Bollettino di Geofisica Teorica ed Applicata Vol. 60, n. 4, pp. 583-598; December 2019
DOI 10.4430/bgta0289

Application of projection pursuit multivariate transform to alleviate the smoothing effect in cokriging approach for spatial estimation of cross-correlated variables

N. Madani

School of Mining and Geosciences, Nazarbayev University, Nur-Sultan, Kazakhstan

(Received: 12 October 2018; accepted: 5 June 2019)

ABSTRACT Cokriging allows predicting variables from sampling information (e.g. boreholes and blast holes), taking into account their cross-correlation structures. When the bivariate relations among the variables are non-linear and complex, cokriging results may suffer from reproducing the complexities of interest. Another property of this linear geostatistical algorithm is the smoothing effect in the estimated block model, in which it over and underestimates the original distribution of the variables. To come up with those difficulties, this paper proposes an innovative algorithm to integrate the cokriging approach with a factor-based methodology entitled "projection pursuit multivariate transform" to first reproduce the complexity among the variables and, second to manage the smoothing effect in traditional cokriging algorithms. To do so, six cross-correlated variables obtained from a blast hole campaign belonging to a Nickle-Laterite deposit are presented and tested with the algorithm proposed. The results indicated that this algorithm is dramatically capable of alleviating the smoothing effect while the complexity in cross-correlation characteristics among the variables can be preserved.

Key words: projection pursuit multivariate transform, estimation, cokriging, smoothing effect.

1. Introduction

Linear interpolation of geo-related variables is important to different aspects of mining engineering such as long- and short-term mine planning (Sinclair and Blackwell, 2002; Verly, 2005; Rossi and Deutsch, 2014). The ore deposit description attained from spatial modelling of blast hole data sets, has been widely taken into account, particularly, for short-term mine design and grade control, so as to account for prediction of ultimate destination of a mined block (Davis, 1992; Rossi and Deutsch, 2014). In this step of mine design, ore/waste classification of extractable selective mining units are considered and proper choice should be taken into account. These models also can be applicable as auxiliary models utilised for amending the local estimation of the long-term resources model that obtained already from borehole sampled locations (Rossi and Deutsch, 2014). It is approved in practice that making wrong decision leads to losing a huge amount of money for choosing the improper destination of a mined block.

In the case of complex ore deposits, the issue of estimation of variables in the region is questionable and somehow tedious. Due to this fact, modelling those complications in the co-spatial

behaviour of the grades in multi-element deposits prompts one to employ enhanced geostatistical techniques. Traditional approaches for this purpose such as polygon, inverse distance weighted, and even (co)kriging approaches (Scerbo and Mazzotti, 1991; Wackernagel, 2003; Chiles and Delfiner, 2012; Asghari *et al*., 2016) are incapable of reproduce those complexities such as non-linearity, heteroscedasticity, and geological constraint that frequently exist among the variables of interest (Madani *et al*., 2018). Another difficulty in employing those techniques in grade estimation is smoothing effect in the variability of estimated grades accompanying with over - and underestimation (Journel and Huijbregts, 1978; Goovaerts, 1997; Chiles and Delfiner, 2012). Therefore, application of such approaches might lead to a significant bias in the generated block model and consequently decreases the level of confidence in selection of a proper destination, which directly impact on annual revenue of a mining project.

In this paper, it is of interest to utilise one of the factor-based approaches to populate the cross-correlated variables integrating with the cokriging paradigm, for which whole the abovementioned complexities are present in the data set. This factor-based method is entitled Projection Pursuit Multivariate Transform (PPMT) (Barnett *et al*., 2014), a multivariate Gaussian technique, based upon modified components of projection-pursuit density estimation (Friedman, 1987; Hwang *et al*., 1994) that transforms data of any multivariate characters with any number of variables and samples to an uncorrelated multi-Gaussian global distribution. The rationale of this method is to find the vector that yields the most non-Gaussian projection. This algorithm and Gaussianised associated steps is iterated many times up to the level that converges to the univariate Gaussian model (Barnett *et al*., 2014, 2016). In comparison to some other factor-based methods such as Principle Component Analysis (PCA) (Hotelling, 1933a, 1933b) and Minimum/maximum Autocorrelation Factors (MAF) (Switzer and Green, 1984; Desbaratas and Dimitrakopoulos, 2000), PPMT method improves the reproduction of non-linear multivariate properties while requiring fewer steps and suppositions. This method, through transformation to the factors, is able to handle any complexity among the cross-correlated variables and reproduce its complicated spatial structures. Cokriging as a linear estimation methodology was also applied in this research in combination with PPMT in order to establish an innovative workflow for short-term mine planning in complex ore deposits. The fundamental of this algorithm is that, the data in PPMT can be transferred to Gaussian distribution and through the estimation by cokriging, the estimated values on the block can be back-transferred to original space. The benefit is in alleviation of the smoothing effect and overcoming the problem in under and over estimation that often happen in conventional cokriging. To do so, a data set obtained from a blast hole campaign pertaining to a Nickel-Laterite deposit is considered and six cross-correlated variables (Fe, Ni, MgO, SiO_2, Al_2O_3, and Cr) with complexity in multivariate characteristics are selected for further analysis. The primary inspection of bi-variate relation (scatter-plot) exposed many aspects of complexity such as non-linearity and heteroscedasticity. In order to track the workflow, the summarised items are consecutively considered:

1) declustering: in a first step, the data set is declustered in order to obtain the representative distributions;
2) apply the PPMT to the six variables and yield the six PPMT equivalent-factors;
3) variogram analysis: building the cokriging systems needs the information of covariance matrix. The most straightforward approach for such a variance-covariance matrix calculation is variogram that describes the spatial variability of the underlying variables in the region.

It also contributes to the knowledge of the spatial correlation between different points in a deposit (Journel and Huijbregts, 1978). The experimental variograms are then calculated over the PPMT factors and the proper structures are fitted by semi-automatic approaches. Since there are six variables and the cross-dependency is critical, cross-variograms (linear model of coregionalisation) are also taken into account;

4) estimation: as mentioned above, cokriging as an estimation technique is applied to model six PPMT factors in each block. A moving neighbourhood is considered for simple cokriging with maximum and minimum ranges equal to the variogram ranges containing 50 surrounding data;
5) back-PPMT-transformation: the estimated factors are then back-transformed to the original distribution as the same information was required for PPMT forward transformation.

The spatial distribution of each variable is considered. In order to check whether the cross-correlation in bivariate distribution is reproduced after modelling process, the scatterplots of the desired variables are presented. This consistency in bi-variate distributions corroborates that the goal in modelling the complexities is successfully achieved and one can reproduce the complexity in multivariate distribution by integration of PPMT to the co-estimation approaches. The workflow is also applicable to short-term mine planning wherever the grade control is the subject of making-decision and one needs to check the reliable spatial variability of the variables of interest in the deposits with complex characteristics in ore formation.

