

Geostatistics on Stratigraphic Grid

Antoine Bertoncello¹, Jef Caers¹, Pierre Biver² and Guillaume Caumon³.

¹ ERE department / Stanford University, Stanford CA USA; ² Total CSTJF, Pau France;

³ CRPG-CNRS / Nancy Université, Nancy France.

Abstract

For computational reasons, practical implementations of many geostatistical algorithms are designed for Cartesian grids. However, many applications in folded or faulted geological structures with complex stratigraphy, require unstructured grids containing blocks with varying support. The current practice deals with these grids by defining a physical space, where all structural geological features are incorporated (stratigraphic grid) and an original depositional space, where geostatistical algorithms are applied (typically a Cartesian grid). Therefore, these two spaces need to be linked, which is not straightforward. The traditional method consists of a direct mapping between the two spaces, fast and easy to complete. However, this method does not ensure the respect of the target statistics in the real space. In addition, it assumes that all the cells of the stratigraphic grid have the same volume. Hence, important global measures such as NTG or OOIP can become biased. The method introduced in this paper aims at overcoming these problems. It consists first, of sampling the stratigraphic grid with a regular lattice of points. These points are then mapped in the depositional space into a set of irregularly spaced points (due to the unfolding and unfaulting affecting the grid-geometry). Thus, the repartition of the points in the depositional space reflects implicitly the model geometry. Performing estimation/simulation on this set of points and then mapping back the result ensures reproduction of target statistics in the real space and properly accounts for the support effect. As a consequence, variances are correctly modeled, the estimated/simulated values are smoothed according to the volume of the cells and statistics are respected.

Introduction

Geostatistics is based on the random function concept, whereby the set of unknown values is considered as a set of spatially dependent random variables. The goal of geostatistics is to infer and sample this random function, conditionally to the available data. In addition to this set of data, another parameter influences the characteristics of the random variables: the support on which the variable is represented. Indeed, in petroleum or mining geostatistics, the work unit is a block, with a specific shape and size, whereas the well data are intrinsically smaller (generally considered as point support). It has, however, rarely been considered in the petroleum geostatistical algorithms, because of CPU limitations. Some analytical methods exist to model this change of support, but they are approximative and computationally intensive. In this paper, we define a new method to account for this change of support during property modeling.

Gridding and Reservoir Geometry Representation

1-Definition of the Grids

A critical step in reservoir modeling is to represent accurately the reservoir geometry. The volume of the reservoir is discretized by a set of grid blocks, which is used as a support to integrate data, perform the property modeling algorithms and, finally, to apply upscaling techniques to build a flow simulation grid. Different topologies can be used such as Cartesian grid, stratigraphic grid, unstructured grid... One type of grid commonly employed is the stratigraphic grid (Caumon, 2006). This is an irregular structured grid, locally unstructured when faulted and exclusively composed of hexahedral cells, distorted and indexed with three axis (stratigraphic coordinate): U and V parallel to the layering, and W perpendicular to U and V and representing the age of the deposits (Fig. 2).

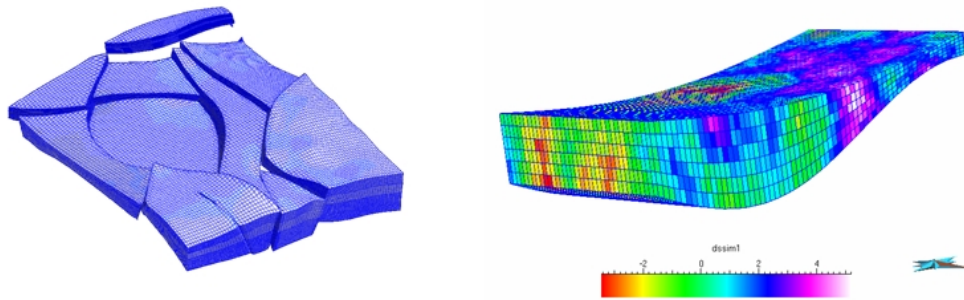


Figure 1: Example of stratigraphic grid. Stratigraphic grids fit the geometry of the reservoir, incorporating faults and respecting the stratigraphic architecture

2-Advantages of a Stratigraphic Grid

Geostatistical algorithms use intensively distance calculations between the data location and the block where the property is interpolated or simulated. This raises the problem of how these distances should be computed. If the subsurface geological structures are neither faulted nor folded after the depositional process, using an Euclidian distance is relevant. This is, however, rarely the case. That's the reason why a curvilinear coordinates system U V W is more appropriate (Mallet, 2004). Such curvilinear coordinates account for the shape of the horizons (which themselves control the geological continuity). It defines a Geo-Chronological (depositional space) model of the subsurface structures. The gridding defines then the stratigraphic heterogeneities of the reservoir and simplifies many problems: realistic petrophysical property modeling, fast flow simulation according to the heterogeneity of the reservoir, correlation between wells (Mallet (2004) and Caumon (2006)). This approach could be considered as a specific numerical representation of "Time-Stratigraphy" concept introduced by Wheeler (Wheeler, 1958), the Geochron model providing just a numerical generalization of this concept.

