Les Cahiers du GERAD ISSN: 0711–2440

High-order block support spatial simulation and application at a gold deposit

J.P. de Carvalho, R. Dimitrakopoulos, I. Minniakhmetov

G–2018–85

November 2018

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2018
– Bibliothèque et Archives Canada, 2018

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Legal deposit – Bibliothèque et Archives nationales du Québec, 2018
– Library and Archives Canada, 2018

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

High-order block support spatial simulation and application at a gold deposit

Joao Pedro de Carvalho
Roussos Dimitrakopoulos
Ilnur Minniakhmetov

GERAD, Montréal (Québec), Canada, H3T 2A7

COSMO–Stochastic Mine Planning Laboratory, McGill University, Montréal (Québec) Canada, H3A 0E8

Department of Mining and Materials Engineering, McGill University, Montréal (Québec), Canada, H3A 0E8

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

November 2018
Les Cahiers du GERAD
G–2018–85

Abstract: High-order sequential simulation methods have been developped as an alternative to existing frameworks to facilitate the modelling of spatial complexity of non-Gaussian variables of interest. The high-order simulation approaches address the modelling of curvilinear features and spatial connectivity of extreme values that are common in mineral deposits, petroleum reservoirs, water aquifers and other geological phenomena. This paper presents a new high-order simulation method that generates realizations directly at the block support, conditioned to the available data at point support scale. Under the context of sequential high-order simulation, the method estimates, at each block location, the cross-support joint probability density function using Legendre-like splines as the set of basis functions. The proposed method adds previously simulated blocks to the set of conditioning data, which initially contains the available data at point support. A spatial template is defined by the configuration of the block to be simulated and related conditioning values in both support scales and is used to infer additional high-order statistics from a training image. Through testing the proposed method with an exhaustive dataset, it is shown that simulated realizations reproduce major structures and high-order relations of data. The practical intricacies of the proposed method are demonstrated in an application at a gold deposit.

Keywords: sequential high-order simulation, block support, cross-support joint probability density function

1 Introduction

Stochastic simulation methods are used to quantify spatial uncertainty and variability of pertinent attributes of natural phenomena in the geo-sciences and geo-engineering. Initial simulation methods were based on Gaussian assumprions and second-order statistics of corresponding random field models (Journel and Huijbregts 1978; David 1988; Goovaerts 1997). To address limits of Gaussian approaches, multiple point statistics (MPS) based simulation methods have been introduced (Guardiano and Srivastava 1993; Strebelle 2002; Zhang et al. 2006; Arpat and Caers 2007; Remy et al. 2009; Mariethoz et al. 2010; Mariethoz and Caers 2014; Mustapha et al. 2014; Chatterjee et al. 2016; Li et al. 2016; Zhang et al. 2017) to remove distributional assumptions, as well as to enable the reproduction of complex curvilinear and other geologic features by replacing the random field model with a framework built upon the extraction of multiple point patterns from a training image (TI) or geological analogue. The main limitations of MPS methods are that they do not explicitly account for high-order statistics, or provide consistent mathematical models while they generate TI-driven realizations. Previous studies have shown resulting realizations that comply with the TI used, but do not necessarily reproduce the spatial statistics inferred from the data (Osterholt and Dimitrakopoulos 2007; Goodfellow et al. 2012). As an alternative to the above limitations, a high-order simulation (HOSIM) framework has been proposed as a natural generalization of the second-order based random field paradigm (Dimitrakopoulos et al. 2010; Mustapha and Dimitrakopoulos 2010a, b, 2011, Minniakhmetov and Dimitrakopoulos 2017a, b; Minniakhmetov et al. 2018; Yao et al. 2018). The HOSIM framework does not make any assumptions about the data distribution, and the resulting realizations reproduce the high-order spatial statistics of the data. Similar to the MPS and most Gaussian simulation approaches, HOSIM methods generate realizations at the point support, whereas, in most major areas of applications, simulated realizations need be at the block support scale. Typically, the change of support needed is addressed by generating simulated realizations at a very dense grid of nodes that is then post-processed to generate realizations at the block support size needed, which is a computationally demanding process, as related configurations may require extremely dense grids that may be in the order of many millions to billions of nodes. Thus, there is a need for computationally efficient methods that simulate directly at the block support scale.

In the context of the conventional second-order geostatistics, direct block support simulation has been proposed. Godoy (2003) presents an approach, termed "direct block simulation," that discretizes each block into several internal nodes, but only stores a single block value in memory for the next group simulation. This mechanism drastically reduces the amount of data stored in memory and saves considerable computational effort. Boucher and Dimitrakopoulos (2009) expand the sequential direct block simulation method to incorporate multiple correlated variables by applying the min/max autocorrelation function (MAF). Emery (2009) uses an explicit change of the support model and directly simulates at block support. Although efficient, these methods carry all the limitations of a Gaussian simulation framework and related spatial connectivity is limited to two-point spatial statistics, thus, remain unable to characterize non-Gaussian variables, complex non-linear geological geometries and the critically important connectivity of extreme values (Journel 2018). Thus, alternatives are needed.

High-order sequential simulation methods use high-order spatial cumulants to describe complex geologic configurations and high-order connecticity. At the same time, simulated realizations remain consistent with respect to the statistics of the available data, while capitalizing on the additional information that TIs can provide. Dimitrakopoulos et al. (2010) describe these high-order spatial cumulants as combinations of moment statistical parameters. Mustapha and Dimitrakopoulos (2010a) propose a high-order simulation algorithm, where the conditional probabilities density functions (cpdf) are approximated by Legendre polynomials and high-order spatial cumulants. A template is defined based on the central node to be simulated and the nearest conditioning data. The replicates of this configuration are obtained from both the data and TI, and are used as input for the calculation of the Legendre coefficients in the cpdf approximation. Advantages of this method lie in the absence of assumption on the distribution of the data and in being a data-driven approach. Minniakhmetov et al. (2018) replace the Legendre polynomial by Legendre-like splines as the basis function for the estimation of conditional probabilities. Results show a more stable approximation of the related cpdf. Improving upon the computational performance, Yao et al. (2018b) propose a new approach, where the calculation of the cpdf is simplified and no explicit calculation of cumulants is required. Although effective, the methods above are performed at point support scale.

This paper presents a new method that generates high-order stochastic simulations directly at the block support scale. The technique considers overlapping grids representing a study area at two support scales, point and block support, where the simulation process is implemented at the latter support. In the sequential simulation pocess followed, only the initial point support data and previously simulated blocks are added to the set of conditioning values, thus drastically reducing the number of elements stored in memory. The block to be simulated and the nearest conditioning data, at the point or block support, define the spatial configuration of the template used. Similarly, the TI is represented in both support scales to provide replicates of related spatial template configuration. The conditional cross-support joint density function estimated at each block is approximated by Legendre-like splines.