2. Methodology: simple cokriging

Simple cokriging is a generalisation of simple kriging, i.e. kriging with a known mean value, and aims to predict primary and secondary variables by taking into account their joint spatial correlation structure (Journel and Huijbregts, 1978; Goovaerts, 1997; Wackernagel, 2003; Chilès and Delfiner, 2012; Madani and Emery, 2019). Provided that these variables are represented by second-order stationary random fields, the cokriging predictor and the variance of the prediction error (known as the simple cokriging variance) for the primary variable (hereafter denoted with index 1) given one secondary variable (denoted with index 2) are defined as (Myers, 1982):

$$Z^*_{SCK}(x_0) = m_1 + \sum_{\alpha=1}^{n_1} \omega^1_\alpha (Z_1(x_{1,\alpha}) - m_1) + \sum_{\alpha=1}^{n_2} \omega^2_\alpha (Z_2(x_{2,\alpha}) - m_2) \quad (1)$$

$$\sigma^2_{SCK}(x_0) = C_{11}(x_0 - x_0) - \sum_{\alpha=1}^{n_1} \omega^1_\alpha C_{11}(x_{1,\alpha} - x_0) - \sum_{\alpha=1}^{n_2} \omega^2_\alpha C_{21}(x_{2,\alpha} - x_0) \quad (2)$$

where ω^i_α (i = 1, 2) is the weight assigned to the data $Z_i(x_{i,\alpha})$ of the i-th variable Z_i at the a-th data location $x_{i,\alpha}$ (a = 1, … n_i) of this variable, x_0 is the location targeted for prediction; m_i is the mean value of the i-th variable Z_i; C_{ij} is the direct ($i = j$) or cross ($i \neq j$) covariance between variables Z_i and Z_j (i, j = 1, 2). The previous equations can be generalised to the case with more than one secondary variable, at the price of heavier notation, which will not be considered in this work. Note that the numbers of data are not necessarily the same for the primary and secondary variables, a case known as a heterotopic sampling design (Wackernagel, 2003) in opposition to the isotopic (equally-sampled) case. The weights ω^i_α required in Eqs. 1 and 2 are obtained by

solving the following system of linear equations:

$$\begin{cases} \sum_{\alpha=1}^{n_1} \omega_\alpha^1 C_{11}\left(x_{1,\beta} - x_{1,\alpha}\right) + \sum_{\alpha=1}^{n_2} \omega_\alpha^2 C_{12}\left(x_{1,\beta} - x_{2,\alpha}\right) = C_{11}\left(x_{1,\beta} - x_0\right), & \beta = 1,\ldots n_1 \\ \sum_{\alpha=1}^{n_1} \omega_\alpha^1 C_{21}\left(x_{2,\beta} - x_{1,\alpha}\right) + \sum_{\alpha=1}^{n_2} \omega_\alpha^2 C_{22}\left(x_{2,\beta} - x_{2,\alpha}\right) = C_{21}\left(x_{2,\beta} - x_0\right), & \beta = 1,\ldots n_2 \end{cases} \quad (3)$$

Different neighbourhood strategies can be used to reduce the number of data for cokriging. For instance, a single search strategy selects the data locations that are geographically the closest to the target location x_0, irrespective of which variables are known at those locations, whereas a multiple search strategy consists in selecting the closest data of each (primary or secondary) variable.

3. Projection pursuit multivariate transform

One of the new developed factor-based approaches, is PPMT which is targeted to handle with multivariate complexities such as non-linearity and heteroscedasticity that exist intrinsically among the variables (Barnett *et al*., 2014; Barnett, 2017) and is suitable in the cases wherever the traditional normal score transformation is not able to manage those mentioned complexities. The general steps for implementation of PPMT is based on two forward and backward transformations (Barnett *et al*., 2014; Battalgazy and Madani, 2019). Forward transformation converts the original data to uncorrelated multi-Gaussian distribution taking into account any complexity existing among the variables. This transformation is based on the Projection Pursuit Density Estimation algorithm (PPDE) (Friedman, 1987). Provided that these variables are represented by second-order stationary random fields, PPMT transformation steps for two variables can be defined as:

1. transform the original variables to normal score values with a mean of zero and variance one N(0, 1): this can be implemented by normal score transformation methodologies such as Gaussian anamorphosis (Rivoirard, 1994) or quantiles-based approach (Deutsch and Journel, 1998):

$$Z_i(u) = G^{-1}(F_i(Y_i(u))) \qquad i = 1, 2 \quad (4)$$

where $G^{-1}(.)$ is standard normal cumulative distribution function, $F_i(.)$ are the cumulative distribution function of the original variables $Y_i(u)$ and $Z_i(u)$ are the normal score values;

2. data sphering (A): compute the experimental variance-covariance matrix at lag 0 since we are dealing with normal score values, this matrix is identical to the sample correlation matrix. In the case of two variables, this matrix (V) is as:

$$V = Corr\{Z(u),\ Z(u)\} = \begin{bmatrix} \rho_{11}(0) & \rho_{12}(0) \\ \rho_{21}(0) & \rho_{22}(0) \end{bmatrix} \quad (5)$$

where the principal diagonal element equals one which is identical to the total variance, upper and lower diagonal elements $\rho_{12}(0)$ and $\rho_{21}(0)$ equal the linear correlation coefficient between two normal score variables $Z_1(u)$ and $Z_2(u)$, respectively;

3. data sphering (B): perform the spectral decomposition of above matrix (V) to derive the orthonormal eigenvectors matrix (M_1), associated with the underlying diagonal eigenvalues matrix (E_1), such that:

$$V = M_1 E_1 M_1^T \tag{6}$$

It is necessary to check that the entries of E are in decreasing order;

4. data sphering (C): calculate the sphering transformations at locations u by:

$$\varphi_i(u) = Z_i(u) M_1 E_1^{-1/2} M_1^T \qquad i = 1, 2 \tag{7}$$

where $\varphi_i(u)$ are the scores with normal standard distribution due to the *a priori* multivariate Gaussian assumption and are decorrelated;

5. projection pursuit: although the $\varphi_i(u)$ are decorrelated, however, the complexity still manifests itself in bivariate relation. Projection pursuit can transform the decorrelated variables $\varphi_i(u)$ to the multi-Gaussian variables free of the underlying complexity. To do so, taking into account that projection of the data is $\tau = \varphi_i(u)\alpha$, where k*1 is a unit vector α, projection index test statistic is defined as $A(\alpha)$ that determines the univariate non-Gaussianity. When related projection is appropriately Gaussian, projection index, $A(\alpha) = 0$. According to Friedman (1987), the optimised search is used to determine the θ which will identify maximum $A(\alpha)$. $\varphi_i(u)$ is transformed to standard Gaussian, $\varphi_i(u)'$ where related projection is $\tau' = \varphi_i(u)'\alpha$ after the optimum α is identified. The transformation starts with application of Gram-Schmidt algorithm to compute the orthogonal matrix (Eq. 8) (Reed and Simon, 1972):

$$\omega = [\alpha, \phi_1, \phi_2, \ldots, \phi_{k-1,}] \tag{8}$$

and transformation can be reached by multiplication of $\varphi_i(u)$ and U (Eq. 9):

$$\varphi_i(u)\omega = [p, \varphi_i(u)\phi_1, \varphi_i(u)\phi_2, \ldots, \varphi_i(u)\phi_{k-1,}] \tag{9}$$

Then, in order to get the transformation that outputs Gaussian standard projection, τ', normal score transform is computed as in Eq. 10:

$$\breve{G}\left(\varphi_i(u)\omega\right) = [\tau', \varphi_i(u)\phi_1, \varphi_i(u)\phi_2, \ldots, \varphi_i(u)\phi_{k-1,}] \tag{10}$$

and back-transformation to the original basis is computed by Eq. 11:

$$\varphi_i(u) = \breve{G}\left(\varphi_i(u)\omega\right)\omega^T \tag{11}$$

The data transformed $\varphi_i(u)'$ outputs Gaussian projection by α where projection index, $A(\alpha)$ is zero. Additionally, in order to get other complexities, optimised search can be repeatedly used to determine maximum projection index;

6. stopping criteria: this step focuses on the selection of target projection index. Increase of dimensions leads to difficult resolution and discovery of complexities, while the number of observations results in the reliability of the projections for identification of appropriate multivariate structure. Random samples from Gaussian CDF can be used by implementation of bootstrapping algorithm, where m is distribution, k is dimension and n is number of observations for choosing the target projection index for PPMT stopping. After, value of the projection index can be computed;
7. back-transform: Gaussian realisations can then be back-transformed based on the mapped data on an original space, where configuration between simulated and mapped one is preserved.

