Non-parametric diagrams for plurigaussian simulations of lithologies

D. Allard, D. D'Or, P. Biver and R. Froidevaux

Abstract Although truncated plurigaussian models have been proposed more than two decades ago, their practical use has been limited by three major impediments: first, the difficulty to relate underlying Gaussian variables to physically interpretable processes; second, the estimation of a truncation diagram from data or concepts is not intuitive for geologists; third, updating a truncation diagram according to current proportions could be a challenge (current practice is to simplify truncation diagrams with rectangular domains). When bivariate auxiliary information and lithofacies data are available at wells, we present an approach for using a bi-Gaussian model allowing for very general shapes. The general idea is to relate the Gaussian variables to auxiliary variables using kernel regression and to build assignation diagrams by calibration with lithofacies observed at wells.

1 Introduction

The basic idea behind truncated and plurigaussian simulations is to start out by simulating one or more Gaussian random fields at every grid point in the study domain and to use some simple rule to convert these Gaussian values into lithotype indicators. With truncated Gaussian simulations (Matheron et al., 1987), lithofacies must ordered and, on any fine resolution grid, lithofacies that are not successors in the

Denis Allard
INRA, UR546 Biostatistique et Processus Spatiaux (BioSP), Avignon, France,
e-mail: allard@avignon.inra.fr

Dimitri D'Or and Roland Froidevaux
Ephesia Consult, 9 rue Boissonas, Geneva, Switzerland,
e-mail: dimitri.dor@ephesia-consult.com

Pierre Biver
Total SA, CSTJF avenue Larribau, Pau, France,
e-mail: Pierre.biver@total.com

stratigraphic sequence will never be in contact. This feature proves to be too restrictive, for example if there is no ordered sequence in the facies, or if certain facies can be in contact with more than two facies. To overcome these limitations, plurigaussian simulations was proposed (Galli et al., 1994; Le Loc'h et al., 1996). Later, theoretical and algorithmical advances have been made in order to handle many underlying Gaussian random fields (Dowd et al., 2003; Xu et al., 2006; Emery, 2007) and to co-simulate continuous and categorical variables (Emery and Silva, 2009). Applications of plurigaussian simulations are numerous for modeling reservoirs but also in hydrogeology for modeling aquifers (Mariethoz et al., 2009).

In plurigaussian models, a set of p Gaussian random fields are used to define the lithofacies structure. Generally, $p = 2$, and we will stick to this choice in the remaining of this work. Lithofacies are then assigned according to a rock type rule (Armstrong et al., 2003). Typical rock type rules separate the bi-Gaussian domain $\mathbb{R}^2$ along vertical and horizontal lines. The precise layout of these lines define the type of the chosen model. Figure 1 represents one example of a rock type rule along with one realization of the categorical variable "lithofacies". The realization of the lithofacies variable shows a pattern in direct relation with this rock type rule, which is here parametrized by two values. Armstrong et al. (2003) distinguishes four steps in a plurigaussian simulation approach.

1. Choosing the model type, that is the rock type rules that best suit to the problem at hand. Since the correspondence between rock type rules and patterns of simulated lithofacies is not at all obvious, Armstrong et al. (2003) provided a catalog of examples of rock type rules with simulated patterns. This first step is crucial. Yet, at the same time i) quantitative rules for choosing a model type are limited to information on lithofacies contacts; ii) in the 2-dimensional Gaussian plane, boundaries between lithofacies must usually be parallel to the axes.
2. Estimating the parameter values. The exact value of the thresholds (i.e. the exact location of the horizontal and vertical lines on the (Y_1, Y_2) plane) is in straightforward relation to the target proportions, which can vary on the simulation grid. Knowing the mathematical relation between indicator variograms and variograms of the underlying Gaussian variables, it is possible estimate the variogram parameters.
3. Generating Gaussian values at the conditioning points. This task is usually performed using a MCMC simulation algorithm, such as the Gibbs sampler.
4. Simulating Gaussian values at each grid node with a usual Gaussian simulation algorithm.

Finally, the Gaussian values are transformed into the lithofacies categorical variable using the rock type rule applied with the local parameters when necessary. The drawbacks of this approach are multiple: i) when there are more than a very few lithofacies (say, if $K > 5$), it is extremely difficult to chose a model type and rock type rules; ii) estimating the parameters of the variogram becomes extremely cumbersome; iii) since there is no direct interpretation of the underlying Gaussian variables, geologists are sometimes reluctant to use this method.

