

# *Co-simulated Size Number: An Elegant Novel Algorithm for Identification of Multivariate Geochemical Anomalies*

**Nasser Madani & Emmanuel John  
M. Carranza**

**Natural Resources Research**

Official Journal of the International  
Association for Mathematical  
Geosciences

ISSN 1520-7439

Volume 29

Number 1

Nat Resour Res (2020) 29:13-40

DOI 10.1007/s11053-019-09547-9

**Your article is protected by copyright and all rights are held exclusively by International Association for Mathematical Geosciences. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at [link.springer.com](http://link.springer.com)".**



Original Paper

# Co-simulated Size Number: An Elegant Novel Algorithm for Identification of Multivariate Geochemical Anomalies

Nasser Madani<sup>1,3</sup> and Emmanuel John M. Carranza<sup>2</sup>

Received 26 June 2019; accepted 17 August 2019

Published online: 24 August 2019

Identification of geochemical anomalies is of particular importance for tracing the footprints of anomalies. This can be implemented by advanced techniques of exploratory data analysis, such as fractal/multi-fractal approaches based on priori or posteriori distribution of geochemical elements. The latter workflow involves analysis of 2D/3D produced maps, which can be mostly obtained by geostatistical algorithms. There are two challenging issues for such an analysis. The first one corresponds to handling the cross-correlation structures among the data, and the second one relates to the compositional nature of data. To tackle these problems, this paper investigates the application of Gaussian co-simulation for modeling the cross-correlated compositional data in order to recognize the multivariate geochemical anomalies in integration with fractal analysis. In this context, an innovative algorithm, namely co-simulated size number (CoSS-N), is introduced for this purpose. The compositional nature of data is addressed by additive log-ratio transformation of original data while the Gaussian co-simulation handles the reproduction of cross-correlation among the components. The co-simulated outputs are then taken into account for capturing different geochemical populations, showing different levels of backgrounds and anomalies. The algorithm is illustrated via a real case study located in Philippine wherever seven geochemical components are required to be considered. The accuracy of results is examined by statistical validation techniques, indicating the capability of the CoSS-N algorithm for multivariate identification of geochemical anomalies.

**KEY WORDS:** Fractal/multi-fractal analysis, Turning bands co-simulation, Geostatistical analysis, Co-simulated size number, Compositional data.

## INTRODUCTION

Identification of geochemical anomalies by separation of geochemical data into different populations is an essential aspect of mineral exploration. Unlike traditional approaches of anomaly recogni-

tion, which are based only on statistical parameters, fractal/multi-fractal modeling takes into account not only the frequency distribution of data but also spatial deviation of data from background due to “self-similar” fractal geometry (Mandelbrot 1983; Feder 1988; Cheng 2012; Sadeghi et al. 2012, 2015; He et al. 2013; Luz et al. 2014; Zuo and Wang 2016).

Fractal analysis for delineation of geochemical anomalies can be classified mainly into two groups (Madani and Sadeghi 2019). The first group deals with global distribution of a variable, in which the input to fractal analysis is based on sample frequency. For instance, the number–size model de-

<sup>1</sup>School of Mining and Geosciences, Nazarbayev University, Nur-Sultan, Kazakhstan.

<sup>2</sup>University of KwaZulu-Natal, Westville Campus, Durban, South Africa.

<sup>3</sup>To whom correspondence should be addressed; e-mail: [nasser.madani@nu.edu.kz](mailto:nasser.madani@nu.edu.kz)

scribes the relationship between number of objects (i.e., samples) and their sizes, which can be element concentration or metal grade (Mandelbrot 1983; Agterberg 1995). Although this methodology has vast applications in the analysis of geochemical data for mineral exploration, the spatial distributions of sample locations are not absolutely effective in capturing geochemical anomalies. In the particular case of scarce data, the sample distributions may not represent reliable statistical parameters of data populations, leading to bias in the inference of thresholds for defining data populations. In this case, the algorithm proposed by Madani and Sadeghi (2019) using Monte Carlo simulation algorithm (GSS-N) overcomes the problem of scarce data and thus facilitates straight-line fitting on log-log plots, leading to less bias in the identification of thresholds for defining data populations.

The conditional or posteriori distribution of a geochemical variable is the basis of the second group of fractal analysis related to the description of sample data that are mapped in either regular or irregular grid locations. For instance, the concentration-area (C-A) and concentration-perimeter (C-P) fractal models (Cheng et al. 1994, 1996), spectrum-area (S-A) fractal model (Cheng et al. 1999, 2000), concentration-volume (C-V) fractal model (Afzal et al. 2010), spectrum-volume (S-V) fractal model (Afzal et al. 2011), wavelet-number (W-N) fractal model (Chen and Cheng 2018) and simulated size-number (SS-N) fractal model (Sadeghi et al. 2015) pertain to this group. The last model, which was developed based on integration of simulated values and their number (or frequency), combines geostatistical simulation and fractal modeling to identify reliable thresholds. For this model, Gaussian simulation, either by sequential Gaussian simulation (SGS) (Almeida and Journel 1994) or by turning bands simulation (TBSIM) (Emery and Lantuéjoul 2006) as the two widely used geostatistical algorithms in uncertainty quantification, is recommended for construction of a stochastic local distribution that is suitable in distinguishing deposit-related geochemical anomalies. The SGS relies on the sequential paradigm of simulating Gaussian random fields conditioned to the set of data derived from a simple kriging exercise (Chilès and Delfiner 2012). Although this approach is considered reliable in geostatistical modeling, it has major difficulty in ascertaining moving neighborhood and in reproducing short-scale continuity (Lantuéjoul 1994). To address these shortcomings of SGS, it has been

shown that simulation results can be optimized by increasing the size of neighborhood ranges (Emery 2004). The TBSIM, the fundamental concept of which is based on simulation of one-dimensional Gaussian random fields and spanning those to d-dimensional random fields, has the ability to address the issue of increasing neighborhood ranges. However, it may create stripping artifacts in the case of few turning lines, and so the idea is to increase the number of turning lines to weaken this kind of artifacts (Lantuéjoul 1994; Gneiting 1999; Emery and Lantuéjoul 2006). These two geostatistical simulation algorithms are adequate for univariate spatial modeling; however, when one is dealing with more than one cross-correlated variable, the modeling becomes challenging. One reason is that these univariate simulation approaches insufficiently take into account the dependency characteristics among variables and subsequently lead to deficient results. In the case of dealing with cross-correlated variables, however, co-simulation algorithms are highly advocated. Sequential Gaussian co-simulation (Gómez-Hernández and Journel 1993; Pebesma 2004) and turning bands co-simulation or TBCOSIM (Carr and Myers 1985, 1989; Emery 2008), for example, offer much flexibility to construct outcomes (scenarios) that preserve inter-relationships between variables while spatial variability is preserved. Comparing and contrasting the performance of these two algorithms, Paravarzar et al. (2015) demonstrated that TBCOSIM outperforms sequential Gaussian co-simulation with respect to reproducing the cross-correlation calculated by spatial continuity and statistical parameters. Broadly speaking, Gaussian (co)-simulation approaches are the offshoot of deterministic approaches such as kriging. Although this most commonly used interpolation method is robust, it suffers from some impediments such as smoothing effect such that maps of estimated variables show squeezed distribution compared to the distribution of original data (Chilès and Delfiner 2012; Rossi and Deutsch 2014; Afzal et al. 2015; Battalgyzy and Madani 2019). This impacts over- and underestimation of de-clustered original low and high values in data, respectively. Moreover, using this biased result for further examination, such as fractal analysis and capturing of anomalies, leads to serious biased inference of essential thresholds (Sadeghi et al. 2015).

It is quite often that Gaussian co-simulation of cross-correlated variables is dealt with geochemical compositional data. In this case, the implementation

of conventional geostatistical modeling on the original scale of these data may be problematic and produce inconsistent results (Reimann et al. 2008; Pawlowsky-Glahn et al. 2015; Pawlowsky-Glahn and Egozcue 2016). In this respect, two main issues must be addressed: one related to closure problem (Aitchison 1986) and the other concerns spatial continuity analysis of compositional data (Pawlowsky-Glahn 1984; Pawlowsky-Glahn and Egozcue 2016). Considering the spatial structure of data means that the covariance structure of sample points should be taken into account in compositional data, and whenever it is recognized through the Aitchison geometry with Euclidean vector space structure, the additive log-ratio transformation becomes an oblique coordinate system with only limited use in multivariate statistical analysis (Pawlowsky-Glahn and Egozcue 2016). A simple solution, namely logarithmic transformation of original data (Reimann et al. 2008), although well suited for right-skewed distributions, is not appropriate for tackling the closure problem. Instead, log-ratio transformation developed in the 1980s for this purpose (Aitchison and Shen 1980; Aitchison 1982, 1986) has to be applied prior to any geostatistical analysis in the case of dealing with geochemical compositional data (Pawlowsky-Glahn 2003; Pawlowsky-Glahn and Olea 2004; Tolosana-Delgado 2006; Tolosana-Delgado and van den Boogaart 2013; Pawlowsky-Glahn and Egozcue 2016; Tolosana-Delgado et al. 2018). Three log-ratio transformations have been proposed: (1) additive log-ratio (alr) transformation (Aitchison 1986); (2) centered log-ratio (clr) transformation (Aitchison 1986); and (3) isometric log-ratio (ilr) transformation (Egozcue et al. 2003). Other approaches were introduced based on converting original compositional variables by stoichiometric closure characteristic (Mery et al. 2017; Adeli et al. 2018), which is not ascribed to logarithmic transformation.

In this paper, we propose an innovative algorithm, namely co-simulated size-number (CoSS-N), based on integration of geostatistical co-simulation of cross-correlated compositional data with fractal/multi-fractal analysis to enhance characterization of multivariate geochemical anomalies. This approach is illustrated through a case study pertaining to a geochemical exploration campaign in the Philippines. The geochemical compositions consist of seven components (arsenic, copper, zinc, nickel, cobalt, manganese and fill-up (or un-determined component)).

The main aim of this research is sixfold: (1) to present methodologies for fractal analysis, geostatistical co-simulation and compositional data analysis; (2) to propose an innovative algorithm, namely co-simulated size-number (CoSS-N); (3) to trial this algorithm on a real case study and discussing the relevant statistical analysis; (4) to compare results obtained from Gaussian co-simulation with independent simulation through global and local statistical validation techniques; (5) to derive multivariate geochemical anomalies by CoSS-N; and (6) to combine Gaussian co-simulation results and inferred thresholds for probabilistic description of anomalous areas for predictive mapping of epithermal Au deposits.

## METHODOLOGY

### Number-Size Fractal Model (N-S)

The N-S fractal model was proposed by Mandelbrot (1983) for describing the distribution of geochemical data. According to this model, there is a relation between the number and size parameters of evaluated data, which can be expressed as (Mandelbrot 1983):

$$N(\geq \rho) = F\rho^{-D} \quad (1)$$

where  $\rho$  denotes element concentration,  $N(\geq \rho)$  denotes cumulative number of samples with concentration values greater than or equal to  $\rho$ ,  $F$  is a constant and  $D$  is the scaling exponent or fractal dimension of the distribution of element concentrations. According to Mandelbrot (1983), log-log plots of  $N(\geq \rho)$  vs.  $\rho$  show straight-line segments with different slopes,  $-D$ , corresponding to different concentration intervals.

Agterberg (1995) proposed a “concentration-size” multi-fractal model based on the N-S model in order to determine and describe the spatial distribution of geochemical attributes in large mineral deposits. Monecke et al. (2005) used the N-S model to describe geochemical data, which indicate enrichment of minerals by replacement due to metasomatic processes resulting in the formation of hydrothermal deposits in the Waterloo massive sulfide deposit, Australia. The power-law frequency model, which has been suggested based on the N-S model, measures the frequency distribution of element and mineral concentrations based on the

number of samples (Li et al. 1994; Turcotte 1996; Shi and Wang 1998; Zuo et al. 2009; Sadeghi et al. 2012). The first 3D modeling work based on the N–S fractal model was demonstrated by Sadeghi et al. (2012) for the separation of mineralized zones and wall rocks, and the precision, and thus, the applicability of the model was proven by comparison of results with those of the concentration–volume (C–V) fractal model. The most important advantage of this model is that there is no need for preprocessing of data for pre-estimation before modeling. Sadeghi et al. (2015) presented the simulated size-number (SS-N) approach considering a Monte Carlo simulation over the local distribution. In this respect, different geostatistical simulation algorithms such as (TBSIM) (Emery and Lantuéjoul 2006) or SGS (Almeida and Journel 1994) have been recommended for construction of such a stochastic local distribution that can be used subsequently for distinguishing deposit-related anomalies. Sadeghi et al. (2015) used the TBSIM in their study for probabilistic modeling and further analysis of SS-N model, and they showed that the SS-N model outperforms the C–V model in terms of local and global statistical parameters reproduction. However, this algorithm is valid either when correlated covariates do not exist or when the global distribution and local distribution of the underlying variables are independent in the region.

### Turning Bands Co-simulation

The existence of cross-correlation among certain variables motivates one to use Gaussian co-simulation approaches rather than independent simulation (Wackernagel 2003; Chilès and Delfiner 2012; Eze et al. 2019; Abildin et al. 2019). The reason relates to taking into account the inter-dependency characteristic among certain variables, for which it leads to generate results that reproduce the local and global multivariate statistical parameters of data in the area of study. Among the Gaussian co-simulation techniques, TBCOSIM has received wide acceptance among practitioners because of its versatility and straightforwardness and it outperforms other Gaussian co-simulation algorithms with regard to reproducing cross-correlation parameters (Paravarzar et al. 2015). Therefore, in this study, we used TBCOSIM to show the capability of the proposed algorithm.

TBCOSIM is an extension of TBSIM, indicating an approximate algorithm based on the multi-Gaussian distribution assumption of the underlying random field as first introduced by Matheron (1973) and then extended in some organized program codes (Lantuéjoul 1994; Emery and Lantuéjoul 2006). The principal concept of this algorithm is based on, first, drawing plenty of lines with random orientations and, second, simulating a one-dimensional Gaussian random field along each line (Lantuéjoul 1994, 2002). In other words, the concept of the turning bands algorithm is to make simpler the problem of simulation in  $R^3$  or  $R^2$  into an  $R$  problem. The random field  $\{Y(x), x \in R^d\}$  in Eq. 1 has zero mean and isotropic covariance  $C_Y$  (Eq. 2), where  $U$  is a uniform distribution over  $S_d$ , which is unit sphere of  $R^d$ .

$$Y(x) = X(\langle x, U \rangle) \text{ for any location of } \forall x \in R^d \quad (2)$$

$$C_Y(r) = \int_{S_d} C_X(\langle h, u \rangle) \omega_d(du) \quad (3)$$

where  $\langle \cdot \rangle$  and  $\langle \cdot \rangle$  is the inner standard products in  $R^d$ ,  $h$  is a vector of  $R^d$ ,  $\omega_d$  is a uniform distribution over  $S_d$  and  $r$  is modulus of  $R^d$ . Matheron (1973) first introduced and proved Eq. 2 as a relationship between continuous and isotropic covariances in  $R^d$  with continuous covariances in  $R$ . By using this algorithm, one can substitute multi-dimensional simulation by one-dimensional simulation.

Having the covariance model fitted to the primary de-clustered normal score variable, the covariance function is derived from one-dimensional random fields. TBSIM provides a non-conditional multi-dimensional random field compatible with the target covariance model, in which the simulated values are practically standard Gaussian (Emery and Lantuéjoul 2006). In order to generate the conditional realizations, the non-conditional simulation obtained should be progressed through one post-processing of kriging step (Journel and Huijbregts 1978; Emery 2008; Chilès and Delfiner 2012).