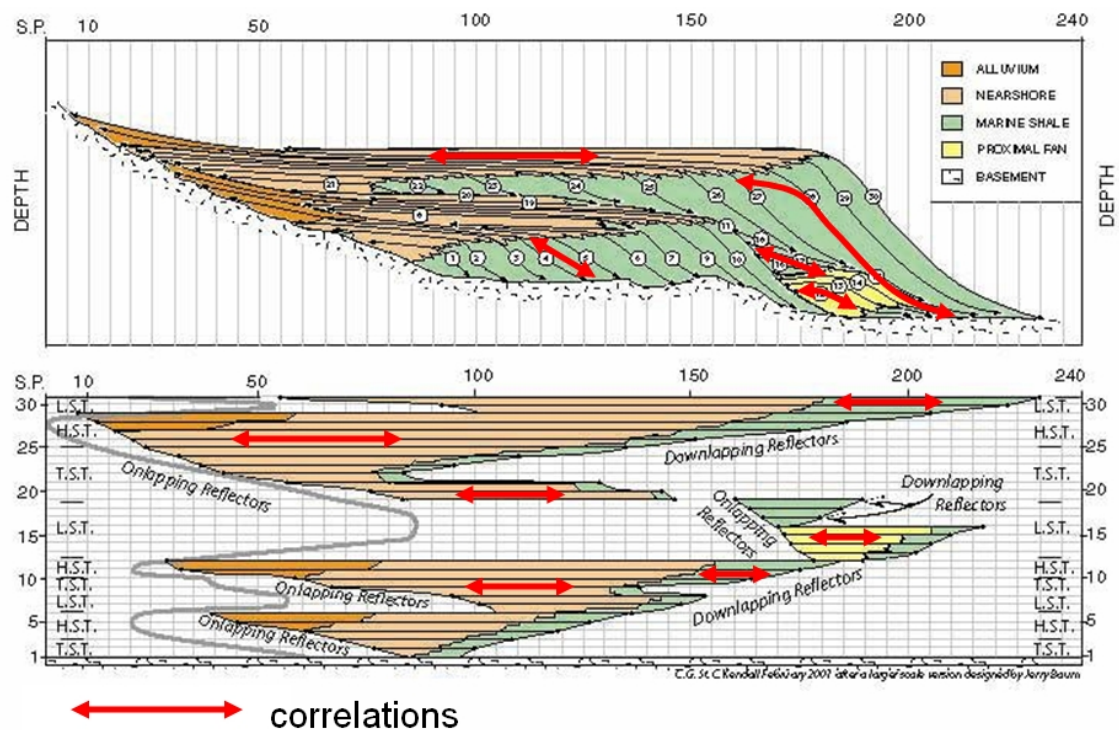


Figure 2: Generally, for reservoirs with complex stratigraphy, the correlations are curvilinear in the physical space. They become Euclidian in the depositional space, so working in this space make their calculation easier. Picture from Kendall (2003)

3-Volume of the supports in property modeling

Once the grid has been defined, all the data must be incorporated to perform stochastic simulation. Several sizes of support need to be considered (Caers, 2005). The data are generally assumed to be of point size (typically 1 inch x 1 feet), the cells of the grid on which the properties are assigned have a larger volume (typically 100 feet x 100 feet x 1 foot), seismic data have their own resolution (typically 300 feet x 300 feet x 10 feet). Well data are generally used to infer the histogram and variogram statistics, and used during the modeling process as hard data. In addition, this grid is not used directly for flow simulation. Upscaling is then needed to obtain a coarser grid, which can be run in reasonable time in a flow simulator. Generally, the size of the cells of the flow simulation grid is 300 feet x 300 feet x 10 feet. Integrating all the data in a single numerical model requires dealing with different scales. It is relatively easy to construct a reservoir model considering only one size of support (well data for instance), but this approach will ignore the contribution of the other data with different scale of observations. The real challenge is to use all the data available while accounting for their various supports.

Change of support: methods used and their limitations

1-Definition of the change of support

The support of $Z(u)$ is defined to be u , the region over which $Z(\cdot)$ is averaged. The change of support problem refers to making inference on block averages whose supports are different from those of the data. Ignoring the different scales of those data when constructing the property model leads to erroneous models. Indeed, in the presence of variables which average linearly, the change of volume from a volume v to a volume V (with $V > v$) entails the following on the property distribution (Journel and Huijbergts, 1978):

- The mean remains unchanged.
- The dispersion variance decreases (small scale variations disappears).
- The shape of the histogram tends to become more symmetric (due to the decrease of the variance).

2-Block Kriging

Block kriging was developped during the 60's for the mining industry (Journel and Huijbergts, 1978) (Goovaerts, 1997). The problem consists in the mineral grade estimation of selective mining unit blocks. Block kriging is a term for estimation of average z -values over a segment, surface or volume. Since the averaging is a linear process, the block value is defined as:

$$z_v(u) = \frac{1}{|v(u)|} \int_{v(u)} z(u') du'. \quad (1)$$

The correlations point-to-block are defined as:

$$C_{PB}(u_\alpha, v(u)) = \frac{1}{|v(u)|} \int_{v(u)} C(u_\alpha - u') du'. \quad (2)$$

The correlations block-to-block are defined as:

$$C_{BB}(v(u_\alpha), v(u_\beta)) = \frac{1}{|v(u_\alpha)||v(u_\beta)|} \int_{v(u_\alpha)} du \int_{v(u_\beta)} C(u - u') du', \quad (3)$$

and the covariance matrix (the inverse matrix needed to solve the kriging system) is the following:

$$\begin{bmatrix} C_{PP} & C_{PB} \\ C_{PB} & C_{BB} \end{bmatrix}$$

Two problems remain with this method. The first one is a computational time issue. Indeed, if the point to estimate is replaced by a block, the covariance matrix should include point to point correlations (redundancy between data), point-to-block correlation (correlations from data-to-blocks), and block-to-block correlation (correlations with the blocks of the neighborhood already simulated). If the cell sizes are not constant over the grid, which is the case for a stratigraphic grid, this covariance matrix must be recomputed for each grid cell.

