

Comparison of Cell-Based Facies Modeling Techniques On a Synthetic Data Volume

YuLong Xie (ylxie@civil.ualberta.ca)

Department of Civil & Environmental Engineering, University of Alberta

W. P. Gouveia (wences_p_gouveia@email.mobil.com)

A. S. Cullick (a_stan_cullick@email.mobil.com)

Mobil Technology Company, Dallas, TX

Clayton V. Deutsch (cdeutsch@civil.ualberta.ca)

Department of Civil & Environmental Engineering, University of Alberta

Abstract

The assessment of subsurface reservoir exploitation alternatives often relies on fluid flow prediction using a three dimensional spatial description of the hydrocarbon reservoir. An accurate prediction of reservoir performance, i.e. the flow of fluids, depends on accurate representations of the subsurface "container" (structure), "plumbing" (flow units and fluid flow properties), and fluid content. The representation of the "plumbing" comes from the knowledge of the geology and rock properties. The rocks are characterized by reservoir facies, or lithologic facies, that are differentiated based on their porosity, permeability, relative permeability, compressibility, and saturations. The depositional environment or depositional process guides the facies connectivity.

The representation of lithofacies should be constrained by all available information about the reservoir, such as geological, geophysical and engineering data. Geostatistical stochastic modeling algorithms have been used extensively to model the subsurface distribution of lithofacies. However, the problem of generating models that are consistent with multiple data types and support volumes, e.g. well logs with high resolution vertical and low resolution horizontal combined with seismic with low resolution vertical and high resolution horizontal, is still a challenge.

We utilize a synthetic, small reservoir facies data model to compare stochastic, cell-based facies modeling algorithms. The purpose is to assess to what extent the algorithms produce models that can be constrained by sampled data at wells or seismic and also predict the spatial distribution of lithofacies as represented by the original exhaustive data. The geostatistical algorithms are the sequential indicator, truncated Gaussian, and Bayes-updating sequential indicator. Each uses facies indicator variograms as the principal source for the facies spatial distribution. We also generate models from an optimization-based algorithm. In this approach, the differences in the model and the data are minimized explicitly in a multi-component objective function.

The methods yield different quality of predictions. We show the dependence of model predictability on the density of sample locations. The optimization method is flexible and

extensible in incorporating different types of data, each with a different support volume. This extensibility is demonstrated with the addition of multiple-point spatial facies correlation.

Introduction

The assessment of subsurface reservoir exploitation alternatives often relies on fluid flow prediction using a three dimensional spatial description of the hydrocarbon reservoir. An accurate prediction of reservoir performance depends on accurate representations of the subsurface "container" (structure), "plumbing" (flow units and fluid flow properties), and fluid content. The representation of the "plumbing" comes from the knowledge of the geology and rock properties. The rocks are characterized by reservoir facies, or lithologic facies, that are differentiated based on their porosity, permeability, relative permeability, compressibility, and saturations. It is the lithofacies model, with its characteristic spatial distribution of the fluid porosity, permeability, and saturation, which governs the flow of fluids in the reservoir. The depositional environment or depositional process guides their connectivity.

The representation of lithofacies should be constrained by all available information about the reservoir, such as geological, geophysical and engineering data. Often, however, subsurface models of lithofacies are built on sparse well control. These data, as measured from core and well log, seismic, and well production, are representative of different scales and have distinct levels of uncertainty.

Geostatistical stochastic modeling algorithms have been used extensively to model the subsurface distribution of lithofacies. The problem of generating models that are consistent with multiple data types and support volumes, e.g. well logs with high resolution vertical and low resolution horizontal combined with seismic with low resolution vertical and high resolution horizontal, has long been a challenge. In this report we utilize a single reservoir facies data model as a means to compare output from several cell-based lithofacies modeling algorithms, in order to assess to what extent they can constrain the spatial distribution of lithofacies to varied sample data.

The geostatistical algorithms are the sequential indicator (Journel and Alabert [13], Alabert and Massonat [1] and Deutsch and Journel [5]), the truncated Gaussian (Matheron [19] and Galli [8]) and the Bayes-updating sequential indicator simulations (Doyen [6, 7]). These algorithms are pixel-based methods which utilize facies variogram models to estimate the presence or absence of a specific lithofacies in space. We do not present any results from an "object-based" (Boolean, Halderson and MacDonald [11], Halderson [12]) procedure as the data set used does not lend itself to defined geometric objects.

We also present an optimization-based algorithm. In this approach, initiated by Deutsch [4] and extended by Gouveia [9], a multi-component objective function is designed in such a way that its minimization results are models that are consistent with the input data sets under consideration. The method is flexible and extensible in incorporating different types of data, each with a different support volume.

The lithofacies modeling algorithms are applied to synthetic reservoir data, which is considered to be a representation of the "true" subsurface model. By having access to the "true" model, we can provide insight on the strengths and weaknesses of each one of the modeling procedures. The synthetic data set is described in the next section. Following that

Reservoir lithofacies	“seismic” facies
Facies 1	Sand
Facies 2	
Facies 3	
Facies 4	
Facies 5	
Facies 6	Shale

Table 1: Relationship lithofacies and “seismic” facies.

we present the results obtained with the geostatistical and optimization-based algorithms. The dependence on the density of sample data is examined for each. We then conclude with a summary on the results.

Data

The “true” reservoir model (Figure 1.a), is based on a segment extracted from a model of a producing oil reservoir, which has a shallow-marine depositional origin. The model contains six facies which are considered to be reservoir lithofacies (represented by different colors in the figure) spatially distributed on a grid of 100 cells in the East-West direction, 100 cells in the North-South direction and 10 cells in the depth direction. The areal dimensions of each cell is 100 meters by 100 meters with a thickness of 6.1 meters. Figure 1.b shows a top view of the 4th layer of the reservoir model. Twenty-five depth profiles along well trajectories, whose surface locations are shown in Figure 1.c, were extracted from this model and used as “well data”. Unless stated otherwise, these 25 depth profiles are considered as “hard” data in the simulation, i.e., their lithofacies values will not be perturbed during the modeling process.

In addition to the lithofacies indicators at the wells, we assume for the study that a seismic volume has provided “seismic” facies, which can be mapped to the reservoir facies and from which are derived sand proportions. As detailed in Gouveia [9] such “seismic” facies (Figure 2) were obtained by computing the proportions of shale and non-shale facies across the depth of the “true” reservoir model. Thus, the seismic data do not have any vertical resolution. The association of “seismic” facies and facies is listed in Table 1.

The spatial correlation of the six lithofacies is incorporated into the reservoir model via variograms. We determined a detailed variogram for each of the lithofacies using the “true” exhaustive reservoir model data. That provides a best case for variogram analysis. For each lithofacies, three major directions of continuity were selected (see Table 2) and associated experimental variograms have been calculated. Table 3 lists the analytic model variograms derived from the experimental variograms. These reproduce the basic features of the experimental variograms, as can be seen in Figure 3.

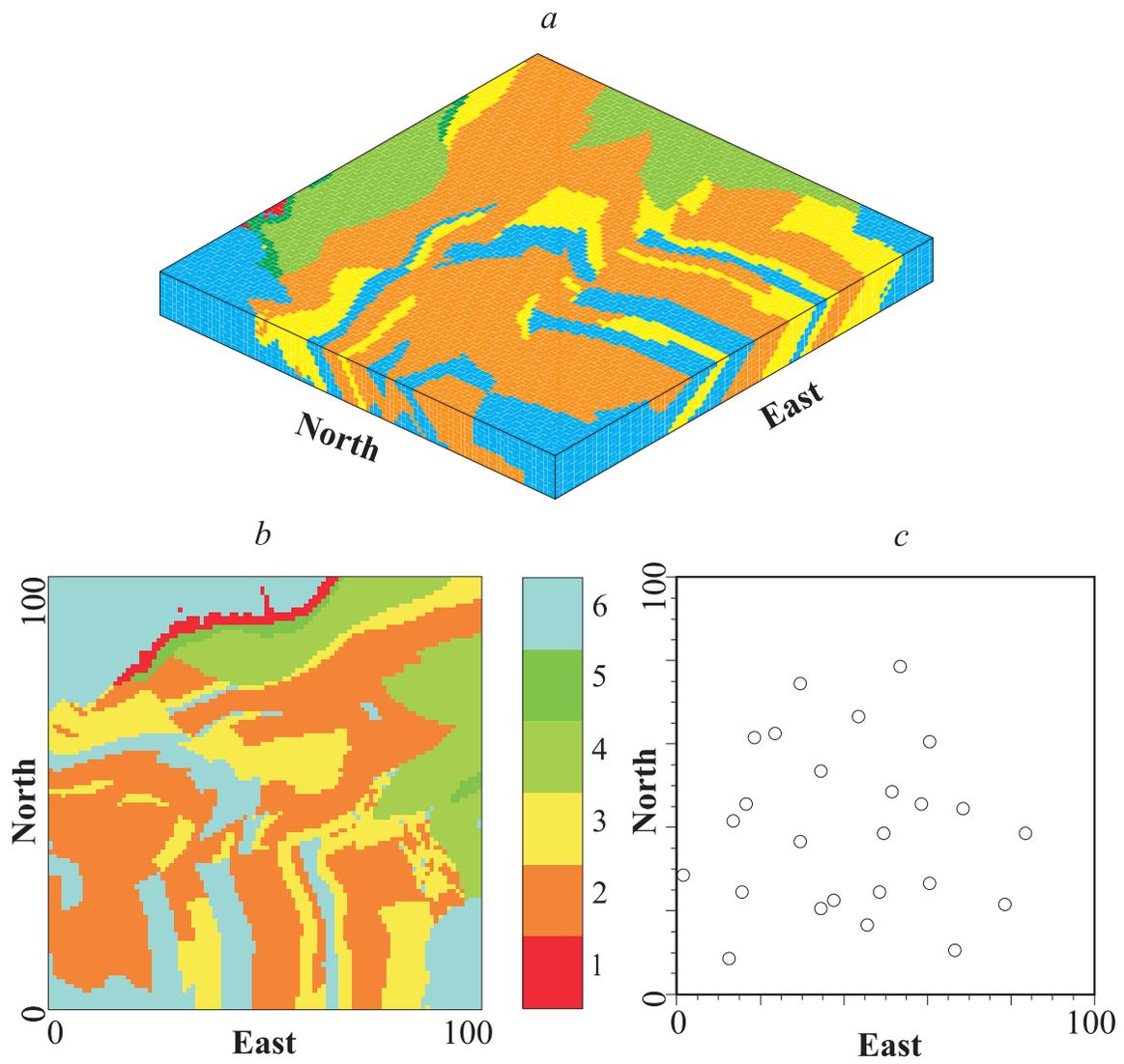


Figure 1: (a) True reference model, (b) its 4th layer and (c) location of the 25 wells used in the simulations.