4. Coregionalisation modelling: Linear Model of Coregionalisation (LMC)

Solving the cokriging system requires the knowledge of the direct and cross-covariances between the primary and secondary variables. In this respect, the linear model of coregionalisation is widely used to fit such covariances, owing to its mathematical simplicity and tractability (Journel and Huijbregts, 1978; Goovaerts, 1997; Wackernagel, 2003). In this model, the direct and cross-covariances $C_{ij}(h)$ $(i, j = 1, 2)$ are defined as weighted sums of L basic covariances, also called basic nested structures:

$$C_{ij}(h) = \sum_{l=1}^{L} b_{ij}^{l} c_l(h) \tag{12}$$

where, for each structure $(l = 1,...L)$, $(b_{ij}^{l})_{i,j=1,2}$ is a 2×2 real-valued, symmetric, positive semi-definite matrix (coregionalisation matrix) and $c_l(h)$ is a permissible stationary covariance model (basic nested structure). In practice, such a model can be fitted to a set of experimental direct and cross-covariances by means of semi-automated algorithms (Goulard and Voltz, 1992; Emery, 2010).

In this study, the PPMT is integrated with (co)kriging approach in order to circumvent the problem of smoothing effect in traditional (co)kriging and also to reproduce the complex bivariate structure among the cross-correlated variable. To do this, the proposed algorithm is presented as a flowchart in Fig. 1.

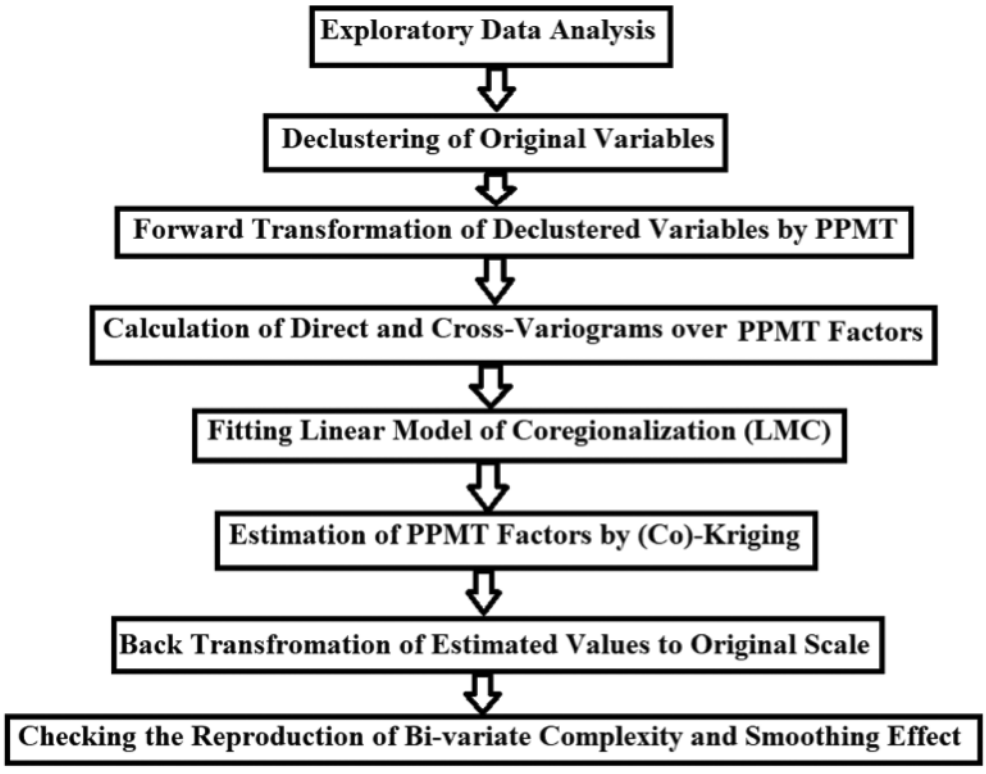


Fig. 1 - Workflow showing the proposed algorithm for modelling the cross-correlated variables with complex bivariate relationship.

5. Case study

The case study pertaining to a Nickle-Laterite deposit including 9990 sample available from blast hole data set with a very dense sampling pattern composited in one metre. Since this grid is dense, in order to show the capability of the proposed methodology, 1000 random data are selected, for which six cross-correlated variables are isotopically assayed at each sample point. Isotopic sampling ensures that all the variables are available through all the sample locations (Wackernagel, 2003). The location map of the blast holes for each variable is presented in Fig. 2. The name and location of data set cannot be disclosed because of confidentiality reasons.

The data set is declustered by cell declustering approach (Deutsch and Journel, 1998). In this method, the area of interest should first be divided into a grid of cells and the weights are required, then, to be assigned to each data according to the number of samples falling in the same cell. The resulted weights in occupied cells are greater than zero and in total sum to one, while the vacant cells receive no weights [see Rossi and Deutsch (2014) and reference herein]. This step is required since the sampling pattern is irregular and assigning the weights obtained from this approach contributes to alleviation of biasedness in univariate and multivariate statistical parameters derivation. Declustering technique is, then, implemented in a dimension of 50×50×15 (in metres) based on the primary pattern of the randomly selected blast holes. Statistical parameters are calculated as shown in Tables 1 and 2. Declustering also ensures that the statistical parameters are representative and they are not impacted from scarcity of data in some regions anymore. The reason of using the blast hole data set relates to the paramount importance of this type of information. One of the crucial step in reserve evaluation and also ore/waste classification includes the reliable modelling of the variable under study in a deposit attained from analysis of blast hole data set. Therefore, employing a trustworthy geostatistical algorithm can boost the quality of estimation and reduces the plausible penalties that may happen because of biased results obtained from traditional algorithms.

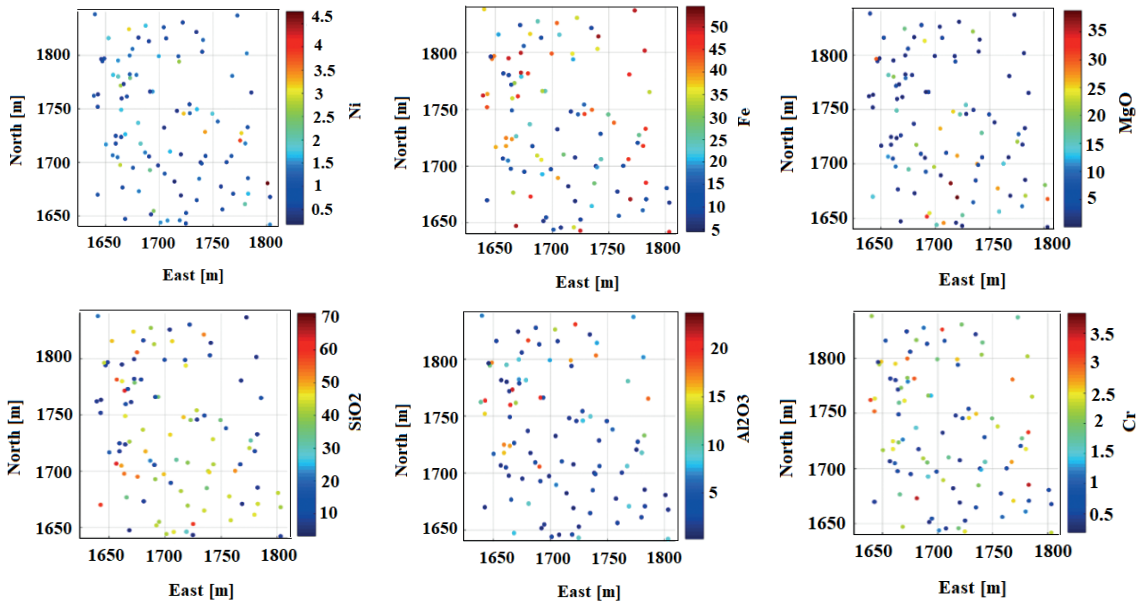


Fig. 2 - Location maps of six cross-correlated variables obtained from a blast hole campaign. The variables are all measured in (%).

