Ordinary Cokriging of Additive Log-Ratios for Estimating Grades in Iron Ore Deposits

M. N. M. BOEZIO[1], J. F. C. L. COSTA[1] and J. C. KOPPE[1]

[1]Laboratory of Mineral Exploration, Mining Engineering Department, Federal University of Rio Grande do Sul (UFRGS), Av. Bento Gonçalves 9500, Bloco IV, Prédio 75, Sala 104, CEP 91501-970, Porto Alegre, Brazil.
maria.morales@ufrgs.br.

Abstract

Risk assessment and economic evaluation of mining projects are mainly affected by the determination of grades and tonnages. In the case of iron ore, multiple variables must be determined for ore characterization which estimation must satisfy the original mass balances and stoichiometry among granulometric fractions and chemical species.

Models of these deposits are generally built from estimates obtained using ordinary kriging or cokriging, most time using solely the global grades and determining the ones present at different granulometric partitions by regression. Alternative approaches include determining the totality of the chemical species and distributing the closing error or leaving one variable aside and determining it by difference afterwards, adding up the error of previous determinations. Furthermore, the estimates obtained are outside the interval of the original variables or even exhibiting negative values. These inconsistencies are generally overridden by post-processing the estimates to satisfy the closed sum condition and positiveness.

In this paper, cokriging of additive log-ratios (*alr*) is implemented to determine global grades of iron, silica, alumina, phosphorous, manganese and loss by ignition and masses of three different granulometric partitions, providing better results than the ones obtained through cokriging of the original variables, with all the estimates within the original data values interval and satisfying the considered mass balances.

Key words: iron ore, additive log-ratios, cokriging, compositional data, geostatistics

1 Introduction

Iron ore quality is characterized by multiple variables: not only the iron grade but also the contaminants that interfere in the subsequent steel manufacturing processes. In addition, granulometric partitions are strictly controlled to meet the products´ specifications. Consequently, multiple variables must be determined, providing block models of estimates that must satisfy the mass balances among granulometric fractions and the stoichiometry in each block.

The classical methodology of multivariate geostatistics, ordinary cokriging (Marechal, 1970), takes into consideration the spatial direct and cross-correlation among variables, leading to models that are more consistent with the natural phenomenon under study.

The inverse correlation of iron and silica present in itabirites can be explained by the genesis of the deposit. However, data from iron ore deposits constitute compositional data, with constrained sums provided both by the mass balances among granulometric partitions and by the stoichiometry. Consequently, most spatial correlations are somehow affected by the constrained sum, being spurious (artificial) (Pearson, 1897; Pawlowsky-Glahn and Olea, 2004).

In 1981, Aitchison introduces the concept that compositional data carry only relative information about the values of the components, so that statements on compositions can be expressed in terms of ratios (or log-ratios, because the logarithmic transformation provides better mathematical properties and ease of manipulation) with the advantage of the log-ratio transformation taking the problem from the simplex (restricted sample space) to the multivariate, unrestricted real space.

The previous concepts are extended to regionalized variables (Matheron, 1965) by Pawlosky-Glahn and Olea (2004), and alert that the cross-covariances matrix, in the case of compositional data, is a singular matrix, such that the application of direct cokriging of the original variables is not possible. Another issue to consider is the negative bias condition introduced when considering the direct covariances. With these previous concepts, and in the presence of compositional data with a constrained sum, cokriging of original data does not provide an unbiased estimate.

General practice for resolving the closure problem when determining grades from iron ore deposits through cokriging of the original variables, consider two alternatives (Goovaerts, 1997): (i) leaving one variable outside the cokriging system and determine it afterwards, (ii) determining all the variables involved and distributing the error to satisfy the constrained sum. Nevertheless, both alternatives lead to an unpredictable amount of obtained estimates that are outside the original data values interval or take negative values. These inadequate values must be post-processed, replaced by valid ones, generally obtained by interpolation methods such as kriging, local means or inverse distance, which either do not take into account the spatial correlation or the and closure constraint.

In this paper, cokriging of additive log-ratios (*alr*) (Pawlowsky-Glahn and Olea, 2004) is implemented to determine global grades of iron, silica, alumina, phosphorous, manganese and loss by ignition and masses of three different granulometric partitions, from an iron ore deposit located in the Ferrous Quadrilateral, Brazil. Comparison with the cokriging estimates obtained from the original data (classical approach) is also presented in the discussion of the results.

2 Methodology

The methodology applied in the present work is ordinary cokriging of additive log-ratios (*alr*) as it is exhaustively presented in Pawlowsky-Glahn and Olea (2004). For further details on multivariate geostatistics fundamentals, refer to Wackernagel (1994), Chilès and Delfiner (1999) and Goovaerts (1997), among others.