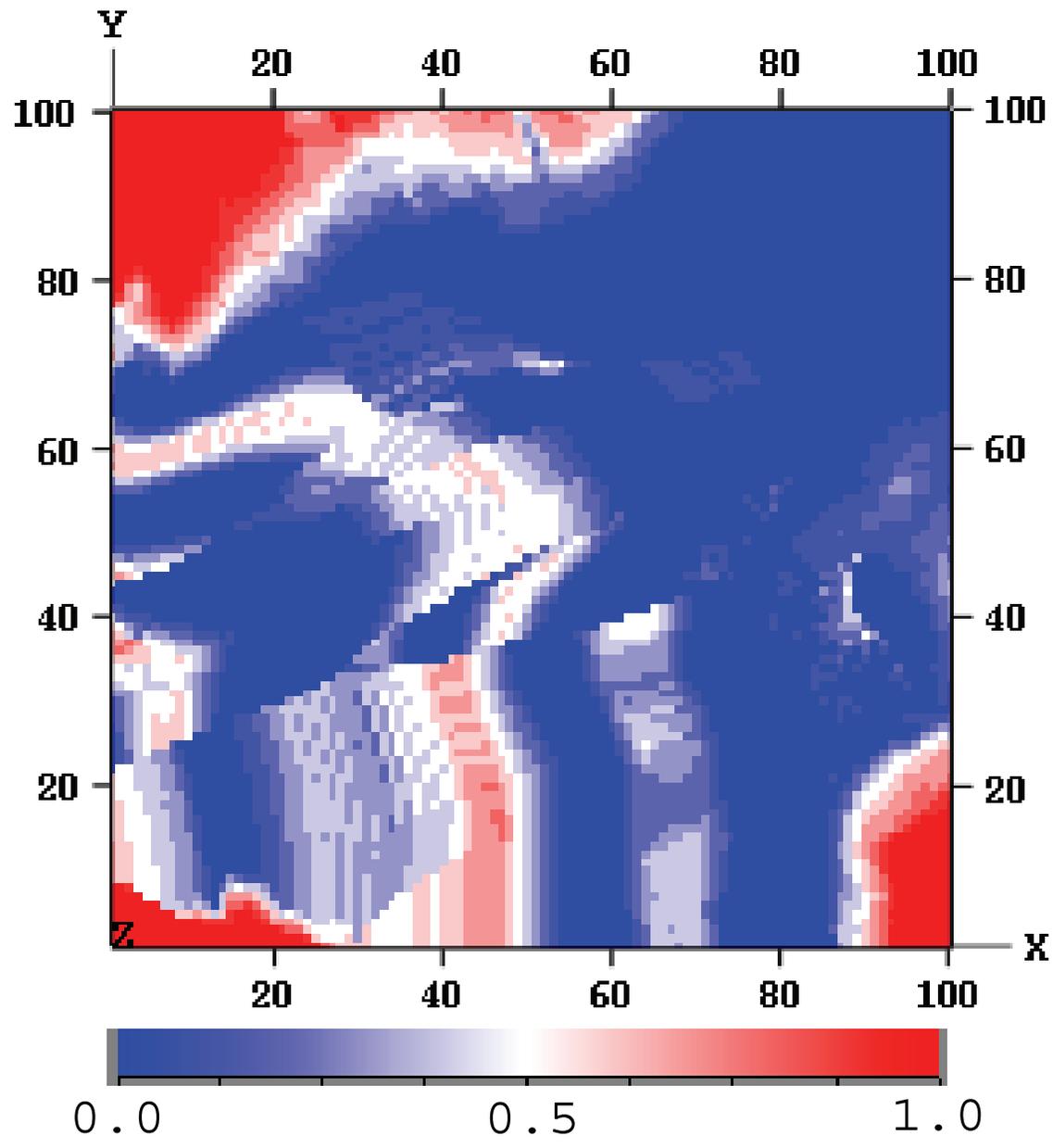


Figure 2: Proportions of “seismic” shale lithofacies obtained from the model shown in Figure 1.

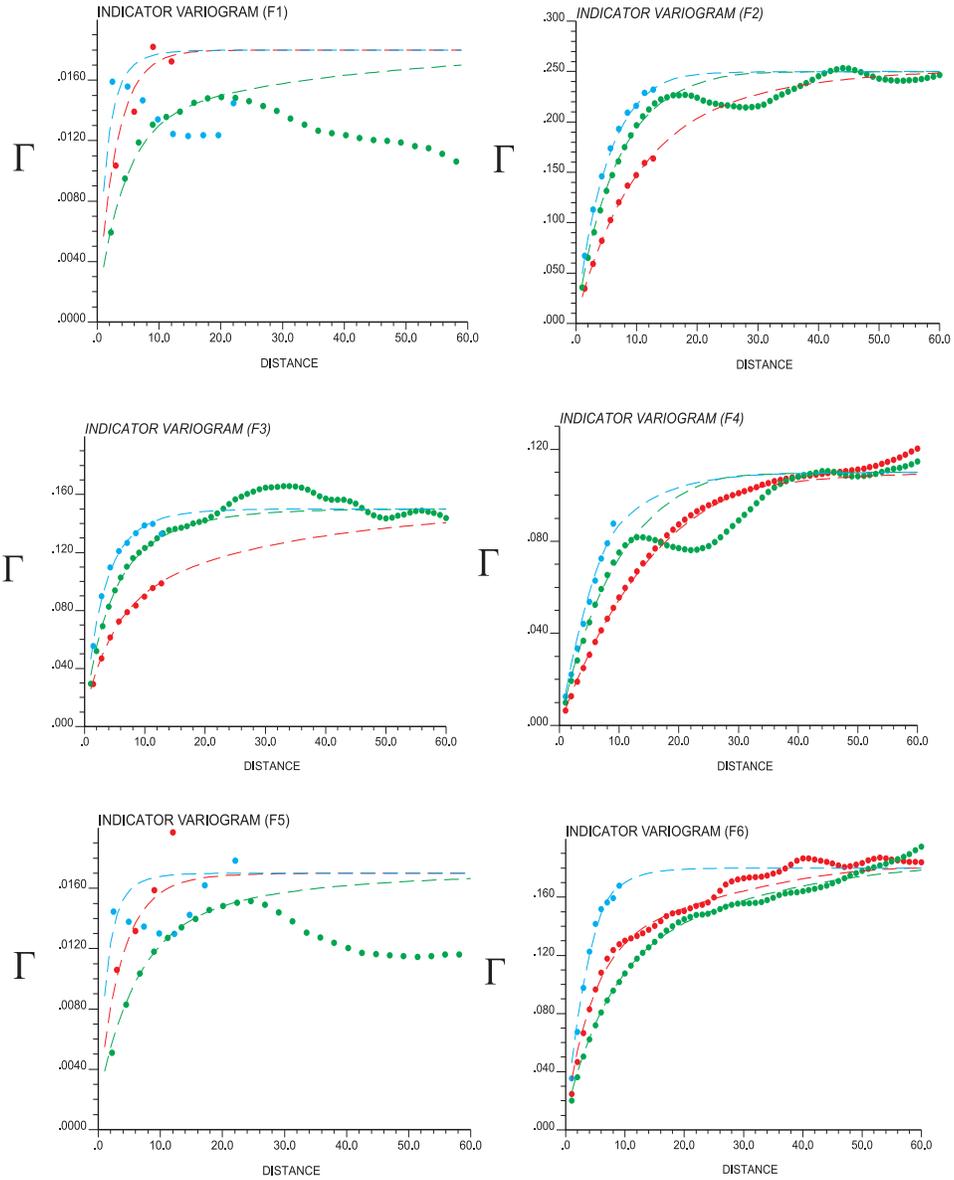


Figure 3: Experimental variograms computed from the “true” model (dots) and correspondent variogram models (dashed).

Reservoir lithofacies	Major directions (Azimuth, dip)
Facies 1	-20°, 45° 70°, 0° 160°, 45°
Facies 2	90°, 45° 180°, 0° 270°, 45°
Facies 3	90°, 45° 180°, 0° 270°, 45°
Facies 4	90°, 0° 0°, 0° 90°, 0°
Facies 5	-20°, 45° 70°, 0° 160°, 45°
Facies 6	900°, 0° 0°, 0° 0°, 90°

Table 2: Major correlation directions for reservoir lithofacies.

Facies Modeling

Sequential indicator simulation

Sequential indicator simulation (Journal and Alabert [15], Alabert and Massonat [1] and Deutsch and Journal [5]) is a commonly used non-Gaussian technique based on the indicator formalism that maps data into indicator variables. The indicator transformation for lithofacies, a discrete (categorical) variable, is defined as:

$$i(u_\alpha; f_k) = \begin{cases} 1 & \text{if } s(u_\alpha) = s_k, \text{ i.e., } s_k \text{ presents at } u_\alpha, k = 1, \dots, N_f \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where, $s(\mathbf{u})$ is the categorical random variable at location $\mathbf{u} \in \mathbf{A}$, A representing the reservoir spatial domain. The possible outcomes for this random variable are $s_k, k = 1, \dots, N_f$, where s_k is the k -th lithofacies and N_f is the number of possible lithofacies, six in this study.

The sequential indicator simulation algorithm proceeds as follows:

1. Calculate the indicator variograms in the major directions for each lithofacies type (Figure 3 and Table 3).
2. Apply the indicator transform (Equation 1) to each “hard” datum at location \mathbf{u} .
3. Define a random path visiting each node of the grid only once.

Facies	Variogram Model
1	$\gamma(h) = \gamma(h_1, h_2, h_3) = 0.001 + 0.012 \text{Exp} \left(\sqrt{\left(\frac{h_1}{10}\right)^2 + \left(\frac{h_2}{13}\right)^2 + \left(\frac{h_3}{4}\right)^2} \right)$ $+ 0.005 \text{Exp} \left(\sqrt{\left(\frac{h_1}{8}\right)^2 + \left(\frac{h_2}{110}\right)^2 + \left(\frac{h_3}{10}\right)^2} \right)$
2	$\gamma(h) = \gamma(h_1, h_2, h_3) = 0.005 + 0.215 \text{Exp} \left(\sqrt{\left(\frac{h_1}{30}\right)^2 + \left(\frac{h_2}{18}\right)^2 + \left(\frac{h_3}{14}\right)^2} \right)$ $+ 0.030 \text{Sph} \left(\sqrt{\left(\frac{h_1}{70}\right)^2 + \left(\frac{h_2}{30}\right)^2 + \left(\frac{h_3}{18}\right)^2} \right)$
3	$\gamma(h) = \gamma(h_1, h_2, h_3) = 0.010 + 0.070 \text{Exp} \left(\sqrt{\left(\frac{h_1}{14}\right)^2 + \left(\frac{h_2}{10}\right)^2 + \left(\frac{h_3}{7}\right)^2} \right)$ $+ 0.070 \text{Exp} \left(\sqrt{\left(\frac{h_1}{90}\right)^2 + \left(\frac{h_2}{28}\right)^2 + \left(\frac{h_3}{16}\right)^2} \right)$
4	$\gamma(h) = \gamma(h_1, h_2, h_3) = 0.000 + 0.080 \text{Exp} \left(\sqrt{\left(\frac{h_1}{40}\right)^2 + \left(\frac{h_2}{23}\right)^2 + \left(\frac{h_3}{24}\right)^2} \right)$ $+ 0.030 \text{Sph} \left(\sqrt{\left(\frac{h_1}{35}\right)^2 + \left(\frac{h_2}{30}\right)^2 + \left(\frac{h_3}{10}\right)^2} \right)$
5	$\gamma(h) = \gamma(h_1, h_2, h_3) = 0.002 + 0.011 \text{Exp} \left(\sqrt{\left(\frac{h_1}{10}\right)^2 + \left(\frac{h_2}{18}\right)^2 + \left(\frac{h_3}{4}\right)^2} \right)$ $+ 0.004 \text{Exp} \left(\sqrt{\left(\frac{h_1}{18}\right)^2 + \left(\frac{h_2}{75}\right)^2 + \left(\frac{h_3}{10}\right)^2} \right)$
6	$\gamma(h) = \gamma(h_1, h_2, h_3) = 0.010 + 0.120 \text{Exp} \left(\sqrt{\left(\frac{h_1}{14}\right)^2 + \left(\frac{h_2}{23}\right)^2 + \left(\frac{h_3}{12}\right)^2} \right)$ $+ 0.050 \text{Sph} \left(\sqrt{\left(\frac{h_1}{60}\right)^2 + \left(\frac{h_2}{70}\right)^2 + \left(\frac{h_3}{8}\right)^2} \right)$

Table 3: Variogram models for lithofacies of reservoir model shown in Figure 1.

4. At each node \mathbf{u}' :

- Estimate N_f conditional probability values $p(s(\mathbf{u}') = \mathbf{s}_k | \mathbf{data})$ using indicator kriging algorithms operating on the available conditioning data set *data*.
- Ensure that each conditional probability value are within the interval $[0, 1]$, and that their summation adds up to unity.
- The N_f conditional probability values define a probability distribution function for the lithofacies type at location \mathbf{u}' . Sampling this function results in the simulated lithofacies at this location, which is added to the conditioning data.
- Proceed to the next node along the random path and repeat steps above.

The indicator kriging procedure in step 1 above uses the variogram models listed in Table 3 to impose a correlation on the lithofacies spatial distribution.

In the simulations resulting from the indicator approach the distribution of heterogeneities of the reservoir lithofacies are quantified by the input variogram models. Therefore, it is essential that variograms associated with the simulated models reproduce the features of the variogram models. Figure 4.*left* shows three model realizations obtained from the indicator kriging simulation when none of the conditioning data, are used for the 4th layer. Figure 4.*right* shows similar results for three realizations when the 25 wells used as hard data. Visual comparison indicates that the simulations bear limited resemblance to the true spatial distribution of lithofacies, although some improvement was achieved when the conditioning data were used.