In TBCOSIM, it is of interest to simulate stochastically the cross-correlated variables (more than two). In this case, the cross-covariance function is needed to construct such one- and multi-dimensional Gaussian random fields in the region. The non-conditional step is the same as TBSIM per variable; however, as part of the conditioning mechanism, co-kriging must be used rather than

kriging in order to keep the multivariate characteristics (Carr and Myers 1985; Myers 1989; Emery 2008); thus:

$$Y_{CCS}(x) = Y^{SCK}(x) + [Y_S(x) - Y_S^{SCK}(x)] \quad (4)$$

where  $Y^{SCK}(x)$  is the simple co-kriging of  $Y(x)$  from the conditioning data;  $Y_S(x)$  is the non-conditional simulation at location  $x$  for variables;  $Y_S^{SCK}(x)$  is the simple co-kriging of the non-conditional simulation from its value at the data locations. This simple co-kriging system can also be substituted for ordinary co-kriging paradigm (Emery 2007).

The general workflow in the latter case is similar to TBSIM, as previously explained; however, in the variogram analysis, since the co-kriging system is established on the basis of the cross-covariance matrix, it is necessary to calculate the direct and cross-variograms. In order to fit the theoretical direct and cross-variograms models, linear model of co-regionalization (LMC) as an alternative can be used to fit all experimental variograms as a linear combination of equivalent structures together with the identical ranges, but different in sills (Wackernagel 2003; Chilès and Delfiner 2012). The most tedious part of this job is to construct the permissible positive semi-definiteness conditions in fitting the sill matrices. Once this constraint is corroborated, the model can be used in the variance-covariance matrix required in the conditioning process (Goovaerts 1997).

### Compositional Data Analysis

In modeling of cross-correlated variables, it is common to deal with compositional data. In this context, these data are variables of a set of components forming a whole such that their sum can be chosen arbitrarily, without loss of information as a positive constant  $\Delta$  on the scale of the underlying random vector, such as 1 (parts per unit), 100%,  $10^6$  (parts per million) and  $10^9$  (parts per billion) (Aitchison 1986; Pawlowsky-Glahn et al. 2015). A vector of  $\delta$  components  $Z(x) = \{Z_1(x), Z_2(x), \dots, Z_\delta(x)\}$  is a composition of (Aitchison 1986; Pawlowsky-Glahn et al. 2015):

$$\forall Z_i(x) > 0 \quad \sum_{i=1}^{\delta} Z_i(x) = \Delta \quad (5)$$

The compositional space satisfying the condition in Eq. (5) is known as simplex  $S^\delta$ .

It is also common that the sum of components may not be bounded by  $\Delta$  and one needs to introduce the sub-composition by complementing the given components with a further component, namely fill-up or filler variable (Tolosana-Delgado et al. 2005, 2018). This characteristic is of paramount importance if the fill-up variable has a physical interpretation, in which it represents the totality of elements that were not assayed (Filzmoser et al. 2009; Tolosana-Delgado et al. 2018).

The modeling strategy of compositional data by geostatistical algorithms entails using some transformation techniques in order to preserve the summation constraint. Because the geostatistical-based algorithms have been mostly developed for original data without constraints, they implicitly follow the assumption of Euclidean geometry. Therefore, in the case of constant sum constraint, the statistical parameters such as global and local correlation coefficients may be deemed unrealistic (Pawlowsky-Glahn 1984; Aitchison 1986; Pawlowsky-Glahn and Egozcue 2016; Tolosana-Delgado et al. 2018). To address this difficulty, several log-ratio transformations have been introduced to remove the positively and sum constant of compositional data. Among others, the alr transformation, as a widely used approach, converts original data into log ratios as (Aitchison 1982, 1986; Pawlowsky-Glahn and Olea 2004):

$$F(x) = \left( \ln \frac{Z_1(x)}{Z_\delta(x)}, \ln \frac{Z_2(x)}{Z_\delta(x)}, \dots, \ln \frac{Z_{\delta-1}(x)}{Z_\delta(x)} \right) \quad (6)$$

where the denominator ( $Z_\delta(x)$ ) can be any of the components, yet the same one should be used for all the points and must be strictly positive. This denominator can be, for instance, the fill-up variable because it does not impact the final results of forward and backward transformation (Job 2010) and subsequently results in one less transformed variable than the original number of components. The reason for choosing alr transformation in this study relates to the suitability of spatial covariance structure analysis of a regionalized composition that is needed for the proposed algorithm (CoSS-N) (Pawlowsky and Burger 1992). The backward alr is represented (Aitchison 1982, 1986; Pawlowsky-Glahn and Olea 2004) as:

$$B(x) = \left( \frac{\exp(F_1)}{\sum_{i=1}^{\delta-1} \exp(F_i) + 1}, \frac{\exp(F_2)}{\sum_{i=1}^{\delta-1} \exp(F_i) + 1}, \dots, \frac{\exp(F_{\delta-1})}{\sum_{i=1}^{\delta-1} \exp(F_i) + 1} \right) \times \Delta \quad (7)$$

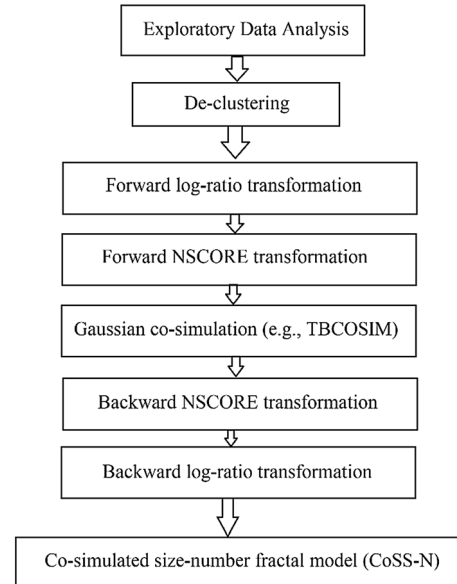
Once the transformed components are obtained, the geostatistical algorithms can be incorporated to the regionalized components  $F(x)$  rather than original data and the modeled results should be back-transformed to the compositional space afterward (Pawłowsky-Glahn and Egozcue 2016). Geostatistical simulation of compositional data in a studied region provides alternative scenarios with equiprobable and representative compositional random function at unsampled locations taking into account the transformed conditioning data (Goovaerts 1997; Chilès and Delfiner 2012; Talebi et al. 2019). Establishment of the co-kriging systems for co-simulation of transformed compositional data, however, requires modeling the direct and cross-spatial continuity (e.g., variogram or covariance). This can be attained from fitting a linear model of co-regionalization carrying direct variograms in the diagonal and cross-variograms off-diagonal (Tolosana-Delgado and van den Boogaart 2013). In order to validate the simulation results, Mueller et al. (2014) illustrate that co-simulation reproduces global statistical parameters of the mean value, variance and distribution of both variables, either in original scale or in log-ratio transformed values, in the case wherever the log-ratio transformed data do not significantly deviate from multi-normality assumption. Furthermore, satisfactory or unsatisfactory reproduction of spatial continuity shows the same characteristics because the conventional co-simulation methods (Van den Boogaart et al. 2017; Tolosana-Delgado et al. 2018). Van den Boogaart et al. (2017) also use the same validation techniques on transformed variables to the original scale for comparing different simulation algorithms.

### Co-simulated Size-Number Fractal Model (CoSS-N)

This proposed idea is based on relating the co-simulated size (CoSS) and number of involved compositional data ( $N$ ) through integration of the TBCOSIM and N-S fractal analysis with the following equation:

$$N(\text{CoSS} \geq \rho) = F\rho^{-D} \quad (8)$$

where  $\rho$  denotes element concentration,  $N(\geq \rho)$  is cumulative number of samples with the average of simulated concentration values greater than or equal to  $\rho$ ,  $F$  is a constant and  $D$  is the fractal dimension of the distribution of co-simulated element concentrations. In this model, a large number of possible scenarios can be reproduced using any Gaussian co-simulation algorithm (e.g., TBCOSIM) taking into account the global and spatial cross-correlation structure among the compositional data. The generated realizations not only reproduce the multivariate statistical parameters up to an acceptable level, but also more precisely distinguish different anomalies through fractal analysis by N-S. In the proposed algorithm (Fig. 1), first, exploratory data analysis must be implemented to reveal the inherent characteristics of compositional data. In this regard, if the dataset is not bounded by a constant  $\Delta$ , such as 1, 100 or  $10^6$ , a fill-up or filler variable, representing the not assayed elements, should



**Figure 1.** Overall workflow for the proposed algorithm based on CoSS-N model.



be considered. Second, the compositional data should be de-clustered prior to log-ratio transformation. Third is the transformation of original compositional random vectors by log ratio to obtain  $F(x)$  as in Eq. 6. Fourth is the transformation to normal score standard  $N(0,1)$ . Fifth is the (co)-spatial continuity analysis in order to infer the linear model of co-regionalization. Sixth, apply the TBCOSIM algorithm over normal score data. Seventh is the back-transformation of realizations from normal score into first, compositional data  $B(x)$  (Eq. 7) and, second, original scale. Finally, apply the CoSS-N models on simulated results to capture geochemical anomalies.

## DATASET

The geochemical dataset used here pertains to 135 stream sediment samples representing a total catchment area of  $\sim 101 \text{ km}^2$  in the Aroroy gold district (Philippines). This is the same dataset used in Carranza (2008, 2010a, b, 2011). The data are concentrations of six elements (As, Cu, Zn, Ni, Co and Mn) determined from the stream sediment samples by atomic absorption spectroscopy after aqua regia decomposition (JICA-MMAJ 1986). Among the uni-element data, the As data contain censored values (i.e., values below the detection limit of 0.5 ppm) associated with 35 of 40 sample catchment basins in the northeastern part of the district underlain mainly by diorite, which is neither genetically nor spatially associated with the epithermal gold deposits in the district (cf. Mitchell and Leach 1991; Carranza 2009), and none of the sample catchment basins associated with censored As values contains any of the known epithermal gold deposits. The censored As values were replaced by one half the detection limit. Then, each step as stated in the “Co-simulated Size-Number Fractal Model (CoSS-N)” section was performed on the basis of the proposed algorithm (CoSS-N) for capturing geochemical anomalies over cross-correlated compositional data.

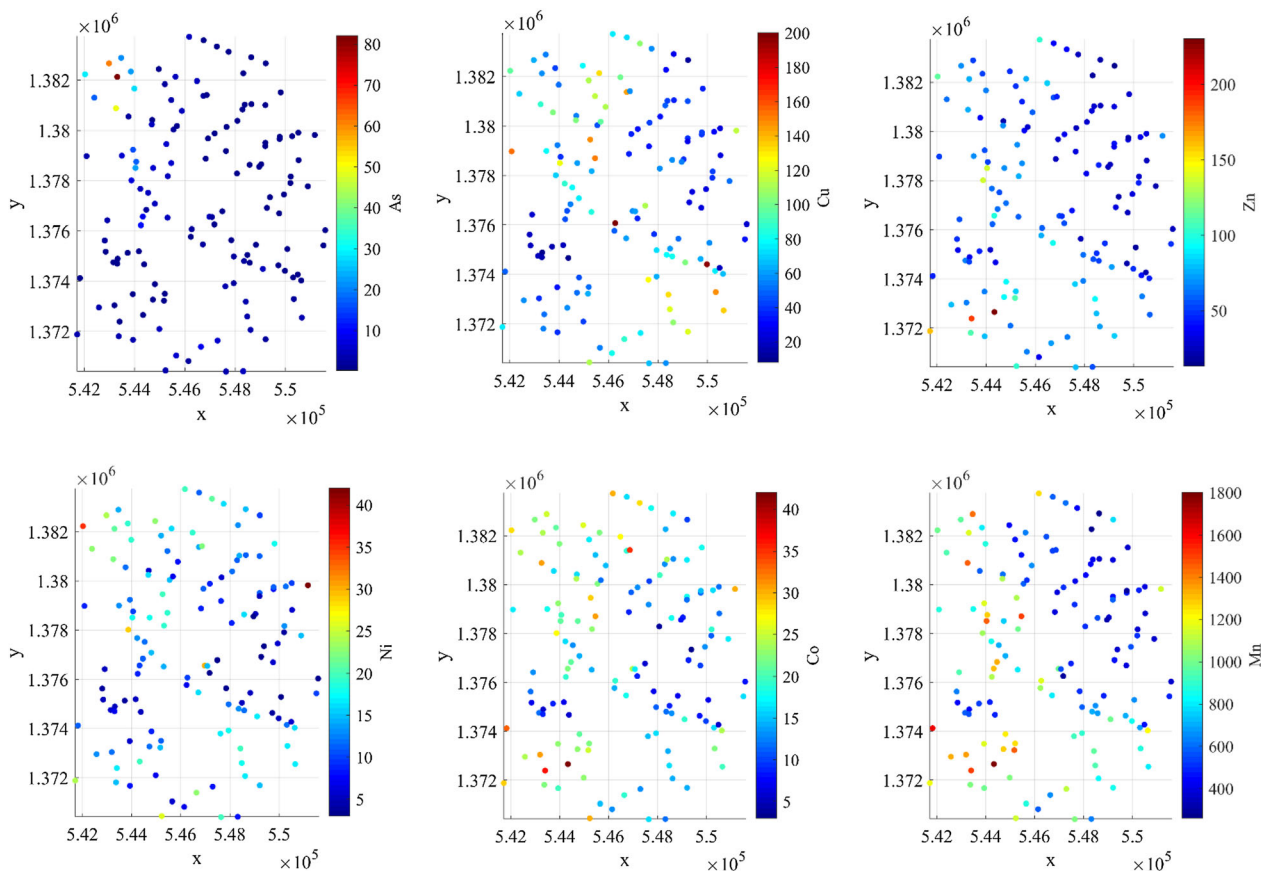
## EXPLORATORY DATA ANALYSIS

Each of the variables in the dataset was assayed in a homotopic pattern, implying that each of them was measured from the same sets of sample points and all sample locations are in common (Wacker-

nagel 2003). The location map of each variable is shown in Figure 2.

A straightforward tool for examining the spatial characteristic of data as a part of exploratory data analysis is trend analysis (Rossi and Deutsch 2014), which illustrates a cross-plot of the data against spatial coordinates. The interpretation of this plot in conjunction with the location map of data contributes to better understanding of the spatial variability of a variable being examined in a region. Figure 3 shows the variability of all six variables vs. principal coordinates, along west–east and south–north directions. Consideration of the plots from west to east (Fig. 3a) reveals that As is significantly decreasing while Mn, Co and Zn are mildly decreasing. However, the variability of the elements is different in the south–north direction, along which As is increasing notably while Mn, Zn and Cu are smoothly decreasing (Fig. 3b). The other variables do not show significant variation along these directions (Fig. 4).

Statistical parameters are crucial for describing the global characteristics of geochemical variables and possibly compositional data. If the sampling pattern is irregular, it is of particular importance to de-cluster the dataset in order to compute more representative statistical parameters (Goovaerts 1997). In this technique, each datum receives a weight between 0 and 1 based on spatial closeness to the surrounding data. Histograms and all summary statistics can then be calculated from the weighted data (Rossi and Deutsch 2014). In this study, cell de-clustering (Deutsch and Journel 1998) as a commonly used de-clustering technique was applied. In this technique, the area should first be divided into a grid of cells. The weights then can be calculated according to the number of data falling in the same cell. Each cell occupied by sample points is assigned the same weight regarding the number of data while an unoccupied cell receives no weight. A dimension of  $3000 \times 3000$  based on average sampling pattern is selected for the cell of interest to obtain the weights of each sample point. This cell size approximately corresponds to the spacing of the data in sparsely sampled regions (Rossi and Deutsch 2014). Table 1 shows the summary statistical parameters before and after de-clustering. For comparing the global statistics of the variables, the coefficient of variation (COV) is also provided. This unit-less measure of variability, which for all, is less than 2.5, indicating that one is not dealing with mixture of very high and very low values (Rossi and Deutsch 2014), and thus,



**Figure 2.** Location maps of values of six variables at Aroroy gold district. All element values are in parts per million (ppm).

