

Original Paper

# Multivariate Mapping of Heavy Metals Spatial Contamination in a Cu–Ni Exploration Field (Botswana) Using Turning Bands Co-simulation Algorithm

Peter N. Eze,<sup>1</sup> Nasser Madani,<sup>2,3</sup> and Amoussou Coffi Adoko<sup>2</sup>

Received 6 January 2018; accepted 6 April 2018

With a mining-driven economy, Botswana has experienced increased geochemical exploration of minerals around existing mining towns. The mining and smelting of copper and nickel around Selibe-Phikwe in the Central Province are capable of releasing heavy metals including Pb, Fe, Mn, Co, Ni and Cu into the soil environments, thereby exposing humans, plants and animals to health risks. In this study, turning bands co-simulation, a multivariate geostatistical algorithm, was presented as a tool for spatial uncertainty quantification and probability mapping of cross-correlated heavy metals (Co, Mn, Fe and Pb) risk assessment in a semiarid Cu–Ni exploration field of Botswana. A total of 1050 soil samples were collected across the field at a depth of ~ 10 cm in a grid sampling design. Rapid elemental concentration analysis was done using an Olympus Delta Sigma portable X-ray fluorescence device. Enrichment factor, geoaccumulation index and pollution load index were used to assess the potential risk of heavy metals contamination in soils. The partially heterotopic nature of the dataset and strong correlations among the heavy metals favors the use of co-simulation instead of independent simulation in the probability mapping of heavy metal risks in the study area. The strong correlation of Co and Mn to iron infers they are of lithogenic origin, unlike Pb which had weak correlation pointing to its source in the area being of anthropogenic source. Manganese, Co and Fe show low enrichment, whereas Pb had high enrichment suggesting possible lead pollution. We, however, recommend that speciation of Pb in the soils rather than total concentration should be ascertained to infer chances of possible bioaccumulation, and subsequent health risk to human by chronic exposure.

**KEY WORDS:** Probability mapping, Gaussian random field, Semiarid soils, Portable XRF device, Uncertainty quantification.

## INTRODUCTION

Quantification of uncertainty has been widely researched in environmental studies including assessment of soils potentially polluted by heavy metals (Liu et al. 2006; Xie et al. 2011; Sakizadeh et al. 2017). Essentially, co-simulation of cross-correlated variables is key to successful mapping of spatial uncertainty (Wackernagel 2003; Chilès and Delfiner 2012). The most straightforward algorithms

<sup>1</sup>Department of Earth and Environmental Science, Botswana International University of Science and Technology, Private Bag 16, Palapye, Botswana.

<sup>2</sup>Department of Mining Engineering, School of Mining and Geosciences, Nazarbayev University, Astana, Kazakhstan.

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

required for the realization of this goal are those that make the use of Gaussian random fields (Pardo-Igúzquiza and Chica-Olmo 1993, 1994; Chilès and Lantuéjoul 2005). Sequential Gaussian Simulation (SGS) (Tran 1994; Gómez-Hernández and Cassiraga 1994) and Turning Bands Simulation (TBSIM) (Matheron 1973; Mantoglou 1987; Emery and Lantuéjoul 2006) are the two geostatistical algorithms widely used in uncertainty quantification. The SGS relies on the sequential paradigm of simulating the 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 a major limitation relating to the difficulty in ascertaining moving neighborhood and reproducing the short-scale continuity (Lantuéjoul 1994).

To address the pointed shortcomings of SGS, it has been shown that the simulation results can be optimized by increasing the size of neighborhood ranges (Emery 2004). The TBSIM, the fundamental concept of which is based on the 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. The reasonable criticism, however, is that artifacts might occur in the case of a few turning lines. The idea therefore is to increase the number of turning lines to weaken this kind of stripping (Lantuéjoul 1994; Gneiting 1999; Emery and Lantuéjoul 2006). In order to co-simulate cross-correlated variables, sequential Gaussian co-simulation (Gómez-Hernández and Journel 1993; Pebesma 2004), and turning bands co-simulation (Emery 2008; Carr and Myers 1985; Myers 1989) offer much flexibility to construct the outcomes (scenarios) preserving the interrelationship between the variables as well as displaying the spatial variability. These two current algorithms were examined and the results showed that turning bands co-simulation outperforms sequential Gaussian co-simulation with regard to reproducing the cross-correlation calculated by spatial continuity and statistical parameters (Parvarzar et al. 2015).

Spatial uncertainty quantification of heavy metal concentrations in the soil environment is critical to our understanding of their possible pollution pathways, biogeochemical cycles (Cloquet et al. 2006) and for devising possible remediation procedures. Statistical (Hani and Pazira 2011) and geochemical methods including enrichment factor (Anderson and Kravitz 2010), sequential extractions (Zimmerman

and Weindorf 2010), geochemical relationships (Bourennane et al. 2010) and geochemical mapping (Reimann and de Caritat 2005), among other methods, have been used for the assessment of heavy metal contamination in soils. Since these methods are based on assumptions of reference values for natural contents, there is always considerable range and discrepancies in the assessed soil contamination results from single application of each and every one of the methods. Invariably, there is no universally recognized method of estimation, but only more or less site-specific approximations (Desaules 2012).

As an interface between the atmosphere and the lithosphere, soils act as a sink for suspended particulate matter. Heavy metals in near-surface soils can be heavily disturbed by anthropogenic activities including mining and smelting from local and regional industrial plants. The deposition of transported particulate matter by wind or water has made the study of the natural status of soil geochemistry very difficult. Elements in biotic and abiotic (soils) environments exist in various forms and show some form of spatial trends. Three criteria including relatively high concentrations, different multivariate relationships from elements of natural origins and spatial patterns related to contaminant sources can be considered in identifying contaminants if significant patterns can be attributed to anthropogenic activities (Zhang 2006).

With the rapid growth of computer technology and new statistical methods of analysis, such as geostatistics (Burgos et al. 2008), geographical information systems (GIS) are becoming one of the most important tools for studying environmental geochemistry problems (Zhang and Selinus 1998). Although the distribution, chemistry and elemental interaction of heavy metals in soils have been largely studied (Udeigwe et al. 2015; Jørgensen and Jensen 2009; Eze et al. 2010), there is still the need for site-specific study of elemental interactions in soils in view of the fact that soils under different climates and environmental conditions would behave differently. Moreover, site-specific information is a great asset for source identification, apportionment of pollutants and remediation planning. The Maibele Airstrip North field, where this study took place, has never been investigated for heavy metal accumulation, which makes it an interesting topic to study. In this paper, focus is made on: (i) the application of turning bands co-simulation model to quantify the spatial uncertainty of four cross-correlated heavy metal elements (Mn, Fe, Co and Pb); (ii) comparing

## Multivariate Mapping of Heavy Metals Spatial Contamination

the output of turning bands co-simulation with the results of independent turning bands simulation of each of these elements by some selected statistical parameters. For comprehensive mathematical details on turning bands simulation (TBSIM) and turning bands co-simulation (TBCOSIM), readers are referred to Emery and Lantuéjoul (2006) and Emery (2008), respectively.

### THE DATA SET

The study area, Maibele Airstrip North (628200 and 631100 Easting; 759120 and 759680 Northing), is located in the east central part of Botswana and covers about 2.64 km<sup>2</sup> (Fig. 1).

It is about 40 km away from the Selebe-Phikwe Ni-Cu mine smelter/concentrator plant. Maibele Airstrip North is found in the greenstone belts and shear zones of the Zimbabwe Craton—a part of the Archean basement that forms the core of the southern Africa subcontinent and it stretches across northeastern Botswana and Zimbabwe. Elevations within the study area are typical of the basement system of Botswana lying at an altitude between 780 and 950 m above sea level with slopes in the study area generally being below 15%. One soil type (Haplic Luvisols) is identified on the available 1:1 000,000 FAO soil map of the area. The soils developed on paragneisses and amphibolites rich ultramafic rocks.



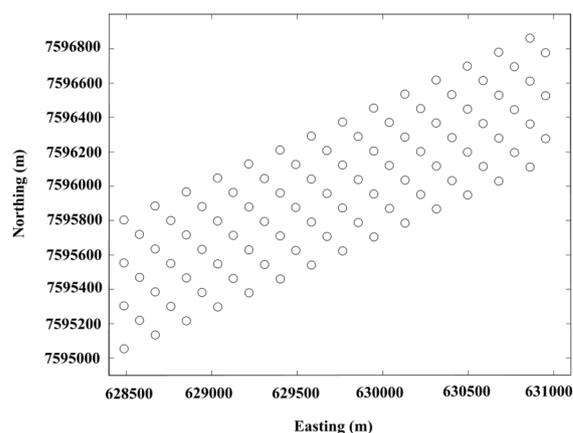
**Figure 1.** Geographical map of the study area (Maibele Airstrip is indicated by filled red star).

The climate is typically semiarid with a mean annual precipitation of 600 mm. It is characterized by a hot and rainy season with subtropical thunderstorms (November–March); a cool dry season (April–August); and a hot season with no occurrences of rain (September–November). Prevailing in the area are the south-easterly quadrant winds. A greater portion of the land is sparsely vegetated with grasses, shrubs and trees including *Aristida conjesta*, *Grewia flava*, and *Colophospermum mopane* and hard veld vegetation (Likuku et al. 2013).

The dataset used in this study to characterize spatial concentrations of heavy metals was obtained during a soil sampling campaign in search of geochemical anomaly in November 2015 to expand Botswana Metal Limited (BML)'s Cu-Ni mining operations in Maibele Air Strip North. A total of 1050 surface soil samples were collected, and analytical results of Mn, Fe, Co and Pb were recorded in terms of total concentration (mg kg<sup>-1</sup>). The sampling followed a grid scheme, by dividing the studied area into a grid of 100 m × 100 m. The location map of sample points is illustrated in Figure 2. Details of the sampling design and the analytical procedures using portable XRF are published in a repository (Eze et al. 2016a, b). The manufacturer's limits of detection (LOD) for the heavy metals are: < 10 ppm for Co, Fe, and Mn; and < 5 ppm for Pb.

