Integration of different-quality data in short-term mining planning

Integração de dados de diferentes qualidades no planejamento de curto prazo

http://dx.doi.org/10.1590/0370-44672015680212

Cristina Paixão Araújo
Ms Candidate, Chemist
Federal University of Rio Grande do Sul- UFRGS
Mining Engineering Department
Porto Alegre - Rio Grande do Sul - Brazil
cristinapaixaoaraujo@yahoo.com.br

João Felipe Coimbra Leite Costa
PhD, Professor
Federal University of Rio Grande do Sul- UFRGS
Mining Engineering Department
Porto Alegre - Rio Grande do Sul - Brazil
jfelipe@ufrgs.br

Abstract

Decisions, from mineral exploration to mining operations, are based on grade block models obtained from samples. This study evaluates the impact of using imprecise data in short-term planning. The exhaustive Walker Lake dataset is used and is considered as the source for obtaining the true grades. Initially, samples are obtained from the exhaustive dataset at regularly spaced grids of 20 × 20 m and 5 × 5 m. A relative error (imprecision) of ±25% and a 10% bias are added to the data spaced at 5 × 5 m (short-term geological data) in different scenarios. To combine these different types of data, two methodologies are investigated: cokriging and ordinary kriging. Both types of data are used to estimate blocks with the two methodologies. The grade tonnage curves and swath plots are used to compare the results against the true block grade distribution. In addition, the block misclassification is evaluated. The results show that standardized ordinary cokriging is a better methodology for imprecise and biased data and produces estimates closer to the true grade block distribution, reducing block misclassification.

keywords: biased samples, grade estimates, kriging, cokriging.

Resumo

Decisões relacionadas à exploração mineral e à operação da mina estão subordinadas aos modelos de blocos obtidos a partir das amostras. Esse estudo avalia o impacto do uso de dados imprecisos no planejamento de curto prazo. O banco de dados exaustivo Walker Lake foi usado e considerado como o teor real do depósito. Inicialmente, as amostras foram obtidas de banco de dados com espaçamento regular de 20 × 20 m e 5 × 5 m. O erro relativo de ±25% (imprecisão) e 10% de viés foram adicionados aos dados espaçados a 5 × 5 m (dados geológicos curto prazo) em diferentes cenários. Para combinar esses diferentes dados (precisos e exatos em 20 × 20 m e imprecisos e enviesados em 5 × 5 m), duas metodologias foram investigadas: cokrigagem e krigagem ordinária. As curvas teor tonelagem e análise de deriva foram utilizadas para comparar os resultados com a distribuição de real dos blocos. Além disso, a classificação errônea dos blocos foi avaliada. Os resultados mostraram que o uso da cokrigagem ordinária estandarizada é a melhor metodologia em situações que existem dados imprecisos e enviesados e as estimativas produzidas são mais próximas da distribuição real dos blocos, reduzindo o erro de classificação dos blocos.

Palavras-chave: amostras com viés, teor das estimativas krigagem, cokrigagem.

1. Introduction

In the mining industry, decisions from mineral exploration through to mining operations are based on grade block models obtained from samples. It is common for data to be collected in various formats, and consequently, have varying precision and accuracy. During the exploration stage, samples are obtained from diamond-made drillholes, which are of high quality and are usually associated with negligible sampling errors. During production, samples are usually collected from blastholes, which may lead to large sampling errors in terms of either bias or precision (GOOVAERTS, 1997: GY, 1998). From a geostatistical perspective, this difference in precision has to be taken into account so that the two types of data can be combined. The aim of this paper is to investigate two geostatistical methodologies for integration of data: ordinary kriging (JOURNEL & HUIJBREGTS, 1978) and standardized ordinary cokriging (GOOVAERTS, 1997).The estimates for each scenario were compared with a reference block grade model. As blasthole sampling is mainly used for short-term grade models, the results emphasize the impact of block misclassification.

2. Material and method

2.1 Data presentation

This study uses the exhaustive Walker Lake dataset (ISAAKS & SRIVASTAVA, 1989) with 78 000 point support samples distributed regularly at 1 × 1 m. The variable V was used and the original unit was rescaled so that it resembled grades from a copper mineral deposit. To obtain the reference block grade distribution, the exhaustive point support dataset was averaged into 3210 blocks of size 5 × 5 m. These blocks represented the true block grades, and were used for comparison.

