Les Cahiers du GERAD ISSN: 0711–2440

Joint simulation of correlated variables using high-order spatial statistics for orebody modeling

I. Minniakhmetov
R. Dimitrakopoulos

G–2015–85

September 2015

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Joint simulation of correlated variables using high-order spatial statistics for orebody modeling

Ilnur Minniakhmetov [a]
Roussos Dimitrakopoulos [a,b]

[a] *COSMO – Stochastic Mine Planning Laboratory, Department of Mining and Materials Engineering, McGill University, Montreal (Quebec) Canada, H3A 2A7*

[b] *and GERAD*

ilnur.minniakhmetov@mail.mcgill.ca
roussos.dimitrakopoulos@mcgill.ca

September 2015

Les Cahiers du GERAD
G–2015–85

Abstract: Geostatistical simulation techniques are used to quantify uncertainty of spatial attributes of interest describing mineral deposits, petroleum reservoirs, hydrogeological horizons, environmental contaminants and so on. The majority of existing methods consider second-order spatial statistics and Gaussian processes, while the more advanced multiple point based simulation approaches are algorithmic and do not consistently account for the high-order spatial relations in data. Recently, simulation techniques for complex and non-Gaussian, spatially distributed variables have been developed, based on high-order spatial cumulants, and make no assumptions on data distribution or require data transformations. In this paper, the previous developments are extended and a new approach for the joint simulation of multiple correlated variables using high-order spatial statistics is proposed. The technique is based on a new algorithm described here for the decorrelation of correlated variables into factors, using the so-termed diagonal domination condition of high-order cumulants. The decorrelated factors are then simulated using high-order simulation and back-transformed into the initial correlated variables. The decorrelation using diagonal domination of high-order statistics is tested with a dataset from a multi-element iron ore deposit and then a fully known two-dimensional dataset with two correlated variables is used to demonstrate the practical intricacies of the proposed method.

Key Words: Stochastic simulation, joint simulation, high-order spatial statistics, cumulant, decorrelation.

1 Introduction

Major source of uncertainty affecting mine planning optimization is that pertinent geological aspects of the orebody model being considered, including grades, metal, material types, other rock properties of interest are interpolated from a limited number of drillholes (Dimitrakopoulos, 2011). The last two decades new, advanced spatial stochastic simulation techniques have been developed to assist with the quantification of the related uncertainty. Since the early efforts (e.g. David, 1988; Goovaerts, 1997; Chiles and Delfiner, 2012), a new generation of algorithms implicitly uses high-order spatial relations. More specifically, multiple point simulation (MPS) algorithms (Journel, 2005; 2007), such as the snesim (Strebelle and Cavelius, 2014), filtersim (Zhang et al., 2006), and simpat (Arpat and Caers, 2007), cdfsim (Mustapha, Chatterjee and Dimitrakopoulos, 2014) and others (De Vries et al., 2008; Chiginova and Hu, 2008; Honarkhah, 2011; Straubhaar, 2011; De Iaco and Maggio, 2011); Markov random field models based multi-point approaches (Tjelmeland, 2013); multi-scale MPS simulations based on discrete wavelet decomposition (Chatterjee and Dimitrakopoulos, 2012). MPS techniques use pattern-based algorithms and are highly depend on the so-termed training image, TI, (or geological analogue) used rather than drill hole data available. As a result, the related simulated representations of attributes of interest may not reproduce the spatial statistics of the data and thus are called TI-driven. These issues become apparent in mining applications as there are, relatively to other applications such as in petroleum reservoir characterization, a reasonable amount of data available (Osterholt and Dimitrakopoulos, 2007; Goodfellow et al., 2012).

The high-order geostatistical simulation framework based on high-order spatial statistics was recently introduced, to provide an alternative, data-driven and mathematically consistent approach (Dimitrakopoulos et al., 2010; Mustapha and Dimitrakopoulos, 2010a, 2011; Mustapha, Dimitrakopoulos and Chatterjee, 2011). Notably, the high-order simulation framework does not require any distribution assumptions or data transformations and it is founded upon the well-established sequential simulation framework, where a non-parametric Legendre polynomial series approximation (Lebedev, 1965) is used to estimate local conditional density functions. This framework provides a consistent mathematical model and, it is important to stress again, ensures that the modeling process is data-driven and thus avoids potential conflicts between the spatial statistics of available data and simulated realizations. The above developments considers up to now univariate simulation and the high-order simulation of multiple correlated attributes of interest in a mineral deposit remains to be addressed.

Joint simulations of multiple correlated elements is available with second-order Gaussian simulation methods and are based on decorrelation approaches, such as minimum-maximum autocorrelation factors (e.g. Fonseca and Dimitrakopoulos, 2003; Boucher and Dimitrakopoulos, 2009; 2012; Mueller and Ferreira, 2012; Rondon, 2012) or principal component analysis, PCA (David, 1988; Wackernagel, 1998). Approaches for decorrelation of non-Gaussian variables may be based on principle component cumulant analisys or PCCA (Morton and Lim, 2009) and independent component analysis or ICA (Comon, 1994). The first method relies on finding principal cumulant components that account for most of the variation in all higher-order cumulants, just as PCA obtains maximum variance components. In ICA approaches, transformation into independent factors is performed by maximization of different independence measures (Hyvarinen 1999), such as likelihood and network entropy, mutual information and Kullback-Leibler divergence, non-linear cross-correlations, non-linear PCA criteria, higher-order cumulant tensors, weighted covariance matrix, negative entropy, general contrast functions. In this paper, a new measure of independence for ICA is first proposed. The measure is based on the diagonal domination of high-order cumulants of factors: The absolute values of diagonal elements are substantially greater than non-diagonal ones. It is not hard to see, that this condition will maximize independence between factors, because for independent variables the high-order cumulants are diagonal. Generally, there can be no transformation into completely independent factors. From this point of view, using diagonal domination condition seems natural, because, in contrast to any strong diagonalization condition, it allows to obtain maximally diagonal cumulants for all orders simultaneously. The advantage of proposed technique is that it uses directly the high-order statistics of a multivariate dataset, which characterize multivariate probability distribution functions.

In the following sections, first the basic definitions for high-order spatial statistics and spatial simulation are given and are followed by the proposed joint high-order simulation approach. Subsequently, a multi-

element dataset from an iron ore deposit is used to test and assess the high-order decorrelation approach proposed and it is followed by an application and testing of the proposed simulation method in an exhaustively known deposit.

