JME

Journal of Mining & Environment,
Vol.8, No.3, 2017, 373-391.
DOI: 10.22044/jme.2017.892

A new method to consider spatial risk assessment of cross-correlated heavy metals using geo-statistical simulation

M. Sakizadeh[1*], M.T. Sattari[2] and H. Ghorbani[3]

1. Department of Environmental Sciences, Faculty of Sciences, Shahid Rajaee Teacher Training University, Tehran, Iran
2. Department of Water Engineering, Agriculture Faculty, University of Tabriz, Tabriz, Iran
3. School of Agricultural Engineering, Shahrood University of Technology, Shahrood, Iran

Received 10 December 2016; received in revised form 3 March 2017; accepted 13 March 2017
**Corresponding author: msakizadeh@gmail.com (M. Sakizadeh).*

Abstract
The soil samples were collected from 170 sampling stations in an arid area in Shahrood and Damghan, characterized by prevalence of mining activity. The levels of Co, Pb, Ni, Cs, Cu, Mn, Sr, V, Zn, Cr, and Tl were recorded in each sampling location. A new method known as min/max autocorrelation factor (MAF) was applied for the first time in the environmental research works to de-correlate these elements before their geo-statistical simulation. The high cross-correlation among some elements, while poor spatial correlation among the others, could have made spectral decomposition of MAFs unstable, resulting in some negative eigenvalues, so it was decided to reduce the dimensionality of the original variables by Principal Component Analysis (PCA). The resultant 6 heavy metals (Cr, Mn, Cu, V, Ni, and Co) were converted to their respective MAFs followed by their geo-statistical simulation using Sequential Gaussian Simulation (SGS) independently. Examination of the cross-variograms of MAFs indicated that the resultant factors had been rigorously de-correlated, especially at zero lag and around Δ lag distance. Several validation checks including reproduction of variograms in data and normal score space, close matching between distribution of MAFs versus simulated realizations, and reproduction of descriptive statistics and data histograms all confirmed that the data values had been honored by this applied method. The results obtained indicated that this method could reproduce the data values as well as the spatial continuity of heavy metals (e.g. semi-variograms) successfully. In addition, this technique is simpler and more computationally efficient than its equivalent sequential Gaussian co-simulation as fitting a linear model of co-regionalization (LMC) is not required in the data-driven MAF method.

Keywords: *Decorrelation, Geo-Statistical Simulation, Min/Max Autocorrelation Factor.*

1. Introduction

An issue in the spatial assessment of soil and environmental attributes is that the information is only available at a limited number of sampling locations, which necessitates adopting methods to quantify the distribution and variability of the property of interest at unsampled locations. For this purpose, univariate kriging methods such as ordinary kriging [1], regression kriging [2], indicator kriging [3], and factorial kriging [4] have been widely proposed to cope with this problem.

In this respect, there are multiple applications of geo-statistical methods available for the analysis of environmental attributes of soil in the previous studies. For instance, Lu et al. [5] used principal component analysis (PCA) and ordinary kriging to analyze the spatial distribution and origin of heavy metals in agricultural soils in Shunyi, Beijing, China, and concluded that soil contamination by Cd, Cu, and Zn was mainly derived from agricultural practices, whereas As and Pb were mainly due to soil parent materials, and Hg contamination was caused by the

atmospheric deposits. Guagliardi et al. [6] applied the multi-Gaussian approach to build the geochemical maps of heavy metals, and proved that high concentrations of potentially toxic elements were available in soils near major roads, indicating that the anthropogenic factors determine the anomalies in these areas. Lv et al. [7] utilized factorial kriging to study the spatial distribution of heavy metals in the soils of China, and showed that the spatial variation of Cr and Ni was related to the parent material at both the local and regional scales, and was derived from natural sources. Mining activity was observed to affect the spatial variation in Cd, Cu, Pb, and Zn at the local scale, while the parent material dominated the spatial variation of those metals at the regional scale.

However, the downside of these techniques is that when multiple contaminants are considered, the correlation coefficients of the data are not taken into account in these methods. A practical alternative is to use multivariate methods (e.g. co-kriging) in order to analyze the relationship among the variables of interest [8] but these methods are computationally intensive since they involve modeling cross-variograms [9]. In addition, the linear model of co-regionalization (LMC) should be satisfied, meaning that the direct and cross variograms have to be a linear combination of basic structures [10], which is difficult to implement for a large number of variables [11].

In order to facilitate the analysis and modeling, de-correlation has been proposed, whereby the features are converted to new variables that are a linear combination of the original attributes. Multiple methods such as Principal Component Analysis [12], Stepwise Conditional Transformation (SCT) [13], Min/Max Autocorrelation Factors (MAF) [14], Independent Component Analysis (ICA) [15], and Uniformly Weighted Exhaustive Diagonalization with Gauss iterations (U-WEDGE) [16] have been utilized in this field.

Among the above-mentioned methods, PCA and MAF are the most applied ones. PCA transforms the attributes to new orthogonal factors in a new space, where they are no longer correlated, thereby each new variable or factor can be treated independently, avoiding the LMC constrain [17]. However, only in cases in which there is an intrinsic correlation between the data, PCA does not guarantee de-correlation for a non-zero separation vector [18].

To the contrary, MAF, similar to PCA, de-correlates a set of variables into new uncorrelated factors. In this field, two techniques have been proposed in the literature to implement MAF. In the first one, known as model-based MAF [e.g. 19-20], the direct and cross-spatial continuity of attributes is modeled mainly through fitting a two-structure linear model of co-regionalization (2SLMC) [11, 21]. In the second approach, known as data-driven MAF [8, 19], cross-correlation between the derived factors is removed for at least short lag distances. The latest method was first developed by Switzer and Green [23] and utilized later by other researchers in different fields of studies [e.g. 22, 24]. The redeeming feature of the data-driven method over the model-based technique is that orthagonalization of factors can be implemented without the need for fitting a LMC, hereby it is more computationally efficient [24].

In this respect, Shakiba et al. [21] applied MAF along with fuzzy logic on a set of seismic attributes. The study was concerned with the discrimination ability of these methods for fault and non-fault areas in oil and gas exploration sites. MAF was proved to be successful in dimensional reduction of seismic attributes and transformation of variables into un-correlated factors for all lag spacings. In a more relevant work to the current study, Lin et al. [16] combined Uniformly Weighted Exhaustive Diagonalization with Gauss iterations (U-WEDGE) along with Sequential Gaussian Simulation (SGS) to de-correlate a number of spatially correlated heavy metals and analyze the uncertainty in the distribution of these contaminants. The results of the uncertainties were used by the Info-Gap Decision Theory (IGDT) to propose remediation regimes for the contaminated area. As a whole, all the applied methods in the latest research works were successful in the implementation of the objectives of the study.

The main objectives of the current research work were (i) to consider the applicability of min/max autocorrelation factor to de-correlate some cross-correlated trace elements in the soil samples of Shahrood and Damghan in the central part of Iran (ii) to assess the reproduction of the original data from simulated MAFs.Min/max autocorrelation factor followed by Sequential Gaussian Simulation To our knowledge, it is it is the first application for the assessment of a contaminated area.

2. Materials and methods

2.1. De-correlation using min/max auto-correlation factor

Let $Z(u) = \left[Z_1(u), Z_2(u), \ldots, Z_K(u)\right]$ be a stationary and ergodic random field (RF) over a region D, representing K correlated heavy metals measured on soil samples in the studied area. The associated variogram matrices of the foregoing spatial samples can be represented by [8]:

$$\Gamma_z(h) = \frac{1}{2}\mathrm{E}\left[\left(Z(u+h)-Z(u)\right)^T .\left(Z(u+h)-Z(u)\right)\right] \quad (1)$$

Since the variogram matrix depends only on the lad distance of the sampling points, at large lag distances (e.g. $h \to \infty$), the variogram matrix approaches the variance-covariance matrix (e.g. $\Gamma_z(h) \to B$), in which B is the variance-covariance matrix at zero lag distance. Now consider the normal score transformation of the above random field, $Y(u) = \left[Y_1(u), \ldots, Y_K(u)\right] = \left[f_1(Z_1(u)), \ldots, f_K(Z_K(u))\right]$, which is multi-Gaussian with zero mean and unit variance. If there is a spatial auto-correlation among the attributes, the variables can be converted to orthogonal, independent factors called min/max autocorrelation factors (MAFs) (Equation 2) that can be simulated independently using the univariate geo-statistical methods followed by back-transformation of the results to the data space by the related back-transformation matrix [23].

$$\mathrm{M_{MAF}(u)} = \mathrm{A^T Y(u)} \quad (2)$$

where A is the transformation matrix for which the coefficients are obtained through:

$$\mathrm{A^T} = \mathrm{Q_2^T \Lambda_1^{-1/2} Q_1^T} \quad (3)$$

In Equation (3), $\mathrm{Q_1}$ and Λ_1 are the matrices of the eigenvectors and eigenvalues derived by the spectral decomposition of the variance-covariance matrix (e.g. B). Matrix $\mathrm{Q_2}$ is obtained through the following equation:

$$\Lambda_2 = \mathrm{Q_2 \Gamma_{P_{PCA}}(\Delta) Q_2^T} \quad (4)$$

in which $\Gamma_{P_{PCA}}(\Delta)$ is the experimental variogram matrix at a short lag distance (e.g. lag Δ) for PCA factors calculated from:

$$\mathrm{P_{PCA}(u)} = \Lambda_1^{-1/2} \mathrm{Q_1^T Y(u)} \quad (5)$$

Obtaining a reliable value for the experimental variogram at a short lag distance is not an easy task to implement. A practical approach is to test different delta values and select the one that provides the best de-correlation stable decomposition results for factors [24].

Considering the de-correlation of the original cross-correlated attributes, the whole procedure for simulation of correlated variables using data-driven MAF, as explained by Dimitrakopoulos and Fonseca [24], is as follows:

1. Normalize input attributes to variables with zero mean and unit variance (e.g. normal score transformation);
2. Transform normalized attributes to new de-correlated variables using data-driven MAF;
3. Convert each factor to Gaussian variables;
4. Calculate the experimental variograms and fit variogram models for each factor in normal score space;
5. Simulate each normalized factor independently conditional to the data values;
6. Back-transform the simulated variables into MAF space and validate the simulation results;
7. Back-transform the simulated MAF to data space and validate the results;

A visual assessment of cross-semi-variograms of each MAF was used as a criterion to detect any remnant correlation among factors. Perfect de-correlation is obtained when all of the MAFs are around zero for all lag distances [25].

2.2. Field and laboratory analysis

Located in an arid zone, the studied region, with an area of 4416 Km^2, is one of the driest provinces of Iran with a long-term average annual precipitation of about 152 mm [26]. This region is bordered by KavirNamak, in which Shahrood and Damghan are the main capital cities with populations of over 200,000 and 57000 people, respectively. The main source of water is groundwater, which is abstracted through wells and Qanats (some ancient underground tunnels for extraction of water in dry parts of Iran). In this field, the average annual groundwater consumption in Shahrood, for instance, was about 545 million cubic meters [27]. Heavy metal pollution of soil has been attributed to the prevalent mining activities and geological formations in the previous research works [e.g. 28-29]. For example, the mean values for Co, Cr, Mn, and Ni in the soil samples considered by HajizadehNamaghi et al. [29] were 58.78, 1025.8, 903.7, and 902.3 mg/kg in the studid area, respectively. Mining activity is common in the

region [30], and there are many active and abandoned mines by which heavy metals are dispersed through disposal of mine tailings. The location of some of these mines has been given in Figure 1. Due to these activities, most of the soils in the area are contaminated by heavy metals, resulting in their subsequent transfer to plants and groundwater resources [31-32].