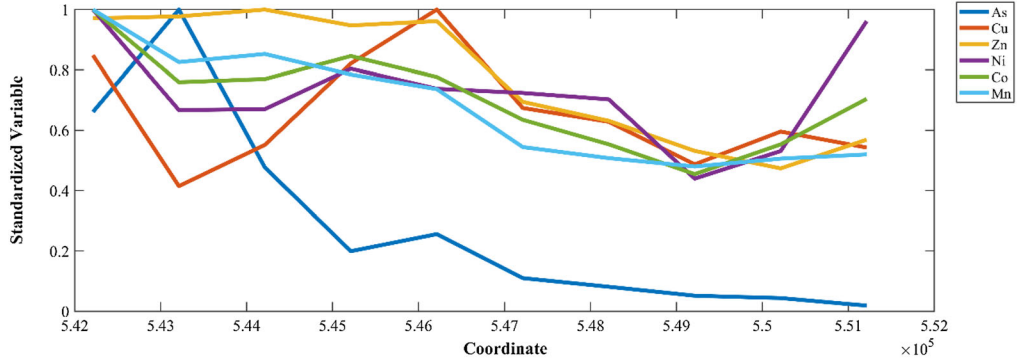
the separation of geological domains is not required, although As shows the highest COV among the variables. As can be seen from Table 1, the de-clustered mean for all the variables increased by 20.61% for As, 1.36% for Cu, 5.93% for Zn, 5.76% for Ni, 3.81% for Co and 0.35% for Mn. These significant ranges of variations indicate that the data are mostly clustered in low-valued areas (Rossi and Deutsch 2014).

Because the scope of this study is multivariate geostatistical modeling of variables, multivariate statistics of the variables are discussed. Correlation coefficients as popular measure for this quantify the global dependency that might exist among variables. The cross-correlation coefficient matrix for both clustered and de-clustered is provided in Table 2 (upper diagonal and lower diagonal, respectively). Note that there are strong correlations (i.e., above 50%) among Zn and Ni, Co, Mn, between pairs of Ni–Co and Co–Mn in both cases, indicating a high level of dependency in global relationship. All the

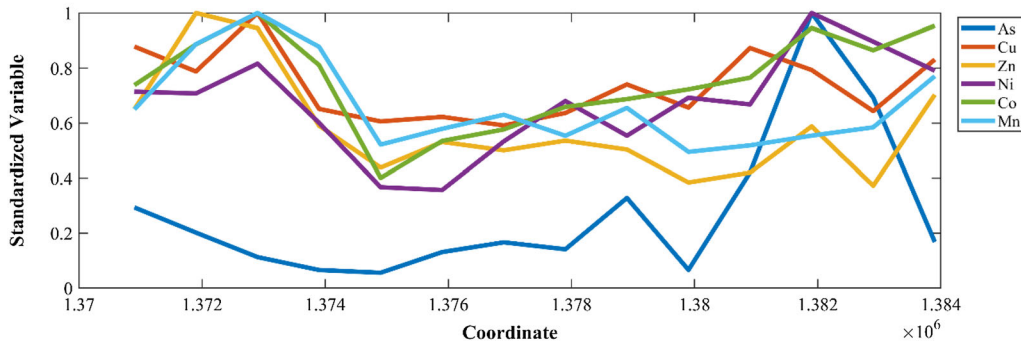
variables also are positively correlated. Since the de-clustering weights are determined on the basis of geometric configuration, only one set of weights is computed for correlation coefficients (Rossi and Deutsch 2014).

## LOG-RATIO TRANSFORMATION

The studied variables were transformed to log ratios for further analysis. Prior to this transformation, a fill-up variable was introduced to constrain the sum of components to  $10^6$  ppm. This variable represents the components in the samples that were not assayed. The descriptive statistical parameters of the compositional data may be uninformative (Aitchison 1986). In particular, statistical correlation of such data may be spurious and unrealistic (Pawlowsky-Glahn 1984; Pawlowsky-Glahn and Olea 2004). One of the reason for this relates to the Aitchison distance between two compositional

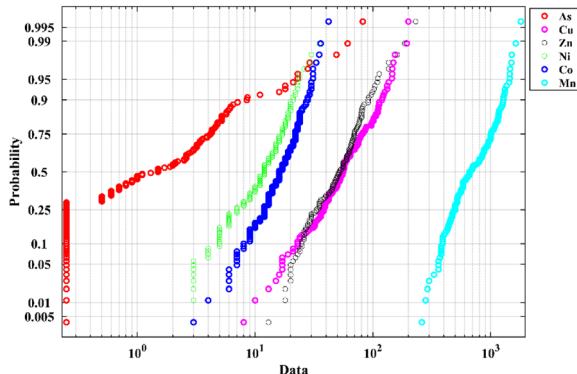


(a) Trend in West–East direction



(b) Trend in South–North direction

**Figure 3.** Trend analysis per variable against the principal axis of coordinates; the ordinate values are shown in standardized measure.



**Figure 4.** Lognormal probability plot of six cross-correlated variables.

observations,  $Z_\alpha(x)$  and  $Z_\beta(x)$ , which does not follow the typical Euclidean geometry (Aitchison 1986):

$$\begin{aligned} \text{Dist}_{\text{Aitchison}}(Z_\alpha(x), Z_\beta(x)) &= \sqrt{\left(\frac{1}{2\delta}\right) \times \sum_{i=1}^{\delta} \sum_{j=1}^{\delta} \left(\ln \frac{Z_{\alpha i}(x)}{Z_{\alpha j}(x)} - \ln \frac{Z_{\beta i}(x)}{Z_{\beta j}(x)}\right)^2} \quad (9) \end{aligned}$$

In this respect, for inferring the central tendency of each component, Aitchison proposed the use of compositional center for a dataset of size  $n$ , regardless of the type of log-ratio transformation:

$$\begin{aligned} \mathbf{M} = [g_1, g_2, \dots, g_\delta] \text{ with } g_i = \left(\prod_{j=1}^n Z_{ij}(x)\right)^{1/n}, \quad (10) \\ i = 1, 2, \dots, \delta. \end{aligned}$$

Furthermore, Aitchison (1986) defined the dispersion measure of a compositional dataset that can be described by variation matrix  $T = [t_{ij}]$ :

**Table 1.** Summary statistics of six variables before and after de-clustering

Before/after de-clustering	As	Cu	Zn	Ni	Co	Mn
Number of data	135	135	135	135	135	135
Mean	4.61/5.56	64.03/64.90	56.86/60.23	12.59/13.36	17.31/17.97	796.92/799.70
Variance	111.92/149.06	1421.63/1358.89	1044.01/1200.43	46.76/53.80	53.90/55.14	118,621.74/118,464.71
Minimum	0.25/0.25	8/8	13/13	3/3	3/3	260/260
Maximum	82/82	200/200	230/230	42/42	42/42	1800/1800
Coefficient of variation	2.29/2.19	0.58/0.56	0.56/0.57	0.54/0.54	0.42/0.41	0.43/0.40

**Table 2.** Cross-correlation matrix among the six variables (upper diagonal: before de-clustering, lower diagonal: after de-clustering)

	As	Cu	Zn	Ni	Co	Mn
As	1	0.095	0.213	0.363	0.285	0.333
Cu	0.082	1	0.304	0.405	0.288	0.304
Zn	0.19	0.298	1	<b>0.496</b>	<b>0.636</b>	<b>0.733</b>
Ni	0.384	0.423	<b>0.519</b>	1	<b>0.717</b>	0.399
Co	0.28	0.288	<b>0.656</b>	<b>0.712</b>	1	<b>0.686</b>
Mn	0.312	0.285	<b>0.718</b>	0.401	<b>0.704</b>	1

Bold values indicate that the correlation among variables are strong

$$t_{ij} = \text{Var} \left[ \ln \frac{Z_i(x)}{Z_j(x)} \right] \quad (11)$$

This can also be an approximation to measure correlation between two components (Buccianti and Pawlowsky-Glahn 2005), although it cannot be seen as a typical correlation coefficient. The low values of this measure represent that all two side observations show about the same ratio, while high variability reflects substantially different ratios of the two parts among the observations (Filzmoser et al. 2009). Its normalized version  $\exp(-t_{ij})$  transforms the values to the interval  $[0, 1]$  (Buccianti and Pawlowsky-Glahn 2005). The compositional center (closed geometric mean) and the variation matrix for all six compositional dataset and one fill-up variable are presented in Table 3.

## NORMAL SCORE TRANSFORMATION

In order to implement the geostatistical co-simulation algorithm, the alr-transformed variables (i.e., alr.As, alr.Cu, alr.Zn, alr.Ni, alr.Co and alr.Mn) should be transformed to standard Gaussian, with mean and variance of 0 and 1, respectively. The reason for Gaussian transformation relates to the multi-Gaussian distribution assumption that makes the inference of the conditional distribution in geo-

statistical algorithms convenient (Goovaerts 1997). However, several variables in the geosciences rarely obey this assumption and rather show asymmetric distributions, positively or negatively skewed (Deutsch and Journel 1998; Van den Boogaart et al. 2017). The transformation can be carried out through Gaussian anamorphosis (Rivoirard 1994) or quantiles-based approach (Deutsch and Journel 1998). In multivariate cases, it is common practice to transform each variable separately to normal standard score and employ one of the widely used functions of Gaussian co-simulation. In this study, the quantiles-based approach was applied to transform the six cross-correlated variables to normal standard score. The normal score transformed variables are depicted in normal probability plots (Fig. 5). The straight-line plots confirm that the distributions of the transformed variables follow the normal standard (ordinate), compared to the standard Gaussian distribution (abscissa).

Univariate Gaussian characteristics of the variables for (co)-simulation are required but not sufficient to ensure that a random function is multivariate Gaussian (Chilès and Delfiner 2012). This examination can be done by checking the experimental variogram of different order ( $w$ ) over the transformed Gaussian variables. In this respect, modogram ( $w = 1$ ) and rodogram ( $w = 1/2$ ) can be defined as (Deutsch and Journel 1998):

Modogram:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} |z(u_\alpha) - z(u_\alpha + h)|^1 \quad (12)$$

Rodogram:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} |z(u_\alpha) - z(u_\alpha + h)|^{1/2} \quad (13)$$

where  $[z(u_\alpha) - z(u_\alpha + h)]$  is an  $h$ -increment of the variable  $z$  and  $N(h)$  is the number of pairs.

**Table 3.** Compositional center of the dataset and variation matrix (lower diagonal: original, upper diagonal: normalized)

Center (ppm)	As	Cu	Zn	Ni	Co	Mn	Fill-up
	1.399	53.67	49.62	10.74	15.68	696.89	$9.9 \times 10^5$
As	0	0.172	0.213	0.181	0.169	0.195	0.109
Cu	1.760	0	0.701	0.699	0.707	0.697	0.678
Zn	1.543	0.354	0	0.736	0.818	0.893	0.760
Ni	1.708	0.357	0.305	0	0.869	0.719	0.703
Co	1.775	0.346	0.200	0.140	0	0.846	0.802
Mn	1.634	0.360	0.112	0.328	0.167	0	0.815
Fill-up	2.211	0.387	0.274	0.352	0.220	0.203	0

Once the two functions are computed, in order to show that the spatial bi-normality distribution assumption between transformed pairs is also honored, the experimental modograms and rodograms should be examined whether they are approximately proportional to the square roots of their experimental variograms (Emery 2005). This can be validated by plotting the former functions vs. the latter functions in a log–log coordinate. The application of this theory to the six cross-correlated variables shows that experimental points follow relatively the theoretical lines (Fig. 6), implying that the transformed variables are in satisfactory agreement with the bi-normality distribution assumption, and thus, the Gaussian (co)-simulation can be performed.

Experimental variograms were calculated for the transformed Gaussian variables in order to model their joint cross-correlation structures. This joint structure is required to solve the co-kriging system in co-simulation algorithm. The computation of direct variogram per variable in different azimuth directions (0°, 45°, 90° and 135°) revealed no significant geometric and zonal anisotropies, indicating that omni-directional variograms are sound enough for modeling such a joint spatial structure. In this regard, two spherical nested structures with ranges of 2000 and 8000, considering a nugget effect, were inferred through the linear model of co-regionalization. This fitting procedure was advocated due to its simplicity and versatility (Journel and Huijbregts 1978; Goovaerts 1997; Wackernagel 2003) and can be obtained by semi-automatic fitting algorithm (Emery 2010). For brevity, the outcomes of fitting for direct variograms are displayed in Figure 7, imposing the semi-definiteness condition, which guarantees the validity of the linear model of co-regionalizations. The complete formula of variogram analysis including the linear model of co-regionalization fitted sills is given in Eq. (14). The

main diagonal of the matrix relates to the fitted sills of direct variograms associated with upper and lower diagonals detailing the fitted sills of cross-variograms.

$$\begin{pmatrix}
 \gamma_{As} & \gamma_{As/Cu} & \gamma_{As/Zn} & \gamma_{As/Ni} & \gamma_{As/Co} & \gamma_{As/Mn} \\
 \gamma_{Cu/As} & \gamma_{Cu} & \gamma_{Cu/Zn} & \gamma_{Cu/Ni} & \gamma_{Cu/Co} & \gamma_{Cu/Mn} \\
 \gamma_{Zn/As} & \gamma_{Zn/Cu} & \gamma_{Zn} & \gamma_{Zn/Ni} & \gamma_{Zn/Co} & \gamma_{Zn/Mn} \\
 \gamma_{Ni/As} & \gamma_{Ni/Cu} & \gamma_{Ni/Zn} & \gamma_{Ni} & \gamma_{Ni/Co} & \gamma_{Ni/Mn} \\
 \gamma_{Co/As} & \gamma_{Co/Cu} & \gamma_{Co/Zn} & \gamma_{Co/Ni} & \gamma_{Co} & \gamma_{Co/Mn} \\
 \gamma_{Mn/As} & \gamma_{Mn/Cu} & \gamma_{Mn/Zn} & \gamma_{Mn/Ni} & \gamma_{Mn/Co} & \gamma_{Mn}
 \end{pmatrix}
 = \begin{pmatrix}
 0.093 & -0.039 & -0.058 & -0.037 & -0.060 & -0.060 \\
 -0.038 & -0.067 & 0.004 & -0.099 & -0.079 & -0.009 \\
 -0.058 & -0.004 & 0.129 & 0.133 & 0.0833 & -0.018 \\
 -0.037 & -0.099 & 0.133 & 0.341 & 0.237 & 0.0519 \\
 -0.060 & -0.079 & 0.083 & 0.237 & 0.284 & 0.102 \\
 -0.060 & -0.009 & -0.018 & 0.051 & 0.102 & 0.122
 \end{pmatrix} \text{nugget}$$

$$+ \begin{pmatrix}
 0.055 & 0.208 & 0.071 & 0.124 & 0.123 & 0.087 \\
 0.208 & 0.803 & 0.276 & 0.448 & 0.454 & 0.317 \\
 0.0715 & 0.276 & 0.096 & 0.147 & 0.153 & 0.105 \\
 0.124 & 0.448 & 0.147 & 0.348 & 0.290 & 0.236 \\
 0.123 & 0.454 & 0.153 & 0.290 & 0.379 & 0.225 \\
 0.087 & 0.317 & 0.105 & 0.236 & 0.225 & 0.166
 \end{pmatrix}$$

$$\text{Sph}(2000 \text{ m}, 2000 \text{ m}, 2000 \text{ m})$$

$$+ \begin{pmatrix}
 1.043 & 0.333 & 0.667 & 0.477 & 0.471 & 0.661 \\
 0.333 & 0.222 & 0.194 & 0.183 & 0.037 & 0.098 \\
 0.667 & 0.194 & 0.911 & 0.283 & 0.442 & 0.832 \\
 0.477 & 0.183 & 0.283 & 0.347 & 0.292 & 0.159 \\
 0.471 & 0.037 & 0.442 & 0.292 & 0.460 & 0.416 \\
 0.661 & 0.098 & 0.832 & 0.159 & 0.416 & 0.930
 \end{pmatrix}$$

$$\text{Sph}(8000 \text{ m}, 8000 \text{ m}, 8000 \text{ m})$$

Once the variograms analysis was achieved, the co-simulation was implemented on a regular 2D grid of 50 m by 50 m. The TBCOSIM algorithm was used with a moving neighborhood in the shape of circle with radius of 5000 m, incorporating up to 200 sample points regardless of considering any quadrant for splitting the neighborhood circles to equal portions. The number of realizations and the turning