The variograms associated with the models shown in Figures 1.a and 1.b are displayed in Figures 5 and 6, respectively. The matches with the model variograms is poor, with some improvement when the 25 wells were used as hard data.

To further investigate the effect of conditioning data on the simulation results we systematically increased the number of available lithofacies log samples. These additional sample data were obtained by simply extracting the lithofacies values from the “true” model along vertical profiles at specific surface sample locations. In addition to the simulation with 25 wells, we performed simulations with a total of 49, 100, 400, 625 and 2500 surface sample locations, which were placed at regular spacing within the reservoir grid. Example results for the 4th layer of the resulting models are shown in Figure 7. With an increase in samples the resemblance of the simulated model to the “true” reservoir model improves. It is somewhat surprising that it takes 400 samples before the model really reproduces the true model closely. More than 400 samples provides little additional information.

Note that the stochastic simulations shown in this section and throughout the paper were post-processed by the *MAPS* algorithm [3] to remove the short-scale variability commonly encountered in cell-based simulation procedures. The algorithm reduces the short-scale variability while preserving a match to the data and maintaining consistency with the input lithofacies global proportions and variogram models.

Truncated Gaussian Simulation

An alternative to sequential indicator simulation is the truncated Gaussian simulation (Matheron [19] and Galli [8]). In this approach the spatial distribution of lithofacies are

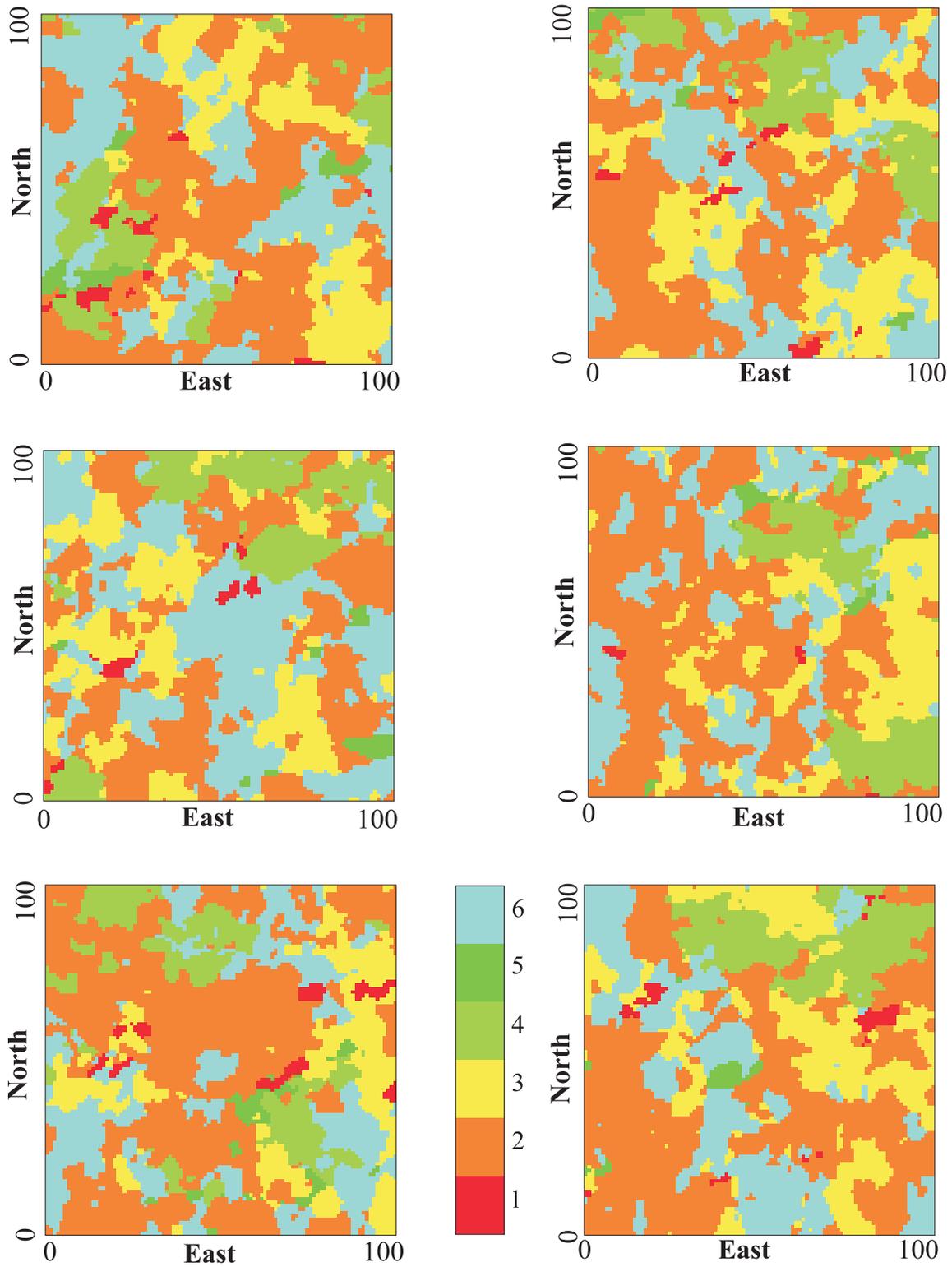


Figure 4: The 4th layer of sequential indicator simulation result without (*left*) and with (*right*) 25 conditioning wells.

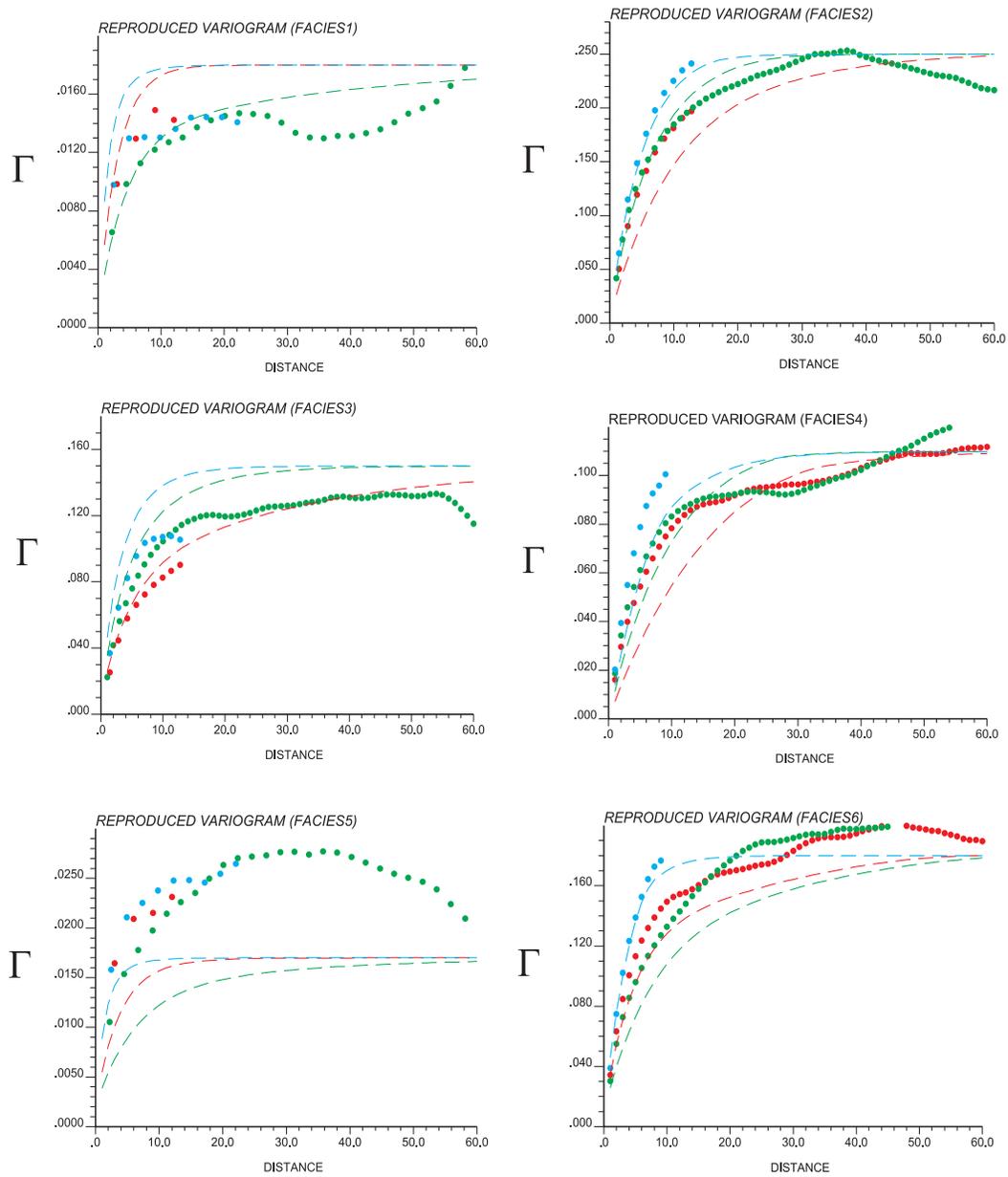


Figure 5: Variograms computed from the sequential indicator simulation without conditioning data (dots) and correspondent variogram models (dashed) from one realization.

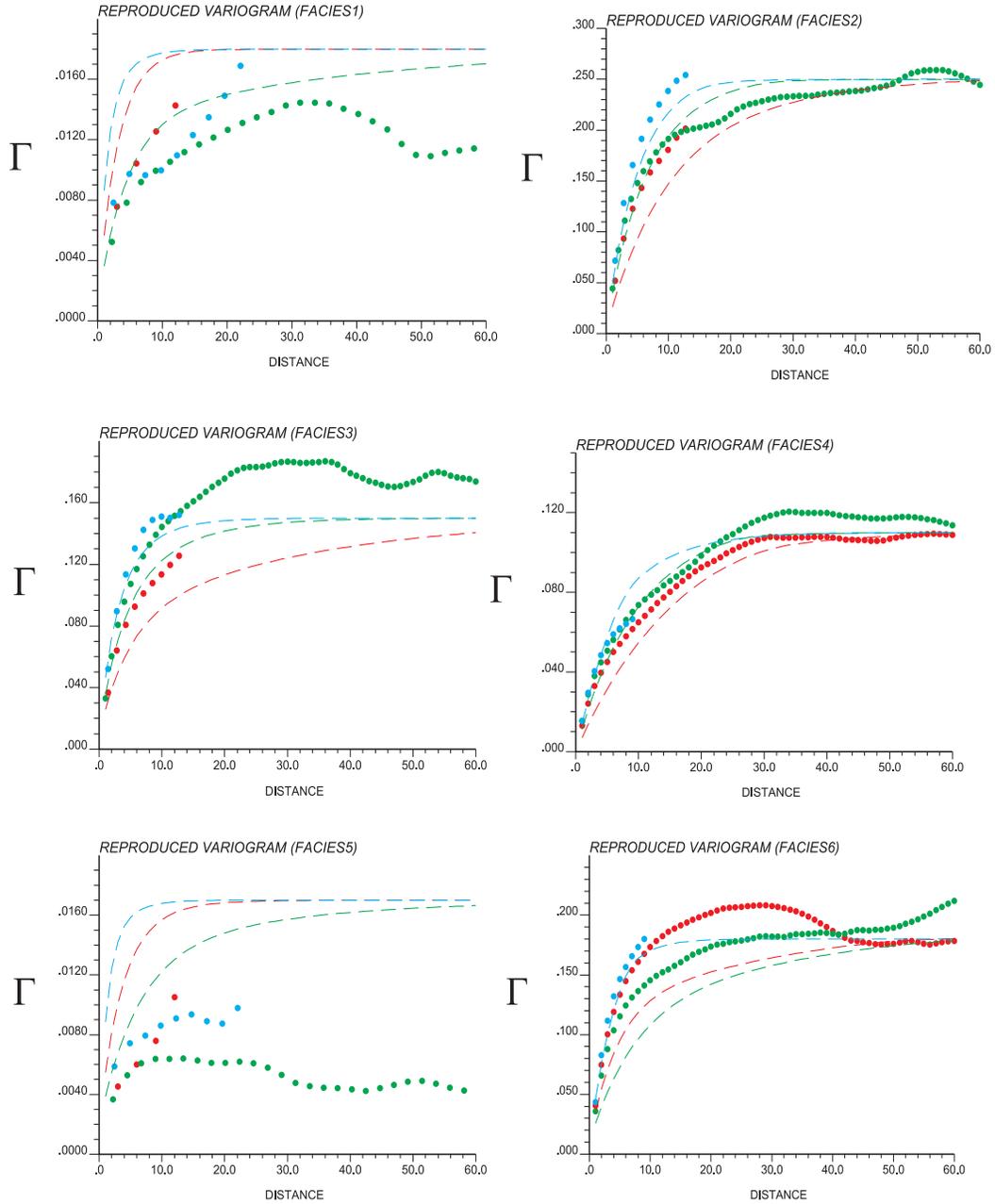


Figure 6: Variograms computed from the sequential indicator simulation with 25 conditioning wells (dots) and correspondent variogram models (dashed) from one realization.