Table 1 - Declustered univariate statistical parameters computed over six cross-correlated variables.

	Mean (%)	Variance	Minimum (%)	Maximum (%)
Ni	1.27	0.662	0.18	4.76
Fe	29.70	257.830	4.50	54.60
MgO	8.32	108.360	0.10	38.72
SiO_2	26.03	339.050	3.20	71.25
Al_2O_3	6.85	35.980	0.20	23.80
Cr	1.58	0.812	0.19	3.83

Table 2 - Multivariate statistics over six cross-correlated variables.

	Ni	Fe	MgO	SiO_2	Al_2O_3	Cr
Ni	1.000	–	–	–	–	–
Fe	-0.171	1.000	–	–	–	–
MgO	0.011	-0.860	1.000	–	–	–
SiO2	0.301	-0.939	0.708	1.000	–	–
Al2O3	-0.384	0.657	-0.655	-0.749	1.000	–
Cr	-0.090	0.855	-0.771	-0.817	0.606	1.000

6. Projection Pursuit Multivariate Transform

The next step after inference of exploratory data analysis is to transfer the six cross-correlated variables [Fe(%), Ni(%), MgO(%), SiO_2(%), Al_2O_3(%), and Cr(%)] into the factors free of complexity characteristics (hereafter PPMT). After this forward transformation, the correlation among the PPMT factors may not be perfectly zero and decorrelated. Fig. 3 shows the bivariate relation between Fe(%) and other five variables before and after PPMT transformation. Barnett *et al*. (2014) recommended to employ the MAF to decorrelate the forward transformed factors. However, in this study, the idea is to take into account those small dependency i.e. correlation that exist among the forward transformed factors and use cokriging system to co-estimate the obtained factors in the region, while the spatial cross-correlation can be quantified through their direct and cross-variograms. Although this measure of correlation may be very small, however considering the cross-variograms in this respect and subsequently cokriging system, contributes taking into account the relationship among the variables for the process of modelling, which is of paramount importance for the purpose of ore/waste classification in this study. After implementation of cokriging, the last step is, then, to back-transfer the estimated factors to the original scale by backward PPMT transformation in order to return their intrinsic complex cross-correlation.

7. Variogram analysis

The experimental direct and cross-variograms are calculated in omni-directional and, then, one-structured spherical variogram with 50 m of range is fitted to the experimental ones as the theoretical model. Variogram analysis in different direction with confined tolerance showed no

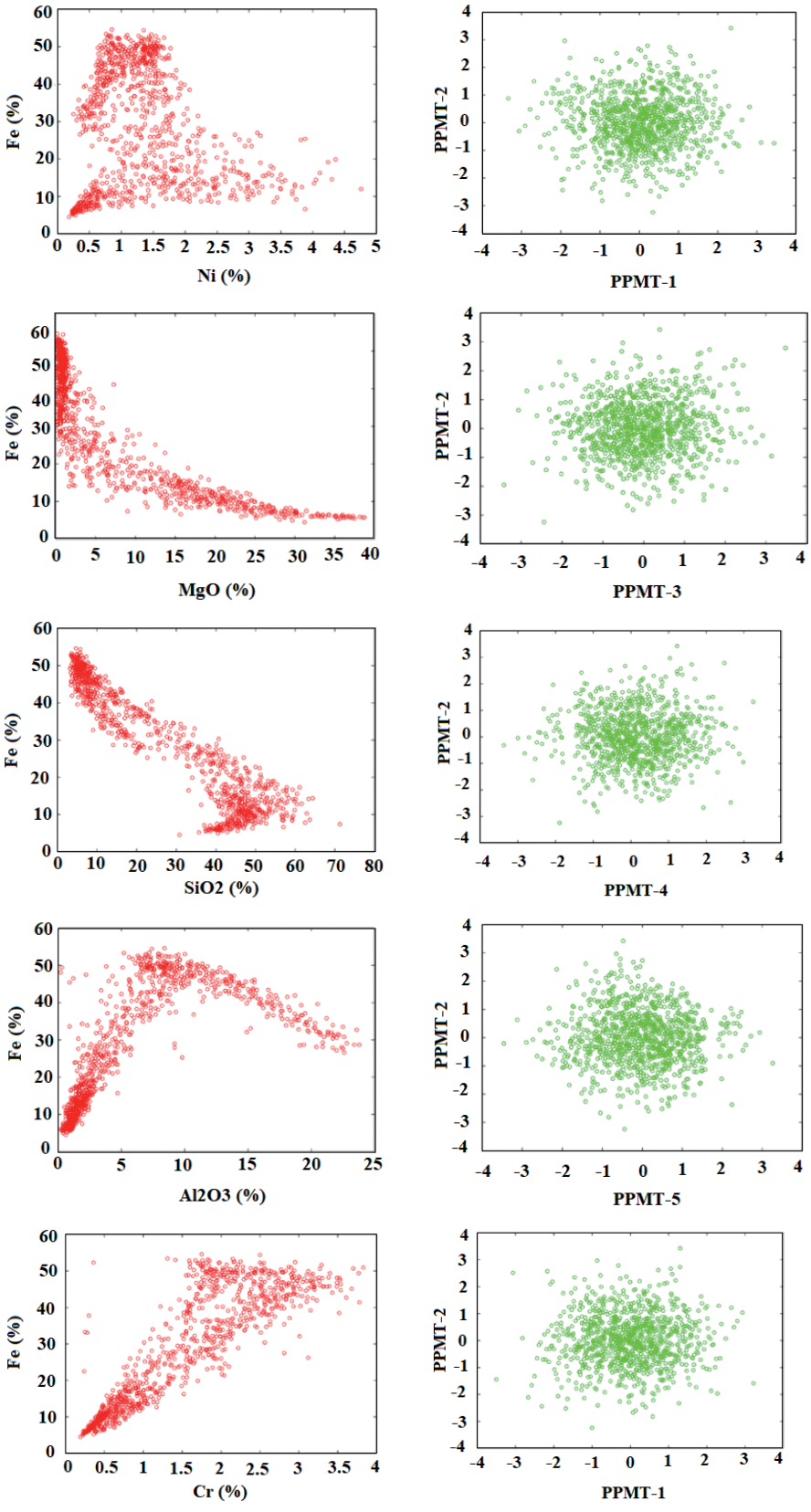


Fig. 3 - Bivariate relation before and after PPMT transformation among Fe and Ni, MgO, SiO_2, Al_2O_3, and Cr.

