Mathematical Geology, Vol. 30, No. 7, 1998

Deriving Constraints on Small-Scale Variograms due to Variograms of Large-Scale Data[1]

Hans Kupfersberger,[2,3] Clayton V. Deutsch,[2,4] and André G. Journel[2]

The application of kriging-based geostatistical algorithms to integrate large-scale seismic data calls for direct and cross variograms of the seismic variable and primary variable (e.g., porosity) at the modeling scale, which is typically much smaller than the seismic data resolution. In order to ensure positive definiteness of the cokriging matrix, a licit small-scale coregionalization model has to be built. Since there are no small-scale secondary data, an analytical method is presented to infer small-scale seismic variograms. The method is applied to estimate the 3-D porosity distribution of a West Texas oil field given seismic data and porosity data at 62 wells.

KEY WORDS: coregionalization model, variogram inference, cosimulation.

INTRODUCTION

Consider the distribution over a 3-D field A of a *primary* attribute of interest $z(\mathbf{u})$, $\mathbf{u} \in A$. A *secondary* attribute y, related to the primary variable, is available at a large scale denoted V. In the context of petroleum geostatistics, z is typically a petrophysical property or a lithofacies indicator variable and y is a seismic attribute correlated to volume averages of the primary z variable. As illustrated in Figure 1, the horizontal scale of the seismic data is comparable to the scale of fluid flow modeling; however, the seismic data has much less vertical resolution than the well data and *primary* z variable. The scale of the z variable will be considered the modeling scale.

Most approaches for integrating large-scale soft data in this context can be classified into three general categories. The first approach consists of assuming that the soft y data provides only information on large-scale trends of the variable

[1] Received 30 June 1997; accepted 29 October 1997.

[2] Stanford Center for Reservoir Forecasting, Department of Petroleum Engineering, Stanford University, Stanford, California 94305-2220.

[3] Present address: Joanneum Research Forschungsgesellschaft mbH, Institute of Hydrogeology and Geothermics, A-8010 Graz, Austria. e-mail: hans.kupfersberger@joanneum.ac.at

[4] Present address: Department of Civil and Environmental Engineering, University of Alberta, Edmonton, Alberta T6G 1T2, Canada.

0882-8121/98/1000-0837$15.00/1

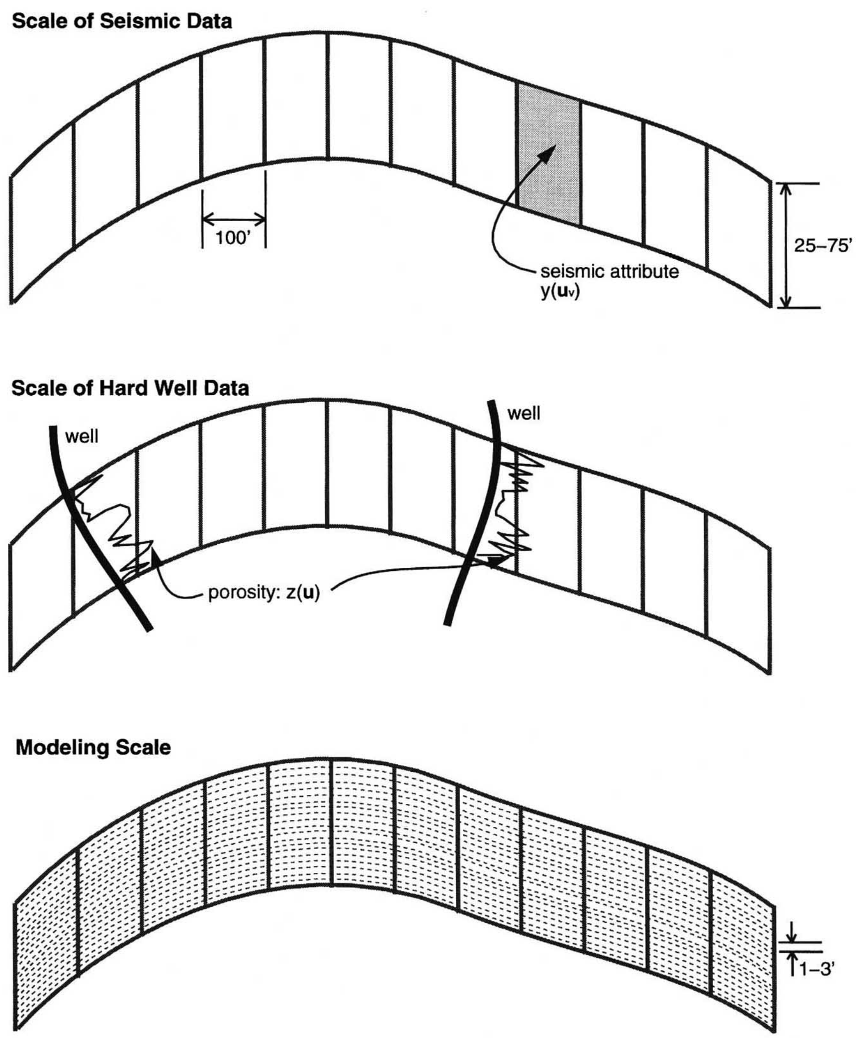


Figure 1. A schematic illustration of the available large-scale seismic data and small-scale well log or core derived porosity. The vertical modeling scale is close to the resolution of the well data.

z. This trend information is accounted for by constraints added to the kriging equations, e.g., the external drift or locally varying mean algorithms (Marechal, 1984; Deutsch and Journel, 1992). A shortcoming of this approach is that the information extracted from the soft y data may not all relate to trends of the z variable. The second approach amounts to replicating the soft y data for all

3-D locations within the vicinity of the soft data and treat it as co-located with the locations of the small-scale z modeling cells. Co-located cokriging (Almeida, 1993) or the Markov–Bayes formalism (Zhu, 1991) are two methods that require the soft y data at the same scale as the hard z data. The significant shortcoming of this approach is the difficulty of inferring the required small-scale (co-located) calibration statistics; the only statistics usually available are large-scale. Finally, the third class of methods involves the use of iterative simulated annealing-type methods (Deutsch and Cockerham, 1994; Deutsch, Srinivasan, and Mo, 1996). These methods directly account for the scale and precision of the soft data but can be CPU-demanding and require delicate adjustment of a number of tuning parameters.