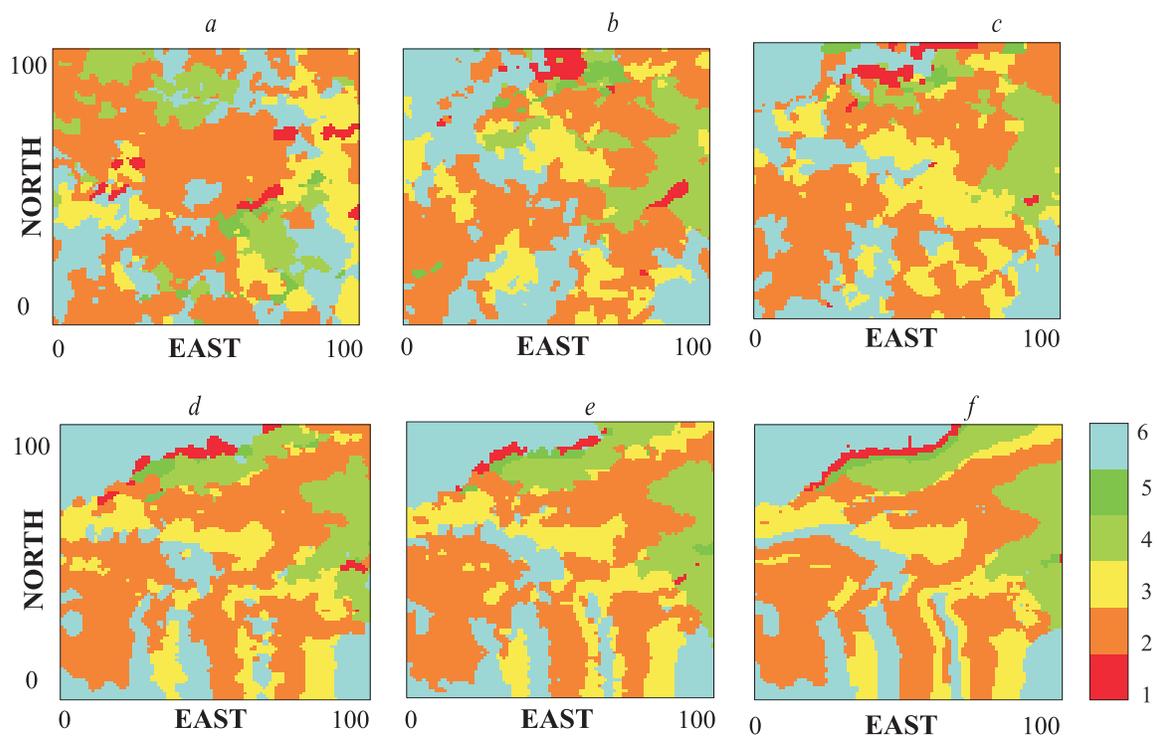


Figure 7: 4th layer of sequential indicator simulation results with increasing surface sample locations. (a) None, (b) 49, (c) 100, (d) 400, (e) 625 and (f) 2500.

modeled by truncating an underlying Gaussian multidimensional function according to the proportions of each lithofacies.

A required step to applying the truncated Gaussian method is to ensure that at least the underlying univariate probability function of the conditioning data is Gaussian. This is accomplished via the normal score transform [18], a non-linear transformation that maps any continuous uni-dimensional random variable into a Gaussian one. In addition, the covariance of the multi-dimensional Gaussian function is determined in such a way that it will yield the required indicator variogram [16].

A limitation of the truncated Gaussian approach is that it cannot be used to reproduce more than one indicator variogram, since the unique parameter of the underlying Gaussian distribution is its covariance model. If multiple categories are to be modeled, the usual procedure is to truncate the Gaussian realization at a number of thresholds defined by the global proportions of the lithofacies. Although the correct lithofacies proportions will be obtained, their spatial structures (with the exception to the one used to define the Gaussian covariance) will not reflect the indicator variogram models, rather they will be arbitrarily controlled by the Gaussian function.

Basically, the truncated Gaussian simulation proceeds as follows:

1. Conditioning data are mapped to normal scores ($y(\mathbf{u}) \in (-\infty, +\infty)$) via the normal score transform.
2. Estimate a normal score variogram based on a representative indicator lithofacies variogram. Such a representative variogram can be the one associated with the largest lithofacies proportions or an “average” indicator variogram.
3. Build a Gaussian random field by sequential Gaussian simulation:
4. Define a random path visiting each node of the grid only once.
5. At each node \mathbf{u}' :
 - Estimate the mean and variance of the conditional Gaussian PDF $p(y(\mathbf{u}')|\mathbf{data})$ by kriging the normal score conditioning *data*.
 - Sample the Gaussian PDF at node \mathbf{u}' to generate the simulated value at this node. Add the simulated value to the conditioning *data* set.
 - Proceed to the next node along the random path and repeat steps above.
6. From the global lithofacies proportions establish N_f+1 thresholds $t_1 = -\infty, t_2, \dots, t_{N_f+1} = +\infty$.
7. Assign lithofacies f_k wherever the Gaussian simulated field is within the interval $[t_k, t_{k+1}]$.

Here, we use the variogram of lithofacies two, the one with the largest global proportion, to construct the variogram associated with the Gaussian function.

$$\gamma(h) = \gamma(h_1, h_2, h_3) = 0.0 + 0.5 \text{Gau} \left(\sqrt{\left(\frac{h_1}{35}\right)^2 + \left(\frac{h_2}{22}\right)^2 + \left(\frac{h_3}{15}\right)^2} \right)$$

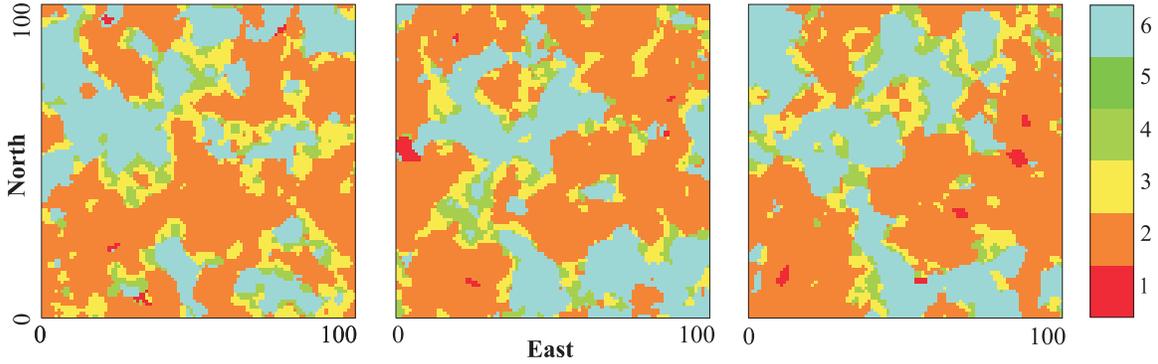


Figure 8: The 4th layer of three truncated Gaussian simulation realizations with 25 conditioning wells.

$$+ 0.5 \text{Sph} \left(\sqrt{\left(\frac{h_1}{60}\right)^2 + \left(\frac{h_2}{30}\right)^2 + \left(\frac{h_3}{22}\right)^2} \right). \quad (2)$$

Figure 8 shows three example truncated Gaussian realizations when the 25 wells are used as “hard” data. Clearly, the realizations fail to reproduce the basic spatial structure of the “true” model (Figure 1). This result, much inferior to these of the sequential indicator procedure (Figure 7), is not unexpected due to the fact that the truncated Gaussian approach cannot account for all lithofacies variograms. By increasing the number of conditioning surface sample locations to 49, 100, 400, 625 and 2500, we obtained the results shown in Figure 9. Again, conditioning data denser than 400 provides little additional improvement on the lithofacies spatial distribution.

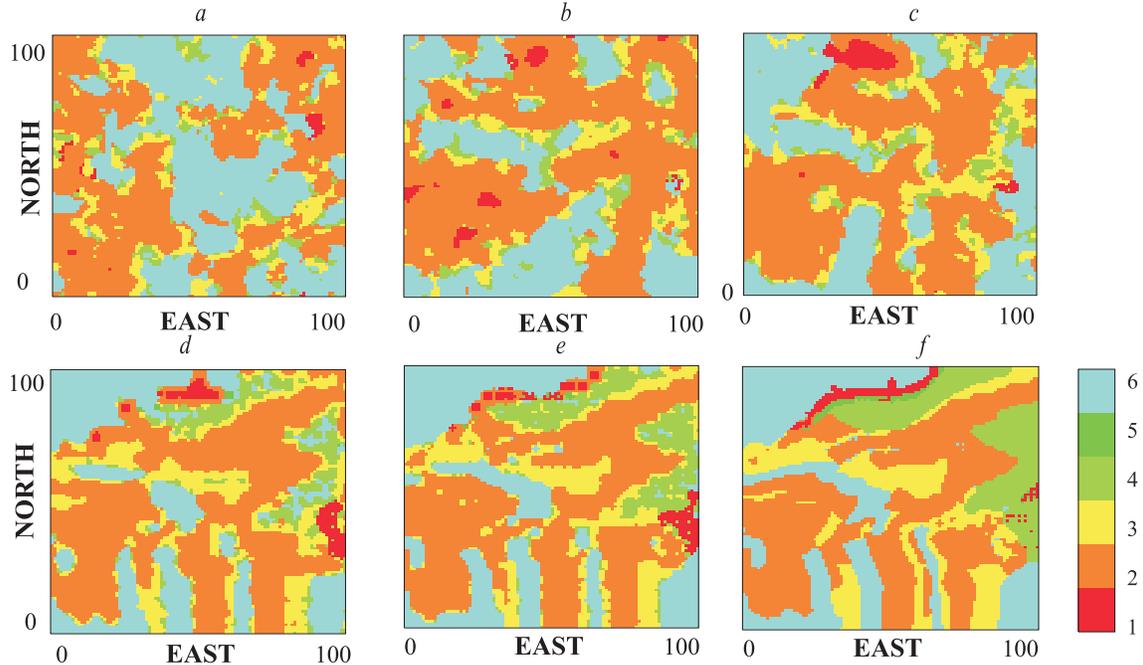


Figure 9: 4th layer of truncated Gaussian simulation results with increasing number of surface sample locations. (a) None, (b) 49, (c) 100, (d) 400, (e) 625 and (f) 2500.

Bayes-updated indicator simulation

The facies simulation can be imposed with the addition of the seismic data as secondary data (Figure 2). A Bayes updating procedure [6, 7] provides for the addition of the “seismic” facies proportion. A block co-kriging (Deutsch [17] and Behrens [2]) or sequential indicator simulation with local mean also provide for incorporating secondary data and should yield equivalent results to the Bayes approach. The Bayes approach biases the indicator-derived probability $p(s(\mathbf{u})|\text{data})$ with the seismic-derived probability $p(s(\mathbf{u})|\text{seismic data})$ in order to produce the aposteriori probability $p(s(\mathbf{u})|\text{data, seismic data})$, given by:

$$\begin{aligned}
 p(s(\mathbf{u})|\text{data, seismic data}) &= \frac{p(s(\mathbf{u})|\text{data}) \mathbf{p}(\text{seismic data}|\mathbf{s}(\mathbf{u}))}{p(\text{seismic data})} \\
 &= \frac{p(s(\mathbf{u})|\text{data}) \mathbf{p}(\mathbf{s}(\mathbf{u})|\text{seismic data})}{p(s(\mathbf{u}))}, \quad (3)
 \end{aligned}$$

where $p(s(\mathbf{u}))$ is a probability on the facies values for location $s(\mathbf{u})$, and can be regarded as the well-data derived global proportions of lithofacies. The probability $p(s(\mathbf{u})|\text{data, seismic data})$ reduces to $p(s(\mathbf{u})|\text{data})$ when the seismic-derived probability $p(s(\mathbf{u})|\text{seismic data})$ equals $p(s(\mathbf{u}))$, meaning that the seismic data do not bring any additional information on the lithofacies than the one contained in the well data.