### Soil Pollution Assessment

Three indices including enrichment factor (EF), geoaccumulation index (I<sub>geo</sub>) and pollution load index (PLI) were used to assess the pollution



**Figure 2.** Location map of sample points; black circles are sample locations.

thresholds of the heavy metals in the near-surface soils. The background values of Fe, 128,840; Mn, 1150; Co 21; and Pb, 2 (mg kg<sup>-1</sup>) used in the pollution assessment indices were obtained from the compilation of the major and trace element geochemistry of the regional plutonic rocks underlying the study area by Kampunzu et al. 2003.

*Enrichment Factor (EF)*

Iron was used as a normalizing element in the calculation of EF. Fe was used because of its relatively more abundance in Earth's crust and it functions as a conservative tracer to differentiate between natural and anthropogenic heavy metal components in soil samples. Thus, EF was first proposed by Buat-Menard and Chesselet (1979):

$$EF = \frac{\left[\frac{C_x}{C_{Fe}}\right]_{\text{Sample}}}{\left[\frac{B_x}{B_{Fe}}\right]_{\text{Background}}} \quad (1)$$

where  $C_x/C_{Fe}$  and  $B_x/B_{Fe}$  represent the ratios of heavy metal of interest to Fe in the sample over that of the background sample. Five EF thresholds were used to weight pollution level (Addo et al. 2012); EF < 2 denoted minimal enrichment, EF from 2 to 5 for moderate enrichment, EF from 5 to 20 for severe enrichment, EF from 20 to 40 very high enrichment, and EF > 40 extremely high enrichment.

*Geoaccumulation Index (Igeo)*

Geoaccumulation index (Igeo) was proposed by Müller (1969):

$$I_{geo} = \log_2 \left[ \frac{C_n}{1.5B_n} \right] \quad (2)$$

where  $C_n$  = measured concentration of the element in the soil sample,  $B_n$  = geochemical background value of the element of interest, and the constant 1.5 minimizes the effect of possible variations in the background values which may be attributed to lithologic variations. The Igeo thresholds ranges were: Igeo ≤ 0 (unpolluted), 0 < Igeo < 1 (unpolluted to moderately polluted), 1 < Igeo < 2 (moderately polluted), 2 < Igeo < 3 (moderately to strongly polluted), 3 < Igeo < 4 (strongly polluted),

4 < Igeo < 5 (strongly to extremely polluted) and Igeo > 5 (extremely polluted), respectively.

*Pollution Load Index (PLI)*

The PLI is first computed from the contamination factor (CF), which is the ratio between the heavy metals in the soils and in the background sample (Tomlinson et al. 1980)

$$CF = \frac{C_m \text{ sample}}{C_m \text{ background}} \quad (3)$$

The PLI is then computed from Eq. 4, where n is the total number of heavy metals assessed.

$$PLI = (CF_1 \times CF_2 \times CF_3 \times \dots \times CF_n)^{\frac{1}{n}} \quad (4)$$

The interpretations of PLI were given by: PLI < 1 denotes perfection, PLI = 1 only baseline levels of pollutants is present, and PLI > 1 polluted soil.

**METHODOLOGY**

Turning bands conditional co-simulation algorithms were carried out following the steps in the flowchart (Fig. 3).

**Exploratory Data Analysis**

Exploratory data analysis was done to find possible errors, outliers and statistical parameters. This can be considered by univariate or multivariate

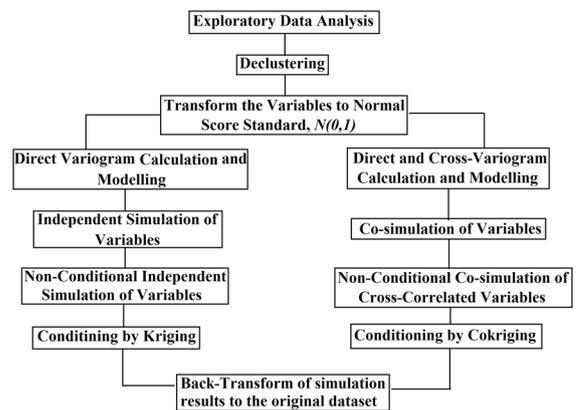


Figure 3. Flowchart for implementing TBSIM and TBCOSIM.

## Multivariate Mapping of Heavy Metals Spatial Contamination

**Table 1.** Statistical description after declustering the heavy metal concentrations.

Statistical parameters	Mn	Fe	Co	Pb
Number of data	1153	1204	1152	1142
Mean value	647.483	43,928.524	26.749	26.876
Variance	59,996.710	191,752,907.954	88.700	635.160
Minimum	168.000	14,431.000	6.000	9.000
Lower quartile	473.771	35,040.402	20.000	17.000
Median	677.790	45,362.147	27.537	21.000
Upper quartile	824.000	53,681.148	33.000	27.000
Maximum	1219.000	75,675.000	50.000	397.000

**Table 2.** Correlations between the declustered heavy metal concentrations.

	Mn	Fe	Co	Pb
Mn	1	–	–	–
Fe	0.939	1	–	–
Co	0.883	0.936	1	–
Pb	0.218	0.218	0.197	1

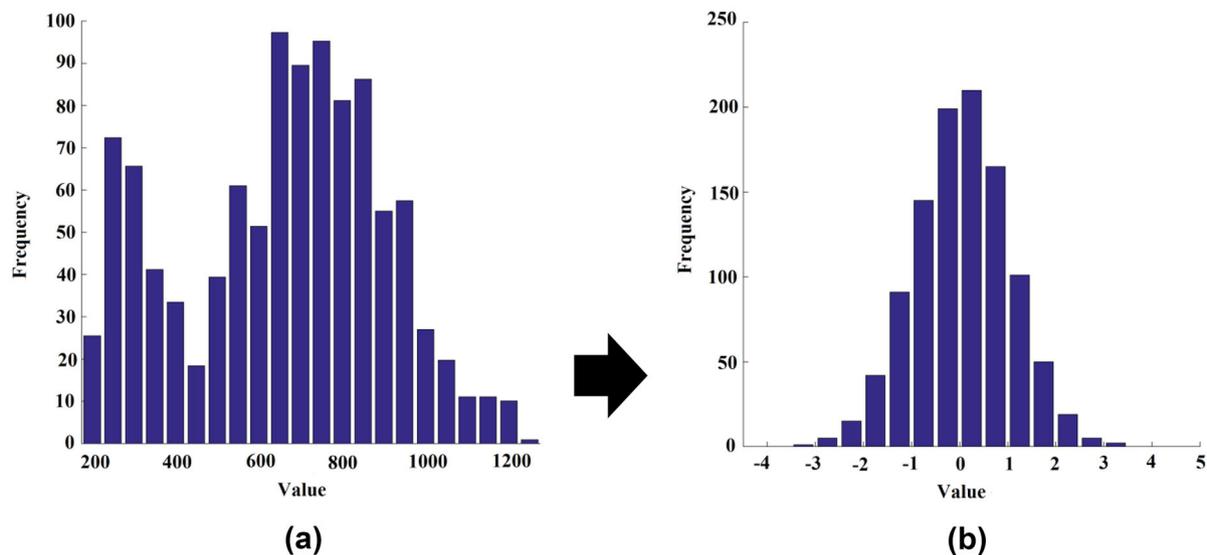
tools. The latter shows to what extent the variables are related to each other, which can be measured by correlation coefficient and drawing the scatter plot between the co-regionalized variables.

### Declustering

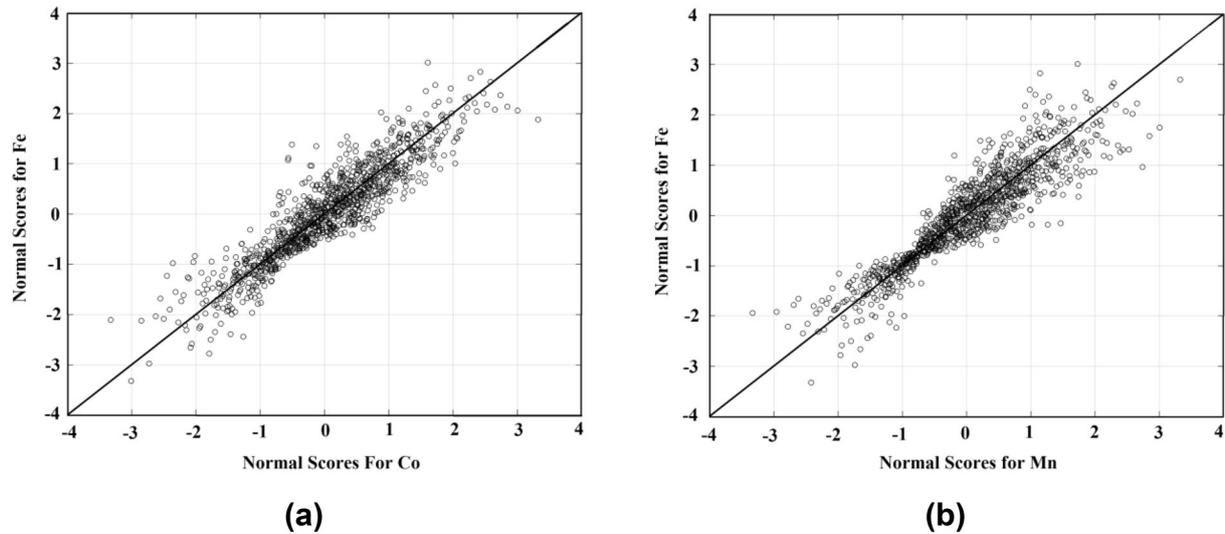
The scarcity of data in some parts of a study area makes the sampling pattern irregular and sta-

tistical parameters possibly biased. One idea to remedy this is to account for the weights of each location by cell-declustering technique to correct the pseudo-skewness in the global distribution of the dataset (Goovaerts 1997; Deutsch and Journel 1998). Yet, this technique is suitable for the univariate case and avoids the cross-correlation among the variables. Following (Bourgault 1997; Bogaert 1999), the co-kriging approach can be applied for co-regionalized variables with simple and cross-variograms to assign consistent weights to each location. Table 1 summarizes the statistical description of metal concentrations after declustering of the dataset.

