

Joint simulation of cross-correlated ore grades and geological domains: an application to mineral resource modeling

Nasser MADANI (✉)¹, Mohammad MALEKI²

¹ School of Mining and Geosciences, Nazarbayev University, Astana 010000, Kazakhstan

² Department of Metallurgical and Mining Engineering, Universidad Católica del Norte, Antofagasta 1270709, Chile

© Higher Education Press 2022

Abstract Spatial modeling of ore grades is frequently impacted by the local variation in geological domains such as lithological characteristics, rock types, and geological formations. Disregarding this information may lead to biased results in the final ore grade block model, subsequently impacting the downstream processes in a mining chain project. In the current practice of ore body evaluation, which is known as stochastic cascade/hierarchical geostatistical modeling, the geological domain is first characterized, and then, within the geological model, the ore grades of interest are evaluated. This practice may be unrealistic in the case when the variability in ore grade across the boundary is gradual, following a smooth transition. To reproduce such characteristics, the cross dependence that exists between the ore grade and geological formations is considered in the conventional joint simulation between continuous and categorical variables. However, when using this approach, only one ore variable is considered, and its relationship with other ore grades that may be available at the sample location is ignored. In this study, an alternative approach to jointly model two cross-correlated ore grades and one categorical variable (i.e., geological domains) with soft contact relationships that exist among the geological domains is proposed. The statistical and geostatistical tools are provided for variogram inference, Gibbs sampling, and conditional cosimulation. The algorithm is also tested by applying it to a Cu deposit, where the geological formations are managed by the local and spatial distribution of two cross-correlated ore grades, Cu and Au, throughout the deposit. The results show that the proposed algorithm outperforms other geostatistical techniques in terms of global and local reproduction of statistical parameters.

Keywords geostatistical simulation, categorical variable, continuous variable, geological domain, variogram inference

1 Introduction

In different areas of geoscience and for the purpose of geostatistical modeling, more than one variable is usually encountered, and each variable has a different statistical nature. For instance, in the mining industry, one often encounters different ore grades and geological domains, such as rock types, geological formations, alterations and mineralizations (Journel and Huijbregts, 1978; Rossi and Deutsch, 2014). In the petroleum industry, the most important variables are permeability, porosity and facies (Caers, 2005; Pyrcz and Deutsch, 2014). In the hydrogeology industry, different variables, such as connectivity and geological domains, are of paramount importance. These variables can be continuous (such as ore grades) or categorical (such as geological domains) (Kitanidis, 1997). Most importantly, a high level of dependence is usually observed not only between continuous variables but also between continuous and categorical variables. For instance, the local and global distributions of the ore grade may depend on the local and spatial variability of the geological domains. These characteristics encourage the use of this dependence during the geostatistical modeling of geological domains and ore grades. Many researchers have discussed stochastic modeling of ore grades and mineral grades, taking the incorporation of geological domains into account (Talebi et al., 2014, 2015, 2016, 2019a, 2019b; Mery et al., 2017; Emery and Maleki, 2019; Madani et al., 2019a, 2019b; Maleki and Emery, 2020; Masoumi et al., 2020). One of the most common and straightforward approaches to achieve this is the hierarchical or cascade approach, in which the area

of each geological domain is first delimited, and then, the ore grades are independently modeled within every single domain, taking only those data within a particular domain into account (Alibert and Massonnat 1990; Haldorsen and Damsleth, 1990; Dubrule, 1993; Boucher and Dimitrakopoulos 2012; Roldão et al., 2012; Jones et al., 2013; Maleki et al., 2020). Despite the fact that the hierarchical method is straightforward, there are some considerable drawbacks. For example, the spatial dependence is not incorporated between the grades crossing the rock-type borders, and as a result, abrupt transitions (i.e., discontinuities) in the ore grade values occur when crossing a boundary, which is referred to as a hard boundary (Duke and Hanna, 2001; Glacken and Snowden, 2001; Wilde and Deutsch, 2012; Rossi and Deutsch, 2014).

To consider the spatial dependence of ore grades across boundaries in a hierarchical approach, one can model the cross-correlation function between ore grade values in each geological domain (Larrondo et al., 2004; Mery et al., 2017). Although the spatial dependence can be reproduced in the outcome models in this case, there are no smooth transitions across boundaries, which is referred to as a soft boundary, when the ore grade values are divided into different categories (as many categories as there are geological domains). Therefore, using this hierarchical technique is questionable in cases when there are gradual ore grade changes near the boundaries and one is interested in reproducing this feature in the outcome model (Maleki and Emery, 2020).

To address this problem, an alternative approach is to jointly and simultaneously model the continuous variable (such as the ore grade) and categorical variable (such as geological formations) to exploit the spatial and local cross correlation of continuous variables through the layout of geological formations. However, simultaneously modeling these variables, which have different natures, is a time-consuming process; this is why few studies have addressed this issue. For instance, Bahar and Kelkar (2000) proposed a method where the categorical variable was obtained through the truncation of a Gaussian random field, and the continuous variable was obtained by transforming and modeling an independent Gaussian random field. Dowd (1994, 1997) and Freulon et al. (1990) considered another approach, in which the categorical variable was obtained through the truncation of a Gaussian random field, and then, the continuous variable was obtained through the transformation of a Gaussian random field. Nevertheless, in these proposed methods, only one Gaussian random field was used for modeling categorical variables, which yielded a spatially ordered sequence of units identical to those in the truncated Gaussian simulation. For a more complex geological context in a contact relationship, these approaches are impractical. To solve this issue, Emery and Silva (2009), Cáceres and Emery (2010), and Maleki

and Emery (2015) used a plurigaussian simulation to truncate two independent Gaussian random fields for modeling categorical variables and back transformation of one Gaussian random field for modeling one continuous variable to model more complicated geological domains.

However, the process of joint modeling can become even more complicated when the number of continuous variables is more than one and there is also a multivariate relationship between several continuous variables and a categorical variable.

In this study, an approach is presented to jointly model two continuous variables and one categorical variable. For this purpose, an algorithm is presented that uses an iterative process for variogram analysis between three variables (a generalization of the concepts proposed in Emery and Silva (2009) and Emery and Cornejo (2010)), followed by a Gibbs sampler algorithm, which is used to convert the categorical data at the sample location to Gaussian values taking into account the cross correlation that exists among them with two other continuous variables. This algorithm then uses a cosimulation approach to model the three Gaussian values at the target location. The first two simulated values can be back transformed to their original scale, and the third value can be truncated into the categorical variable. This concept is illustrated through a case study, where the copper grade (Cu) and gold grade (Au) are correlated and their local and global distributions are controlled by two geological formations: tourmaline breccia and other rocks. In addition, the results of contact analysis reveal that all the boundaries between geological formations are soft, which is the main reason to use a joint modeling approach to model these three variables (i.e., the Cu grade, Au grade and geological formations), and the aim is to reproduce the soft contact between geological formations (Rossi and Deutsch, 2014; Emery and Maleki, 2019). The following section presents the main concepts of the joint modeling approach. Then, the capability of the proposed approach is assessed through a Cu deposit case study. Finally, the results obtained are compared with those using a factorization approach and hierarchical cosimulation, and a discussion of the results follows.

2 Methodology

The joint simulation approach, which was first proposed by Emery and Silva (2009), combines multi-Gaussian simulation and truncated Gaussian or plurigaussian simulation algorithms for generating realizations of one continuous variable and one categorical variable. In the current study, the generalization of this type of joint modeling is addressed using additional continuous variables, while the number of categorical variables remains one.

Let $Z(u) = [Z_1(u), Z_2(u), \dots, Z_m(u)]$ denote m continuous variables and $I(u) = [I_1(u), I_2(u), \dots, I_n(u)]$ denote the indicators of the categorical variables, where n is the number of categories. The indicator $I_n(u)$ is 1 if category n is seen at observation location u and 0 otherwise. It is worth noting that the categories can have either forbidden or allowed contacts.

The proposed workflow in this study includes the joint modeling of two continuous variables and one categorical variable. The former variables are obtained by a multi-Gaussian simulation algorithm, and the latter variable is derived from a truncated or plurigaussian simulation algorithm.

The joint stochastic modeling of $Z(u)$ and $I(u)$ is performed through the following main steps.

2.1 Transformation of the original data

In the first step, the original continuous variables $Z(u) = [Z_1(u), Z_2(u), \dots, Z_m(u)]$ are independently transformed to normal scores $Y(u) = [Y_1(u), Y_2(u), \dots, Y_m(u)]$ with a mean of zero and a variance of one. This can be achieved by using either a Gaussian anamorphosis function (Rivoirard, 1994) or a quantile-quantile approach (Deutsch and Journel, 1998):

$$Y_i(u) = G^{-1}(F_i(Z_i(u))), \quad i = 1, 2, \dots, m, \quad (1)$$

where $G^{-1}(\cdot)$ is the standard normal cumulative distribution function, $F_i(\cdot)$ is the cumulative distribution function of the original variable $Z_i(u)$, and $Y_i(u)$ is the Gaussian random field. This transformation ensures that the univariate data are normal, but the bivariate Gaussianity should be checked (Emery, 2005). The categorical variables are first converted into indicator data $I(u) = [I_1(u), I_2(u), \dots, I_n(u)]$ and then into the Gaussian random field $Y_0(u)$. The latter transformation is described in the following sections.

2.2 Joint spatial structure modeling

The proposed algorithm in this study involves the joint modeling of several standard Gaussian random fields $Y_1(u), \dots, Y_m(u)$ that can be used for generating realizations of continuous variables (such as the ore grade) and a Gaussian matrix random field $Y_0(u)$ that is used for generating realizations of categorical variables (such as geological formations) through a conventional truncated or plurigaussian simulation approach (Armstrong et al., 2011). Since a cosimulation of these Gaussian random fields is applied, the joint spatial structure among them should be identified.

In this respect, the idea is to derive the joint spatial model from experimental direct and cross variograms between $Y_1(u), \dots, Y_m(u)$ and indicator $I(u)$. An alternative approach is to utilize the Monte Carlo simulation to

derive an optimum fit of the underlying linear model of coregionalization. In this approach, an initial model of coregionalization between several Gaussian random fields, including the range, sill and nugget effect, is first proposed. Next, by using the proposed model, a large number of nonconditional realizations (e.g., 1 million) are simulated for each specified lag distance $\Delta(h)$, at which the values of the experimental variograms are available. This can be achieved in different directions of anisotropies if they exist; otherwise, this can be achieved over the omni-directional experimental variograms. The maximum number of lags depends on the area of consideration and the sampling pattern itself.

In this case, one obtains many realizations of $Y^S(h) = [Y_0^S(h), Y_1^S(h), Y_2^S(h), \dots, Y_m^S(h)]$ through an immediate truncation of $Y_0^S(h)$ based on the identified thresholds, and the continuous simulated values can be converted to the indicators $I^S(h)$. Then, the direct and cross variograms between $Y_1^S(h), Y_2^S(h), \dots, Y_m^S(h)$ and $I^S(h)$ are calculated for each realization and averaged over the lag distance $\Delta(h)$. The compatibility of these average values (i.e., $\bar{Y}_1^S(h), \bar{Y}_2^S(h), \dots, \bar{Y}_m^S(h)$ and $\bar{I}^S(h)$) should be intuitively examined with an experimental variogram obtained from $Y_1(u), Y_2(u), \dots, Y_m(u)$ and $I_n(u)$ at each lag distance $\Delta(h)$. If the model is convincing, the proposed linear model of coregionalization can be accepted; otherwise, another model should be iterated to reach a satisfactory theoretical model of coregionalization.

2.3 Conversion of the categorical data into a Gaussian random field

As previously mentioned, the geological formations can be obtained through the truncation of several simulated Gaussian random fields at target locations. Therefore, the indicator data should be converted into the Gaussian matrix random field $Y_0(u)$ at the observation locations, taking the variogram and cross variograms obtained in the previous step into account. For this purpose, an acceptance-rejection method such as the Gibbs sampler (Geman and Geman, 1984; Casella and George, 1992) can be used.

In the Gibbs sampler algorithm, the values at the sample locations are updated several times to obtain convergence to the target distribution, which is a standard Gaussian distribution. In the traditional paradigm of this algorithm, the simulated values at these locations are based on simple kriging estimation and the simple kriging variance that can be obtained by solving a unique neighborhood kriging system. However, in the joint modeling algorithm proposed in this study, simple kriging needs to be replaced by simple cokriging (Emery and Silva, 2009; Emery and Cornejo, 2010) so that the spatial dependence that exists between the Gaussian vector random fields is considered.

In our proposed approach, the Gibbs sampler generates random numbers through a joint multivariate distribution identified from two continuous variables and one variable that is conditioned on inequality constraints at sample locations u (Emery, 2007; Emery et al., 2014):

$$L(u) < Y_0(u) < U(u), \tag{2}$$

where the upper limit $U(u)$ and lower limit $L(u)$ are real numbers or infinite.

