

# HETEROTOPIC BI-CATEGORICAL VARIABLES IN PLURIGAUSSIAN TRUNCATED SIMULATIONS

D. RENARD<sup>1</sup>, H. BEUCHER<sup>1</sup> AND B. DOLIGEZ<sup>2</sup>

<sup>1</sup>Centre de Géosciences, Ecole des Mines de Paris, Fontainebleau, France

<sup>2</sup>Institut Français du Pétrole, Rueil Malmaison cedex, France

## ABSTRACT

*The main characteristics of an ore deposit (i.e. grade) or an oil reservoir (i.e. porosity) can result from the combination of two genetic processes. First, the sedimentation process leads to the definition of the geological variable, referred to as the lithofacies. Then the chemical transformations induce other rock characteristics such as the grade for a chemical element, leading to different grade levels. Both variables are usually considered as categorical variables which are linked by construction. These categorical variables are measured at wells but all variables are not systematically logged and the class definition for each categorical variable is not always the same from one well to another. This may lead to heterotopic information for the two categorical variables. Classically, these deposits are simulated using the variable indicators. So, depending on the link between these two variables, a simple nested simulation can be performed or a more complex correlation must be used. A new Bivariate PluriGaussian (Truncated) model has been developed to process these two categorical variables while accounting for their correlation. The use of heterotopic dataset in the Bivariate PluriGaussian simulations (Bi-PGS) is illustrated on a case study. First the classes of lithofacies and grade levels are defined. Then the complete bivariate model is fitted, where one underlying Gaussian random function is associated to each genetic process: the choice of its parameters is guided by the knowledge of the genetic processes (anisotropy and relationship between the different sets). The two resulting underlying Gaussian random functions can be correlated or not, depending on the correlation between the categorical variables, in other words the dependency between the two genetic processes. Several ways to construct the two underlying Gaussian random*

*functions will be presented. The Bivariate PluriGaussian Truncated Simulation is used to produce several outcomes. It may also be considered in order to derive the missing information at the wells, in a sort of blind test usage.*

## **INTRODUCTION**

Since late 80's, several simulation techniques have been developed to process categorical variables such as a rock lithotype (silt, limestone, sandstone, shale ...), or classes of grade or porosity (rich, medium and poor).

Several techniques for simulating categorical variables can be found in the literature, such as the Sequential Indicator Simulation (SIS) and MultiPoint Statistics Simulation (MPS).

The SIS technique is limited by the choice of model. A valid model must be defined for each specified indicator. Moreover, a multi-indicator approach is required to ensure the consistency of the simulated indicators. This is likely to incur lengthy and tedious modeling.

The limitation for MPS is linked to the choice of the set of training images that match the information available and the conceptual schemes that are to be reproduced.

Another possibility is offered by the Truncated PluriGaussian (PGS) model in which the categories are obtained through variations of underlying Gaussian functions. This paper presents the extension of the PGS to the bivariate case.

### **The PluriGaussian Simulations**

In the original PluriGaussian approach each category is generated by truncating a Gaussian random function (GRF) at different thresholds, hence the use of the term truncated in these simulation methods [Armstrong M., 2003] [Emery X., 2007] [Le Loc'h G., 1994] [Le Loc'h, 1997]. The spatial characteristics of the GRF are deduced from the spatial characteristics of the indicators and the thresholds are computed from their proportions. In the PGS model, there is no restriction on the type of variogram used for modeling the underlying GRF [Lantuéjoul, Ch. 2001].