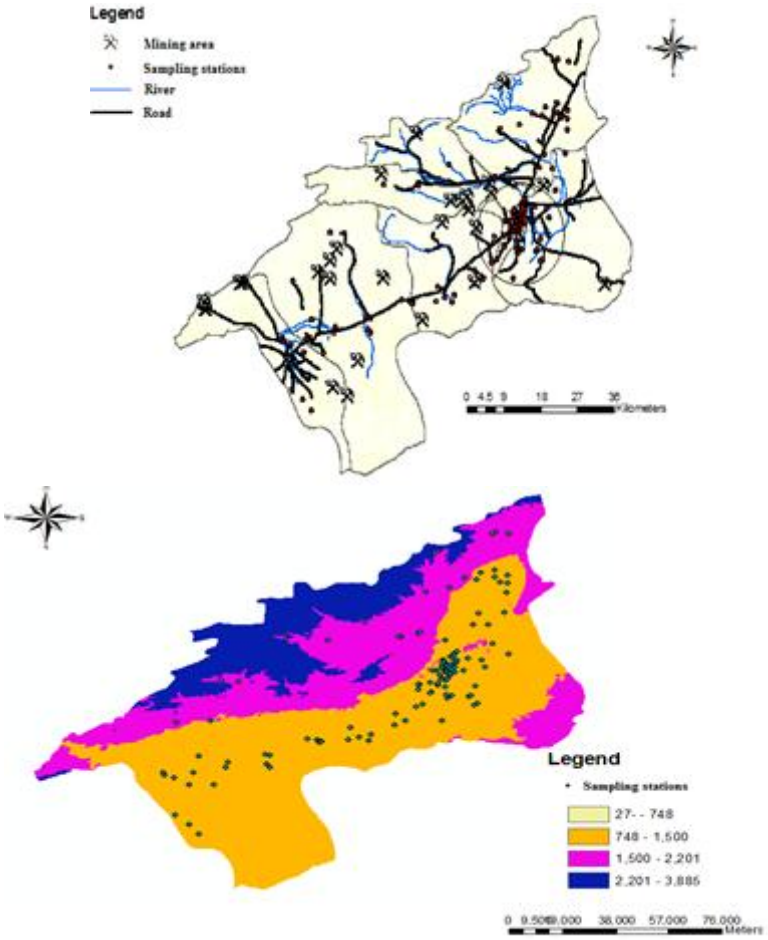

Figure 1. An illustrated map of studied area (above) and a DEM map of region (below). In highlighted region in map of studied area, agricultural activity is prevalent.

For field sampling, a systematic random approach [33] was followed, through which 170 soil samples were taken from upper 10 cm of soil. In laboratory, collected soil samples were air-dried and sieved through a 2-mm stainless steel mesh to remove stones and plant roots. Following digestion of soil samples with nitric acid (HNO_3) and hydrochloric acid (HCl) in a ratio of 3:1 (HNO_3:HCl), total concentrations of trace elements including Co, Pb, Ni, Cs, Cu, Mn, Sr, V, Zn, Cr, and Tl were analyzed by inductively coupled plasma-optical emission spectroscopy (ICP-OES).

2.3. Simulation of heavy metals

The attributes to be simulated should comply with the multi-Gaussian assumption. According to this assumption, the distribution of any value Y(x) conditioned by sampled values is still Gaussian, which is known as Gaussian anamorphosis [34]. In this respect, the attribute under study is viewed as a realization of a random function Z(x) that can be transformed into a Gaussian random field Y(x) with zero mean and unit variance:

$$Z(x) = \phi\left[Y(x)\right] \quad (6)$$

A raw variable can be converted to a Gaussian variable using the inverse of Gaussian anamorphosis function, as follows:

$$Y(x) = \phi^{-1}\left[Z(x)\right] \quad (7)$$

In this research work, Gaussian anamorphosis functions were applied to transform all the attributes to Gaussian variables before simulation.

2.4. Dimensionality reduction by principal component analysis

The distribution and spatial continuity of Co, Pb, Ni, Cs, Cu, Mn, Sr, V, Zn, Cr, and Tl were observed through exploratory data analysis and calculation of experimental variograms, respectively. The diversity of data distribution and spatial continuity were factors that could have made the MAF transformation and the subsequent simulation of all of these variables impractical. Therefore, it was decided to reduce the number of data through principal component analysis (PCA). It has been used earlier by researchers for variable selection [e.g. 35-36]. PCA based on factor analysis was applied on the normal score transformed variables since this method is more robust when the variables follow a Gaussian distribution [25].

2.5. MAF transformation and variography of MAF

Due to the fact that the data had been sampled preferentially in some parts of the studied area (Figure 1), it was first de-clustered using a moving window of 2500 × 2500 m. Experimental variograms were worked out for each six heavy metals (Co, Ni, Cu, Mn, V, and Cr) retained by PCA (refer to section 3.1). The variation in lag sizes and number of lags was based on the general rule of thumb that the product of lag size and number of lags should be less than one half of the largest distance between the data pairs [37]. Accordingly, the value of 2560 m was selected for the lag size with 10 numbers of lags. To consider the possible anisotropy (geometric or zonal) in spatial pattern of the attributes, the calculation of variograms was done in four main directions (e.g. 0, 45, 90, and 135˚) to assess if the spatial variability is the same in all directions [6]. If it is, the distribution is called isotropic; otherwise, it is anisotropic. An example of these calculated variograms is illustrated in Figure 2 for manganese. In this respect, no single direction of greatest continuity was observed given the results of directional variograms, and so isotropy was assumed. The variogram plot was then fitted with a theoretical model for which the best fitting model was selected based on two-fold cross-validation. For this purpose, 70% of the sampling points were used as the training set and the rest as the test set. The criteria for the selection of a model were according to the root mean squared error (RMSE) and root mean square standardized error (RMSSE) of the test set that should be as low as possible and near one for a suitable model, respectively. Also the plot of measured and observed values of the test data was another goodness of the fit criterion. The validation plot of manages, as an example, is given in Figure 3.

The transformation matrix (Equation 3) was used to convert six normal attributes to six MAFs. Derivation of MAFs was implemented using a distance of 1500 m and tolerance on distance of 250 m for de-correlation at short lags (e.g. delta value), selected based on the criterion explained earlier.

Variography on each MAF was then performed independently and the models were fitted to each experimental variogram. Since normal distribution is required for geo-statistical simulation, distribution of MAFs was reconsidered and the required normal transformation was used in case of necessity.

A simulation grid with a 1000 m × 1000 m dimension was then superimposed on normal score transformed MAFs before simulation of factors.

2.6. Conditional sequential Gaussian simulation of MAF

There are multiple conditional simulation methods available in the literature. However, the most applied ones are Turning Bands [38] and Sequential Gaussian Simulation (SGS) [39]. Turning bands algorithm was once used to be popular but due to its disadvantages, it is now less popular among researchers [40]. SGS, in contrast, is more efficient and widely used [41]. The details of the latest method can be found in other published literatures [37]. In short, in this technique, a random path is defined through all the grid nodes containing the conditional data that comprise the original data and the previously simulated values within the defined neighborhood of the simulated point. Simple kriging (most commonly) will then be used to obtain a local conditional distribution. A value is then drawn randomly from normal distribution and added to the nodes in random path. This process is continued by simulation of the next node and so on [10]. Twenty simulations, within a neighborhood circle of 10000 m radius, were generated in this study for each simulated heavy metal, and validated for reproduction of variogram of data in MAF space. It should be

noted, despite kriging methods, that in these kinds of simulations, no cross-validation is performed on the results of simulations because honoring the data values results in a mean prediction error of zero.

The realizations of MAF were then transformed back to normal score space by multiplying a column vector of simulated MAF in each grid node with the corresponding inverse matrix of the MAF loadings [24], and finally, the results were transformed to data space, respectively. Validation of the results of simulations was done according to the comparison of the histogram and experimental semi-variogam of simulation realizations with that of original data to ensure reproduction of the spatial pattern characteristics [42]. In addition, reproduction of the variogram of the simulated MAFs in normal score space and comparison of the distribution of simulated MAFs versus original MAFs through quantile-quantileplots (QQ plots) were also applied for the validation of the results.

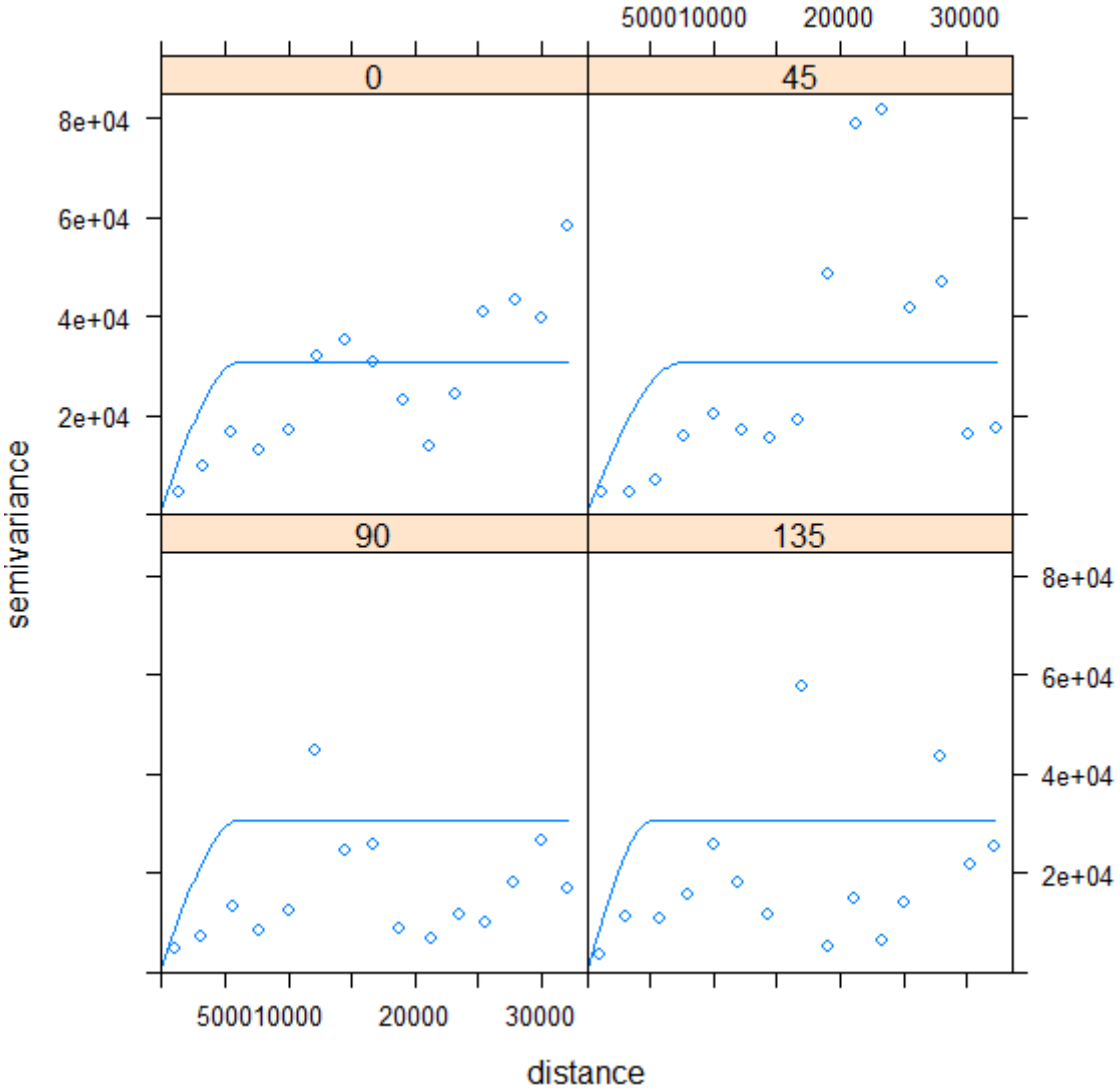

Figure 2. Directional variogram of Mn in four main directions (e.g. 0, 45, 90, and 135°).