The steps for the Gibbs sampler with respect to several Gaussian random fields can be implemented as follows.

I. Initiation: generate random numbers at each sample location by using an acceptance-rejection method (Emery et al., 2014).

II. Iterating for $iter = 1, 2, \dots, L$.

a) Randomly or regularly select an index $i \in \{1, \dots, n\}$ with n sample locations comprising the indicator data.

b) Estimate the condition $Y_0(u_i)$ to $Y_1(u), \dots, Y_m(u)$ and a previously simulated Gaussian random field $Y_0(u)$, which is available at other observation locations, by building a simple cokriging system in a unique neighborhood according to the direct variograms and cross variogram functions obtained in the previous section. Return the cokriging estimated value $Y(u_i)$ and error variance $\sigma_{SCK}^2(u_i)$:

$$Y_{0SCK}(u_i) = \sum_{i \neq k} \Phi_{i,k}^{Y_{1,2,\dots,m}} Y_{1,2,\dots,m}(u_k) + \sum_{i \neq k} \Phi_{i,k}^{Y_0} Y_0(u_k), \tag{3}$$

$$\sigma_{SCK}^2(u_i) = \sigma_{Y_0}^2 - \sum_{i \neq k} \Phi_{i,k}^{Y_{1,2,\dots,m}} C_{Y_{1,2,\dots,m}}(u_{ki}) - \sum_{i \neq k} \Phi_{i,k}^{Y_0} C_{Y_0}(u_{ki}), \tag{4}$$

where $Y_{1,2,\dots,m} = Y_1(u), Y_2(u), \dots, Y_m(u)$ is a vector of Gaussian random fields at data locations corresponding to continuous variables; Y_0 is a previously simulated Gaussian matrix random field corresponding to the categorical variables; $\Phi_{i,k}^{Y_{1,2,\dots,m}}$ is a vector of weights that $Y_1(u), Y_1(u), \dots, Y_m(u)$ Gaussian random vectors assign to estimate Y_{0SCK} ; $\Phi_{i,k}^{Y_0}$ are weights that previously simulated values of Y_0 that are assigned to estimate Y_{0SCK} ; $\sigma_{Y_0}^2$ is the variance of Gaussian random field Y_0 ; C_{Y_0} is the direct covariance of Y_0 ; $C_{Y_{1,2,\dots,m}}$ is the cross covariance between Y_1, Y_2, \dots, Y_m .

c) Simulate a new random vector $Y_0(u_i)$ by utilizing a conditional distribution that possesses a mean and variance equivalent to its simple cokriging estimation and variance of estimation, respectively:

$$Y_0(u_i) = Y_{0SCK}(u_i) + \sqrt{\sigma_{SCK}(u_i)}M, \tag{5}$$

where M is an independent random value drawn from a standard Gaussian random distribution.

d) If $L(u_i) < Y(u_i) < U(u_i)$, update the Gaussian random field by

$(Y^{iter-1}(u_1), \dots, Y^{iter-1}(u_{i-1}), Y^{iter}(u_i), Y^{iter-1}(u_{i+1}), \dots, Y^{iter-1}(u_n))$. Otherwise, set $Y^{iter}(u_i) = Y^{iter-1}(u_i)$.

e) Return to step a) and loop L times.

This iterative algorithm is aperiodic and irreducible, its distribution is invariant and it is similar to the traditional Gibbs sampler algorithm, in which only one variable is used. The chain used in this algorithm guarantees that the target simulated values converge in distribution to a joint multivariate Gaussian random field of $Y_0, Y_1, Y_2, \dots, Y_2(N(0, 1))$, conditioned on the available Gaussian random fields while respecting the restriction imposed by Eq. (2). The number of iterations should be as large as possible to obtain an acceptable convergence (Emery and Cornejo, 2010). The truncation threshold in this algorithm is compatible with the rules applicable for the conventional truncated Gaussian simulation (Armstrong et al., 2011).

It is worth noting that the algorithm proposed above is comprehensive for a case that contains several continuous variables and one categorical variable. This means that in every step of updating the Gibbs process, two or three Gaussian variables (i.e., those variables related to categorical variables) are updated, and their relationship with several other continuous normal score variables is taken into account. The Gaussian random fields and their related thresholds are identified based on the contact relationship among the geological domains. However, it should be mentioned that having numerous continuous variables makes the variogram analysis quite complicated, and up to a point, the inference of a linear model of coregionalization becomes somewhat unfeasible.

The Gibbs sampler presented above is slightly different from the Gibbs sampler proposed by Emery and Silva (2009) and Emery and Cornejo (2010). For the previously proposed Gibbs sampler, two or three Gaussian variables (i.e., those variables related to categorical variables) can be updated by taking only one continuous variable into account.

2.4 Conditional cosimulation

The obtained Gaussian matrix random field $Y_0(u_i)$ from the previously discussed Gibbs sampler holds the joint spatial structure with several other Gaussian random fields $Y_1(u_i), Y_2(u_i), \dots, Y_m(u_i)$ that are also available at sample locations due to embedding the simple cokriging system into the process of Gibbs sampling, where a fitted linear model of coregionalization is used. These cross-correlated variables can then be used as conditioning data for jointly simulating $Y_0, Y_1, Y_2, \dots, Y_m$ at target locations u_0 . There are several Gaussian simulation algorithms that can be used to achieve this. The commonly used methods are the turning bands cosimulation (Emery and Lantuéjoul, 2006) and sequential Gaussian cosimulation (Goovaerts, 1997) methods.

In this study, we use the turning bands cosimulation due to its straightforwardness, simplicity, and its ability to

better reproduce the original statistical variability of the conditioning data compared to the sequential Gaussian cosimulation (Paravarzar et al., 2015). After the simulation, the simulated values of Y_1, Y_2, \dots, Y_m are transformed back to the primary distribution of continuous variables, and simulated values of Y_0 are truncated into the categorical variable through the preidentified truncation thresholds at the target blocks, which refer to the conceptual flag identified based on geological interpretations.

2.5 Workflow

To better understand the algorithm proposed in this study for practical usage in the ore body evaluation process when there are two ore grades as continuous variables and one categorical variable as geological formations available at sample locations (i.e., drill holes) to build a typical geological block model, the previous steps are summarized.

1) Identify the continuous variables (e.g., two ore grades) and categorical variables (e.g., geological formations).

2) Identify the relationship between categories to check the allowed and forbidden contacts between geological formations.

3) Investigate the behavior of continuous variables across the boundary of categories. The proposed workflow requires that both continuous variables show soft boundaries across the category domains. This can be examined by a contact analysis tool.

4) Identify the multivariate relationship between the continuous variables and investigate whether the bivariate

relationship after normal score transformation by Gaussian anamorphosis is linear and follows the multi-Gaussianity assumption.

5) Convert the categorical variable to indicators.

6) Calculate the experimental indicator variograms of the categories and the experimental variograms of two normal score transformed continuous variables.

7) Derive the linear model of coregionalization (LMC) between three Gaussian random fields (Y_1, Y_2, Y_0) based on the discussion provided in Subsection 2.2. The first two fields introduce the Gaussian values of the two continuous variables, and the last field is the Gaussian matrix random field with respect to the categorical variable.

8) Convert the categorical variable into a Gaussian value at the sample location, taking the obtained linear model of coregionalization from the previous step by the Gibbs sampler discussed in Subsection 2.3 into account.

9) Cosimulate the three Gaussian random fields conditioned to the Gaussian values available at sample locations by using the turning bands cosimulation.

10) Back transform the first two simulated Gaussian values to the original scale of the two continuous variables and truncate the other simulated Gaussian values to obtain the corresponding categories at the target locations.

11) Postprocess the realizations to investigate the reproduction of multivariate relationships between continuous variables and check the reproduction of their soft boundaries across the layout of categories.

The workflow for the case when two ore grades as continuous variables and one categorical variable as geological formations are introduced is shown in Fig. 1.

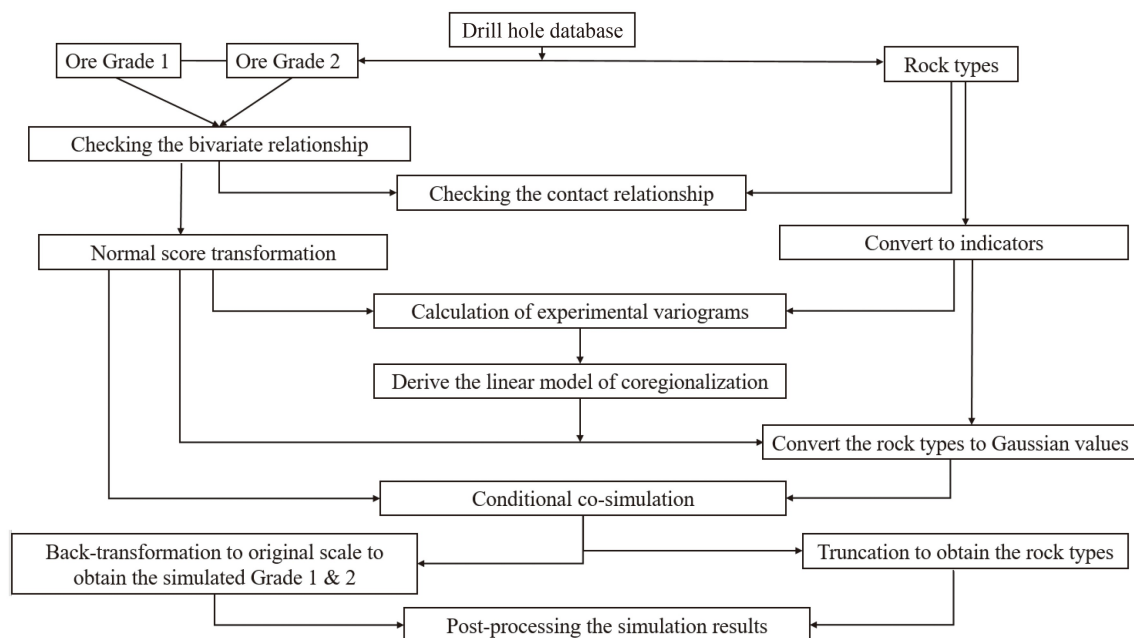


Fig. 1 Flowchart of the proposed joint simulation algorithm for modeling two ore grades and one geological formation.

3 Case study

In this section, the capability of the proposed methodology is assessed through a real case study, where two continuous variables and one categorical variable are jointly simulated through the proposed approach.

3.1 Presentation of the data

The data set, which originates from a drilling campaign of a copper mine (Fig. 2), contains 2376 samples, for which the drill holes are distributed in a volume of approximately 31200000 m³ with the availability of information through geological formations and two continuous variables, the Cu grade $Z_1(u)$ and the Au grade $Z_2(u)$. The name of this deposit cannot be disclosed due to confidentiality reasons. As seen from Fig. 3 and Table 1, there is a strong correlation between the ore grades of Cu and Au, for which their local distributions can be mainly regulated by two geological formations. Therefore, an engaging relation can be deduced among the Cu and Au ore grades and the geological formations. The geological formations that regulate the local and spatial distributions of the Cu and Au ore grades can be recognized as follows.

Tourmaline breccia: this rock unit comprises a substantial amount of the Cu ore grades that are mostly situated across the middle of the area under consideration.

Other rocks: in fact, the tourmaline breccia rock unit in the center of the area under consideration is confined by other rocks in the corner. These geological formations contain various rock types, such as diorite and granodiorite, which have lower amounts of Cu and Au ore grades.

The sample locations of the data are illustrated in Fig. 3. The major concentrations of both Cu and Au ores in tourmaline breccia are expected to be mainly distributed in the center of the deposit and not in the corners.

Prior to any statistical analysis, the data set must be declustered to remove the bias that would affect further

geostatistical analysis. This step is necessary in the case when the sampling pattern is preferential (Rossi and Deutsch, 2014). As seen from Fig. 3, the irregularity in sample points is tangible throughout the area, particularly in the corners and the center of the deposit, where there are few and many observation locations available, respectively. Therefore, the available data were declustered by the cell declustering technique to make the statistical analysis representative (Goovaerts, 1997; Deutsch and Journel, 1998). Table 1 shows declustered statistical parameters of the variables. As seen from this table, 68% of the sample data belong to the tourmaline breccia rock unit, which is the dominant geological formation in the region, and it is expected that the outcome models of the geological formations will be able to reproduce this proportion.