As can be seen in Table 1, the variables are not measured at all the sampling locations. This pattern is so-called partially heterotopic, in which some variables are in common with some sample locations



**Figure 4.** Normal score transformation. (a) Original declustered histogram of Mn, (b) normal score transformed of Mn.



**Figure 5.** Checking the multivariate Gaussian distribution. (a) Relationship between Co & Fe, (b) relationship between Mn & Fe.

(Wackernagel 2003). The partially heterotopic nature of the heavy metals dataset led to the use of co-simulation in the place of independent simulation in addition to the favorable correlations that exist among the variables (Table 2). In this case, the secondary variable (auxiliary) that is available at more locations than the primary variable improves the estimation or simulation for the primary variable (Emery 2008).

### Transform the Data

As TBSIM and TBCOSIM algorithms like other geostatistical simulation techniques are dependent on the simulation of Gaussian random fields, the data should be transformed to normal standard variables to have a Gaussian distribution with mean 0 and variance 1 accounting for declustering weights. The Gaussian anamorphosis is an applicable function for such a transformation (Rivoirard 1994; Chilès and Delfiner 2012). Such a transformation to normal score for Mn, for example, is shown in Figure 4.

### Checking the Multivariate Gaussianity

The presence of interesting positive correlation coefficients among the variables (heavy metals) (Table 2) and their univariate transformation

(Fig. 4) to Gaussian random field does not ensure that the multivariate distributions are also Gaussian (Leuangthong and Deutsch 2003) which is a critical assumption for implementing TBCOSIM. One important specification is to examine the multivariate Gaussianity by checking the homoscedasticity and linearity among the cross-correlated variables (Johnson and Wichern 1998). As an example, Fig. 5 illustrates the scatter plot between three elements (Mn, Fe and Co), and one can see that the bivariate character is somehow in agreement with homoscedasticity and linearity definitions. Therefore, the Gaussian co-simulation approach can be applied here.

### Checking the Bivariate Gaussianity

Gaussian simulation requires the multivariate normality as explained in the previous section. A graphical tool is to look at the interrelationship between two points that are located at a specific distance from each other and checking that the points are distributed according to an elliptic isodensity shape (Rivoirard 1994). As can be seen in Fig. 6, the nearly elliptic shape in distribution of points is somehow corroborated. However, this methodology is somewhat tedious, firstly because it needs examining this visual tool for alternative distances, secondly because detection of elliptical shape is difficult in case of large distances (Emery 2005).

# Multivariate Mapping of Heavy Metals Spatial Contamination

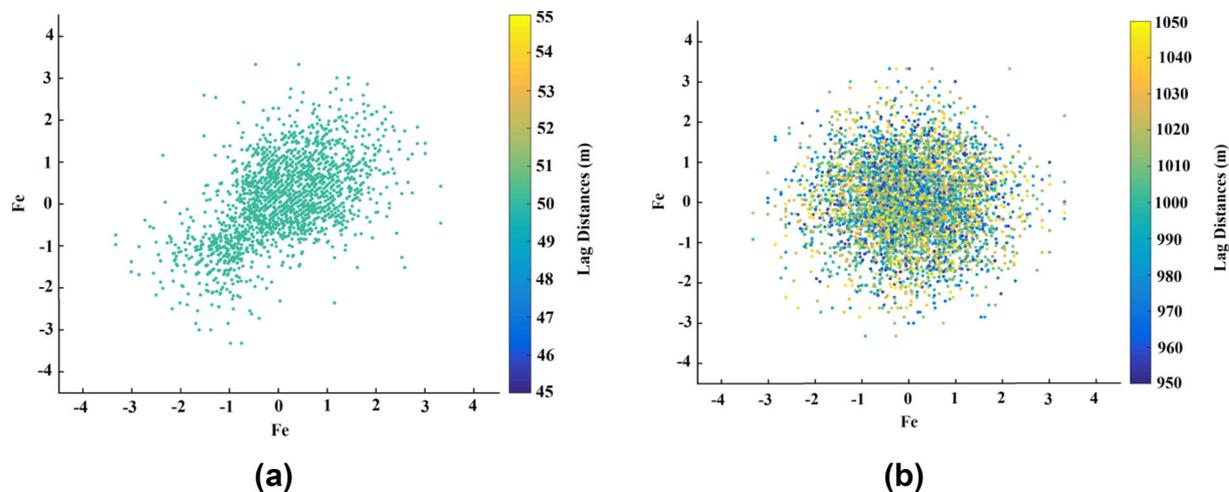


Figure 6. Bivariate Gaussian distribution for Fe. (a) Distance: 50 m (tolerance: 25 m), (b) distance: 1000 m (tolerance: 1000 m).

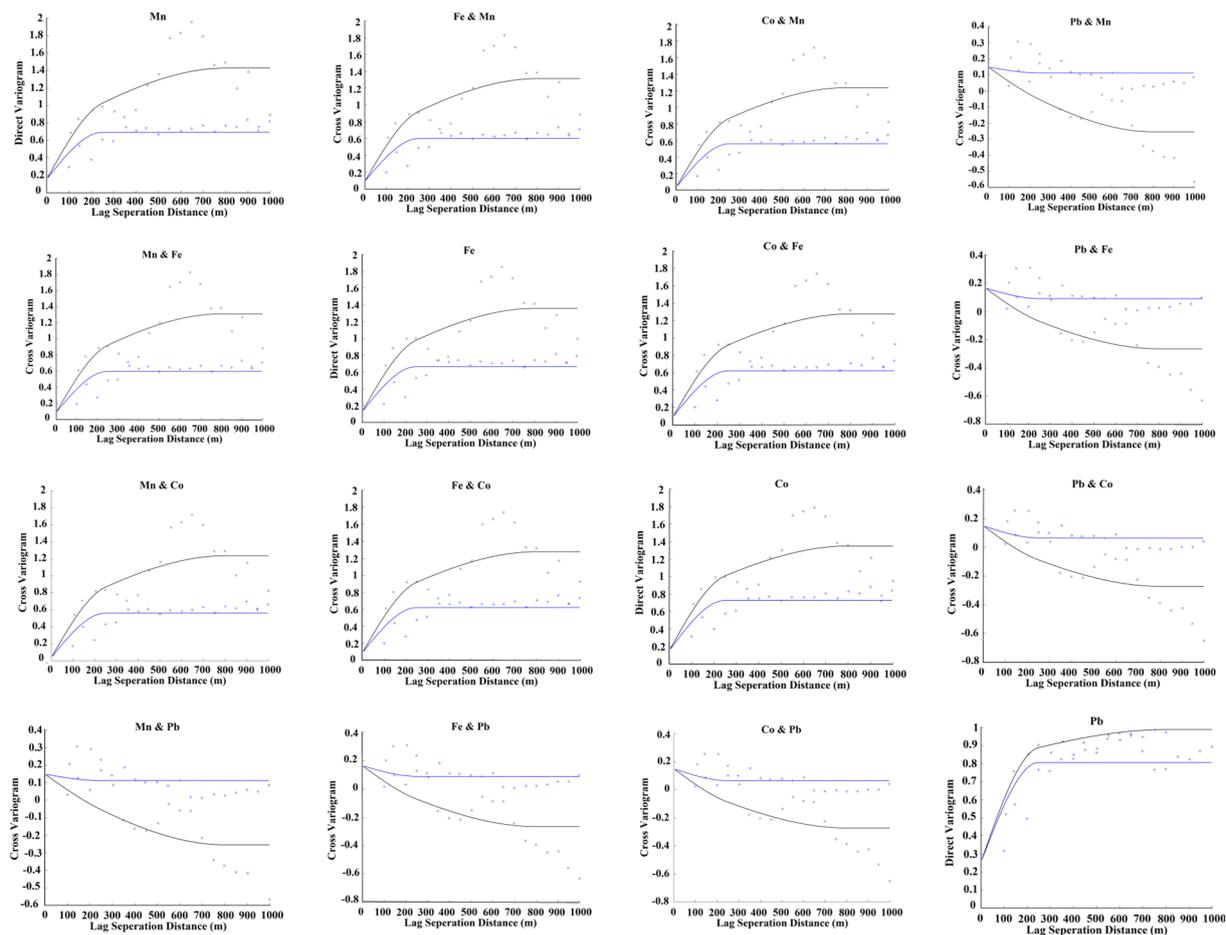


Figure 7. Direct and cross-variograms for four underlying variables (Mn, Fe, Co and Pb). Blue: azimuth 135° & black: azimuth 45°.