The second problem is the estimation of the correlations between blocks. The principal method is a simple subsampling of the blocks into points and an averaging of the point-support covariance values. This approximation is non valid if the averaging of the property is non linear (permeability, acoustics properties). If the property is additive, the correlations point-to-block are defined as:

$$C_{PB} \simeq \frac{1}{n} \sum_{i=1}^n C(u - u_i), \quad (4)$$

and block-to-block are defined as:

$$C_{BB} \simeq \frac{1}{n \cdot N} \sum_{i=1}^n \sum_{j=1}^N C(u_j - u_i). \quad (5)$$

The problems here are (1) an high computational time and, (2) the possibility to obtain a non positive-definite covariance matrix (Journel and Huijbergts, 1978). The other main problem to handle is the change of shape of the histogram. Some analytical methods exist, but they work only for small reduction of the variance and specific distributions (Emery, 2007). In conclusion, the block kriging approach is difficult to implement, restrictive and slow to apply on a stratigraphic grid (where the blocks have different size and shape) and limited to variogram-based geostatistical methods. In the mining industry, simulations are directly used to evaluate the quantity of mineral inside a block and therefore the viability of the project. The notion of dispersion variance and change of shape is important to compute some cut-off or to know some confidence interval. The inference must be precise and in accordance to the volume. In the oil industry, these problems seems to be (misleadingly), at a first sight, not fundamental because geostatistical simulations are used as an input for the flow simulations, and not directly used for the oil recovery computation. Hence, problems of support are not immediately visible to the practitioner .

3-Integration of fine and coarse scale data

In the Earth Sciences, data with different support volumes of large and small scales must be integrated (tomographic data for instance Liu and Journel (2007)). Hence the challenge of integrating data with very different volumes of support has been already addressed. The method consists of calculating the point-to-block correlation by a Fast Fourier Transform (less CPU demanding) and not anymore by a discrete summation. This method proposed by Liu and Journel (2007) is relatively efficient. The technique, however, allows using only one specific algorithm which requires no normal score transform: Direct Sequential Simulation (DSS). With DSS, the simulation is directly performed in the original data space and does not call for any multi-Gaussian assumption. The main limitation is a poor reproduction of the histogram. Moreover, the simulated value is assumed to be point-support

(calculation of point-to-block covariance, not block-to-block covariance). This method can, in theory, be extended to compute block-to-block covariance but it will induce an important increase of the CPU cost. This method is also not applicable to discrete variables (e.g. facies) or multiple point type algorithms (whether discrete or continuous).

Approximation by a point-support simulation

In the actual practice of reservoir modeling, the problem of support has been largely ignored in property characterization, and the property modeling is generally performed on a point support, followed by simple interpolation schemes to create properties on stratigraphic grids.

1-Definition of the depositional space

Contrary to mining geology, petroleum geology often deals with sedimentation processes, and thereby, deals with time of deposition. Defining correlations following formations of the same age (stratigraphic correlations) is more relevant than relying on geographic distances because it links deposits of the same origin (coming from the same paleo-environment), and it by-passes post-sedimentary deformations (Mallet (2002) and Mallet (2004)). In this space, the vertical axis corresponds to the age of deposition (time axis). From a stratigraphic grid, the corresponding grid in the depositional space has constant cell size, the boundary between each layer corresponds to an iso-time line.

2-Link between the physical space and depositional space

The current practice in geostatistics consists in using the parametric domain ($u \ v \ t$) to compute Euclidean distances whose images in the geological domain are curvilinear distances (Mallet, 2002). Two grids are used, one in the physical space, where all the structural and geological features are incorporated (stratigraphic grid) and one in the depositional space (cartesian grid) where geostatistical algorithms are applied (see Fig. 3). Therefore, these two spaces need to be linked, which is not straightforward. More precisely, it raises three main concerns: the first one is the ability to transfer consistently (hard and soft) data from the real space to the depositional space. The second one is to respect some target statistics defined in the physical space while the geostatistical property modeling is executed in the depositional space. The last problem requires accounting for the volumetric distortion in the Cartesian space, even if this is not explicitly represented (there is not information in the depositional space about the geometry of the model).

3-Direct mapping and point-support simulation

The traditional method consists of a direct mapping between the two spaces, which is fast and easy to complete. A one-to-one relationship is established between each grid cell of the stratigraphic (physical space) and cartesian grid (depositional space). The simulation is done on the Cartesian grid, assuming each block to be a point. This method introduces some distortions of volume which bias the results.

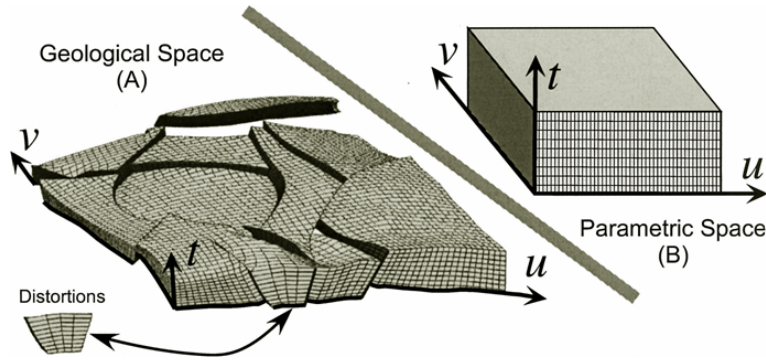


Figure 3: In the current property workflow, two grids are defined. One in the real space, incorporating the structural features of the variogram, and an other one regular in the depositional space (Mallet (2004)).