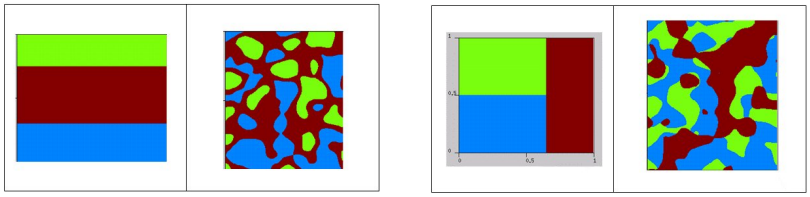

Fig. 1 Examples of rock type rules for truncated bi-Gaussian simulations.

We therefore propose a method which overcomes some of these limitations when auxiliary variables, such as seismic attributes I_p (acoustic impedance of compressive waves) and P_r (Poisson ratio) are available.

1. The underlying Gaussian variables are related to the seismic attributes, when they exist. We thus transform the seismic attributes into Gaussian scores. The parameters of their bivariate variograms will be estimated from these Gaussian scores, hence greatly simplifying this estimation step.
2. The rock-type rule will be built from the simultaneous knowledge of the seismic attributes and the lithofacies at wells and/or drill-holes. By doing so, the geologists do not need anymore to choose among a catalog of rock model. The rock-type rules do not need to follow simple geometry such as lines parallel to the axis, etc. The rock-type rules, also called assignation diagram below, will be automatically tailored to the data and to any vector of target proportions. It can handle a large number of lithofacies (tested with up to 11 lithofacies, on a real case study).

Note that in absence of auxiliary variables, synthetic variables can be used to build assignation diagrams. This approach will be illustrated In Section 6.

In the next Section, we will show below how the rock-type rules are defined from the data. In Section 3, we present an algorithm for honoring any vector of target proportions. An illustration on simulated data with four lithofacies is presented in Section 4. The main steps of plurigaussian simulation with our approach are then detailed in Section 5. Finally, a real case study with 11 lithofacies is presented in Section 6.

2 Non-parametric assignation diagrams

After transformation to uniform scores, the bivariate auxiliary variable (seismic attributes, synthetic variables defined by the user, ...) are denoted $\mathbf{U} = (V,W) \in \mathscr{D} = [0,1] \times [0,1]$. Note that other transforms, such as Gaussian transform could also be used, but this point is of secondary importance. Uniform transforms have been chosen here, because with this transform the area of lithofacies on $\mathscr{D}$ is equal to its probability. We will denote $\mathbf{u} = (v,w)$ a position vector of the variables $(V,W) \in \mathscr{D}$.

We consider that these values are available at every point $\mathbf{x}$ of a geological grid. A lithofacies function F assigns to each point $\mathbf{x}$ a value $k \in \{1,\dots,K\}$ corresponding to the lithofacies. Remember that we do not require the lithofacies to be ordered in any way. Neither do we impose a parametric shape for the arrangement of the lithofacies on the diagram.

The general idea is that the probability of each lithofacies is related to the (transformed) auxiliary variable. Mathematically speaking, we consider that there exist functions

$$P(F(\mathbf{x}) = k \mid \mathbf{U}(x)) = f_k(\mathbf{U}), \tag{1}$$

which are the conditional probabilities of the lithofacies given $\mathbf{U}$.

From the simultaneous knowledge of the auxiliary variable and the lithofacies at data points $\mathbf{x}_1,\dots,\mathbf{x}_n$ (e.g. measured at wells), a kernel regression estimate (Silverman, 1986) of these conditional probabilities is

$$\hat{f}_k(\mathbf{u}) = P(k \mid \mathbf{u}) = \frac{\sum_{i=1}^{n} K_h(\mathbf{u}-\mathbf{u}_i) I_{[F(\mathbf{u}_i)=k]}}{\sum_{i=1}^{n} K_h(\mathbf{u}-\mathbf{u}_i)}, \tag{2}$$

where $i = 1,\dots,n$ is an index on all data and $I_{[A]}$ is the indicator fuction of event A, equal to 1 if A is verified and equal to 0 otherwise. The function $K_h(\mathbf{x})$ is the kernel function

$$K_h(\mathbf{x}) = \frac{K(\mathbf{x}/h)}{h}, \tag{3}$$

where $K(\mathbf{x})$ is an even function such that $\int_{\mathbb{R}^2} K(\mathbf{x}) d\mathbf{x} = 1$ and $\int_{\mathbb{R}^2} K^2(\mathbf{x}) d\mathbf{x} < \infty$. Common kernels are the Gaussian kernel $K(\mathbf{x}) = (2\pi)^{-1} e^{-||\mathbf{x}||^2/2}$ and the bi-squared kernel $K(\mathbf{x}) = 3(1-||\mathbf{x}||^2)^2/\pi I_{(||\mathbf{x}||\leq 1)}$.