In this case study, the data were totally heterotopic and of varying quality. Two types of data were considered. First, samples were obtained from the exhaustive point support dataset at a regular spacing of 20 × 20 m. These samples were precise and accurate and mimicked diamond-drillholed samples. Other samples were obtained from the exhaustive point support dataset at a regular spacing of 5 × 5 m, and imprecision and bias were added. Figure 1 shows the regular spacing between samples in this case study.

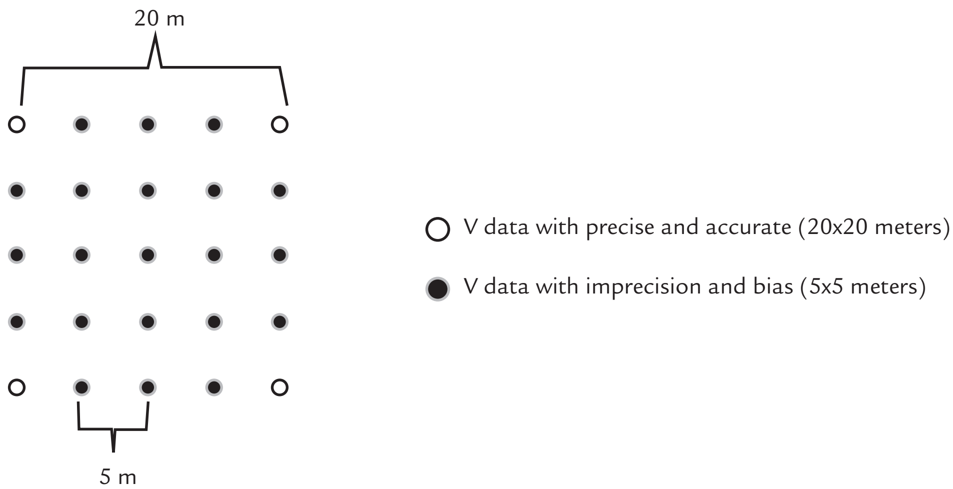

Figure 1
Dataset with a regular spacing grid.

2.2 Adding imprecision

Sampling aims to obtain a representative sample, i.e., one that is accurate and precise. For good-quality data, it is expected that for the measured values, there is no trend (or bias) in the errors; i.e., the errors should have zero mean and follow a Gaussian distribution. When there is a systematic tendency in the errors to over- (or under-) estimate the real values, the estimator begins to produce systematically biased estimates of the true value of the attribute, and the associated error distribution departs from having a zero mean.

2.3 Adding variance (increasing imprecision)

The original samples obtained from the exhaustive dataset at a 5 × 5 m regular spacing were disturbed by adding (or subtracting) a random relative sampling error of 10%. The relative error was assumed to be a standard Gaussian with zero mean and with a standard deviation determined by the product of the relative error (±10%) and the grade (MAGRI & ORTIZ, 2000). These samples mimicked blasthole samples, which have a poorer precision than diamond-drillholed data. The error was assumed to be heteroscedastic, which is frequently the case in practice (GOOVAERTS, 1997; MATHERON, 1963).

2.4 Adding error bias

Not only was the imprecision of the secondary data increased, but also a bias was added. For a 25% bias, the grades had their mean either increased or decreased by 25% (MAGRI & ORTIZ, 2000).

Table I shows the summary statistics of the reference point support dataset (V_Real_points), the reference block support dataset (V_Real_blocks), and the sample dataset with accurate and precise data (V_20×20). The sample datasets had means very close to the true mean, which indicates that there were no biases or precision errors. The data with bias and precision errors (V_5×5_+25% and V_5×5_–25%) had means that were 25% greater (or smaller) than those of the reference block distribution (V_Real_blocks), to mimic the situation in which poor-quality data induce biases that are subsequently transferred to the grade estimation process.