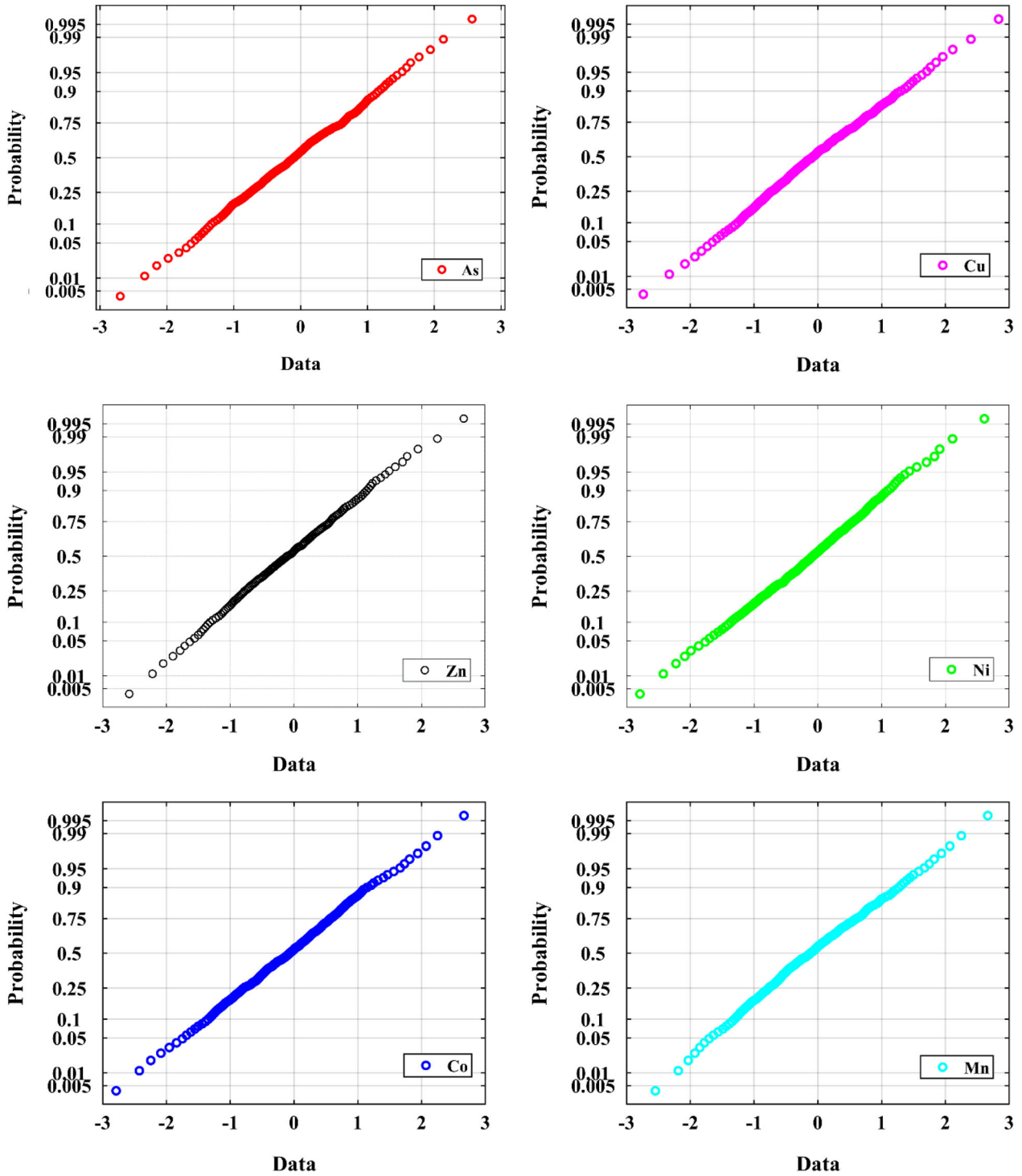


Figure 5. Normal probability plots for normal scores of six alr-transformed variables of interest.

bands were set to 100 and 5000, respectively. Indeed, each of this number of realization mostly corresponds to the efficient characterization of the space of uncertainty (Goovaerts 1999) and should be based on minimum acceptance criteria for geostatistical simulation (Leuangthong et al. 2004). Abildin et al. (2019) showed that 100 realizations are an adequate number in co-simulation of cross-correlated variables for convergence of statistical parameters of the simulation results to the original de-clustered ones. The reason for arranging the latter number of turning bands to 5000 also relates to elimination of stripping effects that should be as large as possible (Emery 2008).

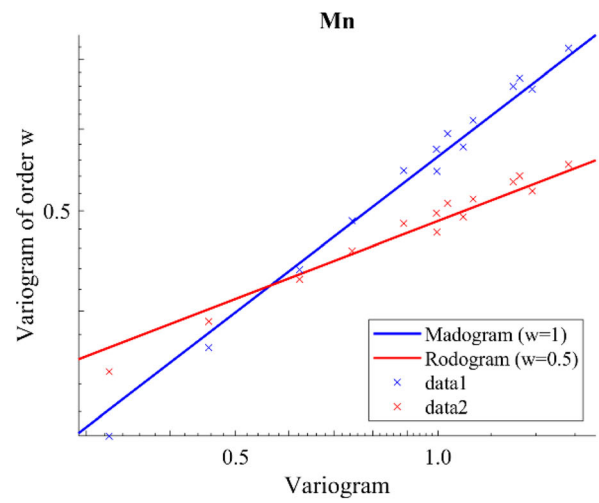
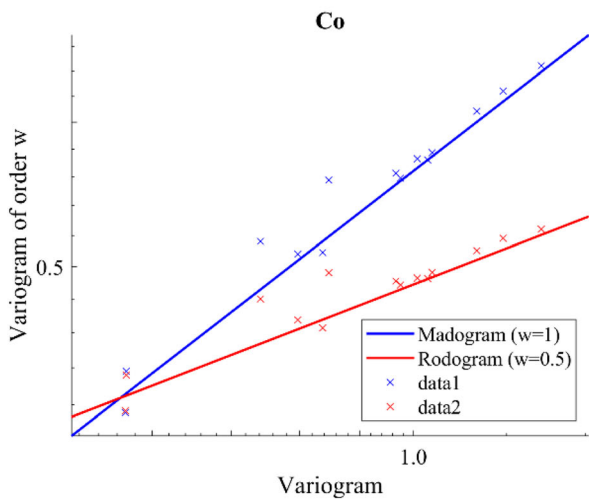
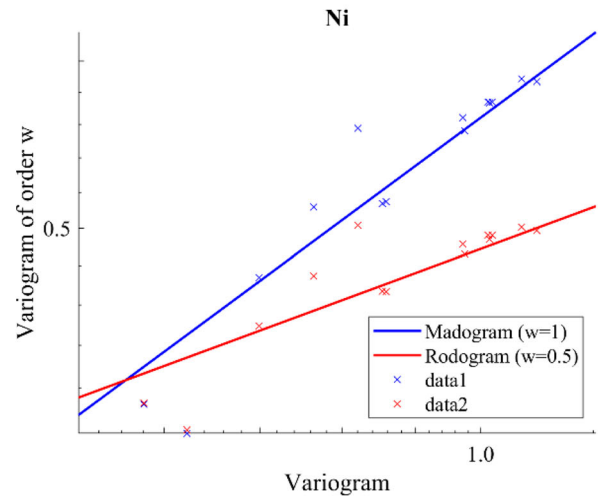
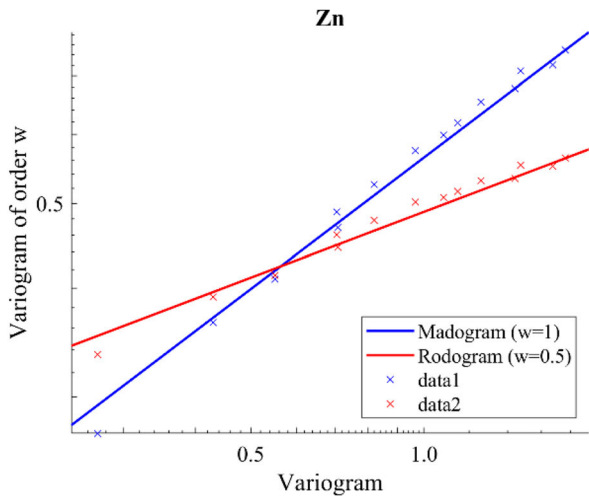
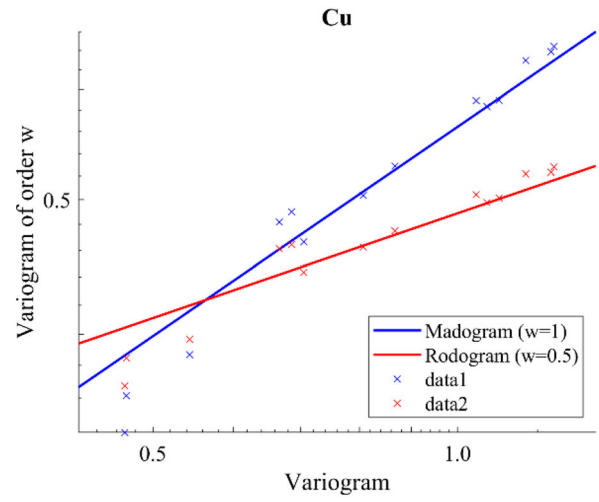
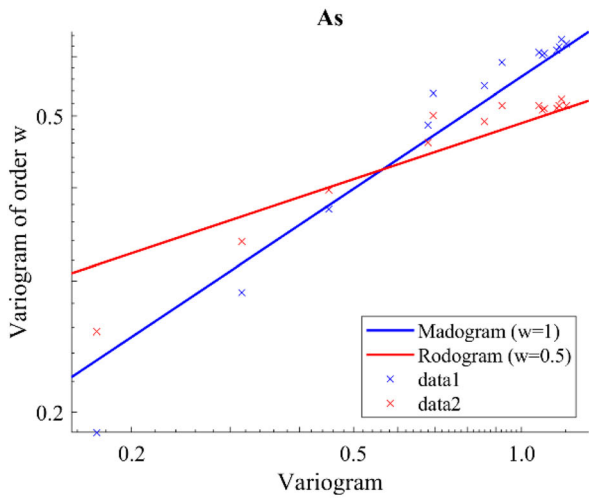
In the process of conditioning, as explained above, simple co-kriging is required to be considered. Multiple searching strategy was also selected, because Madani and Emery (2019) showed that this type of neighborhood is more precise and reliable. This strategy was performed in six parts, in which the neighborhood was searching for 200 closest data of the first, second, third, fourth, fifth and sixth variable in turn, independent of each other. Ordinary co-kriging can likewise be taken into account in the conditioning step, but in such a case the mean values of cross-correlated variables are uncertain (Emery 2007). In order to compare the results of joint simulation based on the methodology proposed (CoSS-N), independent TBSIM was also performed for each variable following Sadeghi et al. (2015) (i.e., SS-N) regardless of cross-correlation among the variables and compositional nature of the dataset. In this workflow, simple co-kriging in TBCOSIM was substituted for simple kriging taking into account the only normal score conditioning data that were obtained from Gaussian transformation of de-clustered original dataset. Then, direct variograms were computed on the normal score dataset and independent simulation (TBSIM) was implemented on each variable with the same number of realizations and turning bands as considered for TBCOSIM. The simulation results were then back-transformed to original scale. In TBCOSIM, however, we needed one extra back-transformation. In this case, the co-simulated realizations that keep the normal score distribution were back-transformed first and then scaled to alr-transformed variables and were then back-transformed to original scale (ppm) by Eq. 5.

E-type maps, obtained by geometric averaging the back-transformed variable of interest across 100 realizations of TBCOSIM per block, are depicted in Figure 8. As can be seen, the results are compatible

with the trend analysis already provided in the above section. For instance, As is concentrated in the left corner of the region and is decreasing toward the east. Another confirmation of the reproduction of dependency among the variables is related to the strong correlations between the pairs of Co–Mn (0.704), Ni–Co (0.712) and Zn–Mn (0.718) and relatively weaker correlations between the pairs of Zn–Ni (0.519) and Zn–Co (0.656), which can be visualized from the produced maps, implying that the TBCOSIM was capable of qualitatively reproducing the de-clustered correlations among the variables through all the realizations.

## VALIDATION

In this section, it is of interest to validate the simulated results obtained from TBCOSIM and compare to the outputs of independent simulation, TBSIM. The scope of this work is mainly on the basis of reproduction of global and local correlation coefficient and its application to enhance characterization of geochemical anomalies. The CoSS-N model is proposed for compositional data as an innovative algorithm for showing the capability of Gaussian co-simulation to fit the latter purpose. In this regard, we decided to make this comparison (i.e., between TBSIM and TBCOSIM) over the conventional statistical parameters for several reasons. First of all, in this paper, the algorithm for TBSIM deems that the original dataset is non-compositional while neglecting the meaningful interrelationship among the variables (i.e., following the algorithm provided by Sadeghi et al. 2015 that is for univariate modeling). Second is the lack of practical guidelines in commercial software from the industrial perspective for dealing with modeling of compositional data (Tolosana-Delgado et al. 2018) and the corresponding statistical analysis, even after providing the monographs (Pawlowsky-Glahn and Olea 2004) despite enhanced development of log-ratio transformation techniques (Pawlowsky-Glahn 2003; Tolosana-Delgado 2006; Tolosana-Delgado and van den Boogaart 2013). Third, subject to the behavior of log-ratio transformed variables and their deviations from normal distribution, alternative examination for reproduction of histogram and variogram on both original scale and log-ratio transformed variables shows similar results (Mueller et al. 2014). Fourth, alr-transformed variables are no longer constrained by constant sum and common





◀ **Figure 6.** Experimental modograms and rodograms of the Gaussian variables as a function of their experimental variograms. In case of bi-normality, the data points should be approximately distributed along the thick solid line.

geostatistical methods can be incorporated to derive the spatial relationships among the variables (Pawłowsky and Burger 1992). Fifth, there may be some sound arguments against using the original correlation and statistical parameters among the primary variables for decision-making that may lead to spurious decisions (Pawłowsky-Glahn and Egozcue 2016 and references therein). This reason is advocated, because the original variables are just a subgroup of a whole family known as compositional dataset, and the original de-clustered correlation still can be considered as an approximation of inherent dependency among the compositional data and may be examined as criteria for comparison between alternative geostatistical simulation methodologies. Sixth, there are several case studies based on geostatistical simulation of log-ratio transformation of original variables and contrasting to other methods that the validation parts are mainly based on conventional statistical parameters (e.g., Rubio et al. 2016; Van den Boogaart et al. 2017; Hosseini and Asghari 2019 and references therein).

