

Improving short-term grade block models: alternative for correcting soft data

<http://dx.doi.org/10.1590/0370-44672016710007>

Cristina da Paixão Araújo

Doutoranda, MSc,
Universidade Federal do Rio Grande do Sul - UFRGS
Departamento de Engenharia de Minas
Porto Alegre – Rio Grande do Sul - Brasil
cristinapaixaoaraujo@yahoo.com.br

João Felipe Coimbra Leite Costa

Professor Titular, Engenheiro de Minas, MSc, PhD
Universidade Federal do Rio Grande do Sul - UFRGS
Departamento de Engenharia de Minas
Porto Alegre – Rio Grande do Sul - Brasil
jfelipe@ufrgs.br

Vanessa Cerqueira Koppe

Professor Auxiliar, Engenheiro de Minas, MSc, Dr
Universidade Federal do Rio Grande do Sul - UFRGS
Departamento de Engenharia de Minas
Porto Alegre – Rio Grande do Sul - Brasil
vkoppe@ufrgs.br

Abstract

Short-term mining planning typically relies on samples obtained from channels or less-accurate sampling methods. The results may include larger sampling errors than those derived from diamond drill hole core samples. The aim of this paper is to evaluate the impact of the sampling error on grade estimation and propose a method of correcting the imprecision and bias in the soft data. In addition, this paper evaluates the benefits of using soft data in mining planning. These concepts are illustrated via a gold mine case study, where two different data types are presented. The study used Au grades collected via diamond drilling (hard data) and channels (soft data). Four methodologies were considered for estimation of the Au grades of each block to be mined: ordinary kriging with hard and soft data pooled without considering differences in data quality; ordinary kriging with only hard data; standardized ordinary kriging with pooled hard and soft data; and standardized, ordinary cokriging. The results show that even biased samples collected using poor sampling protocols improve the estimates more than a limited number of precise and unbiased samples. A well-designed estimation method corrects the biases embedded in the samples, mitigating their propagation to the block model.

Keywords: biased samples, grade estimates, kriging, cokriging, sampling error.

1. Introduction

Mineral deposit mining decisions are based in part on the information provided in the grade block models obtained from samples. To decrease the uncertainties of these grade estimates, short-term mining planning staff increase the sampling density in order to provide accurate and precise predictions. Using a larger number of samples improves the reliability of the resulting estimates. During the exploration stage, sampling is typically obtained from diamond drill holes (DDHs). The DDH technique is costly but provides accuracy and precision. High-quality data is sparse during the exploration stage. In the production stage, sampling is typically performed using other techniques due to budget constraints and the need for rapid data acquisition and model updates. Generally, these production samples are of

lower quality than DDHs and are not rigorously prepared and controlled, as they are not subject to a QAQC protocol (quality assurance / quality control). These results can contain large sampling errors. Given that two sources of information with differing quality levels are available, we want to understand how to use the low-quality data to update the block model while avoiding bias propagation that affects the final results.

The goals of this article are to evaluate the impact of sampling error-affected data quality on block classification by identifying methodologies that can mitigate bias in the data, and to propose a method of correcting imprecision and bias in soft data. In addition, it evaluates the benefits of using soft data in mine planning. The methodologies tested were illustrated us-

ing a case study of a gold mine where Au grades obtained from core samples via diamond drilling (hard data) and from samples obtained via channels (soft data) were available. Four methods for estimating the Au grades at each block to be mined were investigated.

The results of the different methods were compared using three statistical approaches. They were the linear correlation coefficient (ρ), the slope of the linear regression ($y=bx$) measured from the scatter-plot between the estimates and the reference data set, and the absolute error (Costa, 1997). These statistical approaches were used to confirm the accuracy and precision of the estimation methods. The number of misclassified blocks, i.e. ore classified as waste, waste as ore, and their sum was also determined for each of the models.

2. Methodology

Four methodologies were evaluated for block grade estimation: ordinary kriging with hard and soft

data pooled without considering differences in data quality; ordinary kriging with only hard data; standardized

ordinary kriging with pooled hard and soft data; and standardized, ordinary cokriging.

Method 1- Ordinary kriging with pooled hard and soft data:

Ordinary kriging was thoroughly explained by GOOVAERTS (1997).