From Eq. (1), it is easy to check that

$$\sum_{k=1}^{K} \hat{f}_k(\mathbf{u}) = 1$$

for all $\mathbf{u}$ in $\mathscr{D}$. Up to border effects, one can also check that

$$\int_{\mathscr{D}} \hat{f}_k(\mathbf{u}) d\mathbf{u} = n_K/n,$$

the observed frequency of data with lithofacies k.

From these functions, it is quite natural to define an assignation diagram (i.e., a rock-type rule) by assigning to each value of the seismic variables $\mathbf{u}$ the lithofacies with the highest conditional probability, i.e. $F_T(\mathbf{u}) = \arg\max_k(\hat{f}_k(\mathbf{u}))$. Note that this operation is not a truncation of the value $\mathbf{u}$ as it is the case in the usual plurigaussian approach, but it is a maximum probability assignation. It is, however, a rock-type rule. Since this operation is highly non-linear, it is quite clear that the proportion $p_k = \int_{\mathscr{D}} F_T(\mathbf{u}) d\mathbf{u}$ obtained from F_T will be quite different to n_k/n. Let

us thus consider the vector of values $(a_1,\ldots,a_K)$ which will multiply the conditional probabilities, thereby defining intensities $a_k\hat{f}_k(\mathbf{u})$. Note that these intensities are not conditional probabilities since $\sum_{k=1}^{K}\hat{a}_k f_k(\mathbf{u})$ is not necessarily equal to 1. The lithofacies now becomes

$$F_T(\mathbf{u}) = \arg\max_k (a_k\hat{f}_k(\mathbf{u})). \tag{4}$$

In order to verify specified proportions $(\pi_1,\ldots,\pi_K)$ one must thus find the factors $\mathbf{a}=(a_1,\ldots,a_K)$ such that

$$\int_{[0,1]\times[0,1]} I_{[F_T(\mathbf{u})=k]}d\mathbf{u} = \Phi_k(\mathbf{a}) = \pi_k. \tag{5}$$

In the following, we will impose, without loss of generality that $\sum_{k=1}^{K} a_k = 1$.

3 Finding an assignation diagram

Let us denote by $\boldsymbol{\Phi}(\mathbf{a})$ the vector of proportions obtained after assignation according to Eq. (4), i.e. $\boldsymbol{\Phi}(\mathbf{a}) = (\Phi_1(\mathbf{a}),\ldots,\Phi_K(\mathbf{a}))^t$. To solve the inverse problem in Eq. (5) we build the score function

$$S = \sum_{k=1}^{K}(\Phi_k(\mathbf{a})-\pi_k)^2 = (\boldsymbol{\Phi}(\mathbf{a})-\boldsymbol{\pi})^t(\boldsymbol{\Phi}(\mathbf{a})-\boldsymbol{\pi}) \tag{6}$$

which is to be minimized. The inverse problem now becomes a minimization problem which is solved by a modified Gauss-Newton algorithm with a Marquardt regularization (Marquardt, 1963). The algorithm is briefly presented in Appendix A. Numerous experiments, not presented here have shown the following properties:

- The algorithm almost always converges when initialized with $\mathbf{a}=\pi$. The rare cases with lack of convergence were observed when one of the output proportions is imposed to be very close to 0. In this case, the algorithm gets sometimes stuck. When this happens, the algorithm must be started again with a different starting point.
- Perhaps surprisingly, the vector $(a_1,\ldots,a_K)$ corresponding to the target proportion $(\pi_1,\ldots,\pi_K)$ is unique. This ensures that the assignation diagram corresponding to imposed proportions is unique.
- Starting from an initial solution corresponding to a vector of proportions $(\pi_1,\ldots,\pi_K)$, the algorithm converges to a new vector of proportions $(\pi'_1,\ldots,\pi'_K)$ in only very few iterations when $\sum_{k=1}^{K}(\pi_k-\pi'_k)^2$ is small.