In this example, a non-conditional simulation of a GRF is generated using a cubic anisotropic model with range of 2/5 of the field size. This outcome is then truncated into four categories (called A, B, C and D) according to a set of thresholds. The proportions are constant over the field for A (15%), B (30%), C (20%) and D (35%). The order relationship among the different categories can be defined in a simple diagram in which the GRF is represented by the horizontal axis: each category is associated with a rectangle whose surface corresponds to the proportion of the category. In particular, it shows that there is no direct contact between A and C or D: this is referred to as an edge effect. All categories present the same behavior which reflects the spatial characteristics (regularity, anisotropy) of the GRF model (Figure 1)

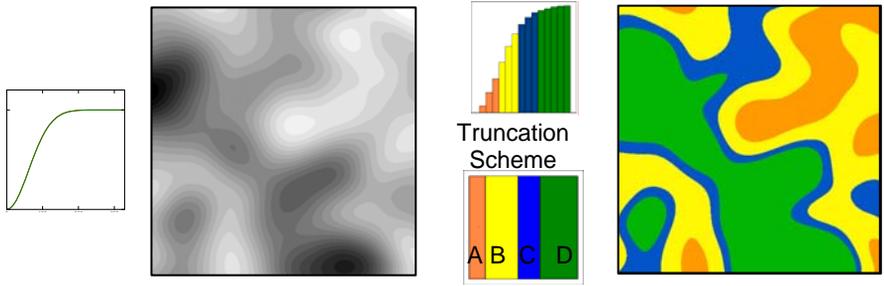


Figure 1: Simulation of a categorical variable (4 categories) obtained by truncating the simulation of an underlying Gaussian random function.

To overcome this edge effect, the Truncated PluriGaussian (PGS) Simulation method was introduced in which the categorical variable is obtained by truncating simultaneously two independent GRF. The first GRF is the same as in the previous example; the second GRF corresponds to an anisotropic cubic model whose range is equal to  $\frac{1}{2}$  of the field along East-West direction and to  $\frac{1}{10}$  along the North-South direction. Each GRF is conventionally represented along the main axis of the Truncation Scheme.

In Figure 2, the Truncation Scheme shows that there is no direct contact between A and C or D; instead B is in contact with A, C and D. It also shows that A and B are defined by thresholds applied to the first GRF and therefore reproduce behavior derived from the model of this GRF, in particular isotropic shapes. C and D are defined by thresholds applied to both GRF and therefore reflect a mixture of behaviors.

The thresholds are derived from the proportions which are the same as in Figure 1.

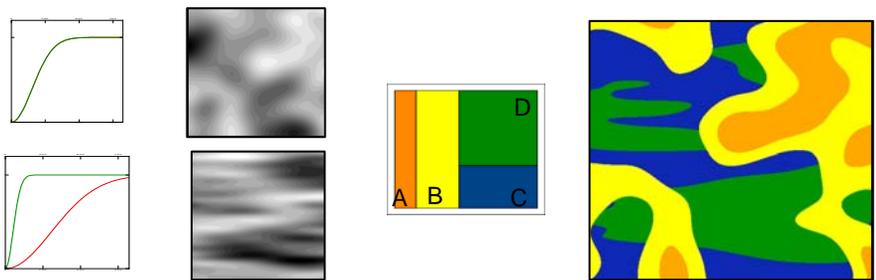


Figure 2: Simulation of a categorical variable (4 categories) obtained by truncating two simulated GRFs with different spatial characteristics.

The PGS model can be further enhanced by including a correlation between the two underlying GRF. Figure 3 shows one PGS outcome obtained with the same parameters as in Figure 2, and introducing a strong correlation coefficient (0.9) between the two GRF. Categories C and D display characteristics that result from

the combination of the anisotropies of the two underlying GRF. This correlation is reflected in the truncation scheme.

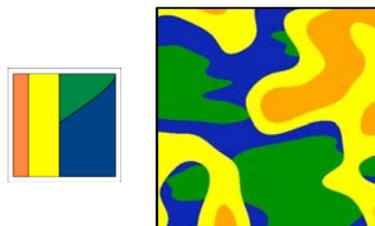


Figure 3: PluriGaussian simulation using a correlation coefficient between the two underlying GRFs.

The PluriGaussian simulations can be conditioned so that any outcome will reproduce the observed categories at sampled locations, which are called conditioning data. A Gibbs sampler iterative procedure is used first to turn the categorical observations into gaussian values which are then used to condition the simulations of the GRF. In the previous examples, only data belonging to categories A, B, C and D can be used as conditioning data.