The remainder of the paper is organized as follows. First, the proposed model for high-order block support simulation is presented. Subsequently, a case study in a controlled environment assesses the performance of the current approach. Next, the method is applied to an actual gold deposit to demonstrate its practical aspects. Conclusions follow.

2 High-order block support simulation

2.1 Sequential simulation

In the following description, the index V relates to elements at the block support, while Prepresents point support. Consider a stationary and ergodic non-Gaussian random field (RF) $Z_P(u_j)$ in R^n, where u_j defines the location of nodes j in the domain $D \subseteq R^n$. Now consider a transformation function that takes the above point support RF to the block support RF

$$Z_V(v) = \frac{1}{|V|} \int_{u_j \in v} Z_P(u_j)\, du_j. \tag{1}$$

Now $Z_V(v_i)$ is also a RF, indexed as $v_i \in D \subseteq R^n, i = 1, \ldots, N_V$ where N_V represents the total number of blocks to be simulated within the domain $D \subseteq R^n$. $Z_V(v_i)$ is the upscaled RF from $Z_P(u_j)$ considering all nodes u_j that are discretized within the block centred in v_i, where V is the volume.

The outcomes from the above RFs are denoted as $z_j^P = z_P(u_j)$ and $z_i^V = z_V(v_i)$, respectively for the point and the block support RF $Z_P(u_j) = Z_j^P$ and $Z_V(v_i) = Z_i^V$. Herein, the objective is to simulate a realization of the RF Z_i^V given the set of initial conditioning values at point support are denoted as $d_p = \{z_1^P, \ldots, z_{N_P}^P\}$, N_P being the total of conditioning point support values. According to the sequential simulation theory in the geostatistical field, the joint probability density function (jpdf) $f_{Z^V_{1,2,\ldots,N_V}}$ can be decomposed in the products of their respective univariate distributions (Journel and Alabert 1989; Journel 1994; Goovaerts 1997; Dimitrakopoulos and Luo 2004)

$$\begin{aligned} f_{Z^V_{1,2,\ldots,N_V}}\left(z_1^V, z_2^V, \ldots, z_{N_V}^V \left| d_P\right.\right) &= f_{Z_1^V}\left(z_1^V \left| d_P\right.\right) f_{Z_V}\left(z_2^V, \ldots, z_{N_V}^V \left| d_P, z_1^V\right.\right) \\ &= f_{Z_1^V}\left(z_1^V \left| d_P\right.\right) f_{Z_2^V}\left(z_2^V \left| d_P, z_1^V\right.\right) \ldots f_{Z_{N_V}^V}\left(z_{N_V}^V \left| d_P, z_1^V, z_2^V, \ldots, z_{N_V-1}^V\right.\right). \end{aligned} \tag{2}$$

According to the Equation 2, each block v^k is simulated based on the estimation of the conditional cross-support probability density function $f_{Z_k^V}\left(z_k^V \left| d_P, z_1^V, z_2^V, \ldots, z_{k-1}^V\right.\right)$, which according to Bayes' rule (Stuart and Ord 1987) is

$$f_{Z_k^V}\left(z_k^V \left| d_P, z_1^V, z_2^V, \ldots, z_{k-1}^V\right.\right) = \frac{f_{Z_1^V,\ldots,Z_k^V}\left(z_k^V, d_P, z_1^V, z_2^V, \ldots, z_{k-1}^V\right)}{f_{Z_1^V,\ldots,Z_{k-1}^V}\left(d_P, z_1^V, z_2^V, \ldots, z_{k-1}^V\right)}. \tag{3}$$

It is sufficient to approximate only the cross-support joint probability density function

$$f_{Z_1^V,\ldots,Z_k^V}\left(z_k^V, d_P, z_1^V, z_2^V, \ldots, z_{k-1}^V\right)$$

since its marginal distribution can be theoretically obtained from

$$f_{Z_1^V,\ldots,Z_{k-1}^V}\left(d_P, z_1^V, z_2^V, \ldots, z_{k-1}^V\right) = \int_D f_{Z_1^V,\ldots,Z_k^V}\left(z_k^V, d_P, z_1^V, z_2^V, \ldots, z_{k-1}^V\right) dv_k. \tag{4}$$

In this paper, the cross-support joint probability density $f_{Z_1^V,\ldots,Z_k^V}\left(z_k^V, d_P, z_1^V, z_2^V, \ldots, z_{k-1}^V\right)$ is approximated using Legendre-like orthogonal splines (Wei et al. 2013; Minniakhmetov et al. 2018).

2.2 Joint probability density function approximation

For simplicity, let $f(x)$be the pdf of a random variable X defined in $\Omega = [a, b]$ and let $\varphi_1(z), \varphi_2(z), \ldots$be a set of set of orthogonal functions defined in the same space Ω. Then, a fixed number ω of those orthogonal functions can approximate $f(x)$ (Lebedev 1965; Mustapha and Dimitrakopoulos 2010a; Minniakhmetov et al. 2018; Yao et al. 2018), when multiplied by coefficients L_i.

$$f(z) = \sum_{i=0}^{\omega} L_i \varphi_i(z). \tag{5}$$

Since the sets of functions are orthogonal,

$$\int_a^b \varphi_i(z)\, \varphi_j(z)\, dz = \delta_{ij}, \tag{6}$$

where δ_{ij} is the Kronecker delta indexed by i and j, such that it gets a unitary value if $i = j$; and 0 otherwise. By using a definition of the expected value of one of a basis function

$$E[\varphi_i(z)] = \int_a^b \varphi_i(z)\, f(z)\, dz. \tag{7}$$

Replacing $f(z)$ as in Equation 5, it is obtained

$$E[\varphi_i(z)] \approx \int_a^b \varphi_i(z) \sum_{j=0}^{\omega} L_j \varphi_j(z)\, dz = \sum_{j=0}^{\omega} L_j \int_a^b \varphi_j(z)\, \varphi_i(z)\, dz = \sum_{j=0}^{\omega} L_j \delta_{ij} = L_i. \tag{8}$$

L_i coefficient can be experimentally obtained from an available sample, thus $f(z)$ is approximated by Equation 5.