In our proposed method, we assume the multi-Gaussian framework for both cokriging and sequential simulation. Myers (1984) discusses the benefits of using the cokriging method. Doyen (1988) applied cokriging to estimate porosity from seismic data. Aboufirassi and Mariño (1984) and Ahmed, de Marsily, and Talbot (1988) used cokriging to estimate transmissivities with measurements of specific capacity and electrical aquifer properties, respectively.

Cokriging with large-scale soft data is not often used because of the difficulty of obtaining a licit model of coregionalization. The required 3-D variogram of the primary z variable at the modeling scale may be inferred from well data; however, the required large-scale Z-Y variogram is much more difficult to infer from the available data. The cross-variogram between a z modeling cell and y large-scale data depends on the vertical position of the cell within the co-located y large-scale data (Fig. 2).

The various cross variogram values depicted in Figure 2 could be calculated experimentally. However, the z-variable at each of the vertical "slices" must be considered as a different variable in fitting a licit model of coregionalization, e.g., given 25 z-slices within a y-block, a coregionalization model with 26 variables would need to be fitted.

All of the required variogram values could be obtained by averaging the variogram values making up a licit small-scale model of coregionalization. There are, however, no small-scale y data. This is the central problem tackled by this paper: inference of small-scale variograms with a combination of small-scale hard z data and large-scale y data.

COKRIGING WITH LARGE-SCALE SOFT DATA

The central feature of the sequential simulation algorithm is the calculation of a conditional distribution of the primary variable given surrounding original primary data, values at previously simulated modeling cells, *and* relevant secondary data. The mean and variance of these conditional distributions could be

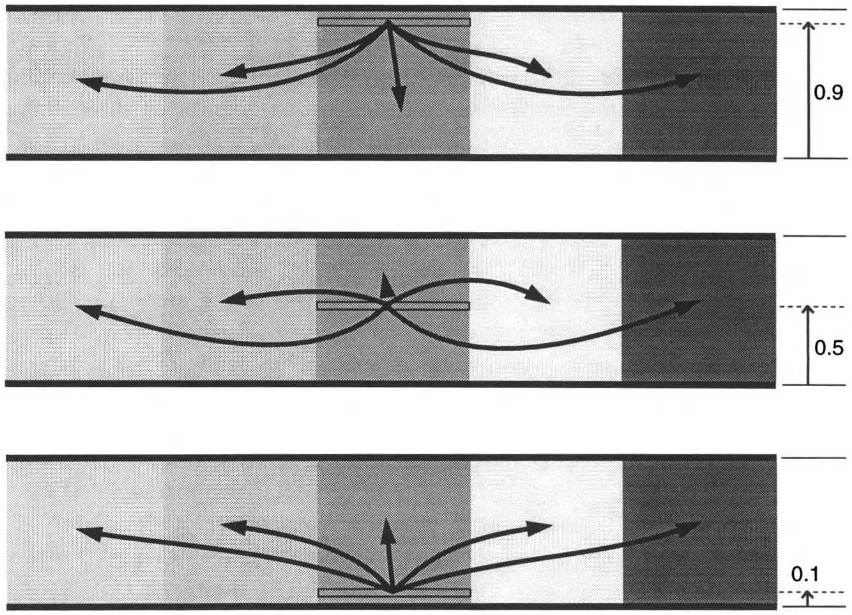


Figure 2. An illustration of the required covariance values: (top) the cell is near the top of the layer, (middle) the cell is near the middle of the layer, and (bottom) the cell is near the bottom of the layer.

established by simple cokriging. The cokriging estimator is written:

$$z^*(\mathbf{u}) - m_z = \sum_{\alpha_1=1}^{n_1(\mathbf{u})} \lambda_{\alpha_1}(\mathbf{u})[z(\mathbf{u}_{\alpha_1}) - m_z] + \sum_{\alpha_2=1}^{n_2(\mathbf{u})} \lambda'_{\alpha_2}(\mathbf{u})[y(\mathbf{u}'_{\alpha_2}) - m_y] \quad (1)$$

where m_z and m_y are the stationary primary and secondary mean values, $n_1(\mathbf{u})$ and $n_2(\mathbf{u})$ are the numbers of primary and secondary data for estimation at $\mathbf{u}$. Weights are calculated by the simple cokriging system of equations:

$$\begin{cases} \sum_{\beta_1=1}^{n_1(\mathbf{u})} \lambda_{\beta_1}(\mathbf{u}) C_Z(\mathbf{u}_{\alpha_1} - \mathbf{u}_{\beta_1}) + \sum_{\beta_2=1}^{n_2(\mathbf{u})} \lambda'_{\beta_2}(\mathbf{u}) C_{ZY}(\mathbf{u}_{\alpha_1} - \mathbf{u}'_{\beta_2}) \\ \quad = C_Z(\mathbf{u}_{\alpha_1} - \mathbf{u}), \quad \alpha_1 = 1, \ldots, n_1(\mathbf{u}) \\ \sum_{\beta_1=1}^{n_1(\mathbf{u})} \lambda_{\beta_1}(\mathbf{u}) C_{YZ}(\mathbf{u}'_{\alpha_2} - \mathbf{u}_{\beta_1}) + \sum_{\beta_2=1}^{n_2(\mathbf{u})} \lambda'_{\beta_2}(\mathbf{u}) C_Y(\mathbf{u}'_{\alpha_2} - \mathbf{u}'_{\beta_2}) \\ \quad = C_{YZ}(\mathbf{u}'_{\alpha_2} - \mathbf{u}), \quad \alpha_2 = 1, \ldots, n_2(\mathbf{u}) \end{cases} \quad (2)$$