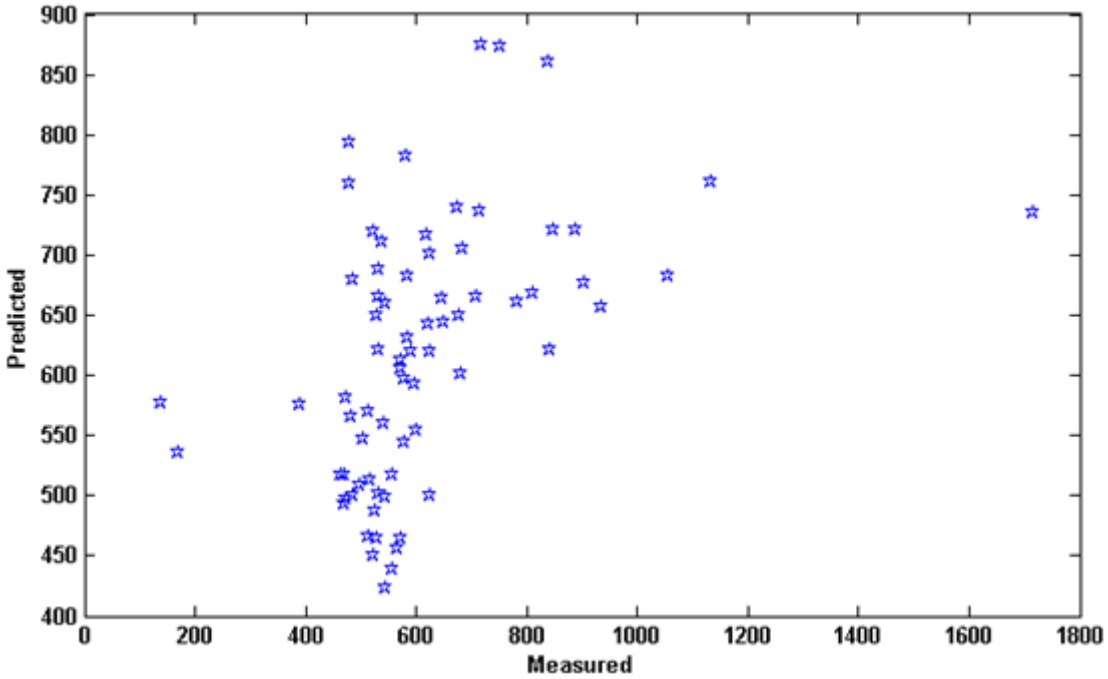

Figure 3. Validation plot of variogram model for Mn. Produced RMSE and RMSSE values for applied model were 185 and 0.99, respectively.

3. Results and discussion

3.1. Dimensionality reduction

The descriptive statistics related to the analyzed elements along with the standard levels for some HMs as well as the number of samples exceeding the cut-off values are rendered in Table 1.

Each HM was first normally transformed by the Gaussian anamorphosis modeling with respect to the de-clustered weight, resulting in normal score values. The theoretical mean and variance of the actual values were compared with that of the transformed data to optimize the number of polynomials used in each Gaussian transformation. In this respect, high cross-correlation among some HMs and poor spatial correlation among others could have made spectral decomposition of MAFs unstable, resulting in some negative eigenvalues [11]. Thus it was decided to reduce the dimensionality of the original variables. For this purpose, a factor analysis along with principal component extraction and Quartimax rotation method (since it resulted in the best rotation of components compared with that of other algorithms) were applied on the normal score factors. The resultant Kaiser–Meyer–Olkin (KMO) test was 0.744, implying the suitability of the factors, considering the proposed cut-off value of 0.5 [43]. Moreover, the generated Chi-square distribution (χ^2) of Bartlett's Test of Sphericity was high (1200) and significant, showing the existence of a common factor between relevant matrices of the parent population [43]. As a whole, three rotated components accounting for 72.02% of the total variance of data were retained based on the criterion of eigenvalues greater than one [44]. The respective eigenvalues of these three components were 4.19, 1.61, and 1.06, while they encompassed 38.06, 14.63, and 9.63 percents of the total variance, respectively. The results of the final rotated component matrix are presented in Table 2.

Considering this table, the importance of each attribute to factors, independent of the other variables, can be considered using factor loadings [25]. Since most of the variabilities of Co, Ni, Cu, Mn, V, and Cr can be explained by factor 1 while other HMs are explained by other factors, these variables were assigned to an independent group and retained for further spatial analysis. The correlation coefficients between the original reduced variables (upper diagonal) versus those of normal score attributes (lower diagonal) are given in Table 3.

Since the normal score transformation is non-linear, the cross-correlation among attributes may not be exactly reproduced [22]. This transformation is performed based on the association of the quantiles of each HM with that of standard normal distribution [42]. It should be noted that due to the non-linearity aspect of this transformation, it does not have any impact on the

subsequent MAF transformation [11]. Some researchers [22] believe that if the correlation coefficients of element before and after normalization are not significantly influenced, their correlation are better projected out in the final results of simulations. The results obtained show that the greatest fluctuations in the amount of correlation coefficients before and after normal score transformation are related to Cu-Mn (from 0.21 to 0.42), Cu-V (from 0.49 to 0.67), and Mn-Ni (from 0.37 to 0.57). There have also been some other changes; however, they are at the same order of magnitude. In cases in which the correlation coefficients have been significantly influenced, the direct method of MAFs [18] can be utilized but in this study, it is not a big issue.

Histograms of original data next to those of transformed attributes are depicted in Figures 4 and 5, respectively.

These histograms were created using de-clustering weights. Given these histograms, some variables (e.g. Co, Mn, Cu) are positively skewed and the others are far from normal distribution (Figure 4). The transformed variables, on the other hand, all have roughly a mean of zero and St. dev. of one, showing a quite standard normal distribution.

Table1. Descriptive statistics, standard levels, and number of samples exceeding cut-off values for analyzed samples.

Heavy metals	Min	Max	Average	SD	Standard level	Reference	Number exceeding threshold
Co	2.53	25.24	10.34	3.15	40	IS	
Pb	5.22	87.74	18.60	9.24	50	IS	2(1.4%)
Ni	11.50	55.81	32.57	7.48	50	IS	4(2.4%)
Cs	1.40	10.47	5.91	1.30			
Cu	6.99	85.22	23.84	10.12	100	IS	
Mn	138.00	1716.21	585.84	159.45	600	USEPA,1983	57(33.5%)
Sr	116.22	970.56	379.11	119.39			
V	22.55	160.96	78.44	20.28	100	IS	21(12.4%)
Zn	14.07	210.49	77.65	28.56	200	IS	2(1.4%)
Cr	27.50	179.79	84.17	22.89	110	IS	18(10.6%)
Tl	0.00	2.16	0.82	0.36	5	IS	

Is: Iranian soil standards.

Table 2. Rotated components matrix of normal score variables. Ana prefix indicates application of anamorphosis function. High loading values are highlighted using bold fonts.

Normal score variables	1	2	3
Ana_Co	**.933**	.118	-.126
Ana_Pb	.214	**.960**	-.058
Ana_Ni	**.830**	.045	.005
Ana_Cs	.421	-.086	.032
Ana_Cu	**.601**	.312	.135
Ana_Mn	**.662**	.150	.059
Ana_Sr	.016	-.022	**.996**
Ana_V	**.954**	.075	.048
Ana_Zn	.382	**.722**	.069
Ana_Cr	**.592**	.015	.127
Ana_Tl	.444	.136	-.051

Table 3. Correlation coefficients between original variables (upper diagonal) and normal score variables (lower diagonal).

	Co	Ni	Cu	Mn	V	Cr
Co	1	0.65	0.39	0.63	0.89	0.43
Ni	0.73	1	0.46	0.37	0.69	0.72
Cu	0.51	0.59	1	0.21	0.49	0.28
Mn	0.66	0.57	0.42	1	0.61	0.28
V	0.87	0.73	0.67	0.67	1	0.53
Cr	0.52	0.71	0.39	0.42	0.58	1

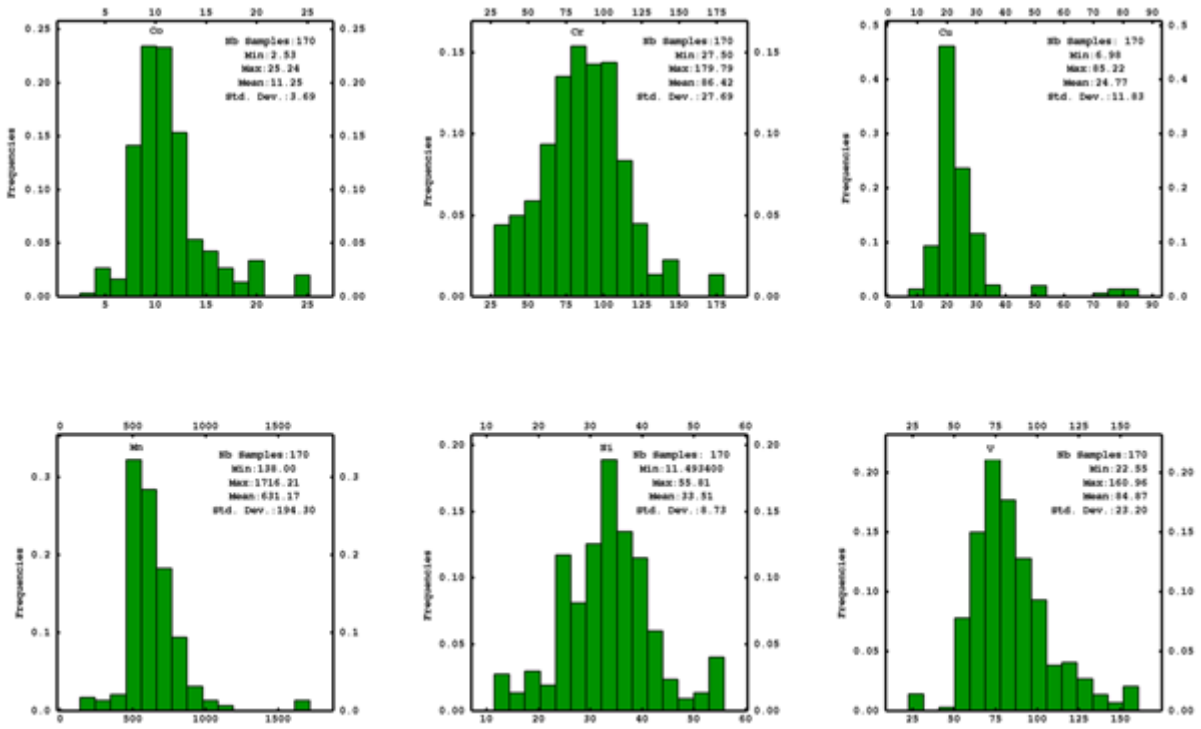

Figure 4. Histograms of original heavy metals.