All the techniques described above correspond to the simulation of a single categorical variable, even when two GRF are involved. These two GRF are defined according to the spatial characteristics of the category, relationships, anisotropies, and qualitative information deduced from the known geology: they carry the same “physical meaning”. In the following approach the idea is to work with sets defined by the classification of two characteristics, hence the bivariate approach.

### The Categorical Bivariate Problem

When a critical property, such as the porosity for a reservoir or the mineral grade for an ore deposit, depends on successive physical processes, it makes sense to consider it as a combination of two random sets, which may or may not be independent.

In the following examples, two categorical variables are considered: the first describes the sedimentation of three lithofacies (A, B and C) and the second results from a chemical transform with three levels (denoted L for low, M for medium and H for high). Moreover, we have a heterotopic dataset: not all chemical levels are present in each sedimentation class: all chemical levels are present in A, only Low and Medium levels in B and Low level in C. Finally, the proportions of all possible associations (A-L, A-M, A-H, B-L, B-M and C-L) are all equal to  $1/6$ .

The intuitive approach is to perform this simulation in two steps. First the sedimentation is processed, conditioned by the sedimentation categorical data, using a PGS model (Figure 4a). This simulation is characterized by the number of GRF, their models and possible correlation coefficient, and the proportions of each sedimentation class ( $1/2$  for A,  $1/3$  for B and  $1/6$  for C). Then, within compartment A of each sedimentation simulation outcome, the second physical

process is simulated conditioned by the subset of chemical categorical data that belong to sedimentation A, using another PGS model. The proportions for the three chemical levels are equal to  $1/3$  (Figure 4b). The same procedure can be repeated for compartment B of each sedimentation simulation outcome, in which the proportions of the two chemical levels (L and M) are equal to  $1/2$  (Figure 4c). Finally, there is no need to simulate the chemical level within compartment C of each sedimentation simulation outcome as it can only be filled with Low chemical level. The simulations of the chemical levels in all sedimentation compartments are glued together (Figure 5).

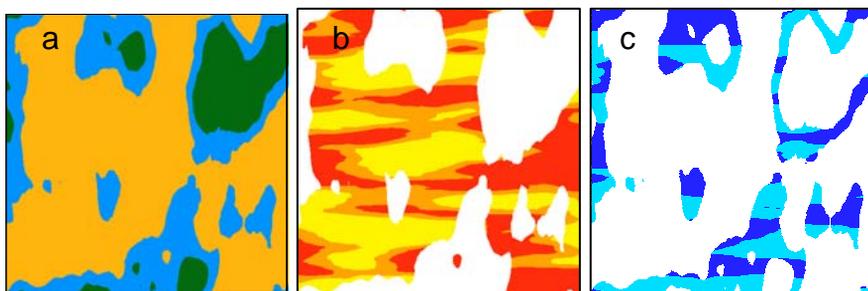


Figure 4: Simulation of the sedimentation (left), simulation of the chemical level in sedimentation A (middle) and B (right).

The drawback of this sequential approach is the lack of continuity of the chemical process throughout the different sedimentation compartments (Figure 5)

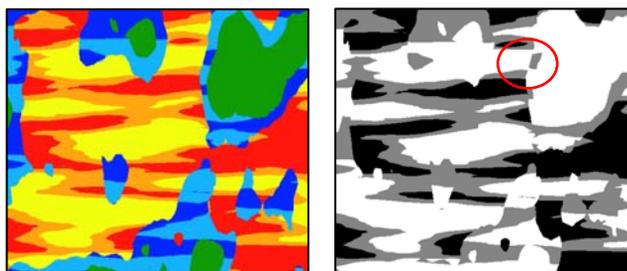


Figure 5: Successive simulation with 6 sedimentation–chemical categories (left). Corresponding simulation of the chemical level (right) with discontinuity (circle).

To avoid these discontinuities, we can return to the traditional PGS technique for this bivariate problem, associating each GRF to a given physical meaning: the first GRF represents the sedimentation and the second GRF the chemical transformation. Figure 6 shows the corresponding truncation scheme and a simulation outcome.

