Integrating Multiple Point Statistics in Sequential Simulation

Julián M. Ortiz and Xavier Emery

Department of Mining Engineering,
University of Chile,

Abstract

Most conventional simulation techniques account for two-point statistics via the modeling of the variogram of the regionalized variable or of its indicators. These techniques cannot control the reproduction of multiple-point statistics or higher order features that may be critical for the performance of the models given the goal at hand (flow simulation in petroleum applications, planning and scheduling for mining applications).

Multiple-point simulation is a way to deal with this situation. It has been implemented for categorical variables, yet the demand of large data sets (training images) to infer the multiple-point statistics has impeded its use in the case of continuous variables. The main problem is that multiple-point statistics are characterized by discretizing the continuous variable by a set of thresholds and coding them as indicators, which results in a loss of resolution between thresholds.

We propose a method to incorporate multiple-point statistics into sequential simulation of continuous variables. Any sequential algorithm can be used. The method proceeds as follows. First, the multiple-point statistics are inferred from a training data set or training image with the typical indicator approach. The conditional probabilities given multiple-points data events enable to update the conditional distributions obtained by the sequential algorithm that uses the conventional two-point statistics. The key aspect is to preserve the shape of the conditional distribution between thresholds after updating the probability for the cutoffs used to infer the multiple-point statistics. Updating takes place under the assumption of conditional independence between the conditional probability obtained from the training set and the one retrieved from the conditional probability defined by the sequential method. The algorithm is presented for any sequential algorithm and then illustrated on a real data set using the sequential indicator and Gaussian simulation methods. The advantages and drawbacks of this proposal are pointed out.

Introduction

Geostatistical simulation is being used increasingly for uncertainty quantification. Traditionally, simulation methods only rely on the inference and modeling of a variogram that characterizes the spatial continuity of the variable of interest (see for example Goovaerts, 1997; Chilès and Delfiner, 1999). However, in most real applications the variogram cannot capture some important features of the true variable. The main reason for the poor performance of models built using conventional simulation tools is that variogram models only control the joint behavior of pairs of points and there is no explicit control on the joint behavior of multiple points. The algorithm used and its underlying assumptions dictates how relationships between multiple points are controlled. Two-point statistics such as the variogram or covariance are not enough to describe some complex features that the real phenomenon may present.

This problem was addressed for categorical variables by Guardiano and Srivastava (1993) in the early nineties. They introduced the idea of going beyond bivariate moments, through the use of *extended normal equations*. The method was based on using a training image for inferring the multiple-point indicator frequencies and then drawing an indicator value for the categorical variable given the probability of the unknown node to belong to each category (facies). This approach was improved by the implementation of Strebelle and Journel (2000) called *single normal equation*, where a search tree was used to find the multiple-point frequencies. Deutsch (1992) proposed the integration of multiple-point statistics in a *simulated annealing* framework. Caers and Journel (1998) used *neural networks* to infer the conditional distributions in a non-linear fashion considering multiple-point statistics. Both authors relied on the use of training images and their applications were oriented to categorical variables.

More recently, Ortiz and Deutsch (2004) suggested the use of multiple-point statistics extracted from production data (blast hole data in mining applications). These statistics are integrated into sequential indicator simulation. The probability of an unsampled location to belong to a class defined by two cutoffs can be approximated using the probability obtained by conventional indicator kriging or using the probability estimated from the training data for the same configuration of class grades in nearby informed locations. These two statistics can be combined under the assumption that they are conditionally independent (Journel, 2002).

In this article, we extend the approach proposed by Ortiz and Deustch (2004) to integrate multiple-point statistics in any sequential simulation algorithm. The key aspect is to infer the conditional distribution and then update it only at few thresholds, preserving its shape as much as possible. The proposed approach is implemented on a case study, where two benches of a copper mine are simulated using sequential Gaussian and indicator sequential simulation and then updated using multiple-point statistics.

Inferring multiple-point statistics

Inference of multiple-point statistics is a difficult problem and requires having abundant data over a large domain. Furthermore, these should be regularly spaced to allow repetition of patterns of several points. In practice, this problem has been solved using training images. Alternatively, in mining applications, the use of abundant and pseudo-regular production data from samples taken in blast holes can replace the training image (Ortiz, 2003; Ortiz and Deutsch, 2004).