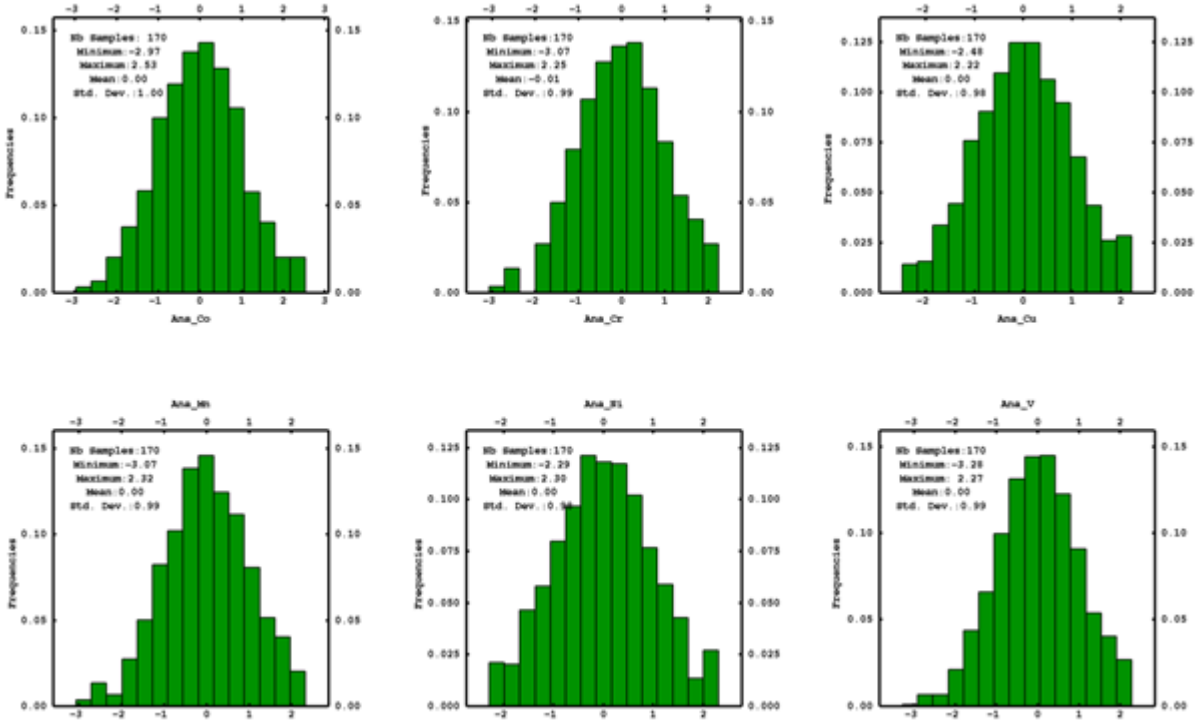

Figure 5. Histograms of normal score variables.

3.2. Spatial relationships using variograms

The experimental variograms were calculated for each retained attribute to identify the degree of spatial continuity of each HM and establish the range of spatial dependence. The variograms were all fitted to spherical models other than Co and Cr, for which, cubic and exponential models were better fitted. The parameters related to the fitted semi-variogram models can be found in Table 4, whereas the fitted models are illustrated in Figure 6 as well.

Table 4. Parameters of variogram models for each HM.

Heavy metals	Variogram model	C_0 (mg/kg)^2	$C+C_0$ (mg/kg)^2	$C_0/C+C_0$	Range(km)
Co	Cubic	2.82	16.36	0.17	54.84
Cr	Exponential	173.20	499.80	0.35	23.71
Cu	Spherical	15.92	137.20	0.12	13.66
Mn	Spherical	700.00	34900	0.02	7.19
Ni	Spherical	16.82	48.74	0.35	15.88
V	Spherical	78.54	499.90	0.16	29.85

C_0: Nugget variance
$C+C_0$: Sill
C: Structural variance

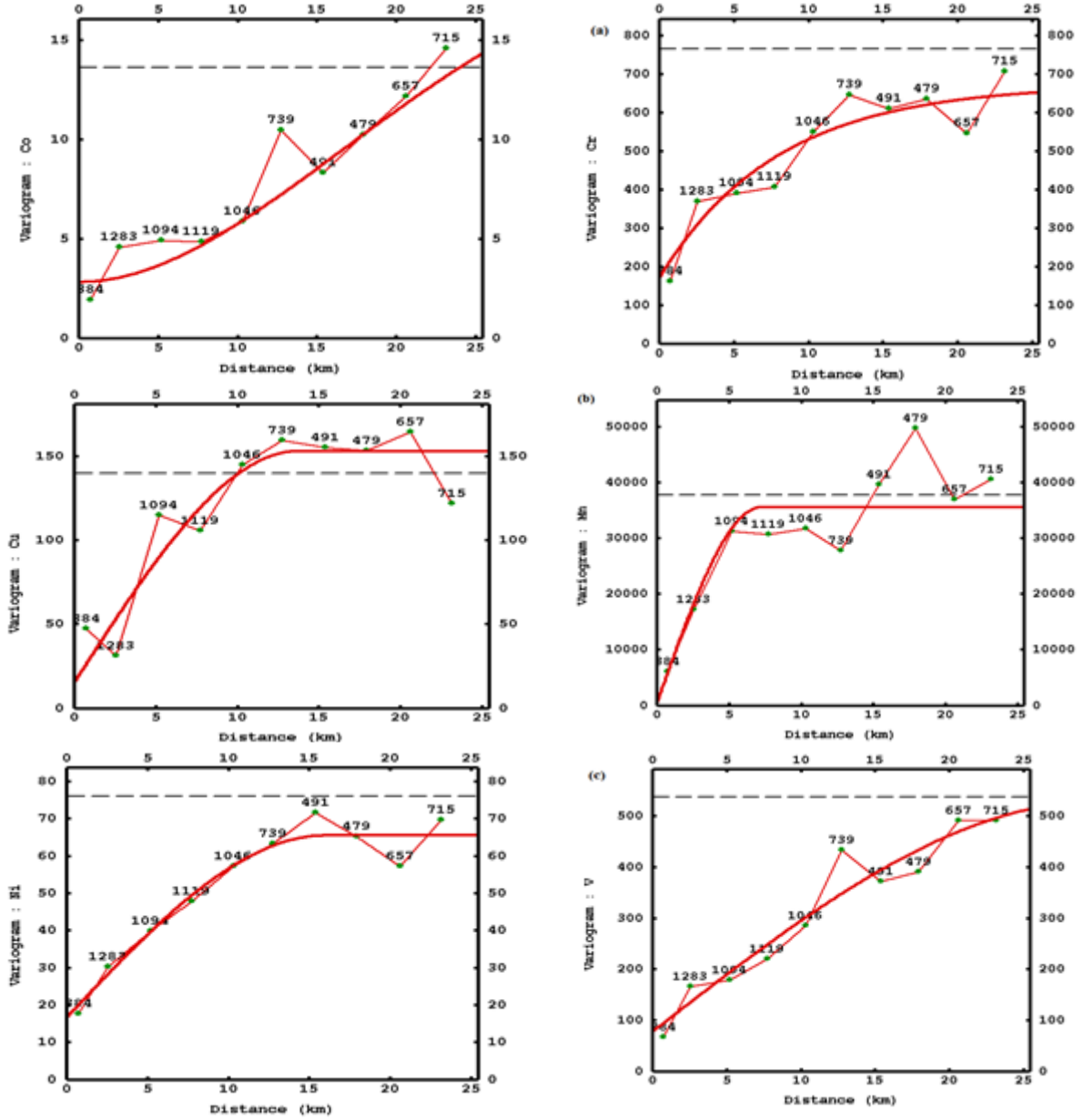

Figure 6. Variogram models fitted to each semi-variogram.

In Table 4, the nugget effect (C_0) indicates the semi-variance value found at the intercept with the y-axis that should be zero for a perfect model. However, a large nugget effect can occur due to sampling error or spatial dependence at finer spatial scales than the sampling resolution [45]. In fact, a variable with no spatial dependence would have a variogram that is a pure nugget. Referring to this table, it is clear that Mn and Cr had high nugget values of 700 and 173.20 (mg/kg)^2, implying that the sampling method followed was not intensive enough to cover a very high

micro-scale variability of these elements, and possibly, more samples should have been collected in the field. The lowest recorded nugget values were related to Co and Cu with levels of 2.82 and 15.92 $(mg/kg)^2$, showing that the sampling method has successfully covered the spatial dependence of these trace elements compared with that of the other metals.

The spatial dependence of Soil HMs can be classified using the criterion of sill/nugget [46]. In this respect, attributes with strong spatial dependence have a ratio less than 25%, whereas for attributes in which the spatial dependence is moderate, this range fluctuates between 25% and 75% and those greater than 75% belong to attributes with weak spatial dependence [47-48]. As a whole, both intrinsic factors (e.g. soil formation factors and parent materials) and extrinsic factors (e.g. anthropogenic sources such as agricultural or industrial activities) may affect the spatial variability of soil properties. Meanwhile, strong spatial dependence can be originated from intrinsic factors, while weak spatial dependence can be attributed to extrinsic factors [49]. Referring to Table 3, it can be concluded that except for Cr and Ni, which can be categorized in the moderate spatial class (both have a nugget/sill ratio of 35%), the spatial dependence of other HMs is strong. In other words, the sources of Cr and Ni have partially emanated from parent material but the anthropogenic source has also played a role in this field. In this field, the high correlation coefficients between Cr and Ni (Table3) corroborate the fact that they have the same origin. Lilli et al. [50] and Cai et al. [51] reached the same conclusion and found a similar source for Cr and Ni on their studies about soils of Asopos Basin in Greece and Shunde in Southeast China, respectively. Although in this work the range value for Cr (23.71 km) was higher than that for Ni (15.88 km), meaning that the variability of chromium was spatially correlated over a larger range than nickel. This may suggest that extrinsic factors (such as mining and agricultural activities in the area) might have weakened the spatial correlation of Ni in some parts of the region. Considering other HMs, the condition for Co and V is the same (e.g. they have roughly similar nugget to sill ratios) but the range value of cobalt is roughly twice that of vanadium. It implies that geogenic sources (e.g. geological formations and soil properties) have contributed to the distribution of these HMs in the soils of the studied area. Natural sources (soil texture and pedogenesis) have been found to be the main contributors of Co in soil in some earlier studies as well [52].