Data	No. of samples	Mean	Standard deviation	CV	Minimum	Maximum
V_Real_points	78 000	2.78	2.50	0.90	0	16.31
V_Real_blocks	3 120	2.78	2.29	0.82	0	13.78
V_20×20	195	2.80	2.48	0.88	0	10.13
V_5×5	3 120	2.77	2.49	0.89	0	16.10
V_5×5_+25%	2 925	3.44	3.12	0.90	0	18.13
V_5×5_−25%	2 925	2.07	1.90	0.91	0	13.61

Table 1
Summary statistics for the original reference and for the biased and imprecise secondary data.

2.5 Estimation Methodologies

Two methodologies were evaluated for the estimation of block grades: ordinary kriging and standardized ordinary cokriging.

2.5.1 Ordinary Kriging

Scenario 1 used only the accurate and precise data (primary) for estimation with ordinary kriging. The dataset (V_20×20) comprised 195 samples that were 20 × 20 m apart. The variogram in this case was defined by equation (1) below. In the other cases, the two types of data were pooled together.

The second alternative that was tested ignored the differences in accuracy and precision between the two sources of information (primary and secondary). Scenario 2 combined accurate and precise data (V_20×20) and data with biases and precision errors added (V_5×5_+25%), whereas scenario 3 pooled together accurate and precise data (V_20×20) and data with systematic lower grades (bias) and greater imprecision (V_5×5_−25%). Accurate data and data with biases and precision errors were combined for estimation purposes. The variograms in this case were defined by equations (2) and (3) below.

Scenario 1: Ordinary kriging with 195 precise and accurate data (V_20×20):

$$\gamma_v(h) = 1.0 + 2.0 \cdot \text{Sph}\,(1) \cdot \left(\frac{N157.5E}{90\text{ m}}, \frac{N67.5E}{40\text{ m}}\right) + 2.92 \cdot \text{Sph}\,(2) \cdot \left(\frac{N157.5E}{120\text{ m}}, \frac{N67.5E}{60\text{ m}}\right) \quad (1)$$

Scenario 2: Ordinary kriging with 3120 data of different quality pooled together (V_20×20 and V_5×5_+25%, imprecise and inaccurate):

$$\gamma_v(h) = 1.50 + 3.0 \cdot \text{Sph}\,(1) \cdot \left(\frac{N157.5E}{35\text{ m}}, \frac{N67.5E}{25\text{ m}}\right) + 5.10 \cdot \text{Sph}\,(2) \cdot \left(\frac{N157.5E}{80\text{ m}}, \frac{N67.5E}{44\text{ m}}\right) \quad (2)$$

Scenario 3: Ordinary kriging with 3120 data of different quality pooled together (V_20×20 and V_5×5_−25%, imprecise and inaccurate):

$$\gamma_v(h) = 0.7 + 1.3 \cdot \text{Sph}\,(1) \cdot \left(\frac{N157.5E}{37\text{ m}}, \frac{N67.5E}{28\text{ m}}\right) + 1.78 \cdot \text{Sph}\,(2) \cdot \left(\frac{N157.5E}{82\text{ m}}, \frac{N67.5E}{41\text{ m}}\right) \quad (3)$$

2.5.2 Standardized ordinary cokriging estimator

Standardized ordinary cokriging was thoroughly explained by GOOVAERTS (1997). This is a suitable framework for incorporating data of variable quality. It takes into consideration the spatial auto- and cross-correlations among the variables involved. The method also filters bias from the inaccurate dataset, proceeding with standardized residuals instead of the original data. It uses standardized ordinary cokriging (GOOVAERTS, 1997), in which the sum of the weights for the primary and secondary variables is 1. Spatial continuity is defined using the linear model of coregionalization (LMC). During the cokriging process, the LMC controls the weights allocated to the secondary data. The cross-correlation is obtained using the cross-covariance, since the cross-variogram requires collocated data (isotopic multivariate dataset). If the cross-covariance resembles the direct variograms, the secondary data are treated as primary in terms of weights during the cokriging procedure (MINNITT & DEUTSCH, 2014).