Moving to the multivariate cross-support case, at every block location v^k the cross-support jpdf $f_{Z_1^V,\ldots,Z_k^V}$ $\left(z_k^V, d_P, z_1^V, z_2^V, \ldots, z_{k-1}^V\right)$ can be defined in a similar sense. Moving forward, the above function is defined as $f\left(z_0^V, \ldots, z_{n_V}^V, z_1^P, \ldots, z_{n_P}^P\right)$ for simplistic notation ensuring a better understanding of variables in both block and point support layers. Also note that, without loss of generality, z_0^V is the value to be simulated at the $v_0$location. The cross-support jpdf is defined in the domain$[a, b]^{n_V+1} \times [a, b]^{n_P}$, where n_P and n_V are the maximum number of points and blocks considered, respectively, for the approximation. Similarly to the univariate case, the cross-support jpdf can be approximated as

$$\begin{aligned} f\left(z_0^V, \ldots, z_{n_V}^V, z_1^P, \ldots, z_{n_P}^P\right) \approx \\ \sum_{k_0^V}^{\omega} \cdots \sum_{k_V^{n_V}}^{\omega} \sum_{k_P^1}^{\omega} \cdots \sum_{k_P^{n_P}}^{\omega} \left[L_{k_0^V \ldots k_{n_V}^V k_1^P \ldots k_{n_P}^P} \varphi_{k_0^V}\left(z_0^V\right) \ldots \varphi_{k_{n_V}^V}\left(z_{n_V}^V\right) \varphi_{k_1^P}\left(z_1^P\right) \ldots \varphi_{k_{n_P}^P}\left(z_{n_P}^P\right) \right]. \end{aligned} \tag{9}$$

Those $L_{i\ldots jk\ldots l}$ coefficients can be calculated experimentally since they can be obtained from the orthogonality property of the basis functions. Following the definition of the expected value of a basis function, it is expressed as:

$$E\left[\varphi_i\left(z_0^V\right)\ldots\varphi_j\left(z_{n_V}^V\right)\varphi_k\left(z_1^P\right)\ldots\varphi_l\left(z_{n_P}^P\right)\right] = \int_a^b\cdots\int_a^b\int_a^b\cdots\int_a^b \varphi_i\left(z_0^V\right)\ldots\varphi_j\left(z_{n_V}^V\right)\varphi_k\left(z_1^P\right)\ldots \varphi_l\left(z_{n_P}^P\right) f\left(z_0^V,\ldots,z_{n_V}^V,z_1^P,\ldots,z_{n_P}^P\right) dz_0^V\ldots dz_{n_V}^V dz_1^P\ldots dz_{n_P}^P. \tag{10}$$

Replacing $f\left(z_V^0,\ldots,z_V^{n_V},z_P^1,\ldots,z_P^{n_P}\right)$ as in Equation 9, it is obtained

$$E\left[\varphi_i\left(z_0^V\right)\ldots\varphi_j\left(z_{n_V}^V\right)\varphi_k\left(z_1^P\right)\ldots\varphi_l\left(z_{n_P}^P\right)\right] \approx \int_a^b\cdots\int_a^b\int_a^b\cdots\int_a^b \varphi_i\left(z_0^V\right)\ldots\varphi_j\left(z_{n_V}^V\right)\varphi_k\left(z_1^P\right)\ldots\varphi_l\left(z_{n_P}^P\right)\sum_{k_0^V}^{\omega}\cdots\sum_{k_{n_V}^V}^{\omega}\sum_{k_1^P}^{\omega}\cdots \sum_{k_{n_P}^P}^{\omega}\left[L_{k_0^V\ldots k_{n_V}^V k_1^P\ldots k_{n_P}^P}\varphi_{k_0^V}\left(z_0^V\right)\ldots\varphi_{k_{n_V}^V}\left(z_{n_V}^V\right)\varphi_{k_1^P}\left(z_1^P\right)\ldots\varphi_{k_{n_P}^P}\left(z_{n_P}^P\right)\right]$$

$$dz_0^V\ldots dz_{n_V}^V dz_1^P\ldots dz_{n_P}^P = \sum_{k_0^V}^{\omega}\cdots\sum_{k_{n_V}^V}^{\omega}\sum_{k_1^P}^{\omega}\cdots\sum_{k_{n_P}^P}^{\omega}\left[L_{k_0^V\ldots k_{n_V}^V k_1^P\ldots k_{n_P}^P}\int_a^b\cdots\int_a^b\int_a^b\cdots\int_a^b\varphi_i\left(z_0^V\right)\varphi_{k_0^V}\left(z_0^V\right)\ldots \varphi_j\left(z_{n_V}^V\right)\varphi_{k_V^{n_V}}\left(z_{n_V}^V\right)\varphi_k\left(z_1^P\right)\varphi_{k_P^1}\left(z_1^P\right)\ldots\varphi_l\left(z_{n_P}^P\right)\varphi_{k_{n_P}^P}\left(z_{n_P}^P\right)\right]$$

$$dz_0^V\ldots dz_{n_V}^V dz_1^P\ldots dz_{n_P}^P = \sum_{k_0^V}^{\omega}\cdots\sum_{k_{n_V}^V}^{\omega}\sum_{k_1^P}^{\omega}\cdots\sum_{k_{n_P}^P}^{\omega}\left[L_{k_0^V\ldots k_{n_V}^V k_1^P\ldots k_{n_P}^P}\delta_{ik_0^V}\ldots\delta_{jk_{n_V}^V}\delta_{kk_1^P}\delta_{lk_{n_P}^P}\right] = L_{i\ldots jk\ldots l}. \tag{11}$$

Now, to determine $L_{i\ldots jk\ldots l}$, the expected value from Equation 11 is calculated from replicates of the training image according to a template defined from simulation grid and sampling data.

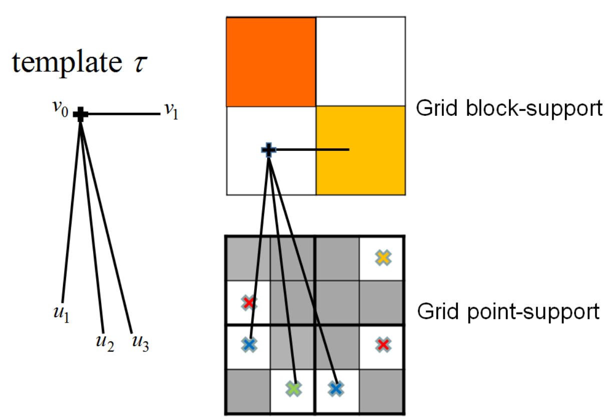

Figure 1: An example of a template τ with conditioning data capturing values in both point and block support sizes.

Let $\tau = [v_0,\ldots,v_{n_V},u_1,\ldots,u_{n_P}]$ be a template as in Figure 1, where $v_0$and $v_1$represent locations at block support, and u_1,u_2 and $u_3$represent point support locations. v_0 is the location of the block to be simulated and n_P and n_V are, respectively, the total number of points and blocks used as conditioning. Note that limited conditioning values are chosen in order of Euclidean proximity from the central block to be simulated. Having the specified templateτ, the TI is scanned, and the replicates of such template are retrieved. Note that τ has elements that belong to the point and block support sizes. Similarly, the TI requires to be available both scales. Therefore, assuming a TI input in the point support scale, it is rescaled to block support, and both are retrieved during the simulation process, each one in its respective layer.