Inference of the multiple-point statistics is done using the indicator coding. First, a number of thresholds are defined and a multiple-point pattern is used to scan the training image or training data set. For each threshold, the training data are coded as one if they belong to the corresponding class, that is, if the value is lower than or equal to the corresponding threshold value, and zero otherwise. From the scanning of the training image or data set, the probability of a node being less than or equal to the threshold can be calculated based on the experimental frequencies of that event.

Since there is no modeling of the experimental multiple-point frequencies, an important limitation of this procedure is that the training data set and modeling scale must be equivalent. That is, the spacing of the (pseudo-)regular data in the training set must be identical to the spacing between nodes that are simulated subsequently. A second problem of this data-driven approach is that mathematical inconsistencies between statistics inferred from the training data and from the sample data used to condition the simulation may exist.

Updating conditional distributions with multiple-point statistics

The proposed approach to update conditional distributions with multiple-point statistics consists in the following steps:

1. Define a random path to visit the nodes in the simulation grid.
2. At every visited node, determine the conditional distribution by simple kriging of the (coded) sample data and previously simulated nodes.
3. Discretize the conditional distribution by a set of thresholds, which are interpreted as the conditional probability of the variable at that location not to exceed the corresponding threshold value.
4. Update the conditional probabilities originated from discretizing the conditional distribution by assuming conditional independence between them and the probability of a node to exceed the corresponding threshold given the multiple-point configuration of (coded) original sample data and previously simulated nodes.
5. Fill in the discretized conditional distribution using some interpolation rule and, more importantly, extrapolation of the tails.
6. Draw a uniform random value in [0,1] to read from the conditional distribution a simulated value.
7. Proceed to the next node in the random path until all nodes have been simulated.

The updating technique described in step 4 was presented by Journel (2002) under the name of *permanence of ratios assumption*, but it is equivalent to the well-known conditional independence assumption used in the Naïve Bayes classifiers (Warner et al, 1961; Anderson, 1974; Friedman, 1997).

This methodology can be applied to any sequential simulation algorithm where the conditional distribution at the simulation nodes has been defined. The most straightforward approach would be to apply it in an indicator context (Ortiz and Deutsch, 2004). In the following sections we present the details of implementing this method using indicator, Gaussian, isofactorial and direct simulation.

Implementation

Let the event **A** be the probability of a node not to exceed a threshold. Event **B** is defined by the information provided by n single point events: $\{I(\mathbf{u}_1)=i_1, I(\mathbf{u}_2)=i_2,...,I(\mathbf{u}_n)=i_n\}$. Finally, event **C** is the multiple-point event defined by the values of the indicators of m points: $\{I(\mathbf{u}'_1)=i'_1, I(\mathbf{u}'_2)=i'_2,...,I(\mathbf{u}'_m)=i'_m\}$ (some of the n points belonging to **B** may also be part of **C**).

Indicator, Gaussian, isofactorial, or direct simulation can provide a conditional distribution that allows the calculation of $P(\mathbf{A} \mid \mathbf{B})$. The training dataset provides an estimate of $P(\mathbf{A} \mid \mathbf{C})$. Obtaining $P(\mathbf{A} \mid \mathbf{B}, \mathbf{C})$ requires knowing the relationship between **B** and **C**, which is generally extremely difficult to get. Some assumption is required. These probabilities are combined assuming they are conditionally independent given **A**, that is, considering the expression for $P(\mathbf{A} \mid \mathbf{B}, \mathbf{C})$ and $P(\overline{\mathbf{A}} \mid \mathbf{B},\mathbf{C}) = 1 - P(\mathbf{A} \mid \mathbf{B},\mathbf{C})$:

$$P(\mathbf{A} \mid \mathbf{B},\mathbf{C}) = \frac{P(\mathbf{A}) \cdot P(\mathbf{B} \mid \mathbf{A}) \cdot P(\mathbf{C} \mid \mathbf{A},\mathbf{B})}{P(\mathbf{B},\mathbf{C})} \quad P(\overline{\mathbf{A}} \mid \mathbf{B},\mathbf{C}) = \frac{P(\overline{\mathbf{A}}) \cdot P(\mathbf{B} \mid \overline{\mathbf{A}}) \cdot P(\mathbf{C} \mid \overline{\mathbf{A}},\mathbf{B})}{P(\mathbf{B},\mathbf{C})}$$

