Porosity prediction using cokriging with multiple secondary datasets

Hong Xu, Jian Sun, Brian Russell, Kris Innanen

ABSTRACT

The prediction of porosity is essential for the identification of productive hydrocarbon reservoirs in oil and gas exploration. Numerous useful technologies have been developed for porosity prediction in the subsurface, such as multiple attribute analysis, kriging, and cokriging. Kriging allows us to create spatial maps from point information such as well log measurements of porosity. Cokriging combines well log measurements of porosity with seismic attributes recorded between the wells to improve the estimation accuracy of the overall map. However, the traditional cokriging for porosity estimation is limited to only one seismic attribute. To introduce more geological information and improve the accuracy of prediction, we develop a new cokriging system that extends traditional cokriging to two secondary variables. In this study, our new cokriging system is applied to the Blackfoot seismic data from Alberta, and the final estimated map is shown to be an improvement over kriging and traditional single attribute cokriging. To show this improvement, "leave-one-out" cross-validation is employed to evaluate the accuracy of porosity prediction with kriging, traditional cokriging, and our new approach. Compared to kriging and traditional cokriging, an improved porosity map, with higher lateral geological resolution and smaller variance of estimation error, was achieved using the new cokriging system. We believe that the new approach can be considered for porosity prediction in any area of sparse well control.

INTRODUCTION

Porosity prediction plays an essential role in process of predicting elastic rock properties and planning production operations (Doyen, 1988). Many techniques have been introduced to predict porosity in subsurface reservoirs, for instance, kriging, cokriging, multi-attribute analysis. The kriging system uses only high vertical resolution well log data in the spatial interpolation, but well logs are poorly sampled laterally. However, the advantage of kriging is that the well values are honored perfectly. On the other hand, multi-attribute analysis gives good spatial resolution if 3D seismic data is used, but it is hard to match the exact well values, since these values are predicted using a least-squares algorithm. Cokriging, a geostatistical technique, has been considered in porosity prediction since it's introduced into the geophysical industry by Doyen (1988) based on theory developed by Matheron (1965). The objective of the cokriging technique is to use attributes, such as acoustic impedance, amplitude or travel time extracted from 3D seismic data, as a secondary variable to guide the interpolation of related well log data, called the primary variable, such as porosity, shale volume or depth. Doyen (Doyen, 1988) applied cokriging to predict porosity by using acoustic impedance extracted as secondary variable from 3D seismic data. Cokriging produces maps that contain the spatial trends constructed by the spatial correlation function to model the lateral variations of the reservoir properties (Doyen et al., 1996).

The traditional cokriging system combines well log data and seismic attribute data, but only one secondary dataset is allowed in calculation. It is necessary to corroborate more than one seismic attribute to support the prediction because every attribute has a particular useful information about reservoir and to predict rocks properties (Guerrero et al., 1996). To optimize the secondary data, numerous methods have been proposed. Russell et al. (2002) combine cokriging and multi-attribute transforms. As Russell et al. (2002) illustrated, the secondary input of cokriging is an improved map generated by multi-attribute analysis. Babak and Deutsch (1992) improved the cokriging model by merging all secondary data into a single super secondary dataset and then implementing the cokriging system with the single merged secondary dataset. Nevertheless, those super secondary data were obtained under assumptions which are unpractical. For example, the multi-attribute algorithm assumes that the predicted areas are highly correlated to the well tie locations, and the linear combination of all secondary data is assumed to generate the super data by merging.

In this paper, to satisfy those assumptions and improve the estimation, we present a new approach that introduces two secondary variables in the cokriging. Two advantages are achieved with the new cokriging system. First, the lateral geological resolution of the final produced maps is increased at the locations away from the well locations because the secondary variable brings in extra geological information. Secondly, the addition of the second seismic data offers an opportunity to decrease the variance of the estimation error.

METHODOLOGY

The traditional cokriging method consisted of one primary and one secondary variable. To introduce more seismic attributes into the estimation, a new cokriging system consisting of one primary and two secondary variables is implemented. The new algorithm exploits the cross-correlation not only between the primary and secondary variables, but also between the two secondary variables.

As with the traditional cokriging algorithm (Isaaks and Srivastava, 1989), the cokriging system containing one primary and two secondary variables is defined as,

$$\hat{u}_0 = \sum_{i=1}^{n} a_i \cdot u_i + \sum_{j=1}^{m} b_j \cdot v_j + \sum_{k=1}^{p} c_k \cdot x_k \tag{1}$$

where $\hat{u}_0$ is the estimate of $\mathbf{U}$ at location 0. u_1, u_2, . . . , u_n are the primary data at n locations; v_1, v_2, . . . , v_m and x_1,x_2, . . . , x_p are the secondary data at m locations and k locations. a_1,a_2, . . . , a_n, b_1,b_2, . . . , b_m, and c_1, c_2, . . . , c_p are cokriging weights to be determined.

Then the estimation error can be written as

$$\mathbf{R} = \hat{u}_0 - u_0 = \sum_{i=1}^{n} a_i \cdot u_i + \sum_{j=1}^{m} b_j \cdot v_j + \sum_{k=1}^{p} c_k \cdot x_k - u_0 \tag{2}$$

where u_1, u_2, . . . , u_n are variables representing the $\mathbf{U}$ phenomenon at the n locations where $\mathbf{U}$ has been sampled. v_1, v_2, . . . , v_m are variables representing the $\mathbf{V}$ phenomenon at the m locations where $\mathbf{V}$ has been sampled. x_1,x_2, . . . , x_p are variables representing the $\mathbf{X}$ phenomenon at the p locations where $\mathbf{X}$ has been sampled.