The seismic-derived probabilities are needed for each reservoir cell and for each lithofacies. For the “true” model, those were obtained in the following way. The seismic shale (Figure 2) and sand facies proportions were directly assigned to probabilities. The connec-

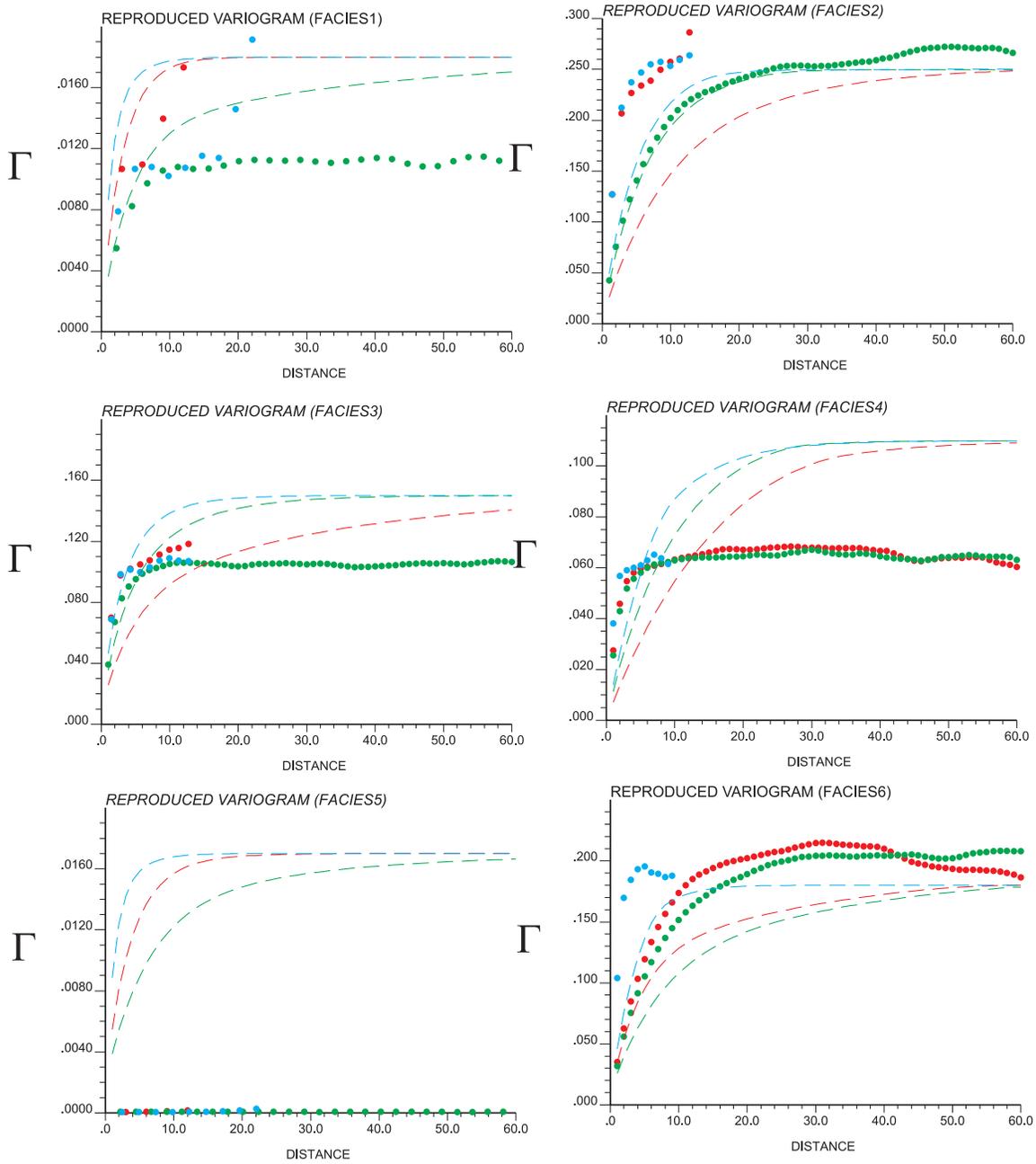


Figure 10: Variograms computed from the truncated Gaussian simulation with 25 conditioning wells (dots) and correspondent variogram models (dashed) from one realization.

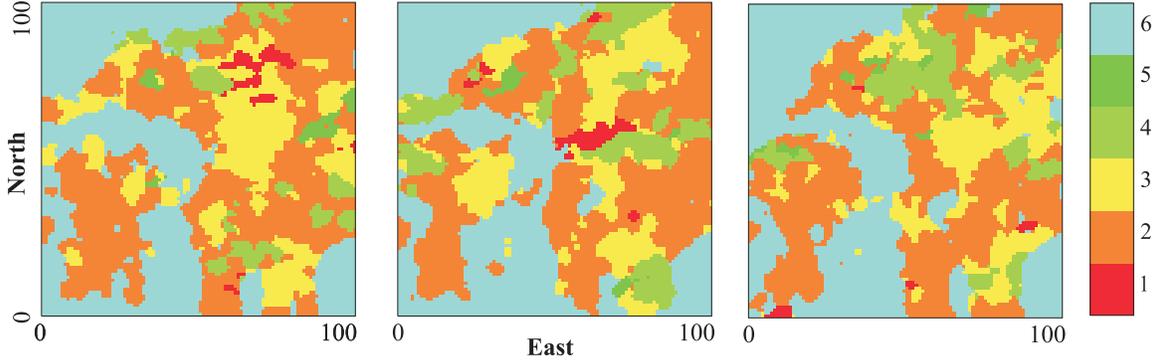


Figure 11: The 4th layer of three Bayes-updating simulation realizations with 25 conditioning wells.

tion between “seismic” facies and reservoir facies was obtained with the facies association in Table 1, i.e.:

$$\begin{aligned}
 p(s(\mathbf{u}) = \text{facies 6} | \text{seismic data}) &\propto \mathbf{p}(s(\mathbf{u}) = \text{shale}) \\
 p(s(\mathbf{u}) = \text{facies 1} | \text{seismic data}) &\propto \mathbf{p}(s(\mathbf{u}) = \text{sand}) \mathbf{p}(\text{facies 1}) \\
 p(s(\mathbf{u}) = \text{facies 2} | \text{seismic data}) &\propto \mathbf{p}(s(\mathbf{u}) = \text{sand}) \mathbf{p}(\text{facies 2}) \\
 p(s(\mathbf{u}) = \text{facies 3} | \text{seismic data}) &\propto \mathbf{p}(s(\mathbf{u}) = \text{sand}) \mathbf{p}(\text{facies 3}) \\
 p(s(\mathbf{u}) = \text{facies 4} | \text{seismic data}) &\propto \mathbf{p}(s(\mathbf{u}) = \text{sand}) \mathbf{p}(\text{facies 4}) \\
 p(s(\mathbf{u}) = \text{facies 5} | \text{seismic data}) &\propto \mathbf{p}(s(\mathbf{u}) = \text{sand}) \mathbf{p}(\text{facies 5}), \tag{4}
 \end{aligned}$$

where $p(s(\mathbf{u}) = \mathbf{sf}_1)$ is the seismic-derived probability of presence of seismic lithofacies sf_1 at location \mathbf{u} , $p(s(\mathbf{u}) = \mathbf{rf}_1)$ is the probability of presence of reservoir lithofacies rf_1 at location \mathbf{u} , and $p(\text{facies})$ represents the global proportion of this reservoir lithofacies. The probabilities on the left hand side of Equation 4 were then used to modify the indicator-derived probabilities as described in Equation 3.

Figure 11 shows three realizations of the 4th layer when the 25 wells were used as conditioning data in addition to the seismic data. The variograms associated with this simulation are shown in Figure 12. Comparison of the indicator kriging simulation results without seismic data (Figure 4) with Bayes-updated simulations show that incorporation of the seismic data brought some improvement to the spatial distribution of the reservoir lithofacies. As before, we progressively increased the number of surface sample location conditioning data to 49, 100, 400, 625 and 2500. Surface sample locations numbering greater than 100 (sample spacing of 1000 m) contribute little to further improvement.

Optimization

Gouveia [9] presented a formulation in which different input data types are used as components of an objective function for which the final model is consistent within a tolerance of the data. The attraction of this technique is the flexibility and extensibility available to add different data types, each at its appropriate support volume. This comes at some additional computational cost. In this section we show reservoir lithofacies models obtained from this

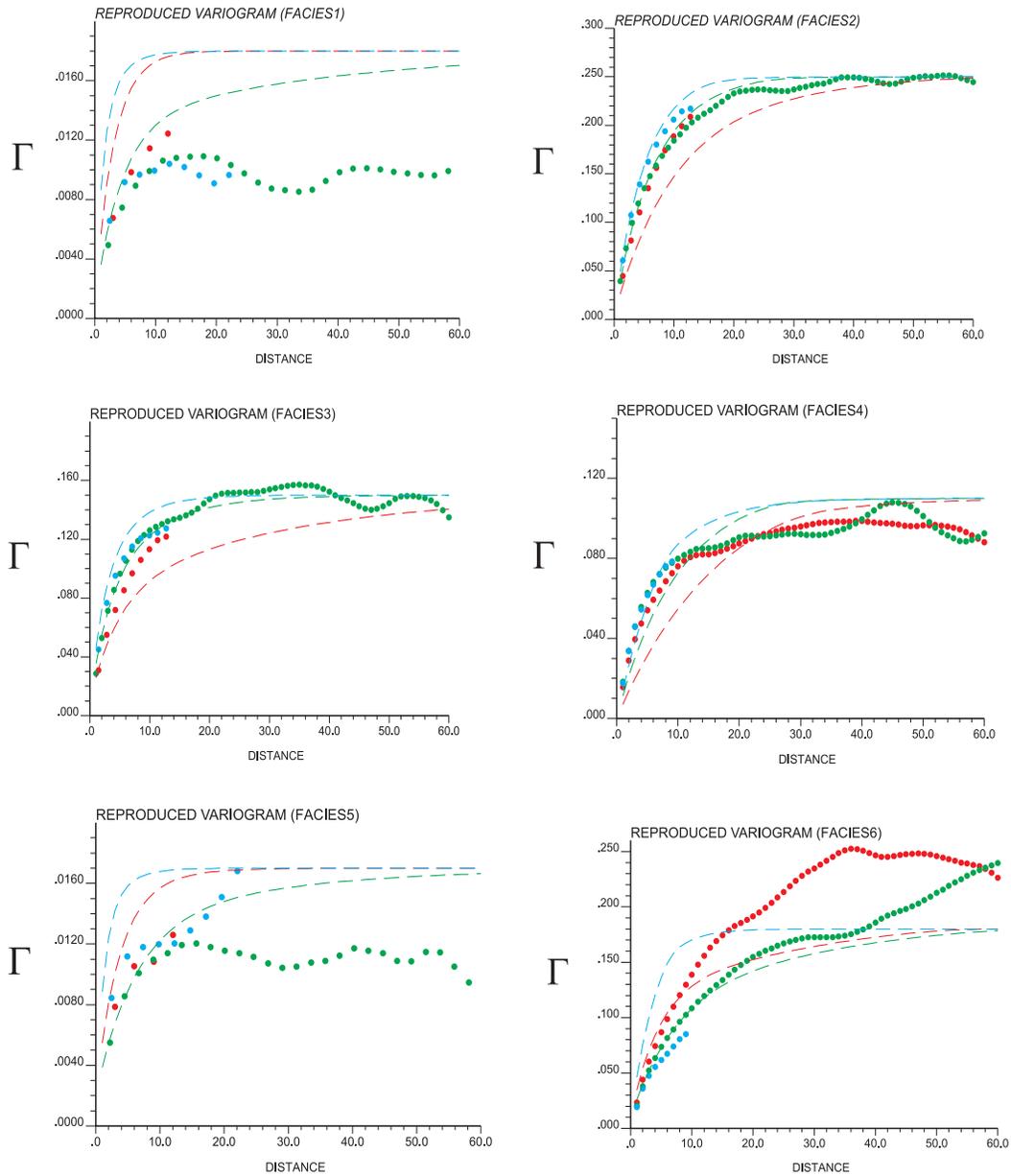


Figure 12: Variograms computed from the Bayes-updated simulation with 25 conditioning wells (dots) and corresponding variogram models (dashed) from one realization.

optimization technique for the following two cases. In Case 1, the input data are the global lithofacies proportions, indicator variograms, and “seismic” -facies proportions. Case 1 can be compared directly with the Bayes-updating approach. In Case 2, additional data on the lithofacies spatial arrangement in the form of transition probabilities [9] is added. Case 2 demonstrates the extensibility of the optimization approach and improved model with added data.