In many applications, the target proportions cannot be considered as stationary. A different vector $(a_1(\mathbf{x}),\ldots,a_K(\mathbf{x}))$ must thus be computed for each different vector of target proportions $(\pi_1(\mathbf{x}),\ldots,\pi_K(\mathbf{x}))$. Clustering similar target proportions in

groups and applying the above minimization problem in some optimized way allows to use this approach on large to very large grids.

4 Synthetic data

500 data location are drawn uniformly in $[0.1,0.9] \times [0.1,0.9]$. A category k, from 1 to 4 is assigned randomly with a probability proportional to the inverse squared distance to, respectively, $(0.25,0.25)$, $(0.75,0.25)$, $(0.25,0.75)$, $(0.75,0.75)$. The resulting conditional densities $f_1(\mathbf{u}),\ldots,f_4(\mathbf{u})$, computed with a bi-squared kernel and $h = 0.48$ are shown in the top panels of Figure 2. The solution vectors $\mathbf{a}$ corresponding to four different vectors were computed with $S < 10^{-5}$. Results are reported Table 1. The corresponding assignation diagrams are shown in the bottom panels of Figure 2.

Table 1 Target proportions and solution vectors for the simulated data

$\boldsymbol{\pi} = \boldsymbol{\Phi}(\mathbf{a})$	$\mathbf{a}$
$(0.1,0.2,0.2,0.5)$	$(0.069,0.364,0.321,0.246)$
$(0.2,0.1,0.5,0.2)$	$(0.127,0.365,0.404,0.104)$
$(0.7,0.1,0.1,0.1)$	$(0.563,0.202,0.178,0.057)$
$(0.25,0.25,0.25,0.25)$	$(0.139,0.394,0.353,0.114)$

There is no direct correspondence between a value a_k and its corresponding output proportion $\Phi_k(\mathbf{a})$. For example an output proportion of 0.1 for the second category is obtained with $a_2 = 0.365$ in the second case while it is obtained with $a_2 = 0.202$ in the third case. The value a_k does not only depend on the corresponding target proportions but it depends on the complete vector of target proportions.

It is interesting to note how each domain is stretched or shrunk as the proportions vary. See for example how the purple lithofacies is mainly located in the bottom left corner. As its proportions increases, it expands in the regions where the data points are mostly found, i.e. in the upper left corner.