2 High-order spatial statistics

Let $(\Omega, \Im, P)$ be a probability space and $Z(x)$ be a real random field in Rn defined at the locations, $x_i \in D \subseteq \mathbb{R}^n (n = 1, 2, 3)$ for $i = 1 \dots N$, where N is the number of points in a discrete grid $D \subseteq \mathbb{R}^n$. Assuming $Z(x)$ is a zero-mean ergodic stationary random field in Rn, then the cumulants of $Z(x)$ are defined by the MacLaurin expansion of the cumulant generating function (Rosenblatt 1985):

$$K(\omega) = \ln\left(E\left[e^{\omega Z}\right]\right) \tag{1}$$

The high-order spatial moment of order r is

$$Mom\left[Z(x), Z(x+h_1), \dots, Z(x+h_{r-1})\right] = E\left[Z(x)Z(x+h_1)\dots Z(x+h_{r-1})\right] \tag{2}$$

Similarly, the cumulants of the random field $Z(x)$ up to order r can be expressed as

$$c_r^z(h_1, \dots, h_{r-1}) = Cum\left[Z(x), Z(x+h_1), \dots, Z(x+h_{r-1})\right] \tag{3}$$

For example, cumulant order 1 is a mean of random field $Z(x)$

$$c_1^z = E\left[Z(x)\right] \tag{4}$$

Second-order cumulant of the random field $Z(x)$ is known as covariance

$$c_2^z(h) = E\left[Z(x)Z(x+h)\right] - E\left[Z(x)\right]E\left[Z(x+h)\right] \tag{5}$$

Its third-order cumulant is given by

$$\begin{aligned} c_3^z(h_1, h_2) = {} & E\left[Z(x)Z(x+h_1)Z(x+h_2)\right] - E\left[Z(x)\right]E\left[Z(x+h_1)Z(x+h_2)\right] \\ & - E\left[Z(x+h_1)\right]E\left[Z(x)Z(x+h_2)\right] - E\left[Z(x+h_2)\right]E\left[Z(x)Z(x+h_1)\right] \\ & - E\left[Z(x)\right]E\left[Z(x+h_1)\right]E\left[Z(x+h_2)\right] \end{aligned} \tag{6}$$

For calculation of higher-order cumulants the reader is referred to Dimitrakopoulos et al. (2010) and Mustapha and Dimitrakopoulos (2010b).

Let us summarize related properties of cumulants:

1. Cumulants are multi-linear

$$Cum(A_{\alpha i}Z(x_i), A_{\beta j}Z(x_j), A_{\gamma k}Z(x_k)) = A_{\alpha i}A_{\beta j}A_{\gamma k}Cum(Z(x_i), Z(x_j), Z(x_k)) \tag{7}$$

   where A is an arbitrary matrix.
2. Cumulants of independent variables are diagonal.
3. Cumulants of Gaussian variables with order greater than two are equal to zero.

3 High-order simulation

Let $Z(x_i)$ or Z_i be random field indexed in $\mathbb{R}^n$, $x_i \in D \subseteq \mathbb{R}^n (n = 1, 2, 3)$, where N is the number of points in a discrete grid $D \subseteq \mathbb{R}^n$. The focus of high-order simulation techniques is to simulate the realization of random field $Z(x_i)$ for all nodes of a discrete grid $D \subseteq \mathbb{R}^n$ with given set of conditioning data $d_M = \{Z(x_\alpha), \alpha = 1 \dots M\}$. The high-order simulation method proposed by Mustapha and Dimitrakopoulos (2010a; 2011) uses Legendre polynomials and coefficients in terms of cumulants to approximate conditional probability density

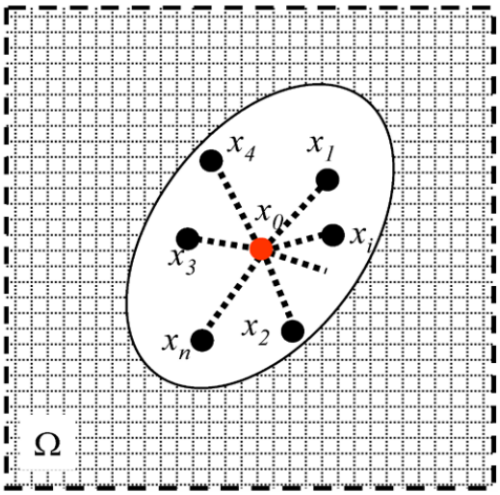

Figure 1: An unknown value is at the location x_0, and the values at the locations $x_1, x_2, \ldots x_n$ in the neighborhood of x_0 are assumed to be known.

function at each node of simulation grid. The algorithm runs sequentially for every node of the simulation grid similarly to sequential simulation algorithms.

On the first step a random path for visiting all unsampled nodes is defined. Assume that x_0 is the first node to be visited with neighbours $x_1, x_2, \ldots x_n$ represented in Figure 1.

Then, the conditional probability density function (cpdf) $f_{Z_0}(z_0/d_M)$ is given by:

$$f_{Z_0}(z_0/d_M) = \frac{1}{\int f_Z(x)dz_0} f_Z(z_0, z_1, \ldots, z_M) = \frac{1}{\int f_Z(x)dz_0} \sum_{i_0=0}^{\infty} \sum_{i_1=0}^{\infty} \cdots \sum_{i_M=0}^{\infty} \bar{L}_{i_0,i_1,\ldots i_M} \bar{P}_{i_0}(z_0) \tag{8}$$

where $z_1, \ldots, z_M$ are values of conditioning data set d_M, $f_Z(z_0, z_1, \ldots, z_M) = f_Z(x)$ is joint probability density function of random value at node x_0 and conditioning data set d_M, $\bar{P}_{i_0}(z_0) = \sqrt{2i_0+1}P_{i_0}(z_0)/\sqrt{2}$ is normalized Legendre polynomial, $L_{i_0,i_1,\ldots i_N}$ and $\bar{L}_{i_0,i_1,\ldots i_N} = L_{i_0,i_1,\ldots i_N}\bar{P}_{i_1}(z_1)\ldots\bar{P}_{i_M}(z_M)$ are coefficients defined in terms of high-order spatial cumulants in (7). Then, a value from the estimated cumulative distribution function is drawn randomly and the simulated value at location x_0 is added to the set of conditioning data. This approach is repeated for all points in the random path defined above.

