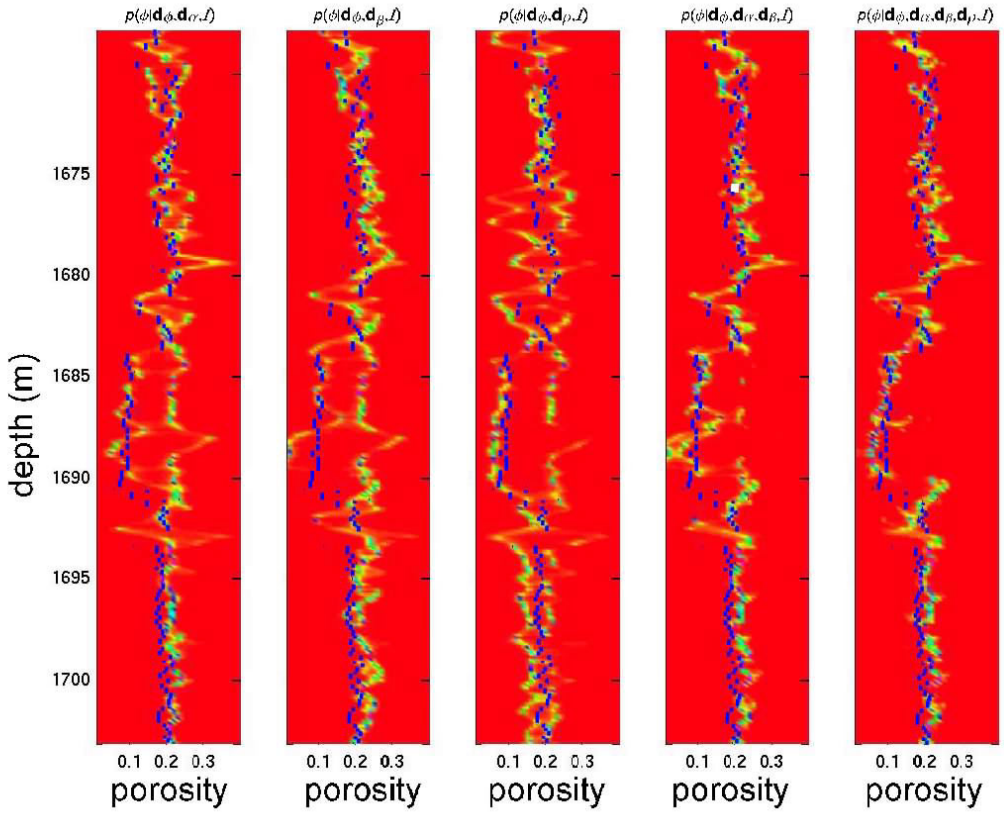

FIG 10: Results of the inversion from combinations of logs. The images from left to right represent the results from pairs of logs: porosity neutron with compression wave logs ($p(\phi|\mathbf{d}_\phi,\mathbf{d}_\alpha,I)$), porosity neutron with shear wave logs ($p(\phi|\mathbf{d}_\phi,\mathbf{d}_\beta,I)$), and porosity neutron with density logs ($p(\phi|\mathbf{d}_\phi,\mathbf{d}_\rho,I)$), respectively, from porosity neutron with sonic (compression and shear wave) logs ($p(\phi|\mathbf{d}_\phi,\mathbf{d}_\alpha,\mathbf{d}_\beta, I)$) and with all log data put together ($p(\phi|\mathbf{d}_\phi,\mathbf{d}_\alpha,\mathbf{d}_\beta,\mathbf{d}_\rho, I)$). The black dots in these images represent the porosity estimates from core samples laboratory analysis.

CONCLUSION

The Bayesian methodology of inference has resources to handle a considerable amount of information in a simple join-inversion process of well-log data. Results of the application of this methodology for porosity inference in synthetic and real example show that:

1. the most informative data about porosity is the shear wave velocity, followed by compression wave velocity and density;

2. the joint inversion of the synthetic well-logs reduced the effect of systematic error in idealized situations;

3. the joint inversion of real well-logs has provided treatment of the effect of complex combination of uncertainties from independent sources in porosity estimates.

FUTURE WORK

Investigations to extend this methodology for inference of other petrophysical properties has been started. This evaluation of the viability of using other rock physics models and other well-log data to infer fluid properties and permeability has been started with a new Bayesian formulation.

ACKNOWLEDGEMENT

We wish to thanks the sponsors of CREWES and to CNPq, a Brasilian Federal Agency for scientific and technological development. We would like to thank the members of CREWES, especially for the Professors Gary Margrave and Larry Lines, researcher, Charles Ursenbach, and the PhD student, Jonathan Downton, for helpful discussions.

REFERENCES

Duijndam, A.J.W., 1988, Bayesian estimation in seismic inversion. Part I: Principles. Geophys. Prosp., **36**, 878-898.

Eberhart-Philips, D., Han, D. and Zoback, M.D. 1989. Empirical relationships among seismic velocity, effective pressure, porosity and clay content in sandstone. Geophysics, **54**, 82-89.

Gouveia, W. and Scales, J.A., 1998, Bayesian seismic waveform inversion: Parameter estimation and uncertainty analysis, J. Geophys. Res., **103**, 2759-2779.

Jeffreys, H., 1936, Theory of Probability: Oxford University Press, London.

Zellner, A., 1996. An Introduction to Bayesian Inference to Econometrics, Wiley Interscience.