Ordinary cokriging would call for similar variogram/covariance input (Deutsch and Journel, 1992; Goovaerts, 1997).

When dealing with large-scale information, the y-variable at location $\mathbf{u}'$ may relate to an average of the primary z- variable over a volume $v(\mathbf{u}')$ centered at $\mathbf{u}'$:

$$y(\mathbf{u}') = \frac{1}{|v|} \int_{v(\mathbf{u}')} z(\mathbf{u})\, d\mathbf{u} \tag{3}$$

In such cases, the y-covariance is actually an average of the z-covariance, that average calling for small-scale z-covariance values within the average volumes v:

$$\begin{aligned} C_y(\mathbf{u}'_\alpha, \mathbf{u}'_\beta) &= \overline{C}_z(v(\mathbf{u}'_\alpha, \mathbf{u}'_\beta)) \\ &= \frac{1}{|v|^2} \int_{v(\mathbf{u}'_\alpha)} d\mathbf{u} \int_{v(\mathbf{u}'_\beta)} C_z(\mathbf{u} - \mathbf{u}')\, \mathbf{du}' = C_v(\mathbf{u}'_\beta - \mathbf{u}'_\alpha) \end{aligned} \tag{4}$$

In order to ensure positive definiteness of the cokriging matrix, the linear model of coregionalization has to be built up from the small-scale C_z-covariance. Based on the development by Journel and Huijbregts (1978, p. 77 ff) a volume V-averaged covariance $C_V(h)$ can be written as

$$C_V(h) = C_{0,V} + C_V \cdot \rho(\mathbf{h}; \mathbf{a}_V) \tag{5}$$

where $C_{0,V}$ is the nugget term and $\rho(\mathbf{h}; \mathbf{a}_V)$ is a basic covariance structure (e.g., spherical or Gaussian) with range a_V and sill 1. In presence of anisotropy $\mathbf{a}_V$ would be a vector. The nugget effect being inversely proportional to the volume of averaging, the small-scale nugget can be computed as

$$C_{0,v} = C_{0,V} \cdot \frac{|V|}{|v|} \tag{6}$$

V denotes the volume associated with the large-scale datum and v represents the volume of the small-scale measurement. Similarly, the small-scale range can be obtained as

$$a_v = a_V - (|V| - |v|) \tag{7}$$

with a being much smaller than v and V. The small-scale sill is calculated as

$$C_v = C_V \cdot \frac{\overline{\rho}(\mathbf{v}, \mathbf{v}; \mathbf{a})}{\overline{\rho}(V, V; a)} \tag{8}$$

where, similarly to relation (3), the $\overline{\rho}$-terms are volume-averaged of the small-scale (quasi-point support) correlogram $\rho(\mathbf{h}; \mathbf{a})$ with range a:

$$\overline{\rho}(V, V; a) = \frac{1}{|V|^2} \int_V du \int_{\mathbf{V}} \rho(\mathbf{u}' - \mathbf{u})\, \mathbf{du}' \tag{9}$$

In practice, one would infer from large-scale (V) data the covariance model $C_V(\mathbf{h})$ as defined in relation (5), then deduce from relations (6), (7), and (8), the nugget, range, and sill parameters of the small-scale covariance model $C_v(\mathbf{h})$. Assuming for $C_v(\mathbf{h})$ the same basic structure $\rho(\cdot)$ used for $C_V(\mathbf{h})$, the small-scale covariance model $C_v(\mathbf{h})$ is then given by expression (4) with the "down-scaled" parameters $C_{0,v}$, a_v, and C_v. Note that $C_v \cdot \bar{\rho}(v, v; a)$ is the dispersion (spatial) variance of quasi-point values within a volume of dimension v (Journel and Huijbregts, 1978, p. 61 ff).

The same methodology is applied to deduce the small-scale zy cross-covariance from a prior large-scale (V) model.

The key assumption underlying the process of regularization is that all averagings are linear. Consequently, within a Gaussian framework for sequential simulation, we assume that the fictitious small-scale normal-score y-variable averages linearly.

Within the analytical model of small-scale covariance inference, two additional assumptions are involved. First, it is assumed that the same type of basic covariance structure that is used to model the large-scale data is also appropriate at the small scale. Second, it is required that a range for the large-scale data can be computed in every direction. Given 2-D horizontal data, a vertical range can obviously not be determined. In this case, a vertical range value has to be arbitrarily chosen to complete the samll-scale covariance model. In practice, however, this vertical range is taken from the small scale z data so that the final linear model of coregionalization is licit.