4 Joint high-order simulation

Let $Z(x) = \{Z_1(x), Z_2(x), \ldots, Z_p(x)\}$ be spatially-correlated random variables given data set $d_M = \{Z_k(x_\alpha), k = 1 \ldots p, \alpha = 1 \ldots M\}$, where k is the index of variable and α is index of location. Due to correlations between variables, the high-order simulation technique described above cannot be applied independently for all $Z(x)$ variables. Thus, the initial variables $Z(x)$ are decomposed into uncorrelated factors $Y(x) = \{Y_1(x), Y_2(x), \ldots, Y_q(x)\}$ using the linear transformation A.

$$Y(x) = AZ(x) \tag{9}$$

Then, independent simulations are carried out for each variable $Y_k(x)\{k = 1 \ldots q\}$ using high-order simulation technique described above. The resulting jointly simulated grades $Z(x)$ are obtained by back-transformation of factors $Y(x)$

$$(x) = \tilde{A}Y(x) \tag{10}$$

where $\tilde{A}$ is a back-transformation matrix. In this paper, the high-order spatial statistics of $Z(x)$ and $Y(x)$ are used to find the transformation matrix A.

4.1 Diagonal dominant spatial cumulants

Let $Cum(Y_{k_1}(x), Y_{k_2}(x), \dots Y_{k_r}(x))$ be joint-cumulants of factors $Y(x)$. Suppose all factors are independent, then from cumulant properties described above, only diagonal elements of high-order joint-cumulants are not equal to zero

$$Cum(Y_{k_1}(x), Y_{k_2}(x), \dots Y_{k_r}(x)) \neq 0, \textit{ if only } k_1 = k_2 = \dots = k_r \tag{11}$$

In general, the joint-cumulants of different orders cannot be diagonalized simultaneously. Therefore, in this work we do not assume that joint-cumulants of factors $Y(x)$ are diagonal, but they have strong diagonal domination. For second-order tensor, i.e. matrix, a diagonal domination condition is

$$|a_{ii}| \geq \sum_{j \neq i} |a_{ij}| \ \ \forall i \tag{12}$$

For high-order tensor diagonal domination condition is domination of diagonal elements over sum of non-diagonals:

$$|a_{i,i,\dots,i}| \geq \sum_{i_1} \sum_{i_2} \cdots \sum_{i_{r-1}} a_{i,i_1,\dots,i_{r-1}} \ \ \forall i \tag{13}$$

here the sums are taken for all indexes $i_1, i_2, \dots, i_{r-1}$ except the case when all indexes are equal i.

4.2 Objective function

To obtained transformation matrix A into the factors $Y(x)$ with diagonal dominant cumulants the following objective function is used

$$F(A) = \sum_d \alpha_d F_d(A) \tag{14}$$

where d is order of cumulant, α_d are constants, and $F_d(A)$ are defined by

$$\begin{aligned} F_d(A) &= \sum_{k_0} \frac{\sum_{k_1} \sum_{k_2} \cdots \sum_{k_{d-1}} \left\| Cum(Y_{k_0}(x), Y_{k_1}(x), \dots, Y_{k_{d-1}}(x)) \right\|_2}{\left\| Cum(Y_{k_0}(x), Y_{k_0}(x), \dots, Y_{k_0}(x)) \right\|_2} \\ &= \sum_{k_0} 1 + \frac{\sum_{non-diagonal} \left\| Cum(Y_{k_0}(x), Y_{k_1}(x), \dots, Y_{k_{d-1}}(x)) \right\|_2}{\left\| Cum(Y_{k_0}(x), Y_{k_0}(x), \dots, Y_{k_0}(x)) \right\|_2} \end{aligned} \tag{15}$$

here factors Y_{k_0} are functions of matrix$A$, because $Y(x) = AZ(x)$.

Coefficients α_d are decreasing for high-order cumulants to guarantee first of all decorrelation for low-order statistics. In this work, we use the same weights $\alpha_d = O\left(1/d!\right)$ as suggested by Morton and Lim (2009) for principle cumulant component analysis technique. It is not hard to see, that the second term in (15) is the inverse diagonal domination condition (13), and minimization of objective function (15) will maximize diagonal domination of joint cumulant. To minimize objective function (15) limited-memory Broyden–Fletcher–Goldfarb–Shanno optimization algorithm was used (Perry 1977). The gradients for objective function were calculated analytically from (15).

5 Case studies

This section explores the application of the proposed diagonal domination technique for joint simulation in two testing cases. Firstly, a dataset from a multiple element iron ore deposit is used to explore the proposed high-order decorrelation method and, secondly, the widely known Walker Lake data set provides a fully known environment to assess the technique (Isaaks and Srivastava, 1990).

5.1 Iron ore deposit data set

The dataset from the iron ore deposit consists of grades from drillhole samples for five elements: iron (Fe), phosphorus (P), Aluminum oxide (Al_2O_3), Silicon dioxide (SiO_2) and loss of ignition (LOI). According to histograms (Figure 2), all variables have non-Gaussian distributions. To decorrelate the variables the proposed high-order decorrelation technique is used.

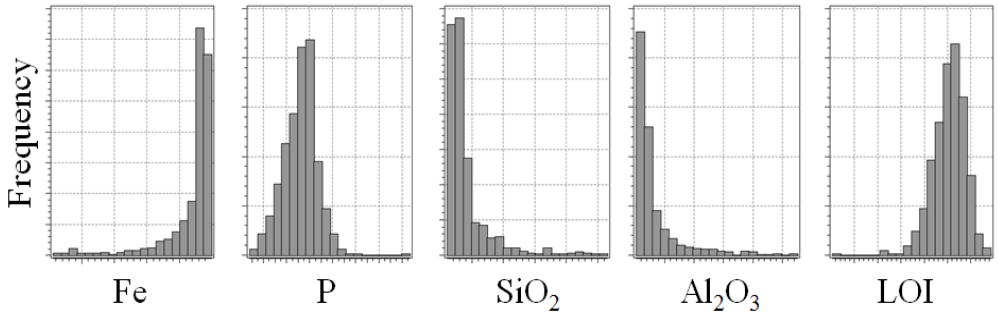

Figure 2: Histograms of different material types distribution.