4-Problem of a direct mapping between the two spaces

Biasing of the statistics

The data are defined in the physical space. The histogram is valid therefore only in this space. On the contrary, the algorithms are performed in the depositional space, so the target histogram must be representative for this space. The transformation of the distribution is tedious. We can keep the same histogram (assuming that in both spaces, the histogram is the same (Fig. 4) or modify the histogram in accordance with the volume distortion at the well location (Fig. 5). In both cases, the volume distortions are not accounted for in the whole grid, hence some biases may still exist when the properties are mapped back into physical space. More over, in most existing software, the displayed statistics are computed in the depositional space and not in the real one. The statistics seems to be respected, which is misleading because the back-transform to the physical space has not yet been applied.

Biasing of the dispersion variance

The variance decreases when the volume of the block increases: the small scale heterogeneities are not anymore represented (Journel and Huijbergts, 1978). In the current workflow, this is not taken into account. The variance used by default is the quasi point support one (from the well data), which is very high compared to the blocks theoretical variance. This may have repercussions on Oil In Place calculations where are inflated variance of porosity and may lead to unrealistic P10 and P90 estimation.

Categorical variable: facies and support

A facies is defined by specific textural, chemical and biologic characteristics. This definition makes sense only reported to a specific volume of support. Then facies defined at the centimetric and metric scale are obviously not the same. Defining a facies with different sizes of support could also lead to inconsistencies. A property representing the proportion of each facies within a volume seems to be more appropriate.

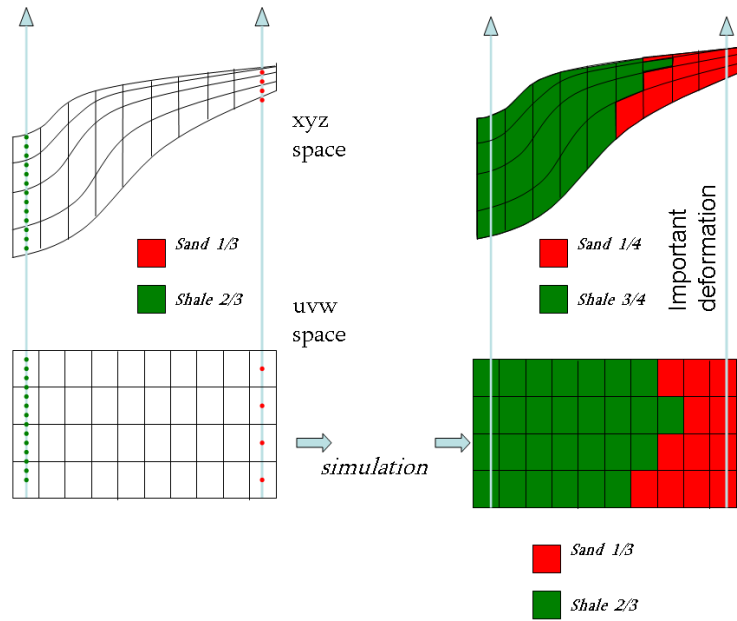


Figure 4: Geostatistical simulations performed in the depositional space. The data comes from the physical space. Assuming that the histograms are the same in the two spaces will not guarantee the respect of the target histogram in the real space

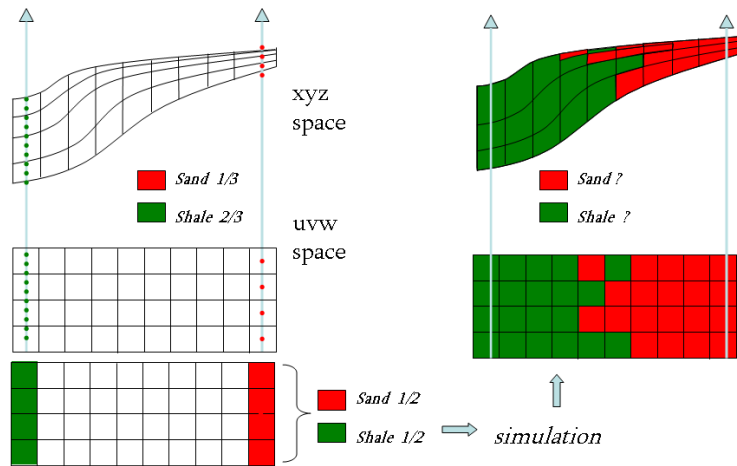


Figure 5: Correcting the histogram in accordance to the volume distortion at the wells locations will bias the results because it ignores the global distortion (it is just calculated from a few points).

Approach proposed

The next section introduces a new approach to perform property modelling in the depositional space while accounting for the volume distortion.

1-Idea

To properly account for support, our idea consists of a sampling of the stratigraphic grid with points and a mapping of the sampled points in the depositional space. Once in the depositional space, the points carry information about the stratigraphic grid geometry, which is important for including geometry and structural information into property modeling. Performing geostatistical simulations on these points allows to take into account the change of support and volume distortion. The main problem to solve is then to keep an acceptable computing time.