Also, the matrix formation of equation (2) can be rewritten as

$$\mathbf{R} = \mathbf{w^t Z} \tag{3}$$

where $\mathbf{w^t}$ = (a_1, a_2, ..., a_n, b_1,b_2, ..., b_m, c_1, c_2, ..., c_p, -1) and $\mathbf{Z^t}$ = (u_1, u_2, ..., u_n, v_1, v_2, ..., v_m, x_1,x_2, ..., x_p, $\mathbf{u_0}$).

Then, the variance of $\mathbf{R}$ can be expressed as

$$\begin{aligned}
\mathbf{Var}\{\mathbf{R}\} &= \mathbf{w^t C_z w} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j Cov\{u_i u_j\} + \sum_{i=1}^{m}\sum_{j=1}^{m} b_i b_j Cov\{v_i v_j\} + \sum_{i=1}^{p}\sum_{j=1}^{p} c_i c_j Cov\{x_i x_j\} \\
&+ 2\sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j Cov\{u_i v_j\} + 2\sum_{i=1}^{n}\sum_{j=1}^{p} a_i c_j Cov\{u_i x_j\} + 2\sum_{i=1}^{m}\sum_{j=1}^{p} b_i c_j Cov\{v_i x_j\} \\
&- 2\sum_{i=1}^{n} a_i Cov\{u_i u_0\} - 2\sum_{i=1}^{m} b_i Cov\{v_i u_0\} - 2\sum_{i=1}^{p} c_i Cov\{x_i u_0\} \\
&+ Cov\{u_0 u_0\}
\end{aligned} \tag{4}$$

where $Cov\{u_i u_j\}$ is the auto-covariance between u_i and u_j, $Cov\{v_i v_j\}$ is the auto-covariance between v_i and v_j, and $Cov\{x_i x_j\}$ is the auto-covariance between x_i and x_j, $Cov\{u_i v_j\}$ is the cross-covariance between u_i and v_j, $Cov\{u_i x_j\}$ is the cross-covariance between u_i and x_j, and $Cov\{v_i x_j\}$ is the cross-covariance between v_i and x_j.

The traditional cokriging system must satisfy two conditions for the set of cokriging weights. First, the weights must be unbiased in equation (1). Secondly, the error variances must be as small as possible in equation (2). The new cokriging system is under the same conditions.

To tackle the unbiasedness condition, the expected estimation value in Equation (1) is computed as below,

$$\begin{aligned}
E(\hat{U}_0) &= E\{\sum_{i=1}^{n} a_i u_i + \sum_{j=1}^{m} b_j v_j + \sum_{k=1}^{p} c_k x_k\} \\
&= \sum_{i=1}^{n} a_i E\{u_i\} + \sum_{j=1}^{m} b_j E\{v_j\} + \sum_{k=1}^{p} c_k E\{x_k\} \\
&= \widetilde{m}_U \sum_{i=1}^{n} a_i + \widetilde{m}_V \sum_{j=1}^{m} b_j + \widetilde{m}_X \sum_{k=1}^{p} c_k
\end{aligned} \tag{5}$$

where $E\{U_i\}$ = $\widetilde{m}_U$, $E\{V_j\}$ = $\widetilde{m}_V$, and $E\{X_K\}$ = $\widetilde{m}_X$. To make this function to be unbiased, we need $\sum_{i=1}^{n} a_i = 1$, $\sum_{j=1}^{m} b_j = 0$, and $\sum_{k=1}^{p} c_k = 0$ as the unbiased conditions.

The next step is to minimize the error variance (equation (2)). The Lagrange multiplier method (Ito and Kunisch, 2008) is used to minimize a function with three constraints. We equate each non-biased condition to be zero, multiply by a Lagrange multiplier, and then add the result to equation (4). The following equation gives the mathematical algorithm behind Lagrange multipliers:

$$\mathbf{Var}\{\mathbf{R}\} = \mathbf{w^t C_z w} + \mu_{\mathbf{1}}(\sum_{i=1}^{n} a_i - 1) + \mu_{\mathbf{2}}(\sum_{j=1}^{m} b_j) + \mu_{\mathbf{3}}(\sum_{k=1}^{p} c_k) \tag{6}$$

where μ_1, μ_2, and μ_3 are the Lagrange multipliers. Considering the unbiased conditions we gave above, three additional terms in equation (6) are all equal to zero and do not contribute to the error variance $\mathbf{Var}\{\mathbf{R}\}$.