3 Case study

The case study comes from a BIF (banded iron formation) iron ore deposit, located in the Ferrous Quadrilateral, Brazil (Figure 1). The original UTM coordinates were rotated both for the data set and geological model, aligned with the principal direction of the orebody.

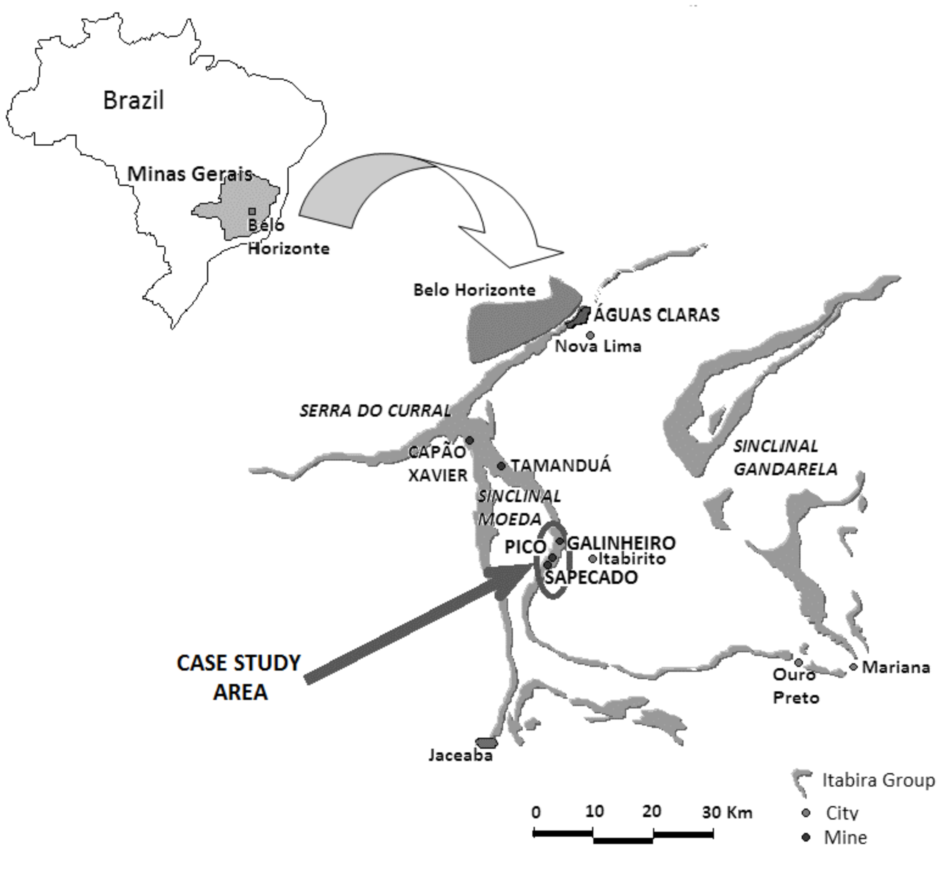


Figure 1 – Map of the case study area, located in the Ferrous Quadrilateral, Brazil, in the Pico Complex, that host the Pico, Galinheiro and Sapecado Mines.

The data set comes from itabirites of various types, of economic importance, with iron grades from 30 to 64% that constitute a geostatistical domain, arbitrarily called IB.

The location map of the samples projected in the XY plane (Figure 2),shows a sample spacing of 50m x 50 m for values of the X coordinate from -7600m to -5500m, and of 200m x 100m along X and Y directions respectively, for X coordinate values greater than -5500 meters.

Evaluation of grain size partition through crushing and screening tests results in three products: Lump Ore (fraction1), Sinter Feed (fraction 2) and Pellet Feed (fraction 3).

Analysis performed at each granulometric partition and at the total sample (global) lead to the concentration of the attributes of interest: mass of the granulometric fractions, grades of iron, alumina, silica, phosphorous and manganese and loss by ignition, namely Wi, FEi, Ali, SIi, Pi, MNi and PPCi, respectively, where index i corresponds to the granulometric partitions i=1, 2, 3 and for the total grade i = T. In the present work, only global grades (i=T) and masses in granulometric partitions (W1, W2 and W3) are considered.

Table 1 presents the basic statistics of the original data, including declustered mean and variance, obtained through moving window method (Isaaks e Srivastava, 1989).

For determining the masses of each granulometric fraction and global grades, it is considered that the closed sum is given by the mass balance and by the stoichiometry, respectively. The masses of the granulometric fractions 1, 2 and 3 add up to 100% (Equation 1).

$$W_T(u) = W_1(u) + W_2(u) + W_3(u) = 100\% \tag{1}$$

Stoichiometry among the chemical species of interest is given by Equation 2.