### Calculation of Spatial Continuity

Direct and cross-variograms were calculated over these four normal scored heavy metals. First of all, the experimental direct variograms are calculated along different directions to detect possible anisotropy. Different measures of range in variogram indicated that the maximum axis of anisotropy is in azimuth of 135 degrees, and consequently the minimum anisotropy is extended along the azimuth of 45 degrees. Therefore, in total, 16 experimental variograms (direct and cross) were computed along these anisotropy directions. However, TBSIM requires just the direct variograms, while TBCOSIM needs direct and cross-variograms as well. The most demanding procedure beyond this step is to infer the linear model of co-regionalization (Wackernagel 2003; Emery 2008; Chilès and Delfiner 2012). Henceforth, a semiautomatic fitting function was applied to fit two spherical nested structures and a nugget effect (Fig. 7) according to the following formulae:

$$\begin{aligned}
 & \begin{pmatrix} \gamma_{\text{Mn}}(h) & \gamma_{\text{Mn-Fe}}(h) & \gamma_{\text{Mn-Co}}(h) & \gamma_{\text{Mn-Pb}}(h) \\ \gamma_{\text{Fe-Mn}}(h) & \gamma_{\text{Fe}}(h) & \gamma_{\text{Fe-Co}}(h) & \gamma_{\text{Fe-Pb}}(h) \\ \gamma_{\text{Co-Mn}}(h) & \gamma_{\text{Co-Fe}}(h) & \gamma_{\text{Co}}(h) & \gamma_{\text{Co-Pb}}(h) \\ \gamma_{\text{Pb-Mn}}(h) & \gamma_{\text{Pb-Fe}}(h) & \gamma_{\text{Pb-Co}}(h) & \gamma_{\text{Pb-Pb}}(h) \end{pmatrix} \\
 = & \begin{pmatrix} 0.098 & 0.043 & -0.005 & 0.081 \\ 0.043 & 0.071 & 0.036 & 0.089 \\ -0.005 & 0.036 & 0.096 & 0.076 \\ 0.081 & 0.088 & 0.076 & 0.140 \end{pmatrix} \text{nugget} \\
 + & \begin{pmatrix} 0.609 & 0.569 & 0.587 & 0.007 \\ 0.569 & 0.614 & 0.603 & -0.014 \\ 0.577 & 0.603 & 0.649 & -0.02 \\ 0.007 & -0.014 & -0.028 & -0.752 \end{pmatrix} \text{Sph}(250\text{m}, 250\text{m}) \\
 + & \begin{pmatrix} 0.866 & 0.820 & 0.800 & -0.239 \\ 0.820 & 0.782 & 0.756 & -0.227 \\ 0.800 & 0.756 & 0.741 & -0.220 \\ -0.239 & -0.227 & -0.220 & -0.066 \end{pmatrix} \text{Sph}(800\text{m}, \infty)
 \end{aligned} \tag{5}$$

In the above equation, the first and second ranges show the anisotropy along the directions of azimuth 135° and 45°, respectively.

### Independent Simulation and Co-simulation of Heavy Metals

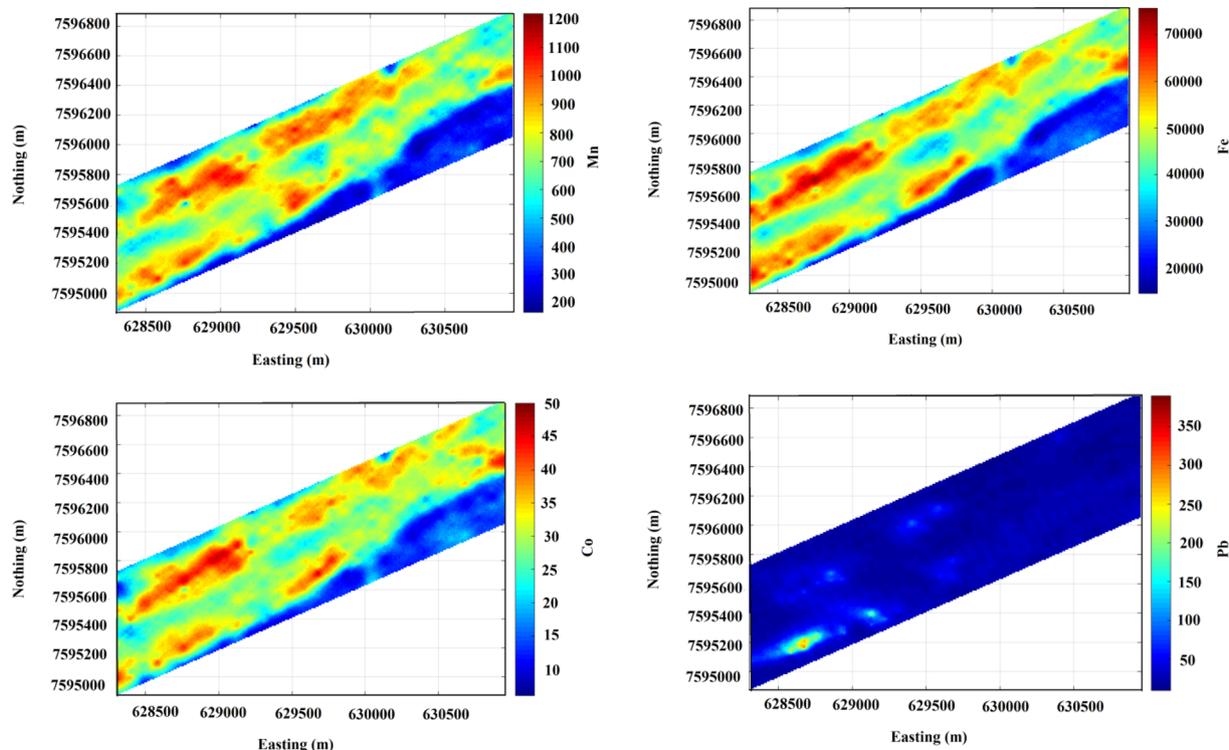
Simulation was performed on a regular grid with dimension of 10 m × 10 m × 10 m, and ordinary kriging and co-kriging were utilized for the

process of conditioning to the hard data (sampling locations), respectively. The proposed approaches can be substituted for simple kriging and co-kriging where the uncertainty is significant in mean value of the random field (Emery 2007, 2012). The neighborhood is moving with conditioning to 20 surrounding data characterized by maximum and minimum anisotropy equal 1000 and 800 m, respectively, derived from the variogram analysis. The number of lines for turning bands should be large as much as possible (Emery 2008). Henceforth, it was set to 1000 lines for elimination of stripping effects. The fitted linear model of co-regionalization was incorporated to both algorithms, which consists of the characteristics of the direct and cross-variograms (applicable only for co-simulation) models (Fig. 7). The number of realizations was considered to be 100 for both cases. E-type maps were produced by averaging the 100 realizations of the co-simulated elements in each block. As can be seen, the desired correlation coefficient considered beforehand is reproduced very well and manifest themselves in the maps (Fig. 8).

### Checking the Validity of the Results

The purpose of this section is to verify whether or not the desired correlation coefficients are treated well between the back-transformed variables elements through the realizations. It is expected that these coefficients fluctuate around the experimental correlations (Table 2). This condition stands for TBCOSIM applying ordinary co-kriging, but not for TBSIM considering ordinary kriging (Fig. 9). The latter case shows a very significant bias due to the low correlations, on average, reproduced by the realizations. This is due to the fact that, compared to TBCOSIM, ordinary kriging applied in TBSIM does not consider linear correlation among the elements and consequently leads to a poor regeneration of the interdependence among the elements to be simulated. Figure 9 comprises results of methodologies for two elements (Mn and Fe) showing their effect on the reproduction of correlation coefficients. To check the validity of either model, a cross-validation technique was employed whereby it co-simulates the underlying variables at every data location, conditionally to the information available at the other locations to compare the actual values with the average of the simulated values that can be consid-

## Multivariate Mapping of Heavy Metals Spatial Contamination

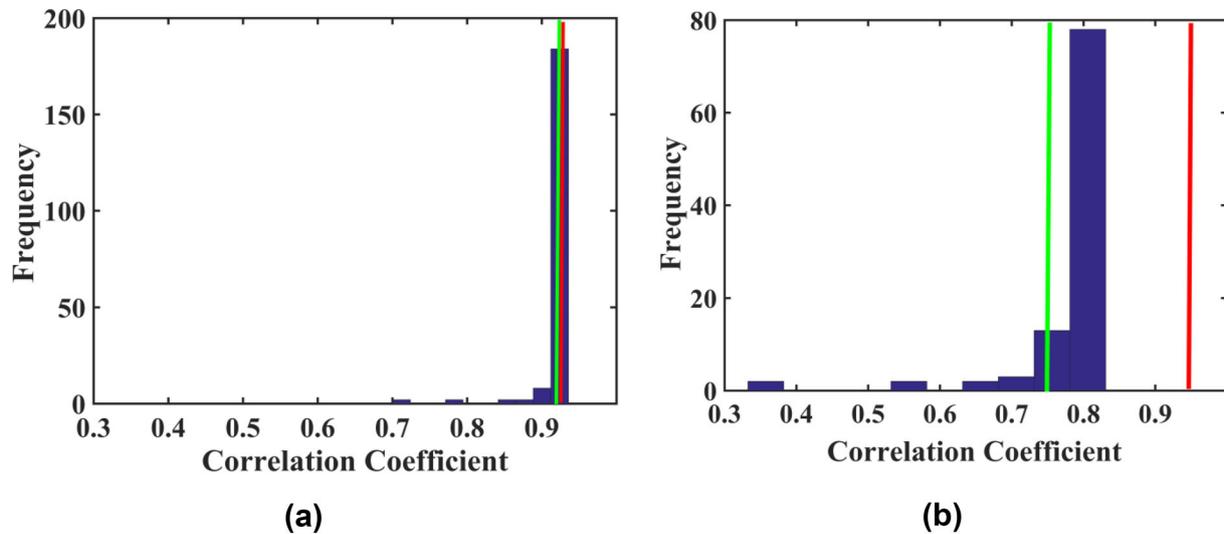


**Figure 8.** E-type maps of Mn, Fe, Co and Pb generated from 100 realizations of co-simulation results.

ered as the best prediction of the true values (Deutsch and Journel 1998). As can be seen from Figure 10, the average of simulated values fluctuates around the true value, and the conditional regression between true and simulated values is close to identity, which corroborates conditional unbiasedness. These results demonstrate that the TBCOSIM algorithm is sufficient in this geo-environmental study.