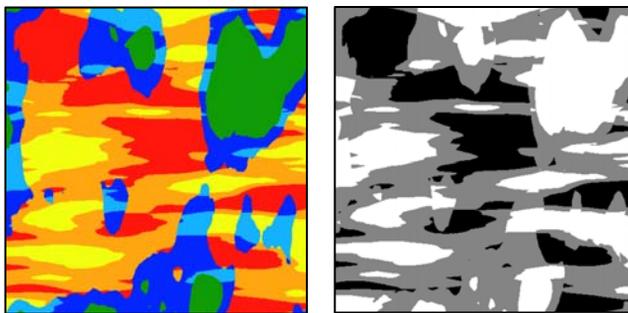


Figure 6: PGS with 6 sedimentation–chemical categories (left). Corresponding simulation of the 3 chemical levels (right).

The simulated chemical levels no longer show any discontinuity other than those due to abrupt changes in chemical level proportions (linked to sedimentation changes). Moreover, using this approach, a correlation between the two physical processes can be taken into account (Figure 7). The link between the two processes can be characterized by correlations other than the linear one chosen for the illustration.

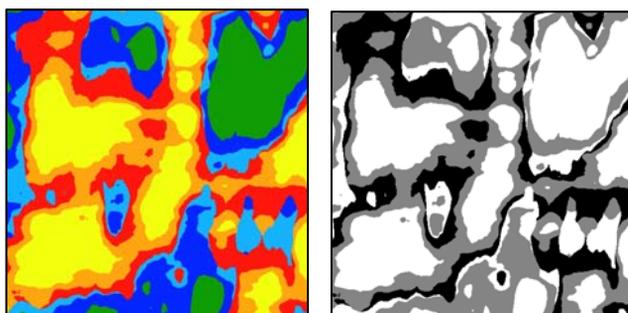


Figure 7: PGS with 6 sedimentation–chemical categories with correlated GRF (left). Corresponding simulation of the 3 chemical levels (right).

This bivariate application of the PGS technique has a limitation for conditioning. Despite the bivariate nature of the problem, the conditioning data consist of a single variable, which combines the sedimentation and the chemical information. It is not possible, therefore, to handle heterotopic information.

## THE BIVARIATE PLURIGAUSSIAN MODEL

A new model, the Bivariate PluriGaussian (Bi-PGS), has been developed to address the problem of heterotopic bivariate conditional simulation. The sedimentation process is governed by a complete PGS model, classically characterized by the number of GRF, their models and possible correlation, the proportions of the different sedimentation categories. The chemical process is governed by a second complete PGS model. The link between the two PGS is

given through the proportions of all the associations of a sedimentation class with a chemical level.

The information consists of nine samples (Figure 8) classified into three sedimentation classes (A, B and C). The chemical level is measured in three classes (Low, Medium or High) for six samples (Table 1). As in the previous examples, the proportions of the associations (A-L, A-M, A-H, B-L, B-M and C-L) are all equal to 1/6; any other association has a zero proportion.

Table 1: List of the samples with the sedimentation and chemical information.

<i>Rank</i>	<i>Sedimentation</i>	<i>Chemical Level</i>
1	A	M
2	B	L
3	C	L
4	A	H
5	B	-
6	A	L
7	B	M
8	C	-
9	A	-

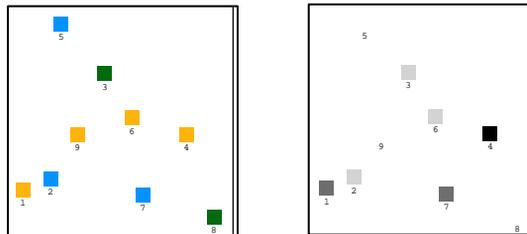


Figure 8: Sedimentation categorical data (left) and Chemical level (right). When no information is available, only the sample number is printed.

For the conditioning step, all heterotopic data are now included. For example, sample #5, which belongs to sedimentation class B, but has no recorded chemical level is used as conditional information in the first PGS and will not constrain the second PGS. An analogous situation would pertain for a sample with a recorded chemical level but no sedimentation information.