5.2 High-order cumulant visualization

For high-order visualization of cumulants the unfolding technique for high-order tensor is used. High-order tensors are unfolded into the matrixes, which tend to have minimum difference between the number of columns and rows. For example, fourth-order tensor with dimensions $3 \times 3 \times 3 \times 3$ and sixth-order tensor with dimensions $3 \times 3 \times 3 \times 3 \times 3 \times 3$ are unfolded into matrixes 9×9 and 81×81, respectively. Third-order tensor with dimensions $3 \times 3 \times 3$ and fifth-order tensor with dimensions $3 \times 3 \times 3 \times 3 \times 3$ are unfolded into matrixes 3×9 and 9×81, respectively. Unfolded fourth-order tensor with dimensions $3 \times 3 \times 3 \times 3$ is shown on Figure 3. Each element of its matrix representation with dimensions 9×9 is depicted by grid-cell. The correspondence between indexes of tensor i, j, k, l and its matrix representation ij, kl is given by

$$\begin{aligned} ij &= 3i + j \\ kl &= 3k + l \end{aligned} \tag{16}$$

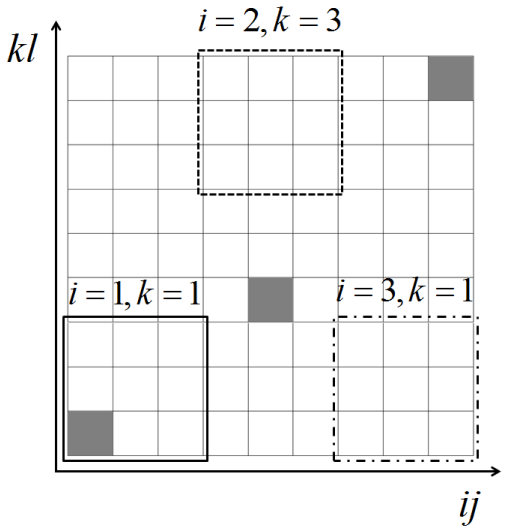

Figure 3: Unfolded tensor $3 \times 3 \times 3 \times 3$. Each cell is element of the tensor. Solid, dashed, and dot-dashed squares are slices. Grey cells are diagonal elements of the tensor.

For example, group of cells highlighted by solid square represents slice of fourth-order tensor $T_{1,j,1,l}$ with fixed indexes $i = 1, k = 1$. Group of cells highlighted by dashed square represents slice of fourth-order tensor $T_{2,j,3,l}$ with fixed indexes $i = 2, k = 3$ and group of cells highlighted by dash-doted square represents slice of fourth-order tensor $T_{3,j,1,l}$ with fixed indexes $i = 3, k = 1$. The grey cells are diagonal elements of tensor: $T_{1,1,1,1}$, $T_{2,2,2,2}$, and $T_{3,3,3,3}$.

5.2.1 Decorrelation results

For the purpose of present study, only the high-order cumulants up to order four are calculated, as this seems sufficient. The second-order cumulants shown on Figure 4a is the covariance matrix of initial variables and on Figure 4b is covariance of factors. Grey levels represent correlation levels; white color is for high correlation, black – for low correlation. According to Figure 4a, the first variable Z_1 is highly correlated with the third Z_3 and fourth Z_4 ones, second and fifth variables have low correlations with other variables. In the factor space (Figure 4b) all non-diagonal elements have values less than 0.01, which means they have almost no correlation between each other.

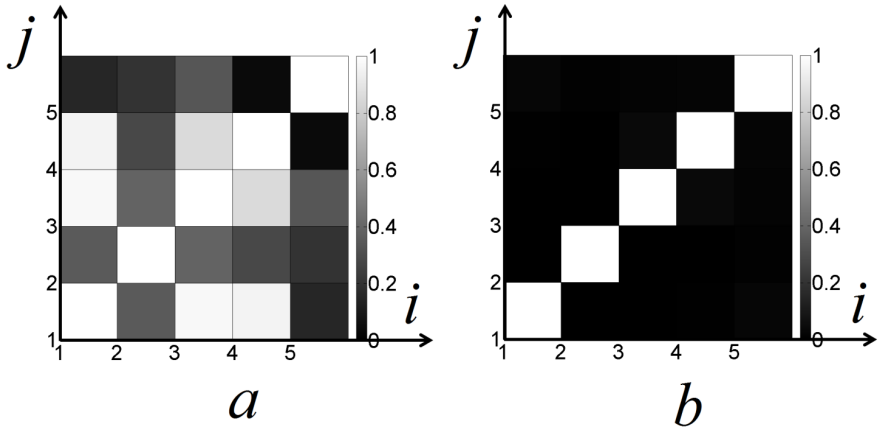

Figure 4: Covariance matrix of (a) initial variables and (b) factors. Each cell is element of the tensor. Grey levels are the levels of correlation: white color – maximum correlation, black color – no correlation.

If second-order cumulants are responsible for correlation between two variables, then the third-order cumulants on Figure 5 represent the dependence between values of three variables. The group of white and light grey cells on the left side of Figure 5a shows that the values of variable Z_1 are highly correlated with values of Z_2 and Z_3, Z_3 and Z_4, or Z_2 and Z_4.The cells referred to variable Z_5 have dark grey color, which means that the values of variable Z_5 have low correlation with values of other variables. In the factor space, there is a low correlation between values of first Y_1 and fifth variables Y_5 represented by dark grey cells in

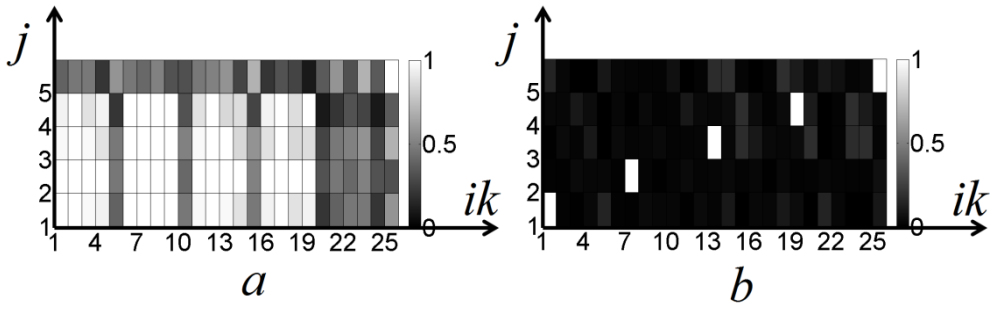

Figure 5: Unfolded third-order joint cumulant of (a) initial variables and (b) factors. Each cell is element of the tensor. Grey levels are the levels of correlation: white color – maximum correlation, black color – no correlation.

the first column and first row on Figure 5b. The high values are only on diagonal cells, therefore the factors are decorrelated for order 3.

The fourth-order cumulants are shown on Figure 6. Similarly the above, there is a dependence between variables Z. However, it should be noticed, that for order 2 and 3 there is almost no correlation between variable Z_5 and others, and only for order 4, it is not hard to see, that the high level of dependence between variablesZ_5, Z_1 and Z_3 is revealed. In the factor space (Figure 6b), the fourth-order cumulant has values about $0.2 - 0.4$ for elements Y_3, Y_4, Y_5. The high values are only on diagonal cells, therefore the factors are decorrelated also for order four.