The algorithm for the block support high-order simulation method is as follows:

1. Upscale the TI from point to block support.
2. Define a random path to visit all the unsampled block locations on the simulation grid.
3. At each v^0 block location:
   (a) Find the nearest conditioning point and block support values.
   (b) Obtain the template τ according to the configuration of the central block and related conditioning values in both support sizes.
   (c) Scan the training images, searching for replicates of the template τ and corresponding values.
   (d) Calculate all the spatial cross-support coefficients $L_{i\ldots jk\ldots l}$ using Equation 11.
   (e) Derive the conditional cross-support jpdf $f_{Z_0^V}\left(z_0^V \,\middle|\, d_P, z_1^V, z_2^V, \ldots, z_k^V\right)$ according to Eqations 9 and 11.
   (f) Draw a uniform value from $[0,1]$ to sample z_0^V from the conditional cumulative distribution derived from the above.
   (g) Add z_0^V to the simulation grid at block support scale so that it can be a conditioning value for the next block.
4. Repeat steps 2 and 3 for additional realizations.

2.3 Approximation of a joint probability density using Legendre-like orthogonal splines

The current paper uses the Legendre-like splines (Wei et al. 2013; Minniakhmetov et al. 2018) as means to obtain the basis function above mentioned. In short, those splines are a combination of Legendre polynomials (Lebedev 1965) up to the r^{th}order and linear combinations of B-splines(De Boor 1978). B-splines are a particular class of piecewise polynomials (splines) connected by some condition of continuity, and by itself do not form an orthogonal basis. Thus, as introduced in Wei et al. (2013), the first $r+1$ splines are the Legendre polynomials, which can be defined as in (Lebedev 1965)

$$\varphi_r = \frac{1}{2^r r!}\left(\frac{d^r}{dz^r}\right)\left[\left(z^2-1\right)^r\right], \quad -1 \le z \le 1. \tag{12}$$

The additional functions are constructed given the domain T

$$T = \{\underbrace{a, a, \ldots, t_0 = a}_{r+1} < t_1 \le t_2 \le \ldots \le t_{m_{\max}} < \underbrace{t_{m_{\max}+1} = b, b, \ldots, b}_{r+1}\}, \tag{13}$$

where the t_i elements are referred to as knots and $m_{\max}$represents the maximum number of knots. The final Legendre-like splines are defined as

$$\varphi_{r+m}(t) = \frac{d^{r+1}}{dt^{r+1}} f_m(t), \quad m = 1 \ldots m_{\max}, \tag{14}$$

The $f_m(t)$ is the determinant of the following matrix

$$f_m(t) = \det\begin{pmatrix} B_{-r,2r+1,m}(t) & B_{-r+1,2r+1,m}(t) & \cdots & B_{-r+m-1,2r+1,m}(t) \\ B_{-r,2r+1,m}(t_1) & B_{-r+1,2r+1,m}(t_1) & \vdots & B_{-r+m-1,2r+1,m}(t_1) \\ \vdots & \vdots & \ddots & \vdots \\ B_{-r,2r+1,m}(t_{m-1}) & B_{-r+1,2r+1,m}(t_{m-1}) & \cdots & B_{-r+m-1,2r+1,m}(t_{m-1}) \end{pmatrix}, \tag{15}$$

which is constructed with the auxiliary splines $B_{i,r,m}(t)$, of r^{th}order, obtained according to the recursive rule

$$B_{i,0,m} = \begin{cases} 1, & t_{i,m} \leq t \leq t_{i+1,m} \\ 0, & \text{otherwise} \end{cases}, \tag{16}$$
$$B_{i,r,m}(t) = \frac{t - t_{i,m}}{t_{i+r-1,m} - t_{i,m}} B_{i,r-1,m}(t) + \frac{t_{i+r,m} - t}{t_{i+r,m} - t_{i+1,m}} B_{i+1,r-1,m}(t).$$

These auxiliary functions are defined on the knot sequence $T_m = \{t_{i,m}\}_{i=-r}^{r+m+1}$, $m = 1 \ldots m_{\max} - 1$ and the $t_{i,m}$ term is defined as

$$t_{i,m} = \begin{cases} a, & -r \leq i \leq 0 \\ t_i, & 1 \leq i \leq m \\ b, & m+1 \leq i \leq m+r+1 \end{cases}. \tag{17}$$

3 Testing with an exaustive dataset

The method outlined above is tested using the two-dimensional image of the Walker-Lake dataset (Isaaks and Srivastava 1989). The exhaustive dataset comprises two correlated variables U and V of sizes 260x300 pixels. Random stratified sampling is used to retrieve 234 values from or 0.3% of the exhaustive image V to be used as the data set in the direct block simulation of V, to test the proposed method. The full dataset V is converted from the point to a block support representation by an averaging over 5x5 pixels. This block support version is referenced here as the fully-known reference image and it is used for comparisons. Figure 2 shows V at the point and block support scale, as well as the dataset to be used. The image U is chosen as the training image in the simulation process. Figure 3 presents the TI in both point and block support (5x5 unit size) scales. To help the method find more meaningful spatial patterns of the potential conditioning templates, the TI is scaled to the range of values of the dataset. Histograms of the exhaustively known image, TI and dataset are displayed in Figure 4 and basic statistics are presented in Table 1.

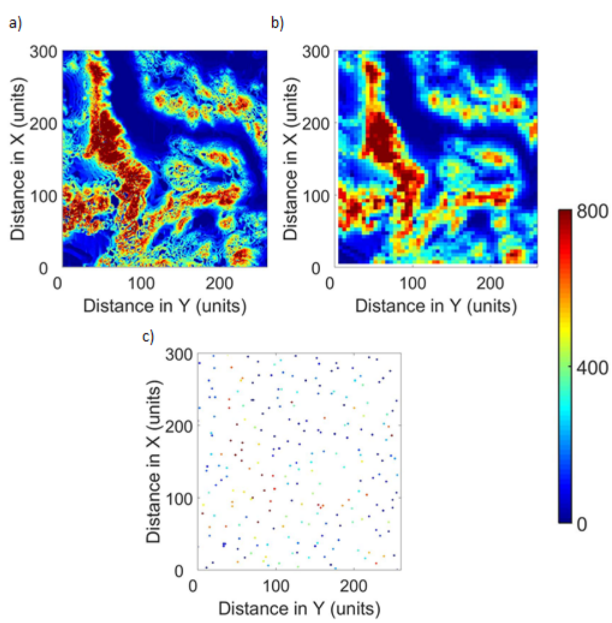

Figure 2: Exhaustive image “V” (a) at point support, (b) at block support, and (c) 234 samples from image in (a) – 234 samples.

The test conducted consists of generating 15 simulated realizations of the V dataset at block support, using the data and the training image mentioned above. Figure 5 presents three of the simulated realizations generated and Table 2 shows the statistics relative to the average of the 15 simulations, training image and

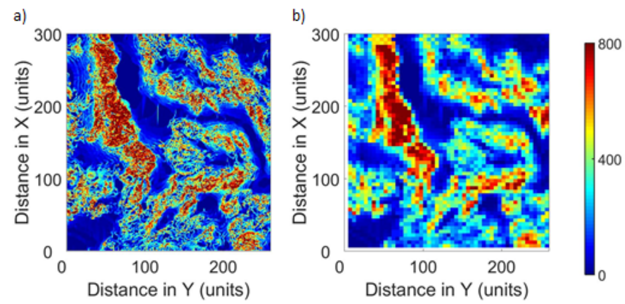

Figure 3: Training image "U" and the (a) point support scale, and (b) block support scale.