Among the considered HMs, Mn had the strongest spatial dependence (nugget to sill ratio of 0.02) and the lowest range (7.19km) followed by Cu($(C_0)/(C+C_0)$ = 0.12 and range of 13.66 km), indicating that they can be attributed solely to geogenic sources. With respect to trace elements in the lithosphere, manganese is one of the most abundant ones. Since Mn is mainly accumulated in the top horizons of soil as a result of its fixation by organic matter [53], its subsequent transport in the lithosphere may be limited, resulting in its accumulation next to the source of production. This may partly justify the strong spatial dependence and low range of this element.

3.3. Data-Driven MAF

After testing different separation distances of Δ, based on the average distance among sampling points in the field, a value of 1500 m was selected for Δ as it assured a suitable de-correlation of factors and stable decomposition of MAF [24].

The variance-covariance matrix of data at zero lag distance (e.g. $B = Q_1^T \Lambda_1 Q_1$) was obtained as:

$$B = \begin{pmatrix} 1 & 0.519 & 0.509 & 0.664 & 0.734 & 0.870 \\ 0.519 & 1 & 0.395 & 0.423 & 0.713 & 0.585 \\ 0.509 & 0.395 & 1 & 0.425 & 0.592 & 0.672 \\ 0.664 & 0.423 & 0.425 & 1 & 0.571 & 0.670 \\ 0.734 & 0.713 & 0.592 & 0.571 & 1 & 0.735 \\ 0.870 & 0.585 & 0.572 & 0.670 & 0.735 & 1 \end{pmatrix}$$

The spectral decomposition of the above matrix makes eigenvectors (Q1) and eigenvalues (Λ_1) of B as follows:

$$Q_1 = \begin{pmatrix} 0.440 & 0.215 & 0.141 & 0.624 & -0.535 & 0.063 \\ 0.365 & -0.775 & 0.323 & 0.188 & 0.192 & 0.411 \\ 0.360 & 0.220 & -0.899 & 0.244 & 0.296 & 0.064 \\ 0.380 & 0.447 & 0.251 & 0.009 & 0.663 & -0.043 \\ 0.442 & -0.287 & 0.008 & -0.291 & -0.019 & -0.820 \\ 0.453 & 0.149 & 0.070 & -0.656 & -0.388 & 0.387 \end{pmatrix}$$

$$\Lambda_1 = \begin{pmatrix} 4.047 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.652 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.602 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.109 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.367 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.224 \end{pmatrix}$$

The associated transformation matrix (A) is used to convert six HMs to their respective min/max autocorrelation factors:

$$A = \begin{pmatrix} 0.914 & 1.216 & -1.160 & 0.607 & -0.924 & 0521 \\ -0.497 & 0.027 & -1.320 & -0.137 & 0.274 & -0.306 \\ 0.173 & 0.186 & -0.261 & 0.110 & -0.856 & -1.114 \\ 0.143 & -1.158 & -0.188 & 0.720 & -0.114 & -0.023 \\ -0.990 & 0.298 & 1.393 & 0.713 & 0.040 & 0.397 \\ 0.362 & -0.389 & 1.115 & -1.081 & 1.942 & -0.194 \end{pmatrix}$$

Experimental cross-variograms between MAFs (in supplementary material) indicated that MAFs have been rigorously de-correlated and the maximum cross-variogram values in most of the plots is around 0.1. However, values up to 0.3 have also been observed at longer lag distances. This small amount of remnant correlation may be due to the different number of pairs used to calculate variograms in MAF transformation matrix [11], as it can be observed in Figure 6.

In the majority of plots, MAFs have become completely de-correlated at short lags and around the separation distance of Δ. Although the de-correlation step was satisfactory, MAFs did not follow the standard normal assumption required by SGS. Therefore, a second normal score transformation was performed on MAFs to convert MAFs to normal score values with standard normal distribution, and then the variogram models were fitted to each normal score transformed MAF followed by simulation of the resultant factors by SGS. The applied isotropic models consisted of a nugget structure and a spherical structure. Twenty realizations of MAFs were generated in normal score space, which then back-transformed to MAF space. There were some validation checks to examine the quality of SGS [43]. Given in Figure 7, the cumulative variograms of 20 simulated realizations in normal score space overlaid with the respective fitted variogram models of MAFs for each factor, separately. As a whole, the variograms have all been well reproduced but it seems as if the best results have been obtained for factor#5 and factor#6 as the fitted variograms are close to the mean of realizations.

Referring to other variograms, the mean of realizations have roughly underestimated the fitted variogram. However, the predictions are reasonable anyhow. The other validation check consisted of comparing the distribution of the first realization in the simulated MAFs (in MAF space) versus the distribution of the original MAF by q-q plots. Q-Q plots in geo-statistics are generated by dividing a population into equal numbers of values and averaging the values inside each bin. In case of matching the quantile values of two populations, the distribution of points lays along the bisector (e.g. X=Y) line. The close match of these two distributions confirms the success of simulations too (Figure 8). The minor upward deviation of some factors (such as Factor#1 and Factor#5) from the line of unit slop can be attributed to the back-transformation process from normal score to MAF space.

The realizations of MAF generated during SGS were back-transformed to normal score variables by multiplying columns of MAF with the inverse matrix of transformation matrix (A). The normal score variables were then transformed to data space to produce Co, Cr, Cu, Mn, Ni, and V from the related factors. There were some validations to examine the quality of back-transformation of HMs at this stage of SGS. The first test was reproduction of the descriptive statistics (including min, max, mean, standard deviation) and histogram of the original attributes from simulated values. For each variable, a realization was selected randomly out of 20 simulated realizations [55]. The respective minimum and maximum values of original data were reproduced well (Figure 9) but except for Ni in which the mean value of original nickel has been slightly overestimated, the mean and st. dev. of other HMs a little exceeded that of simulated values. The histograms of the simulated HMs have also roughly followed those of the original variables as well.

Reproduction of the target variograms was considered by comparison of the experimental variogram of simulated attribute against that of original HM (Figure10). With respect to this Figure, the variograms of simulated values have not been exactly reproduced, and there are some underestimation in variogram reproduction but the spatial continuity of original attributes have been captured all in all. This may be due to the fact that, in this work, soil samples have been preferentially taken in some parts of the studied region (e.g. mainly clustered in the agricultural land use), thus complicating the exact honoring of spatial variability of attributes. In addition, back-transformation from MAFs to normal score and its subsequent conversion to original values might have also contributed in this respect. To sum up, considering the simulation results, the common smoothing problem of kriging methods resulting in overestimation of small values and underestimation of large values [55] have been obviated through this method.

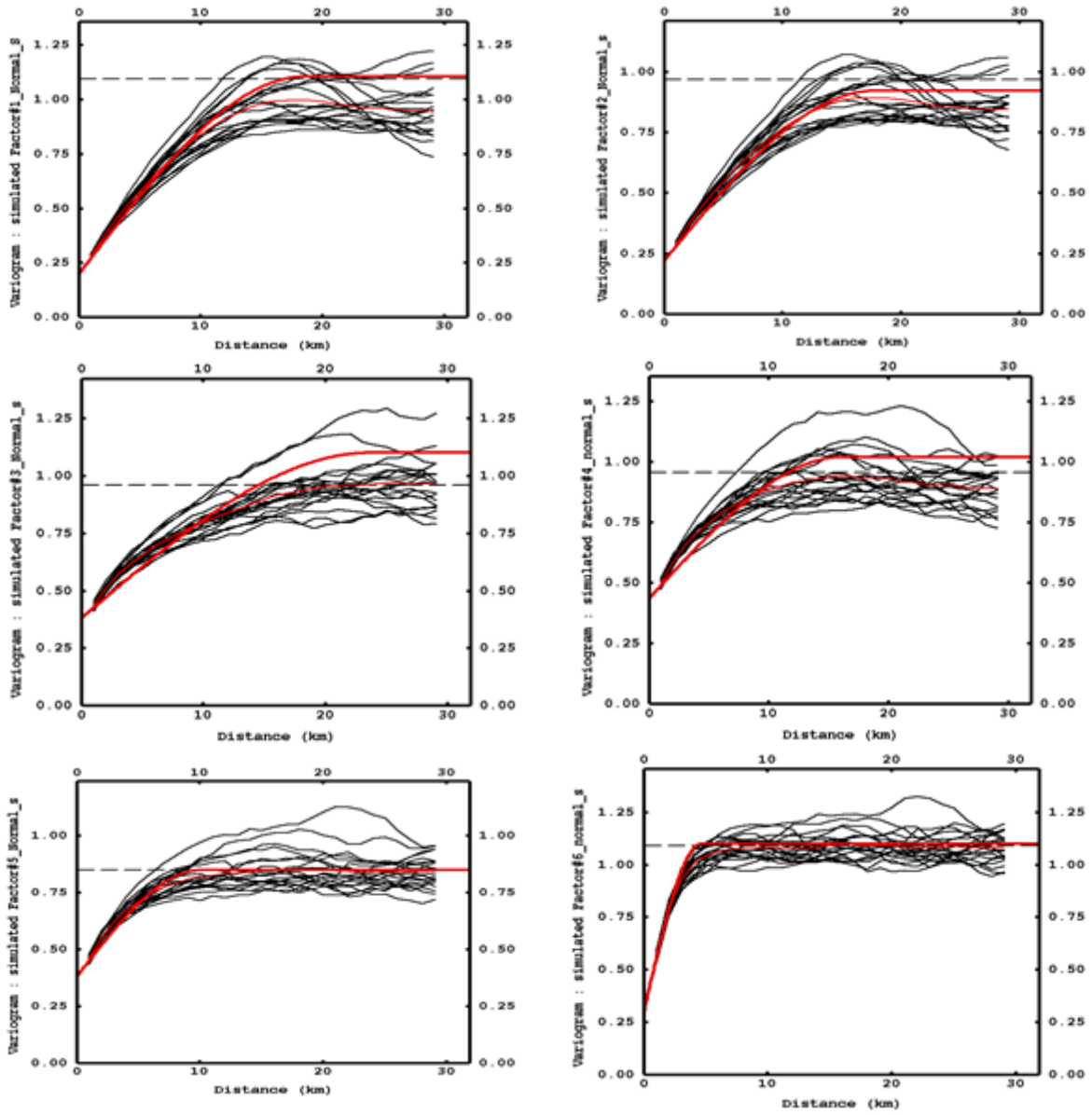

Figure 7. Variogram models of 20 simulated MAFs overlaid with original MAF model in normal score space. (Highlighted red variogram is target variogram and thin red variogram is mean of realizations).