significant anisotropy, the reason why omni-directional variograms are computed for measuring the spatial continuity. In order to obtain the linear model of coregionalizations, it is necessary to fit a model that honors the semi-positive definiteness condition in the sill matrices (Journel and Huijbregts, 1978; Goovaerts, 1997; Wackernagel, 2003). This step is followed by a semi-automatic approach (Goulard and Voltz, 1992; Emery, 2010). For brevity, the direct variograms of PPMT factors are illustrated in Fig. 4 and the relevant formula provided in Eq. 13. All the cross-variograms are provided in the Appendix.

$$\begin{pmatrix} \gamma_{Ni} & \gamma_{Fe/Ni} & \gamma_{MgO/Ni} & \gamma_{SiO2/Ni} & \gamma_{Al2O3/Ni} & \gamma_{Cr/Ni} \\ \gamma_{Ni/Fe} & \gamma_{Fe} & \gamma_{MgO/Fe} & \gamma_{SiO2/Fe} & \gamma_{Al2O3/Fe} & \gamma_{Cr/Fe} \\ \gamma_{Ni/MgO} & \gamma_{Fe/MgO} & \gamma_{MgO} & \gamma_{SiO2/MgO} & \gamma_{Al2O3/MgO} & \gamma_{Cr/MgO} \\ \gamma_{Ni/SiO2} & \gamma_{Fe/SiO2} & \gamma_{MgO/SiO2} & \gamma_{SiO2} & \gamma_{Al2O3/SiO2} & \gamma_{Cr/SiO2} \\ \gamma_{Ni/Al2O3} & \gamma_{Fe/Al2O3} & \gamma_{MgO/Al2O3} & \gamma_{SiO2/Al2O3} & \gamma_{Al2O3} & \gamma_{Cr/Al2O3} \\ \gamma_{Ni/Cr} & \gamma_{Fe/Cr} & \gamma_{MgO/Cr} & \gamma_{SiO2/Cr} & \gamma_{Al2O3/Cr} & \gamma_{Cr} \end{pmatrix} = \tag{13}$$

$$\begin{pmatrix} 0.992 & -0.032 & -0.040 & 0.006 & -0.104 & -0.050 \\ -0.032 & 1.041 & -0.029 & -0.014 & -0.023 & 0.003 \\ -0.040 & -0.029 & 0.993 & 0.014 & -0.031 & 0.009 \\ 0.006 & -0.014 & 0.014 & 0.956 & -0.008 & 0.035 \\ -0.104 & -0.023 & -0.031 & -0.008 & 0.982 & 0.016 \\ -0.050 & 0.003 & 0.009 & 0.035 & 0.016 & 0.954 \end{pmatrix} \text{Sph}(50\text{m}, 50\text{m}, 50\text{m})$$

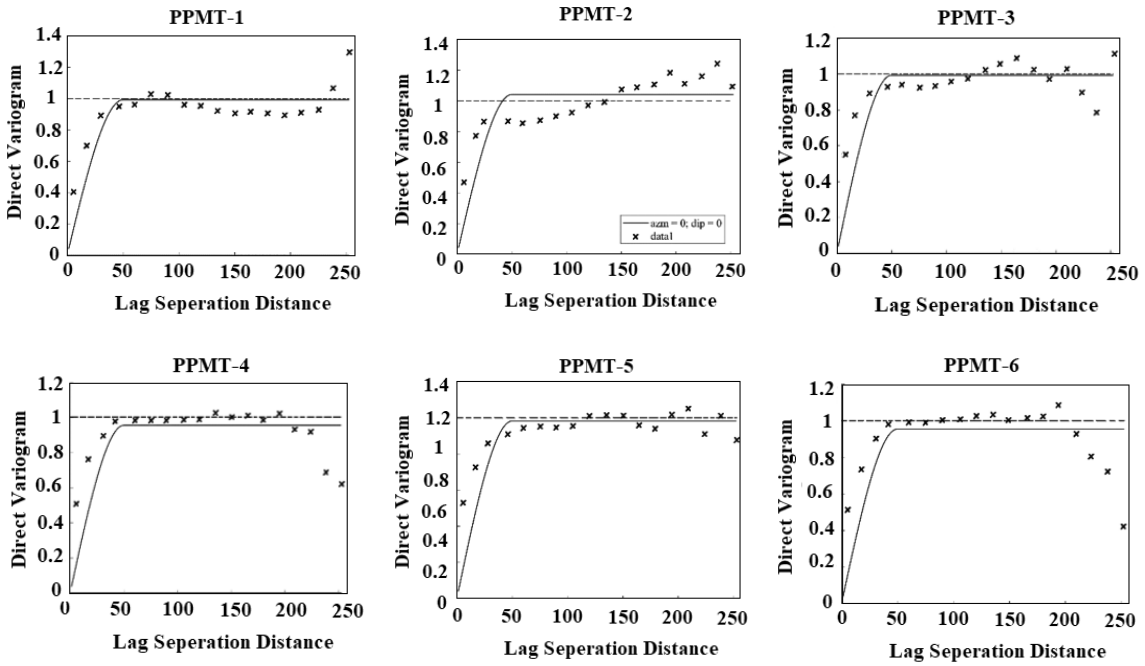


Fig. 4 - Direct experimental and theoretical variograms over six PPMT factors, the dashed lines show the variance of PPMT factors.

8. Spatial 3D estimation

Once the direct and cross-variogram formulae are derived, one can establish the simple cokriging system. A grid with mesh dimension of 5×5×5 (in metres) with mesh size of 38, 41 and 15 along east, north, and elevation coordinates are considered. The neighbourhood is selected as moving and the parameters for the range of search neighbourhood are set to 200 m with up to 8 number of data per octant. This number is chosen arbitrarily. One of the method to define this number optimally can be based on cross-validation or jack-knife approach (Deutsch and Journel, 1998) as it is a common practice in geostatistical contexts. However, since the scope of this study is mainly on the basis of examination of smoothing effect, the chosen number of data for estimation deems not very crucial at this stage. The estimated PPMT factors are then back-transformed to the original distribution. The produced maps are shown in Fig. 5 for six back-transformed cross-correlated variables.

The results are quite satisfying from visual inspection in terms of reproducing the cross-correlation magnitude in the final maps (Fig. 5). For instance, there is a strong negative correlation between Fe and MgO (-0.86). This can be corroborated through the maps, since the high values of Fe is spatially distributed in the areas, for which the Al_2O_3 shows low amount of concentration (north and east part of the region) and vice versa.

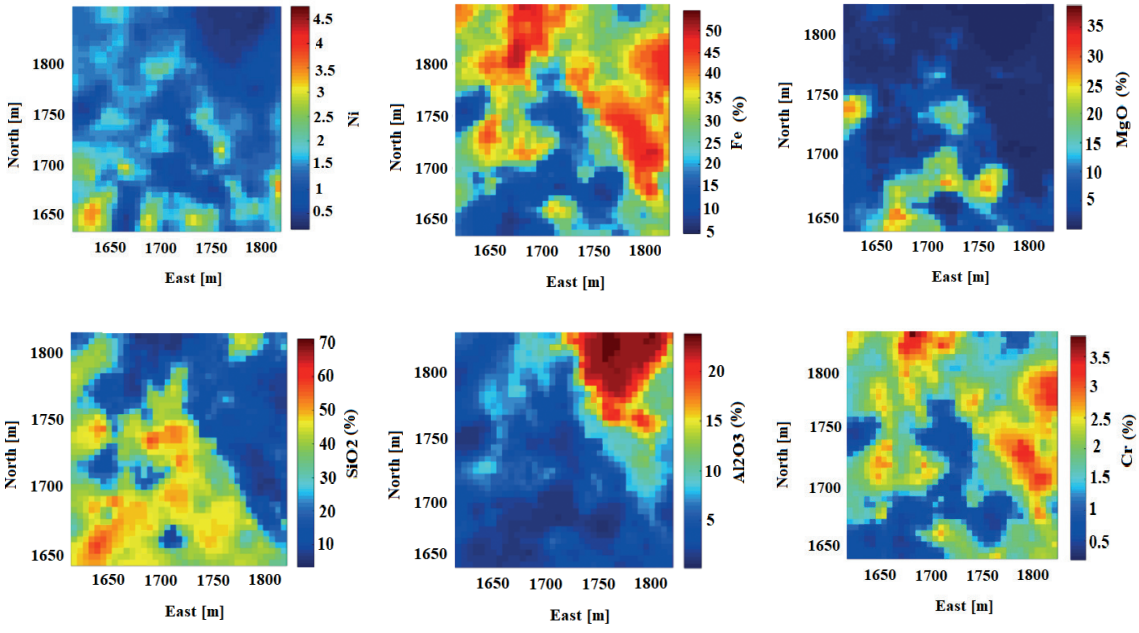


Fig. 5 - Estimated maps produced by integration of cokriging and PPMT approach.