Other ways for quality checking of simulation results are to check the reproduction of local and global distributions. For the former, spatial continuity as direct and cross-variograms of the simulated realizations should be compared in average with the fitted theoretical model (Emery 2004). If the average model is not compatible with the theoretical model, it implies that the simulation algorithm produced biased results. In this study, the non-conditional results corroborate that the spatial continuity is well reproduced by TBCOSIM (Fig. 11). Slight differences that appear in large lags are typically impacted by the extensiveness of the underlying area, which is bounded regionally, termed as ergodic fluctuations (Matheron 1989). The reason why non-conditional

realizations are employed is that the conditioning information distorts the prior model due to the restriction in the size of a domain (Lantuéjoul 2002; Emery 2007, 2008). Since ergodic fluctuations are also observed for the marginal distribution, the global statistical parameters of simulation results can be affected through conditioning dataset as well (Goovaerts 1997). Cumulative distribution function is a satisfying measure of global distribution, in which the abscissa shows the data values ordered from smallest to largest and the ordinate represents the cumulative probability assigned to each data (Davis 1986). Examination of this graph provides an interesting comparison of the global distribution for different random fields. The cumulative distribution functions are calculated over the non-conditional realizations and are compared with the primary declustered cumulative distribution (Fig. 12). The closeness of the model statistics to the sample statistics marks a superb reproduction in global distribution. There are some methodologies such as simulated annealing (Journel and Xu 1994) that allows improving the reproduction of the target



**Figure 9.** Correlation coefficients for Mn and Fe over 100 realizations (red: primary correlation coefficient; green: average reproduced correlation coefficient). (a) Co-simulation, (b) simulation.

statistics respecting the conditioning data. Nevertheless, this is not the scope of the current study.

### Spatial Uncertainty Quantification of the Heavy Metals

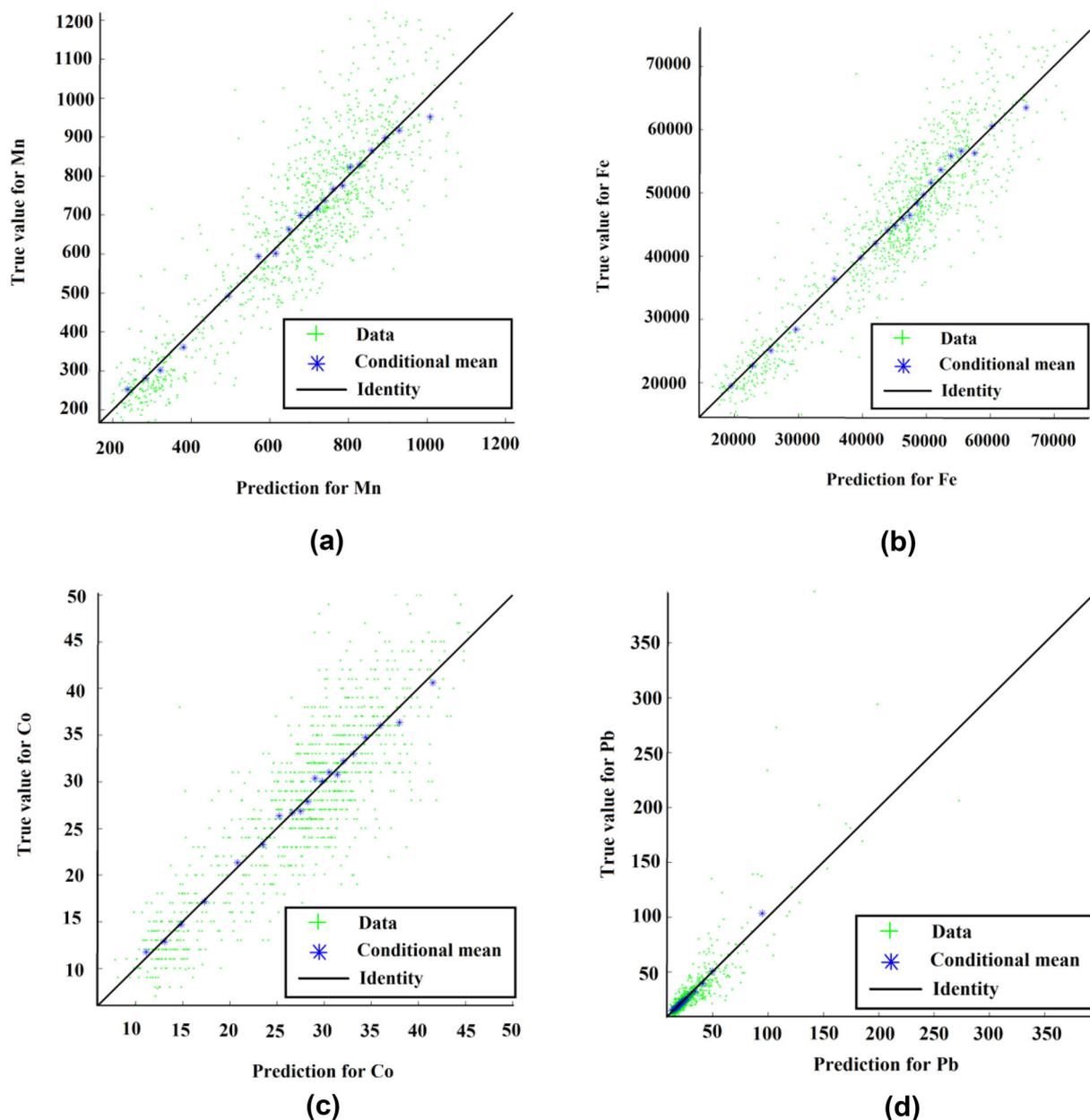
This step consists of calculating the uncertainty within the realizations. The concepts of Igeo (Fig. 13) and EF (Fig. 14) show varying levels of heavy metals contaminations in the near-surface soils. Therefore, according to the different thresholds of contamination levels applicable to each index, the probability can be evaluated for heavy metal concentrations above, below or between the thresholds to identify polluted areas that need remediation. Figure 13 shows the probability maps for the Igeo thresholds (deteriorating, extremely polluted, strongly to extremely polluted, strongly polluted, moderately to strongly polluted and moderately polluted) of Pb. Areas with little uncertainty are those associated with high probability for a range of threshold (shown in red in Fig. 13); areas marked in dark blue indicate little risk of not finding this range of established thresholds, or those associated with very low probability of having contamination depending on the pre-specified thresholds, while the other areas (in light blue, green or yellow in Fig. 13) have higher uncertainty of being contaminated or not.

### DISCUSSION

Heavy metals in soils pose great ecological risk in various ways. Plants could take up heavy metals and incorporate them in their tissues. Such plants would then become sources of food poisoning in the ecosystem food web. Human are also capable of being affected by soil heavy metals in three major ways including direct ingestion (this pathway is more common among children), direct inhalation through the mouth or nose, and dermal adsorption of soils adhered to the skin (Qu et al. 2012; Qi et al. 2016). By using Igeo and EF (Figs. 13 and 14) the spatial distribution of the contamination levels of heavy metals has been documented for the Maibele North exploration site. The probability maps shows range of indicated thresholds at a local (block-by-block) scale that estimate, for each block, the frequency of occurrence of each range of indicated contamination threshold over the 100 conditional realizations (Fig. 14). These show the risk of occurrence of the range of indicated threshold from the one that has already been specified.

Both EF and Igeo indicate high level of Pb in the soils. At some points, Pb had EF greater than 40, a remarkable enrichment. Natural concentration of Pb in the earth's crust varies from 15 to 20 mg/kg (ATSDR 2007), while most of the Igeo values are negative, Pb had positive Igeo values as high as 5. However, Pb had relatively weaker spatial variations

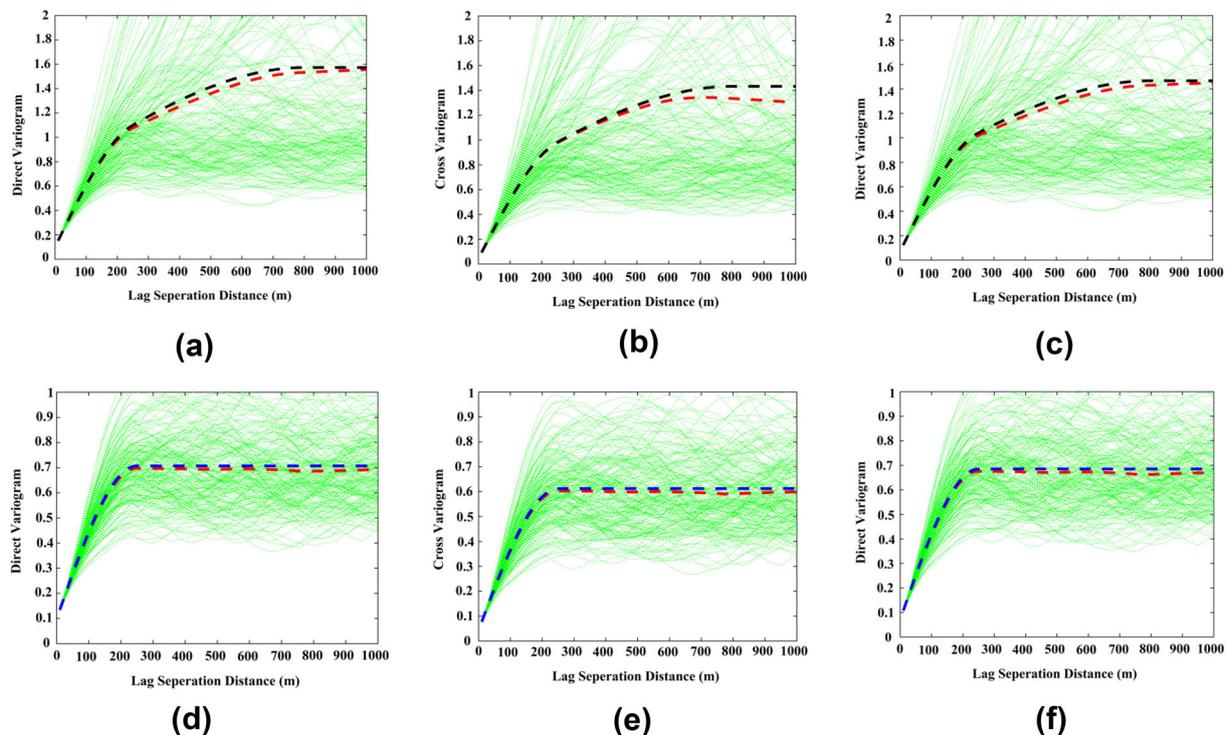
## Multivariate Mapping of Heavy Metals Spatial Contamination