and the conditional independence conditions

$$P(\mathbf{B} \mid \mathbf{A},\mathbf{C}) = P(\mathbf{B} \mid \mathbf{A}) \text{ and } P(\mathbf{C} \mid \mathbf{A},\mathbf{B}) = P(\mathbf{C} \mid \mathbf{A})$$

the ratio between $P(\mathbf{A} \mid \mathbf{B}, \mathbf{C})$ and its complement $P(\overline{\mathbf{A}} \mid \mathbf{B},\mathbf{C})$ can be calculated as:

$$\frac{P(\mathbf{A} \mid \mathbf{B},\mathbf{C})}{P(\overline{\mathbf{A}} \mid \mathbf{B},\mathbf{C})} = \frac{P(\mathbf{A}) \cdot P(\mathbf{B} \mid \mathbf{A}) \cdot P(\mathbf{C} \mid \mathbf{A},\mathbf{B})}{P(\overline{\mathbf{A}}) \cdot P(\mathbf{B} \mid \overline{\mathbf{A}}) \cdot P(\mathbf{C} \mid \overline{\mathbf{A}},\mathbf{B})} = \frac{P(\mathbf{A}) \cdot P(\mathbf{B} \mid \mathbf{A}) \cdot P(\mathbf{C} \mid \mathbf{A})}{P(\overline{\mathbf{A}}) \cdot P(\mathbf{B} \mid \overline{\mathbf{A}}) \cdot P(\mathbf{C} \mid \overline{\mathbf{A}})}$$

This expression can be simplified to:

$$P(\mathbf{A} \mid \mathbf{B},\mathbf{C}) = \frac{\dfrac{1-P(\mathbf{A})}{P(\mathbf{A})}}{\dfrac{1-P(\mathbf{A})}{P(\mathbf{A})} + \dfrac{1-P(\mathbf{A} \mid \mathbf{B})}{P(\mathbf{A} \mid \mathbf{B})} \cdot \dfrac{1-P(\mathbf{A} \mid \mathbf{C})}{P(\mathbf{A} \mid \mathbf{C})}}$$

It can be seen that, under the assumption of conditional independence, the probability of event **A** can be calculated with relative ease, since it does not require knowing the relationship between **B** and **C**. We now present four cases where this approximation can be implemented.

INDICATOR SIMULATION

Consider the usual indicator coding:

$$i(\mathbf{u}_\alpha; z_k) = \begin{cases} 1, \text{ if } z(\mathbf{u}_\alpha) \le z_k \\ 0, \text{ otherwise} \end{cases} \quad k = 1,...,K$$

where $z(\mathbf{u}_\alpha)$ is the value at location $\mathbf{u}_\alpha$. This can be interpreted as a probability:

$$P(\mathbf{A}) = i(\mathbf{u}_\alpha; z_k) = \text{Prob}\{z(\mathbf{u}_\alpha) \le z_k\} = F(z_k)$$

For a simulated node located at $\mathbf{u}_0$, the conditional probability given the data in a search neighborhood can be calculated by simple indicator kriging (Journel, 1983; Alabert, 1987):

$$\begin{aligned} P(\mathbf{A} \mid \mathbf{B}) = \left[i(\mathbf{u}_0; z_k)\right]^*_{SIK} &= \left[\text{Prob}\{z(\mathbf{u}_0) \le z_k \mid (n)\}\right]^*_{SIK} \\ &= \sum_{\alpha=1}^{n} \lambda_\alpha^{SIK}(\mathbf{u}_0; z_k) \cdot i(\mathbf{u}_\alpha; z_k) + \left[1 - \sum_{\alpha=1}^{n} \lambda_\alpha^{SIK}(\mathbf{u}_0; z_k)\right] \cdot F(z_k) \end{aligned}$$

where (n) represents the conditioning information provided by n samples and previously simulated nodes in the search neighborhood, $\lambda_\alpha^{SIK}(\mathbf{u}_0; z_k)$ for $\alpha = 1,\ldots,n$ are the simple indicator kriging weights, and $F(z_k)$ is the global proportion below threshold z_k.