The distribution of both Cu and Au ore grades is controlled by the geological formations (i.e., high concentrations of Cu and Au ores can be observed in the tourmaline breccia rock unit, and lower concentrations can be observed in the other rock types (Fig. 3 and Table 1)). In addition, there is a strong relationship between the ore grades of Au and Cu whether geological formations are considered (Table 1). These characteristics justify joint modeling between the Au and Cu ore grades and geological formations. Prior to modeling, a contact analysis was implemented to specify the best suited technique for considering and integrating the role of geological formations in the modeling process of the Cu and Au ore grades. This analysis was used to determine the behavior of variables near the boundary between two geological domains from the ore grade transition perspective. For this purpose, there are two types of contact analysis: the mean grade analysis and cross-correlation contact analysis (Maleki and Emery, 2015). In the former type of analysis, the arrangement of data belonging to one geological formation is addressed, and its distance from the border with another geological formation is considered. Usually, only the pairs of datapoints in which the tail is in one geological formation

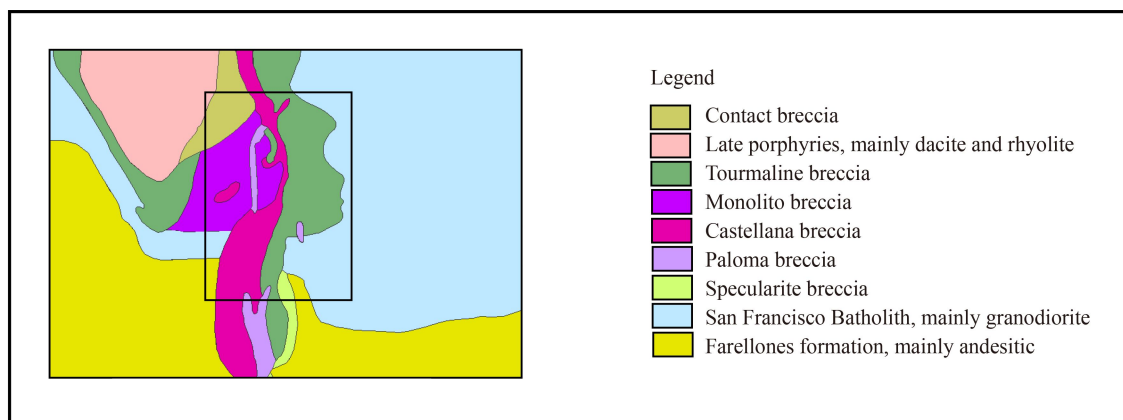


Fig. 2 Regional geological map; the approximate area of the copper mine is indicated by a black rectangle (the coordinates are local).

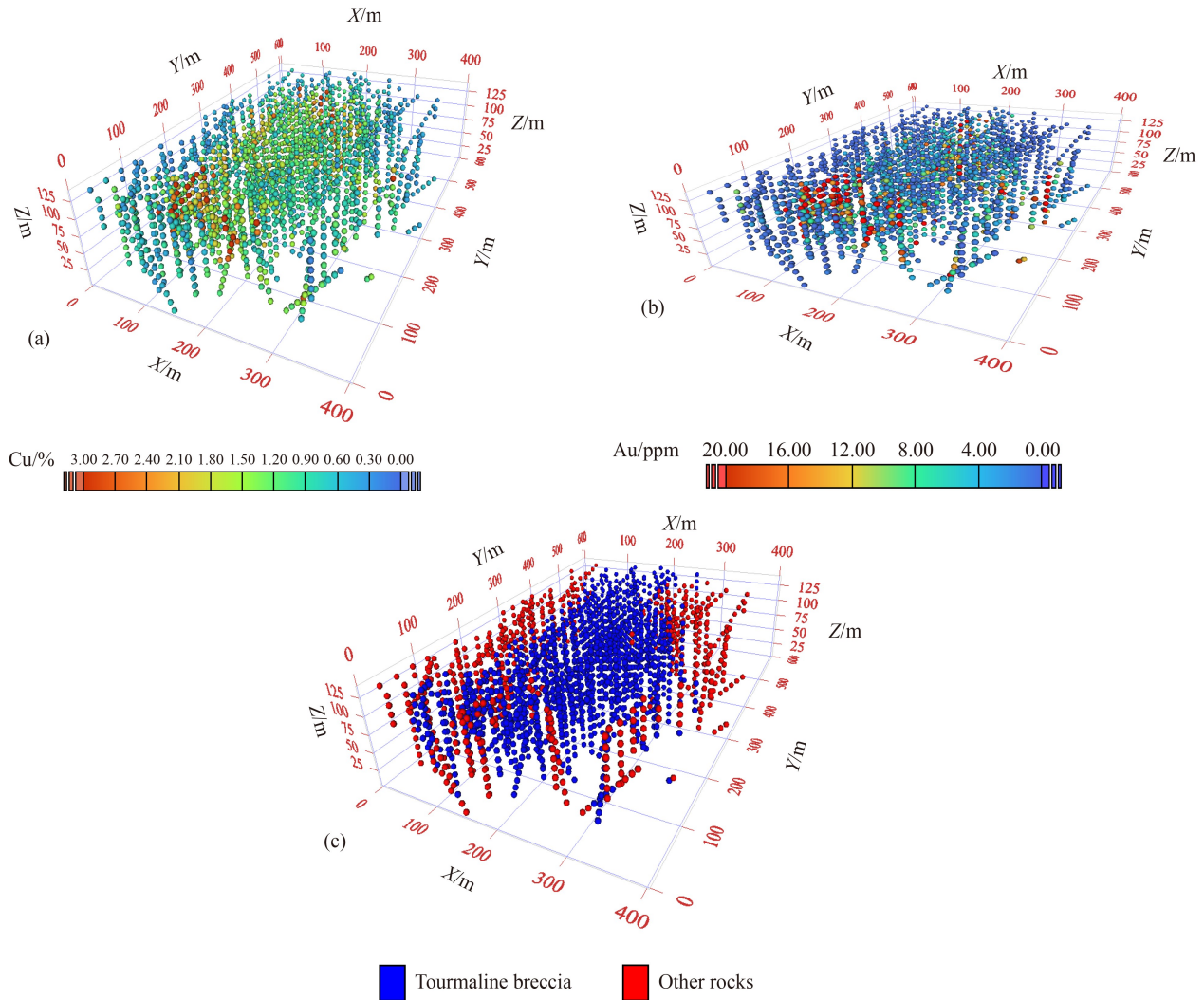


Fig. 3 A 3D location map of the sample data for Cu ore grades (a), Au ore grades (b), and the geological formations (c).

Table 1 Main declustered statistical parameters of variables

Parameter	Global		Tourmaline breccia		Other rocks	
	Cu/%	Au/ppm	Cu/%	Au/ppm	Cu/%	Au/ppm
Number of data	2376	2376	1635	1635	741	741
Mean	1.05	3.95	1.19	4.87	0.74	1.9
Maximum	7.24	69.85	7.24	68.83	3.71	69.85
Minimum	0.12	0	0.16	0	0.12	0
Variance	0.41	52.11	0.437	58.04	0.22	33.02
Correlation	0.71		0.71		0.65	

and the head is in another geological formation is considered. The mean grade contact analysis can then be calculated by plotting the mean grade of each rock group as a function of the distance to the border between two geological formations (Rossi and Deutsch, 2014). In the latter, the cross correlation between data pairs is calculated and plotted as a function of distance (lag separation).

As seen in Figs. 4(a) and 4(c), by approaching the boundary, gradual changes can be observed in the ore grade values for both variables. Moreover, from Figs. 4(b) and 4(d), it can be deduced that there is a spatial cross correlation between ore grades on two sides of the boundaries. As a result, it can be concluded that there is a soft boundary of Au and Cu between geological formations. In addition to the strong global cross

dependence that exists among the Au and Cu ore grades and geological formations that is evident from Table 1, the presence of soft boundaries (Fig. 4) also supports using a joint simulation approach since it is capable of reproducing such boundaries and spatial dependencies between ore grades on two sides of the geological boundary.

3.2 Joint modeling of spatial continuity

To determine the spatial joint structure, Cu and Au ore grades were transformed into normal scores by taking the Gaussian anamorphosis function into account. An examination is necessary to determine the multi-Gaussianity assumption and bivariate normality through the normal score values. The reason is that most multi-Gaussian simulation approaches, such as the turning bands cosimulation that we recommended in our proposed algorithm, entail these assumptions to be corroborated within the transformed variables. To do so, the spatial multi-Gaussianity of transformed data was assessed by investigating the experimental variogram of different

orders (ω) through normal score variables. This is based on the comparison between the conventional variogram (i.e., order 2) and the variograms of lower order, the madogram and rodogram (i.e., orders 1 and 2, respectively) (Emery, 2005). The results of this validation verified that spatial multi-Gaussianity can be corroborated since the points in these plots are disseminated along the theoretical lines (Figs. 5(a) and 5(b)). Moreover, the scatterplot between normal score variables verifies the presence of marginal bivariate normality since the cloud of points mimics the common elliptical shape of multivariate normal distributions (Fig. 5(c)). Therefore, the normal scores of the original data can be used for any multi-Gaussian cosimulation algorithm.

Next, the categorical data were converted to normal score values. As previously mentioned, this step is followed by the Gibbs sampler algorithm. In this regard, the covariates Cu and Au are also taken into account in the Gibbs sampler by using the simple cokriging system. This, however, requires inference of a linear model of coregionalization that can be obtained by a Monte Carlo simulation technique following an iterative process to

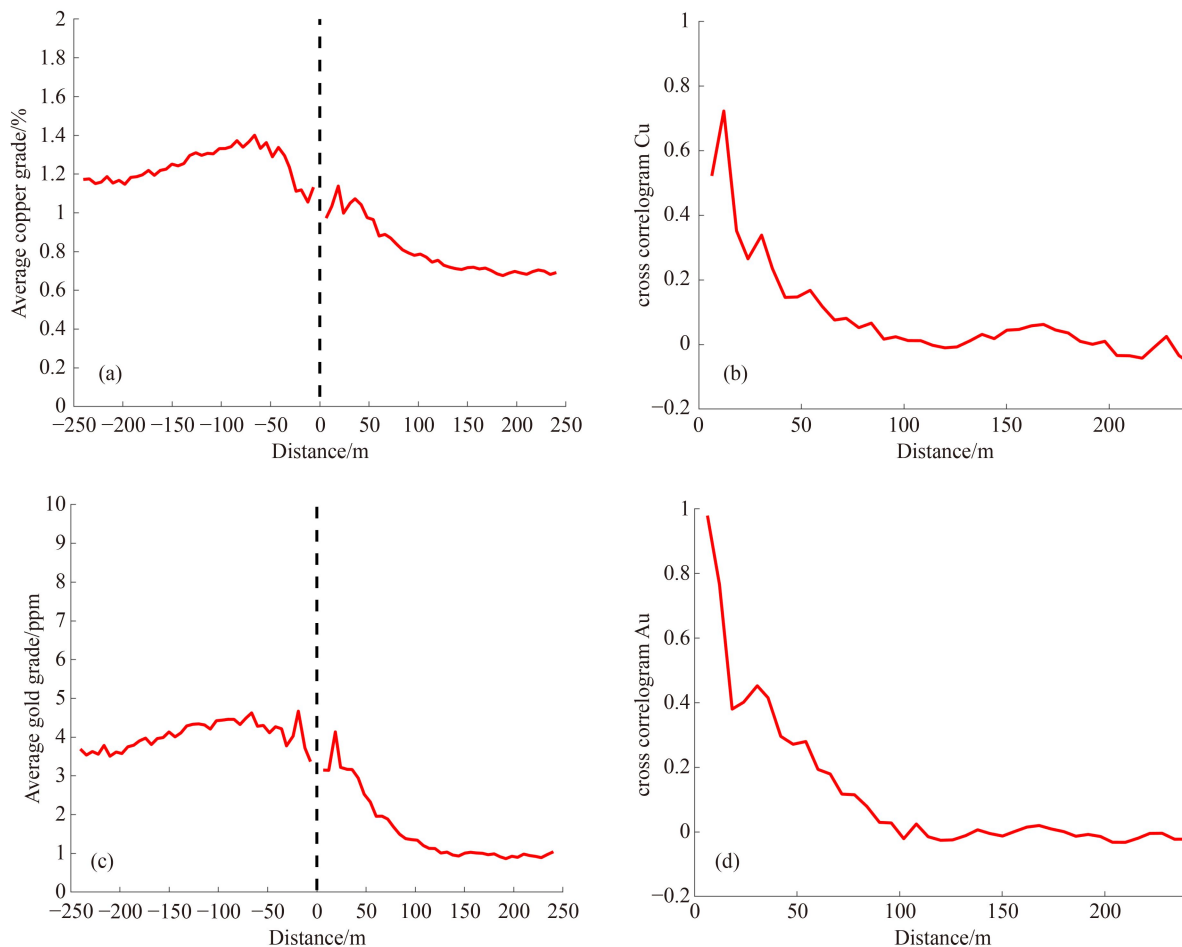


Fig. 4 Contact analysis between geological formations. Left: mean grade as a function of the distance to the geological formation boundary for (a) Cu and (c) Au; right: cross correlation between ore grades measured in different geological formations as a function of the distance between sampling data for (b) Cu and (d) Au.