2-Algorithm Steps

Sampling

The main purpose of the sampling is to handle the geometric variability of the grid cells by points. In general, the main variation in the cells shape and volume is in the vertical direction. Then, it makes sense to sample only in this direction (Fig. 6).

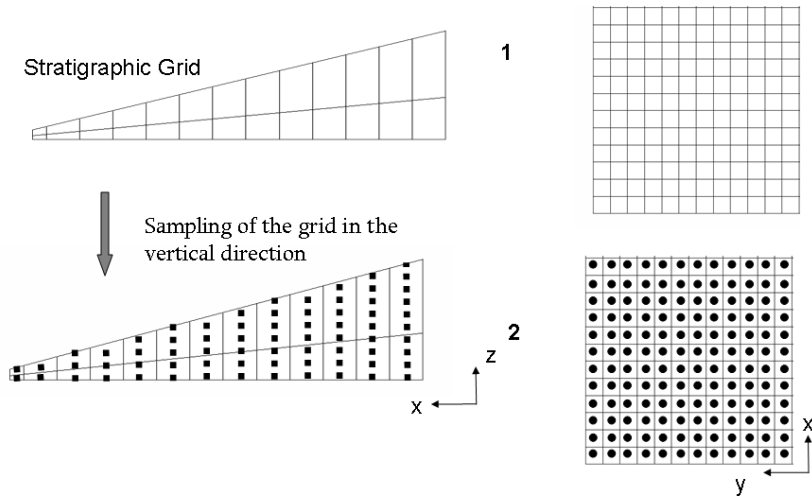


Figure 6: The sampling is done only in the vertical direction, in the direction where the distortion is the most important. For tartan grids, the sampling could be made in al three directions, incurring higher computational time.

Mapping of the points

The sampled points are mapped in the depositional space. The parametric function used here comes from Mallet (2002). It consists of the calculation, from the vertical coordinate of each point, of the corresponding relative time of deposition. The spread of the point cloud and their density in the depositional space are related to the size and the shape of the cells in the real space (Fig. 7).

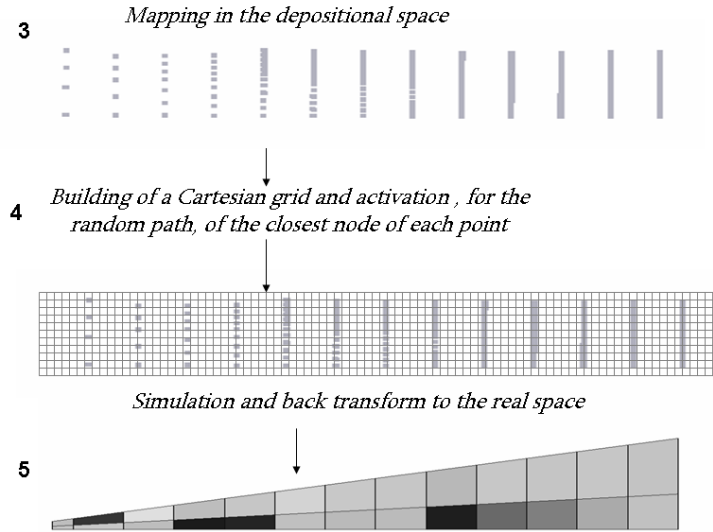


Figure 7: Once mapped, the points are inside a parallelepiped. In the depositional space, the information about the geometry of the stratigraphic grid is represented by the arrangement and the spatial density of the points.

Simulations

A very fine Cartesian grid is built to the background and each sampled point is associated to its closest node on the grid, defining then a sparse random path. Simulations are directly performed on that grid, which is less CPU demanding than simulating directly on the sampled points. The time of simulation depends on the geometry of the grid. Indeed, the sampling step is defined by the thickness of the smallest cells. So, if the grid has both very small and very large cells, simulation time may become large.

Back Transform

Each cell of the structural grid contains a corresponding set of points in the depositional space. Once the simulation is completed, the point values are averaged, linearly or non-linearly, into block value. The method of averaging depends on the property modeled. More precisely, not all the upscaling techniques could be applied here because the blocks are defined by a discontinuous set of points.