There is obviously no nugget effect on the cross-covariance, since there are no collocated data. However, the intersection point on the y-axis can be obtained only by extrapolating the trend of the covariance at nonzero lags. Equations (4), (5), and (6) below show the spatial direct and cross-correlations for scenario 4, where the primary variable is the V sample (accurate and precise), whereas the secondary variable comprises additional samples from V with biases (higher means) and precision errors added (V_5×5_+25%). Scenario 5 includes the primary variable V sample (accurate and precise) (V_20×20), whereas the secondary variable is the V sample with biases (lower means) and precision errors added (V_5×5_−25%). The LMC adjustment is shown in equations (7), (8), and (9).

$$\gamma_{(1)}(h) = 1.8 + 1.8 \cdot Sph\,(1) \cdot \left(\frac{N157.5E}{90\ m}, \frac{N67.5E}{40\ m}\right) + 2.32 \cdot Sph\,(2) \cdot \left(\frac{N157.5E}{120\ m}, \frac{N67.5E}{60\ m}\right) \quad (4)$$

$$\gamma_{(2)}(h) = 2 + 3.5 \cdot Sph\,(1) \cdot \left(\frac{N157.5E}{90\ m}, \frac{N67.5E}{40\ m}\right) + 4.20 \cdot Sph\,(2) \cdot \left(\frac{N157.5E}{120\ m}, \frac{N67.5E}{60\ m}\right) \quad (5)$$

$$\gamma_{(12)}(h) = 0.0 + 2.5\ Sph\,(1) \cdot \left(\frac{N157.5E}{90\ m}, \frac{N67.5E}{40\ m}\right) + 3.00 \cdot Sph\,(2) \cdot \left(\frac{N157.5E}{120\ m}, \frac{N67.5E}{60\ m}\right) \quad (6)$$

$$\gamma(1) = 1.8 + 1.8 \cdot Sph\,(1) \cdot \left(\frac{N157.5E}{90\ m}, \frac{N67.5E}{40\ m}\right) + 2.32 \cdot Sph\,(2) \cdot \left(\frac{N157.5E}{120\ m}, \frac{N67.5E}{60\ m}\right) \quad (7)$$

$$\gamma(2) = 0.7 + 1.8 \cdot Sph\,(1) \cdot \left(\frac{N157.5E}{90\ m}, \frac{N67.5E}{40\ m}\right) + 1.5 \cdot Sph\,(2) \cdot \left(\frac{N157.5E}{120\ m}, \frac{N67.5E}{60\ m}\right) \quad (8)$$

$$\gamma(12) = 0.00 + 1.7 \cdot Sph\,(1) \cdot \left(\frac{N157.5E}{90\ m}, \frac{N67.5E}{40\ m}\right) + 1.8 \cdot Sph\,(2) \cdot \left(\frac{N157.5E}{120\ m}, \frac{N67.5E}{60\ m}\right) \quad (9)$$

3. Results and discussion

Figure 2 shows scatterplots of the estimated (using all tested estimation methods and data) and the true block grades. It also shows the basic statistics of the two distributions and the slope of the regression line between them.

Figure 2a shows the estimates using ordinary kriging with few, but accurate and precise, data. The mean of the estimated grades is close to that of the reference block grades, and the slope of the regression line is close to 1. However, as expected, the standard deviation is smaller than that of the reference block grade distribution because of the smoothing effect of kriging. This scenario uses sparsely spaced samples, which leads to a severe reduction in the variance of the block estimates.

When we used ordinary kriging combining data of different quality, the results were not good. The mean of the estimates was biased and did not reproduce the mean of the reference block

