Spatial Estimation of Reservoir Properties Using Bayesian Wavelet Regression

G. Álvarez, Universidad Simón Bolívar,
B. Sansó, University of California Santa Cruz, R. Michelena, *i*Reservoir,
J.R Jiménez, Universidad Simón Bolívar
Contact-Author: Bruno Sansó, School of Engineering,
University of California, Santa Cruz, CA-95064, USA (bruno@ams.ucsc.edu)

Key Words: Bayesian variable selection, Wavelet regression, reservoir exploration

Abstract:

We consider the problem of estimating the properties of an oil reservoir, like porosity and sand thickness, in an exploration scenario where only a few wells have been drilled. We use gamma ray logs measured from the wells as well as seismic traces around the wells. We fit a linear regression model that accounts for the spatial correlation structure of the observations using an isotropic correlation function. We first transform the predictor variable using discrete wavelets. We then perform a Bayesian variable selection using a Metropolis search. We obtain predictions of the properties over the whole reservoir providing a probabilistic quantification of their uncertainties, thanks to the Bayesian nature of our method. The cross-validated results show that a very high accuracy can be achieved even with a very small number of wavelet coefficients.

1. Introduction

Predicting the properties of a reservoir using the information from well logs is a fundamental issue in petroleum management and exploration. This is usually achieved using and array of geostatistic techniques. In an exploration scenario, log data are usually scarce, as they are only available at few locations where wells have been drilled, whether seismic data are usually available for the whole reservoir. This presents a problem for most geostatiscs methods, since they are difficult to apply and generally fairly imprecise when the number of wells is small and seismic information is difficult to incorporate.

We present a method based on regressing the observations on a wavelet decomposition of a signal, either well logs or seismic traces. The method uses a Bayesian approach to estimate the property of interest on a location in a reservoir and quantify the uncertainty associated with the estimation. We present the model in Section 2. In Section 3 we discuss the methods used to estimate the parameters in the model and produce estimates of the properties at unobserved sites. In Section 4 we present the results. This work has been developed in Álvarez (2003) where a generalization that considers joint estimation of several properties using the information from both signals is undertaken.

2. The model

In this paper we consider data from 14 wells located in a reservoir in South Western Venezuela. For these wells we analyze γ-ray logs obtained at a depth of 9000. A grid of seismic traces covering an area of 100 Km^2 around the wells. The average window length was determined by the petrophysicist as 150 feet, which corresponds to 32 seconds in the seismic trace. Two reservoir properties are available: porosity and clay volume.

Well logs and seismic traces consist of series of 512 and 128 readings respectively. The 14 locations of the wells are irregularly scattered over the whole area of the reservoir. In order to relate the properties to the well logs or the traces, we consider the regression model

$$\boldsymbol{y} = \alpha \mathbf{1} + \boldsymbol{X}\boldsymbol{b} + \varepsilon \tag{1}$$

where $\boldsymbol{y} \in \mathbb{R}^n$ is the vector of properties at the locations of the wells. In our case $n = 14$. $\alpha \in \mathbb{R}$ is an intercept. $\boldsymbol{X} \in \mathbb{R}^{n \times p}$ is the matrix of the signals, either well logs or seismic traces and $p = 512$. $\boldsymbol{b} \in \mathbb{R}^p$ is a vector of coefficients and $\varepsilon \in \mathbb{R}^n$ is the error term. Notice that the model in (1) has a larger number of regression coefficients than data. Thus direct estimation of the $\boldsymbol{b}$ using traditional regression methods is unfeasible. We either have to impose some restrictions or consider prior information.

Motivated by the work of Brown and Vannucci (2001) we consider a wavelet transformation of the signals. The idea of such a transformation is that

a reduced number of wavelet coefficients should be able to capture the information in the signals needed to predict the value of the reservoir property at a given location. A discrete wavelet transformation is given by an orthogonal matrix (see for example Vidakovic, 1999, chap. 4), say $\boldsymbol{W} \in \mathbb{R}^{p\times p}$, such that $\boldsymbol{WW}' = \boldsymbol{I}$, where $\boldsymbol{I}$ denotes the identity matrix. We then have that

$$\boldsymbol{y} = \alpha\mathbf{1} + \boldsymbol{XWW}'\boldsymbol{b} + \varepsilon = \alpha\mathbf{1} + \boldsymbol{Z}\beta + \varepsilon \qquad (2)$$

where $\boldsymbol{Z} = \boldsymbol{XW}$ is the matrix of wavelet coefficients and $\beta = \boldsymbol{W}'\boldsymbol{b}$ the new regression vector.