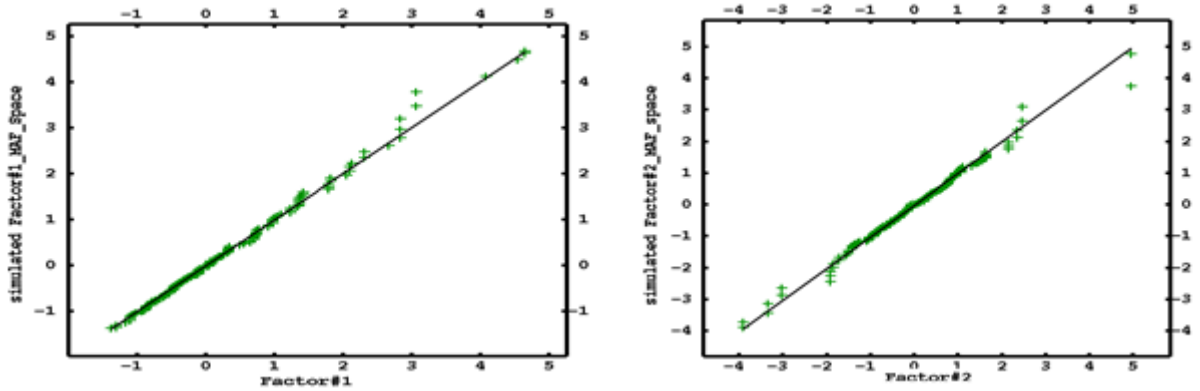

Figure 8. Q-Q plot of simulated MAF against original MAF in MAF space.

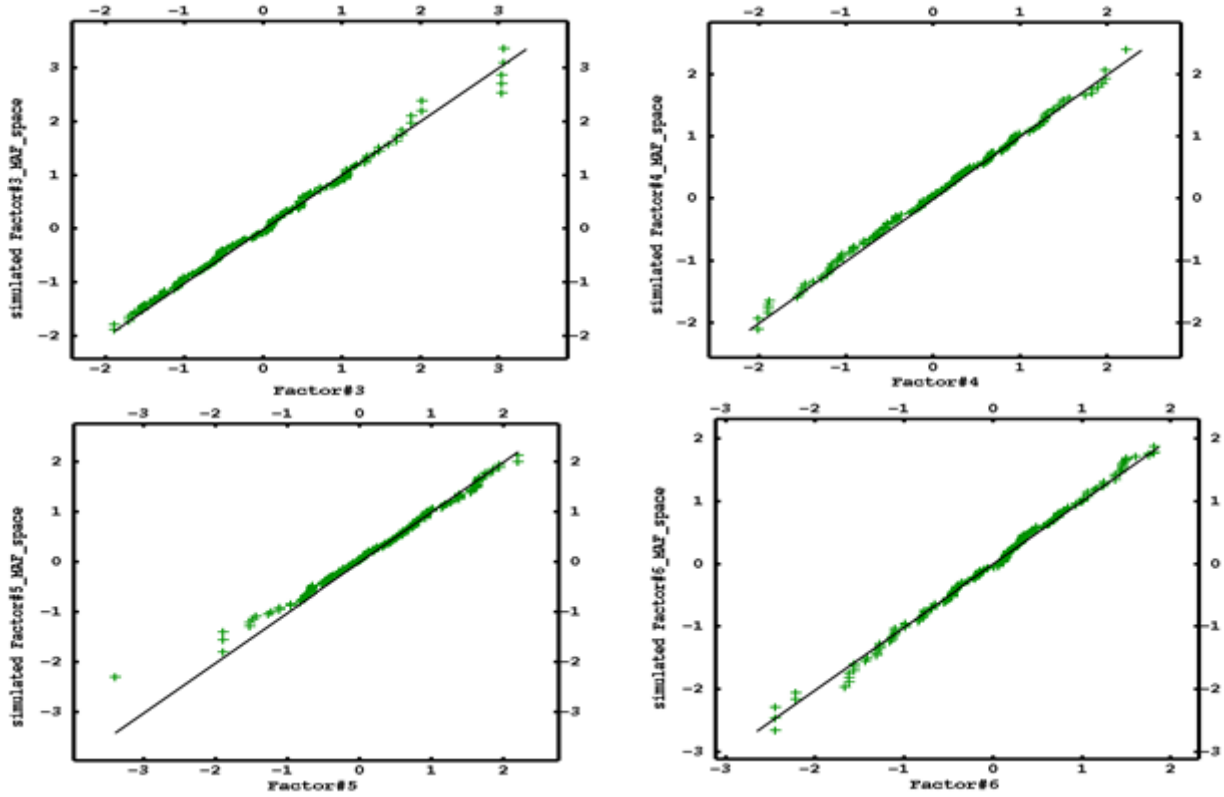

Figure 8. Continued.

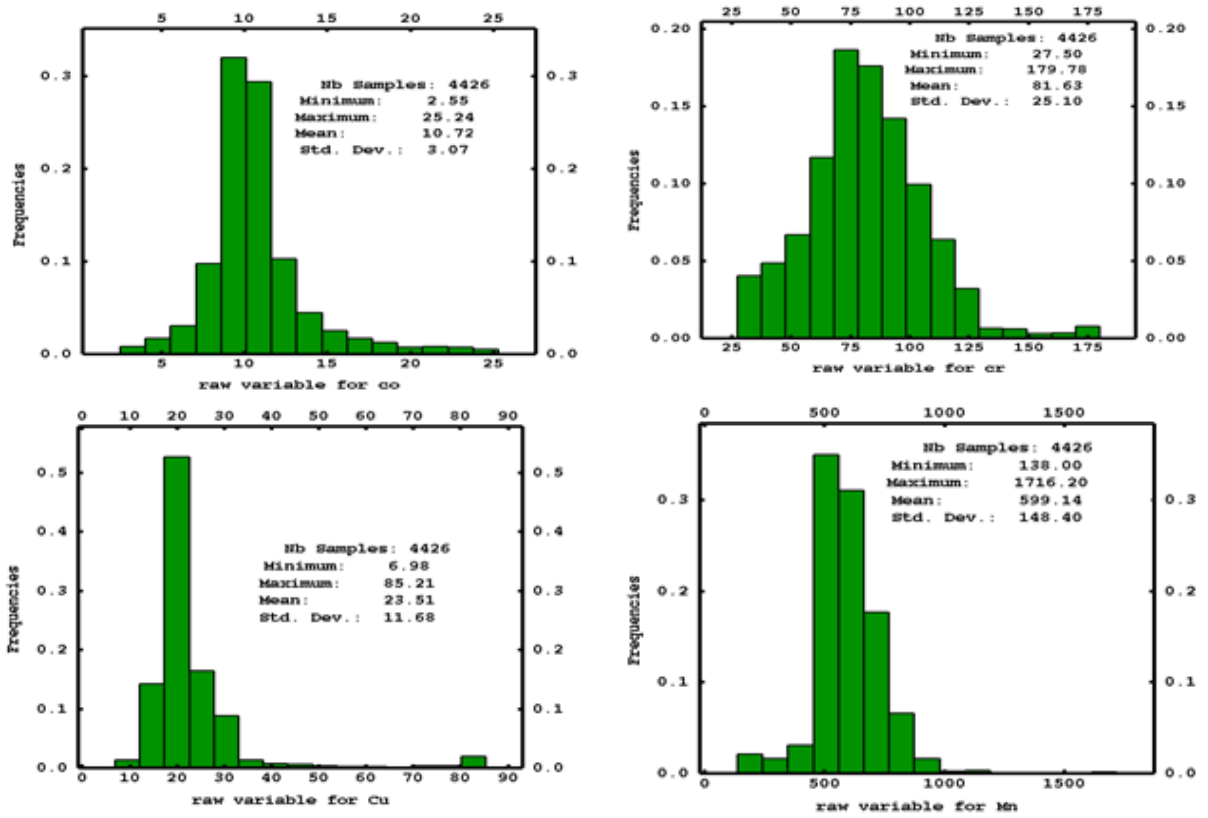

Figure 9. Histogram of simulated values reproduced from SGS along with MAF de-correlation.

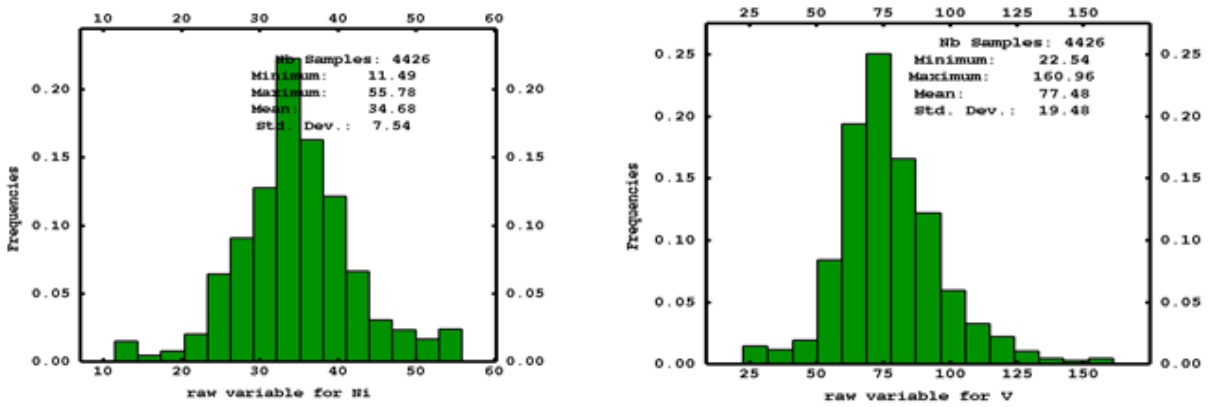

Figure 9. Continued.