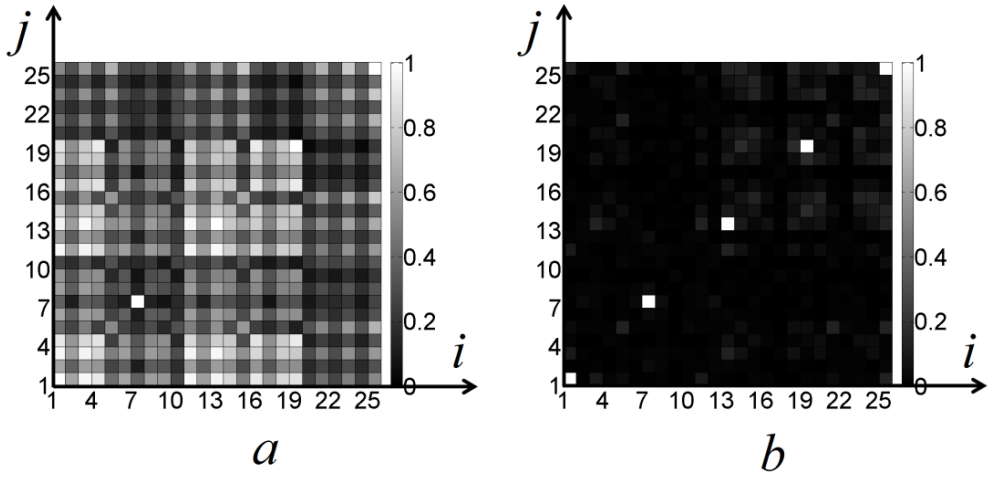

Figure 6: Unfolded fourth-order joint cumulant of (a) initial variables and (b) factors. Each cell is element of the tensor. Grey levels are the levels of correlation: white color – maximum correlation, black color – no correlation.

The correlations between initial variables Z and decorrelated factors Y represented on Figures 7 and 8, respectively. Obviously, grades values of iron and phosphorus are dependent; the same is true for aluminum oxide and silicon dioxide. On the contrary, the scatter plots of factors Y_1 with Y_2 and $Y_3$with Y_4seem to look like a white noise and show no correlations between factors.

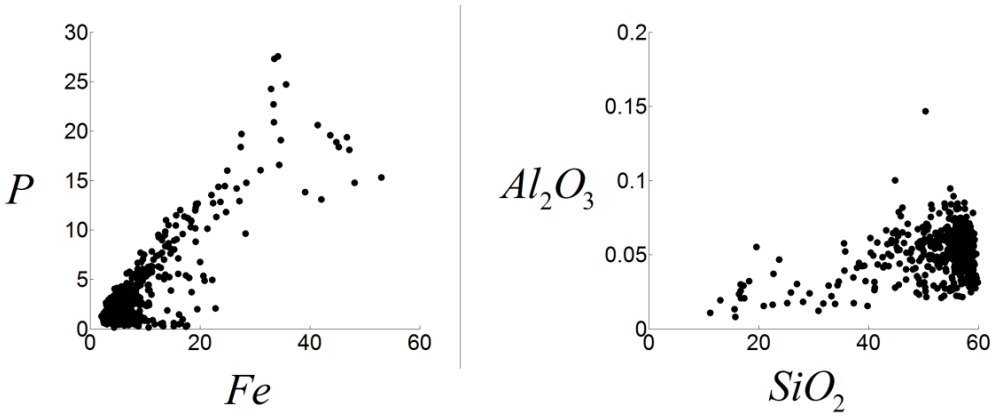

Figure 7: Scatter plot of initial variables. Left figure represents the dependence between phosphorus and iron grades. Right figure represents the dependence between Aluminum oxide and Silicon dioxide grades.

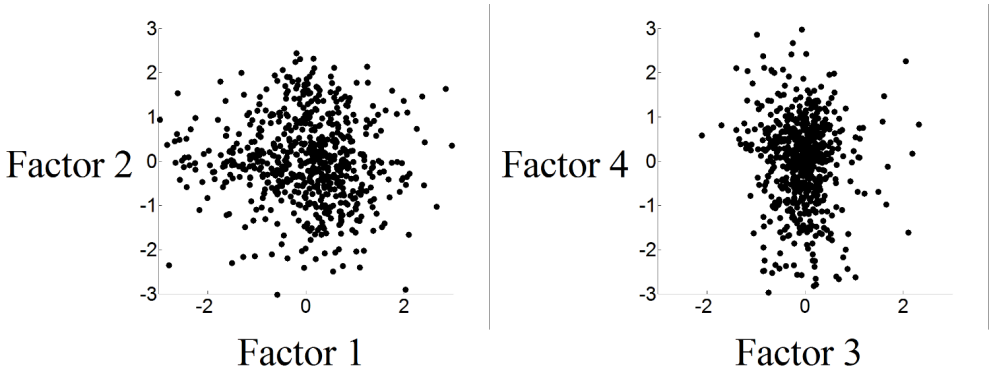

Figure 8: Scatter plot of factors. Left figure represents the dependence between first and second factors. Right figure represents the dependence between third and fourth factors.

5.3 Walker Lake case study

As an initial data for Walker Lake case study drill holes data and training images of two variables U and V are used (Figures 9 and 10, respectively). Drill holes data are sampled uniformly random from training images: 0.5%, 1%, 5% and 10% of training image grid-points (Figure 9). All initial data is transformed into decorrelated factors, simulated independently using HOSIM technique, and back-transformed into initial data space.

5.3.1 Decorrelation results

Similarly to the dataset from the iron ore deposit, high-order cumulants up to order four for data variables and decorrelated factors is calculated. The covariance matrixes of variables U and V are shown on Figure 11a. The correlation between variables is high, approximately 0.8. The same behavior is for third and fourth order cumulants, Figures 12a and 13a, respectively. In the factor space there is no correlation on the second-order cumulant (Figure 11b) and low values, approximately 0.2–0.3, for third-order and forth-order cumualnts, Figures 12b and 13b, respectively. Thus, the factors are decorrelated up to order four using the proposed diagonal domination technique.