In order to minimize equation (6), the partial derivatives of $\mathbf{Var}\{\mathbf{R}\}$ with respect to the n+m+p weights (a, b, c) and three Lagrange multipliers (μ_1, μ_2, μ_3) have to be equal to zero because the minimum occurs at zero. Those functions are expressed as,

$$\begin{aligned}\frac{\partial Var\{R\}}{\partial a_j} &= 2\sum_{i=1}^{n} a_i Cov\{u_i u_j\} + 2\sum_{i=1}^{m} b_i Cov\{v_i u_j\} \\ &+ 2\sum_{i=1}^{p} c_i Cov\{x_i u_j\} - 2Cov\{u_0 u_j\} + 2\mu_1 \qquad for \quad j = 1, \dots, n\end{aligned} \tag{7}$$

$$\begin{aligned}\frac{\partial Var\{R\}}{\partial b_j} &= 2\sum_{i=1}^{m} b_i Cov\{v_i v_j\} + 2\sum_{i=1}^{n} a_i Cov\{u_i v_j\} \\ &+ 2\sum_{i=1}^{p} c_i Cov\{x_i v_j\} - 2Cov\{u_0 v_j\} + 2\mu_2 \qquad for \quad j = 1, \dots, m\end{aligned} \tag{8}$$

$$\begin{aligned}\frac{\partial Var\{R\}}{\partial c_j} &= 2\sum_{i=1}^{p} b_i Cov\{x_i x_j\} + 2\sum_{i=1}^{n} a_i Cov\{u_i x_j\} \\ &+ 2\sum_{i=1}^{m} b_i Cov\{v_i x_j\} - 2Cov\{u_0 x_j\} + 2\mu_3 \qquad for \quad j = 1, \dots, p\end{aligned} \tag{9}$$

$$\frac{\partial Var\{R\}}{\partial \mu_1} = 2\sum_{i=1}^{n} a_i - 1 \tag{10}$$

$$\frac{\partial Var\{R\}}{\partial \mu_2} = 2\sum_{i=1}^{m} b_i \tag{11}$$

$$\frac{\partial Var\{R\}}{\partial \mu_3} = 2\sum_{i=1}^{p} c_i \tag{12}$$

Recording equation 7-12, we get the final cokriging system,

$$\begin{aligned}\sum_{i=1}^{n} a_i Cov\{u_i u_j\} + \sum_{i=1}^{m} b_i Cov\{v_i u_j\} + \sum_{i=1}^{p} c_i Cov\{x_i u_j\} + \mu_1 = Cov\{u_0 v_j\} \\ for \quad j = 1, \dots, n\end{aligned} \tag{13}$$

$$\sum_{i=1}^{n} a_i Cov\{u_i v_j\} + \sum_{i=1}^{m} b_i Cov\{v_i u_j\} + \sum_{i=1}^{p} c_i Cov\{x_i v_j\} + \mu_2 = Cov\{u_0 v_j\} \quad for \quad j = 1, \ldots, m \tag{14}$$

$$\sum_{i=1}^{n} a_i Cov\{u_i x_j\} + \sum_{i=1}^{m} b_i Cov\{v_i u_j\} + \sum_{i=1}^{p} c_i Cov\{x_i x_j\} + \mu_3 = Cov\{u_0 x_j\} \quad for \quad j = 1, \ldots, p \tag{15}$$

$$\sum_{i=1}^{n} a_i = 1, \tag{16}$$

$$\sum_{i=1}^{n} b_i = 0, \tag{17}$$

$$\sum_{i=1}^{n} c_i = 0 \tag{18}$$

Note that $Cov\{U_i V_j\} = Cov\{V_i U_j\}$, $Cov\{U_i X_j\} = Cov\{X_i U_j\}$ and $Cov\{V_i X_j\} = Cov\{X_i V_j\}$

We write the matrix form of those equations (from Eq. 13 to 18) as,

$$\begin{pmatrix} C_{uu} & C_{vu} & C_{xu} & 1 & 0 & 0 \\ C_{uv} & C_{vv} & C_{xu} & 0 & 1 & 0 \\ C_{ux} & C_{vx} & C_{xx} & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{a} \\ \mathbf{b} \\ \mathbf{c} \\ \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix} = \begin{pmatrix} C_{u_0 u} \\ C_{u_0 v} \\ C_{u_0 x} \\ 1 \\ 0 \\ 0 \end{pmatrix} \tag{19}$$

where C_{uu} is the auto-covariance of primary variable, C_{vv} is the auto-covariance of first secondary variable, and C_{xx} is the auto-covariance of the second secondary variable. C_{uv} is the cross-covariance between primary and first secondary variables, C_{ux} is the cross-covariance between primary and second secondary variables, C_{xv} is the cross-covariance of two secondary variables, μ_1, μ_2, and μ_3 are the Lagrange multipliers and a, b, and c are weight vectors of primary, first secondary, and second secondary variables to be determined. Note that $C_{uv} = C_{vu}$, $C_{ux} = C_{xu}$, and $C_{xv} = C_{vx}$.

In next section, matrix equation (19) is considered as a new cokriging system, which involves two seismic attributes to be combined with well log data. This system will be implemented for the porosity prediction using a real seismic data from Alberta.

CASE STUDY

This case study predicts porosity using the new cokriging estimation system described in the previous section and compares the result with maps generated by kriging and traditional cokriging. The procedure for implementing the new cokriging porosity prediction is as follows: (1) Prepare input data. (2) Calculate variograms. (3) Perform cokriging (4) Compare the performance. (5) Apply cross-validation.

Input data

The survey data was recorded over the Blackfoot field located in southern Alberta in 1995 for PanCanadian Petroleum. There are twelve wells involved in this study area, all of which contain calculated porosity logs. The porosity is treated as the primary variable, which is computed using an average value between the picked top and base of the zone of interest in each well. Figure 1 shows the well locations in the survey area and the porosity value at each location.