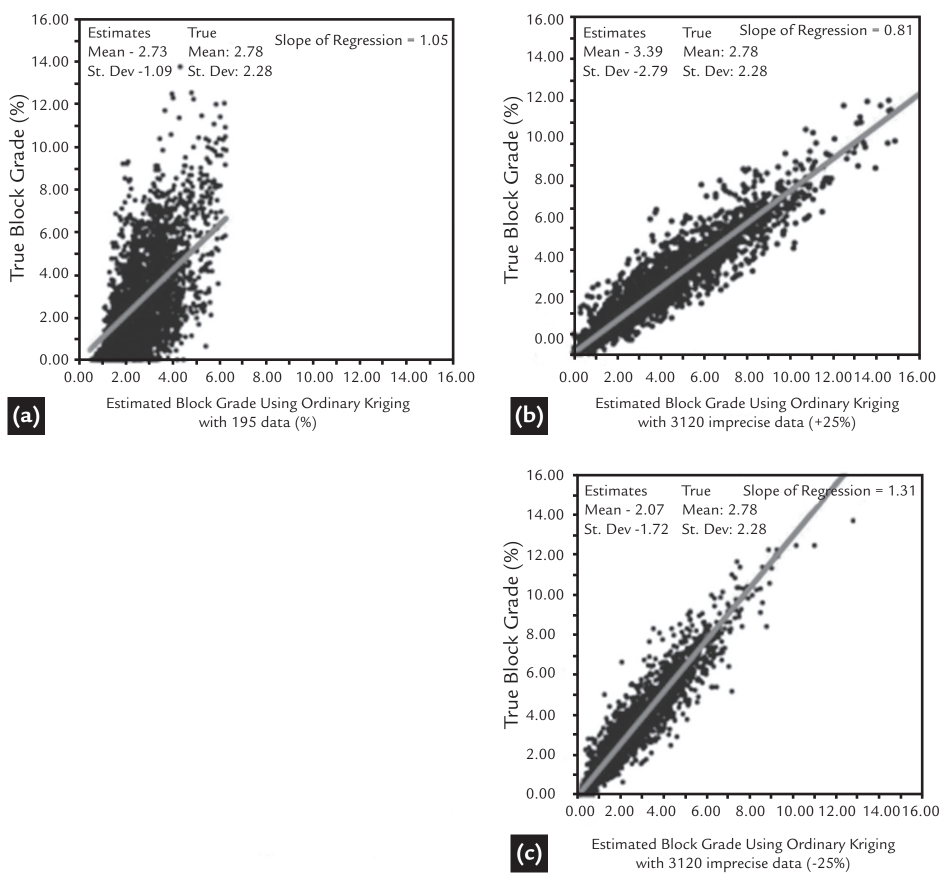

Figure 2
Scatterplot between the estimated block grades and the true block grades using ordinary kriging:
(a) accurate and precise data;
(b) imprecise and inaccurate data combined with the secondary data having systematically higher readings (V_20×20 and V_5×5_+25%);
(c) imprecise and inaccurate data combined with the secondary data having systematically lower readings (V_20×20 and V_5×5_−25%).

grade distribution. These results were to be expected, and they are shown here to highlight the effect on the model estimates that one can expect by ignoring and combining data of different precision and accuracy in a blind manner. The slope of the regression line of the estimates versus the true value shows a clear bias. In scenario 2 (Figure 2b), the mean and the standard deviation of the estimates are higher than the mean of the reference block grades. The slope of the regression line is less than 1, showing that the results systematically overestimate the true grades.

Scenario 3 (Figure 2c) shows the opposite: the mean and the standard deviation of the estimates are underestimated compared with the reference block grades. The slope of the regression line is greater than 1, showing that the results under-estimate the real grades. The proposed solution is inappropriate, since it can lead to either over- or under-estimation of the blocks, depending on how the bias in the secondary data affects the grades.

When ordinary cokriging was used to estimate and integrate data of different quality (scenarios 3 and 4; Figure 3a and 3b), the results were far superior. The mean of the estimates reproduced the mean of the reference block grades, and the slope of the regression line was close to 1.

The standard deviation of the estimates was lower than that of the reference block grades.