CASE STUDY

The proposed method is demonstrated on a producing West Texas field to estimate thte porosity distribution which may be used for subsequent flow modeling. The dataset consists of porosity logs from 62 wells within an area of approximately 10 * 10 km. Figure 3 shows the well locations and the vertically averaged porosities. The vertical sampling interval of the porosity data is about 30 cm. In addition, about 4000 regularly spaced low frequency seismic data are available covering the entire model domain (Fig. 4). The depositional environment consists of very fine grained aeolian sands reworked by littoral and middle neritic processes. The productive interval is composed of a shallowing-upward, prograding carbonate shelf sequence. Lithologically, the formation primarily consists of alternating dolomites and siltstones. The siltstones generally lack crossbedding and present reasonably sharp transition to the overlying and underlying dolomites.

The variograms of porosity normal scores in two horizontal and the vertical directions are displayed in Figure 5. These variograms reveal a zonal anisotropy

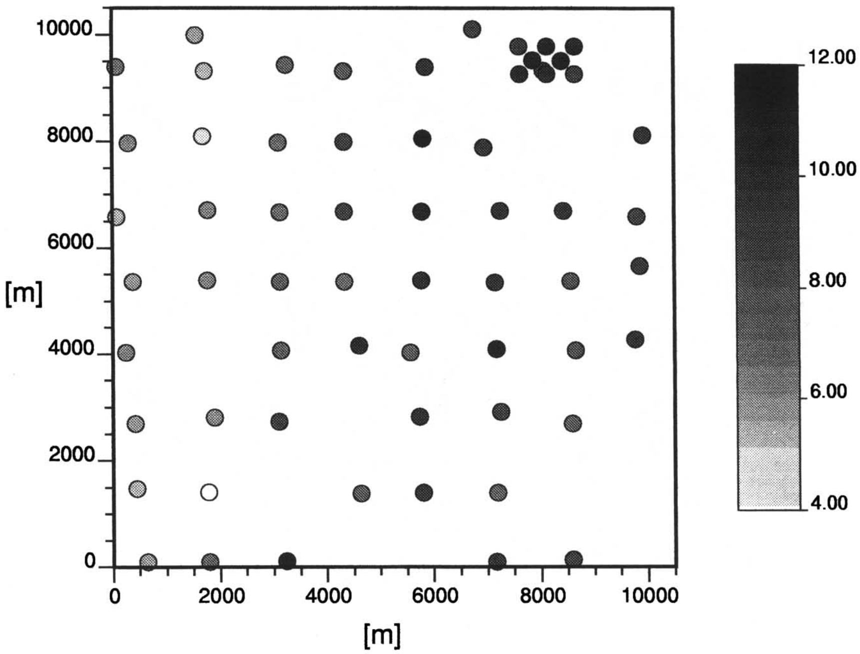


Figure 3. Location of the 62 wells and the vertical averaged porosities.

in the horizontal direction: the variograms have different ranges and one variogram does not reach the unit sill. Equation (10) represents the experimental variogram model fitted to the porosity normal scores. *Exp* and *Sph* denote an exponential and a spherical covariance structure, respectively. Figure 6 shows (a) the horizontal variograms of the normal scores of the low seismic data, and (b) the cross-variograms between the vertically averaged porosity and the seismic attribute (normal scores). The fitted sill values for each variogram component are computed as 0.29 and 0.71 for the (v) seismic normal scores and as 0.11 and 0.52 for the (v) cross-variogram.

$$\gamma_Z = 0.0 + 0.4 Exp\left(\sqrt{\left(\frac{h_x}{1000.0}\right)^2 + \left(\frac{h_y}{3000.0}\right)^2 + \left(\frac{h_z}{12.0}\right)^2}\right) + 0.6 Sph\left(\sqrt{\left(\frac{h_x}{25000.0}\right)^2 + \left(\frac{h_y}{5000.0}\right)^2 + \left(\frac{h_z}{50.0}\right)^2}\right) \tag{10}$$

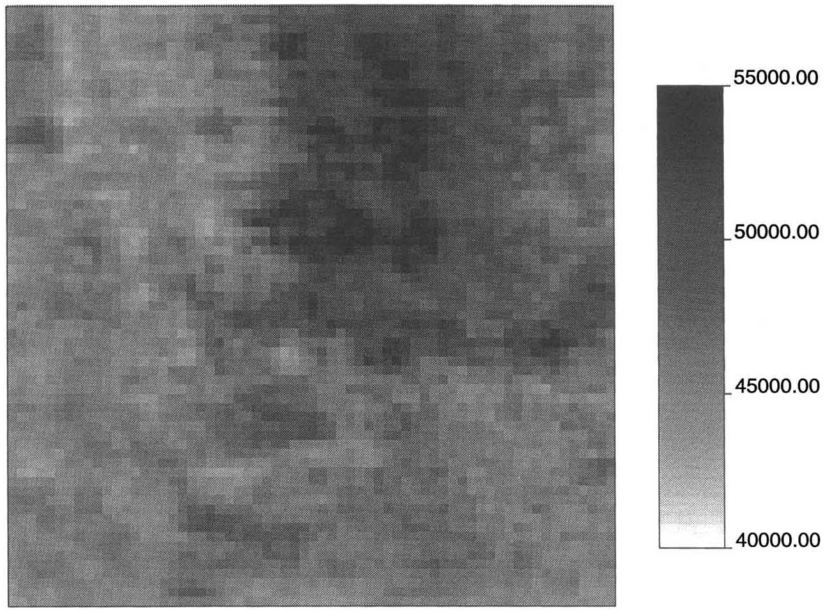


Figure 4. Horizontal distribution of seismic data.

The linear model of coregionalization requires that the nested structures (types, ranges), fitted to the small-scale primary data (10), be kept for the small-scale (v) variogram of the seismic data and for the small-scale (v) cross-variogram. Thus, only the nugget and the sill components of the seismic variogram and of the cross-variogram have to be determined by the analytical model.