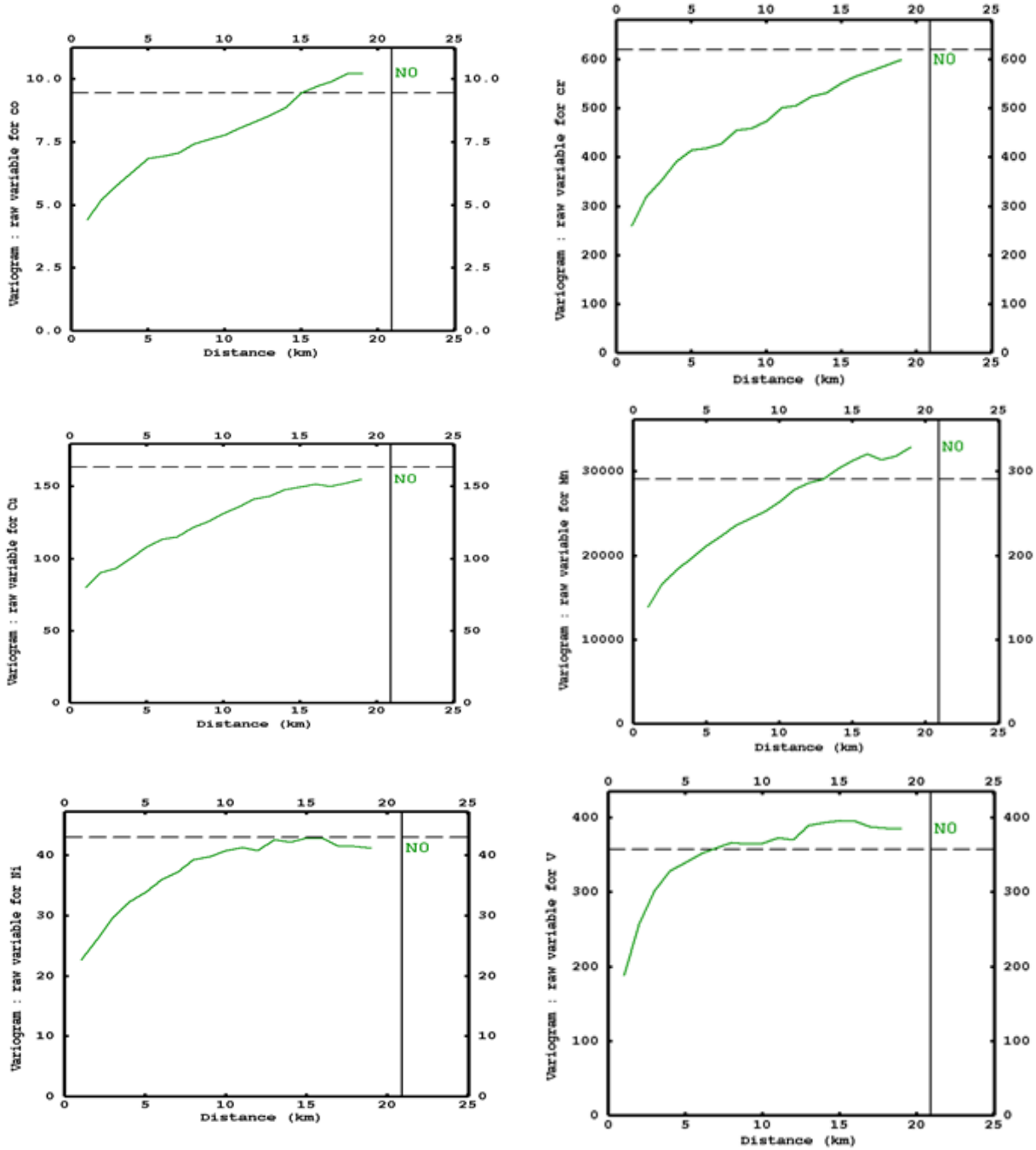

Figure 10. Reproduction of experimental variograms from a simulated realization selected at random.

Finally, the uncertainty in simulation results was estimated through considering 3 out of 20 produced realizations of chromium picked randomly (as an example) (Figure11). Given the simulation maps, there are some uncertainties associated with the predications at high values of chromium. However, these uncertainties are not at a level to influence the whole area, and the simulation maps are acceptable as a whole.

The de-correlation method used in the current research work is simpler and more computationally efficient than its equivalent sequential Gaussian co-simulation as fitting a linear model of coregionalization (LMC), to account for the correlation coefficients among attributes, is not required in the data-driven MAF method. This note has been emphasized in similar research works in different fields of studies [54]. Therefore, this method can be applied by other researchers for production of simulation maps of a contaminated area, while working with multiple cross-correlated pollutants.

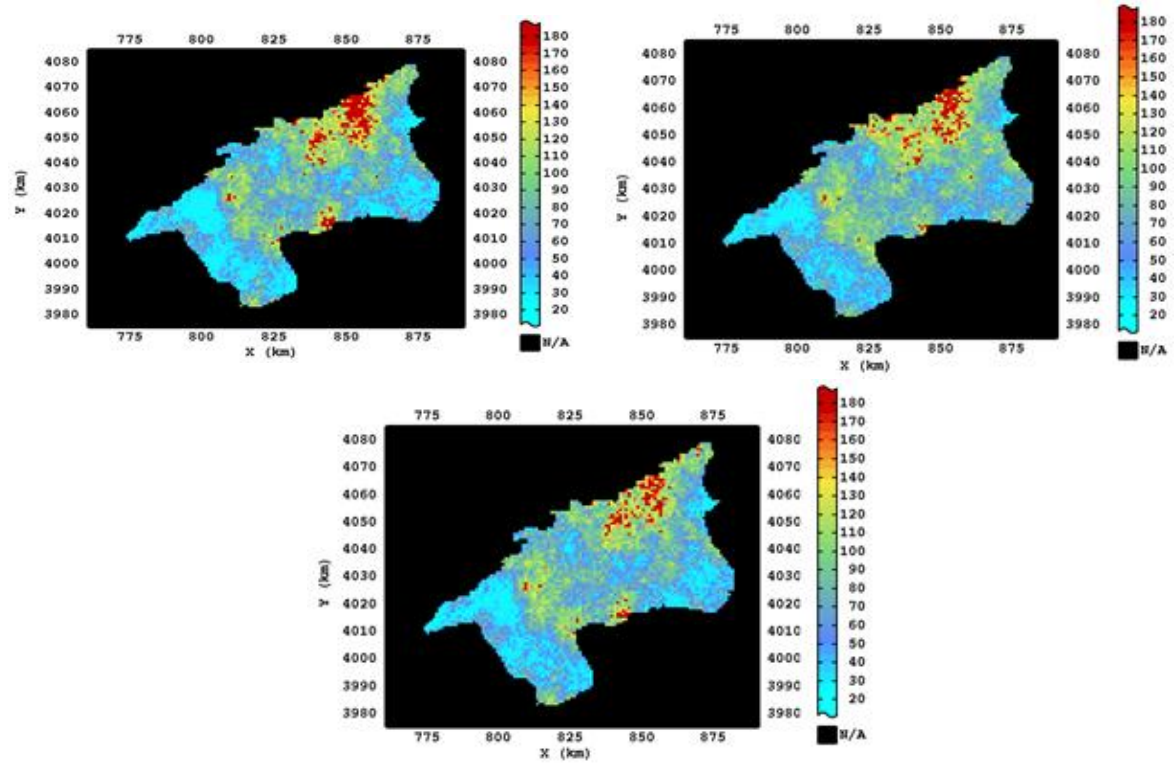

Figure 11. Uncetainty in simulation results of Cr constructed from multiple realizations.

4. Conclusions

Co-simulation of cross-correlated attributes (e.g. in an area in which multiple soil contaminants have been collected) is not an easy task to implement. De-correlation methods have been proposed to handle this kind of data while simulating a high number of correlated data records. In this research work, a method known as min/max autocorrelation factors (MAFs) was applied to de-correlate six HMs, selected out of eleven elements, followed by their subsequent simulation independently using SGS. Despite PCA, this technique assures de-correlation at zero lag as well as Δ lag distance. Due to de-correlation of factors, they can be simulated independently, and so it is less computationally intensive compared with its co-simulation counterpart. The validation of simulation results also confirmed that not only data values are honored but also spatial continuity of attributes was reproduced in this research work. All in all, this method was successful, and reached the objective of this study. The generated simulation results through different realizations facilitate the uncertainty and spatial risk assessment of the contaminated area.

Acknowledgments

The authors are grateful to Geological Survey of Iran for the help in analysis of heavy metals. The financial support of this project was provided by the grant number100-2164 offered by Geological Survey of Iran.

References

[1]. Mirzaei, R. and Sakizadeh, M. (2016). Comparison of interpolation methods for the estimation of groundwater contamination in Andimeshk-Shush

Plain,Southwest of Iran. Environ. Sci. Pollut. Res. 23: 2758-2769.

[2]. Pásztor, L., ZsuzsannaSzabó, K., Szatmári, G., Laborczi, A. and Horváth, A. (2016). Mapping geogenic radon potential by regression kriging. Sci. Total Environ. 544: 883-891.

[3]. Chica-Olmo, M., Luque-Espinar, J.A., Rodriguez-Galiano, V., Pardo-Igúzquiza, E. and Chica-Rivas, L. (2014). Categorical Indicator Kriging for assessing the risk of groundwater nitrate pollution: The case of Vega de Granada aquifer (SE Spain). Sci. Total Environ. 470-471: 229-239.

[4]. Lv, J., Liu, Y., Zhang, Z. and Dai, J. (2013). Factorial kriging and stepwise regression approach to identify environmental factors influencing spatial multi-scale variability of heavy metals in soils. J. Hazard. Mater. 261: 387- 397.

[5]. Lu, A., Wang, J., Qin, X., Wang, K., Han, P. and Zhang, S. (2012). Multivariate and geostatistical analyses of the spatial distribution and origin of heavy metals in the agricultural soils in Shunyi, Beijing, China. Sci. Total. Environ. 425: 66-74.

[6]. Guagliardi, I., Cicchella, D. and De Rosa, R. (2012). A geostatistical approach to assess concentration and spatial distribution of heavy metals in urban soils. Water Air Soil Pollut. 223: 5983-5998.

[7]. Lv, J., Liu, Y., Zhang, Z. and Dai, B. (2014). Multivariate geostatistical analyses of heavy metals in soils: Spatial

multi-scale variations in Wulian, Eastern China.Ecotox. Envirn. Safe. 107: 140-147.

[8]. Chen, T., Chang, Q., Liu, J., Clevers, J.G.P.W. and Kooistra, L. (2016). Identification of soil heavy metal sources and improvement in spatial mapping based on soil spectral information: A case study in northwest China. Sci. Total Environ. 565: 155-164.

[9]. Lopes, J.A., Rosas, C.F., Fernandes, J.B. and Vanzela, G.A. (2011). Risk quantification in grade-tonnage curves and resource categorization in a lateritic nickel deposit using geologically constrained joint conditional simulation. J. Min. Sci. 47 (2): 166-176.

[10]. Goovaerts, P. (1997). Geostatistics for natural resource evaluation. Oxford University Press.

[11]. Desbarats, A.J. and Dimitrakopoulos, R. (2000). Geostatistical simulation of regionalized pore-size distributions using min/max autocorrelation factors. Math. Geol. 32 (8): 919-942.

[12]. DeCaritat, P. and Grunsky, E.C. (2013). Defining element associations and inferring geological processes from total element concentrations in Australian catchment outlet sediments: Multivariate analysis of continental-scale geochemical data. Appl. Geochem. 33: 104-126.

[13]. Hosseini, S.A. and Asghari, O. (2015). Simulation of geometallurgical variables through stepwise conditional transformation in Sungun copper deposit, Iran. Arab. J. Geosci. 8 (6): 3821-3831.

[14]. Mueller, U.A. and Grunsky, E.C. (2016). Multivariate spatial analysis of lake sediment geochemical data; Melville Peninsula, Nunavut, Canada. Appl. Geochem. 75: 247-262.

[15]. Boluwade, A. and Madramootoo, C.A. (2015). Geostatistical independent simulation of spatially correlated soil variables. Comput. Geosci. 85: 3-15.

[16]. Lin, W.C., Lin, Y.P. and Wang, Y.C. (2016). A decision-making approach for delineating sites which are potentially contaminated by heavy metals via joint simulation. Environ. Pollut. 211: 98-110.

[17]. Da Silva, C. and Costa, J. (2014). Minimum/maximum autocorrelation factors applied to grade estimation. Stoch. Environ. Res. Risk Assess. 28: 1929-1938.

[18]. Bandarian, E., Bloom, L. and Mueller, U. (2008). Direct minimum/maximum autocorrelation factors within the framework of a two structure linear model of coregionalization. Comput. Geosci. 34 (3): 190-200.

[19]. Vargas-Guzmán, J. and Dimitrakopoulos, R. (2003). Computational properties of min/max autocorrelation factors. Comput. Geosci. 29 (6): 715-723

[20]. Boucher, A. and Dimitrakopoulos, R. (2009). Block-support simulation of multiple correlated variables. Math. Geosci. 41 (2): 215-237.