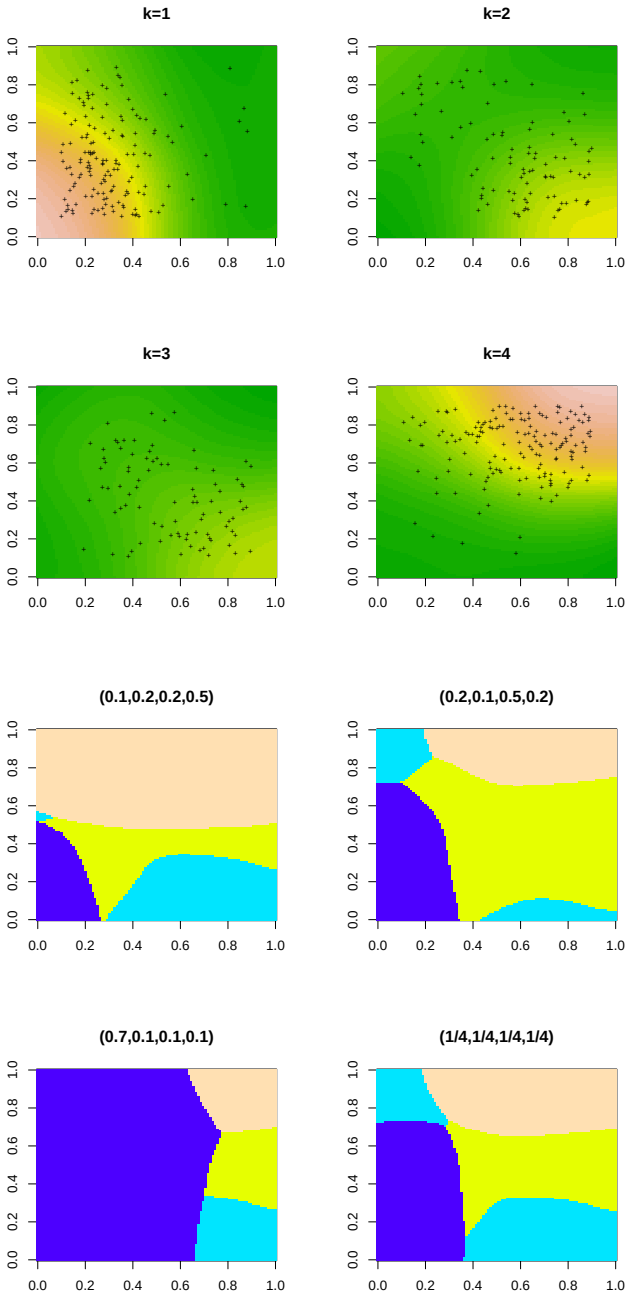

Fig. 2 Conditional densities of simulated data (top) and assignation diagram for four contrasted vector of proportions (bottom). Purple: $k = 1$; Blue: $k = 2$; Green: $k = 3$. Orange: $k = 4$.

5 Plurigaussian simulation with assignation diagram

The full work-flow for performing a plurigaussian simulation with the non-parametric assignation diagram described above is not very different to the usual work-flow.

1. The seismic attributes are first transformed into normal scores, from which a bivariate variogram model is estimated/fitted.
2. The normal scores are then transformed into uniform scores.
3. In absence of auxiliary variables, step 1. and 2. are skipped. Instead, the user defines a set of synthetic data points in $\mathbf{U} \in \mathscr{D}$, along with the desired lithofacies attribute.
4. From the simultaneous knowledge of the auxiliary variables and the lithofacies information at the data points, the conditional probabilities $\hat{f}_k$ are computed with kernel regression estimate (Eq. (2)).
5. Bi-variate Gaussian values are simulated at lithofacies conditioning points with the use of a Gibbs Sampler. The simulated Gaussian value must correspond to the conditioning lithofacies for the local assignation diagram (if proportions are non-stationary).
6. A bivariate Gaussian Random field is then simulated conditionally to all Gaussian values, transformed from data, or simulated at the previous step.
7. Lithofacies values resulting from the application of the local assignation diagram to the simulated Gaussian values are found.

6 Case study

The approach has been tested on a real field case in Western Africa. It is a complex facies model in a mixed carbonated platform environment influenced by complex progradational and retrogradational mechanisms. Eleven litho-facies are depicted on cores and propagated through the log interpretations of the large number of available wells.

Relevant geophysical attributes are not available for this reservoir; but however, from geological concepts, we have a precise idea of possible contacts between the litho-facies. For this reason, it has been decided to build from scratch the assignation diagram; in the unit square of uniform scores, synthetic data points are digitalized (Fig. 3, left panel). Kernel regression (Eq. 2) allows us to compute the eleven functions f(x). Figure 3 (right panel) shows the initial assignation diagram corresponding to $a_k = 1$ for $k = 1, \ldots, K$.

The probabilities of occurrence of the different litho-facies are also provided by the project; it consists in a set of eleven 3D properties depicting the 3D trends of probability for the eleven litho-facies; these 3D cubes are issued from geological conceptual zones definition mixed with litho-facies well data statistical analysis.

Using the algorithm of Appendix A (Gauss-Newton iterative process with Marquardt regularization), it is possible to modify the initial assignation diagram in

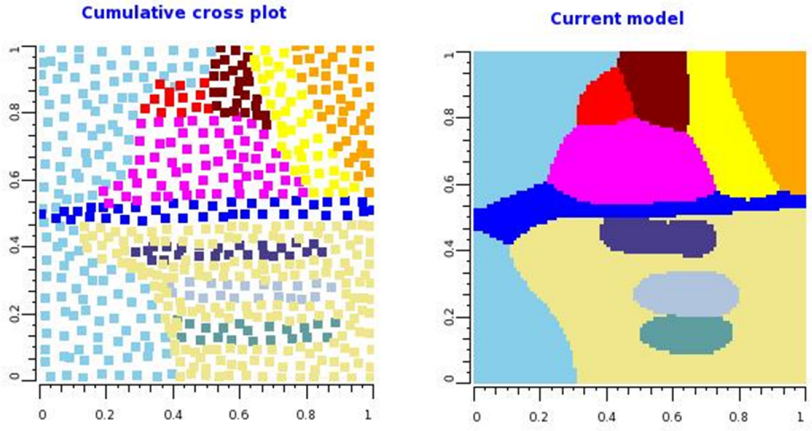

Fig. 3 Initial assignation diagram construction. Left: digitized point sets with facies assignation. Right: initial litho-facies polygons corresponding to $a_k = 1$ for $k = 1, \ldots, K$.