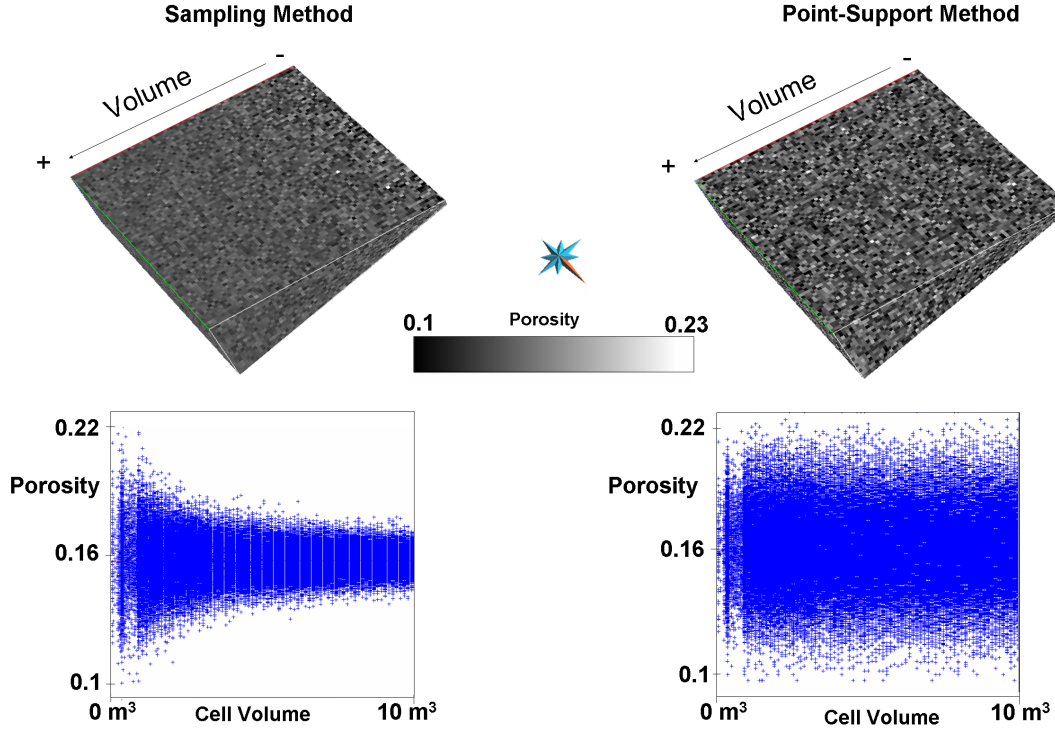


Figure 8: The fluctuations of the block values decrease as the block size increases. The visual result is a smoothing of the property (in accordance with the block size). The point-support method (SGS) does not reproduce this tendency.

Results

This method has been implemented for Sequential Gaussian Simulation (SGS, Goovaerts (1997)) and for Sequential Indicator Simulation (SIS, Goovaerts (1997)).

1-Decrease of the dispersion variance

An increase in the cells size leads to a smoothing of the property (decrease of the dispersion variance). The averaging performed inside each block tends to dampen the fluctuations between extremes values. The intensity of this smoothing effect is directly correlated to the number of sample points inside the block (i.e. the volume). Two tests, using SGS, have been performed on a specifically designed grid, where the cells become increasingly thinner toward the pinch-out. The first one, a simulation with a pure nugget effect, shows a fast decrease of the dispersion variance as the cells size increases (Fig. 8). The visual effect is a strong smoothing of the property. For the second example (Fig. 9), the variables are now correlated through an exponential variogram. The smoothing effect is still clearly visible, but less marked. This is induced by the presence of the variogram. Indeed, when the range of the variogram increases, the relative size of the support (compared to the scale of the geological phenomena) decreases. Inside each block, the sample points values are then, more correlated and

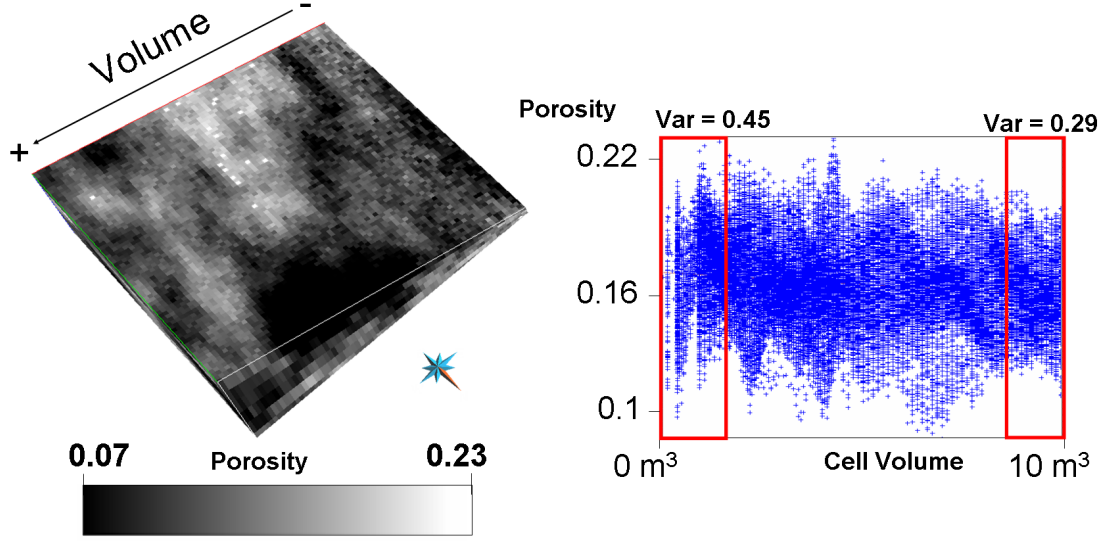


Figure 9: For a exponential variogram, the smoothing effect due to the increase of the block size is clearly visible. The point support variance is 1, the final block support variance is 0.4.

the values are closer from one to each other. The smoothing effect of the averaging is reduced, the impact of the change of support becomes less significant.

2-Change in the histogram shape

In theory, the shape of the distribution associated to each block is defined partially by its volume. If a large number of geostatistical simulations (100 simulations in this case) are performed and a histogram of the cell values is computed, we see clearly that the histogram tends to be more symmetric for high volume blocks (Fig. 10). This is due to the decrease of the dispersion variance.

3-Influence on the OIP calculation

The oil in place has been computed over 400 unconditional simulations (Fig. 11).

$$OIP = \sum_{grid} V_{block} S_o \Phi_{block}, \quad (6)$$

With $S_o = 0.8$ and Φ_{block} the simulated porosity for the block.

On the right, the algorithm used to obtain this histogram is a point-support algorithm (SGS). On the left, the proposed sampling method has been performed. It turns out that ignoring the volume of the supports leads to an overestimate of the P10-P90 interval. For SGS, the variance associated to each block is the point variance, which is higher than the realistic block variance. Hence, more variability is induced between realizations, which explains the larger dispersion between Oil In Place volumes.