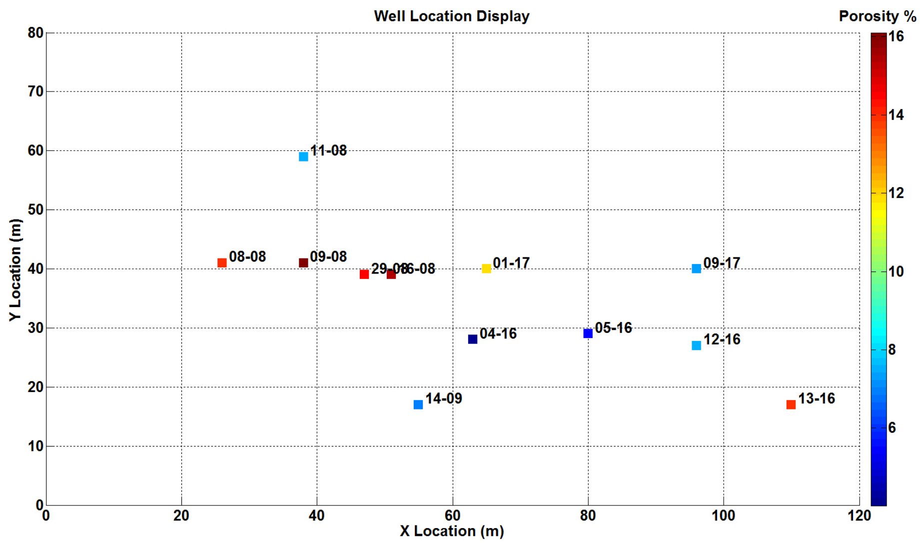


FIG. 1: Well location display with porosity value

The two secondary datasets consist of two structure slices extracted from the acoustic impedance inversion of the initial seismic volume and the stacked P-wave seismic data. To obtain the inversion volume, we build an initial model from the well logs and pick horizon on the seismic section and stop perturb this model until the synthetic seismogram for each trace in the volume has a best least-squares match with the original data. Figure 2a shows crossline 18 from the seismic volume, showing correlated sonic logs from two intersecting wells, 14-09 and 13-16, and the picked channel top. Figure 2b shows crossline 18 from the inverted volume. The color key indicates impedance.

The horizon slice of the P-wave impedance inversion was computed by using an arithmetic average over a 10 ms window below the picked channel top from the 3D inverted volume. Similarly, we extracted three data slices, seismic amplitude, amplitude envelope, and instantaneous phase, by calculating a 10ms RMS average over the zone of interest. The cokriging estimation system requires a strong correlation between the primary and secondary variables. Thus, we calculated correlation coefficients between the porosity values and all four data slices. The best two correlation coefficients are calculated from the inversion slice and seismic amplitude slice, which are -0.65 and 0.41, respectively. Thus, we use the inversion slice (Figure 3) and seismic amplitude slice (Figure 4) as the two secondary inputs.

Variograms and Covariance

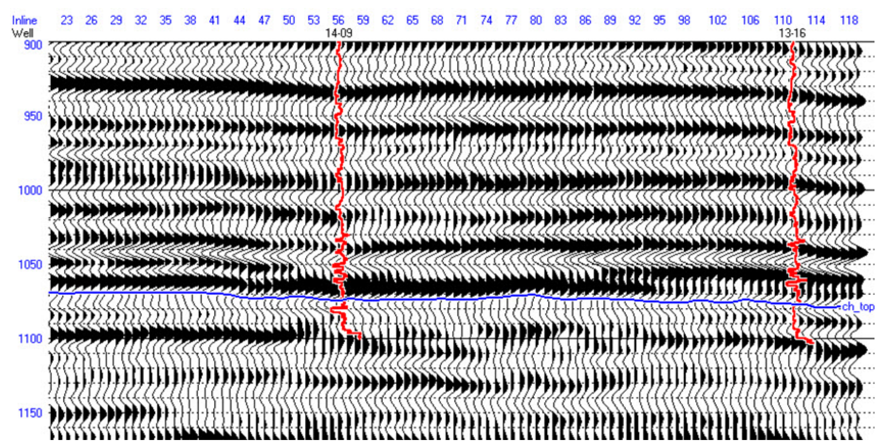
A variogram is a concise way to describe the degree of spatial dependence between the input data and is calculated by,

$$\gamma_{uv}(h) = \frac{1}{2N(h)} \sum_{(i,j)|h_{ij}=h} (u_i - u_j)(v_i - v_j) \tag{20}$$

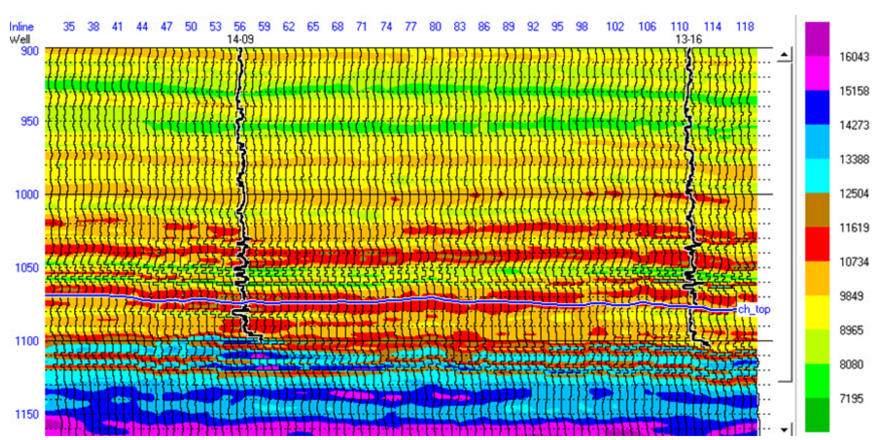
where h is the lag distance. $N(h)$ is the number of data pairs whose locations are separated by h. $\gamma_{uv}(h)$ is the cross variogram related of lag h. u and v are input data. The spatial interpolation is based on the principle that close samples tend to be more similar than distant samples.