The spatial correlation of the observation in $\boldsymbol{y}$ is captured by assuming that the error term ε corresponds to an isotropic random field with an exponentially decaying correlation function. Thus

$$\mathrm{cov}(\varepsilon_i, \varepsilon_j) = \sigma^2 \exp\left(-\frac{1}{\lambda}||s_i - s_j||\right),$$

where s_k denotes the location of the kth site. The proposed correlation can easily be substituted for any other parametric family of correlations offering wider flexibility, like the Matèrn class, as described for example in Stein (1999).

We take a Bayesian approach to estimate the parameters in the model and, following Brown and Vannucci (2001), we specify a prior for the original regression coefficients $\boldsymbol{b}$ as a p-variate normal with mean 0 and covariance matrix $\boldsymbol{H}$, denoted as $N_p(0, \boldsymbol{H})$, for $\boldsymbol{H}$ corresponding to the covariance matrix of an autoregressive process of order one. Such a distribution is used to guarantee that the components of $\boldsymbol{b}$ vary smoothly and that the variances of the transformed coefficient β show the typical decay of wavelet coefficients. The selection of the relevant wavelet coefficients is achieved by considering a prior distribution for the parameter β_i given by

$$p(\beta_i) \propto (1-\gamma_i)\delta_0 + \gamma_i N_1(0, \tilde{h}_i) \quad ,$$

where δ_0 is a point mass at 0 and γ_i is a binary variable that indicates whether the i-th coefficient is 0 or not. $\tilde{h}_i$ corresponds to the diagonal of the matrix $\tilde{\boldsymbol{H}} = \boldsymbol{WHW}'$. So, a priori, the i-th coefficient has probability $1-\gamma_i$ of being equal to zero and probability γ_i of being distributed as a normal with zero mean and variance $\tilde{h}_i$. To complete our model we consider the following prior distributions: $\alpha \sim N_1(0, h\sigma^2)$, $\gamma_i \sim Ber(\omega)$, $i = 1, 2\cdots, p.$, where $Ber(p)$ denotes a Bernoulli with parameter p, σ^2 follows an inverse gamma with parameters a_σ and b_σ and λ follows a gamma with parameters a_λ and b_λ. a_σ and b_σ were chosen to obtain a diffuse prior on σ^2. a_λ was taken as 0.4 for clay volume and 0.5 for porosity, b_λ was set to 4 in both cases. These values are compatible with the covariograms of the observations, but reflect a fair level of uncertainty.

3. Estimation and prediction

To obtain inferences on the parameters in our model we explore their joint posterior distribution using a Markov chain Monte Carlo method (MCMC) as proposed, for example, in Gamerman (1997). The MCMC that we use consists of sampling iteratively from the distributions of each of the parameters or blocks of parameters conditional on all the remaining ones. Thus, samples of α are obtained from a univariate normal. Samples of β are obtained from a multivariate normal and samples of σ^2 correspond to an inverse gamma. The spatial range λ is sampled by considering a Metropolis-Hastings step that consists of accepting or rejecting a proposed value with a probability that depends on the conditional density of λ evaluated at the current state of the chain.

Generating samples of $\gamma = (\gamma_1, \ldots, \gamma_{2^p})$ presents the challenge of dealing with a highly multivariate distribution, since in our case p is 512 when using the γ-ray logs as predictor and 128 when using the seismic traces. We proceed by considering a random initial configuration $\gamma^{(0)}$. Then, at each iteration, one the following two ways of choosing a candidate configuration is chosen with (fixed) probability ϕ: (a) Generate a new candidate by choosing at random a component. This component is deleted if it is part of the current configuration and added if it is not; (b) Select two components i and j such that $\gamma_i = 0$ and $\gamma_j = 1$ and swap their values. The proposed configuration is rejected or accepted following a Metropolis-Hastings rule. Experience shows that good predictions can be obtained with about 20 wavelet coefficients. Thus our prior distribution for γ is such that the prior expected number of coefficients, $p\omega$, is equal to 20. Thus $\omega = 0.16$ when the seismic traces are used, and $\omega = 0.04$ when the well logs are used.

The posterior predictive density of a new observation y_N, given the observed data $\boldsymbol{y}_{obs}$, can be estimated from the m simulated values from the MCMC for the joint parameter vector, say $\theta^{(j)}$, using the approximation

$$\hat{p}(y_N|\boldsymbol{y}_{obs}) = \frac{1}{m}\sum_{j=1}^{m} p(y_N|\theta^{(j)}).$$

In our case, in order to predict the value of a property for a specific location s_N, we use the information provided by the wavelet transformation of the signal