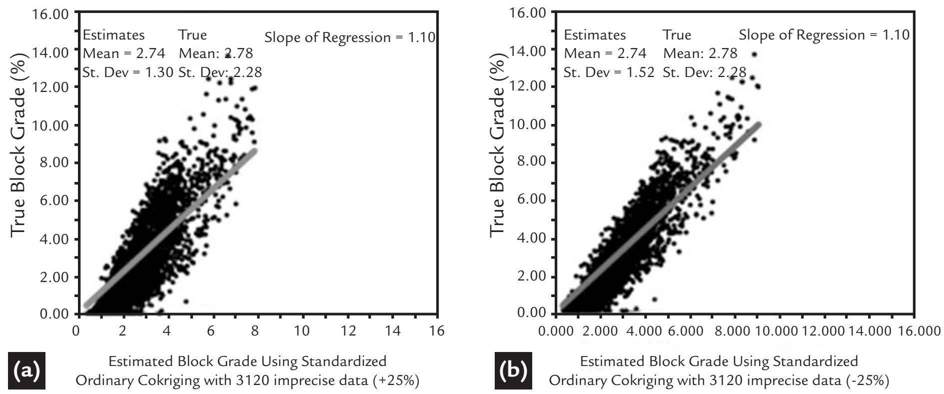

Figure 3
Scatterplot between the estimated block grades and the reference block model using standardized ordinary cokriging: (a) primary variable precise and accurate data (V_20×20) combined with secondary variable biased and imprecise error data (V_5×5_−25%); (b) primary variable precise and accurate data (V_20×20) combined with secondary variable biased and imprecise error data (V_5×5_−25%).

Figure 4 shows the grade tonnage curves for the reference block grade model and for the estimates. As expected, the greater the smoothing effect, the greater the deviation from the reference grade tonnage curve. Consequently, the estimates using ordinary kriging with few precise and accurate data produced a poorer grade tonnage curve. The grade above cutoff predicted by ordinary kriging was under-estimated. Also, the largest deviations of the predicted tonnage from the true model occurred with the ordinary kriging block model. The best results were achieved with ordinary cokriging. For all the cutoffs, the ordinary cokriging grade tonnage curve is the closest to the reference curve.

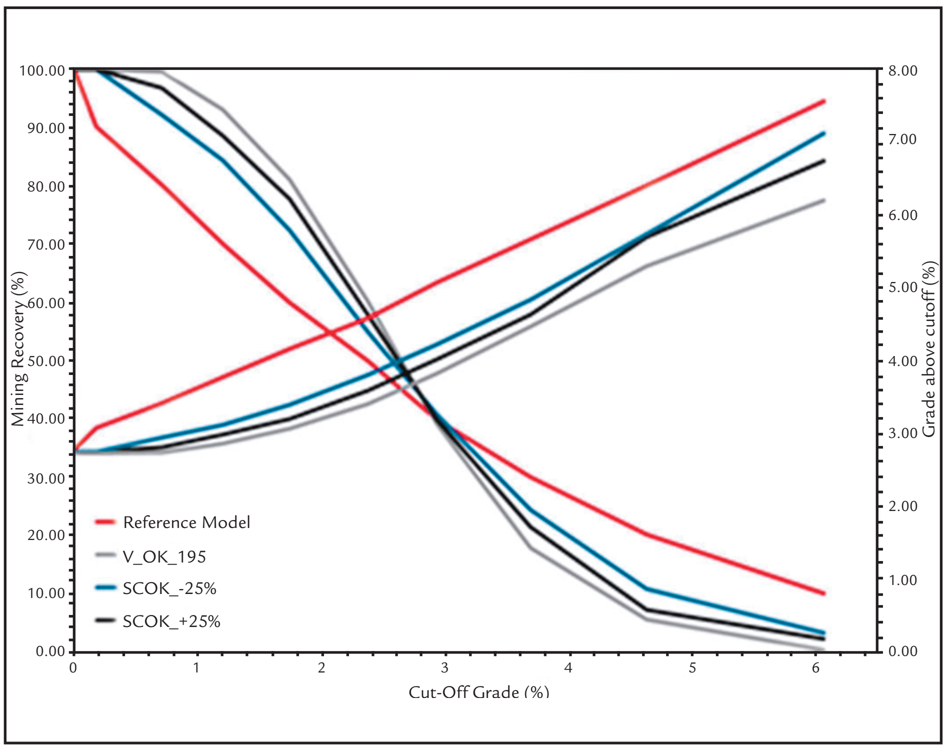

Figure 4
Grade tonnage curves.

Figure 5 shows the swath plots with different estimation methodologies, and it can be seen that the results estimated using ordinary cokriging are closest to the reference model.