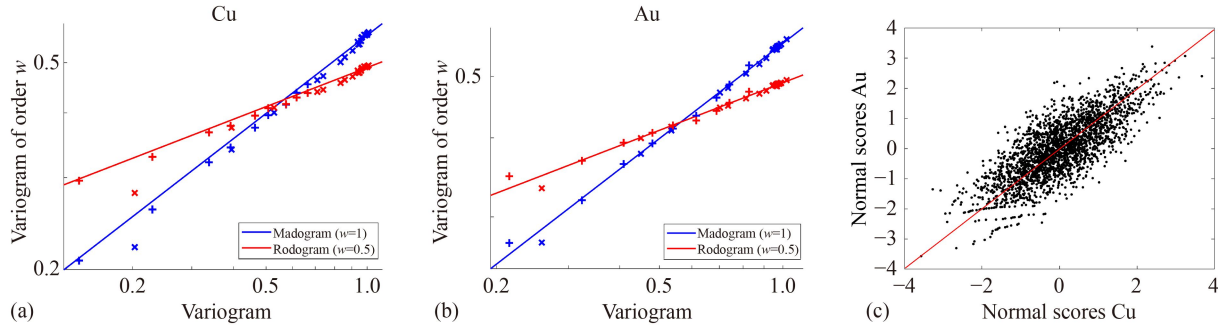


Fig. 5 Experimental madogram and rodogram of the normal score variables as a function of their experimental variograms for (a) Cu and (b) Au. In the case of multi-Gaussianity, the points are disseminated along the solid line. (c) Scatterplot of the normal scores of Cu and Au.

derive the ideal ranges, matrices of sills and nugget effects. Since there are only two geological formations in this case study, deducing a flag is quite straightforward. As a result, one Gaussian random field introduces two categories: tourmaline breccia and other rocks, which can be obtained by a declustered truncation threshold as $G^{-1}(P) = -0.47$, where P is the declustered proportion of tourmaline breccia (0.32). Therefore, in the first step of variogram analysis, the experimental variograms of a Gaussian vector random field (i.e., the normal score values of Cu and Au) and one indicator along two directions of anisotropy were computed.

These directions were obtained from geological interpretation and calculation of the variograms in various directions with small tolerances in lag and angles. The results show that the variogram represents isotropy in the horizontal plane and anisotropy in the vertical direction, with higher ranges in the former and lower ranges in the latter. Subsequently, a linear model of coregionalization is proposed for the three Gaussian random fields (Y_1, Y_2, Y_0). As seen in Eq. (6), some parameters of the proposed model are approximately known and are associated with direct and cross variogram models between Gaussian random fields of Cu and Au (Y_1, Y_2). Therefore, an iterative process was used to determine the whole linear model of coregionalization (Coefficients a, b, c, d, e, f):

$$\begin{aligned}
 & \begin{pmatrix} Y_{11} & Y_{12} & Y_{10} \\ Y_{21} & Y_{22} & Y_{20} \\ Y_{01} & Y_{02} & Y_{00} \end{pmatrix} = \begin{pmatrix} 0.175 & 0.128 & a \\ 0.128 & 0.247 & b \\ a & b & c \end{pmatrix} \text{nugget} \\
 & + \begin{pmatrix} 0.577 & 0.498 & d \\ 0.498 & 0.605 & b \\ d & b & f \end{pmatrix} \text{Sph}(140 \text{ m}, 140 \text{ m}, 140 \text{ m}) \\
 & + \begin{pmatrix} 0.2449 & 0.13 & g \\ 0.13 & 0.139 & e \\ g & e & i \end{pmatrix} \text{Sph}(250 \text{ m}, 250 \text{ m}, 9999 \text{ m}).
 \end{aligned} \tag{6}$$

One million realizations for the three Gaussian random fields in each lag separation were generated. The third Gaussian random field was truncated according to the

truncation rule to obtain the indicator variogram model. Then, the obtained model was compared with experimental variograms between the Cu and Au ore grades and the geological formations indicator. This process was repeated until a satisfactory model was obtained.

The final model was derived comprising two nested structures of spherical models and a nugget effect by the following coefficients (Fig. 6): $a = 0.015, b = 0.017, c, d = -0.15, e = -0.2, f = 0.1, g = -0.43, h = -0.32, i = 0.9$.

3.3 Joint modeling of continuous and categorical variables

Following the inference of the linear model of coregionalization that was obtained in the previous section, the Gibbs sampler algorithm was used to convert the categorical data into normal scores at the sample locations. To achieve this objective, 100 realizations were drawn from a multivariate distribution among the Cu and Au ore grades and a Gaussian random field subject to the inequality constraint as in Eq. (2) by imposing the truncation threshold. A large number of iterations was also considered to ensure that the simulation results converged to the target multivariate distribution $N(0, 1)$. However, inclusion of the two underlying variables that are correlated with the underlying Gaussian random field through establishing a simple cokriging system led to the acceleration of the rate of convergence.

The three Gaussian random fields at the sample locations were conditionally cosimulated in the study area by the turning bands cosimulation algorithm.

At each sample location, the normal scores of the Au and Cu ore grades and 100 realizations of the normal scores obtained from the Gibbs sampler were obtained, which can be used to identify the corresponding geological formations by applying the truncation threshold. Since all three Gaussian random fields are cross correlated, the turning bands cosimulation algorithm was employed for the simulation of these variables at the target locations. The simulation grid had a block dimension of $5 \text{ m} \times 5 \text{ m} \times 12 \text{ m}$, and a moving neighborhood containing up to 60 conditioning data was defined

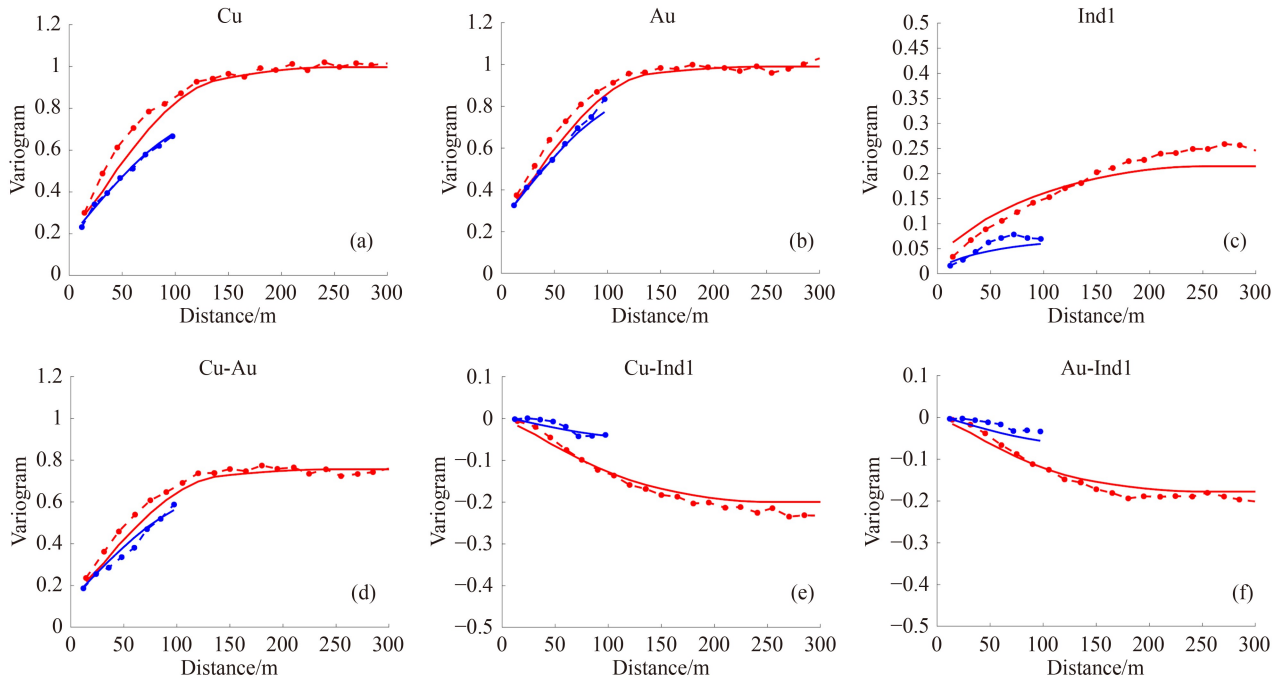


Fig. 6 Linear model of coregionalization along the main direction of anisotropy (red and blue lines) for direct variograms (a, b, and c) and cross variograms (d, e, and f) between normal scores of the Cu and Au ore grades and the indicator data (dashed lines).

for establishing the simple cokriging system with the linear model of coregionalization that was previously obtained. The number of turning lines was set to 1000 to prevent the striping effect in the simulation results. Finally, the realizations of continuous (the Cu and Au ore grades) and categorical variables (geological formations) were obtained by back transformation of Y_1, Y_2 and truncation of simulated values Y_0 , respectively.

Figure 7 shows the results of the simulation for one of the realizations. The reproduced maps are compatible with the location maps presented and discussed in Fig. 3. For instance, tourmaline breccia appears mostly in the center of the deposit, where the locations with high concentrations of Cu and Au ore grades are reproduced well and are compatible with the geological interpretation of the deposit.

In addition, by having one hundred generated realizations, the probability of occurrence of two geological formations accompanying E-type maps of the Cu and Au ore grades were calculated in the study area (Fig. 8). In fact, the E-type maps suffer from a smoothing effect, which is a typical characteristic of these types of maps. These maps are mostly applicable for visualization and should never be used in further analysis of a mine evaluation; instead, one should work with all realizations together.

In addition, the results of contact analysis calculated for three different generated realizations of the Cu and Au ore grades (which were selected randomly) are presented in Fig. 9 (i.e., the blue, green, and purple lines) over the simulated grid. As seen from this figure, through all

selected realizations, one can observe a gradual change in the ore grade by crossing the boundary between geological formations. Moreover, in all selected realizations, there is a correlation between ore grade values located on two sides of the boundary, which is in satisfactory agreement with the conditions observed in the original data. Therefore, it can be deduced that the proposed approach is capable of reproducing the soft boundaries and spatial correlation between ore grade values on two sides of the boundary.

3.4 Validation of the results

The proposed methodology can also be evaluated against actual data. A split-sample method, known as jackknife (Deutsch and Journel, 1998), is chosen for this purpose. The original data set consisting of 2376 observation locations was split into two main groups. One set (30% of data) was considered the test set, and another nonoverlapping set (70% of data) was selected for analysis. The analysis data were then used to cosimulate the Au and Cu ore grades and geological formations at the observation locations of the test data through the proposed approach. Since the true values of the Cu and Au ore grades are available at test observation locations, a comparison can be made with the average of simulated results conditioned to the analysis data set. For this purpose, a scatterplot is generated between the true and predicted (average of simulation results) values at test observation locations for both the Cu and Au ore grades (Figs. 10(a) and 10(b)). The tendency of the conditional mean to the identity