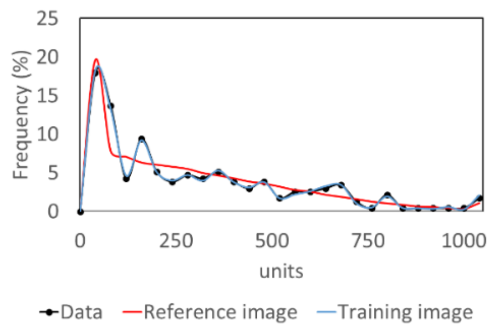

Figure 4: Histogram of data, reference and training image at point support scale.

Table 1: Basic stats of dataset, training image and fully known image at point support scale.

Basic stats	Dataset	Reference image	Training image
Average	277.0	278	278.5
Median	193.5	221.3	193.6
Variance	68926	62423	70385

reference image at block scale. A comparison of Figures 2(b) and 5 suggest that the simulations reproduce the main structures of low and high values of the fully-known reference image V. The histograms and variograms presented in Figures 6 and 7 respectively reasonably follow the behaviour exhibited by the variogram model from the data and the training image. Note that the variograms of data is computed at point scale and rescaled to represent the corresponding volume-variance relation (Journel and Huijbregts 1978).

Table 2: Basic stats of the average of the simulations, training image and reference image at block support scale.

Basic stats	Simulations	Training image	Reference image
Average	280.6	278.5	278.0
Median	234.6	237.0	235.4
Variance	51764	46518	52304

Spatial cumulants (Dimitrakopoulos et al. 2010) quantify the spatial relationships between three and more points and are used herein to assess high-order spatial patterns. The third-order cumulant maps are presented along with the template used for its calculation in Figure 8. Figure 9 shows the fourth-order cumulant map, where three slices of the complete cumulant map and the related template are displayed. In both figures, the colors range from blue to red, representing lower to higher spatial inter-correlation between values. Note that the reference and training image high-order maps are calculated on the block support scale, while the cumulant maps related to each simulation is averaged to a single map using the 15 stochastic simulated realizations at block support. During the calculation of the high-order spatial statistics from the data, only a few replicates are obtained and Figure 8(a) presents a smooth interpolation using B-splines. Regarding the third-order maps, the average of the simulations match the spatial features observed in the data and fully-known dataset. The fourth-order cumulant map reproduces the characteristics that are closer

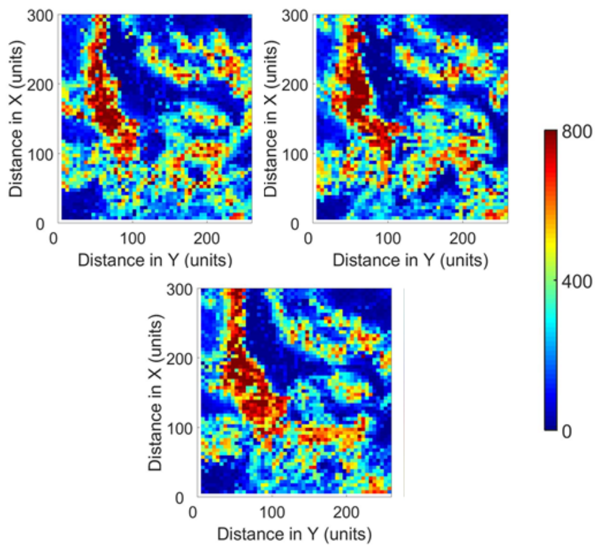

Figure 5: Example of 3 simulated realizations of the Walker Lake reference image V.

to the TI than the fully-known image, as expected. Note that, by explicitly calculating the spatial high-order cumulants, the information received from the training image to infer local cross-support distributions is conditioned to the data.

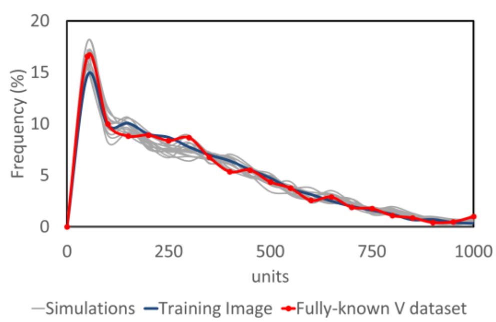

Figure 6: Histograms of the simulations at block support, and comparison with reference and training image also at block support.

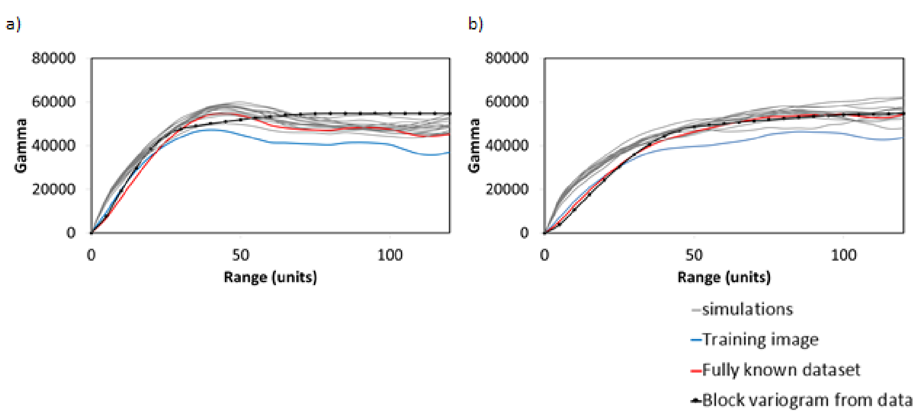

Figure 7: Variograms of simulated realizations, exhaustive image, TI, and variogram from data rescaled to block support variance; figure (a) shows the WE direction, and (b) the NS direction.

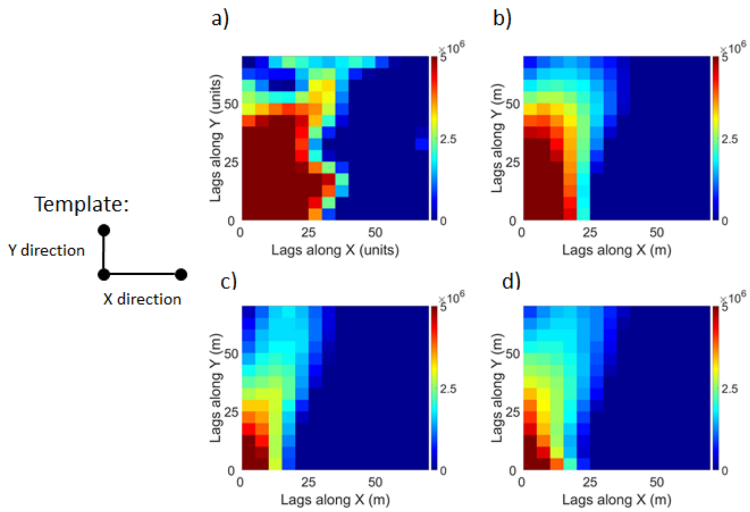

Figure 8: Third-order cumulant maps for (a) point support data used, (b) fully-known block support image V, (c) training image, and (d) the average map of the 15 simulated realizations.