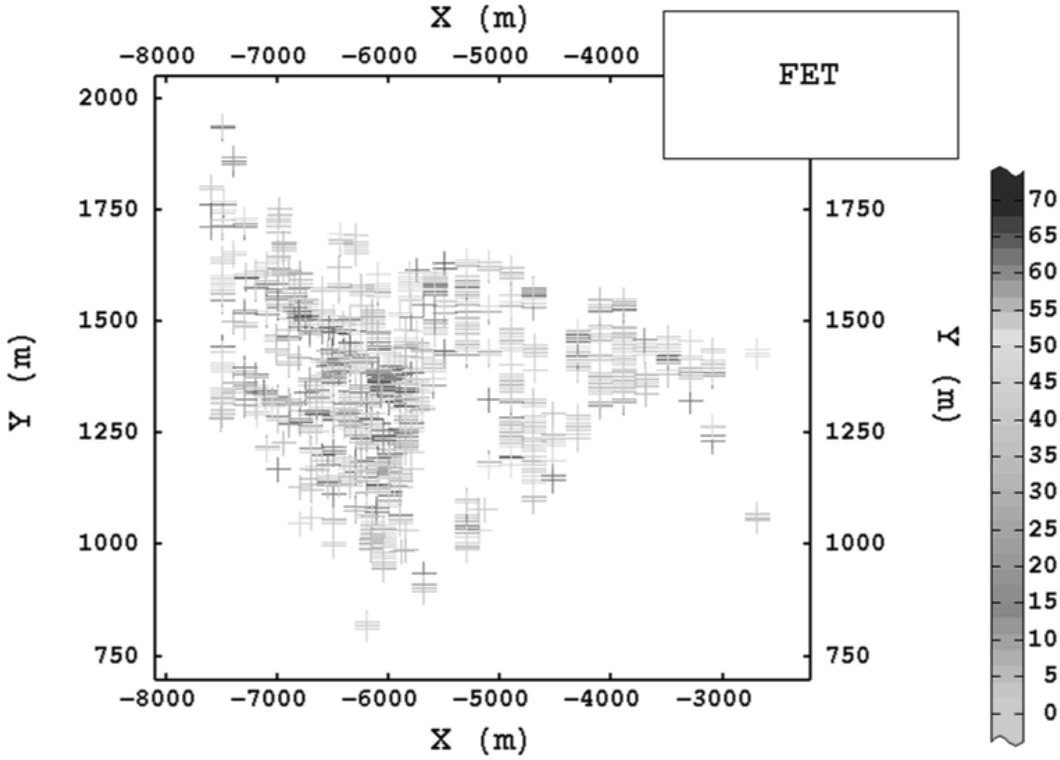


Figure 2 - Location map for samples in IB geostatistical domain (plan view at XY), showing iron global grade (FET).

$$\frac{FE_T(u)}{0.69825} + \frac{P_T(u)}{0.43638} + \frac{MN_T(u)}{0.63193} + AL_T(u) + SI_T(u) + PPC_T(u) = 100 \tag{2}$$

which is rearranged to leave the iron grade as an independent term (Equation 3).

$$\frac{0.69825 \cdot P_i(u)}{0.43638} + \frac{0.69825 \cdot MN_i(u)}{0.63193} + 0.69825 \cdot AL_i(u) + 0.69825 \cdot PPC_i(u) + \\ + 0.69825 \cdot SI_i(u) + FE_i(u) = 69.825 \tag{3}$$

For obtaining the additive log-ratios (*alr*), the last terms in Equation (1) and Equation (3) are arbitrarily chosen to be the divisor in the ratios. Subsequently, spatial continuity of the additive log-ratios is analyzed and modeled through the Linear Model of Corregionalization (LMC) (Goovaerts, 1997; Wackernagel, 1994; Chilès and Delfiner, 1999), as the *alr* transformation removes the spurious correlation, but does not guarantee the spatial decorrelation of the resulting transformed variables.

Estimation of the additive log-ratios *alr* is done through ordinary cokriging at 50 x 20 x 10 m blocks, and back-transforming the estimates to the original sample space (the simplex) through the additive generalized logistic transformation (*agl*) (Pawlowsky-Glahn and Olea, 2004).

3.1 Mass of granulometric fractions

The mass of the granulometric fractions are determined through the additive log-ratios *alr*, Y1BW(u) and Y2BW(u) (Equation 4).

Table 1 – Basic statistics of the original data including declustered mean and variance.