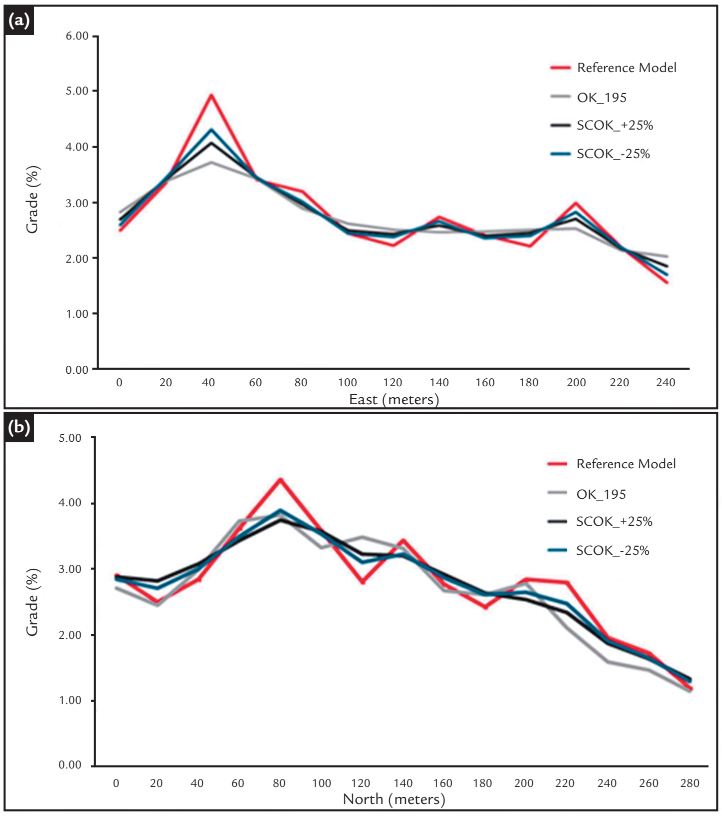

Figure 5
Swath plots
between reference model and estimates:
(a) eastings;
(b) northings.

Figure 6 shows the total number of misclassified blocks, the number of ore blocks classified as waste, and the number of waste blocks classified as ore with the grade models obtained with each methodology. Five cutoffs were considered: 0.93%, 1.73%, 2.35%, 4.24%, and 5.34%. It can be seen from Figure 6 that standardized ordinary cokriging generated the best result in terms of block misclassification for all the cutoffs considered. The difference is evident. At 1.73%, the number of misclassified blocks is approximately 620 using standardized ordinary cokriging. In the case of ordinary kriging with accurate and precise data, the number increases to 875. However, ordinary kriging erroneously sent much more waste to the plant, causing dilution. Even worse, the ordinary kriging model sent far more ore blocks to the waste pile. The better results with regard to block misclassification shown by ordinary cokriging are consistent with the scatterplots between true and estimated block values (Figure 3).