From the training information, a different conditional probability can be obtained for $z(\mathbf{u}_0)$ to be less than or equal to z_k, given the information of an m-point configuration:

$$P(\mathbf{A} \mid \mathbf{C}) = \left[\text{Prob}\{z(\mathbf{u}_0) \le z_k \mid (m)\}\right]^*_{MP}$$

The conditional probabilities obtained by simple indicator kriging can be updated with the conditional probabilities obtained from the training dataset, allowing the calculation of the conditional distribution that accounts for both, the set of *n* single point events and the single *m*-points event. Since the discretization of the conditional distributions by the indicator approach is generally coarse, the updated conditional distribution will also be a coarse approximation of the conditional distribution. The usual interpolation between the estimated indicators and extrapolation beyond the first and last thresholds is necessary (Deutsch and Journel, 1998).

Gaussian Simulation

A natural extension to the implementation presented above is to update conditional probabilities obtained via a multigaussian sequential simulation. The method requires the transformation of the original distribution into a standard Gaussian distribution:

$$Z(\mathbf{u}) = \phi(Y(\mathbf{u}))$$

It is widely known that under the assumption of multivariate gaussianity, the conditional distributions are fully defined by the mean and variance obtained by simple kriging:

$$m_Y(\mathbf{u}_0) = \sum_{\alpha=1}^{n} \lambda_\alpha^{SK}(\mathbf{u}_0) \cdot y(\mathbf{u}_\alpha) \qquad \sigma_Y^2(\mathbf{u}_0) = 1 - \sum_{\alpha=1}^{n} \lambda_\alpha^{SK}(\mathbf{u}_0) \cdot C(\mathbf{u}_0, \mathbf{u}_\alpha)$$

The expression for the conditional distribution is:

$$P(\mathbf{A} \mid \mathbf{B}) = \left[\text{Prob}\{z(\mathbf{u}_0) \le z_k \mid (n)\}\right]_{SK}^{*} = \left[\text{Prob}\{y(\mathbf{u}_0) \le y_k \mid (n)\}\right]_{SK}^{*} \sim N\left(m_Y(\mathbf{u}_0), \sigma_Y^2(\mathbf{u}_0)\right)$$

An indicator-based approach to update these conditional distributions would consist in discretizing them into a series of thresholds, which can be easily done numerically, and then updating these conditional probabilities with the conditional probabilities obtained from the training set. Again, a decision about how to interpolate between the discrete points and beyond the first and last thresholds is necessary; however, in this case, the thresholds can discretize the conditional distribution more precisely than in indicator simulation. For instance, instead of taking ten to fifteen thresholds, over a hundred thresholds can be easily taken, provided sufficient training information is available for reliable estimation of the conditional probabilities $P(\mathbf{A} \mid \mathbf{C})$.

Isofactorial Simulation

The next case of interest corresponds to sequential isofactorial simulation (Emery, 2002). Again, this method relies on a transformation of the original distribution into a new variable that follows an isofactorial distribution with marginal pdf *f*(.). Notice that the transformation function may differ from the one used in the case of sequential Gaussian simulation and that *f*(.) is not necessarily the standard Gaussian pdf. Typical applications consider transforming the raw variable to a Gaussian or Gamma distributions, although other cases can be considered, such as a Beta, Poisson, Binomial or Negative Binomial distribution. The conditional probability can be obtained by disjunctive kriging of the indicator function at a given threshold z_k:

$$P(\mathbf{A} \mid \mathbf{B}) = \left[i(\mathbf{u}_0; z_k)\right]_{DK}^{*} = \left[i(\mathbf{u}_0; y_k)\right]_{DK}^{*} = \sum_{p=0}^{\infty} \psi_p \cdot \left[\chi_p\left(Y(\mathbf{u}_0)\right)\right]_{SK}^{*}$$

where the coefficients are calculated as:

$$\psi_p = \int_{-\infty}^{y_k} \chi_p(y) \cdot f(y) \cdot dy$$

and each factor $\{\chi_p(Y(\mathbf{u}_0)), p \in \mathsf{N}\}$ is estimated by simple kriging from its values at the neighboring data locations:

$$\left[\chi_p(Y(\mathbf{u}_0))\right]^*_{SK} = \sum_{\alpha=1}^{n} \lambda^{SK}_{\alpha,p}(\mathbf{u}_0) \cdot \chi_p(Y(\mathbf{u}_\alpha))$$