	Num. of Samples	Minimum	Maximum	Mean	Variance	Declustered Mean	Declustered Variance
ALT	1191	0.11	1.99	1.03	0.218	1.05	0.213
FET	1191	30.29	63.98	50.51	69.994	49.46	71.366
MNT	1191	0.01	2.98	0.18	0.166	0.17	0.137
PT	1191	0.01	0.20	0.05	0.001	0.05	0.001
PPCT	1180	0.14	7.09	1.93	1.423	2.04	1.502
SIT	1191	1.44	54.45	24.31	149.516	24.02	154.258
W1	1191	0.01	94.53	18.24	168.718	17.71	162.414
W2	1191	4.46	54.16	28.47	61.285	28.08	58.957
W3	1191	1.00	94.48	53.29	213.411	54.21	199.768

$$Y_{1BW}(u) = \ln\left(W_1(u) \middle/ W_3(u) \right)$$
$$Y_{2BW}(u) = \ln\left(W_2(u) \middle/ W_3(u) \right) \quad (4)$$

The spatial continuity analysis of these transformed variables leads to an anisotropic ellipsoid with its principal and intermediate axes along N68º and N158º in the XY plane, and the minor one perpendicular to this plane, D-90º. The anisotropy of the transformed variables is rotated 32 degrees counter-clockwise from the anisotropic ellipsoid obtained for the original variables.

The Linear Model of Corregionalization (LMC) is presented in

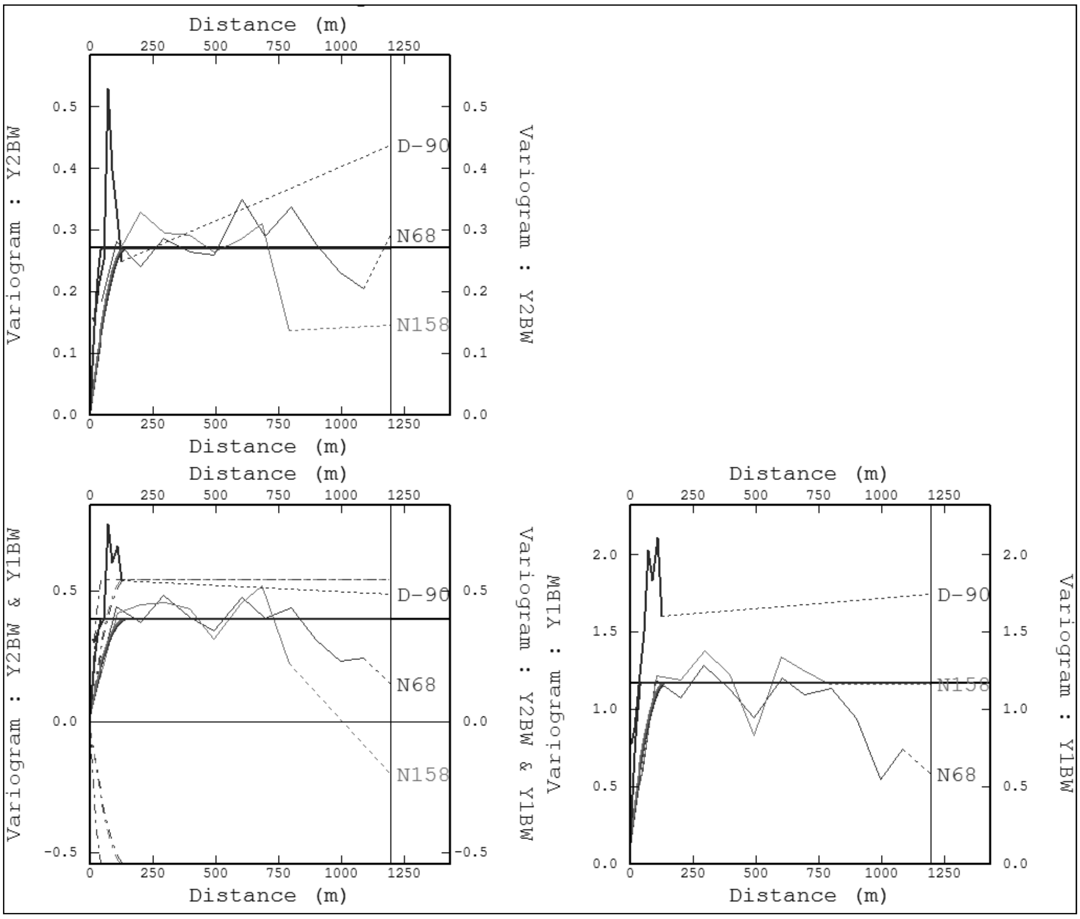
Figure 3 showing the experimental and modeled direct and cross variograms, the LMC equation $\boldsymbol{\gamma}_{BW}$ that includes a nugget effect and two spherical structures *Sph* (Armstrong, 1989; Chilès and Delfiner, 1999), with the corresponding ranges along the three principal anisotropic directions.

Ordinary cokriging is performed in the transformed variables using the neighborhood and search strategies parameters presented in Table 2, which were determined helped by using cross validation (Isaaks and Srivastava, 1989).

Estimates Y*1BW(u) e Y*2BW(u) are back transformed to the original sample space using the inverse transformation *agl* (Pawlowsky-Glahn and Olea, 2004) obtaining the estimated values for W1(u), W2(u) and W(3), W1Y(u), W2Y(u) e W3Y(u), respectively.