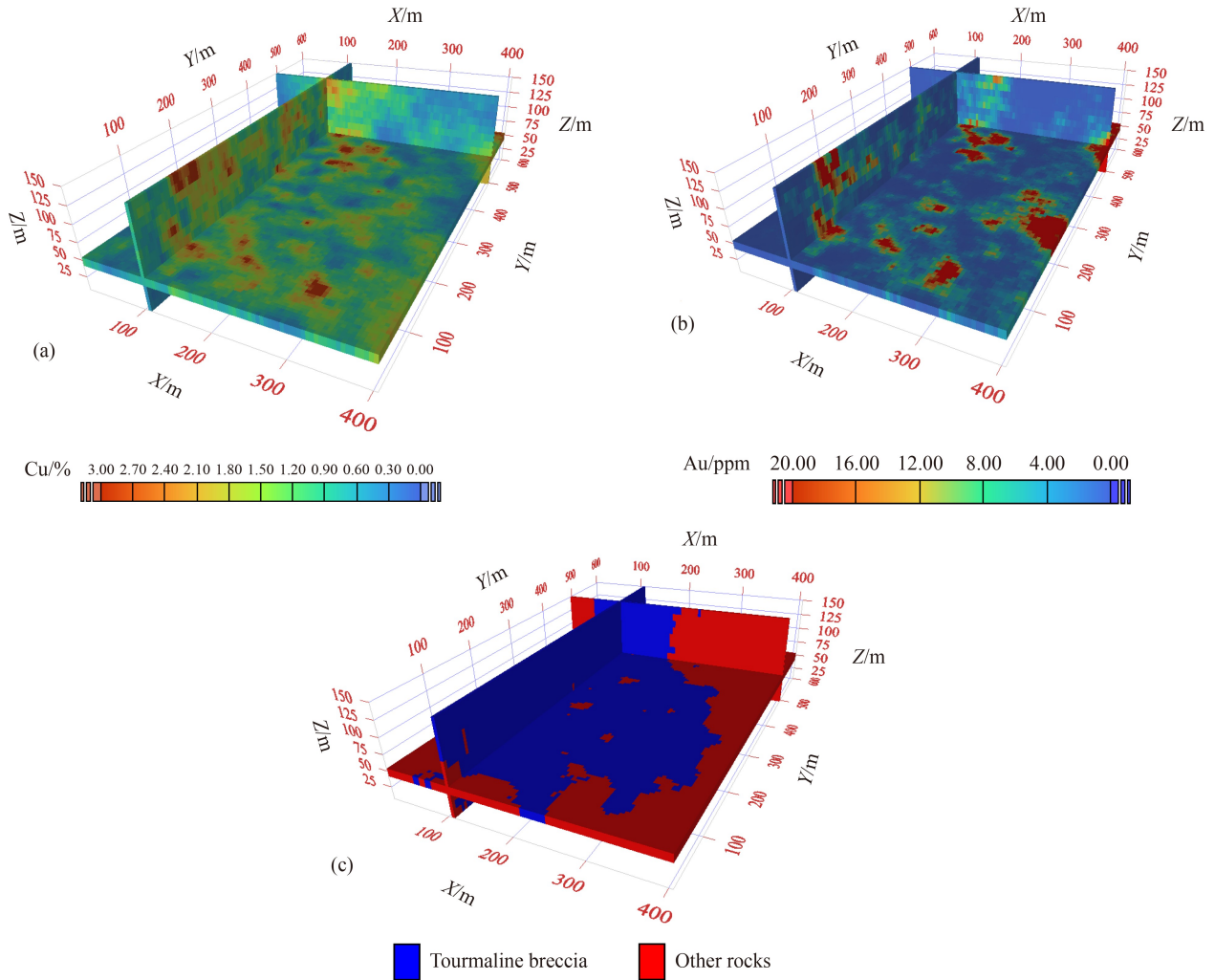


Fig. 7 One realization of the (a) Cu ore grade (b) Au ore grade and (c) geological formations. This realization is randomly chosen to show the variability of simulation in the region.

shows that the average of simulated results is conditionally unbiased and precise. Another criterion is also to check the symmetric probability intervals (SPIs) at each test observation location known as the accuracy plot (Deutsch, 1996). The bounds of the probability intervals (quantiles (p)) were computed over the simulated values; then, the probability p against the proportion of test observation locations containing the p -SPI was calculated and drawn by a scatterplot. The plots indicate that a proportion p of the probability intervals contains the true Cu and Au ore grades, and there is an acceptable agreement within the observed and expected proportions based on p due to slight departure from the identity line (Figs. 10(c) and 10(d)).

Thus far, some jackknife validation results have been discussed that are suitable for investigating the suitability of the method proposed for the reproducibility of the ore grades for Cu and Au. However, it might also be of interest to examine the ability of the proposed joint simulation approach to regenerate geological domains (i.e., geological formations) versus actual data. A

confusion matrix (Stehman, 1997) was calculated. This matrix, which is also known as the error matrix, is a specific table that allows validation of the performance of any predictor algorithm. This approach is a widely used technique to describe the performance of machine learning methods (e.g., classification algorithms). In fact, it is a contingency table with two dimensions (i.e., “actual” versus “predicted”) that allows the identification of the correct classification and misclassification of the categories. In this study, we also established this matrix to explore how many times the proposed algorithm misclassified the actual geological formations at test sample locations to assess the accuracy of the proposed joint simulation approach. The results obtained for jackknife method are once again used for this purpose. As previously described, by taking the analysis data set into account, the ore grades and geological formations are jointly simulated at test locations. The predicted ore grades versus true ore grades are thoroughly discussed in Fig. 9.

Since true geological formations are also available at

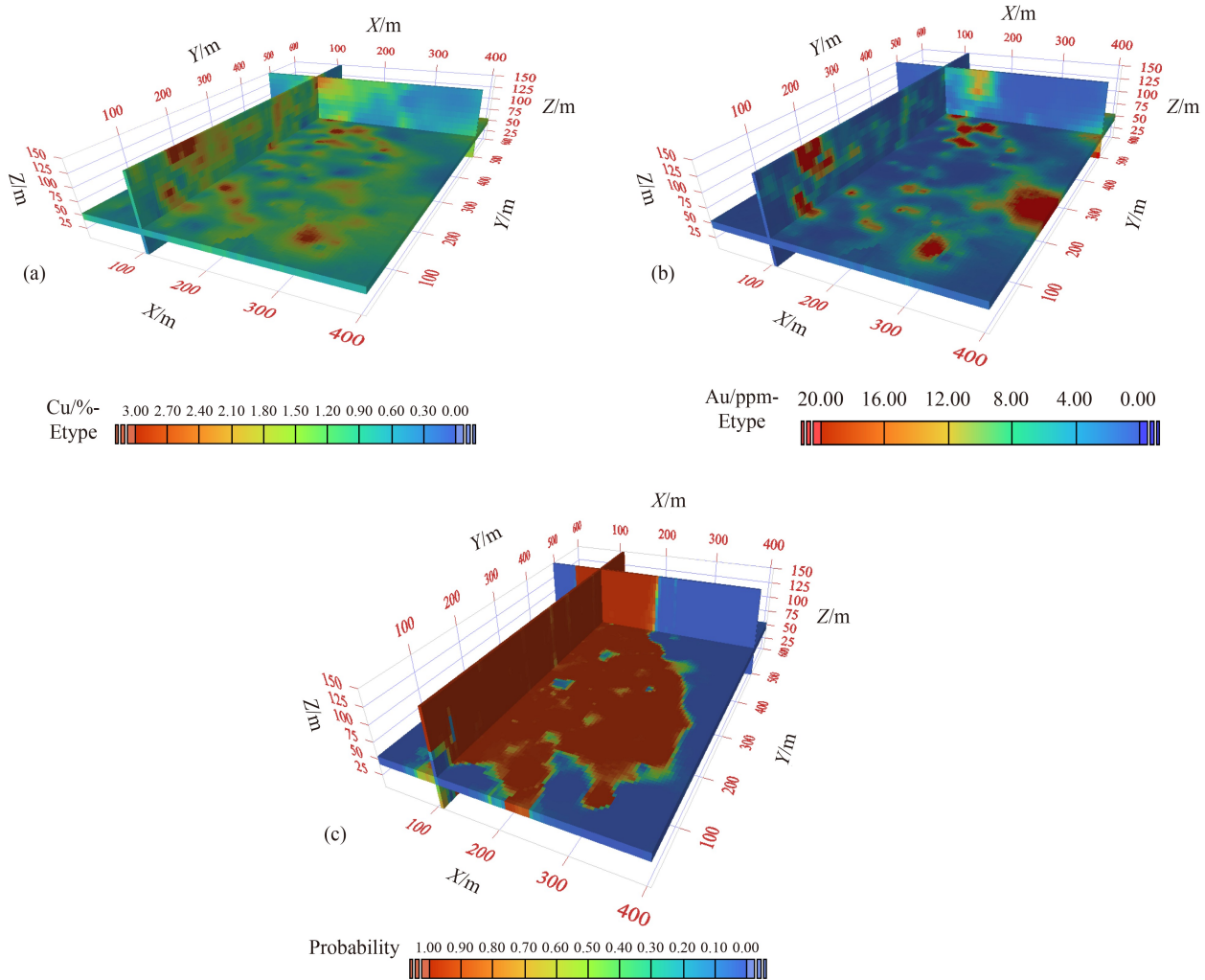


Fig. 8 E-type maps of (a) Cu ore grade (b) Au ore grade, and the probability map of (c) tourmaline breccia obtained from one hundred realizations.

test locations, one needs to derive the predicted geological formations at test locations to establish the confusion matrix. The most likely geological formations are obtained at test locations through 100 simulated values of geological formations obtained by the jackknife method. First, the probability of the tourmaline breccia and other rock geological formations were calculated at each test location. Then, a category was assigned based on the geological formations that produced the most likely geological formation at that location. For instance, if the probability of tourmaline breccia and other rocks at location “A” is 0.75 and 0.25, respectively, then tourmaline breccia is assigned to location “A” as the most likely geological formation. These are the predicted values in the confusion matrix. As seen in Table 2, out of 1664 test locations (overall), there are 460 sample locations where the true geological formation is tourmaline Breccia and the predicted geological formation is also tourmaline breccia; there are 1048 locations where the true geological formation is other rocks and the predicted geological formation is also other rocks; there

are 62 sample locations where the true geological formation is tourmaline breccia and the predicted geological formation is other rocks; there are 94 locations where the true geological formation is other rocks and the predicted geological formation is tourmaline breccia. The list of rates that can be calculated from the confusion matrix are as follows.

1) Accuracy: overall, the accuracy identifies how often the proposed algorithm is correct in predicting the geological formations: 90.63%;

2) Misclassification: overall, misclassification how often the proposed algorithm is incorrect in predicting geological formations: 9.37%;

3) Precision: how often the proposed algorithm is correct when predicting tourmaline breccia: 83.03%;

4) Prevalence: how often tourmaline breccia actually is predicted in the proposed methodology: 33.29%, which is close to the original declustered proportion of tourmaline breccia in the original data set (i.e., ~32%).

As can be observed from the obtained rates (high accuracy and low misclassification rate), the proposed

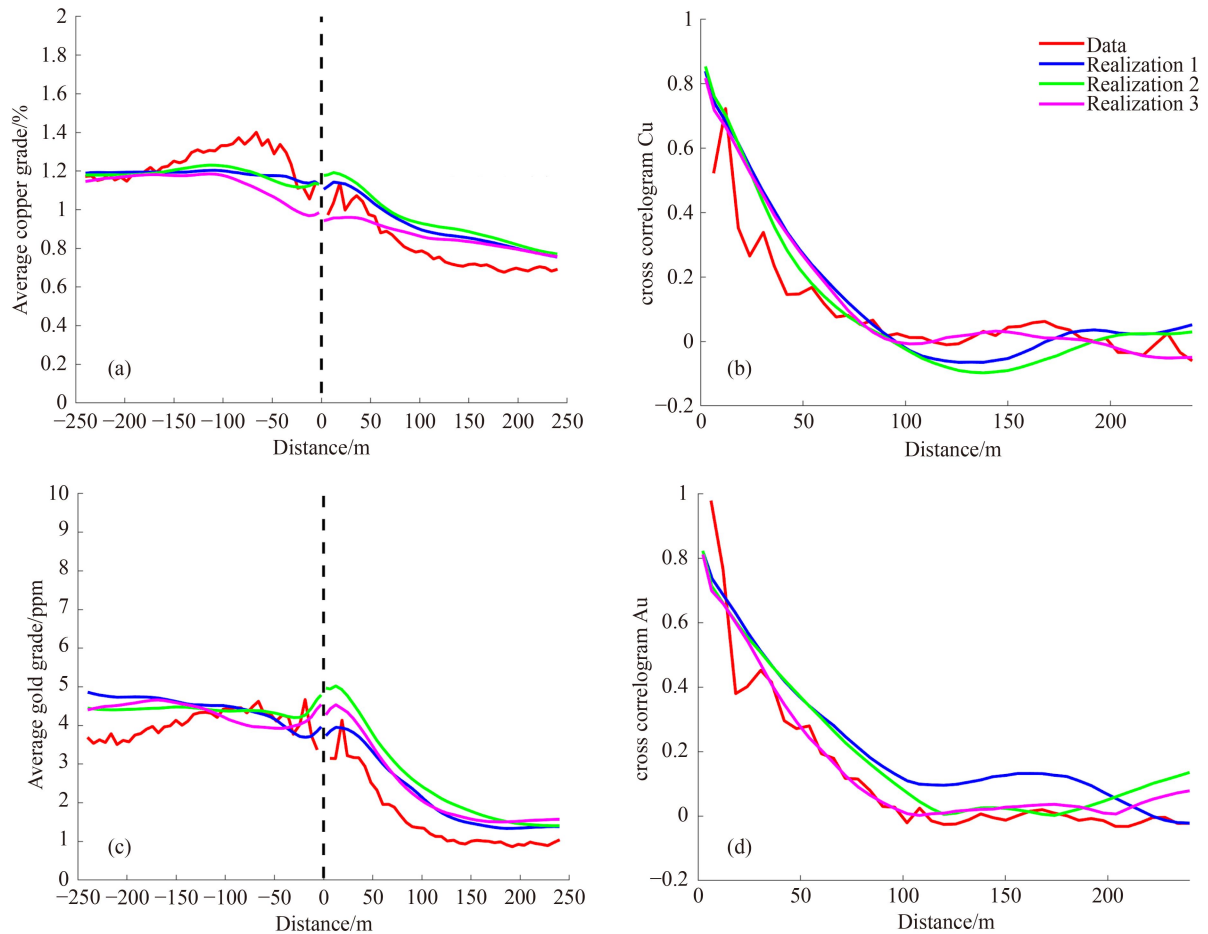


Fig. 9 Contact analysis between geological formations. Left: mean grade as a function of the signed distance to the geological formation boundary for (a) Cu ore grade and (c) Au ore grade; right: cross correlation between ore grades measured in different geological formations as a function of the distance between the sampling data for (b) Cu and (d) Au. Red line: original data; blue, green and purple lines: simulated ore grades.

algorithm for joint simulation can successfully predict the geological formations at test locations.