[21]. Shakiba, S., Asghari, O., KeshavarzFarajKhah, N., SarallahZabihi, S. and Tokhmechi, B. (2015). Fault and non-fault areas detection based on seismic data through min/max autocorrelation factors and fuzzy classification. J. Nat. Gas Sci. Eng. 26: 51-60.

[22]. Rodon, O. (2012). Teaching aid: minimum/maximum autocorrelation factors for joint simulation of attributes. Math. Geosci. 44 (4): 469-504.

[23]. Davis, J.C. (1986). Statistics and data analysis in geology. 2nd ed. New York: Wiley.

[24]. Dimitrakopoulos, R. and Fonseca, M.B. (2003). Assessing risk in grade-tonnage curves in a complex copper deposit, northern Brazil, based on an efficient joint simulation of multiple correlated variables. Application of Computers and Operations Research in the Minerals Industries. South African Institute of Mining and Metallurgy.

[25]. Ferreira, J. (2010). Comparison and application of three decorrelation methods PCA, MAF and ACDC. BSc (Hons) Mathematics Thesis. Edith Cowan University. Australia.

[26]. Kazemi, G.A. and Mehdizadeh, H. (2003). Trend, characteristics and chemical composition of atmospheric precipitation at Shahrood, Northeast Iran.

In: Abstracts of the 33rd general assembly of the International Union of Geodesy and Geophysics (IUGG). Sapporo. Japan.

[27]. Bakhshi, M. (1998). A short report on the climate and water resources of Shahrood [in Persian]. In: Proc. of Shahrood and development symposium. Shahrood. Iran. pp. 143–164.

[28]. Doulati Ardejani, F., JodieriShokri, B., Moradzadeh, A., Shafaei, S.Z. and Kakaei, R. (2011). Geochemical characterisation of pyrite oxidation and environmental problems related to release and transport of metals from a coal washing low-grade waste dump, Shahrood, northeast Iran. Environ. Monit. Assess. 183: 41-55.

[29]. Hajizadeh Namaghi, H., Karami, G.H. and Saadat, S. (2011). A study on chemical properties of groundwater and soil in ophiolitic rocks in Firuzabad, east of Shahrood, Iran: with emphasis to heavy metal contamination. Environ. Monit. Assess. 174: 573-583.

[30]. Jodeiri Shokri, B., Ramazi, H., DoulatiArdejani, F. and Moradzadeh, A. (2014). A statistical model to relate pyrite oxidation and oxygen transport within a coal waste pile: case study, AlborzSharghi, northeast of Iran. Environ. Earth Sci. 71: 4693-4702.

[31]. Nazemi, S. (2012). Concentration of heavy metal in edible vegetables widely consumed in Shahroud, the north east of Iran. J. Appl. Env. Bio. Sci. 2 (8): 386-391.

[32]. Sakizadeh, M., Mirzaei, R. and Ghorbani, H. (2016). Accumulation and soil-to-plant transfer factor of lead and manganese in some plant species in Semnan Province, central Iran. Iran J. Toxicol. 3 (10): 29-33.

[33]. IAEA. (2004). Soil sampling for environmental contaminants. International atomic energy agency. Austria. 81 P.

[34]. Wackernagel, H. (2003). Multivariate geostatistics: an introduction with applications. Berlin Heidelberg New York: Springer.

[35]. Pacheco, J., Casado, S. and Porras, S. (2013). Exact methods for variable selection in principal component analysis: Guide functions and pre-selection.Comput. Stat. Data An. 57 (1): 95-111.

[36]. Sakizadeh, M., Mirzaei, R. and Ghorbani, H. (2016). Support vector machine and artificial neural network to model soil pollution: a case study in Semnan Province, Iran. Neural Comput. Applic.doi:10.1007/s00521-016-2231-x.

[37]. Sengupta, D. (2014). Recent trends in modeling of environmental contaminants., Springer.

[38]. Emery, X. (2008). A turning bands program for conditional co-simulation of cross-correlated Gaussian random fields. Comput. Geosci. 34 (12): 1850-1862.

[39]. Guagliardi, I., Cicchella, D., De Rosa, R. and Buttafuoco, G. (2015). Assessment of lead pollution in topsoils of a southern Italy area: Analysis of urban and peri-urban environment. J. Environ. Sci-China. 33: 179-187.

[40]. Webster, R. and Oliver, M.A. (2007). Geostatistics for environemtalsciences, 2nd edition. John Wiley and Sons.

[41]. De Vitry, C., Vann, J. and Arvidson, H. (2007). A guide to selecting the optimal method of resource estimation for multivariate iron ore deposits. in Proceedings Iron Ore 2007. (The Australasian Institute of Mining and Metallurgy: Melbourne). pp. 67-78.

[42]. Tajvidi, E., Monjezi, M., Asghari, O., Emery, X. and Foroughi, S. (2015). Application of joint conditional simulation to uncertainty quantification and resource classification. Arab. J. Geosci. 8: 455-463.

[43]. Wu, E.M. and Kuo, S.L. (2012). Applying a multivariate statistical analysis model to evaluate the water quality of a watershed. Water Environ. Res. 84: 2075-2085.

[44]. Kowalkowski, T., Zbytniewski, R., Szpejna, J. and Buszewski, B. (2006). Application of chemometrics in river water classification. Water Res. 40 (4): 744-752.

[45]. Leon, E., Vargas, R., Bullock, S., Lopez, E., Panosso, A.R. and La ScalaJr., N. (2014). Hot spots, hot moments, and spatio-temporal controls on soil CO_2 efflux in a water-limited ecosystem. Soil Biol. Biochem. 77: 12-21.

[46]. Rodriguez Martin, J.A., Ramos-Miras, J.J., Boluda, R. and Gil, C. (2013). Spatial relations of heavy metals in arable and greenhouse soils of a Mediterranean environment region (Spain). Geoderma. 200-201: 180-188.

[47]. Liu, X., Xu, J., Zhang, M. and Zhou, B. (2004). Effects of Land Management Change on Spatial Variability of Organic Matter and Nutrients in Paddy Field: A Case Study of Pinghu, China. Environ. Manag. 34 (5): 691-700.

[48]. Lin, W.C., Lin, Y.P., Wang, Y.C., Chang, T.K. and Chiang, L.C. (2014). Assessing and mapping spatial associations among oral cancer mortality rates, concentrations of heavy metals in soil, and land use types based on multiple scale data. Int. J. Environ. Res. Public Health. 11 (2): 2148-2168.

[49]. Cambardlla, C.A., Moorman, T.B., Novak, J.M., Parkin, T.B., Turco, R.F. and Konopka, A.E. (1994). Field-scale variability of soil properties in central Iowa soils. Soil Sci. Soc. Am. J. 58: 1501-1511.

[50]. Lilli, M.A., Moraetis, D., Nikolaidis, N.P., Karatzas, G.P. and Kalogerakis, N. (2014). Characterization and mobility of geogenic chromium in soils and river bed sediments of Asopos basin. J. Hazard. Mater. 282: 12-19.

[51]. Cai, L., Xu, Z., Bao, P., He, M., Dou, L., Chen, L., Zhou, Y. and Zhu, Y.G. (2015). Multivariate and geostatistical analyses of the spatial distribution and source of arsenic and heavy metals in the agricultural soils in Shunde, Southeast China. J. Geochem. Explor. 148: 189-195.

[52]. Ha, H., Olson, J.R., Bian, L. and Rogerson, P.A. (2014). Analysis of heavy metal sources in soil using kriging interpolation on principal components. Environ Sci. Technol. 48 (9): 4999-5007.

[53]. Kabata-Pendias, A. (2011). Traceelements in soils and plants. Taylor and Francis Group.

[54]. Mueller, U.A. and Ferreira, J. (2012). The U-WEDGE Transformation Method for Multivariate Geostatistical Simulation. Math. Geosci. 44: 427-448.

[55]. Qu, M., Li, W. and Zhang, C. (2013). Assessing the risk costs in delineating soil nickel contamination using sequential Gaussian simulation and transfer functions. Ecol. Inform. 13: 99-105.

روشی جدید برای ارزیابی ریسک مکانی فلزات سنگین دارای همبستگی متقاطع با استفاده از روش شبیه‌سازی زمین‌آماری

محمد ساکی زاده[1*]، محمد تقی ستاری[2] و هادی قربانی[3]

۱- گروه محیط‌زیست، دانشکده علوم، دانشگاه تربیت دبیر شهید رجایی، ایران
۲- گروه مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز، ایران
۳- دانشکده مهندسی کشاورزی، دانشگاه صنعتی شاهرود، ایران

ارسال ۲۰۱۶/۱۲/۱۰، پذیرش ۲۰۱۷/۳/۱۳

* نویسنده مسئول مکاتبات: msakizadeh@gmail.com

چکیده:

در این تحقیق، نمونه‌های خاک از ۱۷۰ ایستگاه نمونه‌برداری در یک منطقه خشک در شاهرود و دامغان که در آن فعالیت‌های معدنکاوی در جریان است برداشت شد. مقادیر فلزات کبالت، سرب، نیکل، سزیم، مس، منگنز، استرانسیوم، وانادیوم، روی، کروم و تالیوم در هر ایستگاه ثبت شد. روش جدیدی تحت عنوان فاکتورهای خودهمبستگی بیشینه/کمینه برای اولین بار در تحقیقات زیست‌محیطی جهت همبستگی‌زدایی از عناصر پیش از اعمال تکنیک زمین مورد استفاده قرار گرفت. همبستگی بالای بین برخی از عناصر از یک سو و همبستگی مکانی ضعیف بین برخی دیگر می‌توانست منجر به تولید مقادیر ویژه منفی گردد؛ لذا تصمیم گرفته شد که ابعاد متغیرهای اولیه به وسیله روش آنالیز مؤلفه‌های اصلی کاهش داده شود. بر این اساس، شش فلز سنگین (کروم، منگنز، مس، وانادیم، نیکل و کبالت) به فاکتورهای بیشینه/کمینه مربوطه تبدیل و به دنبال آن شبیه‌سازی توسط روش متوالی گوس بر روی هر یک از فلزات به صورت جداگانه صورت پذیرفت. بررسی واریوگرام‌های فاکتورهای تولیدی شده حاکی از این بود که همبستگی متقاطع بین فاکتورها به ویژه در فاصله صفر و فاصله دلتا بین متغیرها از بین رفته است. اعتبارسنجی‌های متعدد از جمله بازتولید واریوگرام‌ها در فضای داده و فضای نرمال، همخوانی بین فاکتورهای اولیه و فاکتورهای شبیه‌سازی شده، بازتولید آمارهای توصیفی مربوط به داده‌ها و هیستوگرام داده‌های اولیه همگی حاکی از موفقیت این روش داشت. نتایج این تحقیق نشان داد که روش به کار گرفته شده قادر به بازتولید موفقیت‌آمیز داده‌ها و پیوستگی مکانی (واریوگرام) آن‌ها است. از این گذشته، استفاده از این روش ساده‌تر و به لحاظ محاسباتی آسان تر از روش معادل آن یعنی روش هم‌شبیه‌سازی متناوب گوس است که به دلیل عدم نیاز به برازش مدل خطی هم ناحیه‌ای است.

کلمات کلیدی: همبستگی زدایی، شبیه‌سازی زمین آمار، فاکتورهای خودهمبستگی بیشینه/کمینه.