order to fit to local proportions. Two Gaussian random fields with Gaussian variograms (kilometric horizontal range and decametric vertical range) are then simulated on a geological grid. At each grid node, the local assignation diagram is used to select the simulated litho-facies as a function of the unit-square transformation of the local Gaussian values, thus obtaining a full non-stationary 3D geological model. Two 3D probability cubes and one realization of the litho-facies variable are illustrated in Figure 4. Non-stationarity of the proportions is fairly well reproduced.

To understand more clearly the deformation of the initial assignation diagram, a 2D view is presented for one particular layer of the reservoir (Figure 5). Note how the assignation diagram varies in order to reach target proportions that are completely different from those prevailing in initial assignation diagram. In particular, a striking feature of our approach is that, preferential contact between lithofacies are respected for all input vector of proportions, even though many of the 11 lithofacies have an imposed proportion set to 0.

7 Conclusion and discussion

We proposed a new approach for performing plurigaussian simulations. Instead of building tuncation diagram with very few parameters and limits between litho-facies mostly parallel to the axes, we propose to relate the rock type rules to auxiliary variables such as sesimic attributes, with the use of kernel regression. In abscence of relevant geophysical variables, synthetic assignation diagram can be built from conceptual ideas concerning litho-facies contacts (facies substitution diagrams), control points are digitized manually.

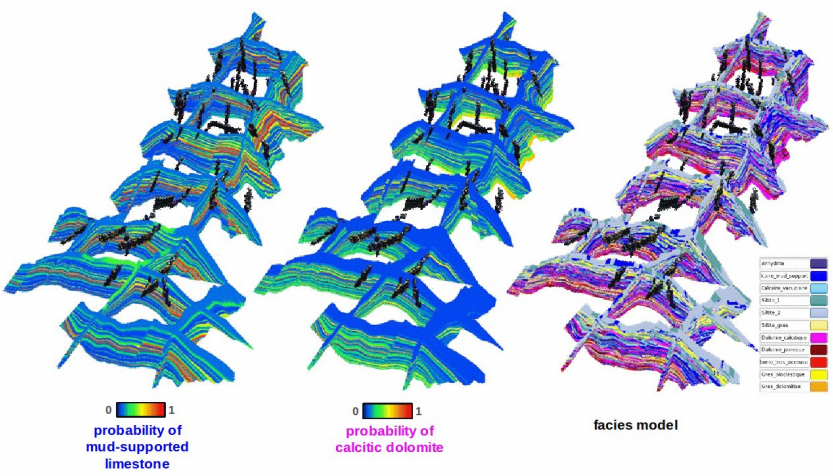

Fig. 4 3D view of the reservoir probability cubes (left and middle); and the corresponding litho-facies model (right).

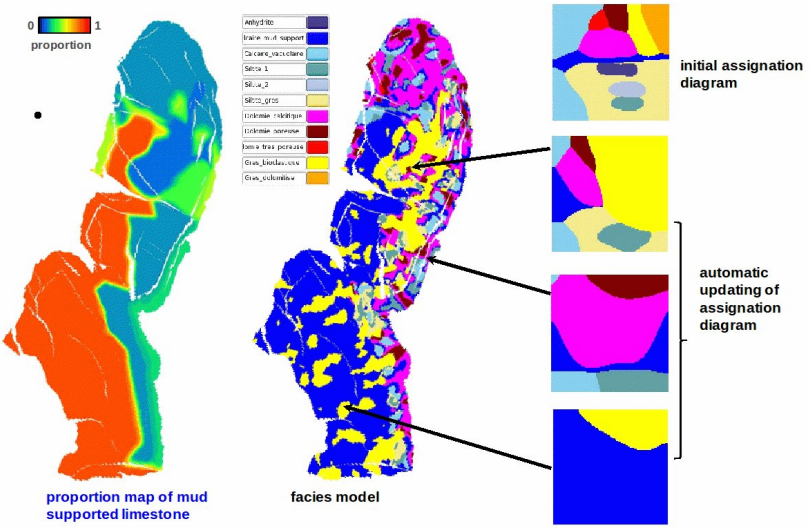

Fig. 5 2D view of the reservoir probability map and corresponding litho-facies model; corresponding deformations of the assignation diagram are depicted on the right.