Table 2 - Neighborhood and search strategies parameters for ordinary cokriging of the alr Y1BW(u) e Y2BW(u), obtained from the mass balance among granulometric partitions.

Type of neighborhood:	Moving
Search ellipsoid:	N68º, N158º, D-90º
Search distance along N68º:	750m
Search distance along N158º:	220m
Search distance along D-90º:	60m
Number of angular sectors:	8
Minimum number of samples in neighborhood:	3
Optimum number of samples in each angular sector:	2
Heterotopic search:	Yes
Block discretization in X:	5
Block discretization in Y:	5
Block discretization in Z:	1



$$\gamma_{BW} = \mathbf{C}_0 + \mathbf{C}_1 \cdot Sph\left(\frac{70m}{N68^\circ}\ \frac{50m}{N158^\circ}\ \frac{20m}{D-90^\circ}\right) + \mathbf{C}_2 \cdot Sph\left(\frac{140m}{N68^\circ}\ \frac{130m}{N158^\circ}\ \frac{45m}{D-90^\circ}\right)$$

Corregionalization matrix $\mathbf{C}_0$

	Y1BW	Y2BW
Y1BW	0.14	0.03
Y2BW	0.03	0.01

Corregionalization matrix $\mathbf{C}_1$

	Y1BW	Y2BW
Y1BW	0.13	0.04
Y2BW	0.04	0.01

Corregionalization matrix $\mathbf{C}_2$

	Y1BW	Y2BW
Y1BW	0.91	0.32
Y2BW	0.32	0.25

Figure 3 – Linear Model of Corregionalization (LMC) γ_{BW}, with a nugget effect and two spherical structures (*Sph*) to model experimental direct and cross variograms of the additive log-ratios *alr* Y1BW(u) and Y2BW(u), obtained from the mass balance among granulometric partitions, with the corresponding corregionalization matrices $\mathbf{C}_0$, $\mathbf{C}_1$ and $\mathbf{C}_2$.

ISBN: 978-84-87867-76-7

3.2 Global grades

For determining the global grades, the constant sum that is considered comes from the stoichiometry relation in Equation 3, with i=T to consider the global grades.
The additive log-ratios *alr* are determined using the global iron grade as the divisor (Equation 5).

$$\begin{aligned} Y_{1EQT}(u) &= \ln\left(\frac{0.69825 \cdot P_T(u)}{0.43638 \cdot FE_T(u)} \right) \\ Y_{2EQT}(u) &= \ln\left(\frac{0.69825 \cdot MN_T(u)}{0.63193 \cdot FE_T(u)} \right) \\ Y_{3EQT}(u) &= \ln\left(\frac{0.69825 \cdot AL_T(u)}{FE_T(u)} \right) \\ Y_{4EQT}(u) &= \ln\left(\frac{0.69825 \cdot PPC_T(u)}{FE_T(u)} \right) \\ Y_{5EQT}(u) &= \ln\left(\frac{0.69825 \cdot SI_T(u)}{FE_T(u)} \right) \end{aligned} \qquad (5)$$

Analyzing the spatial continuity of the transformed variables, it was mapped the anisotropic ellipsoid with its principal and intermediate axes along directions N100º and N190º at the XY plane, respectively, and the minor axe is perpendicular to this plane at D-90º, being these the anisotropic directions of original data.

In this case, modeling of the LMC (γ_{EQT}) has the difficulty of simultaneously fitting five direct variograms and the corresponding ten cross variograms satisfying the positive definiteness conditions of the corregionalization matrices. Figure 4 shows the model equation and corresponding corregionalization matrices.

Ordinary cokriging is then performed using the same neighborhood and search strategies parameters in Table 2, but with the search ellipsoid rotated to the new anisotropic directions (N100º, N190º and D-90º), and the estimates are back-transformed to the original sample space, using the *agl* inverse transformation.

$$\gamma_{EQT} = \mathbf{C}_0 + \mathbf{C}_1 \cdot Sph\left(\frac{190m}{N100^\circ} \frac{100m}{N190^\circ} \frac{20m}{D-90^\circ} \right) + \mathbf{C}_2 \cdot Sph\left(\frac{950m}{N100^\circ} \frac{270m}{N190^\circ} \frac{50m}{D-90^\circ} \right)$$

Corregionalization matrix $\mathbf{C}_0$

	Y1EQT	Y2EQT	Y3EQT	Y4EQT	Y5EQT
Y1EQT	0.13	0.09	0.06	0.10	0.00
Y2EQT	0.09	0.39	-0.02	0.07	0.01
Y3EQT	0.06	-0.02	0.13	0.07	0.06
Y4EQT	0.10	0.07	0.07	0.10	-0.01
Y5EQT	0.00	0.01	0.06	-0.01	0.25