9. Reproduction of statistical parameters

Validation of produced models are necessary, because the generated values at each block obtained from blast holes should be taken into account for further analysis of a mining project such as decision-making about the destination of block to be either sent to mill or stockpiled

and also for short-term mine planning. The imprecise estimation of the grades inside the blocks leads to loose a valuable part of a mine and impacts the economic issues. In this section, it is of interest to examine the reproduction of statistical parameters to analyse the proficiency of the model. In particular case, bivariate relation is of interest to be taken into account for endorsement of reproduction of primary declustered correlation among the variables. For saving the space, Fig. 6 shows the scatterplot between Fe and other five variables (Ni, MgO, SiO_2, Al_2O_3, and Cr) associated with bivariate distribution of the declustered original values. This bivariate relation is well reproduced in all the cases. Table 3 also shows the correlation coefficients values among all the variables along with the original correlation coefficients, corroborating that this bivariate characteristics among all the variables are suitably reproduced.

Table 3 - Reproduced multivariate statistics over six cross-correlated variables. Lower diagonal is original correlation coefficients and upper diagonal is the correlation coefficients obtained from cokriging modelling by PPMT approach.

	Ni	Fe	MgO	SiO_2	Al_2O_3	Cr
Ni	1.000	-0.009	-0.088	0.240	-0.404	0.030
Fe	-0.171	1.000	-0.829	-0.907	-0.559	0.898
MgO	0.011	-0.86	1.000	-0.701	-0.612	-0.795
SiO_2	0.301	-0.939	0.708	1.000	-0.688	-0.828
Al_2O_3	-0.384	0.657	-0.655	-0.749	1.000	0.566
Cr	-0.090	0.855	-0.771	-0.817	0.606	1.000

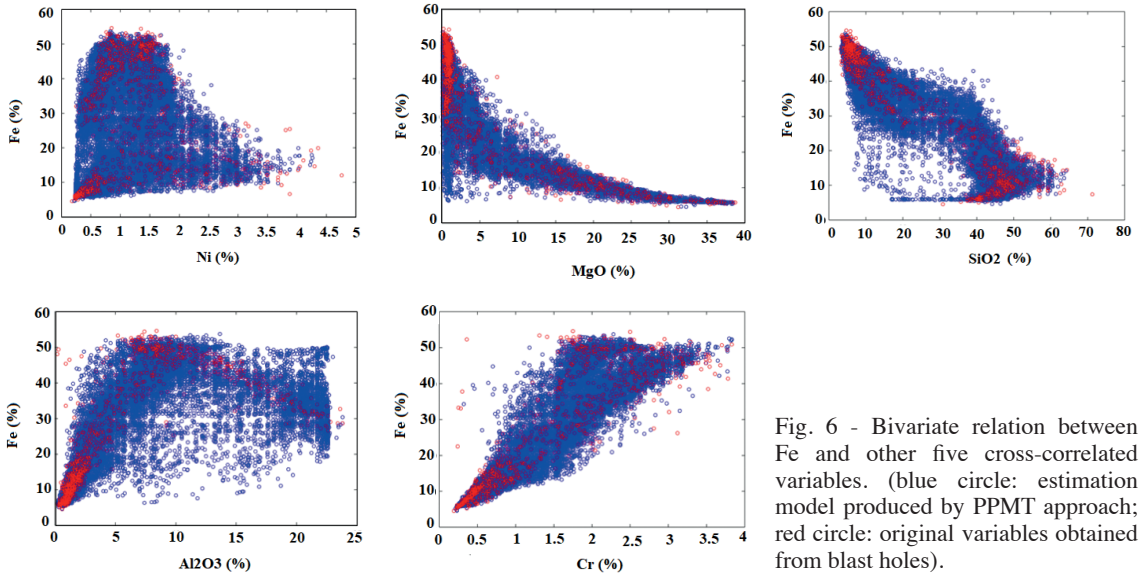


Fig. 6 - Bivariate relation between Fe and other five cross-correlated variables. (blue circle: estimation model produced by PPMT approach; red circle: original variables obtained from blast holes).

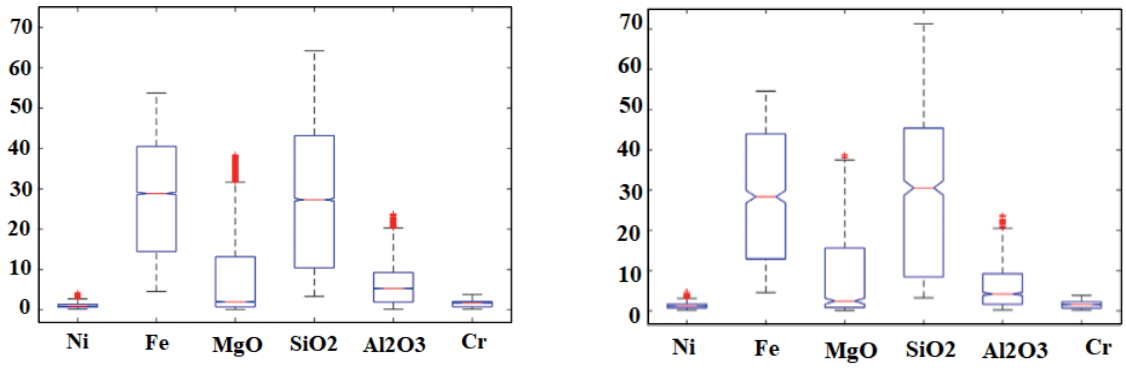
Another criterium for validation of the estimation model, is to consider and compare the univariate statistical parameters of the produced models with original declustered ones. Since the (co)kriging approaches suffer from smoothing effect, it is very significant to consider the variance and ranges in the estimate block models. This characteristic of the linear geostatistics, sometimes prevents the practitioners to employ the right algorithm and advocates them to use some post-

modification procedures (Deutsch and Journel, 1998; Chiles and Delfiner, 2012). For instance, producing the negative values in kriging outputs, which is the results of negative weights assigned to the sampling points, can be either manipulated or neglected in the process of block model analysis. This might make some bias when one is considering the maximum and minimum values of the underlying variables.

The boxplots for both estimated values and original data set are presented in Fig. 6 and produced statistical parameters are also presented in Table 4. The distributions of the estimated values bear resemblance to the original global variability, in which it means that the proposed algorithm is capable of resolving the problem of smoothing effect in conventional techniques of (co)kriging paradigms (maximum and minimum values in Table 4). However, the variance is a bit different from the original data: this deviation can be influenced from the cokriging algorithm.