The assignation diagram is then updated according to algorithm described in Section 3 to verify any given set of proportions.

This approach has been successfully applied for the geostatistical modelling of lithofacies on a complex mixed carbonated reservoir in Western Africa. Due to the high number of lithofacies and to the complexitity of the relationship between these

lithofacies, a usual truncated plurigaussian approach was not possible. Our non-parametric approach offered the flexibility to define complex facies contacts through the design of the assignation diagram shown in Figure 3.

References

1. Armstrong, M., Galli, A., Beucher, H., Loc'h, G., Renard, D., Doligez, B., Eschard, R., Geffroy, F.: Plurigaussian Simulations in Geosciences. Springer, Berlin (2003)
2. Dowd, P.A., Pardo-Igzquiza, E., Xu, C.: Plurigau: a computer program for simulating spatial facies using the truncated plurigaussian method. Computers & Geosciences, **29**(2), 123–141 (2003)
3. Emery, X.: Simulation of geological domains using the plurigaussian model: New developments and computer programs. Computers & Geosciences, **33**(9), 1189–1201 (2007)
4. Emery, X. Silva, D.A.: Conditional co-simulation of continuous and categorical variables for geostatistical applications. Computers & Geosciences, **35**(6), 1234–1246 (2009)
5. Galli, A., Beucher, H., Le Loc'h, G., Doligez, B., Heresim Group: The pros and cons of the truncated Gaussian method. In: Armstrong, M., Dowd, P.A. (eds.) Geostatistical Simulations, pp. 217–233. Kluwer, Dordrecht (1994)
6. Le Loc'h, G., Galli, A.: Truncated plurigaussian method: theoretical and practical points of view. In: Baafi, E.Y., Schofield N.A. (eds.) Proceedings of the Fifth International Geostatistics Congress, Wollongong '96. pp. 211–222. Kluwer, Dordrecht (1996)
7. Mariethoz, G., Renard, P., Cornaton, F., Jaquet, O.: Truncated Plurigaussian Simulations to Characterize Aquifer Heterogeneity. Ground Water; **47**(1), 13–24 (2009)
8. Marquardt, D.: An algorithm for least squares estimation of non-linear parameters. SIAM Journal on Applied Mathematics, **11**, 431–441 (1963)
9. Matheron, G., Beucher, H., De Fouquet, C., Galli, A., Guerillot, D., Ravenne, C.: Conditional simulation of the geometry of fluvio-deltaic reservoirs. SPE16753 (1987)
10. Nocedal, J. and Wright, S.: Numerical optimization. Springer, New-York (1999)
11. Siverman, B.: Density Estimation for Statistics and Data Analysis. Chapman & Hall, London. (1986)
12. Xu, C. Dowd, P.A., Mardia, K.V., Fowell, R.J.: A flexible true plurigaussian code for spatial facies simulations. Computers & Geosciences, **32**(10), 1629–1645 (2006)

Appendix

Let us define

$$S = \sum_{k=1}^{K} (\Phi_k(\mathbf{a}) - \pi_k)^2 = (\boldsymbol{\Phi}(\mathbf{a}) - \boldsymbol{\pi})^t (\boldsymbol{\Phi}(\mathbf{a}) - \boldsymbol{\pi}) \tag{7}$$

the quadratic deviation between the target proportions $\boldsymbol{\pi}$ and the proportions obtained after assignation of the functions $a_k f_k(\mathbf{u})$.

A Gauss-Newton algorithm is an iterative algorithm in which, at iteration s, the vector $\mathbf{a}^{(s)}$ is modified in the direction of the gradient at $\mathbf{p}^{(s)} = \boldsymbol{\Phi}(\mathbf{a}^{(s)})$, with an increment $\boldsymbol{\Delta}$ related to the local curvature. The first order Taylor expansion of $\boldsymbol{\Phi}$ around the solution is

$$\boldsymbol{\Phi}(\mathbf{a}) \simeq \boldsymbol{\Phi}(\mathbf{a}^{(s)}) + \mathbf{J}\boldsymbol{\Delta},$$

where $\mathbf{J}$ is the Jacobian matrix with element

$$J_{ij} = \frac{\partial \Phi_i(\mathbf{a})}{\partial a_j}, \quad i, j = 1, \ldots, K. \tag{8}$$