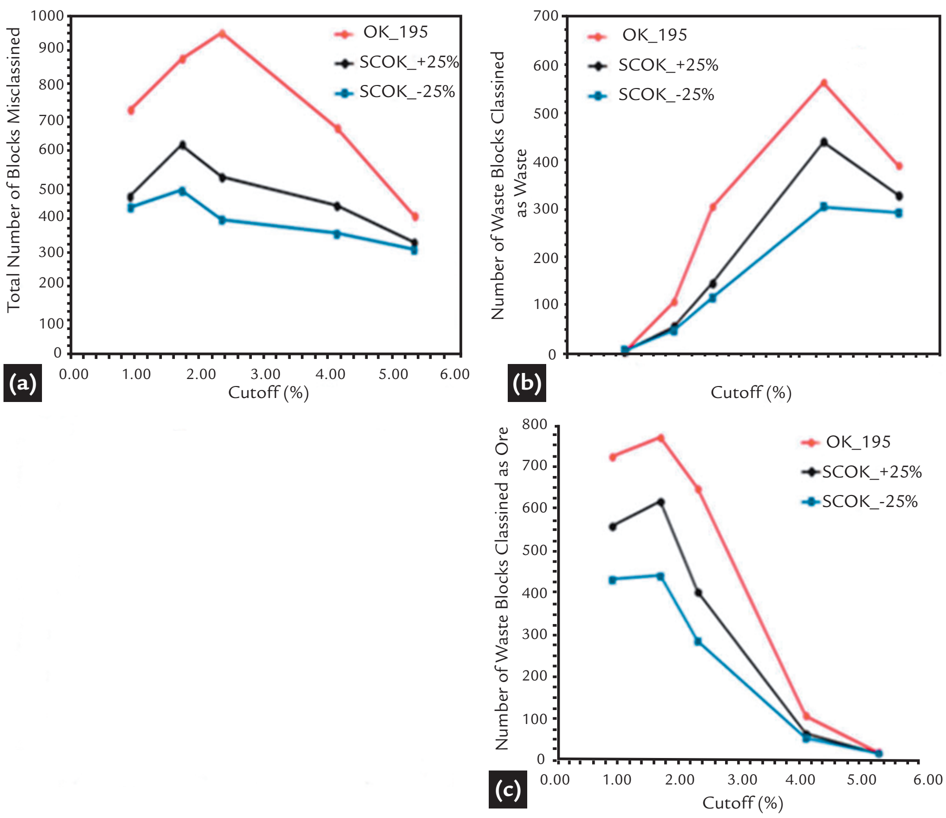

Figure 6
Block misclassification:
(a) total number of blocks misclassified;
(b) number of ore blocks
classified as waste;
(c) number of waste blocks
classified as ore.

4. Conclusion

Samples obtained by different methods lead to data with different error levels (bias and imprecision). This dissimilarity has to be considered when integrating the two sources of information. In the study reported here, two methodologies, namely, ordinary kriging and ordinary cokriging, were used to incorporate data of different quality.

For samples with bias and precision error, ordinary kriging led to poor results strongly affected by sample bias. In the case of ordinary kriging, the best option was to use only those samples that were accurate and precise, discarding biased data.

Ordinary cokriging produced good results, since the poor-quality samples were considered secondary information and their mean was filtered (standardized values). The results of the estimates were less smooth than those obtained with ordinary kriging, which led to estimates closer to the true block grades. In a comparison of ordinary kriging using only data that were accurate and precise (but few in number), ordinary cokriging combining biased and imprecise data, and ordinary standardized cokriging, the latter is favored.

Using the grade block model resulting from standardized cokriging led to grade tonnage curves and swath plots similar to those derived from the reference block grade model. The use of this methodology also drastically reduced the number of blocks misclassified.

Even with less precise and biased samples, improvements in block grade estimates, mining recovery, and block misclassification were achieved, since an appropriate methodology was used to include this source of data. Cokriging provided better estimates of recoverable resources at local and global scales, even when poor-quality samples were used (as secondary data). Bias and imprecision in the secondary data were not transferred to the estimates.

5. Acknowledgments

The authors would like to thank the Brazilian National Research Agency (CNPq) for supporting the Mineral Exploration and Mining Planning Research Unit (LPM) at the University Federal do Rio Grande do Sul.

6. References

GOOVAERTS, P. ***Geostatistics for natural resources evaluation.*** New York: Oxford University Press, 1997.

GY, P. ***Sampling for analytical purposes.*** Chichester: John Wiley & Sons, 1998.

ISAAKS, E. & SRIVASTAVA, M. An introduction to applied geostatistics. New York: Oxford University Press, 1989.

JOURNEL, A., HUIJBREGTS, C. ***Mining geostatistics.*** New York: Academic Press, 1978.

MAGRI, E., ORTIZ, J. Estimation of economic losses to poor blast hole sampling in open pits. In: KLEINGELD, W. J, KRIGE, D. G. eds. Geostatistics 2000: Proceedings of the 6th International Geostatistics Congress: Cape Town. 2000. V. 2, p. 732–741.

MATHERON, G. Principles of geostatistics. ***Economic Geology***, El Paso, v. 58, p. 11246-11266, 1963.

MINNITT, R.C.A., DEUTSCH, C.V., Cokriging for optimal recoverable reserve estimates in mining operations. ***Journal of the Southern African Institute of Mining and Metallurgy*** (SAIMM), v. 114, p. 189–203, March 2014.

Received: 18 November 2014 - Accepted: 9 February 2015.