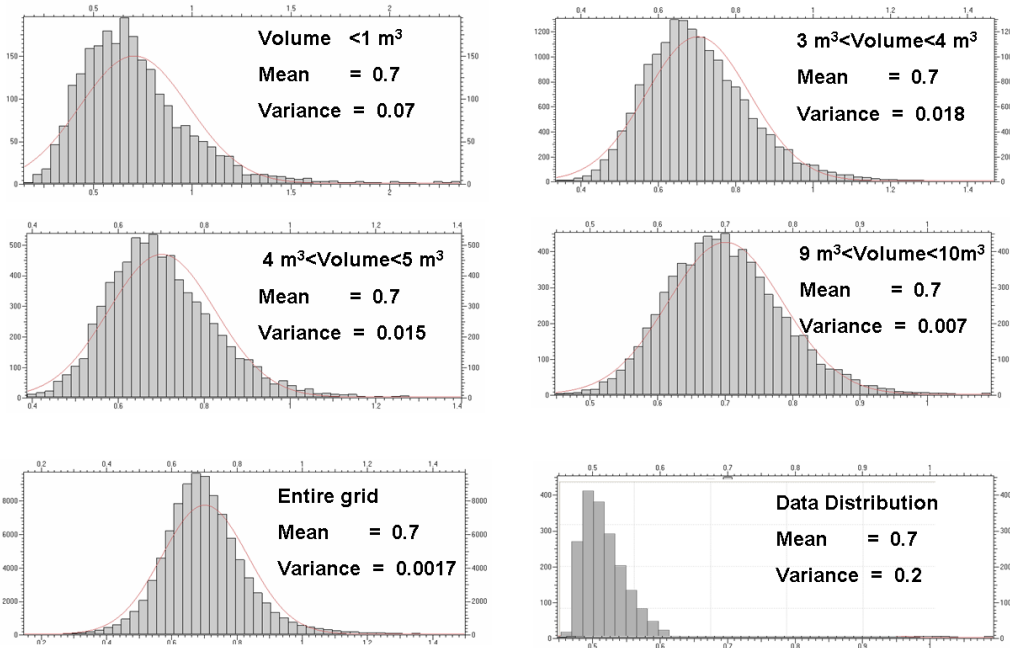


Figure 10: Change of the histogram shape. When the volume of the block increases, the associated distribution becomes more symmetric. The target and final distribution are different, because the variable is not represented at the same scale (quasi point support for the target distribution, block support for the final distribution)

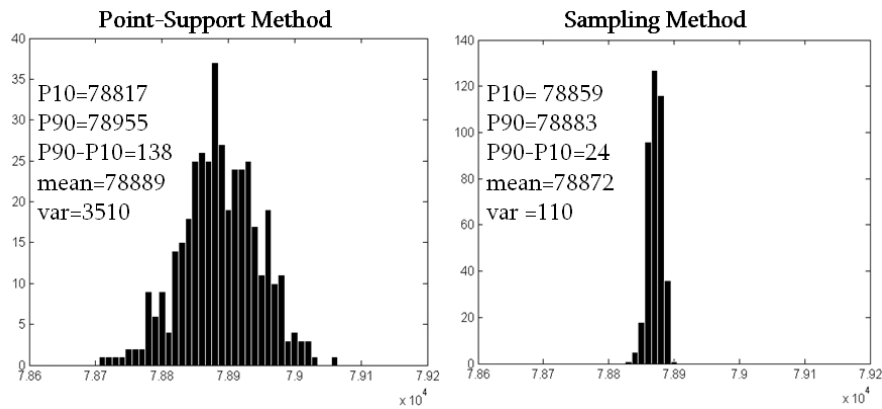


Figure 11: Oil In Place Calculation for four hundred runs. The classic method gives a extreme P10 and P90 because the global variance on the grid is assumed similar as the point-support one.

4-Reproduction of the Statistics

We construct another example where the same grid is modeled with a binary variable indicating sand and shale (using SIS). In this case, due to the configuration of the wells, the cells filled by the sand are globally thinner than the cells filled by the shale (Fig.12). The point support algorithm does not account for this trend and then, induces a systematic bias in the final proportion of sand. Indeed, the back-transform (from the depositional space to the real one) tends to decrease preferentially the volume of the cells filled by the sand (Fig.4 and Fig.5). The target proportion, honored in the depositional space, is no longer reproduced in the real space. With the sampling method the statistics are reproduced at the point support level. Because the arrangement of the points is representative of the geometry of the grid, the statistics are respected according to the volume of the cells (Table 1). More over, the facies are now defined at the point-support scale. The block property value is an average of the corresponding sample-point values. Hence, in this case, the sampling method provides in each block a proportion of shale and sand (continuous variable).