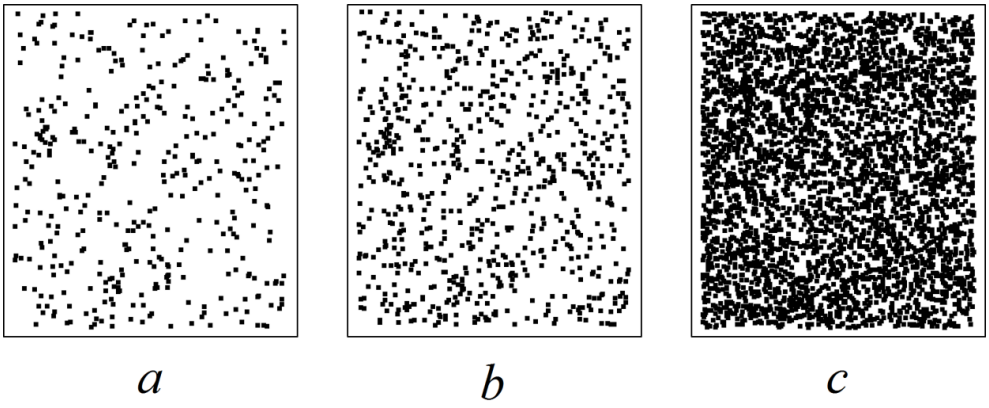

Figure 9: Walker Lake data set. Data points sampled uniformly from training image for V variable: (a) 390 points, (b) 780 points, (c) 3900 points.

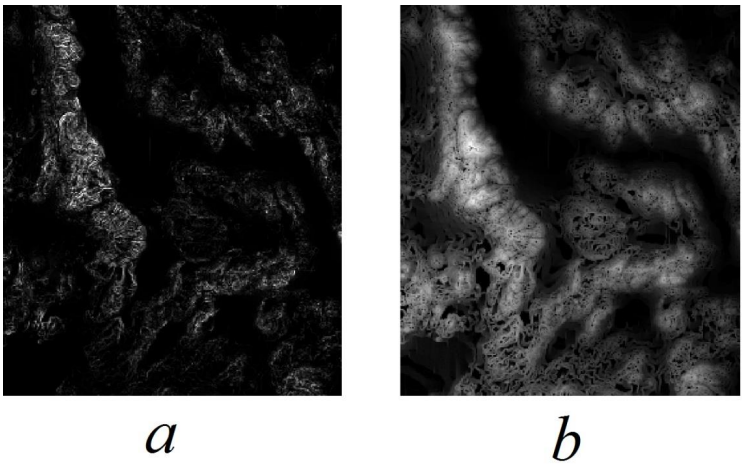

Figure 10: Walker Lake data set. Training images for variables (a) U and (b) V.

5.3.2 Simulation results

The simulations for different number of data samples are shown on Figures 14, 15, and 16. The high-order simulation approach used here (Mustapha and Dimitrakopoulos 2010a, 2011) is data-sample driven, contrary to pattern based methods as noted in the introduction, and it is not hard to see the improvement of simulation results with the increasing of the number of data points.

To quantitative estimation of joint simulation results, the relative errors ε_d for high-order statistics of simulations are calculated

$$\varepsilon_d(i_1, i_2, \dots i_d) = \left| \frac{Cum(\tilde{Z}_{i_1}(x), \dots, \tilde{Z}_{i_d}(x)) - Cum(Z_{i_1}(x), \dots, Z_{i_d}(x))}{Cum(Z_{i_1}(x), \dots, Z_{i_d}(x))} \right| \tag{17}$$

where d is the order of cumulants, $Cum(\tilde{Z}_{i_1}(x), \dots, \tilde{Z}_{i_d}(x))$ and $Cum(Z_{i_1}(x), \dots, Z_{i_d}(x))$ are d^{th}-order cumulants, calculated for simulations and training images, respectively. The relative errors demonstrate how accurate correlation properties of variables U and V we reproduced. Errors are calculated for simulations with different number of data points. The maximal relative error for all cases does not exceed 10%. Thus, the fields U and V are simulated using the proposed diagonal domination technique and high-order simulations, and the high-order statistics of the simulations have approximately the same values as in the initial data sets used.

6 Conclusions

This paper presented a new approach for high-order joint simulation of non-Gaussian spatially correlated variables, as frequently required in orebody modeling applications. The proposed technique combines a new approach for decorrelation random variables into factors using diagonal domination condition for high-order cumulants and the high-order simulation approach for the independent simulation of factors. Simulated factors are then back transformed to independent high-order correlated realizations of pertinent attributes of interest. The algorithm presented herein is tested on drillhole data set of real iron ore deposit and Walker Lake exhaustively known dataset. The results show that the newly proposed method transforms initial variables into factors with low level of dependence. Using information about high-order statistics both on decorrelation step and simulation step allows generating jointly simulated grades with the same statistical properties as the initial data. Further research will be addressed to applications for different case studies and decorrelation of variables for non-zero lags.

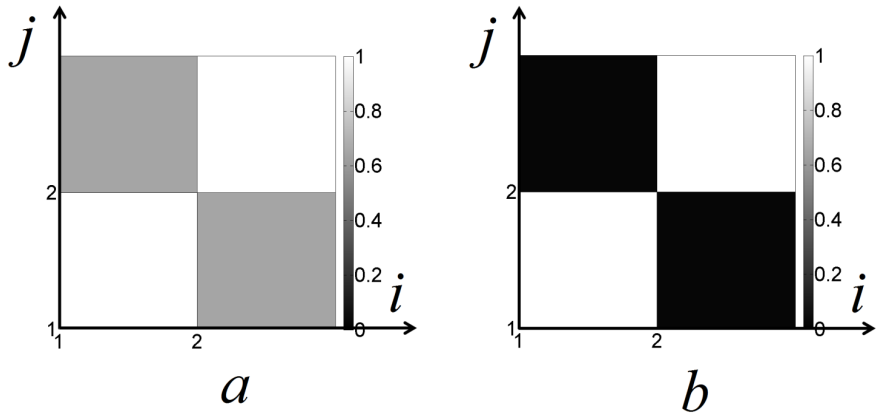

Figure 11: Covariance matrix of (a) initial variables and (b) factors. Each cell is element of the tensor. Grey levels are the levels of correlation: white color – maximum correlation, black color – no correlation.

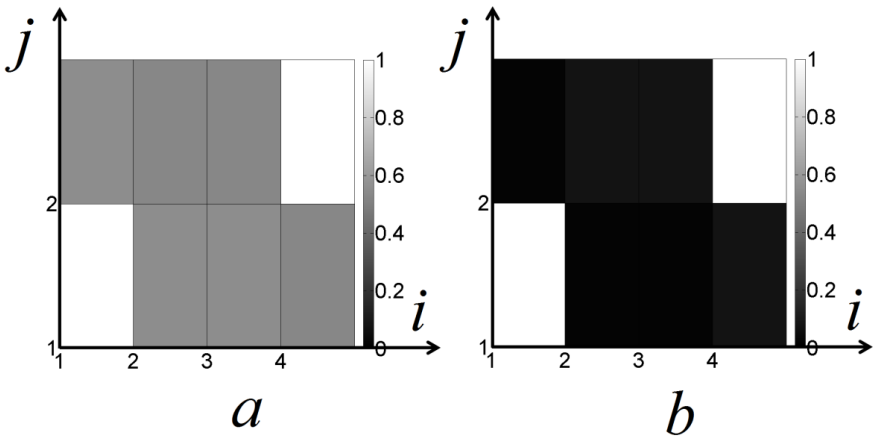

Figure 12: Unfolded third-order joint cumulant of (a) initial variables and (b) factors. Each cell is element of the tensor. Grey levels are the levels of correlation: white color – maximum correlation, black color – no correlation.