Table 4 - Univariate statistical parameters obtained from estimated block models (proposed algorithm) and original declustered distribution obtained from blast holes.

		Mean (%)	Variance	Minimum	Maximum
Ni	Blast holes	1.277	0.662	0.180	4.760
	Estimate model	1.099	0.394	0.212	4.260
Fe	Blast holes	29.790	257.830	4.500	54.600
	Estimate model	28.180	196.530	4.500	53.730
MgO	Blast holes	8.320	108.360	0.100	38.720
	Estimate model	7.640	98.120	0.100	38.410
SiO_2	Blast holes	26.030	339.053	3.200	71.250
	Estimate model	27.320	269.300	3.300	64.160
Al_2O_3	Blast holes	6.852	35.980	0.200	23.800
	Estimate model	7.130	40.150	0.200	23.800
Cr	Blast holes	1.580	0.812	0.190	3.830
	Estimate model	1.550	0.603	0.200	3.830



a) Proposed algorithm (cokriging and PPMT). b) Original blast hole data.

Fig. 7 - Boxplot of the six cross-correlated variables obtained from proposed model (a) and original data set (b).

10. Conclusions

Cokriging is a widely used technique in spatial prediction problems. Its implementation becomes prohibitive in the cases when there exists a complexity among the interdependency of the variables. It also suffers from smoothing effect like other linear geostatistical approaches. An innovative algorithm is proposed in this study based on transformation of original cross-correlated variables to semi non-correlated factors by PPMT. Then, through inference of linear model of coregionalisation (LCM), the co-spatial continuity is fitted over the experimental variograms calculated from six cross-correlated variables obtained from a blast hole campaign. The obtained factors from PPMT are, then, taken into account for establishing the cokriging system considering the simple cokriging. On the other side, the rationale of the proposed algorithm in this study is based on integration of a factor based approach (i.e. PPMT) with a traditional (co)kriging approach. In this sense, transformed variables (PPMT factors) do not necessarily required to be decorrelated after forward transformation. Instead, applying (co)kriging gives this privilege to the process of modelling to take into account even small correlations that exist among these PPMT factors.

The algorithm showed that the smoothing effect (over - and underestimation) in the blocks are alleviated. This also corroborates that the complexity among the variables in bivariate relations (through the scatterplot) for all six variables is well reproduced. It is also recommended to use the algorithm for the feasibility study of a mining project for mineral resource estimation and economic consideration, due to the reliability of the model. It should be noted that using the proposed algorithm may not outperform the results obtained from simulation approaches.

In the case that calculation of complex multivariate uncertainty is of paramount importance in mineral resource modelling, it is highly recommended to use Gaussian (co)-simulation algorithms instead (Hosseini *et al*., 2017; Khorram *et al*., 2017; Battalgazy and Madani, 2019; Abildin *et al*., 2019; Eze *et al*., 2019). One may concern to compare these results with some other factorisation technique such as PCA and/or MAF such as the one presented in Da Silva and Costa (2014). It is worth mentioning that those approaches are suitable for modelling the linearity characteristics among the variables and not very common for modelling the complex bi-variate relations (Abildin *et al*., 2019). However, integration of MAF with PPMT is shown to be capable of producing the satisfying results in this complex bi-variate context (Barnett *et al*., 2014).

The future work for this study can be based on employing the proposed algorithm integrating with estimation of compositional components that have closure problems. In this respect, examination of correlation among the variables might be spurious and draws the careful attention for its interpretation.

Acknowledgements. The author acknowledges the Nazarbayev University for funding this work via ''Faculty development competitive research Grants for 2018-2020'' under Contract No. 090118FD5336. The author is also grateful to editorial team and two anonymous reviewers for their valuable comments, which substantially helped improving the final version of the manuscript.

REFERENCES

Abildin Y., Madani N. and Topal E.; 2019: *A hybrid approach for joint simulation of geometallurgical variables with inequality constraint*. Miner., **9**, 24, doi: 10.3390/min9010024.

Asghari O., Sheikhmohammadi S., Abedi M. and Norouzi G.H.; 2016: *Multivariate geostatistics based on a model of geo-electrical properties for copper grade estimation: a case study in Seridune, Iran*. Boll. Geof. Teor. Appl., **57**, 43-58.

Barnett R.M.; 2017: *Projection Pursuit Multivariate Transform*. In: Deutsch J.L. (ed), Geostatistics Lessons, <http://www.geostatisticslessons.com/lessons/lineardecorrelation.html>.

Barnett R.M., Manchuk J.G. and Deutsch C.V.; 2014: *Projection pursuit multivariate transform*. Math. Geosci., **46**, 337-359.

Barnett R.M., Manchuk J.G. and Deutsch C.V.; 2016: *The projection-pursuit multivariate transform for improved continuous variable modelling*. SPE J., **21**, 2010-2026.

Battalgazy N. and Madani N.; 2019: *Categorization of mineral resources based on different geostatistical simulation algorithms: a case study from an iron Ore deposit*. Nat. Resour. Res., in press, doi: 10.1007/s11053-019-09474-9.

Chiles J.P. and Delfiner P.; 2012: *Geostatistics: modelling spatial uncertainty, 2 ed*. Wiley, New York, NY, USA, 734 pp.

Da Silva C.Z. and Costa J.F.; 2014: *Minimum/maximum autocorrelation factors applied to grade estimation*. Stochastic Environ. Res. Risk Assess., **28**, 1929-1938.

Davis J.G.; 1992: *Grade control for Australian open pit gold mines*. Geol. Soc., London, Special Publ., **63**, 219-232, doi: 10.1144/GSL.SP.1992.063.01.22.

Desbarats A. and Dimitrakopoulos R.; 2000: *Geostatistical simulation of regionalized pore-size distribution using min/max autocorrelation factors*. Math. Geol., **32**, 919-942.

Deutsch C.V. and Journel A.G.; 1998: *GSLIB: geostatistical software library and user's guide*. 2nd ed. Oxford University Press, New York, NY, USA, 369 pp.

Emery X.; 2010: *Iterative algorithms for fitting a linear model of coregionalization*. Comput. Geosci., **36**, 1150-1160.

Eze P.N., Madani N. and Adoko A.C.; 2019: *Multivariate mapping of heavy metals spatial contamination in a Cu-Ni exploration field (Botswana) using turning bands co-simulation algorithm*. Nat. Resour. Res., **28**, 109-124, doi: 10.1007/s11053-018-9378-3.

Friedman J.H.; 1987: *Exploratory projection pursuit*. J. Am. Stat. Assoc., **82**, 249-266.

Goovaerts P.; 1997: *Geostatistics for natural resources evaluation*. Oxford University Press, New York, NY, USA, 483 pp.

Goulard M. and Voltz M.; 1992: *Linear coregionalization model: tools for estimation and choice of cross-variogram matrix*. Math. Geol., **24**, 269-286.

Hosseini S.A., Asghari O., Emery Z. and Maleki M.; 2017: *Forecasting the grade-tonnage curves and their uncertainty at the Mehdiabad deposit-Yazd, central Iran*. Boll. Geof. Teor. Appl., **58**, 217-232.

Hotelling H.; 1933a: *Analysis of a complex of statistical variables into principal components*. J. Educ. Psychol., **24**, 417-441.

Hotelling H.; 1933b: *Analysis of a complex of statistical variables into principal components*. J. Educ. Psychol., **24**, 498-520.

Hwang J., Lay S. and Lippman A.; 1994: *Nonparametric multivariate density estimation; a comparative study*. IEEE Trans. on Signal Processing, **42**, 2795-2810, doi: 10.1109/78.324744.