The uncertainty can also be reported for simulating categorical variables over the confusion matrix. For this purpose, the true values at test locations are considered versus the simulated values of geological formations at the same locations. For this purpose, the confusion matrix was separately computed in each realization. Since we produced 100 realizations in this study, the same number of confusion matrices was produced. The of accuracy and misclassification are shown in Fig. 11. As can be observed, the accuracy is high, ranging from 0.90 to 0.94, while the misclassification is quite low, ranging from 0.05 to 0.09. This shows that the proposed approach is capable of producing reliable results when simulating the categorical variable (geological formations) in the study area.

4 Comparison of the results

To show the capability of the proposed algorithm, the

results of the proposed approach were compared to another approach, where the original Cu and Au ore grades are independently transformed to normal scores by using projection pursuit multivariate transform (PPMT) (Barnett et al., 2014) instead of Gaussian anamorphosis as in our proposed approach. PPMT leads to two decorrelated factors that are normal standards. Therefore, these factors can be used in the simulation process without the need to infer the cross variogram between them. This facilitates variogram inference in the joint simulation workflow and simplifies the process of cosimulation. For this purpose, we used each factor together with the geological formations for a joint simulation. This means that in the simulation of Factor 1, one continuous variable (Factor 1) and one categorical variable (geological formation) were used, and in the simulation of Factor 2, another continuous variable (Factor 2) and the same categorical variable (geological formation) were used. These two simulations were performed independently. Once the simulation results for Factors 1 and 2 were completed, they were back transformed to the original scale of the Cu and Au ore

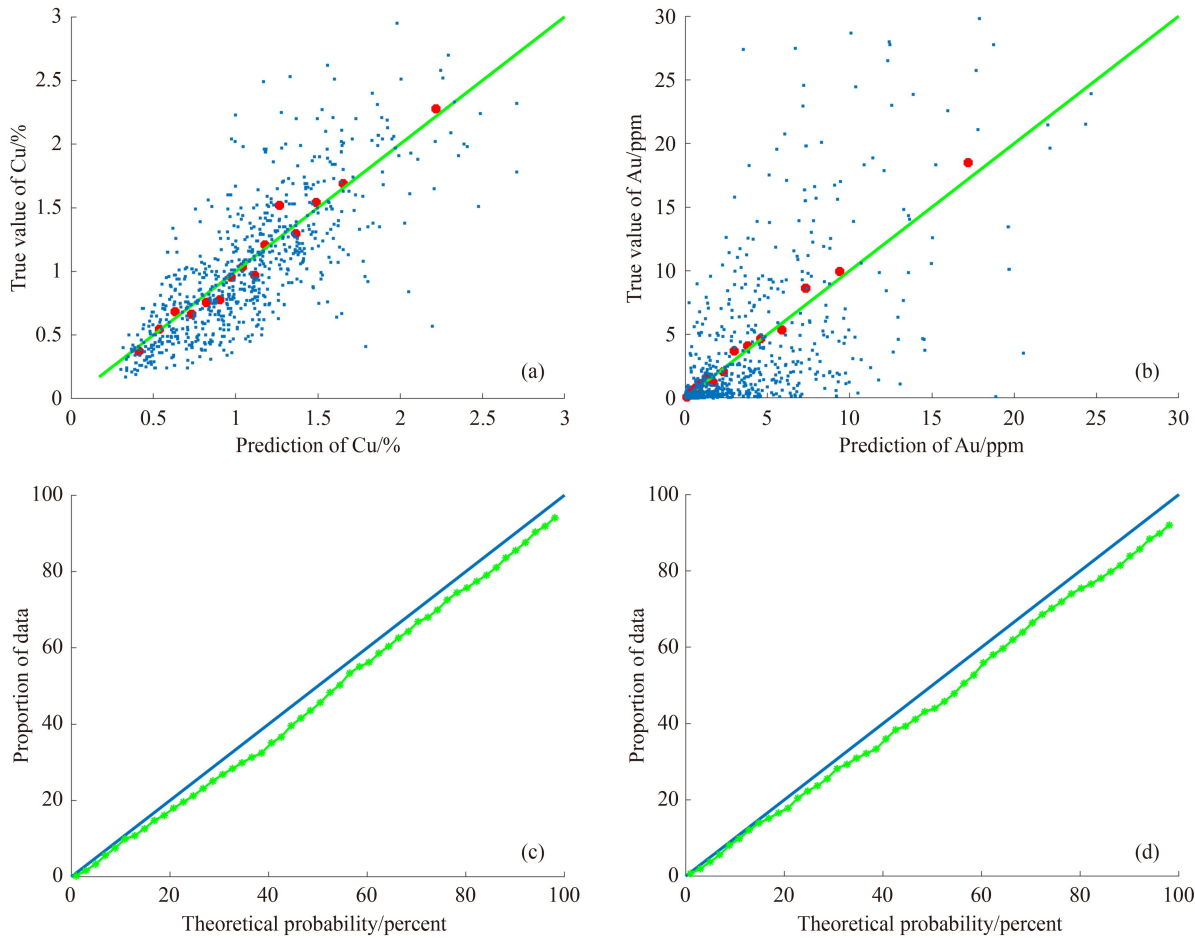


Fig. 10 (a) and (b) Jackknife validation of the simulated Cu and Au ore grades: scatter plot between true Cu and Au ore grades and the average of simulation results; (c) and (d) accuracy plot comparing the proportion of test data set within SPIs versus the nominal interval probability.

Table 2 Confusion matrix. The total number of test samples is 1664

		Predicted	
		Tourmaline breccia	Other rocks
True	Tourmaline breccia	460	62
	Other rocks	94	1048

grades through the PPMT back transformation technique. Another alternative is also considered following the hierarchical resource modeling approach. In this technique, the domain was split into two domains (i.e., tourmaline breccia and other rocks), and the original Cu and Au ore grades were separately cosimulated in each domain.

These joint simulation techniques were compared:

Case A: Joint simulation by the proposed approach;

Case B: Joint simulation by projection pursuit multi-variate transform (PPMT);

Case C: Hierarchical cosimulation.

Figure 12 shows the distribution of the conditional variance (CV) that was measured across 100 realizations for the aforementioned three cases by means of a boxplot. As seen, the interquartile range, which is the distance

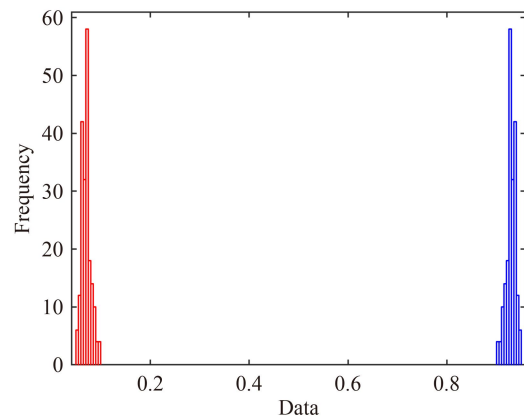


Fig. 11 Frequency of accuracy (blue histogram) and misclassification (red histogram) obtained from the confusion matrix containing over 100 realizations.

between the upper and lower quartiles in Case A, is less than the others for the Cu ore. This can be visually inspected as the width of the blue rectangles. The red line in this figure, which identifies the median of the variance distribution, is also quite low compared to Case B and Case C. This signifies that the proposed approach, in

which geological formations are simultaneously incorporated into the joint simulation workflow, substantially reduces the conditional variance for the Cu ore. However, this interpretation is different for the Au ore since the median of CV is almost the same in Case A and Case B but significantly higher in Case C. In addition, the range of CV for the Au ore is significantly reduced for the results obtained from PPMT in Case B. The reason for this result is still unknown to the authors.

It is also of interest to investigate which case is better in terms of reproducing the dependence between the Cu and Au ore grades and whether the type of normal score transformation (either through Case A & C or Case B) leads to any contributions in resource modeling. This can be checked by calculating the correlation coefficients between simulated continuous variables at the original scale within 100 simulations. Figure 13 shows the histogram of correlation coefficients for each case separately for 100 realizations. As seen, the average value in Case A ($\rho = 0.69$) is much closer to the original declustered correlation coefficient ($\rho = 0.71$), while in other cases, the deviation from the original correlation is significantly practical when the continuous variables are decorrelated and then modeled jointly with the categorical variable though PPMT (Case B) or hierarchically cosimulated in each domain separately (Case C).

The abovementioned measure of correlation provides a global perspective for making comparisons among different cases. Another benchmark for evaluating the quality of the proposed approach is to examine the reproducibility of the original bivariate relation between the Cu and Au ore grades. This can be verified by the scatterplot between the Cu and Au ore grades (Fig. 14). As can be observed, Case A shows a high compatibility of reproduced shape of disseminated points with original ones, whereas Case B and Case C poorly illustrate this relationship.

Figure 15 shows the contact analysis over the first realization obtained from three cases. Case A and Case B are able to reproduce the original soft contact boundary

(Fig. 4), while Case C fails to reproduce this boundary. This is compatible with what was proposed in Maleki et al. (2021) and Madani et al. (2021), who showed that contact analysis is necessary in mineral resource evaluation and that one should not be used in one another without justification.

5 Discussion

Using a reliable workflow for geostatistical modeling of geological domains and its implementation in ore body assessment is of paramount importance. The incorrect placement of geological boundaries in an ore deposit will lead to biased results in downstream activities of a mining project. This, for instance, can increase dilution, lead to biased ore/waste classification in the short term excavation of materials, invalid classification of mineral resources and false strategic mine planning, which also impacts medium- and short-term production scheduling. It can also lead to the improper valuation of economical parameters, and inefficient revenue and profit during the life of the mine. Indeed, using a reliable geostatistical modeling workflow is a necessity for the mining industry as the complexity of ore deposits has considerably increased. From another point of view, the spatial prediction of geological domains is not only important for further modeling of local grade distribution but also crucial for further optimization of mineral processing plants. For instance, it might be of interest to investigate the type of minerals and shape of clasts in the geological formation domains that significantly impact the design of a proper milling plant in a mine life cycle. Therefore, a proper explanation of these geological domains can directly regulate the processing route of the excavated materials. As a result, proper incorporation of the geological domains in resource estimation and geological block modeling is a key step in every mining project. The innovative technique proposed in this study is able to simultaneously take two metal grades into account for

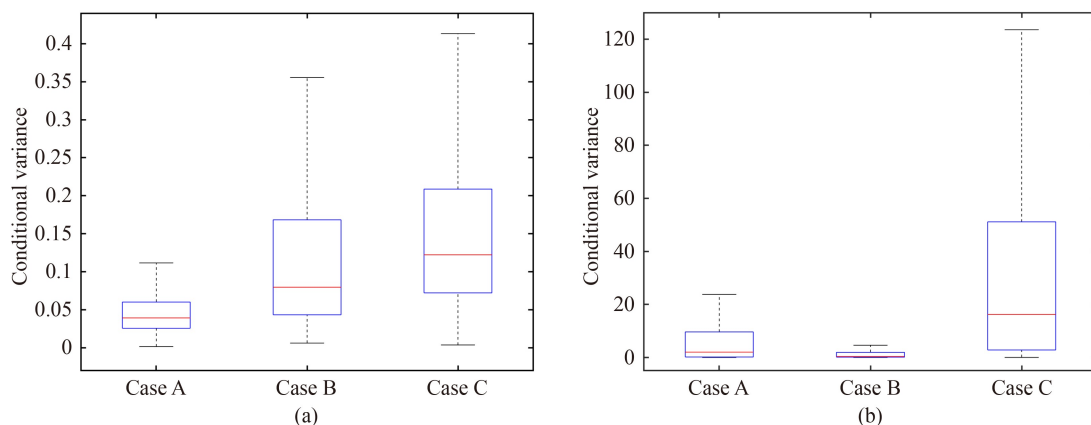


Fig. 12 Boxplot showing the distribution of the conditional variance across the realizations obtained from different cases; (a) Cu ore and (b) Au ore.