It is then straightforward to show that the gradient S is

$$\mathbf{g} = \nabla S = 2\mathbf{J}^t (\boldsymbol{\Phi}(\mathbf{a}) - \boldsymbol{\pi})$$

with ith component

$$g_i = 2 \sum_{k=1}^{K} \Phi_k \frac{\partial \Phi_k(\mathbf{a})}{\partial a_i}.$$

Nocedal and Wright (1999) showed that the optimal step is

$$\boldsymbol{\Delta} = -\mathbf{H}^{-1}\mathbf{g} = -\mathbf{H}^{-1}\mathbf{J}^t(\mathbf{p}^{(s)} - \boldsymbol{\pi}) = -\mathbf{H}^{-1}\mathbf{J}^t(\boldsymbol{\Phi}(\mathbf{a}^{(s)}) - \boldsymbol{\pi}), \tag{9}$$

where $\mathbf{H}$ is the Hessian matrix of $\boldsymbol{\Phi}$ with element (i, j):

$$H_{ij} = 2 \sum_{k=1}^{K} \left(\frac{\partial \Phi_k(\mathbf{a})}{\partial a_i} \frac{\partial \Phi_k(\mathbf{a})}{\partial a_j} + (\Phi_k(\mathbf{a}) - \pi_k) \frac{\partial^2 \Phi_k(\mathbf{a})}{\partial a_i \partial a_j} \right), \quad i, j = 1, \ldots, K.$$

First and second order derivatives are evaluated numerically. At each step, this would in theory require $K(K+1)$ call to the truncation function $\boldsymbol{\Phi}$, which is prohibitively too slow if there is more than a very few lithofacies. The Gauss-Newton algorithm amounts to approximate the Hessian by ignoring the second order derivatives:

$$\mathbf{H} \simeq 2 \sum_{k=1}^{K} \frac{\partial \Phi_k(\mathbf{a})}{\partial a_i} \frac{\partial \Phi_k(\mathbf{a})}{\partial a_j} = \mathbf{J}^t\mathbf{J}. \tag{10}$$

This approximation is exact when $\boldsymbol{\Phi}$ is linear, which is not the case here. However, on simulations not reported here, close to linear variations of $\boldsymbol{\Phi}(\mathbf{a})$ with respect to

a could be observed, except at some specific locations where small variations in the input variables leads to very high variations of the output variable. With this approximation, it is only necessary to compute numerically the Jacobian matrix **J**. A further approximation consists in approximating the gradient by only computing the derivatives in the N directions with the highest difference between the target proportion and the current output proportions. The Levenberg-Marquardt (Marquardt, 1963) algorithm then consists in adding a regularization term. We thus finally consider the matrix

$$\mathbf{H} = 2(\mathbf{J}^t\mathbf{J} + \lambda \mathbf{I}_K), \tag{11}$$

where $\mathbf{I}_K$ is the identity matrix of dimension K at λ is the Marquardt regularization factor. We finally get that

$$\mathbf{a}^{(s+1)} = \frac{1}{2}(\mathbf{J}^t\mathbf{J} + \lambda \mathbf{I}_K)^{-1}\mathbf{J}^t(\boldsymbol{\Phi}(\mathbf{a}^{(s)}) - \boldsymbol{\pi}). \tag{12}$$

Iterations are stopped when S is less than pre-specified value, ε.

Derivatives are computed numerically:

$$\partial \Phi_i(\mathbf{a})/\partial a_k \simeq (\Phi_i(\mathbf{a}') - \Phi(\mathbf{a}))/\delta. \tag{13}$$

In the direction k, the outpout proportions are computed for the vector $\mathbf{a}'$ with $a'_k = a_k + \delta$ and $a'_j = a_j - \delta/(K-1)$, for $j \neq k$. The approximate derivative is then

$$\partial \Phi_i(\mathbf{a})/\partial a_j \simeq (\Phi_i(\mathbf{a}') - \Phi(\mathbf{a}))/\delta. \tag{14}$$

On all tested situations, we found $N = 3$ to be enough to ensure convergence. Increasing N was found to be, on average, unncessary because the decreasing number of iterations does not compensate for the increasing number of calls to the truncation routine when evaluating the gradient. The algorithm is not very sensitive to λ. It was found that $\lambda = 1$ provides good and fast convergence.