Case 1

The lithofacies objective function consists of a weighted sum of three components, global lithofacies proportions (Table 1), indicator variograms (Table 3) and “seismic” -facies proportions (Figure 2). The expression for the objective function is given by

$$\begin{aligned}
 O[floc] = & w_{gp} \sum_{f=1}^{N_f} \|P_f^{strata} - P_f^*\| \\
 & + w_{var} \sum_{f=1}^{N_f} \sum_{lag=1}^{N_{lags}} \|I_f^{strata}(lag) - I_f^*(lag)\| \\
 & + w_{seis} \sum_{i=1}^{N_q} \|H_{shale}^*(i) - \frac{1}{N_q}\|, \tag{5}
 \end{aligned}$$

where P_f is the global proportions of lithofacies f and I_f is the indicator variogram (defined for N_{lags} lags) for this lithofacies. The superscripts *strata* and \star indicate that these quantities are computed from the “true” reservoir model (Figure 1) and the reservoir model at a given iteration in the optimization procedure, respectively. H_{shale} is the quantile histogram (defined for N_q bins) associated with the reservoir shale proportions at seismic scale [9]. Recall that the seismic data do not provide any resolution vertically. In the optimization approach this difference in resolution is handled by the definition of multiple grids [9]. Each seismic grid cell is associated with a probability density on shale proportions obtained from the seismic inversion procedure which quantifies the uncertainty of the seismic estimates. At any given iteration it is possible to compute the shale proportions of the reservoir model at any given cell of the seismic grid which, in turn, represents a quantile of the probability density function associated with this cell. As explained in Gouveia [9] and [9], the extent that the histogram of such quantiles (H_{shale}) is uniform indicates that the reservoir lithofacies model is consistent with the seismic data information and associated uncertainty.

Optimization of the objective function 5 using the initial model shown in Figure 13 results in the model realization shown in Figure 14. The initial model is not fully random, rather it is constructed in such a way that its shale proportions at seismic scale are consistent with the ones obtained from the seismic data. The convergence of the objective function and its components is shown in Figure 15 for 400000 iterations. The behaviour of each one of the components is illustrative of the non-linearity of this optimization problem. The variograms are shown in Figure 16.

A close look at the variograms shows that the ones from the optimization generally have a better match with the input variogram models than the ones from Bayes-updating. An

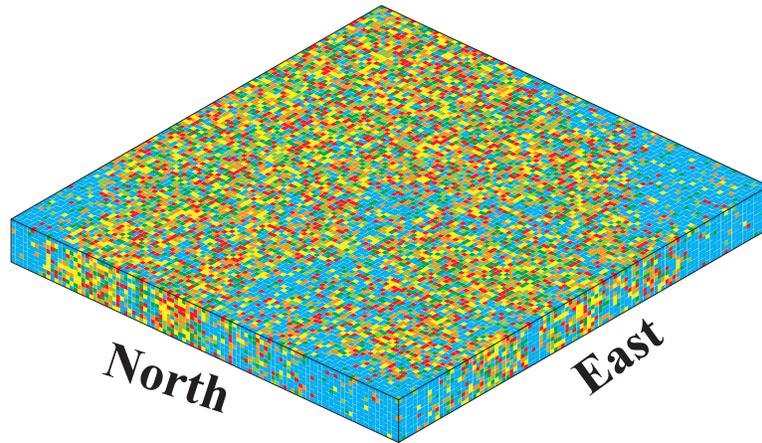


Figure 13: Initial reservoir lithofacies model used in the optimization.

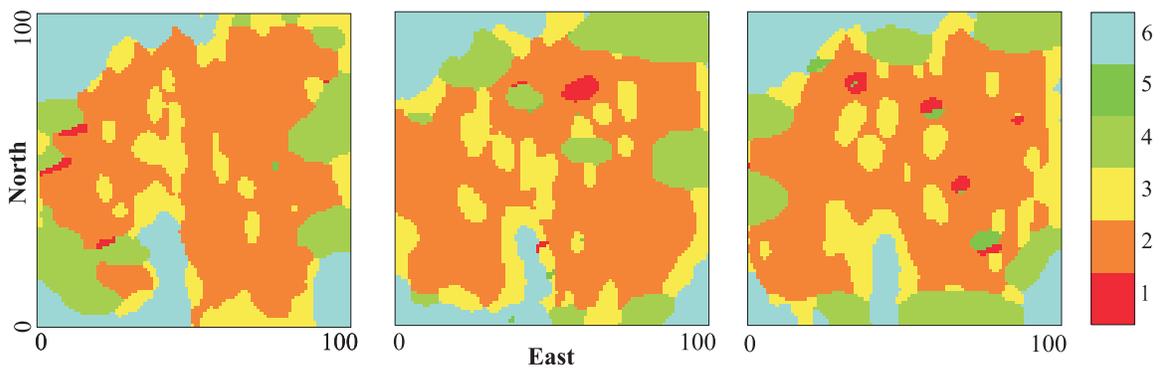


Figure 14: The 4th layer of three optimization simulation realizations with 25 conditioning wells.

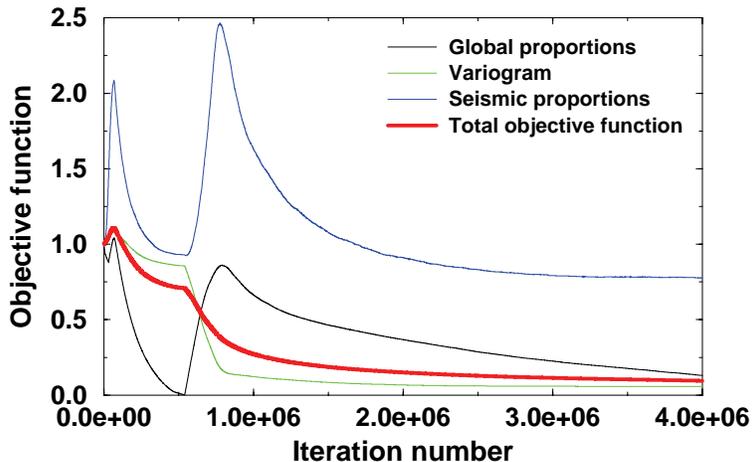


Figure 15: Convergence of the composite objective function and each one of its components. Recall that the initial model used in the optimization was already consistent with the seismic data.

exception to this happens at the smaller lags for the shale lithofacies, where the Bayes-updated variograms present a closer fit. In terms of assessing the consistency with the seismic data information (local PDFs for shale proportions defined at seismic scale) the optimization algorithm provides a quantitative measure, namely the quantile histogram. The extent that this quantile histogram is uniform is indicative of the degree to which the reservoir model reproduces the seismic-derived input data proportions. The quantile histogram for the seismic proportions will be less than uniform to the degree to which other data types (i.e., global proportions and variograms) with different proportions must be accommodated simultaneously by the optimization algorithm. Nonetheless, the quantile histogram associated with the optimum model does indicate that most of the seismic-scale shale proportions lie within the limits of the seismic-derived PDFs.

Case 2

We incorporate the additional data on lithofacies transition probabilities as information on the spatial arrangement between different lithofacies (Gouveia [9]). These are computed along three directions in the reservoir model: East-West, North-South and along the general dip of the reservoir layers. Transition probabilities are one possible way to incorporate more detailed geologic depositional information, as might be available from such sources as outcrop analogs, sequence stratigraphic interpretation, or log motifs. The added objective function component is:

$$w_{tp} \sum_{i=1}^{N_{TP}} \sum_{j=1}^{N_{comp}} \|TP_i^{strata}(j) - TP_i^*(j)\|, \quad (6)$$

where w_{tp} is the weight associated with the transition probability objective function component and $TP_i^{strata}(j)$ is the j -th component of the transition probability computed along

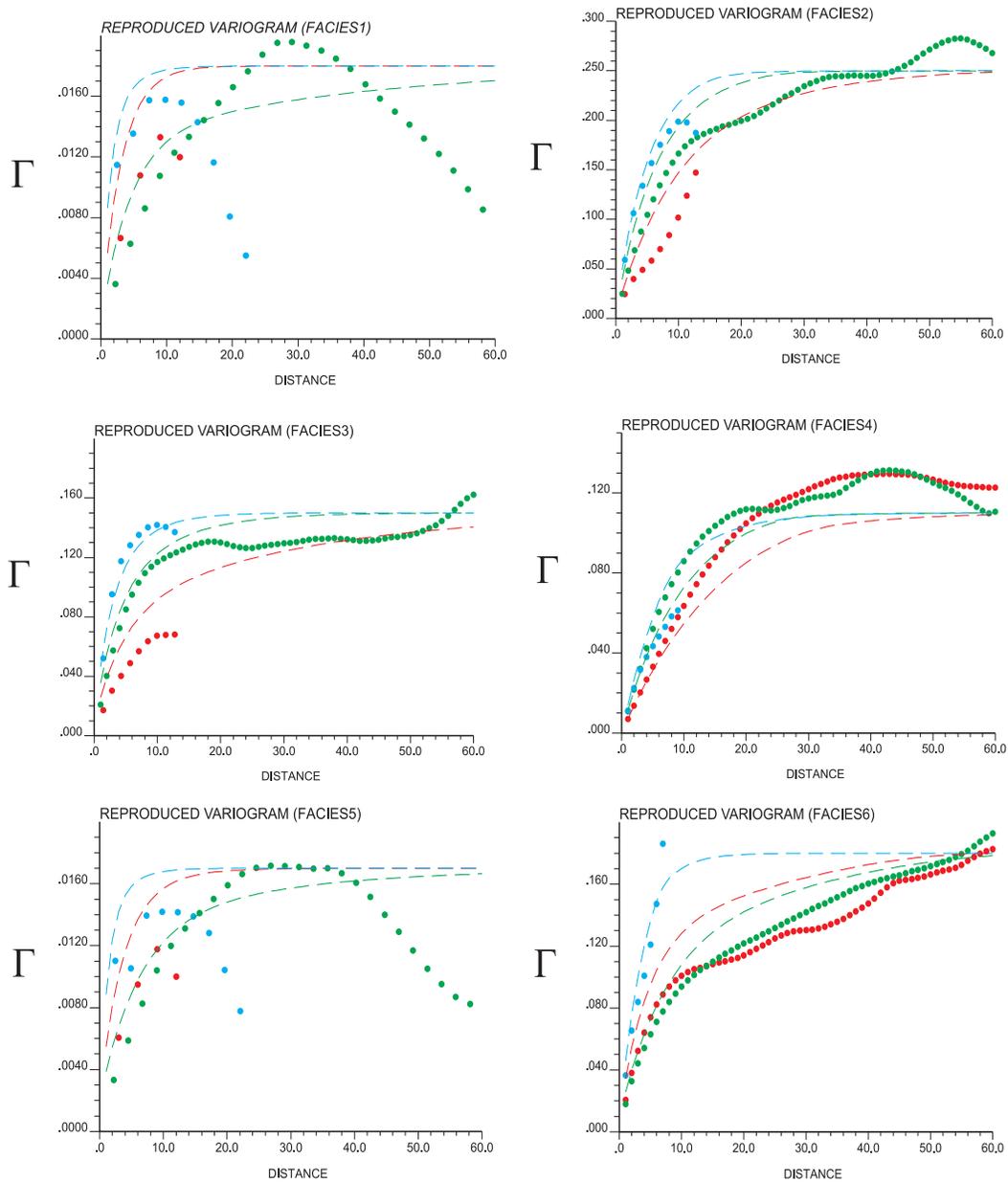


Figure 16: Indicator variograms associated with the optimum model in Figure 14 (dotted lines) and model variograms defined in Table 3 (dashed lines).