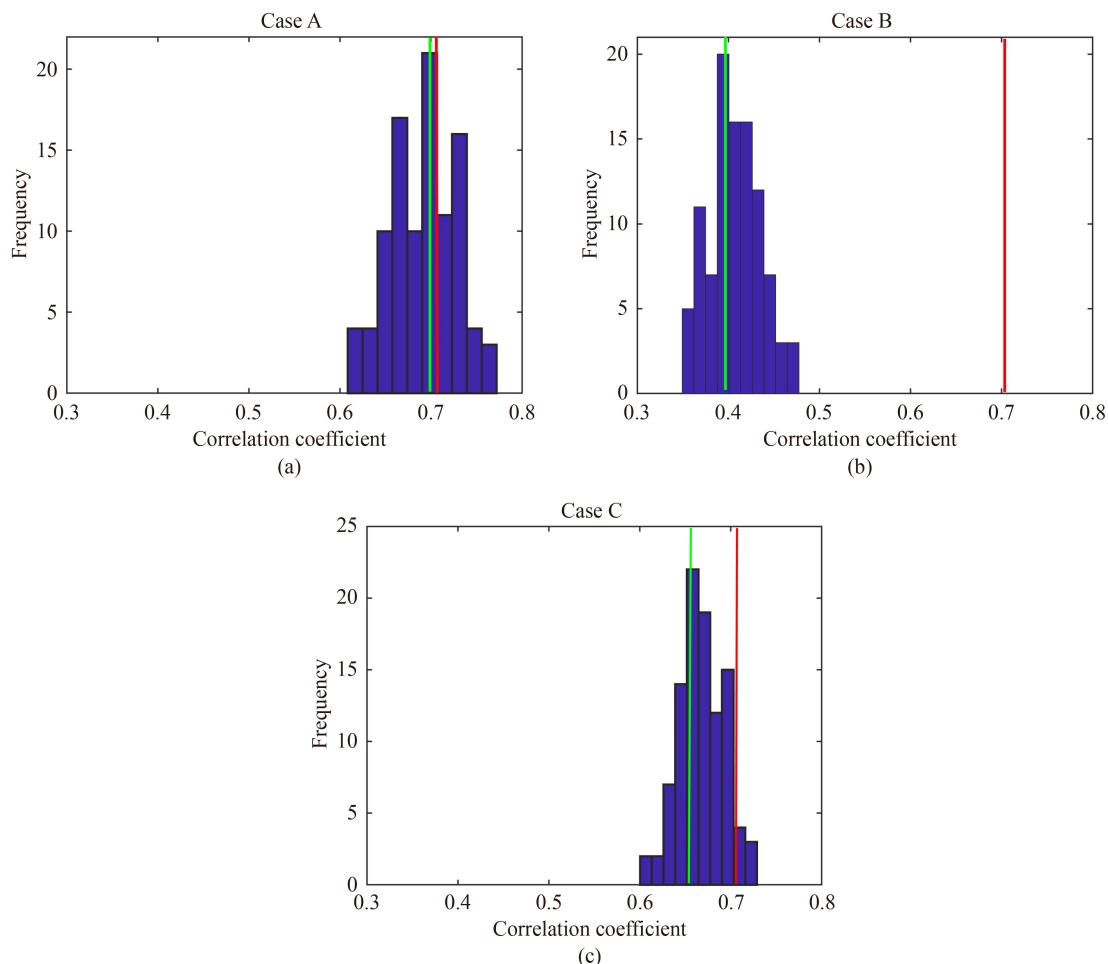


Fig. 13 Reproduction of the correlation coefficient between the back transformed simulation and the original Cu and Au ore grades on average for (a) Case A, 0.69 (green line) and for (b) Case B, 0.40 (green line), and for (c) Case C, 0.66 (green line). The original declustered correlation is 0.71 (red line).

joint simulation with geological domains that have been previously substantially overlooked. This will highly impact producing results that are both reliable in the local distribution of ore grades and trustworthy in the spatial modeling of geological domains.

The idea of this study was to employ a multi-Gaussian simulation for modeling of continuous variables from one side and a plurigaussian simulation for modeling the categorical variable from the other side. This algorithm is an alternative to the methodology proposed by [Emery and Silva \(2009\)](#), which was developed for modeling one continuous variable and a few categories set up in one categorical variable. In their methodology, two Gaussian random fields were used for modeling the categorical variable. Although this joint simulation is versatile, in practice, it becomes very complicated in the cases with more than two or three Gaussian random fields. As a result, one continuous variable will result in six and ten direct and cross variograms modeling to infer the linear model of coregionalization. This is also tedious when the thresholds of categories are numerous. However, in this study, we placed more emphasis on the continuous

variables rather than the categories. In the majority of ore deposits, it is common practice to deal with one important covariate of the main element, which significantly impacts the long- and short-term mine planning of a project. For instance, these can be byproducts, coproducts, or even deleterious elements. These variables usually show a good correlation with the main element and with the geological domains, for which their incorporation into the calculation of the economic block value for mine planning is of paramount importance and affects the net present value of the underlying mining project. It is inevitable that the proposed methodology in this study is restricted to the number of geological domains to two or more categories. In fact, our proposed joint simulation in this study is highly practical in the case of having two geological domains, such as ore and waste, poor ore zones and rich ore zones, low grade domains and high grade domains, and two cross-correlated continuous variables, where one variable is the main element and the other is a byproduct, coproduct or deleterious element, if the boundary of both continuous variables is of course soft across the geological domains.

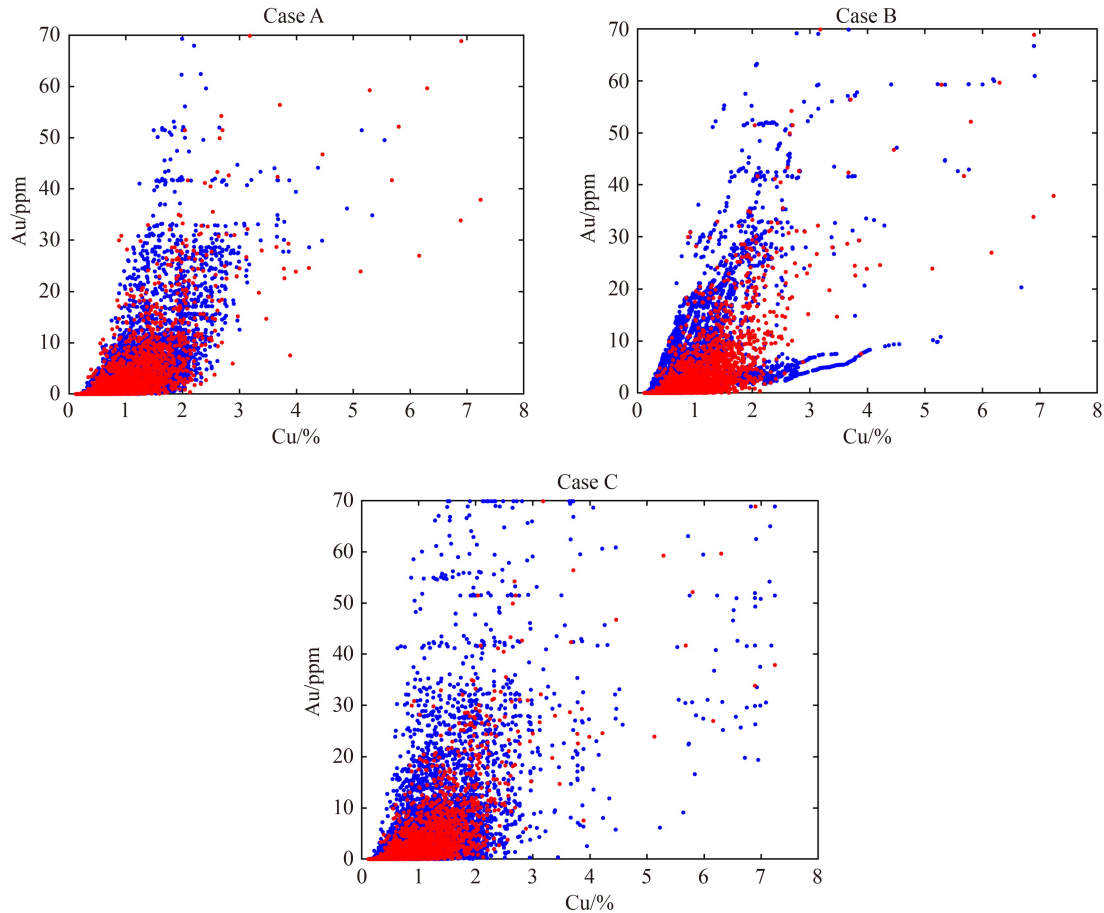


Fig. 14 Reproduction of the bivariate relations obtained from Case A, Case B, and Case C. Red points: original data and blue points: simulation results (Realization #1).

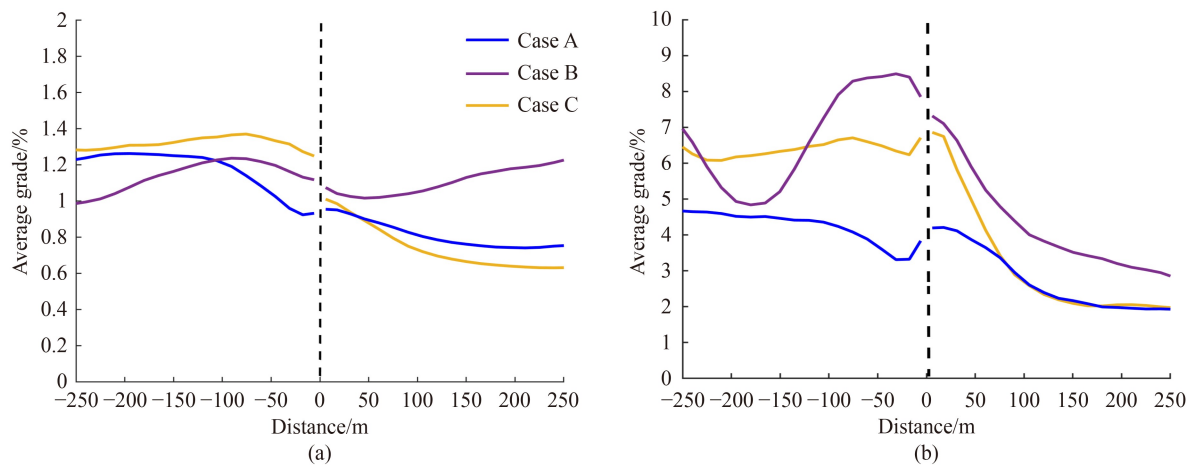


Fig. 15 Contact analysis between the geological formations; mean grade as a function of the distance to the geological formation boundary for (a) Cu and (b) Au ore grades.

In the case of hard boundaries, the hierarchical simulation algorithms are still effective for resource modeling and evaluation. These two methods (i.e., joint simulation and hierarchical simulation methods) cannot be used interchangeably, and contact analysis should always be implemented to select the proper simulation technique (Maleki et al., 2021).

6 Conclusions

An algorithm was proposed in this study for jointly and stochastically modeling two continuous variables and one categorical variable and was tested on a real Cu deposit, where the Cu and Au ore grades are controlled by two geological formations (tourmaline breccias and other rock

types), showing a soft boundary. To incorporate the cross dependence among all the continuous and categorical variables, the required statistical tools and algorithms were presented in this study for a cosimulation paradigm. In the first step, an iterative approach based on Monte Carlo simulation was presented for variogram analysis and inference of an optimum linear model of coregionalization. Next, following a plurigaussian simulation for modeling geological formations, where the categorical variable should be converted into normal scores, an alternative Gibbs sampler was discussed. The rationale of this iterative approach is also to incorporate the influence of continuous covariates, the Cu and Au ore grades. Therefore, the simple kriging in this simulation algorithm was replaced by a simple cokriging. The resulting simulated values at the data location show that they are in compliance with the corresponding geological formation (by applying the truncation threshold) and are also cross-correlated with the normal score transformed variables of the Cu and Au ore grades. In the output of this technique, 100 realizations of the converted categorical data and two normal score variables were available at sample locations. Then, a turning bands cosimulation algorithm was employed to cosimulate three Gaussian random vectors at the target block location. The final simulation model of the Cu and Au ore grades was obtained by a back transformation of two Gaussian random vectors to the original scale, and the final model of geological formations was derived by truncation of another Gaussian random field, each belonging to the corresponding conditioning data. The obtained models were then examined through a cross-validation technique and compared to other common geostatistical (co)-simulation algorithms through the reproduction of contact analysis, correlation coefficients, bivariate relations between the Cu and Au ore grades and the distribution of the conditional variance. These validations indicate that incorporation of the spatial variability of geological formations across the boundary (Cases A and B) substantially improves the modeling of the ore grades, resulting in a lower conditional variance. In contrast, the results also show that using decorrelation methods to first decorrelate the continuous variables and then jointly model them with categorical variables does not improve the modeling process (Case B). It is also shown that the hierarchical cosimulation technique (Case C) that is suitable for hard boundaries is suboptimal in the case of soft boundaries.