For purposes of legibility, we use a single GRF for each PGS model, as illustrated by the truncation schemes (Figure 9)

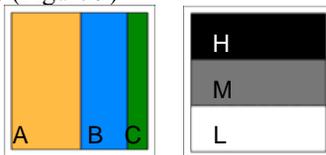


Figure 9: The truncation schemes for sedimentation (left) and chemical levels (right).

The Bi-PGS model is used to perform several simulations conditioned by the set of heterotopic data (Figure 10). At a data point, we note that if the two variables are known, the sedimentation-chemical class is the same for all outcomes. When only the sedimentation is known, the chemical level varies within the possible values.

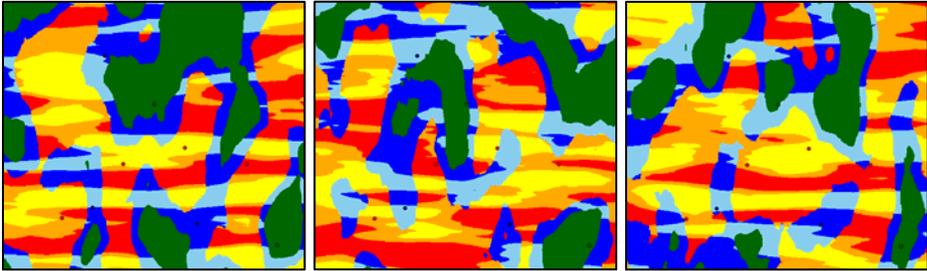


Figure 10: Bi-PGS outcomes with 3 sedimentation classes and 3 chemical levels.

In addition to the outcomes, we can derive the proportion maps which give the probability that a grid node belongs to a sedimentation class (Figure 11). The color scale varies from dark (0%) to light (100%). The sedimentation data are honored as the probability maps show four light spots in the map for A, three in the map for B and two in the map for C.

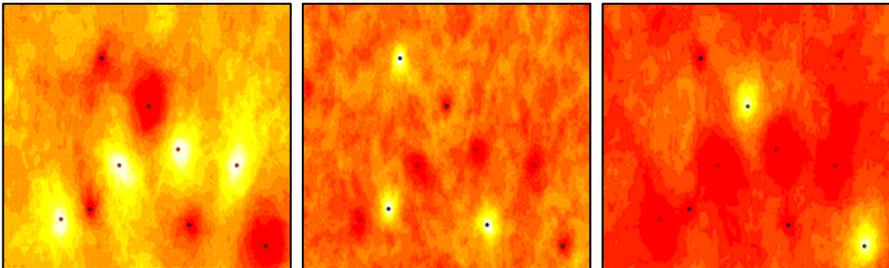


Figure 11: Probability of Sedimentation Class A (left) B (middle) and C (right).

The same procedure for the chemical level shows three spots in the Low map, two in the Medium map and one in the High map which corresponds to categorical data (Figure 12). An interesting feature is the light spot visible in the lower right-hand corner of the Low probability map, at the location of sample#8 where no Chemical level information is provided).

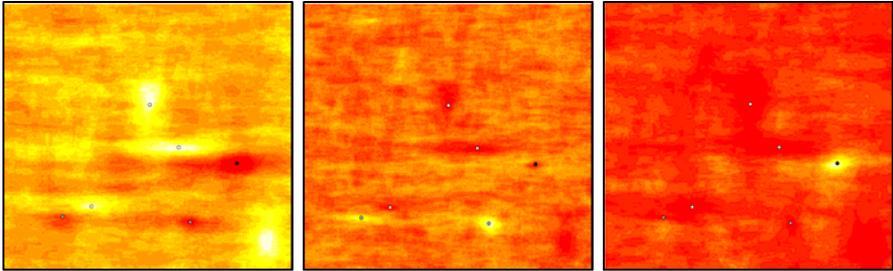


Figure 12: Probability of Chemical level Low (left) Medium (middle) and High (right).