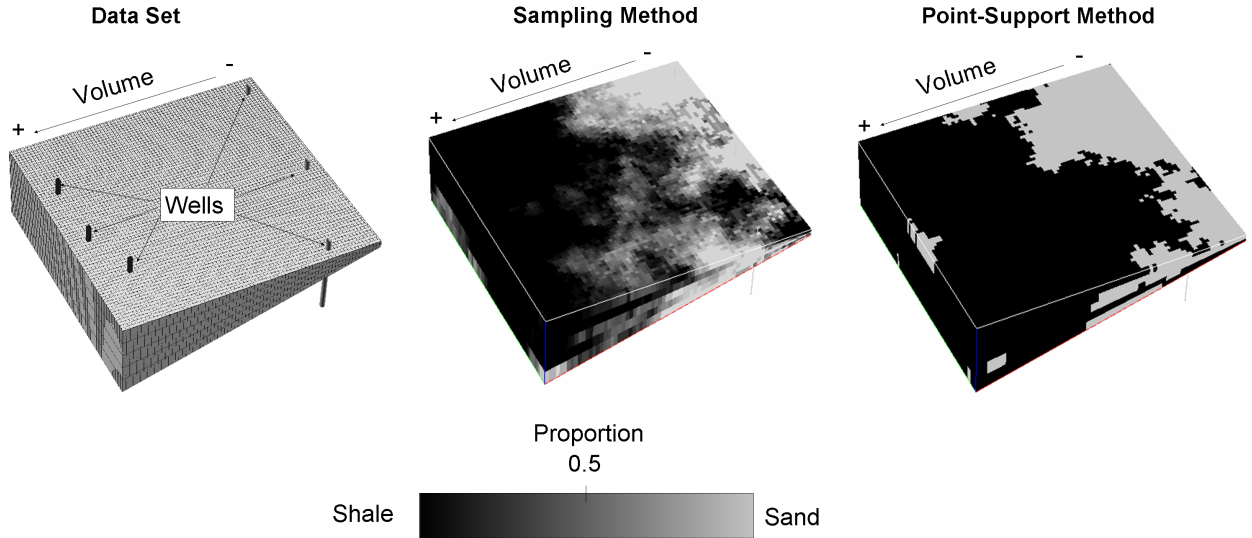


Figure 12: The volumes of the cells filled by the shale are larger than the ones filled by the sand. During the property modeling, considering or not this trend will modify the final result (Table 1).

5-CPU Costs

This method is more CPU demanding than when simulation is directly performed on the grid. For example, the number of cells in the grid is 33750 and the simulation time is less than 5 seconds. The number of points used to sample the grid is 427500 and it increases the simulation time to 79

Sampling Method	real 1 0.248	real 2 0.251	real 3 0.249	real 4 0.247	real 5 0.250	real 6 0.249
Target Sand Proportion = 0.250	real 7 0.248	real 8 0.252	real 9 0.261	real 10 0.255	real 11 0.260	real 12 0.255
Point Support Method	real 1 0.132	real 2 0.128	real 3 0.121	real 4 0.127	real 5 0.129	real 6 0.135
Target Sand Proportion = 0.250	real 7 0.122	real 8 0.113	real 9 0.124	real 10 0.131	real 11 0.141	real 12 0.124

Table 1: Comparison of the sand proportion obtained with the sampling and point-support method. In this case, the cells filled with shale are preferentially thicker (Fig 12). The sampling method honors the target statistics, which is not the case of the commonly used point-support method. It induces indeed a systematic bias, controlled by the global geometric deformation between the stratigraphic grid and the corresponding Cartesian grid in the depositional space.

seconds. The simulation time is directly correlated to the number of points. This number is defined by the size and the geometric complexity of the grid.

6-Limitations of the method

These results emphasize the importance of considering the volume distortion during the property modeling: ignoring these problems leads to a bias of the property variance and mean, which are key parameters in a reservoir uncertainty assessment. However, because the simulation is performed in the depositional space, the properties simulated depend on the interpreted depositional process. Any properties related to a post-sedimentation event can not be represented directly by this method. More-over, multiple-point statistics simulations have not been tackled. In this case, the training image will have to represent the geological phenomena in the depositional space and not anymore in the physical one.

Conclusion

We have started addressing the problem of support in petroleum geostatistics by noticing that the use of stratigraphic grid to represent complex geometry introduces some bias in the statistics. This sampling method allows to solve the problem for variogram-based geostatistics. However, further investigation will be needed to integrate multipoints statistics in this project.

Acknowledgment

This research work was performed in the frame and with the financial support of the Stanford Center for Reservoir Forecasting (SCRF) and Total, the affiliate companies are hereby acknowledge.

Bibliography

- Caers, J. (2005). *Petroleum Geostatistics*. Society of Petroleum Engineers. 96 pages.
- Caumon, G. (2006). Stratigraphic Modeling : Review and Outlook. In *paper presented at the 26th Gocad Meeting (Nancy, France, 6-9 June)*.
- Emery, X. (2007). On Some Consistency Conditions for Geostatistical Change of Support. *Mathematical Geology*, 39(2).
- Goovaerts, P. (1997). *Geostatistics For Natural Resources Evaluation*. Oxford University Press, NY.
- Journal, A. and Huijbergts, C. (1978). *Mining Geostatistics*. Academic Press, NY.
- Kendall, C. (2003). Course Note Sedimentological Processes Modeling. University of South Carolina.
- Liu, Y. and Journal, A. G. (2007). Geostatistical Integration of Coarse and Fine Scale Data, BGEOST: Applications and Results. In *Proc. 20th Annual SCRF Meeting, Stanford CA, May, 7-8*.
- Mallet, J. (2002). *Geomodeling*. Oxford University Press, NY, NY.
- Mallet, J.-L. (2004). Space-Time Mathematical Framework for Sedimentary Geology. *Mathematical Geology*, 36(1):1–32.
- Wheeler, H. E. (1958). Time-Stratigraphy. *Bull. of the AAPG*, 42(5):1047–1063.