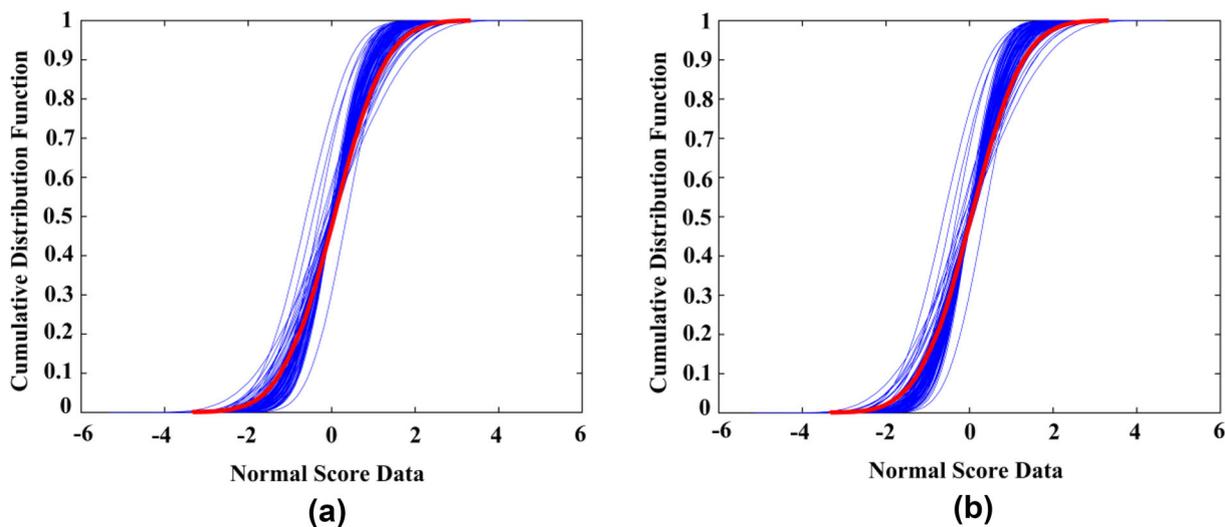
**Figure 10.** (Color line) Scatter plot between true values (ordinate) and predicted values (abscissa) at data locations. Predicted values are calculated as the average of 100 realizations obtained with TBCOSIM. (a) Mn, (b) Fe, (c) Co, (d) Pb.

(Fig. 8). Lead, a nonessential and toxic element, is released from natural and anthropogenic sources. Common sources of Pb in soils are manure, sewage sludge, pesticides, vehicle exhausts and industrial fumes. The studied area is currently used as a grazing field. Cattle dung could therefore possibly be a source of Pb in the soils. Moreover, the use of inorganic fertilizers and agro-chemicals is a common

practice in the catchment areas of the Sekgopye river, close to the study area, which could contribute significantly to the presence of Pb in the river. Open burning of domestic and industrial waste products and disposal of sewage and car batteries in the environment by local inhabitants could constitute a potential source of Pb contamination in the catchment. The co-simulation result (Fig. 8) shows Mn,



**Figure 11.** Direct and cross-variograms of simulated realizations with turning bands co-simulation (TBCOSIM). For brevity, only the direct and cross-variograms of Fe and Mn are displayed (black line: fitted model along azimuth (135°); blue line: fitted model along azimuth (45°) red line: average of the realizations; green lines: individual realizations). (a) Mn (azimuth: 135°), (b) Mn & Fe (azimuth: 135°), (c) Fe (azimuth: 135°), (d) Mn (azimuth: 45°), (e) Mn & Fe (azimuth: 45°), (f) Fe (azimuth: 45°).



**Figure 12.** Global distribution reproduction of the simulation results. For brevity, only the cumulative distribution functions of Fe and Mn are displayed. Blue lines: individual realizations and red line: declustered primary dataset. (a) Mn, (b) Fe.

Co, and Fe had similar spatial variations different from Pb. Manganese is an element of low toxicity having considerable biological importance. It is one

of the biogeochemically active transition metals in the aquatic environment (Evans et al. 1977; Hasan et al. 2013). On the other hand, Fe is often used as

## Multivariate Mapping of Heavy Metals Spatial Contamination

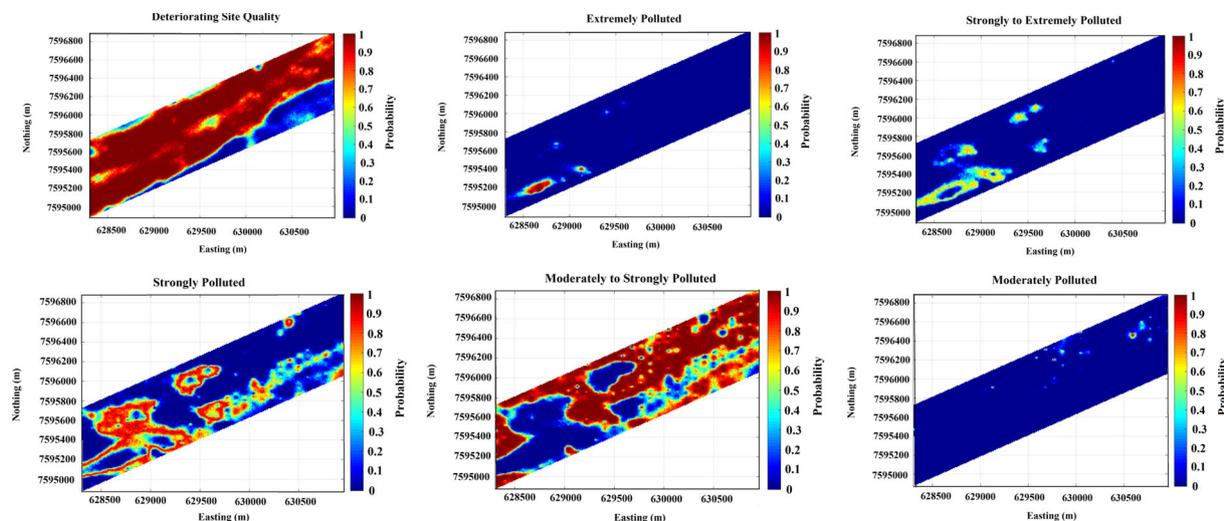


Figure 13. Probability maps for the Igeo thresholds of Pb.

an indication of natural changes in the heavy metal carrying capacity of soils and sediments (Rule 1986), and its concentrations were related to the abundance of metal reactive compounds not significantly affected by human activities (Luoma 1990). The high concentrations of Pb in the area suggest possible soil environmental pollution, but their speciation and bio-accessibility, rather than simply their total concentrations, have to be ascertained (Gupta et al. 1996).

The PLI did not show much fluctuation over the soilscape. The values ranged from 0.71 to 1.03 (results not shown in this paper). PLI values in this range could be attributed to little anthropogenic activities since these areas are nearby residential plots and industrial establishments. The EF values for Co ranged from 2 to 5 in the entire study area (100%), indicating moderate enrichment. For this, we presented the first, second and third quartiles as a better way to graphically show the distribution of the enrichment. Cobalt enrichment in the soils (Fig. 14) falls within an environmentally safe zone. Cobalt concentrations in the soils are, however, still within the tolerable limit and pose no threat of environmental pollution. Cobalt is consistently attributed to weathering processes and its amounts in soils are usually too low to reach contamination levels (Manta et al. 2002; Lee et al. 2006). Manganese and Fe, on the other hand, were below the contamination thresholds in the study area. Therefore, no uncertainty quantification maps were produced

for them. However, intermittent monitoring would be in order.

Whenever the concentration levels of a pollutant exceed the minimum allowable threshold, it becomes very important to consider future remediation measures. In this study, Pb shows high enrichment at several locations and efforts should be made to put this under control. The cost of remediation measures are usually high and this is why spatial distribution of the pollutants is very important in planning what goes to where in the remediation exercise. With stochastic turning bands co-simulation maps and uncertainty maps of Pb created, one can delineate areas that need urgent correction before any agricultural use can be made of the location again. We suggest the use of heavy metal-accumulating plants (metallophytes) in cleaning up Pb from the soils.

This study further affirms that in geo-environmental assessment and monitoring of heavy metals contamination in soils it is better to use turning bands co-simulation approach where there is a strong positive correlation among the variables and particularly, where the sampling is unequal as well. However, it does not imply that this approach is inefficient in equally sampled case. This conclusion is in agreement with the work of Emery (2008), whose study indicated that a turning bands co-simulation algorithm is more accurate for interpolating co-regionalized variables monitored at sampling locations including soil contamination dataset.