$z_N \in \mathbb{R}^p$ corresponding to s_N. Thus, using the j-th iteration from the MCMC,

$$p(y_N|\theta^{(j)}) = N\left(\alpha^{(j)} + \boldsymbol{z}_N'\beta^{(j)} + \frac{\boldsymbol{V}^{(j)}}{v_{N,2}^{(j)}} (\boldsymbol{y} - \alpha^{(j)}\mathbf{1} - \boldsymbol{Z}\beta^{(j)}), \boldsymbol{V}_r^{(j)}\right), \quad (3)$$

where

$$\boldsymbol{V}_N^{(j)} = \begin{pmatrix} \boldsymbol{V}^{(j)} & \mathbf{v}_{1,N}^{(j)} \\ \mathbf{v}_{N,1}^{(j)} & v_{N,2}^{(j)} \end{pmatrix} \in \mathbb{R}^{(n+1)\times(n+1)}$$

and

$$\boldsymbol{V}_r = \boldsymbol{V}^{(j)} - \frac{\mathbf{v}_{N,1}^{(j)}\mathbf{v}_{1,N}^{(j)}}{v_{N,2}^{(j)}} \in \mathbb{R}^{n\times n},$$

$\mathbf{v}_{1,N} \in \mathbb{R}^{(n\times 1)}$, $\mathbf{v}_{N,1} \in \mathbb{R}^{(1\times n)}$ and $[\boldsymbol{V}_N^{(j)}]_{k,l} = \sigma^2 \exp(-\frac{1}{\lambda^j}||s_l - s_k||)$. In words, to obtain a prediction at a location s_N we calculate the wavelet decomposition of the signal at that location and then, for each set of simulated values from the the MCMC, we calculate the spatial correlation matrix $\boldsymbol{V}_N^{(j)}$ and sample the normal distribution specified in (3). The result is a set of samples $y_N^{(1)}, \ldots, y_N^{(m)}$ from the posterior predictive distribution of y_N.

4. Results

We fitted model (2) separately for each property using first the γ-ray logs and then the seismic traces. We considered a wavelet transformation based on the Haar basis. We present results that were obtained from 5,000 iterations of a MCMC after a burn in period of 500 iterations. To explore the predictive capability of the model we adopted a "leave one out" approach, consisting on obtaining the posterior predictive distribution for each of the 14 locations using the remaining 13.

Figure 1 shows the predicted values of porosity and clay volume interpolated over the convex hull of the locations of the wells. The predictions for each well are given by the medians of the simulated values obtained from the 4,500 samples from the MCMC. In Figure 2 we compare the predictive distribution of each of the 14 wells, based on the remaining 13 wells to the actual observed values of clay volume. Similar results are obtained for porosity. Notice that the observations are very central to the predictive densities. The former shows that the method has a very high level of predictive accuracy. Also, given the Bayesian nature of the method, we are not only providing an estimate of the properties at each location, but a precise assessment of the uncertainties

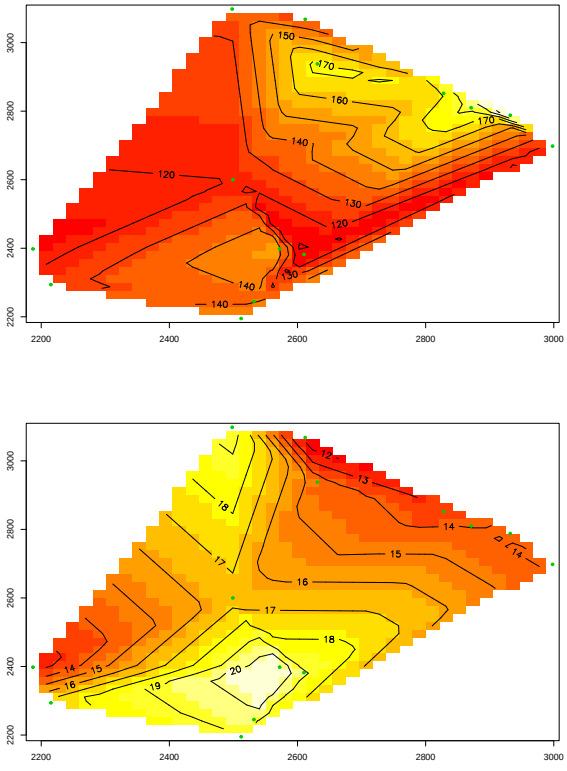

Figure 1: Linear interpolations of the predicted properties obtained using a "leave one out" estimation. The predictions at each well are obtained using the γ-ray logs. Top panel corresponds to clay volume and bottom panel corresponds to porosity

involved in such estimation, given by the predictive distribution.

The number of non-zero coefficients can vary from one iteration to the other of the MCMC. Nevertheless we observed that no more than 10 coefficients where different from zero at any given iteration. This implies that less than 2% of the wavelet coefficient contain enough information to accurately predict the values of porosity and clay volume.