In the West Texas field, the large scale (V) comprises 55 m in the vertical direction (average well length), and the small scale (v) represents 1 m in the vertical direction. From relation (8), the sill values for each variogram component are computed as 0.592 and 0.408 for the (v) seismic normal scores and as 0.27 and 0.33 for the (v) cross-variogram. The $\bar{\rho}$-terms in (8) for (v) and (V) are calculated from Equation (9) using the corresponding (V) variogram structure shown in Figure 6. In the case of the seismic normal scores, the $\bar{\rho}$-terms are calculated as 0.922 and 0.135 for the exponential and as 0.996 and 0.517 for the spherical variogram structure. For the cross-variogram, the corresponding $\bar{\rho}$-values are 0.922, 0.135, 0.996, and 0.517, respectively. The resulting (v) seismic variogram and cross-variogram is given in (11).

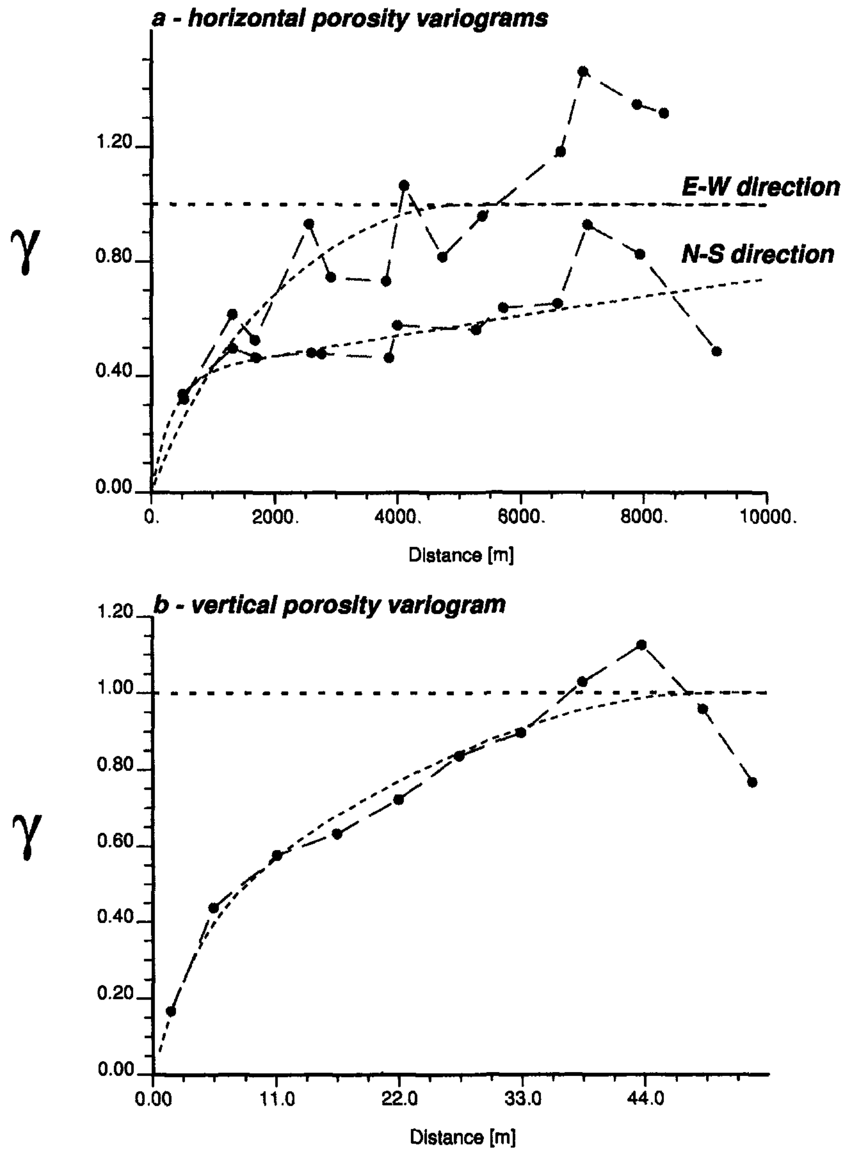


Figure 5. Horizontal (East–West and North–South) and vertical variograms of the porosity data after normal score transform.