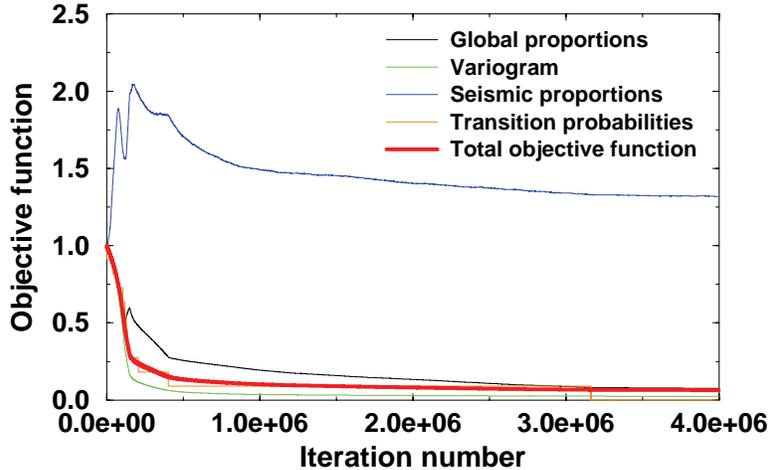


Figure 17: Convergence of the composite objective function and each one of its components. The initial model used in the optimization was consistent with the “seismic” facies proportions.

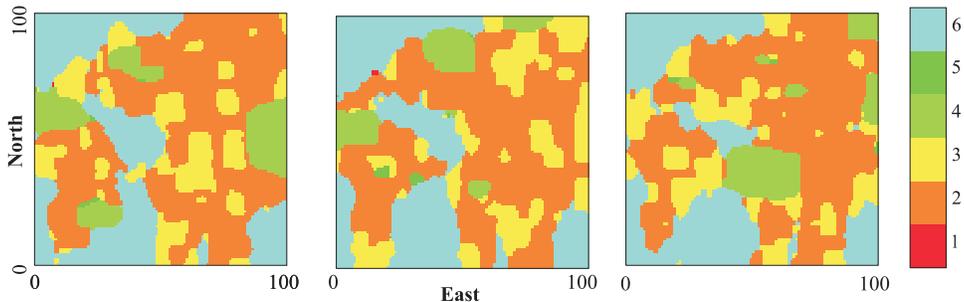


Figure 18: The 4th layer of three optimization realizations. The same initial model shown in Figure 13 was used in the optimization of Equation 6

the i -th direction. A fourth-order transition probability component is defined by any combination of lithofacies present in four successive grid cells along one of the specific directions. In this case, since six lithofacies are being modeled, there are a total of $N_{comp} = 6^4$ components. A low or high probability value associated with a specific lithofacies pair indicates that that association is unlikely or highly likely, respectively, and would be present only to that probability.

Three realizations are shown in Figure 18 and the variograms in Figure 19. The transition probability match is shown in Figure 20.

Again, large-scale features presented in the Bayes-updated model (Figure 11) and in the previous optimization result (Figure 14) are generally similar to the ones present in the model shown in Figure 18. The variograms in Figure 19 show an improved fitting to the input model (Figure 16). This is specially true for the smaller variogram lags, which may be a consequence of the fact that the transition probabilities, basically a constraint on the small-

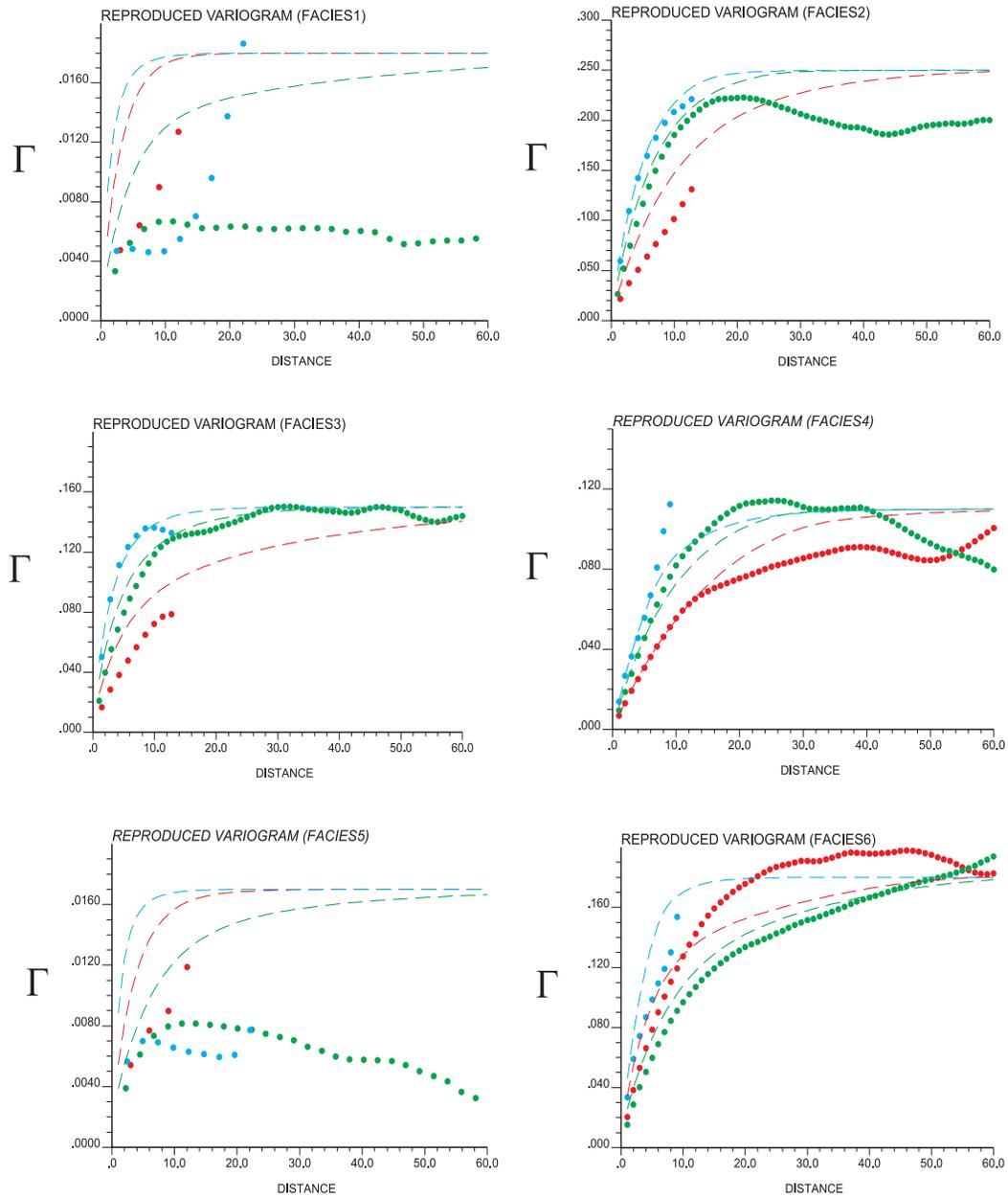


Figure 19: Indicator variograms associated with one optimum model (dotted lines) and model variograms defined in Table 3 (dashed lines).

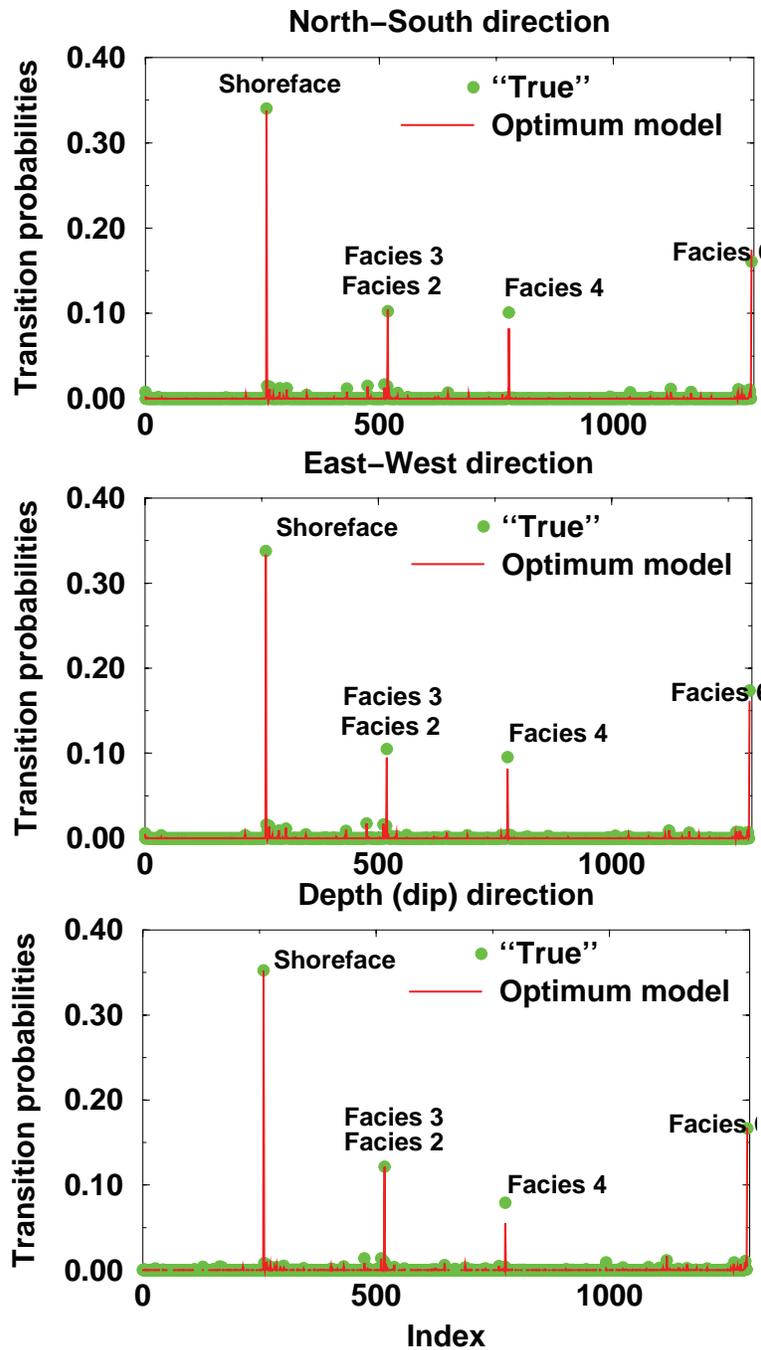


Figure 20: Transition probabilities associated with the optimum model in Figure 18 and the transition probabilities calculated from the lithofacies model in Figure 2. Most of the features in the input transition probabilities have been reproduced by the optimum model.

scale facies variability, are being used. In terms of consistency with the seismic data both the models obtained by the optimization approach are equivalent, since their respective quantile histograms similarly approximate an uniform histogram. Finally, the transition probabilities of the final model are in quite good agreement with the input transition probabilities. It is not clear how a procedure such as the Bayes-updated would incorporate transition probabilities.

Summary

We have investigated with a synthetic data set the performance of cell-based stochastic simulation algorithms for modeling the spatial distribution of the categorical variable lithofacies, given data for well log reservoir lithofacies, "seismic" facies, variograms and facies associations. Table 4 summarized the global facies proportions and their relative standard deviations for each of the methods computed from 100 realizations. The optimization approach is clearly superior. The internal consistency of the model can be compared through entropy calculation (Gouveia [9]). Figure 21 compares the entropies for each of the methods computed from 100 realizations.

In a relative sense, low entropy at a location means less variability over the realizations for that location. The entropy is expected to be low near well and increase away from well control.