All these validation techniques corroborate the fact that the incorporation of the influence of the local and spatial variability of geological formations and its cross dependence with other continuous variables are necessary to produce satisfying results and decrease the uncertainty of the final grade model at unsampled locations. This notion is a necessity in any downstream activities of a mining project, where an erroneous evaluation of the

grade of interest for a block may negatively impact the capital investment of a project. However, there are some further lines of research to improve the proposed approach. The algorithm can be generalized to include more than two continuous variables and more categories with more complex contact relationships. In this case, a hierarchical plurigaussian simulation can be integrated with this technique to model more versatile contact relationships. Another difficulty that can arise in this case may be related to the variogram inference. One alternative approach is to adopt an iterative semiautomatic manner in which to infer the desired linear model of coregionalization. For instance, simulated annealing might be an option for this purpose. Another research avenue is concerned with a case when the multi-Gaussianity assumption cannot be corroborated between the continuous variables. For instance, a more complex shape of bivariate relation such as inequality constraint, nonlinearity and heteroscedasticity, may be shown. In these particular cases, a hierarchical cosimulation algorithm with the acceptance rejection technique (Madani and Abulkhair, 2020; Abulkhair and Madani, 2021), flow anamorphosis (van den Boogaart et al., 2017), and stepwise conditioning transformation (SCT) (Leuangthong and Deutsch, 2003) may be potential options for unlocking the complexity among these Gaussian variables. Nevertheless, these algorithms are limited to the sampling configuration of observations that are required to be isotopic, and their integration with joint modeling of categorical variables needs further investigation.

In the methodology proposed in this work or the initial idea proposed by Emery and Silva (2009), Emery and Cornejo (2010), the Gaussian assumption of the input variables is restricted. The reason is that the notion of the plurigaussian simulation that is proposed in these frameworks is based on the assumption of truncation of the Gaussian random fields. This means that the continuous variables should also possess a Gaussian distribution so that further analysis can be implemented. Therefore, other simulation algorithms that work based on non-Gaussian assumptions cannot be integrated with the plurigaussian simulation in this algorithm. More investigation is needed to determine the problem of modeling the non-Gaussian random fields when they show soft boundaries along the categorical domains.

Acknowledgment The first author is thankful to Nazarbayev University for funding this work via “Faculty Development Competitive Research Grants for 2018–2020 under Contract No. 090118FD5336 and 2021–2023 under Contract No. 021220FD4951”. This work is supported by Faculty Development Competitive Research Grants for 2018–2020 under Contract No. 090118FD5336 and 2021–2023 under Contract No. 021220FD4951.

References

Abulkhair S, Madani N (2021). Assessing heterotopic searching

- strategy in hierarchical cosimulation for modeling the variables with inequality constraints. *C R Geosci*, 353(1): 115–134
- Alabert F G, Massonnat G J (1990). Heterogeneity in a complex turbiditic reservoir: stochastic modelling of facies and petrophysical variability. In: SPE Annual technical conference and exhibition. Society of Petroleum Engineers
- Armstrong M, Galli A, Beucher H, Le Loc'h G, Renard D, Doligez B, Eschard R, Geffroy F (2011). *Plurigaussian Simulations in Geosciences*. Berlin: Springer
- Bahar A, Kelkar M (2000). Journey from well logs/cores to integrated geological and petrophysical properties simulation: a methodology and application. *SPE Reservoir Eval Eng*, 3(5): 444–456
- Barnett R M, Manchuk J G, Deutsch C V (2014). Projection pursuit multivariate transform. *Math Geosci*, 46(3): 337–359
- Boucher A, Dimitrakopoulos R (2012). Multivariate block-support simulation of the Yandi iron ore deposit, Western Australia. *Math Geosci*, 44(4): 449–468
- Cáceres A, Emery X (2010). Conditional co-simulation of Cu grades and lithofacies in the Rio Blanco-Los Bronces Cu deposit. In: Castro R, Emery X, Kuyvenhoven R, eds. *Proceedings of the IV international conference on mining innovation MININ 2010*. Gecamin Ltd, Santiago, Chile, 311–320
- Caers J (2005). *Petroleum Geostatistics*. Richardson, TX: Society of Petroleum Engineers
- Casella G, George E I (1992). Explaining the Gibbs sampler. *Am Stat*, 46(3): 167–174
- Deutsch C V (1996). Direct assessment of local accuracy and precision. In: Baafi E Y, Schofield N A, eds. *Geostatistics Wollongong '96*. Kluwer Academic Publishers, 1: 115–125
- Deutsch C V, Journel A G (1998). *Geostatistical Software Library and Users Guide*. New York: Oxford University Press
- Dowd P A (1994). Geological controls in the geostatistical simulation of hydrocarbon reservoirs. *Arab J Sci Eng*, 19(2B): 237–247
- Dowd P A (1997). Structural controls in the geostatistical simulation of mineral deposits. In: Baafi E Y, Schofield N A, eds. *Geostatistics Wollongong '96*. Dordrecht: Kluwer Academic: 647–657
- Dubrule O (1993). Introducing more geology in stochastic reservoir modelling. In: Soares A, ed. *Geostatistics Troia 92*. Dordrecht: Kluwer Academic: 351–369
- Duke J H, Hanna P J (2001). Geological interpretation for resource modelling and estimation. In: Edwards AC, ed. *Mineral resource and ore reserve estimation—the AusIMM guide to good practice*. The Australasian Institute of Mining and Metallurgy, Melbourne: 147–156
- Emery X (2005). Variograms of order ω : a tool to validate a bivariate distribution model. *Math Geol*, 37(2): 163–181
- Emery X, Cornejo J (2010). Truncated Gaussian simulation of discrete-valued, ordinal coregionalized variables. *Comput Geosci*, 36(10): 1325–1338
- Emery X, Lantuéjoul C (2006). Tbsim: a computer program for conditional simulation of three-dimensional gaussian random fields via the turning bands method. *Comput Geosci*, 32(10): 1615–1628
- Emery X (2007). Using the Gibbs sampler for conditional simulation of Gaussian-based random fields. *Comput Geosci*, 33(4): 522–537
- Emery X, Arroyo D, Peláez M (2014). Simulating large Gaussian random vectors subject to inequality constraints by Gibbs sampling. *Math Geosci*, 46(3): 265–283
- Emery X, Maleki M (2019). Geostatistics in the presence of geological boundaries: application to mineral resources modeling. *Ore Geol Rev*, 114: 103124
- Emery X, Silva D A (2009). Conditional co-simulation of continuous and categorical variables for geostatistical applications. *Comput Geosci*, 35(6): 1234–1246
- Freulon X, de Fouquet C, Rivoirard J (1990). Simulation of the geometry and grades of a uranium deposit using a geological variable. In: *Proceedings of the XXII International Symposium on Applications of Computers and Operations Research in the Mineral Industry*, Technische Universität Berlin, Berlin, 649–659
- Geman S, Geman D (1984). Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images. *IEEE Trans Pattern Anal Mach Intell*, 6(6): 721–741
- Glacken I M, Snowden D V (2001). Mineral resource estimation. In: Edwards AC, ed. *Mineral Resource and Ore Reserve Estimation—the AusIMM Guide to Good Practice*. Australasian Institute of Mining and Metallurgy, 189–198
- Goovaerts P (1997). *Geostatistics for Natural Resource Evaluation*. New York: Oxford University Press
- Haldorsen H H, Damsleth E (1990). Stochastic modeling. *J Pet Technol*, 42(4): 404–412
- Jones P, Douglas I, Jewbali A (2013). Modeling combined geological and grade uncertainty: application of multiple-point simulation at the Apensu Au deposit, Ghana. *Math Geosci*, 45(8): 949–965
- Journel A B, Huijbregts C J (1978). *Mining Geostatistics*. New York: Academic
- Kitanidis P K (1997). *Introduction to Geostatistics: Applications in Hydrogeology*. London: Cambridge University Press
- Larrondo P, Leuangthong O, Deutsch C V (2004). Grade estimation in multiple rock types using a linear model of coregionalization for soft boundaries. In: Magri E, Ortiz J, Knights P, Henriquez F, Vera M, Barahona C, eds. *Proceedings of the 1st international conference on mining innovation*. Gecamin Ltd, Santiago, Chile, 187–196
- Leuangthong O, Deutsch C V (2003). Stepwise conditional transformation for simulation of multiple variables. *Math Geol*, 35(2): 155–173
- Madani N, Abulkhair S (2020). A hierarchical cosimulation algorithm integrated with an acceptance–rejection method for the geostatistical modeling of variables with inequality constraints. *Stochastic Environ Res Risk Assess*, 34(10): 1559–1589
- Madani N, Biranvand B, Naderi A, Keshavarz N (2019a). Lithofacies uncertainty modeling in a siliciclastic reservoir setting by incorporating geological contacts and seismic information. *J Pet Explor Prod Technol*, 9(1): 1–16
- Madani N, Maleki M, Emery X (2019b). Nonparametric geostatistical simulation of subsurface facies: tools for validating the reproduction of, and uncertainty in, facies geometry. *Nat Resour Res*, 28(3): 1163–1182
- Madani N, Maleki M, Sepidbar F (2021). Integration of dual border effects in resource estimation: a cokriging practice on a copper porphyry deposit. *Minerals (Basel)*, 11(7): 660
- Maleki M, Emery X (2015). Joint simulation of grade and rock type in

- a stratabound Cu deposit. *Math Geosci*, 47(4): 471–495
- Maleki M, Emery X (2020). Geostatistics in the presence of geological boundaries: exploratory tools for contact analysis. *Ore Geol Rev*, 120: 103397
- Maleki M, Jélvez E, Emery X, Morales N, (2020). Stochastic open-pit mine production scheduling: a case study of an iron deposit. *Minerals (Basel)*, 10(7): 585
- Maleki M, Madani N, Jélvez E (2021). Geostatistical algorithm selection for mineral resources assessment and its impact on open-pit production planning considering metal grade boundary effect. *Nat Resour Res*, 30(6): 4079–4094
- Masoumi I, Kamali G, Asghari O, Emery X (2020). Assessing the impact of geologic contact dilution in ore/waste classification in the gol-gohar iron ore mine, southeastern Iran. *Minerals (Basel)*, 10(4): 336
- Mery N, Emery X, Cáceres A, Ribeiro D, Cunha E (2017). Geostatistical modeling of the geological uncertainty in an iron ore deposit. *Ore Geol Rev*, 88: 336–351
- Paravarzar S, Emery X, Madani N (2015). Comparing sequential Gaussian and turning bands algorithms for cosimulating grades in multi-element deposits. *C R Geosci*, 347(2): 84–93
- Pyrz M J, Deutsch C V (2014). *Geostatistical Reservoir Modeling*. New York: Oxford University Press
- Rivoirard J (1994). *Introduction to Disjunctive Kriging and Non-Linear Geostatistics*. New York: Oxford University Press
- Roldão D, Ribeiro D, Cunha E, Noronha R, Madsen A, Masetti L (2012). Combined use of lithological and grade simulations for risk analysis in iron ore, Brazil. In: Abrahamsen P, Hauge R, Kolbjørnsen O, eds. *Geostatistics Oslo 2012*. Berlin: Springer, 423–434
- Rossi M, Deutsch C V (2014). *Mineral Resource Estimation*. Berlin: Springer
- Stehman S V (1997). Selecting and interpreting measures of thematic classification accuracy. *Remote Sens Environ*, 62(1): 77–89
- Talebi H, Asghari O, Emery X (2014). Simulation of the lately injected dykes in an Iranian porphyry copper deposit using the plurigaussian model. *Arab J Geosci*, 7(7): 2771–2780
- Talebi H, Asghari O, Emery X (2015). Stochastic rock type modeling in a porphyry copper deposit and its application to copper grade evaluation. *J Geochem Explor*, 157: 162–168
- Talebi H, Mueller U, Tolosana-Delgado R (2019a). Joint simulation of compositional and categorical data via direct sampling technique—application to improve mineral resource confidence. *Comput Geosci*, 122: 87–102
- Talebi H, Mueller U, Tolosana-Delgado R, van den Boogaart K G (2019b). Geostatistical simulation of geochemical compositions in the presence of multiple geological units: application to mineral resource evaluation. *Math Geosci*, 51(2): 129–153
- Talebi H, Sabeti E H, Azadi M, Emery X (2016). Risk quantification with combined use of lithological and grade simulations: application to a porphyry copper deposit. *Ore Geol Rev*, 75: 42–51
- van den Boogaart K G, Mueller U, Tolosana-Delgado R (2017). An affine equivariant multivariate normal score transform for compositional data. *Math Geosci*, 49(2): 231–251
- Wilde B J, Deutsch C V (2012). Kriging and simulation in presence of stationary domains: developments in boundary modeling. In: Abrahamsen P, Hauge R, Kolbjørnsen O, eds. *Geostatistics Oslo 2012*. Berlin: Springer, 289–300