Corregionalization matrix $\mathbf{C}_1$

	Y1EQT	Y2EQT	Y3EQT	Y4EQT	Y5EQT
Y1EQT	0.01	-0.03	-0.01	-0.01	0.00
Y2EQT	-0.03	0.14	0.05	0.01	-0.03
Y3EQT	-0.01	0.05	0.02	0.00	-0.02
Y4EQT	-0.01	0.01	0.00	0.01	0.01
Y5EQT	0.00	-0.03	-0.02	0.01	0.02

Corregionalization matrix $\mathbf{C}_2$

	Y1EQT	Y2EQT	Y3EQT	Y4EQT	Y5EQT
Y1EQT	0.22	0.13	0.05	0.15	0.03
Y2EQT	0.13	1.29	0.11	0.14	0.06
Y3EQT	0.05	0.11	0.14	0.10	0.02
Y4EQT	0.15	0.14	0.10	0.44	0.01
Y5EQT	0.03	0.06	0.02	0.01	0.32

Figure 4 – Linear Model of Corregionalization (LMC) γ_{EQT}, with a nugget effect and two spherical structures (*Sph*) to model experimental direct and cross variograms of the additive log-ratios alr Y1EQT(u), Y2EQT(u), Y3EQT(u), Y4EQT(u) and Y5EQT(u), obtained from the stoichiometry relation among global grades, with the corresponding corregionalization matrices $\mathbf{C}_0$, $\mathbf{C}_1$ and $\mathbf{C}_2$.

4 Discussion of results

In the discussion of results, the unbiasedness of the estimates is analyzed observing the reproduction of the global and local mean. The global mean reproduction is verified comparing the declustered mean of the original data with the mean of the estimates. For the reproduction of the local mean, the conditional expectation of the estimates and original data are plotted together along principal directions X, Y and Z. In this paper, only X direction is shown.

Another aspect to be analyzed is the presence of negative values and the satisfaction of the constant sums.

Comparison with estimates obtained performing ordinary cokriging of the original variables is also presented. In this case, the number of variables to have their spatial correlation jointly modeled through the LMC increases in one unit compared to the ones modeled by ordinary cokriging of the additive log-ratios *alr*.

Basic statistics of the estimates obtained in this case study are relatively similar (Table 3). The global mean of the estimates is similar to the declustered mean of the original data (Table 1), showing unbiasedness of the estimators.

However, it can be observed that the minimum value for manganese global grade (MNT), obtained through cokriging of the original variables, is negative. Further analysis leads to a total of 500 blocks that have negative values for this attribute.

In the case of the local mean, Figure 5 shows the conditional expectation along X coordinate for the original data and estimates obtained both through cokriging of the original variables and cokriging of the additive log-ratios *alr*. Fluctuations of the local means of the estimates follow the ones of the original data.

Table 3 – Basic statistics of the estimates obtained by ordinary cokriging of the original data (CK) and by ordinary cokriging of the additive log-ratios *alr* (CK*alr*).

	CK				CK*alr*			
	Minimum	Maximum	Mean	Variance	Minimum	Maximum	Mean	Variance
ALT	0.35	1.71	1.12	0.05	0.25	1.9	1.07	0.083
FET	33.89	62.16	50.43	18.63	33.5	62.99	51.64	24.484
MNT	-0.03	1.64	0.19	0.05	0.01	2.16	0.11	0.024
PPCT	0.24	4.63	2.31	0.94	0.24	5.53	2.22	1.16
PT	0.02	0.11	0.05	0.01	0.01	0.11	0.05	0.0003
SIT	7.72	48.38	23.87	38.52	4.67	49.2	22.69	62.695
W1	2.88	48.26	17.98	28.02	1.19	58.19	16.02	39.356
W2	14.24	42.13	28.32	12.57	9.09	51.39	29.16	23.185
W3	24.61	79.24	53.70	38.44	19.19	89.18	54.82	66.01

The constant sums are perfectly satisfied for the totality of the estimates in the case of cokriging of the additive log-ratios *alr* (Figure 6), with estimates that are in the original sample space where the constant sum is maintained (the simplex). The constant sums obtained through the estimates obtained by ordinary cokriging of the original data ordinary are adequate, but can vary in the case of another data set, because ordinary cokriging does not guarantee that the constant sum is satisfied.