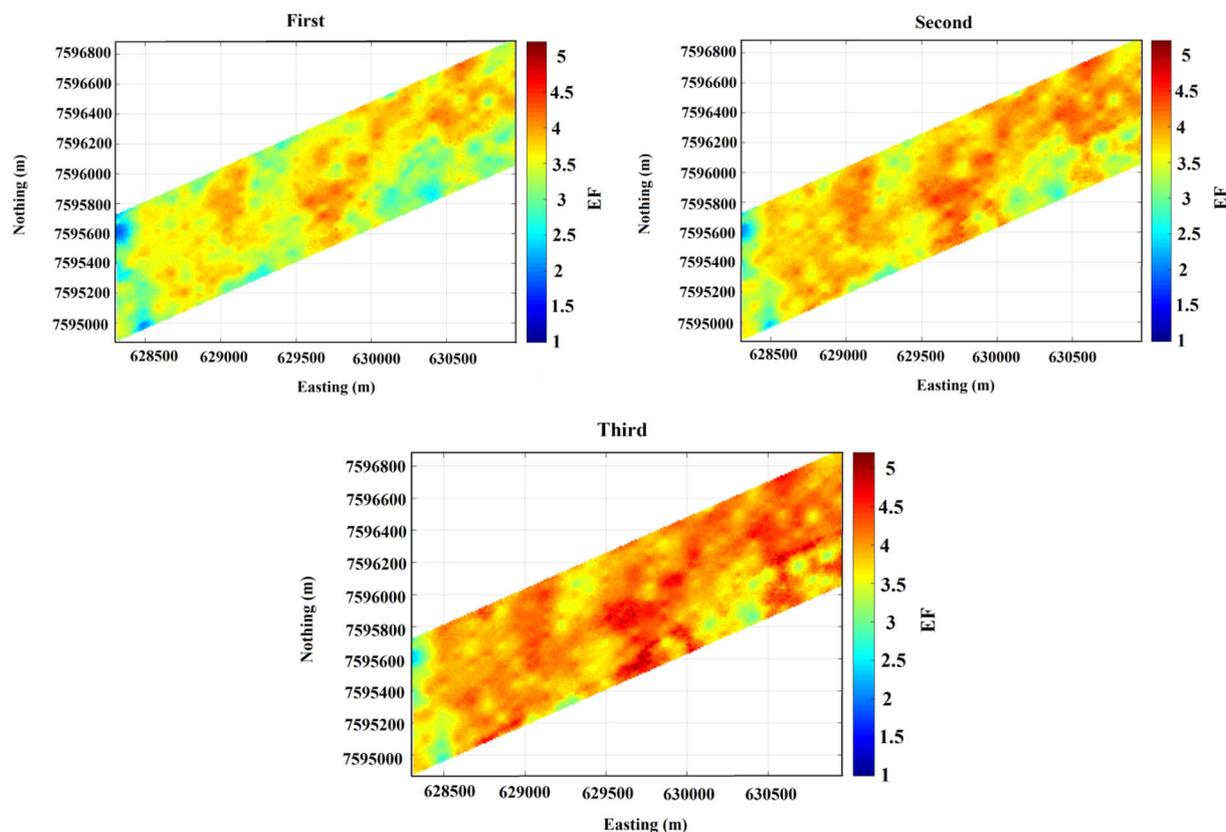


Figure 14. Quartile plots for EF for Co.

## CONCLUSIONS

A turning bands conditional co-simulation approach has been successfully used to quantify and map spatial uncertainty of heavy metals in a semiarid Ni–Cu exploration field in Botswana. In the case of heavy metals dataset with strong positive correlation, multivariate mapping of the cross-correlated variables is highly encouraged rather than univariate mapping of each variable separately because the former method takes into account the intrinsic dependency among variables, and so the post-processing outputs are more reliable. The co-simulation maps of the heavy metals show spatial variations in Co, Mn and Fe distribution. The results of the three indices applied in the study suggest no risk of Co, Mn and Fe contamination at Maibele Airstrip North. Turning bands co-simulation was validated by examining the reproduction of input statistics and by cross-validation. However, the higher spatial concentrations of Pb indicate that, among the four elements studied, it contributed solely to the possible

soil environmental pollution as indicated from PLI, but their speciation and bio-accessibility, rather than simply their total concentrations, have to be ascertained. The results of this study would be invaluable for land use, and management decisions in the area and turning bands co-simulation algorithm can be successfully applied to mapping heavy metal uncertainties in areas having similar geochemical properties across the globe.

## ACKNOWLEDGMENTS

We are grateful to Dr. John Carranza and the two anonymous reviewers for their comments, which substantially helped improving the final version of the manuscript. The second author acknowledges the Nazarbayev University for supporting this work through Faculty Development Competitive Research Grants for 2018–2020 under Contract No. 090118FD5336.

## Multivariate Mapping of Heavy Metals Spatial Contamination

### REFERENCES

- Addo, M. A., Darko, E. O., Gordon, C., Nyarko, B. J. B., Gbadago, J. K., Nyarko, E., et al. (2012). Evaluation of heavy metals contamination of soil and vegetation in the vicinity of a cement factory in the Volta Region, Ghana. *International Journal of Science and Technology*, 2(1), 40–50.
- Agency for Toxic Substance and Disease Registry (ATSDR) (2007). Toxicological profile for lead. <http://www.atsdr.cdc.gov/toxprofiles/tp13.pdf>.
- Anderson, R. H., & Kravitz, M. J. (2010). Evaluation of geochemical associations as a screening tool for identifying anthropogenic trace metal contamination. *Environmental Monitoring and Assessment*, 167(1), 631–641. <https://doi.org/10.1007/s10661-009-1079-2>.
- Bogaert, P. (1999). On the optimal estimation of the cumulative distribution function in presence of spatial dependence. *Mathematical Geology*, 31(2), 213–239. <https://doi.org/10.1023/a:1007513918820>.
- Bourennane, H., Douay, F., Sterckeman, T., Villanneau, E., Ciesielski, H., King, D., et al. (2010). Mapping of anthropogenic trace elements inputs in agricultural topsoil from Northern France using enrichment factors. *Geoderma*, 157(3), 165–174. <https://doi.org/10.1016/j.geoderma.2010.04.009>.
- Bourgault, G. (1997). Spatial declustering weights. *Mathematical Geology*, 29(2), 277–290. <https://doi.org/10.1007/bf02769633>.
- Buat-Menard, P., & Chesselet, R. (1979). Variable influence of atmospheric flux on the trace metal chemistry of oceanic suspended matter. *Earth and Planetary Science Letters*, 42, 398–411.
- Burgos, P., Madejón, E., Pérez-de-Mora, A., & Cabrera, F. (2008). Horizontal and vertical variability of soil properties in a trace element contaminated area. *International Journal of Applied Earth Observation and Geoinformation*, 10(1), 11–25. <http://doi.org/10.1016/j.jag.2007.04.001>.
- Carr, J. R., & Myers, D. E. (1985). COSIM: A FORTRAN IV program for coconditional simulation. *Computers & Geosciences*, 11(6), 675–705. [https://doi.org/10.1016/0098-3004\(85\)90012-3](https://doi.org/10.1016/0098-3004(85)90012-3).
- Chilès, J.-P., & Delfiner, P. (2012). *Geostatistics: Modeling spatial uncertainty* (2nd ed.). New York: Wiley.
- Chilès, J.-P., & Lantuéjoul, C. (2005). Prediction by conditional simulation: Models and algorithms. In M. Bilodeau, F. Meyer, & M. Schmitt (Eds.), *Space, structure and randomness: Contributions in honor of Georges Matheron in the field of geostatistics, random sets and mathematical morphology* (pp. 39–68). New York, NY: Springer.
- Cloquet, C., Carignan, J., Libourel, G., Sterckeman, T., & Perdrix, E. (2006). Tracing source pollution in soils using cadmium and lead isotopes. *Environmental Science and Technology*, 40(8), 2525–2530. <https://doi.org/10.1021/es052232+>.
- Davis, J. C. (1986). *Statistics and data analysis in geology* (2nd ed.). New York: Wiley.
- Desaules, A. (2012). Critical evaluation of soil contamination assessment methods for trace metals. *Science of the Total Environment*, 426, 120–131. <https://doi.org/10.1016/j.scitotenv.2012.03.035>.
- Deutsch, C. V., & Journel, A. (1998). *GSLIB: Geostatistical software and user's guide* (2nd ed.). New York: Oxford University Press.
- Emery, X. (2004). Testing the correctness of the sequential algorithm for simulating Gaussian random fields. *Stochastic Environmental Research and Risk Assessment*, 18(6), 401–413. <https://doi.org/10.1007/s00477-004-0211-7>.
- Emery, X. (2005). Variograms of order  $\omega$ : A tool to validate a bivariate distribution model. *Mathematical Geology*, 37(2), 163–181. <https://doi.org/10.1007/s11004-005-1307-4>.
- Emery, X. (2007). Conditioning simulations of Gaussian random fields by ordinary kriging. *Mathematical Geology*, 39(6), 607–623. <https://doi.org/10.1007/s11004-007-9112-x>.
- Emery, X. (2008). A turning bands program for conditional co-simulation of cross-correlated Gaussian random fields. *Computers & Geosciences*, 34(12), 1850–1862. <https://doi.org/10.1016/j.cageo.2007.10.007>.
- Emery, X. (2012). Cokriging random fields with means related by known linear combinations. *Computers & Geosciences*, 38(1), 136–144. <https://doi.org/10.1016/j.cageo.2011.06.001>.
- Emery, X., & Lantuéjoul, C. (2006). TBSIM: A computer program for conditional simulation of three-dimensional Gaussian random fields via the turning bands method. *Computers & Geosciences*, 32(10), 1615–1628. <https://doi.org/10.1016/j.cageo.2006.03.001>.
- Evans, D. W., Cutshall, N. H., Cross, F. A., & Wolfe, D. A. (1977). Manganese cycling in the Newport River estuary, North Carolina. *Estuarine and Coastal Marine Science*, 5(1), 71–80. [https://doi.org/10.1016/0302-3524\(77\)90074-3](https://doi.org/10.1016/0302-3524(77)90074-3).
- Eze, P. N., Mosokomani, V. S., Udeigwe, T. K., & Oyedele, O. F. (2016a). Quantitative geospatial dataset on the near-surface heavy metal concentrations in semi-arid soils from Maibele Airstrip North, Central Botswana. *Data in Brief*, 8, 1448–1453. <https://doi.org/10.1016/j.dib.2016.08.026>.
- Eze, P. N., Mosokomani, V. S., Udeigwe, T. K., Oyedele, O. F., & Fagbamigbe, A. F. (2016b). Geostatistical analysis of trace elements PXRF dataset of near-surface semi-arid soils from Central Botswana. *Data in Brief*, 9, 764–770. <https://doi.org/10.1016/j.dib.2016.10.010>.
- Eze, P. N., Udeigwe, T. K., & Stietiya, M. H. (2010). Distribution and potential source evaluation of heavy metals in prominent soils of Accra Plains. *Ghana. Geoderma*, 156(3), 357–362. <https://doi.org/10.1016/j.geoderma.2010.02.032>.
- Gneiting, T. (1999). The correlation bias for two-dimensional simulations by turning bands. *Mathematical Geology*, 31(2), 195–211. <https://doi.org/10.1023/a:1007561801981>.
- Gómez-Hernández, J. J., & Cassiraga, E. F. (1994). Theory and practice of sequential simulation. In M. Armstrong & P. A. Dowd (Eds.), *Geostatistical simulations* (pp. 111–124). Dordrecht: Springer.
- 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, Netherlands.
- Goovaerts, P. (1997). *Geostatistics for natural resources evaluation*. New York: Oxford University Press.
- Gupta, S. K., Vollmer, M. K., & Krebs, R. (1996). The importance of mobile, momilisable and pseudo total heavy metal fractions in soil for three-level risk assessment and risk management. *Science of the Total Environment*, 178, 11–20.
- Hani, A., & Pazira, E. (2011). Heavy metals assessment and identification of their sources in agricultural soils of Southern Tehran, Iran. *Environmental Monitoring and Assessment*, 176(1), 677–691. <https://doi.org/10.1007/s10661-010-1612-3>.
- Hasan, A. B., Kabir, S., Selim Reza, A. H. M., Nazim Zaman, M., Ahsan, A., & Rashid, M. (2013). Enrichment factor and geoaccumulation index of trace metals in sediments of the ship breaking area of Sitakund Upazilla (Bhatary-Kumira), Chittagong, Bangladesh. *Journal of Geochemical Exploration*, 125, 130–137. <https://doi.org/10.1016/j.gexplo.2012.12.002>.
- Johnson and Wichern. (1998). *Applied multivariate statistical analysis* (4th ed.). New York: Prentice-Hall.
- Jørgensen, K., & Jensen, L. S. (2009). Chemical and biochemical variation in animal manure solids separated using different commercial separation technologies. *Bioresource Technology*, 100(12), 3088–3096. <https://doi.org/10.1016/j.biortech.2009.01.065>.