### Multivariate Statistical Parameters Reproduction

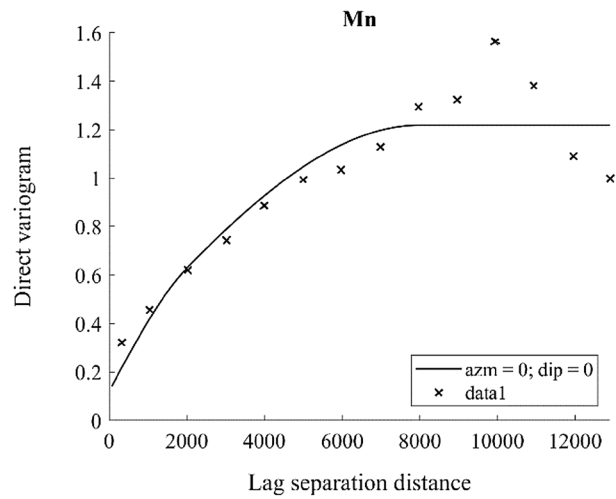
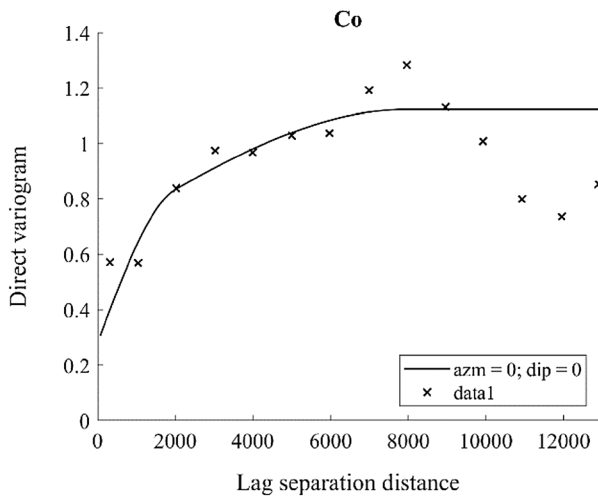
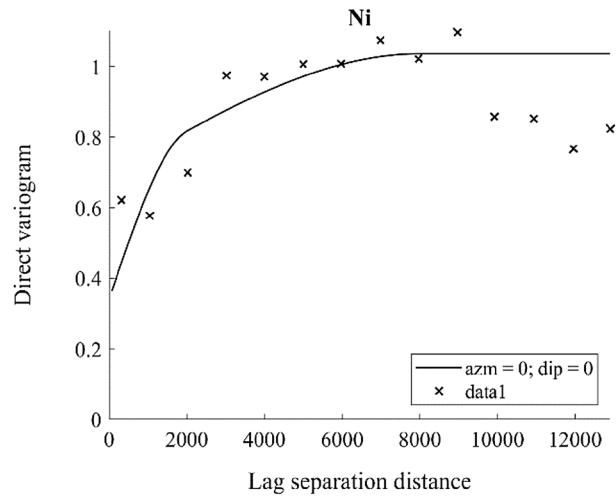
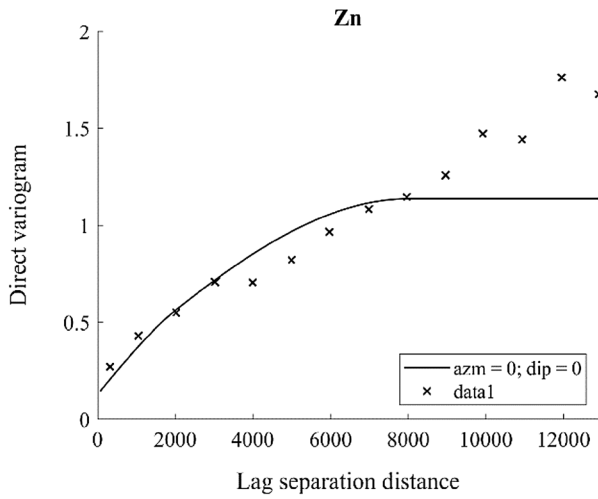
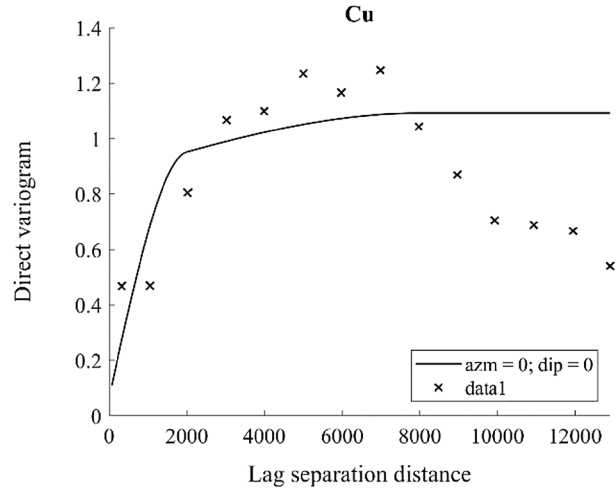
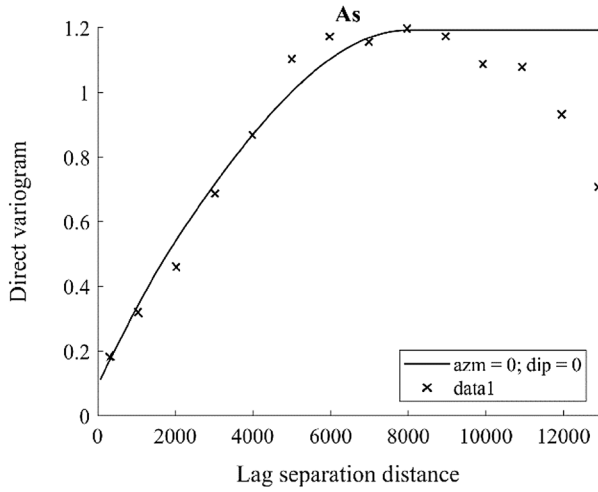
In order to validate the results, it is firstly of interest to check that the realizations reproduce the global dependency relationships among the variables. In this respect, one can compute the correlation coefficients between the back-transformed values of simulated variables in each realization (Eze et al. 2019). It is expected that these coefficients fluctuate around the experimental correlations as provided in Table 2, for instance 0.704 for Co–Mn, 0.712 for Ni–Co, 0.718 for Zn–Ni and 0.656 for Zn–Co. This situation occurs in the case of TBCOSIM, but not in TBSIM (Fig. 9). In the latter algorithm, the produced correlations through 100 realizations are, on average, lower than the expected de-clustered correlations. This bias can be explained because, compared to TBCOSIM, TBSIM does not consider the cross-dependency among variables, leading to poor reproduction of the cross-correlation among the simulated variables. The reproduced correlation coefficients are given in Table 4. Nevertheless, as mentioned earlier, there are some

legitimate arguments concerning computation of Pearson correlation coefficients among compositional variables that may be deemed spurious. Therefore, in order to examine the interrelation dependency regardless of measure of correlation coefficient for comparison of two algorithms, the scatter plot between pairs of Zn–Ni, Co–Zn, Mn–Zn, Co–Ni and Mn–Co are depicted in Figure 10. In this plot, one can examine the reproduction of bivariate relation between the variables whether or not it follows the original de-clustered bivariate relation regardless of any correlation coefficient measures. As can be seen, the plots demonstrate that, compared to TBSIM (Fig. 10), not only is the reproduction of bivariate relations in TBCOSIM improved but also the bivariate relations are roughly in agreement with those obtained from correlation coefficient measures (Fig. 9, Table 4).

Another way to examine the quality of the simulation results is to check the reproduction of spatial continuity in the region, implying whether the local cross-correlation is reproduced well. Computation of direct and cross-variograms of the simulated Gaussian random fields is an option for this purpose. In this respect, spatial continuity measures of simulated realizations should converge, on average, to the fitted theoretical model, obtained either from linear modeling of co-regionalization or from any other fitting procedures (Emery 2004). If this variogram curve, on average, is not in agreement with the theoretical model, it indicates that the simulation algorithm was not appropriate for the studied dataset and leads to produce biased results. In this study, the results of spatial continuity analysis show that although the direct variograms are reproduced well by both TBCOSIM and TBSIM algorithms, the cross-variograms reproduction were not satisfactory for the TBSIM outputs (Fig. 11). The reason why the TBSIM results were biased is that the cross-dependency among the variables was not considered in this algorithm and one should not expect to reproduce the desired cross-correlation among the variables as provided in Table 5 for the lag separation 0.

### Univariate Statistical Parameters Reproduction

It is instructive to examine the reproduction of the means and variances of the simulated realizations by TBCOSIM so as to compare with the original de-clustered means and variances of the



◀ **Figure 7.** Sample (crosses) and modeled (solid lines) direct variograms of Gaussian variables. For brevity, only the direct variograms are displayed, although the fitting has been carried out with all the direct and cross-variograms.

studied variables. While the mean is the acceptable measure for comparison of realizations, and even if the same central value is obtained as the de-clustered mean, the simulation results may show totally different distributions. Thus, the variance as additional measure of variation is required for better interpretation of the simulated data dispersion. Consequently, the mean and variance were calculated per set of realizations. The means and variances of 100 conditional simulated realizations are presented in Table 4. The reproduced means are slightly lower, varying from 0.60 to 11.33% depending on the variable, compared to the de-clustered means, and thus are quite satisfactory for all the six variables. Likewise, the reproduced variances are slightly lower, varying from 5.66 to 25.5% depending on the variable, compared to the de-clustered variances, and thus are quite satisfactory. The slightly lower means and variances can be interpreted as due to the influence of conditioning data (sample values). The simulation results were also checked thoroughly per variable, before back-transformation, to examine whether the realizations follow a normal standard distribution.

### Validation Against Actual Data

The simulated results can also be examined against the actual values of the studied variables. For instance, more samples may be available from some locations as actual data in further steps of mineral exploration that those areas have already been simulated beforehand and some values are available as prediction data. However, in this study, the true values are not available. Therefore, in order to test the proficiency of the co-simulation results, a typical validation technique, which is suitable for validation of geostatistical algorithms, is taken into account for the six studied variables. This test, the so-called cross-validation technique, involves removing one sample at a time and is re-estimated from the remaining neighborhood conditioning data. Each datum is then replaced in the data set once it has been re-estimated (Deutsch and Journel 1998). This type of analysis is basically applicable for evaluation of (a) variogram

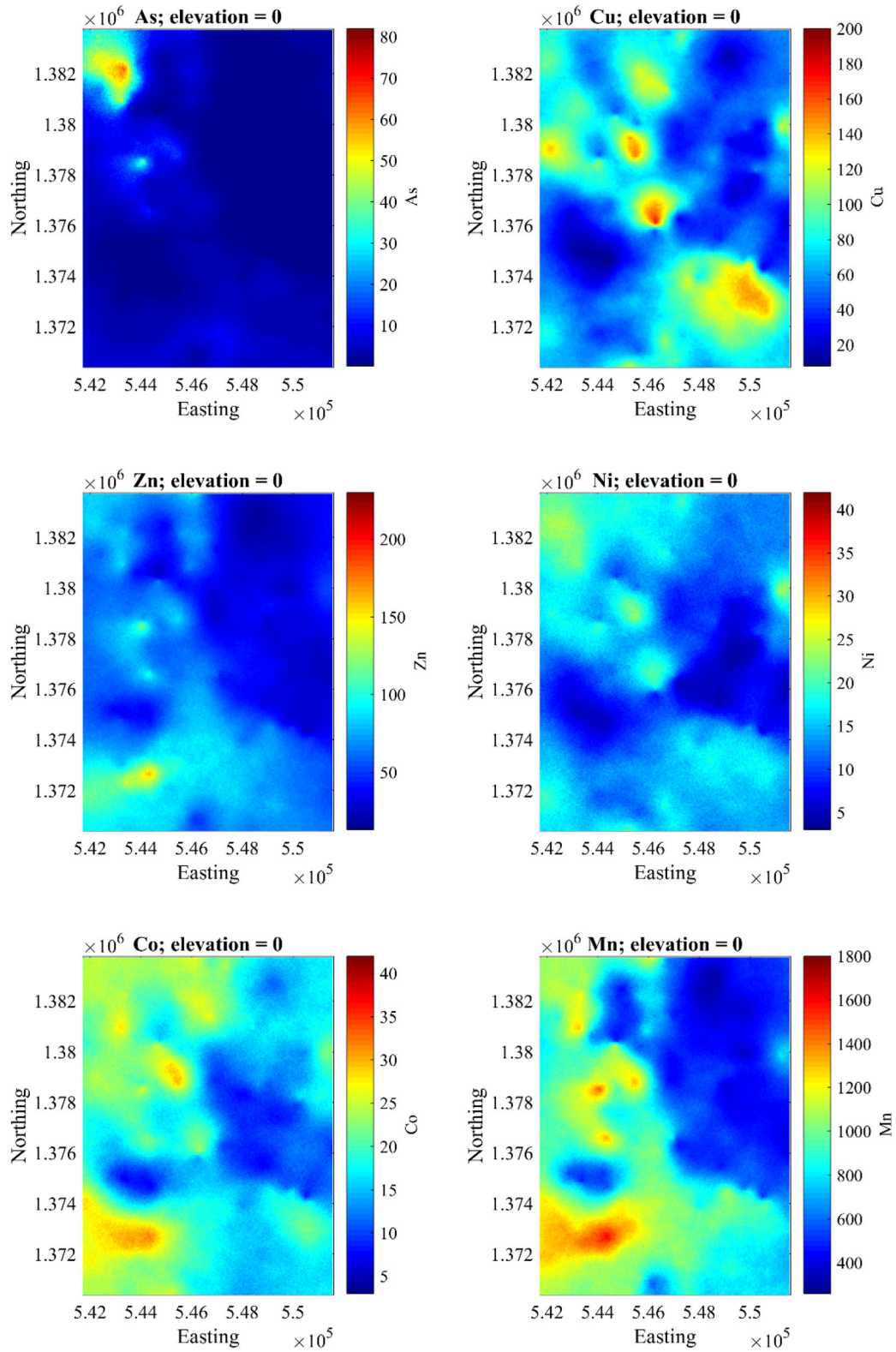
models, (b) type of kriging and (c) the search strategy in geostatistical contexts. Nevertheless, the cross-validation technique is also applicable for evaluation of (co)-simulation results. To do so, TBCOSIM was applied following the procedures that were thoroughly discussed in the previous sections. Because the “true” values of each variable were measured at sample locations, the averages of simulation results at the same locations conditioned to the remaining data provide the prediction values per variable at the corresponding locations. Therefore, a scatter plot can be drawn though the predicted and measured values at sample points. The results are depicted in Figure 12, and the bivariate distributions of plots and their tendencies to the diagonal line (black line) reveal that TBCOSIM provides satisfactory outputs without loss of generality. This also corroborates that this methodology (TBCOSIM) is precise enough in the case of cross-correlated variables whenever the actual data are available.

### Co-simulated Size Number (CoSS-N)

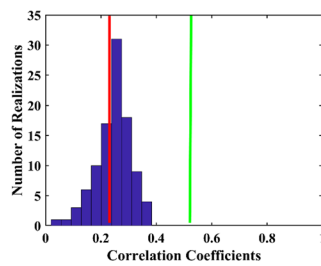
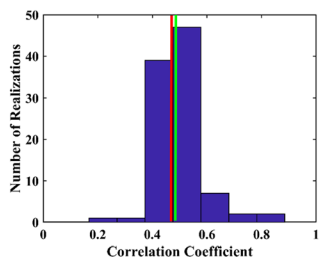
As already discussed, the results of Gaussian co-simulation turn out more reliable outputs due to taking into account the cross-correlation structure among the variables of interest. This concept cannot be corroborated, however, in the case of dealing with independent simulation algorithm for modeling those types of variables with substantial inter-dependency relationship. In this section, it is of interest to capture different thresholds through fractal analysis of TBCOSIM simulated realizations (100 outputs) over the underlying N-S log-log plots. To do so, separate curves were first obtained from each individual realization, and then, the average of all log-log curves was calculated for each variable. Next, the segment lines have been fitted through this average curve graph of realizations and the thresholds were inferred as shown in Figure 11. The corresponding thresholds derived from fractal analysis in Figure 13 are also presented in Table 6.