The weights $\lambda^{SK}_{\alpha,p}(\mathbf{u}_0)$ are obtained by solving a simple kriging system considering a covariance function that depends on the isofactorial distribution and on the degree *p*. In practice only the first few factors are required (Matheron, 1976; Rivoirard, 1994; Chilès and Delfiner, 1999).

As with the conditional probabilities estimated by indicator kriging or under the multigaussian assumption, a probability conditional to the multiple-point event for each threshold can be estimated using the training data set, and subsequently used to update the conditional probability estimated under the isofactorial framework.

Direct Simulation

One last algorithm that could be considered is direct sequential simulation, which basically works by estimating the mean and variance of the conditional distribution by simple kriging. The shape of this distribution is then determined either by sampling the global distribution to match the mean and variance of the local conditional distribution (Soares, 2001), or by defining a conditional distribution lookup table (Oz et al., 2003). The procedure is virtually the same as in Gaussian simulation: obtain the conditional probability from the local distribution and update it with the multiple-point probability inferred from the training dataset.

Case Study

The following case study presents some preliminary results about the application of the proposed methodology to simulate the point-support grades on a copper deposit, based on drill hole (exploration) information. The multiple-point statistics are extracted from production (blast hole) data obtained from two benches already mined out. This information is used to simulate the copper grade on two lower benches. An assumption of strict stationarity is required in order to "export" these multiple-point statistics. The example shows the updating technique implemented for the sequential indicator and Gaussian simulation algorithms.

Figure 1 shows the exploration data for a specific bench and the training information from one of the two benches used for multiple-point statistics inference. These statistics are inferred using a 5 points pattern made of a central node and the four adjacent nodes in the horizontal plane (no vertical data has been used for the multiple-point statistics inference). *Figure 2* displays realizations for a specific bench using sequential indicator and Gaussian simulation and the proposed methods where the conditional distributions are updated with multiple-point statistics extracted from the production data. The typical "patchiness" of indicator simulation appears clearly in the maps. This characteristic appears more strongly when multiple-point information is incorporated under the assumption of conditional independence. The patchiness disappears when using the Gaussian algorithm as a base method for inferring the conditional distributions: in this

case, transitions from high to low grade zones are smoother. However, the integration of multiple-point statistics injects more connectivity to the realization.

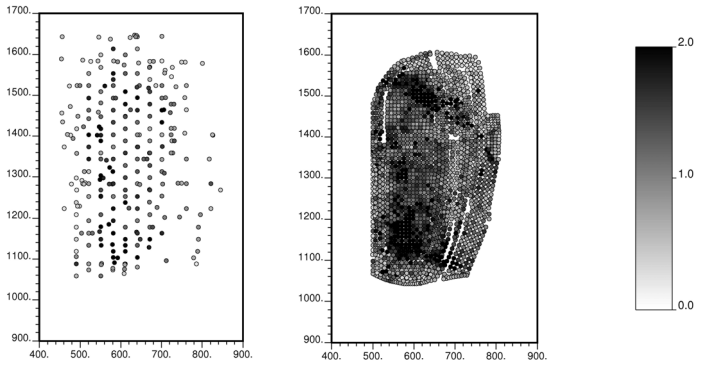

Figure 1. Left: Exploration (drill holes) data. Right: production (blast holes) data. Only the data in one bench are displayed. Production data are used to infer the multiple-point statistics.