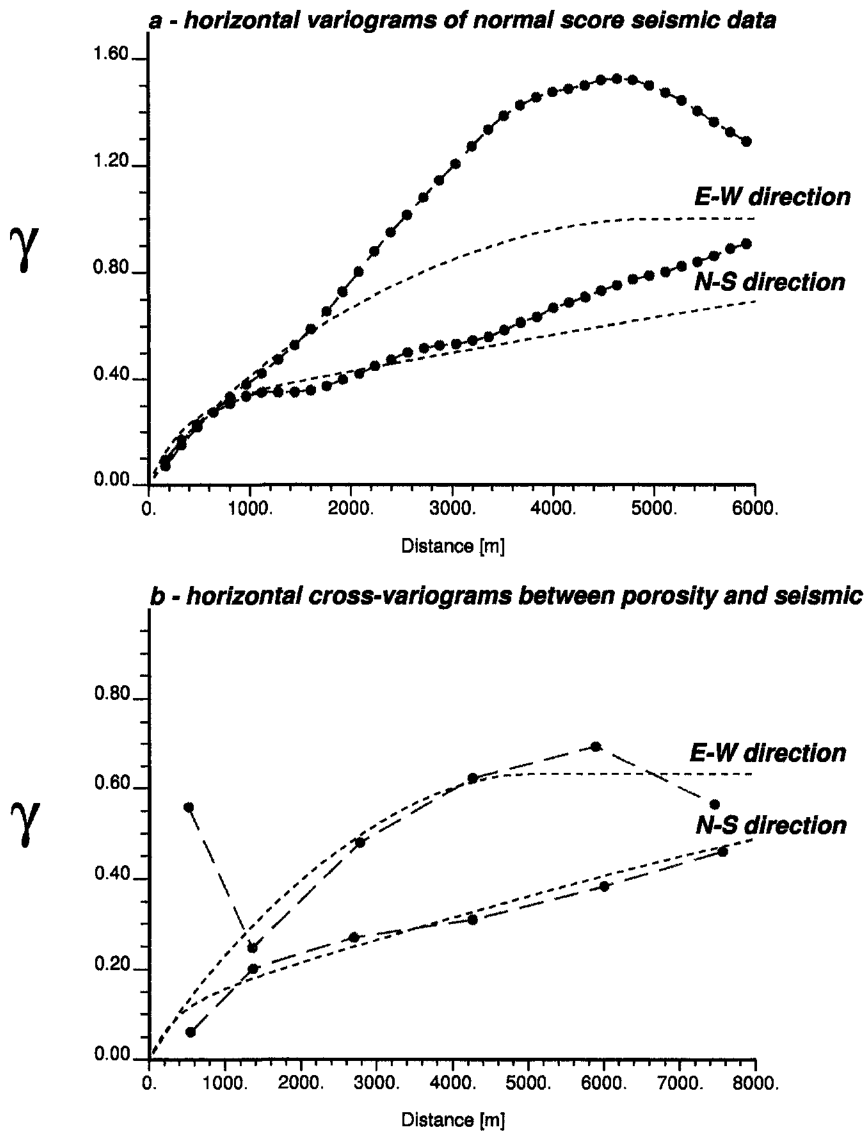


Figure 6. Horizontal (East-West and North-South) variograms of the seismic data after normal score transform (top); horizontal cross-variograms between the vertical averaged normal score porosities and the normal score seismic data (bottom).

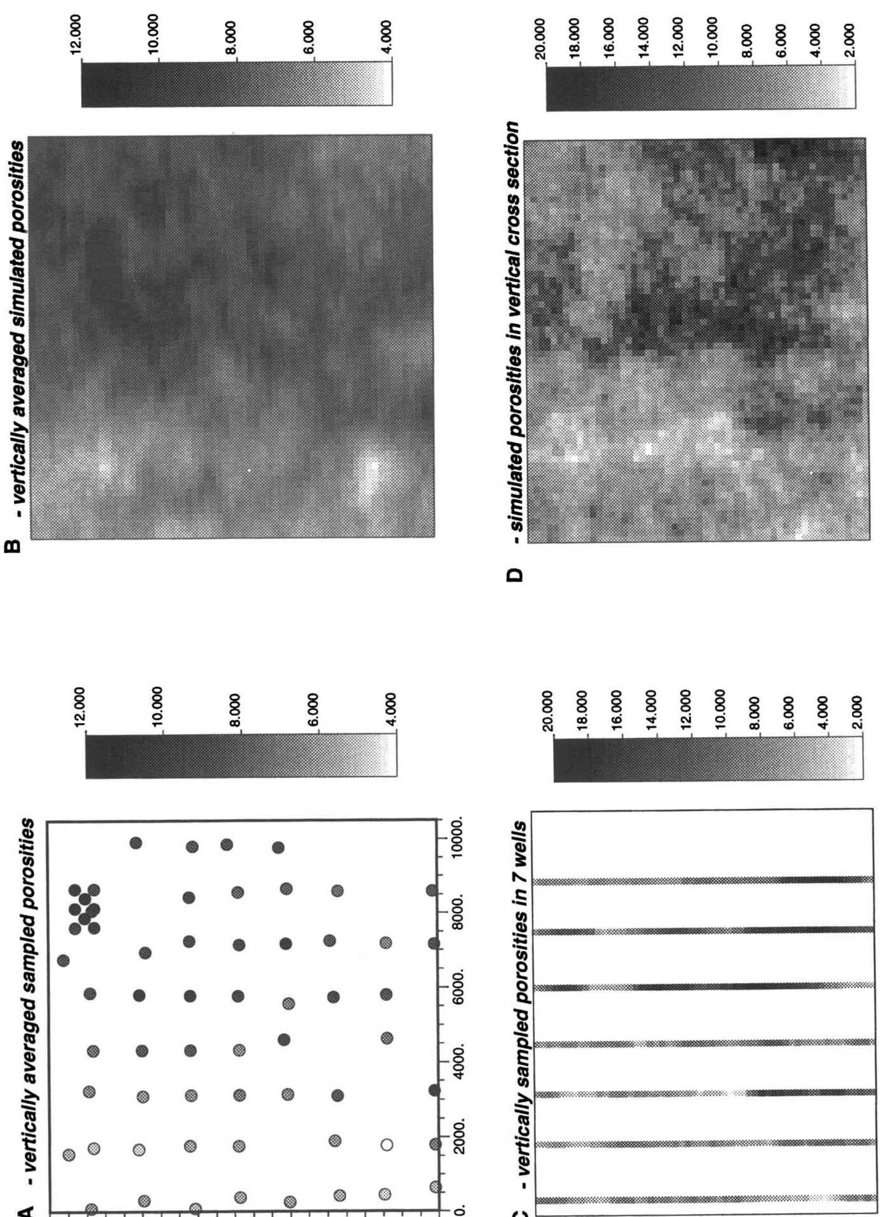
$$\gamma_Y = 0.0 + 0.592\text{Exp}\left(\sqrt{\left(\frac{h_x}{1000.0}\right)^2 + \left(\frac{h_y}{3000.0}\right)^2 + \left(\frac{h_z}{12.0}\right)^2}\right)$$
$$+ 0.408 Sph\left(\sqrt{\left(\frac{h_x}{25000.0}\right)^2 + \left(\frac{h_y}{5000.0}\right)^2 + \left(\frac{h_z}{50.0}\right)^2}\right)$$
$$\gamma_{ZY} = 0.0 + 0.27 Exp\left(\sqrt{\left(\frac{h_x}{1000.0}\right)^2 + \left(\frac{h_y}{3000.0}\right)^2 + \left(\frac{h_z}{12.0}\right)^2}\right)$$
$$+ 0.36 Sph\left(\sqrt{\left(\frac{h_x}{25000.0}\right)^2 + \left(\frac{h_y}{5000.0}\right)^2 + \left(\frac{h_z}{50.0}\right)^2}\right) \qquad (11)$$