### PROBABILISTIC DESCRIPTION OF ANOMALOUS AREA

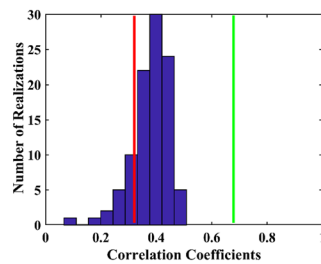
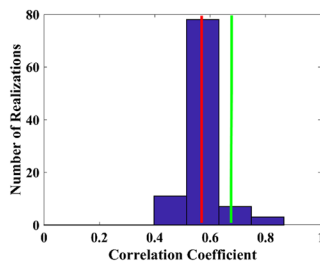
Once the thresholds are derived from fractal analysis on the corresponding realizations obtained from TBCOSIM, one application pertains to the production of maps showing anomalous areas for



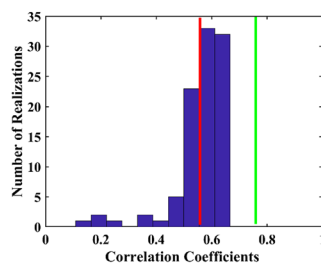
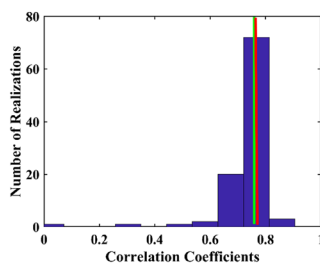
**Figure 8.** E-type maps obtained from 100 realizations of each variable. All values of variables are in ppm.



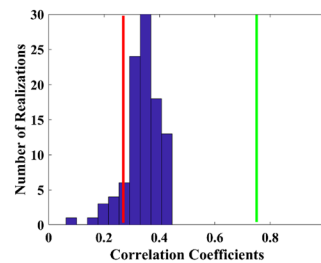
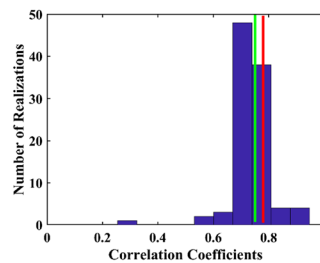
Zn-Ni



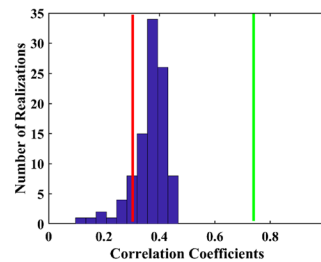
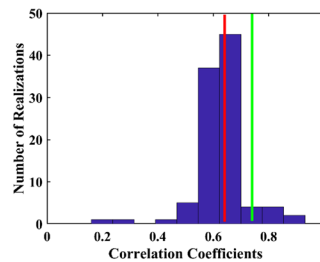
Co-Zn



Mn-Zn



Co-Ni



Mn-Co

◀ **Figure 9.** Correlation coefficients between transformed simulated variables with turning bands co-simulation (left) and TBSIM (right). For brevity, only the correlation between the pairs of Zn–Ni, Co–Zn, Mn–Zn, Co–Ni and Mn–Co is illustrated. The average correlation over the realizations is represented with a red line, while the true de-clustered correlation is represented with a green line.

**Table 4.** Correlation coefficients among the original variables, reproduced by independent and Gaussian co-simulation

	As	Cu	Zn	Ni	Co	Mn
<i>Original de-clustered data</i>						
As	1	0.082	0.19	0.384	0.28	0.312
Cu	0.082	1	0.298	0.423	0.288	0.285
Zn	0.19	0.298	1	0.519	0.656	0.718
Ni	0.384	0.423	<b>0.519</b>	1	0.712	0.401
Co	0.28	0.288	<b>0.656</b>	<b>0.712</b>	1	0.704
Mn	0.312	0.285	<b>0.718</b>	0.401	<b>0.704</b>	1
<i>TBSIM</i>						
As	1	0.126	0.210	0.287	0.231	0.243
Cu	0.126	1	0.259	0.260	0.169	0.200
Zn	0.210	0.259	1	0.244	0.378	0.562
Ni	0.287	0.260	<b>0.244</b>	1	0.339	0.190
Co	0.231	0.169	<b>0.378</b>	<b>0.339</b>	1	0.365
Mn	0.243	0.200	<b>0.562</b>	0.192	<b>0.365</b>	1
<i>TBCOSIM</i>						
As	1	0.197	0.249	0.335	0.256	0.295
Cu	0.197	1	0.359	0.387	0.265	0.338
Zn	0.249	0.359	1	0.488	0.570	0.722
Ni	0.335	0.387	<b>0.488</b>	1	0.735	0.400
Co	0.256	0.265	<b>0.570</b>	<b>0.730</b>	1	0.626
Mn	0.295	0.338	<b>0.722</b>	0.400	<b>0.626</b>	1

Bold values indicate that the correlation among variables are strong

further mineral exploration. This step is crucial, because it makes the investigation more up-scaled and narrower. Carranza (2010a, b) applied principal component analysis (PCA) to explain as much information contained in the data as possible in a few components as possible for Arroyo dataset. PCA searches for the direction in the multivariate space that contains the maximum variability. The two first perpendicular principle components (PCA1) and (PCA2) convey the maximum amount of the remaining data variability (Reimann et al. 2008). PCA can also be seen as a decomposition of the covariance matrix or correlation matrix into its “eigenvectors” and “eigenvalues” that are the loadings of the principle components spanning the new PCA coordinate system. In his study, Carranza (2010a, b) implemented the PCA analysis on the same case study dataset, indicating that PCA1 accounts for 34% of the total variance and represents a

Co–Zn–Mn association, for which it shows possible metal scavenging by Mn oxides in the drainage environments in most part of area, while the PCA2 accounting for 22% of total variance introduces a Ni–Cu–As association with possibly enrichment of epithermal Au deposit in the area due to the weathering and erosion of anomalous sources containing sulfide minerals such as chalcopyrite and arsenopyrite. This association also reflects dacitic/andesitic rocks that host the epithermal Au deposits in the area. In this study, we integrated the two mentioned associations (Co–Zn–Mn and Ni–Cu–As) to produce the probability maps, offering an alternative tool to show the uncertainty of finding probable anomalous areas in the region. In the previous section, the individual maps for each variable obtained through averaging the realizations were presented (Fig. 8). These maps can be integrated to produce the probability maps in the region, based on the corresponding thresholds. For instance, it is of interest to define the probable areas in one unique map instead of three different maps that jointly illustrate the anomalous areas for association of Ni–Cu–As taking into account the thresholds that identify the anomalous area for each individual variable. Based on Table 6, the initiation thresholds for Ni, Cu and As are 19.95 ppm, 39.8 ppm and 2.51 ppm, respectively. Therefore, the idea is to find the probability of common areas based on these variables and, thus, to prepare one map summarizing two or three variables. This is of paramount importance for delineation of areas with possible locations of epithermal Au deposits described by these Ni–Cu–As and Co–Zn–Mn associations. To do so, this concept can be formulated based on probability rules (e.g., Hogg et al. 2012; Olea et al. 2016):

- For two variables,  $A$  and  $B$ :

$$\text{Prob}(A \cap B) = \text{Prob}(A|B) \times \text{Prob}(B) \quad (15)$$

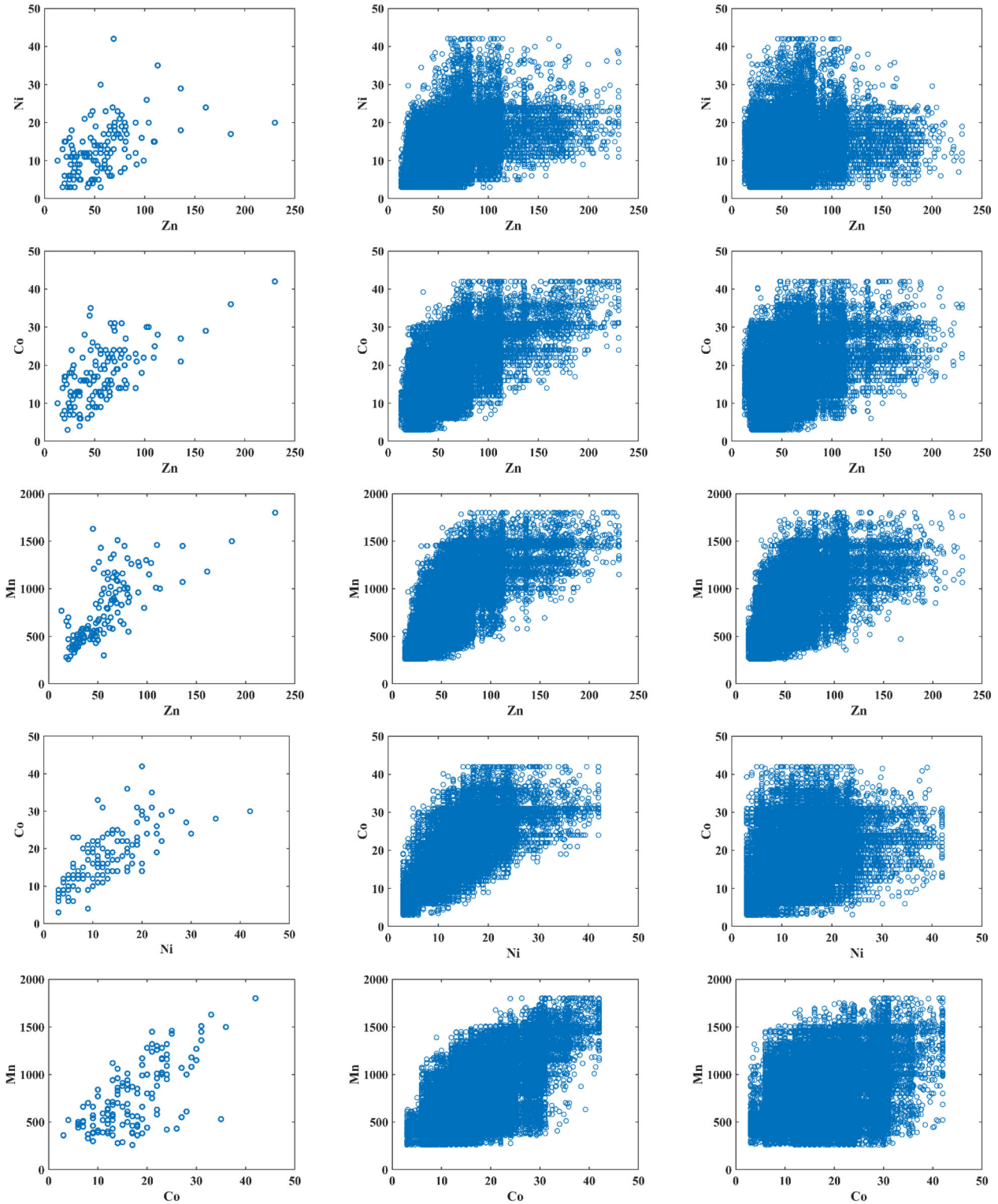
For three variables,  $A$ ,  $B$  and  $C$ :

$$\text{Prob}(A \cap B \cap C) = \text{Prob}(A) \times \text{Prob}(B|A) \times \text{Prob}(C|A \cap B) \quad (16)$$

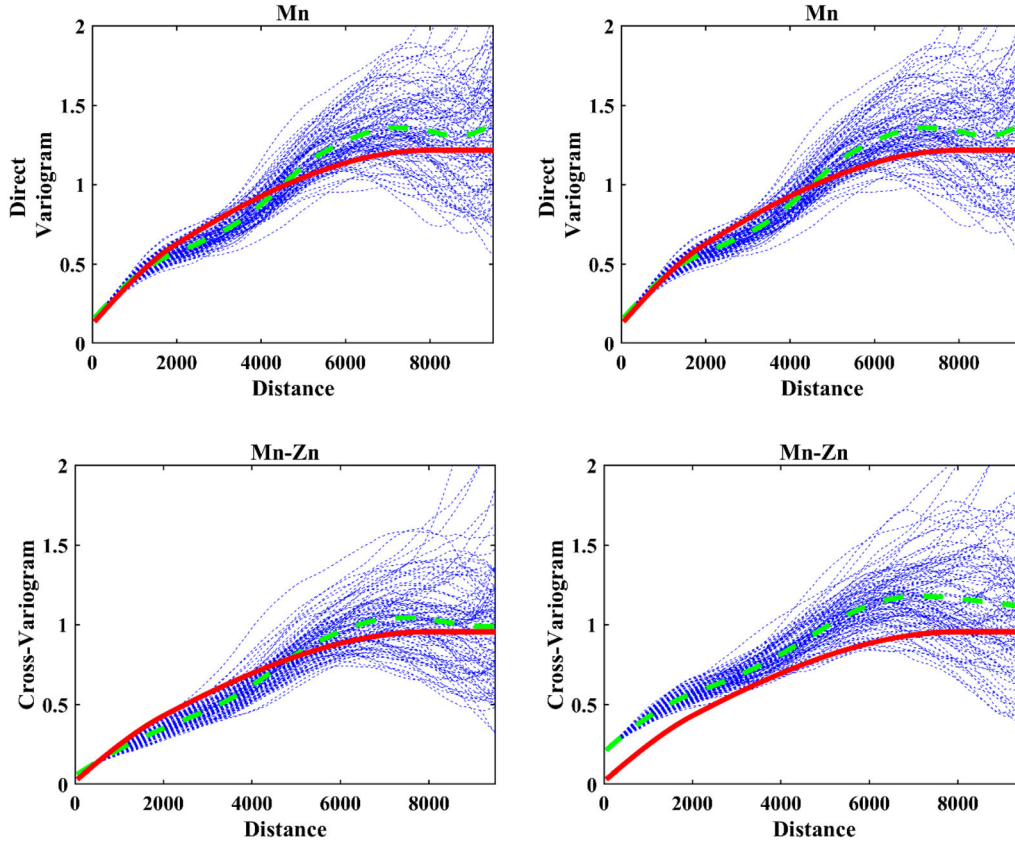
Therefore:

- For likely epithermal Au deposit locations defined by Cu and As variables:

$$\begin{aligned} \text{Prob}(\text{Cu} \geq 39.8 \cap \text{As} \geq 2.51) \\ = \text{Prob}(\text{Cu} \geq 39.8 \cap \text{As} \geq 2.51) \\ \times \text{Prob}(\text{As} \geq 2.51) \end{aligned} \quad (17)$$



**Figure 10.** Bivariate relation between variable pairs of Zn–Ni, Co–Zn, Mn–Zn, Co–Ni and Mn–Co. Left panel: original variables. Middle panel: TBCOSIM. Right panel: TBSIM. For brevity, only a few pairs are illustrated for only realization number 1. TBCOSIM results outperform the TBSIM results, because of the former following the original bivariate shape of correlation.



**Figure 11.** Direct and cross-variograms of Gaussian random fields simulated with turning bands (left panel) and turning bands co-simulation (right panel). For brevity, only the direct variogram for Mn and only cross-variograms for Mn–Zn are displayed. Blue dashed lines: variograms of individual realizations; green dashed line: average variogram of realizations; red solid line: theoretical variogram model obtained from linear model of co-regionalization.