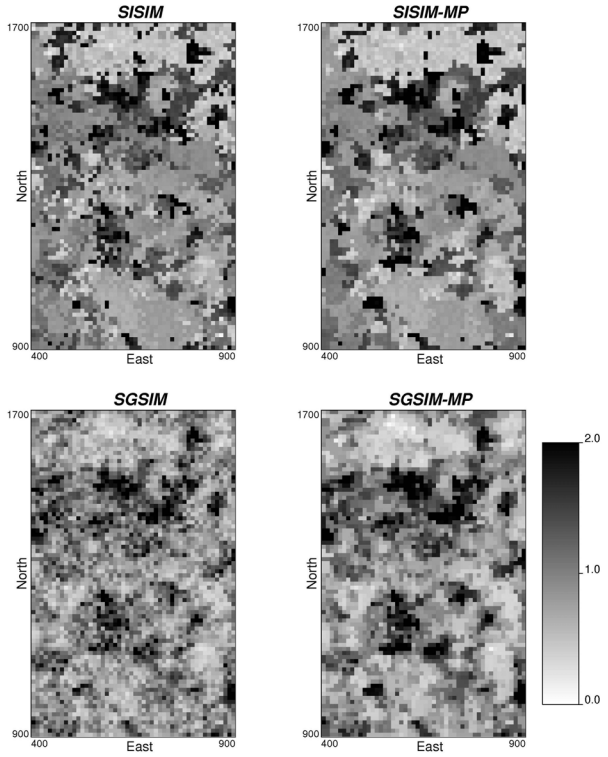

Figure 2. Plan views of the two benches simulated using sequential indicator simulation updated with multiple-point statistics.

Table 1 shows the total copper content and quantity of metal above a cutoff of 0.7 %Cu calculated over a particular area using the four methods for 20 × 20 × 12 m^3 panels. Smoother transitions and the added connectivity explain the higher variance obtained first, between

indicator and Gaussian methods, and second, between the cases without and with multiple-point information. Validation remains a difficult issue and further research is required in this respect.

The implementation of these algorithms has shown some of the possible problems of their application. Numerical approximations are required to interpolate and extrapolate the tails once the discretization in indicators is performed for the updating procedure. Furthermore, the number of thresholds used depends on the quality and size of the training data set, in order to ensure reliable estimation of the multiple-point statistics.

	Cutoff = 0 %Cu		Cutoff = 0.7 %Cu	
	Mean	Std. Dev.	Mean	Std. Dev.
SISIM	69.25	1.43	67.12	1.71
SISIM-MP	70.88	1.72	67.44	2.06
SGSIM	69.85	1.64	67.54	2.00
SGSIM-MP	74.27	1.95	71.30	2.32

Table 1. Total quantity of metal and quantity above a cutoff of 0.7 %Cu from the sets of 100 realizations obtained with each method (in thousands of copper tonnes).

Conclusions

Integrating multiple-point statistics into sequential simulation algorithms can be achieved under some assumption of the relationship (redundancy) between the conditional probability inferred from a training data set, given a multiple-point event, and the conditional probability inferred by a conventional kriging approach (indicator, multigaussian, or disjunctive kriging). We propose assuming conditional independence between these two sources of information, to obtain an estimate of the conditional probability that accounts for the neighboring data (n points) and the closest multiple point configuration (m points). The updating methodology proposed can be applied to any sequential simulation algorithm, provided that a conditional distribution is calculated at each simulation node on the grid.

Implementation of this technique has proven challenging, particularly because of possible inconsistencies between the sources of information (biases) where the statistics are inferred, and because of numerical approximations (particularly when extrapolating the tails) required to obtain the simulated values from the updated conditional distributions. Furthermore, the assumption itself should be investigated. A model that accounts for the redundancy between the sources of information could be easily constructed by defining a parameter τ, such that: $P(\mathbf{C} \mid \mathbf{A})^{\tau} \approx P(\mathbf{C} \mid \mathbf{A}, \mathbf{B})$, hence:

$$\frac{P(\mathbf{A} \mid \mathbf{B},\mathbf{C})}{P(\overline{\mathbf{A}} \mid \mathbf{B},\mathbf{C})} \Big/ \frac{P(\mathbf{A} \mid \mathbf{B})}{P(\overline{\mathbf{A}} \mid \mathbf{B})} = \left(\frac{P(\mathbf{A} \mid \mathbf{C})}{P(\overline{\mathbf{A}} \mid \mathbf{C})} \Big/ \frac{P(\mathbf{A})}{P(\overline{\mathbf{A}})} \right)^{\tau}$$

However, the parameter is difficult to get and, to make things worse, it is location and data dependent.