The positive definiteness of the small-scale linear model of coregionalization created by combining relation (10) and (11) has to be checked. That is, for every variogram component, the product of the sill of the two autovariograms (γ_Z and γ_Y) has to exceed the squared sill of the cross-variogram. From (10) and (11), it can be seen that this condition is fulfilled for both variogram components.

Full cokriging was implemented as an option in the sequential Gaussian simulation program sgsim (Deutsch and Journel, 1992). Figure 7 shows a porosity realization using the developed small-scale linear model of coregionalization. It can be seen that the vertical averaged simulated porosities (Fig. 7B) reproduce very well the vertical averaged sampled porosities (Fig. 7A). Figure 7C displays seven sampled vertical porosity profiles, which are aligned along a cross-section. The simulated porosities at the same cross-section (Fig. 7D) resemble the high and low porosity features in the sample data. Figure 7E and F show the horizontal and vertical porosity variogram model reproduction with the dashed lines depicting the input variogram model and the dotted lines showing the variograms computed from the realization. Again, a very good fit between measured and simulated data can be observed indicating that the analytical approach to infer the small-scale variogram works very well.

The analytical approach is applied within a Gaussian framework. But no assumptions are made which limit the applicability to the Gaussian space. Rather, this is chosen because of convenience. If all necessary variogram parameters are inferred, the method can also be used with non-Gaussian simulation procedures. Implementation of full cokriging is selected to particularly take advantage of the lateral information provided by the large-scale seismic data.

The retained model of coregionalization is built on the assumption that the modeled parameter averages linearly after normal score transformation. In general, large-scale parameter averages must lie in between the arithmetic parameter


A - vertically averaged sampled porosities
B - vertically averaged simulated porosities
C - vertically sampled porosities in 7 wells
D - simulated porosities in vertical cross section
12.000
10.000
8.000
6.000
4.000
20.000
18.000
16.000
14.000
12.000
10.000
8.000
6.000
4.000
2.000
10000.
8000.
6000.
4000.
2000.
0.

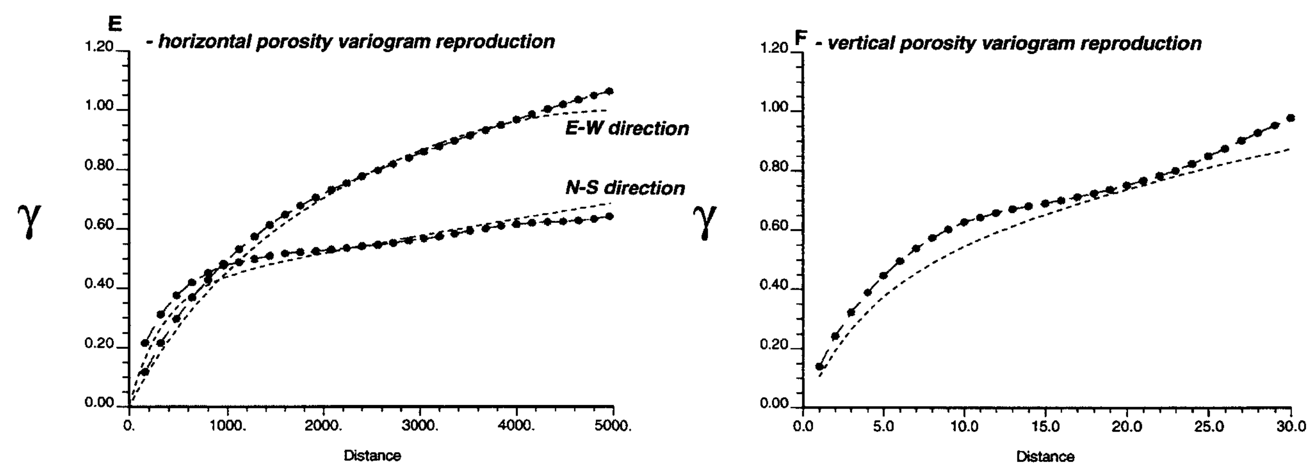


Figure 7. A, Vertical averaged porosity data; B, porosities vertically averaged from realization; C, vertical distribution of porosities in wells at a cross-section; D, simulated vertical distribution of porosities at the same cross-section as in Figure 7C; E, horizontal variogram reproduction between input variograms and variograms calculated from realization; F, vertical variogram reproduction between input variogram and variogram calculated from realization.