Clearly, the predictive ability of the model depends on the number of wells that are used. To assess the robustness of the method with respect to the number of locations used for prediction we chose a well located at the center of the field and predicted its porosity using the remaining 13 locations. We then deleted one location at random at a time and obtained the prediction with the remaining locations. The results using γ-ray logs as predictors are shown in Figure 3 for porosity. A similar behavior is observed for clay volume. As expected, we observe

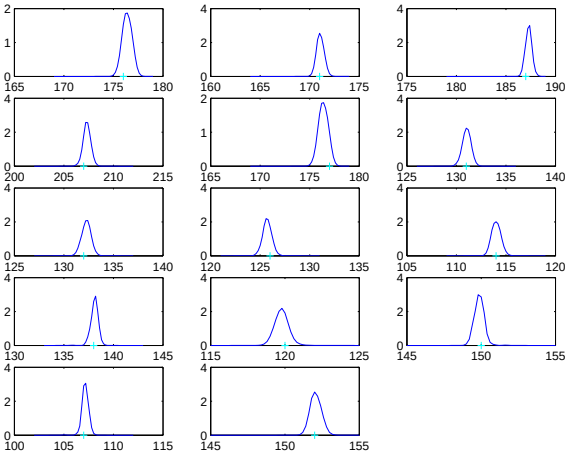

Figure 2: Predictive densities for porosity, obtained using γ-ray logs, at each of the 14 wells using the remaining 13. Actual observations are marked as a dot.

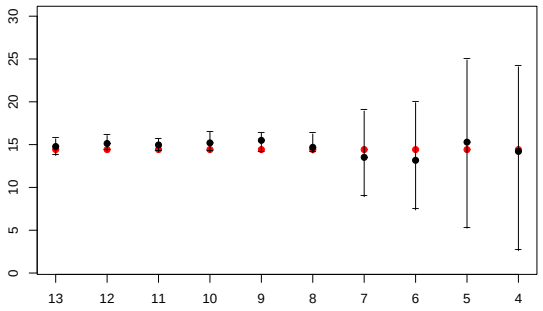

Figure 3: Predictive 95% probability intervals obtained for a centric well, using γ-ray logs, as a function of the number of predictive wells. Left panel corresponds to clay volume and right panel corresponds to porosity.

that the width of the predictive intervals increases as the number of wells decreases. Nevertheless observations and pointwise predictions are fairly close even for as little as six location for the clay volume and four for the porosity.

We repeated the whole analysis using the seismic traces as a predictor of both, clay volume and porosity. Figure 4 shows the interquartile ranges for the leave one out predictions. The accuracy of the predictions is lower than the one obtained when the γ-ray logs are used, but the performance of the model is remarkably good in this case as well.

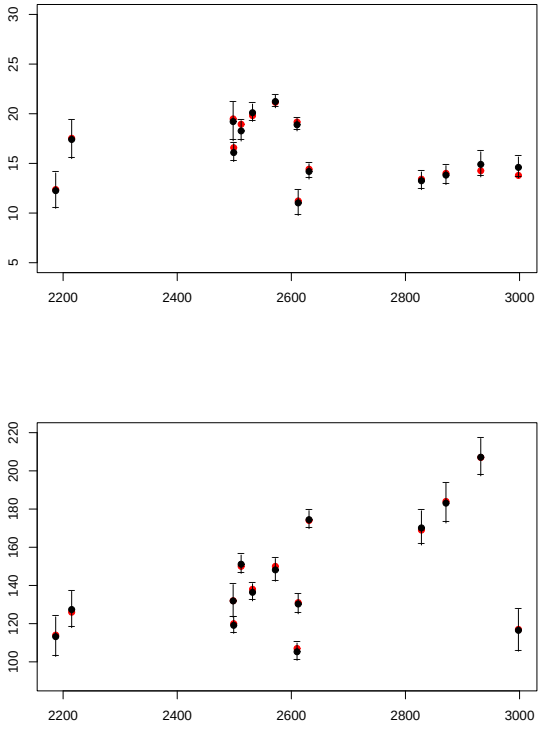

Figure 4: Interquartile intervals of each of the 14 wells using the remaining 13 wells. Predictions are obtained using seismic traces. Top panel corresponds to clay volume and bottom panel corresponds to porosity.

References

Álvarez, G. (2003) *Clasificación de Litologías y Estimación de Propiedades de Yacimientos a Partir de Datos Sísmicos*. Ph.D. thesis, Universidad Simón Bolívar, Caracas, Venezuela.

Brown, P.J. Fearn, T. and Vannucci, M. (2001) Bayesian wavelet regression on curves with application to a spectroscopic calibration problem. *Journal of the American Statistical Association*, **96**, 398–408.

Gamerman (1997) *Markov Chain Monte Carlo. Stochastic Simulation for Bayesian I nference*. U.S.A: Chapman and Hall.

Stein, M. L. (1999) *Interpolation of Spatial Data*. New York: Springer.

Vidakovic, B. (1999) *Statistical Modeling by Wavelets*. New York: John Wiley and Sons.