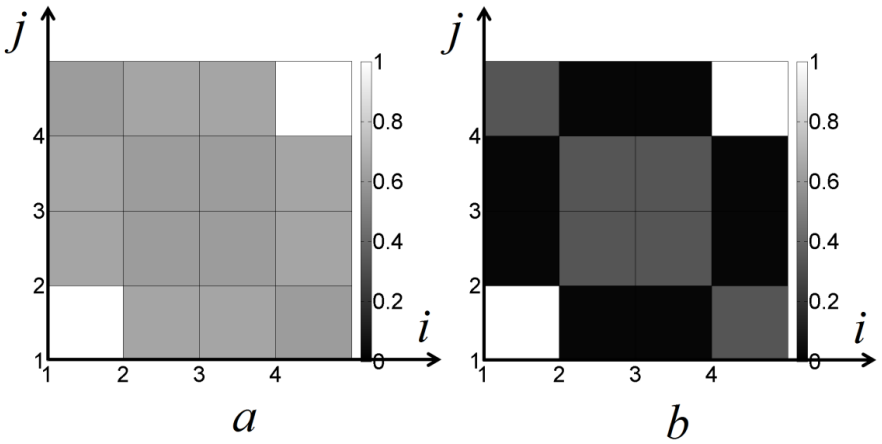

Figure 13: Unfolded fourth-order joint cumulant of (a) initial variables and (b) factors. Each cell is element of the tensor. Grey levels are the levels of correlation: white color – maximum correlation, black color – no correlation.

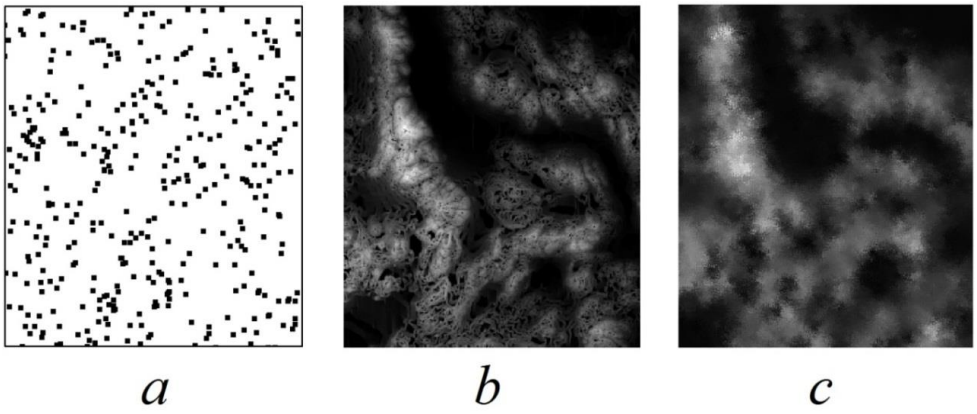

Figure 14: Simulation of variable V using 390 points: (a) data samples, (b) reference image, (c) joint simulation result.

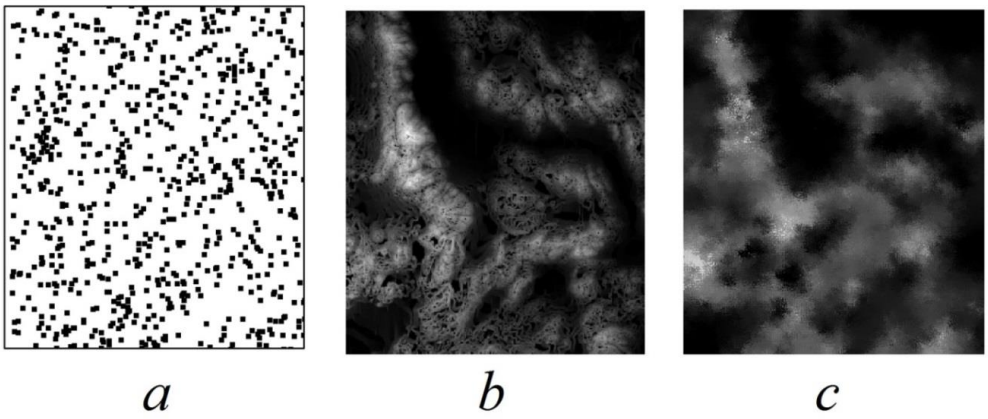

Figure 15: Simulation of variable V using 780 points: (a) data samples, (b) reference image, (c) joint simulation result.

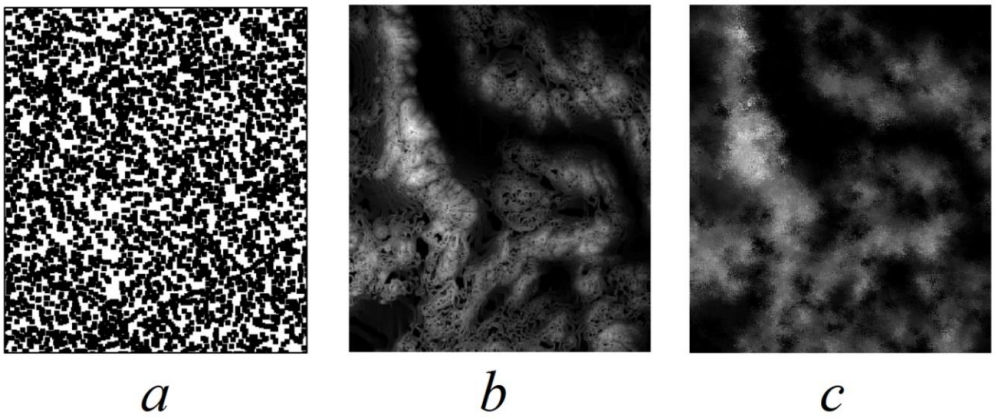

Figure 16: Simulation of variable V using 3900 points: (a) data samples, (b) reference image, (c) joint simulation result.

References

Arpat, B, Caers, J, 2007. Stochastic simulation with patterns. Mathematical Geosciences, 39:177–203.