The simulations at the data locations yield the probability that the data belong to each category, taking into account the models and the link between the two variables, coded in the input proportions. Obviously this only provides useful information at locations where one of the two variables is not recorded. The result can be analyzed in terms of the probability belonging to each category for each variable as shown in Table 2.

Table 2: Probability for sample to belong to a Sedimentation class or a Chemical level.

Rank	Data		Sedimentation Proba. (%)			Chemical Proba. (%)		
	Sedimentation	Chemical	A	B	C	L	M	H
1	A	M	100	0	0	0	100	0
2	B	L	0	100	0	100	0	0
3	C	L	0	0	100	100	0	0
4	A	H	100	0	0	0	0	100
5	B	-	0	100	0	56	44	0
6	A	L	100	0	0	100	0	0
7	B	M	0	100	0	0	100	0
8	C	-	0	0	100	100	0	0
9	A	-	100	0	0	32	38	30

The Sedimentation probabilities are either 0 or 100 as each sample belongs to one of the three classes (A, B or C). The same result holds for the Chemical classes, for the samples where the Chemical level is defined.

- At sample #5, which belongs to Sedimentation class B, the probabilities are 56% for Low and 44% for Medium. The probability for High level is zero: this makes sense as this level is not present for Sedimentation B in the model. The Low and Medium probabilities are close to 50% which corresponds to the theoretical conditional probabilities of  $L|B$  (50%) and  $M|B$  (50%).
- At sample #8, which belongs to Sedimentation class C, the probability of Low chemical level is 100%. This is due to the fact that, in the model, the Chemical level can only be Low for Sedimentation C.
- At sample #9, which belongs to Sedimentation class A, all three Chemical levels can be obtained with a probability of 32% for Low, 38

for Medium and 30 for High, close to the theoretical probabilities of L|A (33%), M|A (33%) and H|A (33%).

The probabilities (that differ from 0% and 100%) are close to the theoretical conditional probabilities, but they also measure the interactions between the different samples which depend upon their distances and the spatial characteristics (range and anisotropy) of the underlying GRFs.

## CONCLUSIONS

The Bi-PGS model is a new model that provides a sound basis for bivariate categorical simulation. It is flexible as each physical process is associated with a complete PGS (possibly using two underlying GRF). A further possible development is to introduce a correlation between the two PGS, but this option still has to be specified more precisely.

The Bi-PGS technique can cope with non-stationarity by using proportions for each sedimentation class and chemical level that vary in space [Beucher, 1993]. This novel technique has given some convincing experimental results that will be applied to model diagenetic imprints on sedimentation. It would also be possible to use the Bi-PGS technique for processing two categorical data sets of different qualities.

## REFERENCES

- Armstrong M., Galli A., Le Loc'h G., Geffroy F. and Eschard R. (2003) *PluriGaussian Simulations in Geosciences*, Springer.
- Beucher H., Galli A., Le Loc'h G., Ravanne C. and Heresim Group (1993) *Including a Regional Trend in Reservoir Modelling Using the Truncated Gaussian Method*. In Soares ed., *Geostatistics Tróia '92*. Vol. 1. Dordrecht : Kluwer. pp. 555-566.
- Emery, X. (2007) *Simulation of Geological Domains using the Plurigaussian Model: New Developments and Computer Programs*, Computer & Geosciences vol. 33, Issue 9. pp. 1189-1201.
- Lantuéjoul, C. (2001) *Geostatistical Simulations : Models and Algorithms*, Springer.
- Le Loc'h, G., Beucher H., Galli A., Doligez B. and Heresim Group (1994) *Improvement in the Truncated Gaussian Method : Combining Several Gaussian Functions*, Proceedings ECMOR IV, 4th European Conference on the Mathematics of Oil Recovery, Røros, Norway. 13p.
- Le Loc'h, G. and Galli, A. (1997) *Truncated Plurigaussian Method: Theoretical and Practical Points of View*. In Baafi et al. eds., *Geostatistics Wollongong '96*. Vol. 1. Dordrecht : Kluwer. pp. 211-222.