- Journel, A. G., & Xu, W. (1994). Posterior identification of histograms conditional to local data. *Mathematical Geology*, 26(3), 323–359.
- Kampunzu, A. B., Tombale, A. R., Zhai, M., Bagai, Z., Majaule, T., & Modisi, M. P. (2003). Major and trace element geochemistry of plutonic rocks from Francistown, NE Botswana: evidence for a Neoproterozoic continental active margin in the Zimbabwe craton. *Lithos*, 71(2), 431–460. [https://doi.org/10.1016/S0024-4937\(03\)00125-7](https://doi.org/10.1016/S0024-4937(03)00125-7).
- 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: Springer.
- Lantuéjoul, C. (2002). *Geostatistical simulation, models and algorithms* (p. 256). Berlin: Springer.
- Lee, C. S.-L., Li, X., Shi, W., Cheung, S. C.-N., & Thornton, I. (2006). Metal contamination in urban, suburban, and country park soils of Hong Kong: A study based on GIS and multivariate statistics. *Science of the Total Environment*, 356(1), 45–61. <https://doi.org/10.1016/j.scitotenv.2005.03.024>.
- Leuangthong, O., & Deutsch, C. V. (2003). Stepwise conditional transformation for simulation of multiple variables. *Mathematical Geology*, 35(2), 155–173. <https://doi.org/10.1023/a:1023235505120>.
- Likuku, A. S., Gaboutloeloe, G. K., & Mmolawa, K. B. (2013). Determination and source apportionment of selected heavy metals in aerosol samples collected from Sebele. *American Journal of Environmental Sciences*, 9(2), 188–200.
- Liu, X., Wu, J., & Xu, J. (2006). Characterizing the risk assessment of heavy metals and sampling uncertainty analysis in paddy field by geostatistics and GIS. *Environmental Pollution*, 141(2), 257–264. <https://doi.org/10.1016/j.envpol.2005.08.048>.
- Luoma, S. N. (1990). Processes affecting metal concentrations in estuarine and coastal marine sediments. In R. W. Furness & P. S. Rainbow (Eds.), *Heavy metals in the marine environment* (pp. 51–66). Boca Raton: CRC Press Inc.
- Manta, D. S., Angelone, M., Bellanca, A., Neri, R., & Sprovieri, M. (2002). Heavy metals in urban soils: a case study from the city of Palermo (Sicily), Italy. *Science of the Total Environment*, 300(1), 229–243. [https://doi.org/10.1016/S0048-9697\(02\)00273-5](https://doi.org/10.1016/S0048-9697(02)00273-5).
- Mantoglou, A. (1987). Digital simulation of multivariate two- and three-dimensional stochastic processes with a spectral turning bands method. *Mathematical Geology*, 19(2), 129–149. <http://doi.org/10.1007/bf00898192>.
- Matheron, G. (1973). The intrinsic random functions and their applications. *Advances in Applied Probability*, 5(3), 439–468. <https://doi.org/10.2307/1425829>.
- Matheron, G. (1989). *Estimating and choosing: An essay on probability in practice*. Berlin: Springer.
- Müller, G. (1969). Index of geoaccumulation in sediments of the Rhine River. *GeoJournal*, 2, 108–118.
- Myers, D. E. (1989). Vector conditional simulation. In M. Armstrong (Ed.), *Geostatistics* (pp. 283–293). Dordrecht: Springer.
- 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. <https://doi.org/10.1016/j.crte.2015.05.008>.
- Pardo-Igúzquiza, E., & Chica-Olmo, M. (1993). The Fourier Integral Method: An efficient spectral method for simulation of random fields. *Mathematical Geology*, 25(2), 177–217. <https://doi.org/10.1007/bf00893272>.
- Pardo-Igúzquiza, E., & Chica-Olmo, M. (1994). Spectral simulation of multivariable stationary random functions using covariance fourier transforms. *Mathematical Geology*, 26(3), 277–299. <https://doi.org/10.1007/bf02089226>.
- Pebesma, E. J. (2004). Multivariable geostatistics in S: The gstat package. *Computers & Geosciences*, 30(7), 683–691. <https://doi.org/10.1016/j.cageo.2004.03.012>.
- Qi, J., Zhang, H., Li, X., Lu, J., & Zhang, G. (2016). Concentrations, spatial distribution, and risk assessment of soil heavy metals in a Zn–Pb mine district in southern China. *Environmental Monitoring and Assessment*, 188(7), 413. <https://doi.org/10.1007/s10661-016-5406-0>.
- Qu, C., Sun, K., Wang, S., Huang, L., & Bi, J. (2012). Monte Carlo simulation-based health risk assessment of heavy metal soil pollution: A case study in the Qixia Mining Area, China. *Human and Ecological Risk Assessment: An International Journal*, 18(4), 733–750. <https://doi.org/10.1080/10807039.2012.688697>.
- Reimann, C., & de Caritat, P. (2005). Distinguishing between natural and anthropogenic sources for elements in the environment: Regional geochemical surveys versus enrichment factors. *Science of the Total Environment*, 337(1), 91–107. <https://doi.org/10.1016/j.scitotenv.2004.06.011>.
- Rivoirard, J. (1994). *Introduction to disjunctive Kriging and non-linear geostatistics*. Oxford: Oxford University Press.
- Rule, J. H. (1986). Assessment of trace metal element geochemistry of Hampton Roads and lower Chesapeake Bay area sediments. *Environmental Geology and Water Sciences*, 8, 209–219.
- Sakizadeh, M., Sattari, M. T., & Ghorbani, H. (2017). A new method to consider spatial risk assessment of cross-correlated heavy metals using geo-statistical simulation. *Journal of Mining and Environment*, 8(3), 373–391. <https://doi.org/10.22044/jme.2017.892>.
- Tomlinson, D. L., Wilson, J. G., Harris, C. R., & Jeffrey, D. W. (1980). Problem in the assessment of heavy metals levels in estuaries and the formation of a pollution index. *Helgoländer Meeresuntersuchungen*, 33, 566–575.
- Tran, T. T. (1994). Improving variogram reproduction on dense simulation grids. *Computers & Geosciences*, 20(7), 1161–1168. [https://doi.org/10.1016/0098-3004\(94\)90069-8](https://doi.org/10.1016/0098-3004(94)90069-8).
- Udeigwe, T. K., Young, J., Kandakji, T., Weindorf, D. C., Mahmoud, M. A., & Stietiya, M. H. (2015). Elemental quantification, chemistry, and source apportionment in golf course facilities in a semi-arid urban landscape using a portable X-ray fluorescence spectrometer. *Solid Earth*, 6(2), 415–424. <https://doi.org/10.5194/se-6-415-2015>.
- Wackernagel, H. (2003). *Multivariate geostatistics—An introduction with applications*. Berlin: Springer.
- Xie, Y., Chen, T.-B., Lei, M., Yang, J., Guo, Q.-J., Song, B., et al. (2011). Spatial distribution of soil heavy metal pollution estimated by different interpolation methods: Accuracy and uncertainty analysis. *Chemosphere*, 82(3), 468–476. <https://doi.org/10.1016/j.chemosphere.2010.09.053>.
- Zhang, C. (2006). Using multivariate analyses and GIS to identify pollutants and their spatial patterns in urban soils in Galway, Ireland. *Environmental Pollution*, 142(3), 501–511. <https://doi.org/10.1016/j.envpol.2005.10.028>.
- Zhang, C., & Selinus, O. (1998). Statistics and GIS in environmental geochemistry—Some problems and solutions. *Journal of Geochemical Exploration*, 64(1), 339–354. [https://doi.org/10.1016/S0375-6742\(98\)00048-X](https://doi.org/10.1016/S0375-6742(98)00048-X).
- Zimmerman, A. J., & Weindorf, D. C. (2010). Heavy metal and trace metal analysis in soil by sequential extraction: A review of procedures. *International Journal of Analytical Chemistry*. <https://doi.org/10.1155/2010/387803>.