Boucher, A, Dimitrakopoulos, R, 2009. Block-support simulation of multiple correlated variables. Mathematical Geosciences, 41(2):215–237.

Boucher, A, Dimitrakopoulos, R, 2012. Multivariate block-support simulation of the Yandi iron ore deposit, Western Australia. Mathematical Geosciences, 44(4): 449–468

Chatterjee S, Dimitrakopoulos R, Mustapha H, 2012. Dimensional reduction of pattern-based simulation using wavelet analysis. Mathematical Geosciences, 44:343–374.

Chiginova, T, Hu, L Y, 2008. Multiple-point simulations constrained by continuous auxiliary data. Mathematical Geosciences, 40:133–146.

Chiles, J P, Delfiner, P, 2012. Geostatistics, modelling spatial uncertainty (2nd ed). Wiley, New York, pp. 238–299

Comon, P, 1994. Independent Component Analysis: A new concept? Signal Processing, 36(3):287–314.

David, M, 1988. Handbook of applied advance geostatistical ore reserve estimation, Elsevier, Amsterdam, p. 216.

De Iaco, S, Maggio, S, 2011. Validation techniques for geological patterns simulations based on variogram and multiple-point statistics. Mathematical Geosciences, 43(4):483–500.

De Vries, L M, Carrera, J, Falivene, O, Gratacos, O, Slooten, L J, 2008. Application of multiple point geostatistics to non-stationary images. Mathematical Geosciences, 41:29–42.

Dimitrakopoulos, R, 2011. Stochastic optimization for strategic mine planning: A decade of developments. Journal of Mining Science, 47(2):138–150.

Dimitrakopoulos, R, Mustapha, H, Gloaguen, E, 2010. High-order statistics of spatial random fields: Exploring spatial cumulants for modelling complex, non-Gaussian and non-linear phenomena. Mathematical Geosciences, 42(1):65–97.

Fonseca, M, Dimitrakopoulos, R, 2003. Assessing risk in grade tonnage curves in a complex copper deposit, northern Brazil, based on an efficient joint simulation of multiple correlated variables. Application of Computers and Operations Research in the Mineral Industries. South African Institute of Mining and Metallurgy. pp. 373–382.

Goodfellow, R, Albor Consuegra, F, Dimitrakopoulos, R, Lloyd, T, 2012. Quantifying multi-element and volumetric uncertainty, Coleman McCreedy deposit, Ontario, Canada. Computers & Geosciences, 42, 71–78.

Goovaerts, P, 1997. Geostatistics for natural resources evaluation. Oxford University Press, Oxford, 483 p.

Hyvarinen, A, 1999. Survey on independent component analysis. Neural Computing Surveys, 2:94–128.

Isaaks, E, Srivastava M, 1990. Applied Geostatistics, Oxford University Press, Oxford, pp. 67–106.

Journel, A G, 2005. Beyond covariance: The advent of multiple-point geostatistics. Geostatistics Banff 2004, Volume 14 of the series Quantitative Geology and Geostatistics, pp. 225–233.

Journel, A G, 2007. Roadblocks to the evaluation of ore reserves – the simulation overpass and putting more geology into numerical models of deposits. Orebody Modelling and Strategic Mine Planning – Uncertainty and Risk Management, The AusIMM, Spectrum Series 14, 2nd Ed., pp. 29–32.

Lebedev, N N, 1965. Special Functions and their Applications. Prentice-Hall Inc., New York, pp. 182.

Honarkhah, M, 2011. Stochastic simulation of patterns using distance-based pattern modeling. Ph.D. thesis, Stanford University, Stanford, Ca.

Morton, J, Lim, L-H, 2009. Principal cumulant component analysis, preprint. http://www.stat.uchicago.edu/~lekheng/work/pcca.pdf

Mueller, U A, Ferreira, J, 2012. The U-WEDGE transformation method for multivariate geostatistical simulation. Mathematical Geosciences, 44(4):427–448.

Mustapha, H, Dimitrakopoulos, R, 2010a. High-order stochastic simulations for complex non-Gaussian and non-linear geological patterns. Mathematical Geosciences, 42(5):455–473.

Mustapha, H, Dimitrakopoulos, R, 2010b. An algorithm for geological patterns recognition using high-order spatial cumulants. Computers and Geosciences, 36(3):313–334.

Mustapha, H, Dimitrakopoulos, R, 2011. HOSIM: A high-order stochastic simulation algorithm for generating three-dimensional complex geological patterns. Computers and Geosciences, 37(9): 1242–1253

Mustapha, H, Dimitrakopoulos, R, Chatterjee, S, 2011. Geologic heterogeneity representation using high-order spatial cumulants for subsurface flow and transport simulations. Water Resources Research, 47(8):W08536.

Mustapha, H, Chatterjee, S, Dimitrakopoulos, R, 2014. CDFSIM: efficient stochastic simulation through decomposition of cumulative distribution functions of transformed spatial patterns. Mathematical Geosciences, 46(1):95–123.

Osterholt, V, Dimitrakopoulos, R, 2007. Simulation of wireframes and geometric features with multiple-point techniques: Application at Yandi iron ore deposit. Orebody modelling and strategic mine planning, AusIMM, Spectrum Series 14, 2nd Ed., 95–124.

Perry, J M, 1977. A class of conjugate gradient algorithms with a two-step variable-metric memory. Discussion Paper 269, Center for Mathematical Studies in Economics and Management Science, Northwestern University.

Rondon, O, 2012. Teaching Aid: Minimum/Maximum Autocorrelation Factors for Joint Simulation of Attributes. Mathematical Geosciences, 44(4):469–504.

Rosenblatt, M, 1985. Stationary Sequences and Random Fields. Birkhäuser, Boston.

Straubhaar J, Renard P, Mariethoz G, Froidevaux R, Besson O, 2011. An improved parallel multiple-point algorithm using a list approach. Mathematical Geosciences, 43(3):305–328.

Strebelle S, Cavelius C, 2014. Solving speed and memory issues in multiple-point statistics simulation program SNESIM. Mathematical Geosciences, 46(2):171–186.

Tjelmeland, H, 2013. Construction of Binary Multi-grid Markov Random Field Prior Models from Training Images. Mathematical Geosciences, 45(4):383–409.

Wackernagel, H, 1998. Principal component analysis for autocorrelated data: a geostatistical perspective. Technical report No 22/98/G, Centre de Géostatistique, École des Mines de Paris.

Zhang, T, Switzer, P, Journel, A, 2006. Filter-based classification of training image patterns for spatial simulation. Mathematical Geology, 38:63–80.