1. Simulations constrained by well data only Even with use of exhaustive anisotropic variogram models the simulations generally did a poor job of reproducing the complex spatial distributions. The level of conditioning required by the sequential indicator simulation is less than the one required by truncated Gaussian simulation. Such a result should be expected in view of the fact that the indicator simulation accounts for each lithofacies variogram, whereas in the truncated Gaussian approach a single "average" or representative variogram is employed in the simulation. One drawback of the Gaussian approach is the assumption that the spatial distribution of lithofacies can be fully characterized by a Gaussian process. The truncated Gaussian, however, by construct, honors the global lithofacies proportions.
2. Simulations constrained by well and seismic data Secondary data improve the reproduction of the true model considerably. When 25 depth profiles were employed as "hard" data, results were generally equivalent for the Bayes-updating and optimization (Figures 11 and 14), with a slightly better variogram reproduction obtained by the optimization approach ((Figures 12 and 16). The internal consistency, as indicated by the entropy plots (Fig. 21) is clearly better for these methods that utilize the interwell seismic facies than for the simulations without the seismic facies. And the Bayes method has somewhat lower entropy than the Optimization method. However, the Bayes method assumed that the seismic facies proportions are known at the resolution of the geologic model grid. This is strictly not the case; the Optimization method incorporates the true resolution of the seismic facies through the multigrid scheme and thus is a better reflection of the true uncertainty. This issue is being investigated further. The computational cost of the optimization method is higher

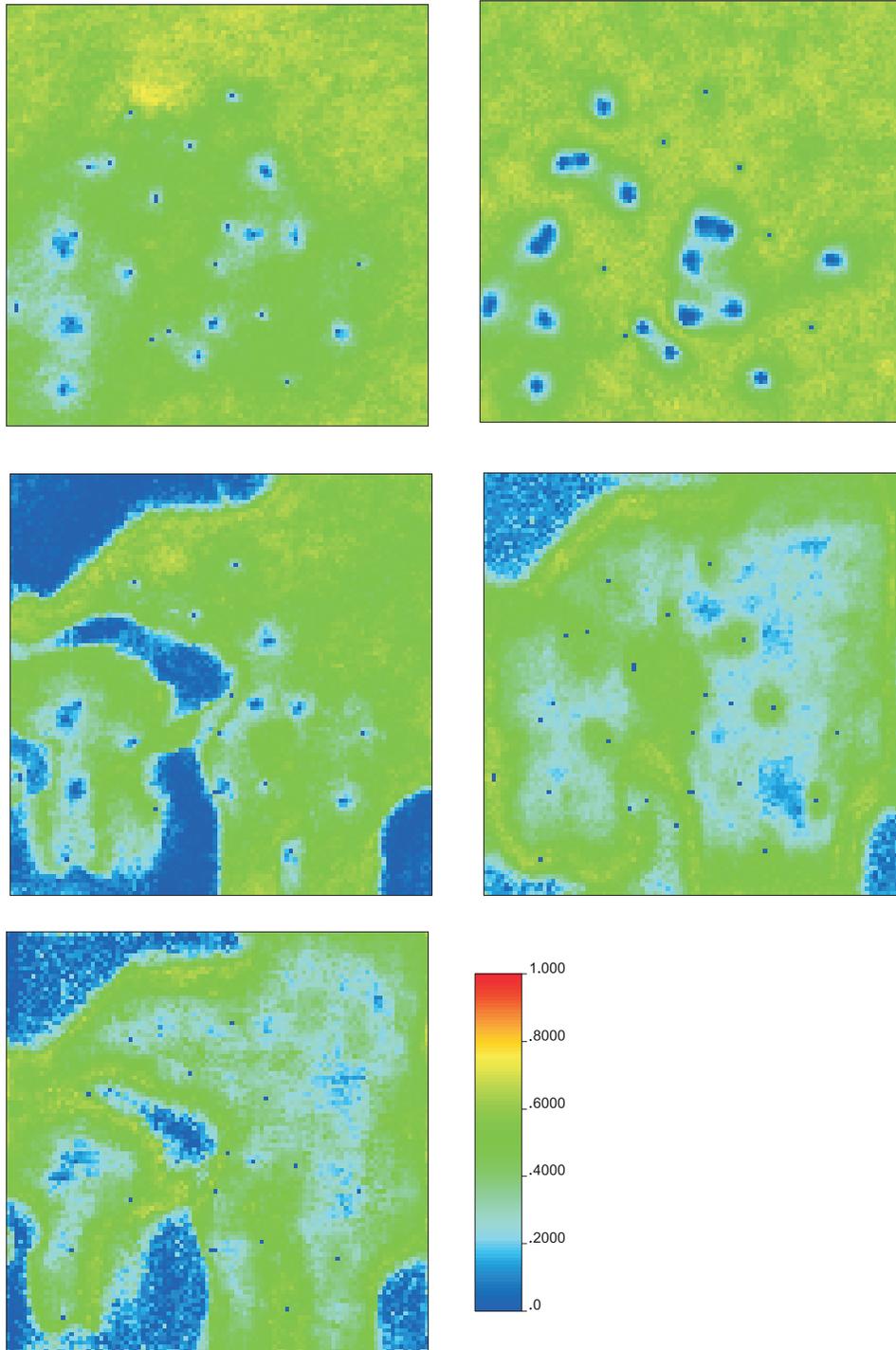


Figure 21: Entropy cube from 100 reservoir models computed (a) sequential indicator simulation, (b) truncated Gaussian simulation, (c) Bayes-updating, (d) optimization constrained by seismic data and (e) optimization constrained by seismic data and four-point transition probabilities

Facies	Reference	SISIM mean/rstd	GTSIM mean/rstd	Bayes mean/rstd	Opt mean/rstd	Opt w/transition mean/rstd
1	0.0178	0.0162/44.4	0.0131/20.6	0.0130/48.5	0.0153/11.8	0.0140/ 7.9
2	0.4282	0.4457/10.6	0.4329/ 3.6	0.3535/10.4	0.4357/ 2.1	0.4319/ 1.3
3	0.1782	0.2575/11.7	0.1832/ 2.5	0.2015/15.9	0.1774/ 1.9	0.1782/ 0.0
4	0.1240	0.0652/41.1	0.1271/ 3.5	0.0515/45.0	0.1204/ 6.0	0.1044/ 9.0
5	0.0165	0.0143/55.9	0.0168/ 5.4	0.0104/61.5	0.0134/11.9	0.0102/ 5.9
6	0.2352	0.2010/19.7	0.2271/ 5.9	0.3701/ 4.8	0.2378/ 4.5	0.2613/ 3.0

Table 4: Mean and relative standard deviation of facies proportion from 100 realizations

than that of the indicator algorithm. For example, a single realization required about 7.5 minutes *CPU* for the Bayes updated indicator method and 12 minutes *CPU* for the optimization method.

3. Simulations constrained by well, seismic data and 4th-order facies associations The transition probabilities further improve the model reproduction. These higher-order statistics could in principle be incorporated in geostatistical simulations via the so-called extended normal equations (Journel and Alabert [14], Guardiano and Srivastava [10] and Srivastava [20]) but it has not been so far applied to reservoir modeling problems.

A strength of the optimization approach is its flexibility to account for a larger number of input data sets than do the common geostatistical algorithms and to preserve the data volume support and uncertainty. The work presented here will continue to be developed along the lines of incorporating additional dynamic multiple-well production data and more information on geological relationships such as facies stacking patterns, nested cyclicities, and complex geometries and secondary data from numerical stratigraphic sections. An ultimate comparison of different techniques should be flow simulation *cdfs* from multiple realizations compared with flow simulation on the “true” model.

References

- [1] F. G. Alabert and G. J. Massonnat. Heterogeneity in a complex turbiditic reservoir: Stochastic modelling of facies and petrophysical variability. In *65th Annual Technical Conference and Exhibition*, pages 775–790. Society of Petroleum Engineers, September 1990. SPE Paper Number 20604.
- [2] R. A. Behrens, M. K. Macleod, T. T. Tran, and A. O. Alimi. Incorporating seismic attribute maps in 3d reservoir models. In *1996 SPE Annual Technical Conference and Exhibition Formation Evaluation and Reservoir Geology*, Denver, CO, October 1996. Society of Petroleum Engineers. SPE paper # 36499.
- [3] C. V. Deutsch. Cleaning categorical variable (lithofacies) realizations with maximum a-posteriori selection. In *Report 11, Stanford Center for Reservoir Forecasting*, Stanford, CA, May 1997.
- [4] C. V. Deutsch and A. G. Journel. The application of simulated annealing to stochastic reservoir modeling. In *Report 4, Stanford Center for Reservoir Forecasting*, Stanford, CA, May 1991.
- [5] C. V. Deutsch and A. G. Journel. *GSLIB: Geostatistical Software Library and User's Guide*. Oxford University Press, New York, 2nd edition, 1998.
- [6] P. M. Doyen, D. E. Psaila, L. D. Den Boer, and D. Jans. Reconciling data at seismic and well log scales in 3d earth modelling. In *1997 SPE Annual Technical Conference and Exhibition Formation Evaluation and Reservoir Geology*, pages 465–474, San Antonio, TX, October 1997. Society of Petroleum Engineers. SPE paper # 38698.
- [7] P. M. Doyen, D. E. Psaila, and S. Strandenes. Bayesian sequential indicator simulation of channel sands from 3-d seismic data in the oseberg field, norwegian north sea. In *69th Annual Technical Conference and Exhibition*, pages 197–211, New Orleans, LA, September 1994. Society of Petroleum Engineers. SPE paper # 28382.
- [8] A. Galli, H. Beucher, G. Le Loc'h, B. Doligez, and the HERESIM group. The pros and cons of the truncated gaussian method. In M. Armstrong and P. Dowd, editors, *Geostatistical Simulation*, volume 1, pages 217–233. Kluwer, 1994.
- [9] W. P. Gouveia, A. S. Cullick, and C. V. Deutsch. An optimization framework for reservoir characterization. In *1998 Annual SEG Convention*, pages 14–17, New Orleans, LA, September 1998. Society of Exploration Geophysicists.
- [10] Felipe Guardiano and R. M. Srivastava. Borrowing complex geometries from training images: The extended normal equations algorithm. In *Report 5*, Stanford, CA, May 1992. Stanford Center for Reservoir Forecasting.
- [11] H. Haldorsen and C. MacDonald. Stochastic modeling of underground reservoir facies. SPE paper # 16751, 1987.

- [12] H. H. Haldorsen, P. J Brand, and C. J Macdonald. Review of the stochastic nature of reservoirs. In S. Edwards and P. R. King, editors, *Mathematics in Oil Production*, pages 109–209. Clarendon Press, Oxford, 1988.
- [13] A. G. Journel and F. G. Alabert. Focusing on spatial connectivity of extreme valued attributes: stochastic indicator models of reservoir heterogeneities. SPE paper # 18324, 1988.
- [14] A. G. Journel and F. G. Alabert. Non-Gaussian data expansion in the earth sciences. *Terra Nova*, 1:123–134, 1989.
- [15] A. G. Journel and A. Arik. Dealing with outlier high grade data in precious metals deposits. In *Proceedings of Computer Applications in the Mineral Industry*, pages 45–72, Holland, 1988. Bolkema Publ.
- [16] A. G. Journel and E. H. Isaaks. Conditional indicator simulation: Application to a Saskatchewan uranium deposit. *Mathematical Geology*, 16(7):685–718, 1984.
- [17] H. Kupfersberger, C. V. Deutsch, and A. G. Journel. Deriving constraints on small-scale variograms due to variograms of large-scale data. *Mathematical Geology*, 30(7):837–852, 1998.
- [18] P. Kyriakidis, C. V. Deutsch, and M. L. Grant. Calculation of the correct normal scores variogram for truncated gaussian lithofacies simulation: Theory and fortran code. Technical report, Stanford Center for Reservoir Forecasting, Stanford, CA, May 1997.
- [19] G. Matheron, H. Beucher, H. de Fouquet, A. Galli, D. Guerillot, and Ch. Ravenne. Conditional simulation of the geometry of fluvio-deltaic reservoirs. SPE paper # 16753, 1987.
- [20] R. M. Srivastava. Iterative methods for spatial simulation. In *SCRF report*, Stanford, CA, May 1992.