In this method, the hard and soft data were pooled without considering dif-

ferences in data quality. The modeled variogram is

$$\gamma_1(h) = 4.48 + 14.57 \cdot Sph(1) \cdot \left(\frac{h_1}{6.00m}, \frac{h_2}{6.00m}, \frac{h_3}{7.00m}\right) + 3.36 \cdot Sph(2) \cdot \left(\frac{h_1}{33.00m}, \frac{h_2}{16.00m}, \frac{h_3}{12.00m}\right) \quad (1)$$

Method 2- Ordinary kriging with only hard data:

This method used only accurate, precise hard data to conduct estimates with

ordinary kriging. The variogram is same as that used in Method_1 (Equation 1).

Method 3- Standardized ordinary kriging with pooled hard and soft data:

The differences in sample support (volume and delimitation) between the channel (soft data) and diamond drilling (hard data) methods indicate global and conditional biases. In our estimate, the hard and soft data were

pooled together, and a correction factor was used to mitigate the bias. This estimation workflow can be understood as a form of co-kriging in a situation where hard and soft data are strongly correlated. In the first step, soft data

($Z_j(u_{\alpha_j})$) was standardized (see Equation 2) using its mean (m_j) and standard deviation (σ_j). The transformation shown in Equation 2 leads to a mean of zero and standard deviation of unity in the transformed data:

$$Z_j(u_{\alpha_j})^* = \frac{[Z_j(u_{\alpha_j}) - m_j]}{\sigma_j} \quad (2)$$

Next, the soft standardized data $Z_j(u_{\alpha_j})^*$ was rescaled to match the hard

data statistics (see Equation 3) using its mean (m_i) and standard deviation (σ_i).

Thus, the means of the hard and soft data are now equal.

$$Z_j(u_{\alpha_j})^*_T = Z_j(u_{\alpha_j})^* \cdot \sigma_i + m_i \quad (3)$$

From a geostatistical viewpoint, this difference in data precision and accuracy must be considered when integrating the two data types. This proposal relies on

standardizing the soft data (Equation 2) using its declustered mean and standard deviation. Next, the soft data is rescaled (Equation 3) using the mean and standard

deviation from the hard data.

The variogram used for these estimates was the same that presented in Equation (1).

Method 4- Standardized Ordinary Cokriging

Standardized ordinary cokriging is thoroughly explained by GOOVAERTS (1997). This is a suitable framework for incorporating data sets of varying quality levels. It considers spatial auto and cross-correlations among the variables involved. This method also filters bias from inac-

curate datasets, and uses standardized residuals instead of the original data. This article uses standardized ordinary cokriging, in which the sum of the weights of the primary and secondary variables is 1. Spatial continuity is defined using the linear model of coregionalization (LMC).

During cokriging, the LMC controls the weights allocated to the soft data. A cross-correlation is obtained using the cross-covariance, since the cross-variogram requires collocated data (an isotopic multivariate dataset). The adjusted LMC is shown in Equations (4), (5), and (6).

$$\gamma_1(h) = 4.40 + 10.25 \cdot Sph(1) \cdot \left(\frac{h_1}{2.72}, \frac{h_2}{0.75m}, \frac{h_3}{2.00m}\right) + 7.76 \cdot Sph(2) \cdot \left(\frac{h_1}{31.48m}, \frac{h_2}{26.37m}, \frac{h_3}{14.00m}\right) \quad (4)$$

$$\gamma_{12}(h) = 3.68 + 3.95 \cdot \left(\frac{h_1}{2.72}, \frac{h_2}{0.75m}, \frac{h_3}{2.00m}\right) Sph(1) + 1.68 \cdot Sph(2) \cdot \left(\frac{h_1}{31.48m}, \frac{h_2}{26.37m}, \frac{h_3}{14.00m}\right) \quad (5)$$

$$\gamma_2(h) = 9.30 + 4.65 \cdot Sph(1) \cdot \left(\frac{h_1}{2.72}, \frac{h_2}{0.75m}, \frac{h_3}{2.00m}\right) + 9.14 \cdot Sph(2) \cdot \left(\frac{h_1}{31.48m}, \frac{h_2}{26.37m}, \frac{h_3}{14.00m}\right) \quad (6)$$

The four methods were tested using the same variogram model (Methods 1, 2, and 3) fitted on the diamond drill hole dataset (Equation 1) to evaluate the efficacy of each methodology. The estimates were performed on

10x10x10 m blocks discretized using 5x5x5 points within each block, arranged along the north (x), east (y), and vertical (z) axes. The search strategy used a maximum of two samples per angular sector of the search ellipsoid.