Figure 5 shows 6 variograms obtained from the well log values and two seismic attributes. Figure 5a, 5e, and 5f are auto-variograms of well log data, inversion, and seismic amplitude, respectively. The cross-variogram between the well log values and inversion attribute is shown in Figure 5b, well to seismic amplitude cross-variogram is shown in Figure 5c, and Figure 5d is the cross-variogram of the inversion and amplitude attributes.

A spherical model was chosen to fit the variogram as shown in Figure 5. Then, covariance



(a) The final CDP stack



(b) P-wave impedance horizon slice

FIG. 2: Crossline 18 from the 3-D seismic volume

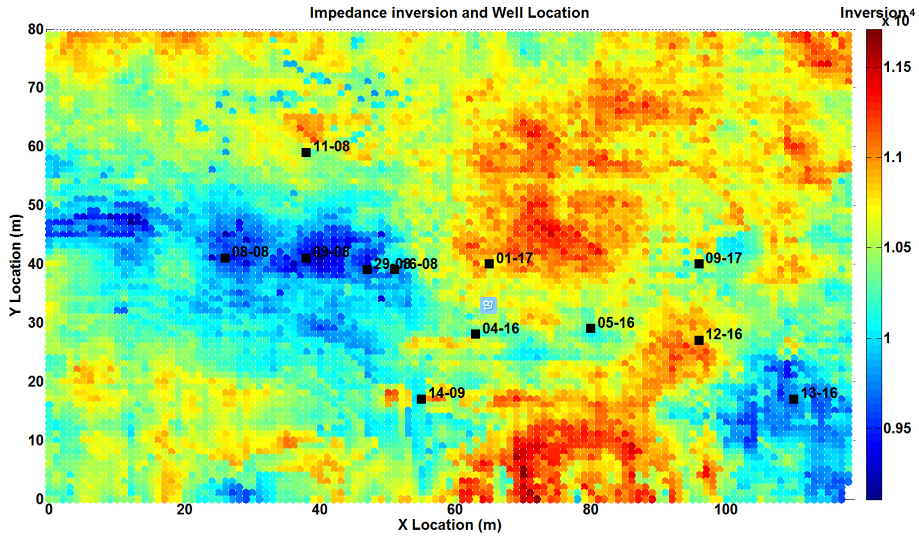


FIG. 3: Inversion slice

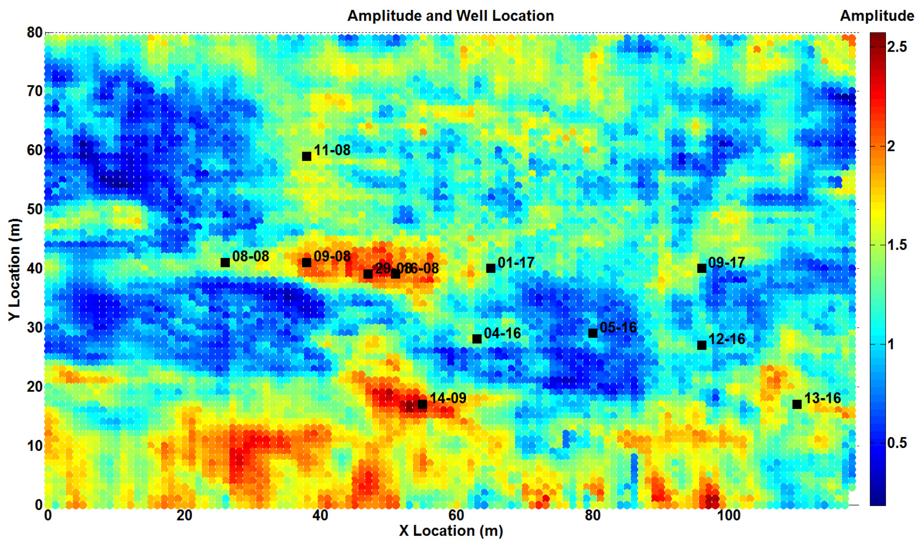


FIG. 4: Seismic amplitude slice

model is converted by following equation,

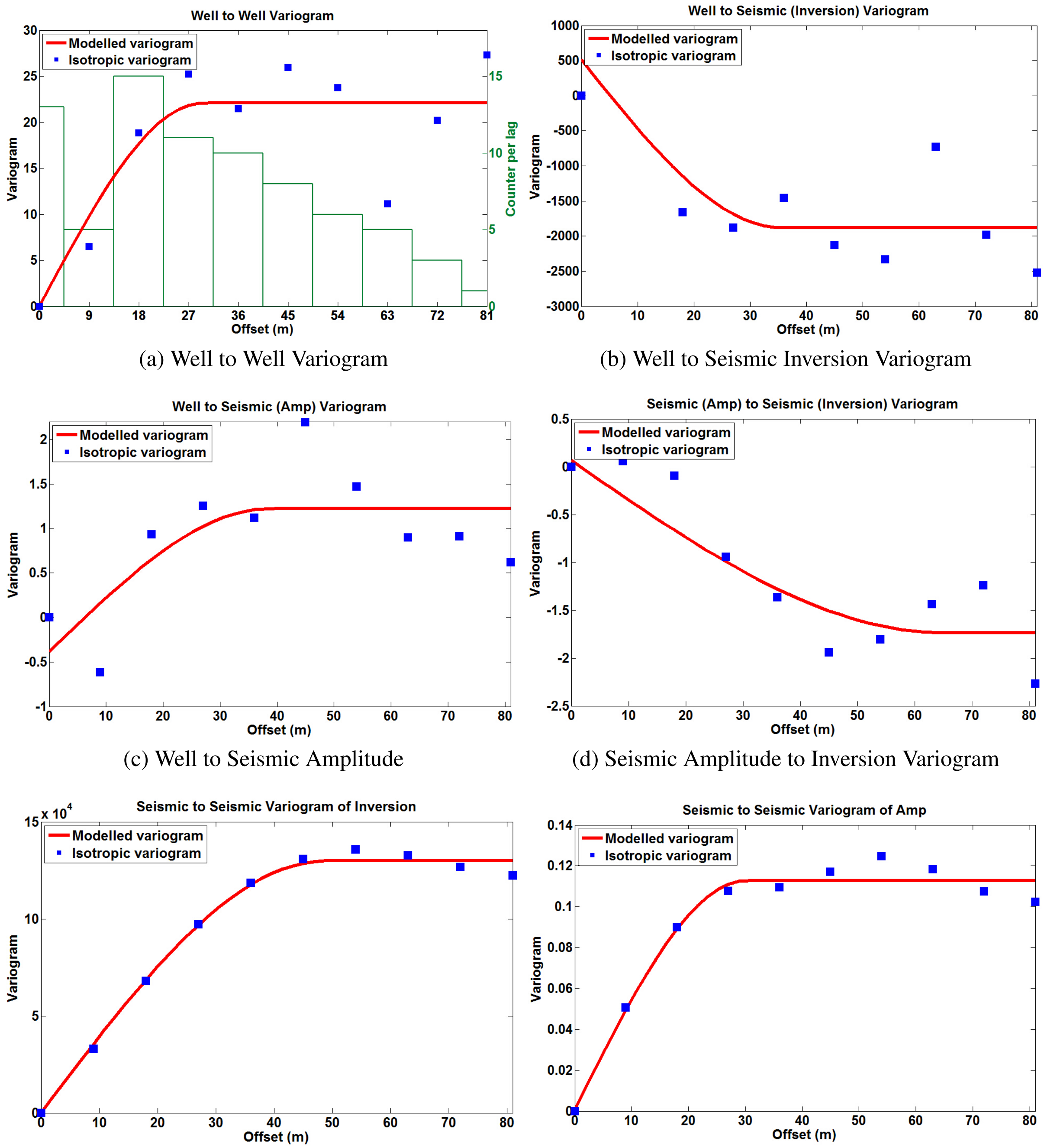
$$\mathbf{Cov}(h) = \gamma(\infty) - \gamma(h) \tag{21}$$

where $\mathbf{Cov}(h)$, the covariance model, is a function of lag h. $\gamma(h)$ is the variogram value of lag h. $\gamma(\infty)$ is the variogram value for very large distances, commonly called the $sill$.

The porosity predicted map is able to be generated after determining the weights (a, b, c) in matrix equation (19) involving the covariance.

Map Results

To evaluate the predicted result under the new cokriging system, the estimates from the kriging and the traditional cokriging were calculated and compared. The kriging interpolation



(a) Well to Well Variogram

(b) Well to Seismic Inversion Variogram

(c) Well to Seismic Amplitude

(d) Seismic Amplitude to Inversion Variogram

(e) Seismic to Seismic Inversion

(f) Seismic to Seismic Amplitude

FIG. 5: Variograms

(Figure 6) looks like a filter away from the data points. Figure 7 shows the result generated by traditional cokriging with only the impedance inversion utilized and Figure 8 shows the estimation with only the amplitude data slice as the secondary input. The final produced porosity map (Figure 9) was constructed by implementing equation (19), and including both impedance inversion and seismic amplitude attributes.

All of the cokriging estimates have more geological information than prediction by kriging