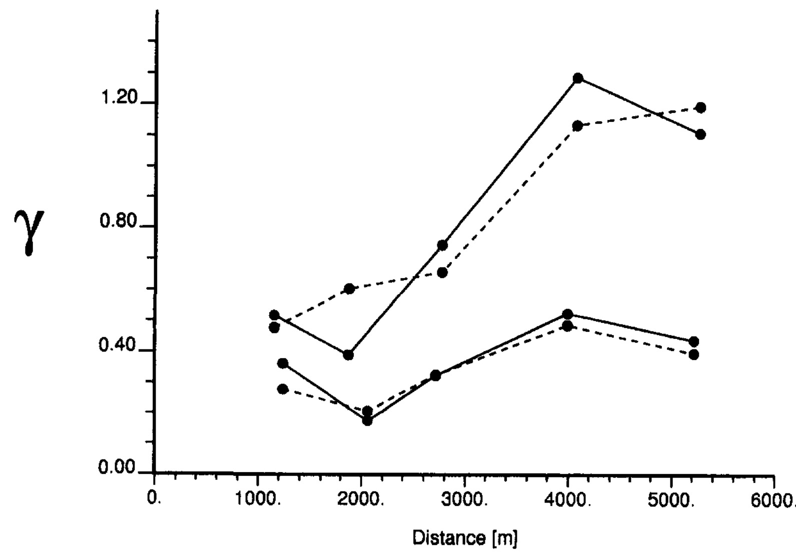


Figure 8. Comparison between horizontal variograms of vertical averaged porosities; solid lines represent arithmetic average and dashed lines represent harmonic average.

average as an upper bound and the harmonic parameter average as a lower bound (Journel, Deutsch, and Desbarats, 1986). The harmonic average is sensitive to values close to 0, whereas the arithmetic average is most influenced by large values. If the spatial structure of the harmonic and the arithmetic average of a given dataset are similar, the assumption of linear averaging after normal score transform is not important. To investigate the limitations of the linear averaging assumption, the vertical harmonic average of the porosity values are computed. Figure 8 shows a comparison between the horizontal variograms of the arithmetic (solid lines) and harmonic (dashed lines) vertical porosity averages. Discrepancies between the variogram values of the two averages are not significant. We conclude that the assumption of linear averaging does not influence the inference of small-scale variogram models for the presented case study.

CONCLUSIONS

Vertically averaged data such as seismic attributes are abundant in the horizontal plane and are valuable for the 3-D modeling of petrophysical parameters which are sampled at only a few well locations. The required cross-variogram values for kriging with large-scale secondary data can be inferred directly from the available data; however, there is no assurance that the resulting coregionalization model will be positive definite. The small-scale variograms and cross-variograms of the primary Z and secondary Y variables are needed to fit rigorously a positive definite model and integrate large-scale secondary data with kriging-based techniques. An analytical approach is presented which allows one to infer a small-scale variogram given only large-scale data. Modeling the porosity distribution of a West Texas field using a full cosimulation technique demonstrates that the analytical model for covariance averaging is capable of capturing sampled data features. Moreover, in terms of variograms of vertically averaged porosities, it is shown that the assumption of linear averaging does not influence the small-scale variograms.

ACKNOWLEDGMENTS

The first author was partly supported by the Fond zur Foerderung der wissenschaftlichen Forschung, Vienna, Austria, project no. J01059-GEO.

REFERENCES

Aboufirassi, M., and Mariño, M. A., 1984, Cokriging of aquifer transmissivities from field measurements of transmissivity and specific capacity: Math. Geology, v. 16, no. 1, p. 19–35.

Ahmed, S., de Marsily, G., and Talbot, A., 1988, Combined use of geostatistical methods for estimating transmissivity using data on specific capacity and transmissivity: Ground Water, v. 26, no. 1, p. 78–86.

Almeida, A. S. D., 1993, Joint Simulation of Multiple Variables with a Markov-Type Coregionalization Model: unpubl. PhD thesis, Stanford University, Stanford, CA.

Deutsch, C. V., and Cockerham, P. W., 1994, Geostatistical modeling of permeability with annealing cosimulation (acs), *in* 69th Annual Technical Conference and Exhibition, SPE 28413, Society of Petroleum Engineers, New Orleans, LA, p. 523–532.

Deutsch, C. V., and Journel, A. G., 1992, GSLIB: Geostatistical software library and user's guide: Oxford University Press, New York, 340 p.

Deutsch, C. V., Srinivasan, S., and Mo, Y., 1996, Geostatistical reservoir modeling accounting for the scale and precision of seismic data, *in* 1996 SPE Annual Technical Conference and Exhibition Formation Evaluation and Reservoir Geology: SPE 35497, Society of Petroleum Engineers, Denver, CO, p. 9–19.

Doyen, P. M., 1988, Porosity from seismic data: A geostatistical approach: Geophysics, v. 53, no. 10, p. 1263–1275.

Goovaerts, P., 1997, Geostatistics for natural resources evaluation: Oxford University Press, New York, 496 p.

Journel, A. G., and Huijbregts, C. J., 1978, Mining geostatistics: Academic Press, New York, 600 p.

Journel, A. G., Deutsch, C. V., and Desbarats, A. J., 1986, Power Averaging for Block Effective Permeability, SPE 15128.

Marechal, A., 1984, Kriging seismic data in presence of faults, *in* Verly, G., et al., eds., Geostatistics for natural resources characterization: Reidel, Dordrecht, Holland, p. 271–294.

Myers, D. E., 1984, Cokriging—new developments, *in* Verly, G., et al., eds., Geostatistics for natural resources characterization: Reidel, Dordrecht, Holland, p. 295-305.

Zhu, H., 1991, Modeling Mixture of Spatial Distributions with Integration of Soft Data: unpubl. PhD thesis, Stanford University, Stanford, CA.