ISBN: 978-84-87867-76-7

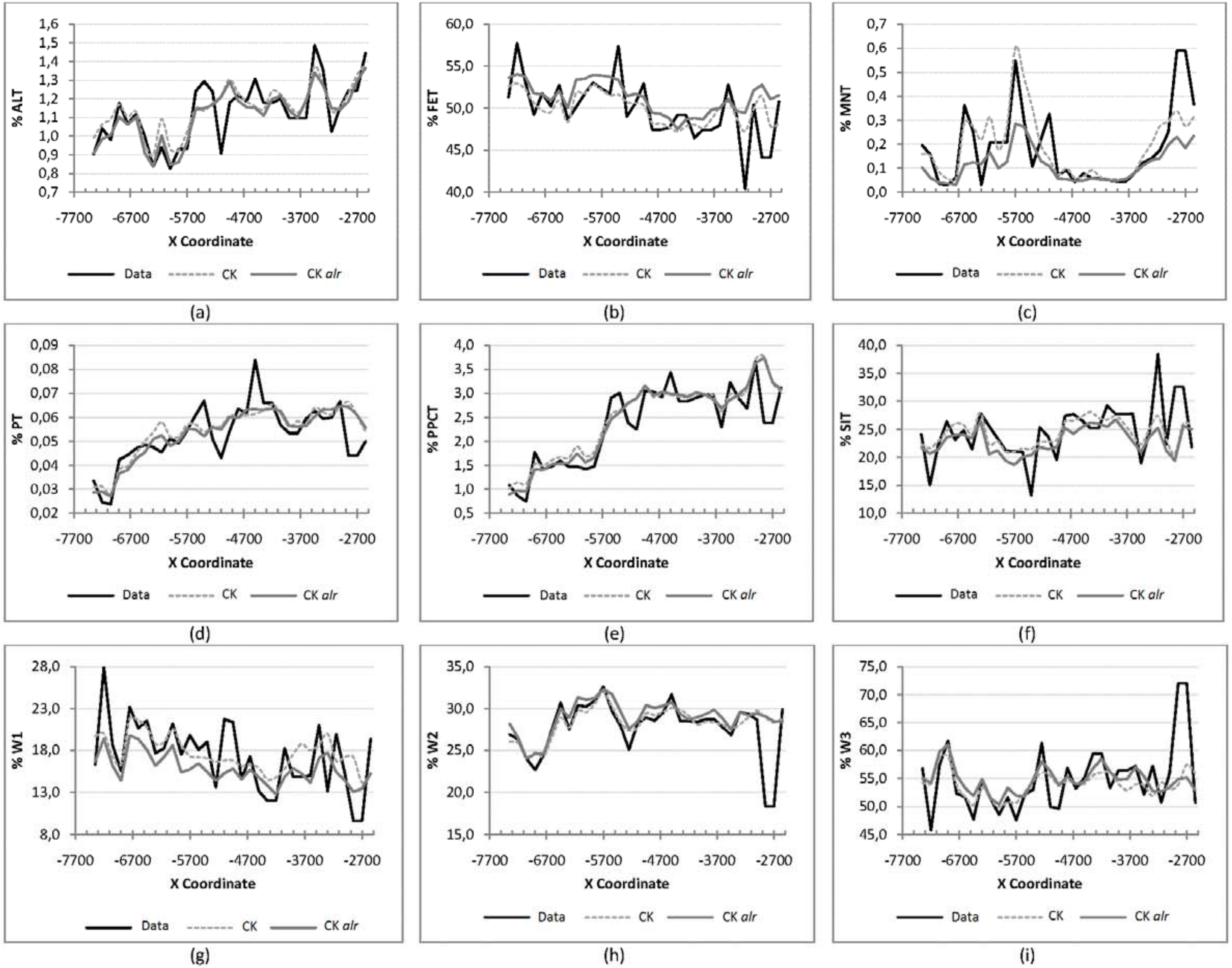
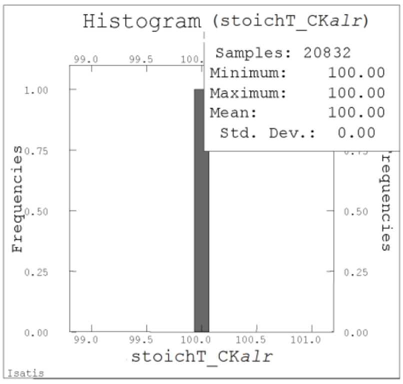
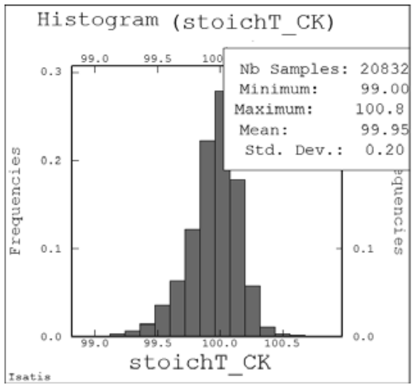


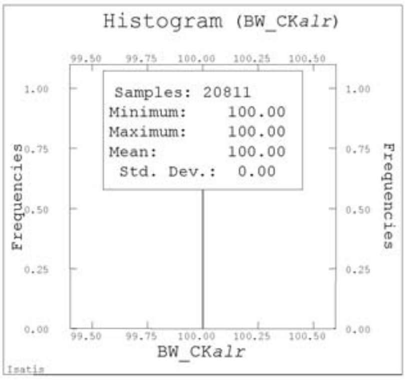
Figure 5 - Conditional expectation diagrams of original data (Data), estimates obtained by cokriging of original data (CK) and by cokriging of additive log-ratios (CK*alr*) along X direction considering bands of 150m, for total grades of (a) alumina, (b) iron, (c) manganese, (d) phosphorous, (e) loss by ignition, (f) silica (ALT, FET, MNT, PT, PPCT and SIT) and mass of granulometric partitions (g) 1, (h) 2 and (i) 3 (W1, W2 and W3).