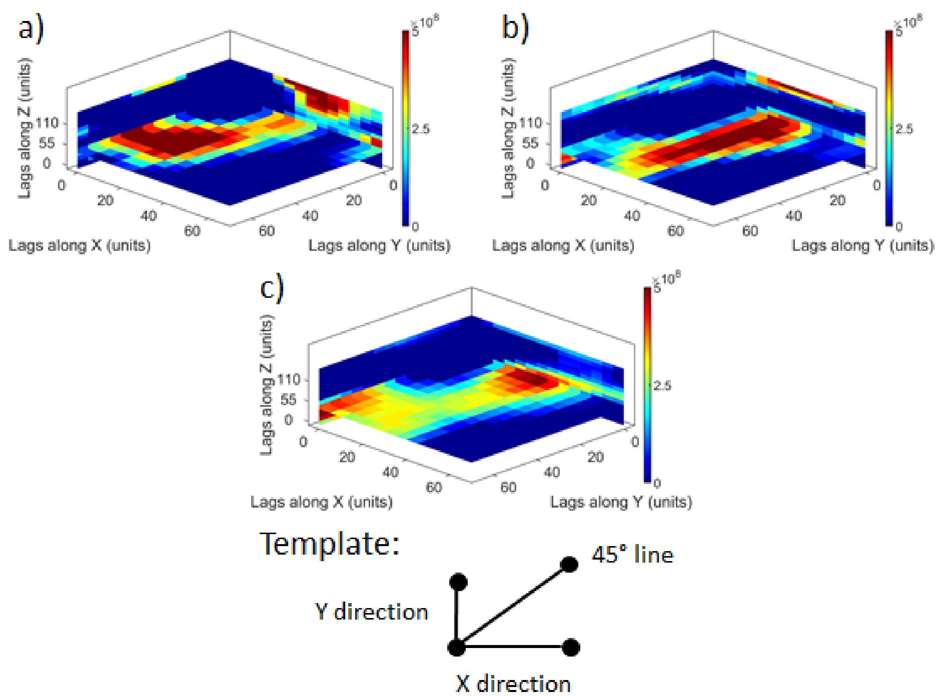

Figure 9: Slices of the fourth-order cumulant maps for (a) fully-known image V, (b) training image, and (c) average map of the 15 simulations, all at block support.

4 Application at a gold deposit

This section applies the proposed method at a gold deposit. The dataset comprises 2,300 drillholes that are spaced approximately at a 35 x 35 m^2 configuration covering an area of 4.5 km^2. The training image is defined on $405\times445\times43$ grid blocks of size $5\times5\times10$ m^3 and is based on blasthole samples. Both inputs are composited in a 10 m bench and are considered to be at point support. Figure 10 presents the drillholes available and the training image at block scale. The deposit to be simulated is represented by 510,800 blocks measuring 10x10x10 m^3 each.

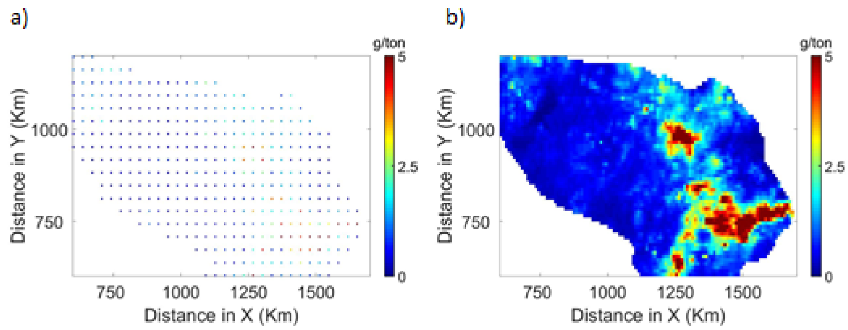

Figure 10: (a) Cross-section of the available drillhole locations, and (b) training image in the block support scale.

Fifteen simulated realizations are generated and cross-sections from two of them are presented in Figure 11 to show similarities with the data and TI in the corresponding cross-section in Figure 10. Notable is the reproduction of a sharp transition from high to low grades. Figure 12 shows the histograms of the simulations and TI at block support. Table 3 provides the related statistics. Variograms at block support are displayed in Figure 13, where the data variogram is regularized to reflect the corresponding volume-variance relation. The second-order spatial statistics of the simulations match reasonably with the pattern followed by the data and are close to those of the TI. Results for third- and fourth-order cumulants and related maps for the data, TI and simulated realizations are shown in Figures 14 and 15, respectively. It is noted that the high-order statistics of the simulated realizations match those of the data and TI.

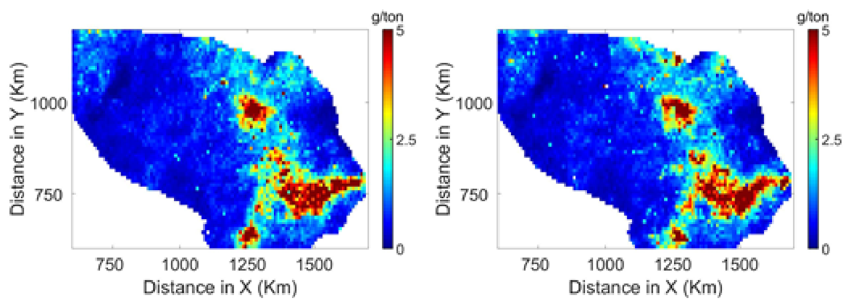

Figure 11: Cross-section of two simulated realizations.

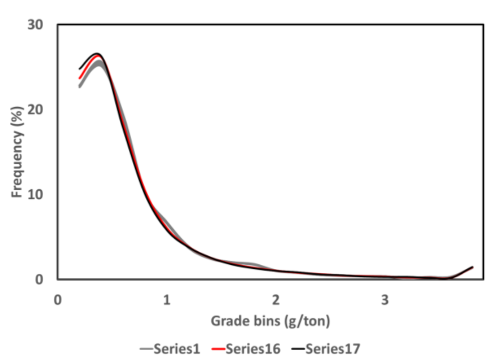

Figure 12: Histograms of simulated realizations and training image.

Table 3: Basic stats of the average of the simulations and training image at block support and dataset at point scale.