The proposal in this article opens an interesting and original research avenue about the use of multiple-point statistics in a data-driven mode. Implementation and applications to real data will offer challenges that have yet to be discovered.

Acknowledgements

We would like to acknowledge the Department of Mining Engineering at the University of Chile and particularly its industry sponsor Codelco Chile for supporting this research.

References

Alabert, F. G., *Stochastic Imaging of Spatial Distributions using Hard and Soft Information*, MSc Thesis, Stanford University, Stanford, CA, 1987.

Anderson, J. A., Diagnosis by logistic discriminant function: further practical problems and results. *Applied Statistics*, vol. 23, no. 3, 1974, p. 397-404.

Caers, J. and Journel, A. G., Stochastic reservoir simulation using neural networks trained on outcrop data, *in* 1998 SPE Annual Technical Conference and Exhibition, New Orleans, LA. Society of Petroleum Engineers. SPE paper # 49026, 1998, p. 321-336.

Chilès, J. P. and Delfiner, P., *Geostatistics: Modeling spatial uncertainty*, Wiley, New York, 1999.

Deutsch, C. V., *Annealing Techniques Applied to Reservoir Modeling and the Integration of Geological and Engineering (Well Test) Data*. PhD thesis, Stanford University, Stanford, CA, 1992.

Deutsch, C.V. and Journel, A.G., *GSLIB: Geostatistical Software Library and User's Guide*, 2nd Edition, Oxford University Press, 1998.

Emery, X., Conditional simulation of non-Gaussian random functions. *Mathematical Geology*, vol. 34, no. 1, 2002, p. 79-100.

Friedman, J. H., On bias, variance, 0/1-loss, and the curse of dimensionality. *Data mining and knowledge Discovery*, vol. 1, 1997, p. 55-77.

Goovaerts, P., *Geostatistics for Natural Resources Evaluation*, Oxford University Press, 1997.

Guardiano, F. and Srivastava, M., Multivariate geostatistics: Beyond bivariate moments, *in* A. Soares, editor, Geostatistics Tróia'92. Kluwer, Dordrecht, vol. 1, 1993, p. 133-144.

Journel, A. G., Nonparametric estimation of spatial distributions, *Mathematical Geology*, vol. 15, no. 3, 1983, p. 445-468.

Journel, A. G., Combining knowledge from diverse sources: An alternative to traditional data independence hypotheses. *Mathematical Geology*, vol. 34, no. 5, 2002, p. 573-596.

Matheron, G., A Simple Substitute for Conditional Expectation: the Disjunctive Kriging, *in* M. Guarascio, M. David and C. Huijbregts, editors, Advanced Geostatistics in the Mining Industry. Reidel, Dordrecht, 1976, p. 221-236.

Ortiz, J. M., *Characterization of High Order Correlation for Enhanced Indicator Simulation*. PhD thesis, University of Alberta, Edmonton, AB, Canada, 2003.

Ortiz, J. M. and Deutsch, C. V., Indicator simulation accounting for multiple-point statistics. *Mathematical Geology*, vol. 36, no. 6, 2004, p. 545-565.

Oz, B., Deutsch, C. V., Tran, T. T. and Xie, Y., DSSIM-HR: A Fortran 90 program for direct sequential simulation with histogram reproduction. *Computers & Geosciences*, vol. 29, no. 1, 2003, p. 39-52.

Rivoirard, J., *Introduction to Disjunctive Kriging and Non-Linear Geostatistics*, Oxford University Press, New York, 1994.

Soares, A., Direct sequential simulation and cosimulation. *Mathematical Geology*, vol. 33, no. 8, 2001, p. 911-926.

Strebelle, S. and Journel, A. G., Sequential simulation drawing structures from training images, *in* 6th International Geostatistics Congress, Cape Town, South Africa, Geostatistical Association of Southern Africa, 2000.

Warner, H., Toronto, H., Veesey, L. and Stephenson, R., A mathematical approach to medical diagnosis. *Journal of the American Medial Association*, vol. 177, 1961, p. 177-183.