Journel A.G. and Huijbregts C.J.; 1978: *Mining geostatistics*. Academic Press, London, UK, 600 pp.

Khorram F., Asghari O. and Memarian H.; 2017: *Change of support based on modified localized uniform conditioning and direct block multivariate simulation on the Sungun porphyry deposit, Iran*. Boll. Geof. Teor. Appl., **58**, 115-136.

Madani N. and Emery X.; 2019: *A comparison of search strategies to design the cokriging neighborhood for predicting coregionalized variables*. Stochastic Environ. Res. Risk Assess., **33**, 183-199, doi: 10.1007/s00477-018-1578-1.

Madani N., Yagiz S. and Adoko A.C.; 2018: *Spatial mapping of the rock quality designation using multi-Gaussian kriging method*. Miner., **8**, 530, doi: 10.3390/min8110530.

Myers D.E.; 1982: *Matrix formulation of cokriging*. Math. Geol., **14**, 249-257.

Rivoirard J.; 1994: *Introduction to disjunctive kriging and non-linear geostatistics*. Oxford University Press, New York, NY, USA, 181 pp.

Reed M. and Simon B.; 1972: *Methods of modern mathematical physics. I: Functional analysis*. Academic Press, 325 pp.

Rossi M. and Deutsch C.V.; 2014: *Mineral resource estimation*. Springer, New York, NY, USA, 348 pp., doi: 10.1007/978-1-4020-5717-5.

Scerbo F. and Mazzotti A.; 1991: *Geostatistical estimates of porosity from seismic data*. Boll. Geof. Teor. Appl., **33**, 85-110.

Sinclair A.J. and Blackwell G.H.; 2002: *Applied mineral inventory estimation*. Cambridge University Press, New York, NY, USA, 399 pp., doi: 10.1017/S0016756803237560.

Switzer P. and Green A.A.; 1984: *Min/max autocorrelation factors for multivariate spatial imagery*. Stanford University, Department of Statistics, Stanford, CA, USA, Technical Report n. SWI NSF 06, 16 pp.

Verly G.; 2005: *Grade control classification of ore and waste, a critical review of estimation and simulation based procedures*. Math. Geol., **37**, 451-475.

Wackernagel H.; 2003: *Multivariate geostatistics: an introduction with applications*. Springer-Verlag, Berlin Heidelberg, Germany, 403 pp., doi: 10.1007/978-3-662-05294-5.

Appendix

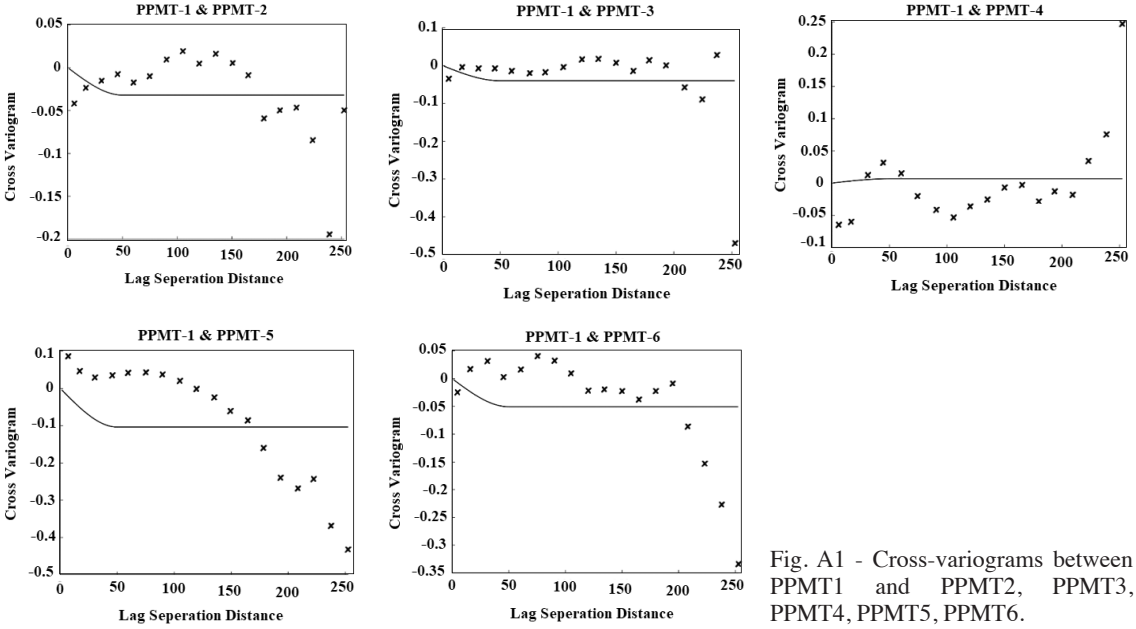


Fig. A1 - Cross-variograms between PPMT1 and PPMT2, PPMT3, PPMT4, PPMT5, PPMT6.

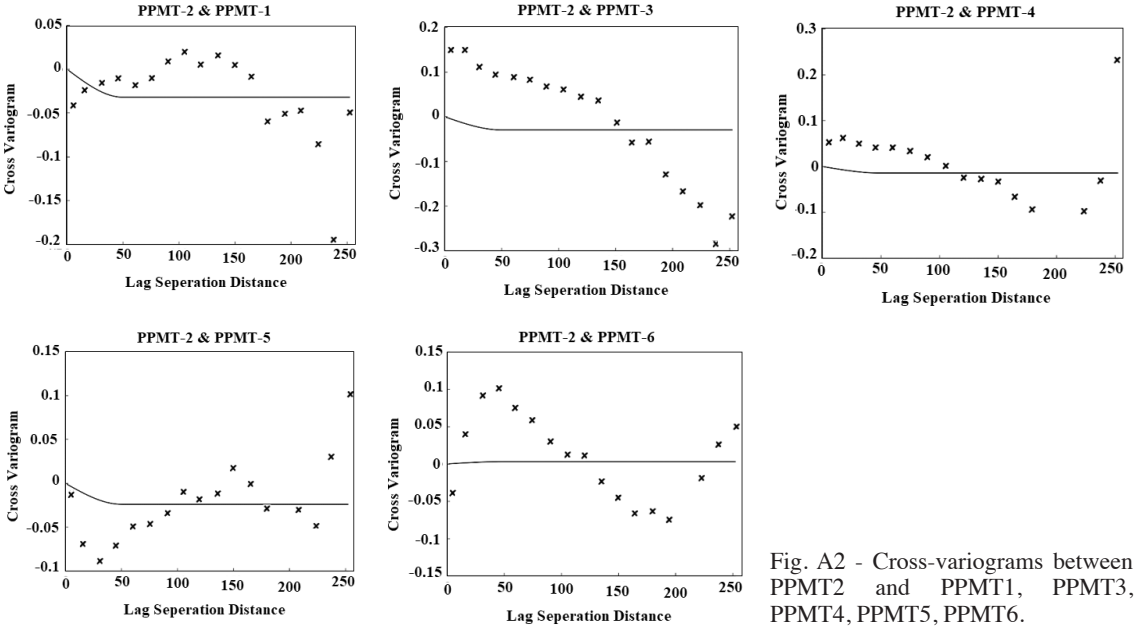


Fig. A2 - Cross-variograms between PPMT2 and PPMT1, PPMT3, PPMT4, PPMT5, PPMT6.

Corresponding author: Nasser Madani
School of Mining and Geosciences, Nazarbayev University
Nur-Sultan 01000, Kazakhstan
Phone: +7 7172 694540; e-mail: nasser.madani@nu.edu.kz