Basic stats	Simulations	TI block support	Data point support
Average	0.63	0.62	0.63
Median	0.40	0.39	0.39
Variance	0.74	0.76	1.81

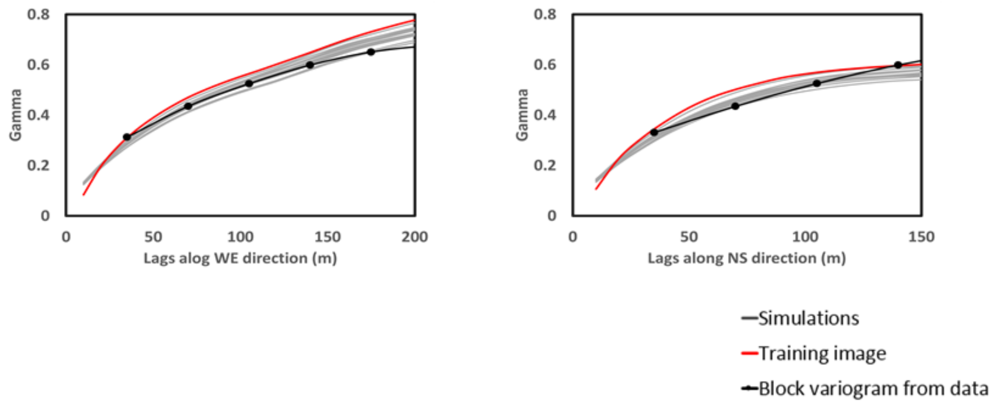

Figure 13: Variograms of simulated realizations and training image and data variograms rescaled to represent block variance; WE direction (left) and NS direction (right).

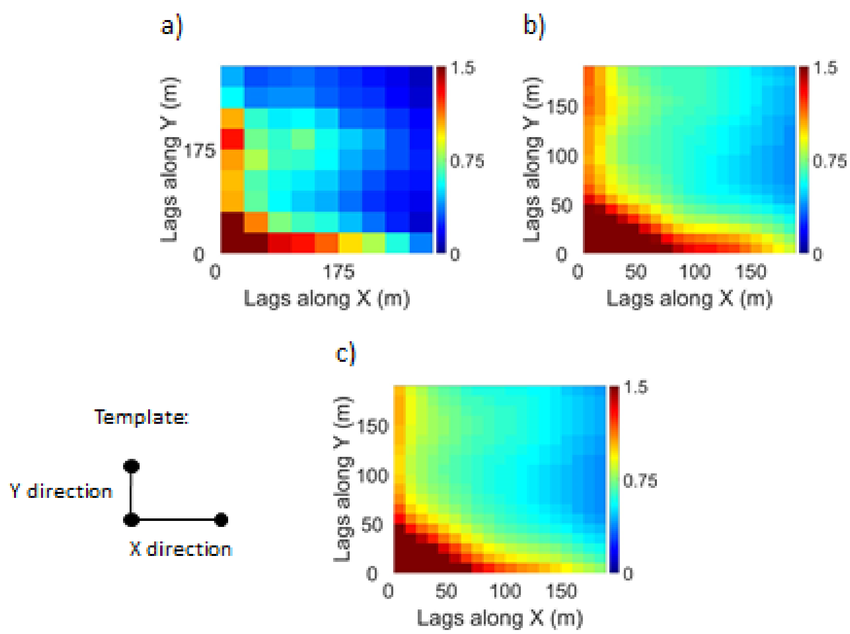

Figure 14: Third-order cumulant maps, obtained with the template in the left, for the (a) dataset, (b) training image at block support, and (d) average map of the 15 simulations.

Further highlighting the advantages of the proposed direct block high-order simulation method, one may note that, for the above case study, the run time of the related algorithm was approximately five and a half hours, while the point high-order simulation requires approximately twenty-four hours. Both approaches are tested with the same specifications and computing equipment: Intel® Core^{TM} i7-7700 CPU with 3.60 GHz, 16GB of RAM and running under Windows 7.

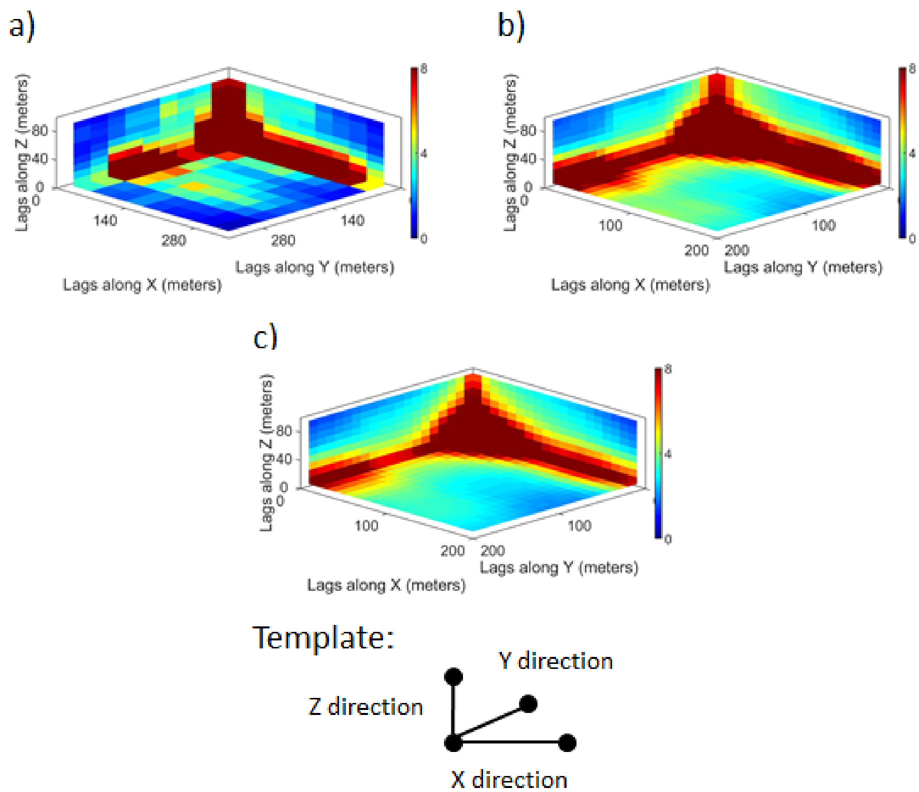

Figure 15: 3 slices of the fourth-order cumulant maps, obtained with the template in the bottom, for the (a) dataset, (b) training image at block support, and c) average map of the 15 simulations.

5 Conclusion

This paper presented a new high-order simulation method that simulates directly at block support scale by estimating, at every block location, the cross-support joint probability density function. Legendre-like splines are the set of a basis function used to approximate the above density function. The related coefficients are calculated from replicates of a spatial template employed. The latter template is generated from the configuration of the block to be simulated and associated conditioning values, whose supports can be both at the point and block support scales. The high-order character of the proposed direct block simulation method ensures that generated realizations reflect complex, non-linear spatial characteristics of the variables being simulated and reproduce the connectivity of extreme values.