**Table 5.** Global means and variances of de-clustered data and 100 realizations by TBCOSIM

Parameters	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
	As		Cu		Zn		Ni		Co		Mn	
De-clustered	5.56	149.06	64.90	1358.89	60.23	1200.43	13.36	53.80	17.97	55.14	799.70	118,464.71
Average TBCO-SIM	4.93	118.30	66.28	1268.40	58.28	894.36	13.19	45.34	17.84	52.02	794.92	108,290.05

- For likely dacitic/andesitic rocks that host the epithermal Au deposits defined by the Ni–Cu–As association:

$$\begin{aligned}
 & \text{Prob}(\text{Ni} \geq 19.95 \cap \text{Cu} \geq 39.8 \cap \text{As} \geq 2.51) \\
 &= \text{Prob}(\text{Ni} \geq 19.95) \\
 &\quad \times \text{Prob}(\text{Cu} \geq 39.8 | \text{Ni} \geq 19.95) \\
 &\quad \times \text{Prob}(\text{As} \geq 2.51 | \text{Ni} \geq 19.95 \cap \text{Cu} \geq 39.8)
 \end{aligned}
 \tag{18}$$



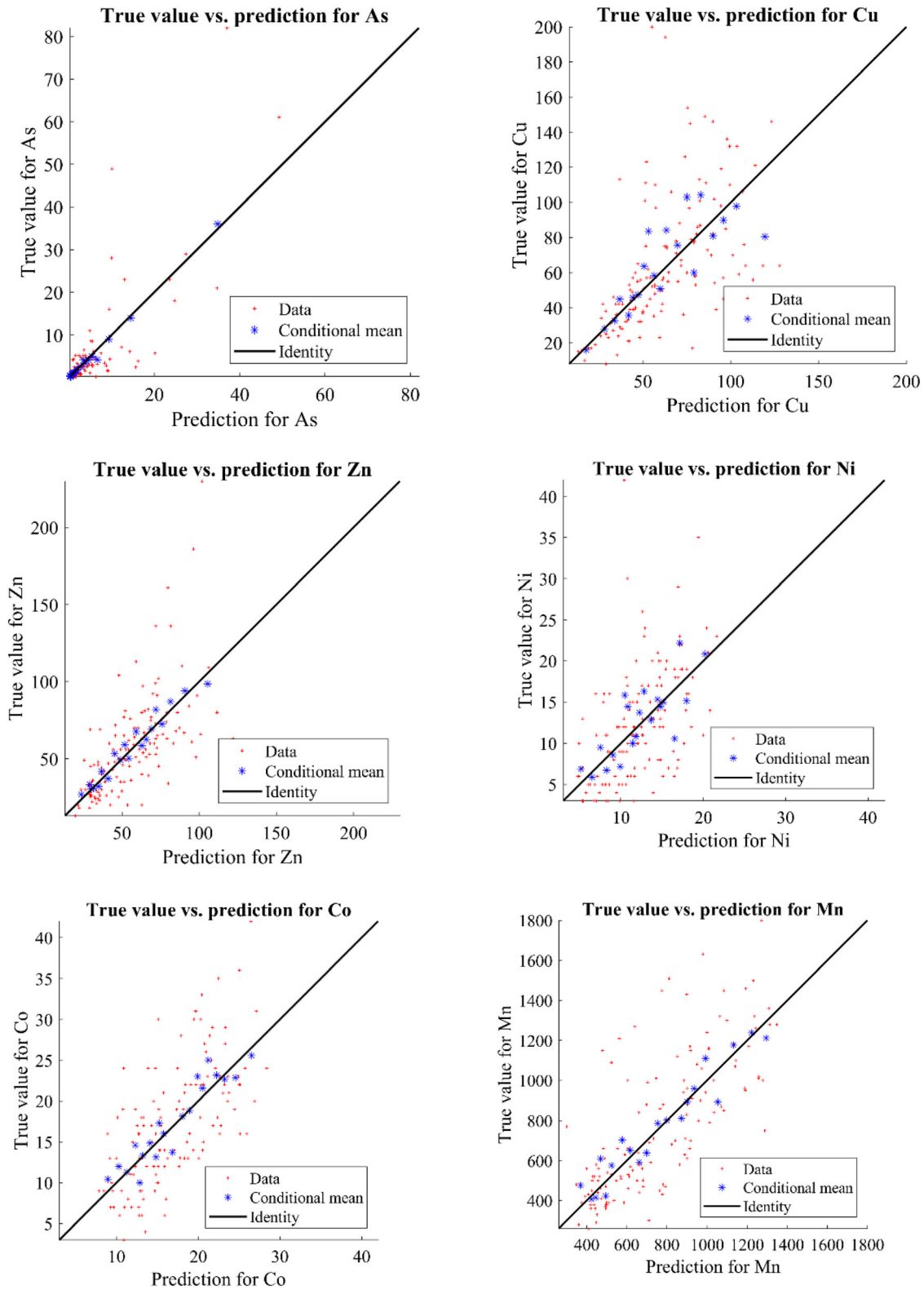


Figure 12. Scatter plots between measured and predicted values at sample locations.

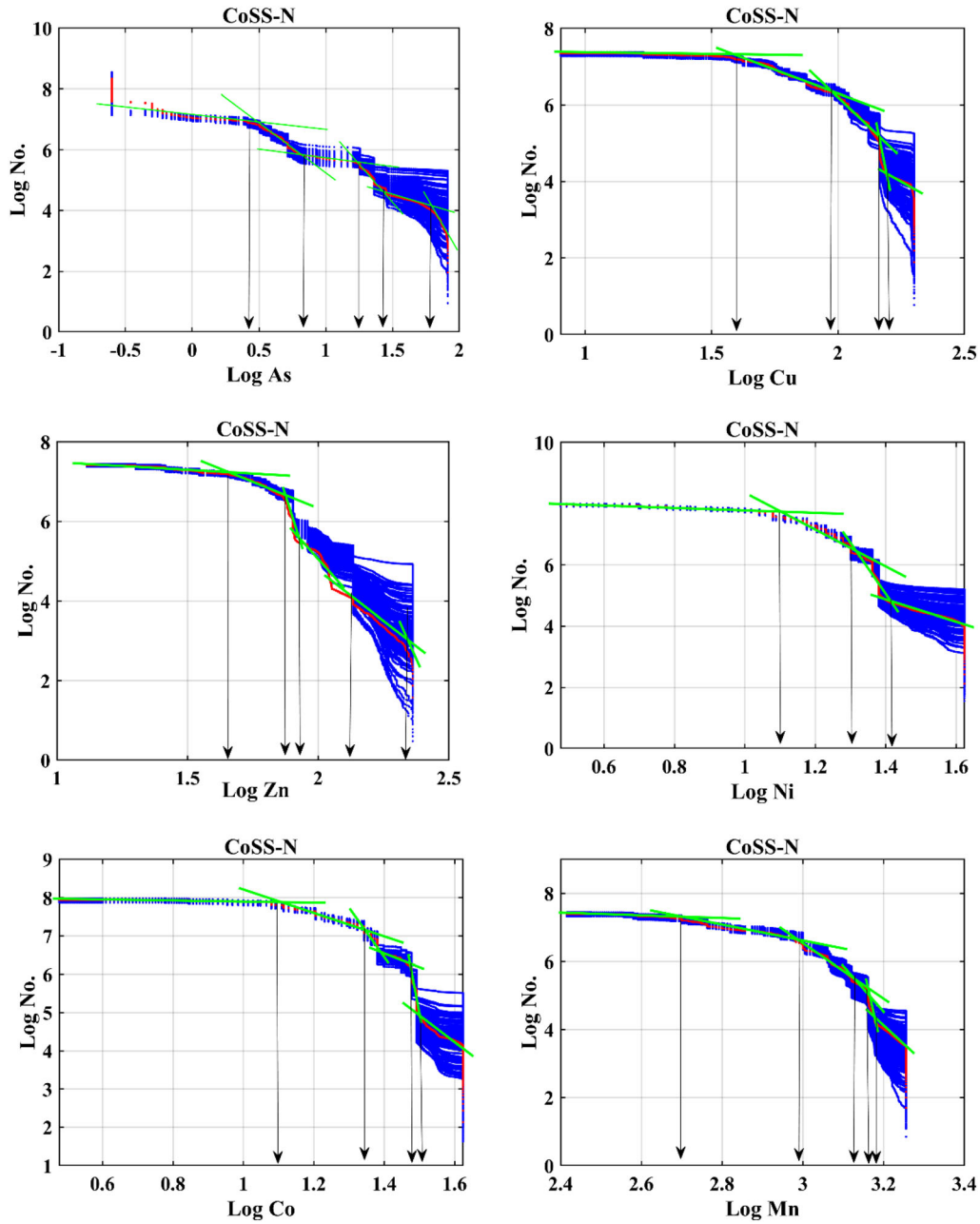


Figure 13. Results of fractal method and the thresholds identified for the CoSS-N model.

- For likely metal scavenging by Mn oxides in the drainage environments defined by the Co–Zn–Mn association:

$$\begin{aligned}
 & \text{Prob}(\text{Co} \geq 30.19 \cap \text{Zn} \geq 44.66 \cap \text{Mn} \geq 1318.3) \\
 &= \text{Prob}(\text{Co} \geq 30.19) \times \text{Prob}(\text{Zn} \geq 44.66 | \text{Co} \geq 30.19) \\
 & \quad \times \text{Prob}(\text{Mn} \geq 1318.3 | \text{Co} \geq 30.19 \cap \text{Zn} \geq 44.66)
 \end{aligned}
 \tag{19}$$

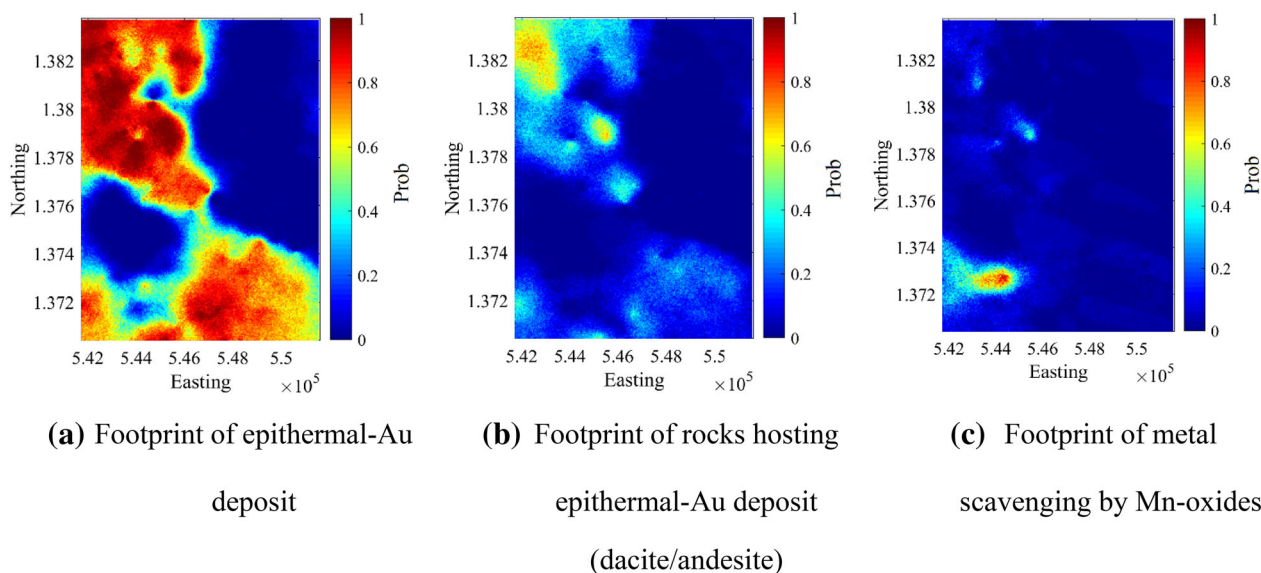
**Table 6.** Different geochemical populations obtained from the proposed CoSS-N approach

Geochemical variable	Population	
As	Background	0.25–2.51
	Very low anomaly	2.51–6.30
	Very low anomaly	6.30–17.78
	Low anomaly	17.78–25.11
	High anomaly	25.11–63.09
Cu	Very high anomaly	> 63.09
	Low background	8–39.8
	High background	39.8–89.12
	Low anomaly	39.8–158.48
	High anomaly	158.48–177.82
Zn	Vary high anomaly	> 177.82
	Background	13–44.66
	Very low anomaly	44.66–63.09
	Low anomaly	63.09–79.43
	High anomaly	79.43–125.89
Ni	Very high anomaly	> 125.89
	Low background	3–12.58
	High background	12.58–19.95
	Low anomaly	19.95–26.91
	High anomaly	> 26.91
Co	Very low background	3–14.12
	Background	14.12–22.38
	High background	22.38–30.19
	Low anomaly	30.19–31.62
	High anomaly	> 31.62
Mn	Very low background	260–489.77
	Low background	489.77–794.32
	High background	794.32–1318.3
	Low anomaly	1318.3–1513.6
	High anomaly	1513.6–1548.8

The probability maps are shown in Figure 14. They are constructed based on Eqs. 15, 16, 17, 18 and 19 for each block through 100 conditional TBCOSIM realizations and show the chance of finding plausible mineralized zones based on the identified thresholds. The sectors with little uncertainty are those associated with a high probability for a given anomaly (portrayed in red in Fig. 13), indicating that there is little risk of not finding an anomaly, or those associated with a very low probability (portrayed in dark blue in Fig. 13), indicating that one is pretty sure of not finding an anomaly, while the other sectors (portrayed in light blue, green, or yellow in Fig. 13) are more uncertain. The footprint of plausible anomalies for enrichment of Cu and As elements (Fig. 14a) shows that the likely locations where epithermal Au deposits can be found trend NW–SE, which is in agreement with the trend of their host rocks (Fig. 14b).

## CONCLUSION

Identification of geochemical anomalies is important for mineral exploration. However, deriving the corresponding elemental thresholds from fractal/multi-fractal analysis is challenging whenever the geochemical variables exhibit dependency among each other and they have the inherent compositional characteristics. In this study, an innovative workflow is proposed for tackling these issues based on a combination of geostatistical Gaussian co-simulation and fractal analysis, namely co-simulated size number (CoSS-N). In this context, the geochemical components should be transformed from original scale to new variables by one of the log-ratio transformation techniques, such as additive log-ratio transformation (alr), after inference of statistical parameters in both original and log-ratio scales, respectively. In order to quantify spatial continuity, direct and cross-variogram can be calculated over alr-transformed variables and linear model of co-regionalization should be fitted respecting semi-definite positive condition. Once the variogram structure is derived, one of the Gaussian co-simulation algorithms, such as turning bands co-simulation (TBCOSIM), can be implemented. The log-ratio simulated results should then be back-transformed to original scale and fractal analysis can then be applied for calculation of thresholds for differentiating of background from anomaly. The proposed approach was performed in a real case study from the Philippines where seven cross-correlated variables were measured at 135 sample points from the surficial environment. In order to validate the proposed algorithm, CoSS-N results were compared with the results of SS-N (Sadeghi et al. 2015), whereby global and local statistical parameters were examined on the simulation results obtained from both approaches. The overall accuracy of the CoSS-N corroborates that it outperforms the SS-N model in the case of dealing with cross-correlated geochemical components. The derived thresholds were then applied to calculate the uncertainty and plausible areas where to find epithermal Au deposits in the region. The probabilistic produced maps showed a NW–SE trend of locations where possible epithermal Au deposit can be found in the region. These results have not been recognized from the geometric average of back-transformed simulation results. In future, this study can be extended based on identification of thresholds by fractal analysis of cross-correlated variables, wherever the underlying



**Figure 14.** Probability maps showing plausible anomalous areas (in warm colors) derived from 100 TBCOSIM realizations.

geochemical components are ascertained by the lithology type in the region.

## ACKNOWLEDGMENTS

The first author acknowledges the Nazarbayev University for supporting this work through Faculty Development Competitive Research Grants for 2018–2020 under Contract No. 090118FD5336.