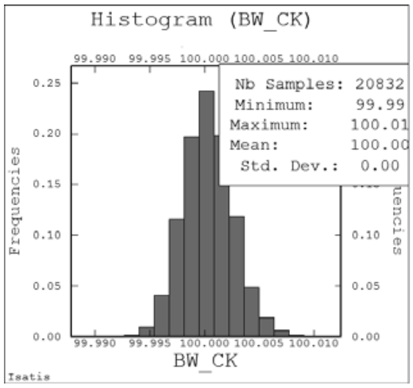
(a)



(b)



(c)



(d)

Figure 6 - Histograms of the constant sum given by the stoichiometry of global grades (stoichT) and by the mass balances BW obtained by cokriging of the additive log-ratios *alr* (CK*alr*) (a) and (c) and (b) and (d) cokriging of the original data (CK), respectively.

5 Conclusions

In the mining industry, data coming from iron ore deposits and other ones such as bauxite and phosphate constitute regionalized compositions as the regionalized variables satisfy mass balances. When various species are involved in the final product quality, as in the case of iron ore, stoichiometry also leads to regionalized compositions. Moreover, most times the same specie simultaneously satisfies mass balances and stoichiometry, although this is not the situation presented in this paper.

Common practice for quantifying grades and tonnages does not consider that the variables involved constitute a regionalized composition. There are various palliative actions used to ensure the constant sums and to eliminate the negative values: (i) all the variables involved are considered to be in intrinsic correlation with the same scale of spatial correlation, which is not necessarily true; (ii) negative values are substituted by arbitrary values, obtained through interpolators with an arbitrary variogram model such as local means or inverse to the distance. This post-processing of the estimates varies from data set to data set, and it is not possible to predict the amount of negative values or the dispersion of the closed sum value.

The sample space of the attributes of this kind of deposits is the D-simplex (D is the number of attributes being considered), where the sum is constrained to a constant value. The transformation into additive log-ratios *alr*, leads to a new sample space being the real space of dimension (D-1), where it is possible to use the classic methodologies of multivariate geostatistics. When back-transforming the estimates to the original sample space through the *agl* transformation, the estimates are again in their sample space, the D-simplex. For this reason, cokriging of the additive log-ratios *alr* guarantees the satisfaction of the balances considered: because the obtained estimates are restricted to the simplex.

The results obtained through cokriging of the additive log-ratios *alr* are highly satisfactory, reproducing the global and local mean, with positive estimates within the original data values interval, satisfying the constant sum condition for the totality of the estimates. In addition, the Linear Model of Corregionalization (LMC) to be modeled is one unit smaller than the one obtained from the original variables, and consequently, easier to model.

In the case of ores that present closed sums, as iron ore, is highly recommended to obtain the estimates that constitute the block models through cokriging of the additive log-ratios alr instead of cokriging of the original variables, or by any other methodology that takes into account that the data set sample space is not the real space but the simplex.

References

Aitchison, J. (1981). A new approach to null correlations of proportions. *Mathematical Geology*, 13(2), 175-189

Armstrong, M. (1989). *Basic Linear Geostatistics*, Springer Verlag, Berlin, 256p.

Chilès, J. P. and Delfiner, P. (1999). *Geostatistics: modeling spatial uncertainty*. Wiley-Interscience, New York, 695p.

Goovaerts, P. (1997). *Geostatistics for Natural Resources Evaluation*. Oxford University Press, New York, 483p.

Isaaks, E. H. e Srivastava, R. M. (1989). *An Introduction to Applied Geostatistics*. Oxford University Press, New York, 561p.

Marechal, A. (1970). *Cokrigeage et regression em correlation intrinsique*. Centre de Geostatistique de Fontainebleau, Fontainebleau, 40p.

Matheron, G. (1965). *Les variables regionalisées et leur estimation*. Ed. Masson, Paris, 306p.

Pawlowsky-Glahn, V. and Olea, R. A. (2004). *Geostatistical Analysis of Compositional Data*. Studies in Mathematical Geology, Oxford University Press, 181p.

Pearson, K. (1897). Mathematical contributions to the theory of evolution on a form of spurious correlation which may arise when indices are used in the measure of organs. *Proceedings of the Royal Society of London*, LX, 489-502

Wackernagel, H. (1994). *Multivariate Geostatistics. An introduction with applications*. Springer-Verlag, Berlin, 275p.