The proposed algorithm was tested using an exhaustive image showing that the different realizations generated can reasonably reproduce spatial architectures observed in the exhaustive image. An application at a gold deposit shows the practical aspects of the method. In addition, it documents that the method works well while simulated realizations are shown to reproduce the spatial statistics of the available data up to the cumulants of the fourth-order that were calculated. Further work will focus on improving computational efficiency, generating training-images that are consistent with the high order relations in the available data, and extending the proposed method to jointly simulate multiple variables.

6 References

Arpat GB, and Caers J (2007) Conditional simulation with patterns. 39:(2):177–203. doi:10.1007/s11004-006-9075-3

Boucher A, and Dimitrakopoulos R (2009) Block simulation of multiple correlated variables. 41:(2):215–237. doi:10.1007/s11004-008-9178-0

Chatterjee S, Mustapha H, and Dimitrakopoulos R (2016) Fast wavelet-based stochastic simulation using training images. 20:(3):399–420. doi:10.1007/s10596-015-9482-y

David M (1988) Handbook of applied advanced geostatistical ore reserve estimation. Elsevier, Amsterdam

De Boor C (1978) A practical guide to splines. Springer-Verlag, Berlin

Dimitrakopoulos R, and Luo X (2004) Generalized sequential Gaussian simulation on group size v and screen-effect approximations for large field simulations. 36:(5):567–590. doi:10.1023/B:MATG.0000037737.11615.df

Dimitrakopoulos R, Mustapha H, and Gloaguen E (2010) High-order statistics of spatial random fields: exploring spatial cumulants for modeling complex non-Gaussian and non-linear phenomena. 42:(1):65–99. doi:10.1007/s11004-009-9258-9

Emery X (2009) Change-of-support models and computer programs for direct block-support simulation. 35:(10):2047–2056. doi:https://doi.org/10.1016/j.cageo.2008.12.010

Godoy M (2003) The effective management of geological risk in long-term production scheduling of open pit mines. University of Queensland, Brisbane, Qld, Australia

Goodfellow R, Albor Consuegra F, Dimitrakopoulos R, and Lloyd T (2012) Quantifying multi-element and volumetric uncertainty, Coleman McCreedy deposit, Ontario, Canada. Computers and Geosciences. doi:10.1016/j.cageo.2012.02.018

Goovaerts P (1997) Geostatistics for natural resources evaluation. Oxford University Press, New York

Guardiano FB, and Srivastava RM (1993) Multivariate geostatistics: beyond bivariate moments. In: Soares A (ed) Geostatistics Tróia '92: Volume 1. Springer Netherlands, Dordrecht, pp 133–144

Isaaks EH, and Srivastava RM (1989) Applied Geostatistics. Oxford University Press

Journel AG (2018) Roadblocks to the evaluation of ore reserves - the simulation overpass and putting more geology into numerical models of deposits. In: Dimitrakopoulos R (ed) Advances in Applied Strategic Mine Planning. Springer, Heidelberg, pp 47–55

Journel AG (1994) Modeling uncertainty: some conceptual thoughts. In: Dimitrakopoulos R (ed) Geostatistics for the Next Century. Springer Netherlands, Dordrecht, pp 30–43

Journel AG, and Alabert F (1989) Non-gaussian data expansion in the earth sciences. 1:(2):123–134. doi:10.1111/j.1365-3121.1989.tb00344.x

Journel AG, and Huijbregts CJ (1978) Mining Geostatistics. Blackburn Press, New York

Lebedev NN (1965) Special functions and their applications. Prentice-Hall Inc., New York

Li X, Mariethoz G, Lu DT, and Linde N (2016) Patch-based iterative conditional geostatistical simulation using graph cuts. 52:(8):6297–6320. doi:10.1002/2015WR018378

Mariethoz G, and Caers J (2014) Multiple-point Geostatistics: Stochastic Modeling with Training Images. Hoboken: Wiley

Mariethoz G, Renard P, and Straubhaar J (2010) The direct sampling method to perform multiple-point geostatistical simulations. 46:(11):1–14. doi:10.1029/2008WR007621

Minniakhmetov I, and Dimitrakopoulos R (2017a) Joint high-order simulation of spatially correlated variables using high-order spatial statistics. 49:(1):39–66. doi:10.1007/s11004-016-9662-x

Minniakhmetov I, and Dimitrakopoulos R (2017b) A high-order, data-driven framework for joint simulation of categorical variables. In: Gómez-Hernández JJ, Rodrigo-Ilarri J, Rodrigo-Clavero ME, et al. (eds) Geostatistics Valencia 2016. Springer International Publishing, Cham, pp 287–301

Minniakhmetov I, Dimitrakopoulos R, and Godoy M (2018) High-order spatial simulation using Legendre-like orthogonal splines. Mathematical Geosciences. doi:10.1007/s11004-018-9741-2

Mustapha H, Chatterjee S, and Dimitrakopoulos R (2014) CDFSIM: efficient stochastic simulation through decomposition of cumulative distribution functions of transformed spatial patterns. 46:(1):95–123. doi:10.1007/s11004-013-9490-1

Mustapha H, and Dimitrakopoulos R (2010a) High-order stochastic simulation of complex spatially distributed natural phenomena. 42:(5):457–485. doi:10.1007/s11004-010-9291-8

Mustapha H, and Dimitrakopoulos R (2010b) A new approach for geological pattern recognition using high-order spatial cumulants. 36:(3):313–334. doi:10.1016/j.cageo.2009.04.015

Mustapha H, and Dimitrakopoulos R (2011) HOSIM: A high-order stochastic simulation algorithm for generating three-dimensional complex geological patterns. 37:(9):1242–1253. doi:10.1016/j.cageo.2010.09.007

Osterholt V, and Dimitrakopoulos R (2007) Simulation of wireframes and geometric features with multiple-point techniques: application at Yandi iron ore deposit. The Australasian Institute of Mining and Metallurgy (AusIMM) Spectrum Series, vol 14, pp 51 - 60

Remy N, Boucher A, and Wu J (2009) Applied geostatistics with SGeMS: a user's guide. Cambridge University Press, Cambridge

Strebelle S (2002) Conditional simulation of complex geological structures using multiple-point statistics. 34:(1):1–21. doi:10.1023/A:1014009426274

Stuart A, and Ord JK (1987) Kendall's advanced theory of statistics

Wei Y, Wang G, and Yang P (2013) Legendre-like orthogonal basis for spline space. 45:(2):85–92. doi:10.1016/j.cad.2012.07.011

Yao L, Dimitrakopoulos R, and Gamache M (2018) A new computational model of high-order stochastic simulation based on spatial legendre moments. Mathematical Geosciences. doi:10.1007/s11004-018-9744-z

Zhang T, Gelman A, and Laronga R (2017) Structure- and texture-based fullbore image reconstruction. 49:(2):195–215. doi:10.1007/s11004-016-9649-7

Zhang T, Switzer P, and Journel A (2006) Filter-based classification of training image patterns for spatial simulation. 38:(1):63–80. doi:10.1007/s11004-005-9004-x