## REFERENCES

- Abildin, Y., Madani, N., & Topal, E. (2019). A hybrid approach for joint simulation of geometallurgical variables with inequality constraint. *Minerals*, 9(1), 24.
- Adeli, A., Emery, X., & Dowd, P. (2018). Geological modelling and validation of geological interpretations via simulation and classification of quantitative covariates. *Minerals*, 8(1), 7.
- Afzal, P., Fadakar Alghalandis, Y., Khakzad, A., Moarefvand, P., & Rashidnejad Omran, N. (2011). Delineation of mineralization zones in porphyry Cu deposits by fractal concentration–volume modeling. *Journal of Geochemical Exploration*, 108, 220–232.
- Afzal, P., Khakzad, A., Moarefvand, P., Rashidnejad Omran, N., Esfandiari, B., & Fadakar Alghalandis, Y. (2010). Geochemical anomaly separation by multifractal modeling in Kahang (Gor Gor) porphyry system. Central Iran. *Journal of Geochemical Exploration*, 104, 34–46.
- Afzal, P., Madani, N., Shahbeik, Sh., & Yasrebi, A. B. (2015). Multi-Gaussian kriging: a practice to enhance delineation of mineralized zones by concentration–volume fractal model in Dardevey iron ore deposit, SE Iran. *Journal of Geochemical Exploration*, 158, 10–21.
- Agterberg, F. P. (1995). Multifractal modeling of the sizes and grades of giant and supergiant deposits. *International Geology Review*, 37(1), 1–8.
- Aitchison, J. (1982). The statistical analysis of compositional data. *Journal of the Royal Statistical Society: Series B (Methodological)*, 44(2), 139–177. **(with discussion)**.
- Aitchison, J. (1986). *The statistical analysis of compositional data. Monographs on statistics and applied probability*. London: Chapman & Hall Ltd. **(Reprinted in 2003 with additional material by The Blackburn Press)**.
- Aitchison, J., & Shen, S. M. (1980). Logistic-normal distributions. Some properties and uses. *Biometrika*, 67(2), 261–272.
- Almeida, A. S., & Journel, A. G. (1994). Joint simulation of multiple variables with a Markov-type coregionalization model. *Mathematical Geology*, 26(5), 565–588.
- Battalgaży, N., & Madani, N. (2019). Categorization of mineral resources based on different geostatistical simulation algorithms: A case study from an iron ore deposit. *Natural Resources Research*. <https://doi.org/10.1007/s11053-019-09474-9>.
- Buccianti, A., & Pawlowsky-Glahn, V. (2005). New perspectives on water chemistry and compositional data analysis. *Mathematical Geology*, 37(7), 703–727.
- Carr, J. R., & Myers, D. E. (1985). COSIM: A FORTRAN IV program for coconditional simulation. *Computers & Geosciences*, 11(6), 675–705.
- Carranza, E. J. M. (2008). *Geochemical anomaly and mineral prospectivity mapping in GIS. Handbook of exploration and environmental geochemistry* (Vol. 11). Amsterdam: Elsevier.
- Carranza, E. J. M. (2009). Controls on mineral deposit occurrence inferred from analysis of their spatial pattern and spatial association with geological features. *Ore Geology Reviews*, 25, 383–400.
- Carranza, E. J. M. (2010a). Catchment basin modelling of stream sediment anomalies revisited: incorporation of EDA and fractal analysis. *Geochemistry: Exploration, Environment, Analysis*, 10, 365–381.

- Carranza, E. J. M. (2010b). Mapping of anomalies in continuous and discrete fields of stream sediment geochemical landscapes. *Geochemistry: Exploration, Environment, Analysis*, 10, 171–187.
- Carranza, E. J. M. (2011). Analysis and mapping of geochemical anomalies using logratio-transformed stream sediment data with censored values. *Journal of Geochemical Exploration*, 110, 167–185.
- Chen, G., & Cheng, Q. (2018). Fractal-based wavelet filter for separating geophysical or geochemical anomalies from background. *Mathematical Geosciences*, 50(3), 249–272.
- Cheng, Q. (2012). Singularity theory and methods for mapping geochemical anomalies caused by buried sources and for predicting undiscovered mineral deposits in covered areas. *Journal of Geochemical Exploration*, 122, 55–70.
- Cheng, Q., Agterberg, F. P., & Ballantyne, S. B. (1994). The separation of geochemical anomalies from background by fractal methods. *Journal of Geochemical Exploration*, 51, 109–130.
- Cheng, Q., Agterberg, F. P., & Bonham-Carter, G. F. (1996). A spatial analysis method for geochemical anomaly separation. *Journal of Geochemical Exploration*, 56(3), 183–195.
- Cheng, Q., Xu, Y., & Grunsky, E. (1999). Integrated spatial and spectral analysis for geochemical anomaly separation. In: S. J. Lippard, A. Naess, & R. Sinding-Larsen (Eds.), *Proceedings of the conference of the international association for mathematical geology* (Vol. 1, pp. 87–92). Trondheim.
- Cheng, Q., Xu, Y., & Grunsky, E. C. (2000). Integrated spatial and spectrum method for geochemical anomaly separation. *Natural Resources Research*, 9, 43–52.
- Chilès, J. P., & Delfiner, P. (2012). *Geostatistics: Modeling spatial uncertainty* (2nd ed.). New York: Wiley.
- Deutsch, C. V., & Journel, A. G. (1998). *GSLIB: Geostatistical software library and user's guide*. New York: Oxford University Press.
- Egozcue, J., Pawłowsky-Glahn, V., Mateu-Figueras, G., & Barceló-Vidal, B. (2003). Isometric logratio transformations for compositional data analysis. *Mathematical Geology*, 35(3), 279–300.
- Emery, X. (2004). Testing the correctness of the sequential algorithm for simulating Gaussian random fields. *Stochastic Environmental Research and Risk Assessment A*, 18(6), 401–413.
- Emery, X. (2005). Variograms of order  $\omega$ : A tool to validate a bivariate distribution model. *Mathematical Geology*, 37(2), 163–181.
- Emery, X. (2007). Conditioning simulations of Gaussian random fields by ordinary kriging. *Mathematical Geology*, 39(6), 607–623.
- Emery, X. (2008). A turning bands program for conditional co-simulation of cross-correlated Gaussian random fields. *Computers & Geosciences*, 34(12), 1850–1862.
- Emery, X. (2010). Iterative algorithms for fitting a linear model of coregionalization. *Computers & Geoscience*, 36(9), 1150–1160.
- Emery, X., & Lantuéjoul, C. (2006). TBSIM: A computer program for conditional simulation of three-dimensional Gaussian random fields via the turning bands method. *Computer & Geoscience*, 32(10), 1615–1628.
- Eze, P. N., Madani, N., & Adoko, A. C. (2019). Multivariate mapping of heavy metals spatial contamination in a Cu–Ni exploration field (Botswana) using turning bands co-simulation algorithm. *Natural Resources Research*, 28(1), 109–124.
- Feder, J. (1988). *Fractals* (p. 283). New York: Plenum Press.
- Filzmoser, P., Hron, K., & Reimann, C. (2009). Principal components analysis for compositional data with outliers. *Environmetrics*, 20, 621–632.
- Gneiting, T. (1999). The correlation bias for two-dimensional simulations by turning bands. *Mathematical Geology*, 31(2), 195–211.
- Gómez-Hernández, J. J., & Journel, A. G. (1993). Joint sequential simulation of multigaussian fields. In A. Soares (Ed.), *Geostatistics Tróia'92* (Vol. 1, pp. 85–94). Dordrecht: Springer.
- Goovaerts, P. (1997). *Geostatistics for natural resources evaluation*. New York: Oxford University Press.
- Goovaerts, P. (1999). Impact of the simulation algorithm, magnitude of ergodic fluctuations and number of realizations on the spaces of uncertainty of flow properties. *Stochastic Environmental Research and Risk Assessment*, 13(3), 161–182.
- He, J., Yao, S., Zhang, Z., & You, G. (2013). Complexity and productivity differentiation models of metallogenic indicator elements in rocks and supergene media around Daijiazhuang Pb–Zn deposit in Dangchang County, Gansu Province. *Natural Resources Research*, 22, 19–36.
- Hogg, R. V., McKean, J., & Craig, A. T. (2012). *Introduction to mathematical statistics* (7th ed.). Harlow: Pearson Education Ltd.
- Hosseini, S. A., & Asghari, O. (2019). Multivariate geostatistical simulation on block-support in the presence of complex multivariate relationships: Iron ore deposit case study. *Natural Resources Research*, 28, 125–144.
- JICA-MMAJ. (1986). *Report on mineral exploration, mineral deposits and tectonics of two contrasting geologic environments in the Republic of the Philippines, Phase II (Masbate and Leyte Areas)*. Tokyo: Japan International Cooperating Agency and Metal Mining Agency of Japan.
- Job, M. (2010). *Application of logratios for compositional data*. Centre for Computational Geostatistics Report 12, 307. University of Alberta, Canada.
- Journel, A. B., & Huijbregts, C. J. (1978). *Mining geostatistics*. New York: Academic.
- Lantuéjoul, C. (1994). Non conditional simulation of stationary isotropic multigaussian random functions. In M. Armstrong & P. A. Dowd (Eds.), *Geostatistical simulations* (pp. 147–177). Dordrecht: Kluwer.
- Lantuéjoul, C. (2002). *Geostatistical simulation: Models and algorithms*. Berlin: Springer.
- Leuangthong, O., McLennan, J. A., & Deutsch, C. V. (2004). Minimum acceptance criteria for geostatistical realizations. *Natural Resources Research*, 13, 131–141.
- Li, C., Xu, Y., & Jiang, X. (1994). The fractal model of mineral deposits. *Geology of Zhejiang*, 10, 25–32. **(in Chinese with English abstract)**.
- Luz, F., Mateus, A., Matos, J. X., & Gonçalves, M. A. (2014). Cu and Zn-soil anomalies in the NE border of the South Portuguese Zone (Iberian Variscides, Portugal) identified by multifractal and geostatistical analyses. *Natural Resources Research*, 23, 195–215.
- Madani, N., & Emery, X. (2019). A comparison of search strategies to design the cokriging neighborhood for predicting coregionalized variables. *Stochastic Environmental Research and Risk Assessment*, 33(1), 183–199.
- Madani, N., & Sadeghi, B. (2019). Capturing hidden geochemical anomalies in scarce data by fractal analysis and stochastic modeling. *Natural Resources Research*, 28(3), 833–847.
- Mandelbrot, B. B. (1983). *The fractal geometry of nature*. San Francisco: W. H. Freeman.
- Matheron, G. (1973). The intrinsic random functions and their applications. *Advances in Applied Probability*, 5(3), 439–468. <https://doi.org/10.2307/1425829>.
- Mery, N., Emery, X., Cáceres, A., Ribeiro, D., & Cunha, E. (2017). Geostatistical modeling of the geological uncertainty in an iron ore deposit. *Ore Geology Reviews*, 88, 336–351.

- Mitchell, A. H. G., & Leach, T. M. (1991). *Epithermal gold in the Philippines: Island arc metallogenesis. Geothermal Systems and Geology*. London: Academic Press.
- Monecke, T., Monecke, J., Herzig, P. M., Gemmel, J. B., & Monch, W. (2005). Truncated fractal frequency distribution of element abundance data: A dynamic model for the metasomatic enrichment of base and precious metals. *Earth Planetary Science Letters*, 232, 363–378.
- Mueller, U., Tolosana-Delgado, R., & van den Boogaart, K. G. (2014). Simulation of compositional data: A nickel laterite case study. In R. Dimitrakopoulos (Ed.), *Advances in ore-body modelling and strategic mine planning*. Melbourne: AusIMM.
- Myers, D. E. (1989). Vector conditional simulation. In M. Armstrong (Ed.), *Geostatistics* (pp. 283–293). Dordrecht: Springer.
- Olea, R. A., Luppens, J. A., Egozcue, J. J., & Pawlowsky-Glahn, V. (2016). Calorific value and compositional ultimate analysis with a case study of a Texas lignite. *International Journal of Coal Geology*, 162, 27–33.
- Paravarzar, S., Emery, X., & Madani, N. (2015). Comparing sequential Gaussian and turning bands algorithms for cosimulating grades in multi-element deposits. *Comptes Rendus Geoscience*, 347(2), 84–93.
- Pawlowsky-Glahn, V. (1984). On spurious spatial covariance between variables of constant sum. *Science de la Terre Information Géologie*, 21, 107–113.
- Pawlowsky, V., & Burger, H. (1992). Spatial structure analysis of regionalized compositions. *Mathematical Geology*, 24, 675–691.
- Pawlowsky-Glahn, V. (2003). Statistical modelling on coordinates. In S. Thió-Henestrosa & J. A. Martín-Fernández (Eds.), *Proceedings of CoDaWork'03, the 1st compositional data analysis workshop*. Girona (E): Universitat de Girona. ISBN 84-8458-111-X. Retrieved April 18, 2019 from <http://ima.udg.es/Activitats/CoDaWork2003/>.
- Pawlowsky-Glahn, V., & Egozcue, J. J. (2016). Spatial analysis of compositional data: A historical review. *Journal of Geochemical Exploration*, 164, 28–32.
- Pawlowsky-Glahn, V., Egozcue, J. J., & Tolosana-Delgado, R. (2015). *Modelling and analysis of compositional data*. New York: Wiley.
- Pawlowsky-Glahn, V., & Olea, R. A. (2004). *Geostatistical analysis of compositional data. Studies in mathematical geology* (Vol. 7). Oxford: Oxford University Press.
- Pebesma, E. J. (2004). Multivariable geostatistics in S: The gstat package. *Computers & Geosciences*, 30(7), 683–691.
- Reimann, C., Filzmoser, P., Garrett, R. G., & Dutter, R. (2008). *Statistical data analysis explained: Applied environmental statistics with R (No. Sirsi) i9780470985816*. Chichester: Wiley.
- Rivoirard, J. (1994). *Introduction to disjunctive kriging and non-linear geostatistics*. Oxford: Oxford University Press.
- Rossi, M. E., & Deutsch, C. V. (2014). *Mineral resource estimation*. Dordrecht: Springer.
- Rubio, R. H., Costa, J. F. C. L., & Bassani, M. A. A. (2016). A geostatistical framework for estimating compositional data avoiding bias in back-transformation. *Rem: Revista Escola de Minas*, 69(2), 219–226.
- Sadeghi, B., Madani, N., & Carranza, E. J. M. (2015). Combination of geostatistical simulation and fractal modeling for mineral resource classification. *Journal of Geochemical Exploration*, 149, 59–73.
- Sadeghi, B., Moarefvand, P., Afzal, P., Yasrebi, A. B., & Saiein, L. D. (2012). Application of fractal models to outline mineralized zones in the Zaghia iron ore deposit, Central Iran. *Journal of Geochemical Exploration*, 122, 9–19.
- Shi, J., & Wang, C. (1998). Fractal analysis of gold deposits in China: Implication for giant deposit exploration. *Earth Sciences Journal of China University of Geosciences*, 23, 616–618. (in Chinese with English abstract).
- Talebi, H., Mueller, U., & Tolosana-Delgado, R. (2019). Joint simulation of compositional and categorical data via direct sampling technique—Application to improve mineral resource confidence. *Computer & Geosciences*, 122, 87–102.
- Tolosana-Delgado, R. (2006). *Geostatistics for constrained variables: Positive data, compositions and probabilities. Application to environmental hazard monitoring*. Retrieved April 18, 2019 from [http://www.tesisenxarxa.net/TDX-0123106-122444/index\\_an.html](http://www.tesisenxarxa.net/TDX-0123106-122444/index_an.html).
- Tolosana-Delgado, R., Mueller, U., & van den Boogaart, K. G. (2018). Geostatistics for compositional data: An overview. *Mathematical Geosciences*. <https://doi.org/10.1007/s11004-018-9769-3>.
- Tolosana-Delgado, R., Otero, N., & Pawlowsky-Glahn, V. (2005). Some basic concepts of compositional geometry. *Mathematical Geology*, 37(7), 673–680.
- Tolosana-Delgado, R., & van den Boogaart, K. G. (2013). Joint consistent mapping of high-dimensional geochemical surveys. *Mathematical Geosciences*, 45, 983–1004.
- Turcotte, D. L. (1996). *Fractals and chaos in geophysics (2nd edn)* (pp. 81–99). Cambridge: Cambridge University Press.
- Van den Boogaart, K. G., Mueller, U., & Tolosana-Delgado, R. (2017). An affine equivariant multivariate normal score transform for compositional data. *Mathematical Geosciences*, 49, 231–251.
- Wackernagel, H. (2003). *Multivariate geostatistics: An introduction with applications*. Berlin: Springer.
- Zuo, R., Cheng, Q., & Xia, Q. (2009). Application of fractal models to characterization of vertical distribution of geochemical element concentration. *Journal of Geochemical Exploration*, 102, 37–43.
- Zuo, R., & Wang, J. (2016). Fractal/multifractal modeling of geochemical data: A review. *Journal of Geochemical Exploration*, 164, 33–41.