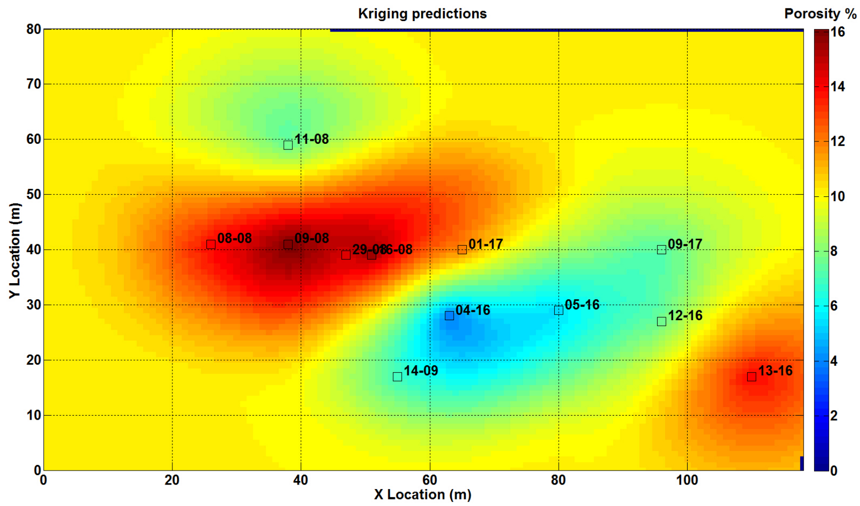


FIG. 6: Kriging interpolation with well log data

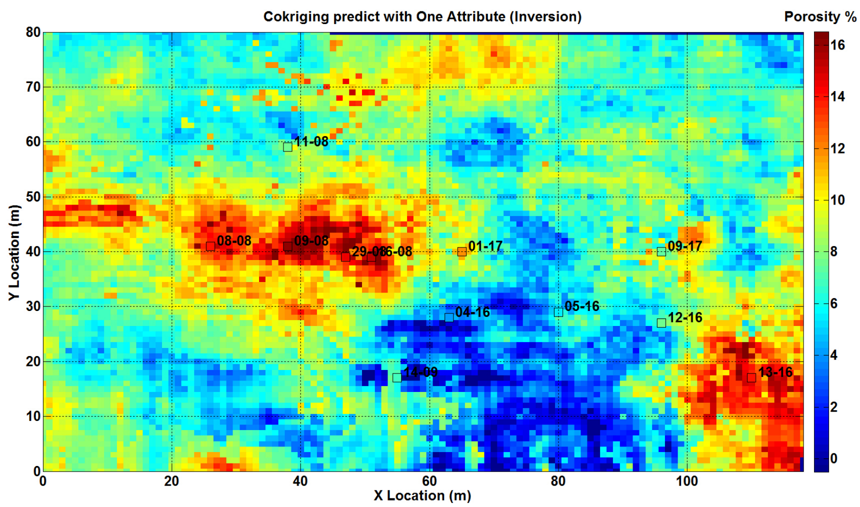


FIG. 7: Traditional cokriging prediction with inversion

alone because kriging uses only well log data, which is poorly sampled spatially. Compared to the traditional cokriging result, there is no significant difference in using two attributes where there is good well distribution. However, the results using the new cokriging approach show higher lateral resolution and a remarkable difference in those areas where there is little well control. To make the analysis be more quantitative, "leave-one-out" cross-validation was employed to calculate RMS errors for kriging, the traditional cokriging, and the new corkiging system.

Cross-validation

Cross-validation is used to validate the accuracy of an interpolation (Voltz and Webster, 1990). "Leave-one-out" cross-validation calculates the difference between the predicted and actual values by removing one well log at a time and computing the root-mean-square error of kriging with the other wells. The average error of leaving each well out is then computed, and is expressed as

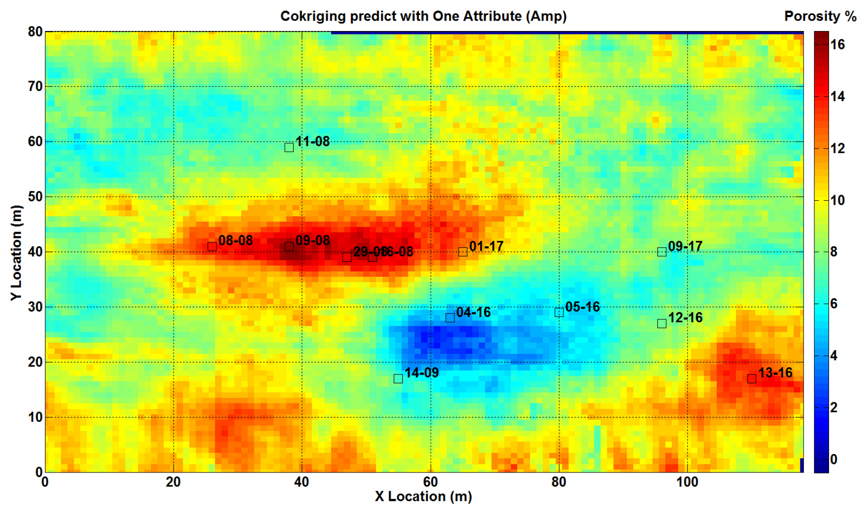


FIG. 8: Traditional cokriging prediction with seismic amplitude

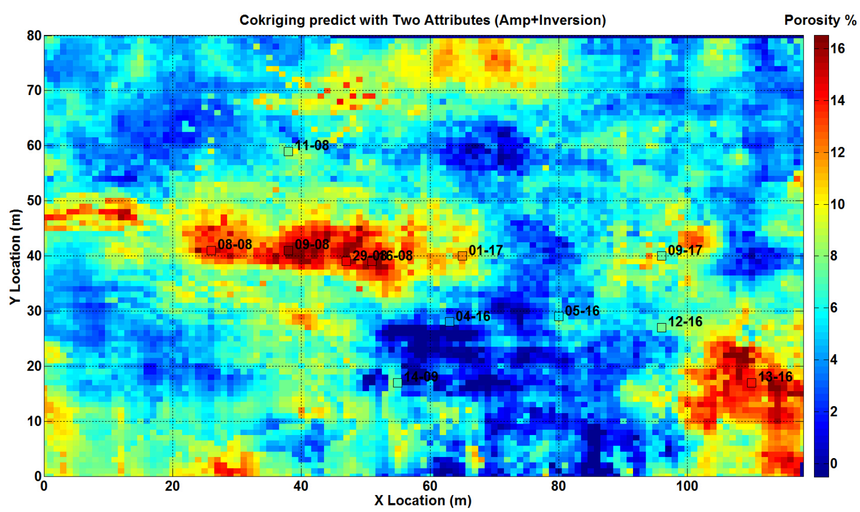


FIG. 9: Cokriging predict with two secondary data (inversion and seismic amplitude)

$$\mathbf{E_{RMS}} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left\{z(x_i) - \hat{z}(x_i)\right\}^2} \tag{22}$$

where $z(x_i)$ is the actual value and $\hat{z}(x_i)$ is the estimated value by leaving one out. N is the numbers of "leave-one-out" calculations implemented, which corresponds to the number of well values in the primary dataset.

RMS errors of kriging, traditional cokriging with inversion, traditional cokriging with seismic amplitude, and the new cokriging system are given by the histograms shown in Figure 10. It is worth noting that the new approach we presented shows a lower RMS error than other approaches. In other words, the new cokriging system, involving two well correlated secondary datasets, gives a better estimation of the porosity.

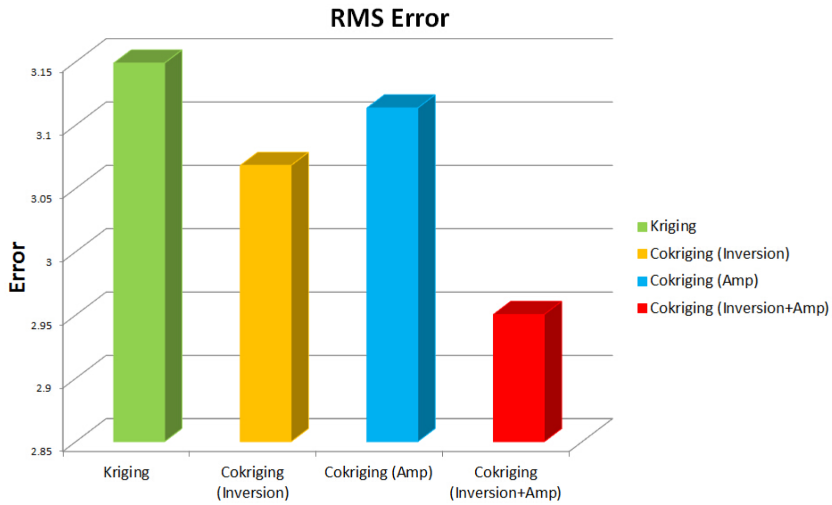


FIG. 10: Leave-one-out Cross-validation

CONCLUSION

In this paper, we have derived and presented a new cokriging estimation system with one primary and two secondary variables, which is designed to bring extra geological information into the estimation process. The case study shows that the new approach is able to improve the spatial lateral resolution at locations away from the well values when compared with the traditional cokriging estimation system.

The "leave-one-out" cross-validation method was applied to validate the accuracy of the new cokriging results. The new cokriging system gives a lower RMS error than the RMS errors of kriging and traditional cokriging. This is due to the additional attribute which was added in the implementation. Furthermore, the new cokriging system offers us a new way to include more than one seismic attribute into the estimation of porosity with cokriging, and could be extended to three or four variables. Finally, it can be concluded that the new cokriging estimation system with one primary and two secondary variables is a step forward for producing improved map estimates from well log data and seismic results.

ACKNOWLEDGMENTS

We would like to thank Jian Sun and Tiansheng Chan for their suggestions, as well as our colleagues at CREWES and Hampson-Russell Software for the support.

REFERENCES

Babak, O., and Deutsch, C., 1992, Collocated cokriging based on merged secondary attributes: Math Geosci, **41**, 921–926.

Doyen, P. M., 1988, Porosity from seismic data -a geostatistical approach: Geophysics, **53**, 1263–1257.

Doyen, P. M., den Boer, L. D., and Pillet, W. R., 1996, Seismic porosity mapping in the ekofisk field using a new form of collocated cokriging: SPE Annual Technical Conference and Exhibition, Denver, Coloradoi.

Guerrero, J., Vargas, C., and Montes, L., 1996, Reservoir characterization by multiattribute analysis: The orito field case: Earth Sci. Res. J., **14**, 173–180.

Isaaks, E. H., and Srivastava, R. M., 1989, An Introduction to Applied Geostatistics: Oxford University Press, New York.

Ito, K., and Kunisch, K., 2008, Lagrange Multiplier Approach to Variational Problems and Applications:

Matheron, G., 1965, Porosity from seismic data -a geostatistical approach.

Russell, B., Hampson, D., Todorov, T., and Lines, L., 2002, Combining geostatistics and multi-attribute transforms:a channel sand case study, black foot oilfield: Journal of Petroleum Geology, **21**, 97–117.

Voltz, M., and Webster, R., 1990, A comparison of kriging, cubic splines and classification forpredicting soil properties from sample information: J. Soil Sci, **41**, 473–490.