Eight angular sectors were used with a minimum distance of 2 m between samples. The same search neighborhood (using octants) and range from the primary variograms were used for each of the four estimates.

3. Case Study

This case study uses data from one geological domain with samples that

mimic two sampling techniques. The data generated via DDH and channel

approaches is referred to as hard and soft data, respectively.

3.1 Creating a reference block

A reference model that mimics a real, unknown deposit was obtained via geostatistical simulation using the Turning Bands algorithm (JOURNAL, 1974). This model was derived from unconditional simulations. No data was used to generate the model, however a histogram and variogram were used as input. The grid nodes were sampled, since the grades within the

mineralized domain reproduced the statistical input. The sampling scheme followed a typical long-term sampling pattern when using DDH and a short-term pattern when using channel samples.

Table 1 shows statistics from the reference point support model derived from unconditional simulations, and the DDH (hard data) samples used to perform

estimations. Also the statistics for the reference block support model obtained by upscaling the point simulated model at 10 m blocks are depicted. A few drill holes were virtually drilled on the reference point simulated model providing 311 samples. The drill hole collars are spaced at an average of 10x10 m and sampled at every vertical meter down the hole.

	Number of Data Points	Mean	Std. Dev.	CV	Min.	Max.
simulated model (point support)	125766	2.27	4.91	2.16	0.03	33.47
DDH samples	311	2.41	4.46	1.85	0.03	33.47
simulated block support model	477	2.31	1.30	0.92	0.30	8.82

Table 1
Statistics from the original and block reference data.

3.2 Creating bias and imprecision

The reference point simulated values were disturbed by adding random Gaussian errors of +/- 25%. Bias was also added increasing the grades by 25%. These biased samples mimic what frequently occurs within channel samples in gold deposits. The error was

assumed to be heteroscedastic, i.e. the variance increases with the mean, as commonly occurs (GOOVAERTS, 1997; MATHERON, 1963).

This imprecise and biased reference point model was used as a source for soft data sampling. An additional set of chan-

nels was sampled virtually in this grade model using a 3 x 3 x 1 m grid.

Table 2 shows statistics from the reference point support biased and imprecise dataset used as the soft data source. Channel samples in this dataset mimic poor-quality, bias-affected data.

Data	Number of Data Points	Mean	Std. Dev.	CV	Min.	Max.
Reference biased and imprecise model	125766	2.83	6.24	2.20	0.00	72.25
Channel samples	1246	3.69	7.91	2.14	0.02	51.56

Table 2
Statistics from the original reference, and from the biased and imprecise soft data.

4. Results and discussion

Figure 1 shows scatter plots that relate the estimates (using all tested estimation methods and data) to the reference model thought to represent the true block grades. In addition, Figure 1 shows the global mean and the standard deviation of each model, as well as the linear correlation coefficients (ρ) and slopes of the linear regressions ($y = bx$) between the estimates and the reference. Figures 1a, 1b, and 1c show estimates determined using ordinary kriging.

Figure 1a shows ordinary kriging results produced when hard and soft data are pooled together, and differences in data quality are ignored. The results are clearly unsatisfactory. The means and standard deviations of the estimates are higher than those of the reference model. The estimates are biased due to systematic overestimation. This solution

is not recommended for use in updating the short-term geological model.

Figure 1b shows estimates produced using ordinary kriging with a small number of accurate, precise data points. The means of the estimated grades are similar to those of the reference block grades, and the slope of the regression (0,48) and linear correlation coefficient (0,29) are low.

When estimates are made using hard and soft data combined with standardized ordinary kriging (Figure 1c), the linear correlation coefficient (0,60) is one of the highest encountered from the methods considered and the slope of the linear regression (0,73) is close to 1. Additional bias-corrected data leads to better estimates, increasing the efficacy of the correlation. This model exhibits the best overall results, as well as reduced conditional bias. Method 3 uses corrected soft

data as a primary variable in the estimation process. It receives the same weight as the hard data, and thus increased its influence on the estimates.

Figure 1d shows that estimates made using standardized ordinary cokriging exhibit a lower correlation (0,32) and slope (0,35) than those produced via standardized ordinary kriging with a combination of hard and soft data. This is probably caused by two factors: poor correlation between hard and soft data and the resulting low weights given to soft data when cokriging is used. In this case study, the cross-correlation is 0,60 when $h = 0$. The minimum correlation coefficient limits required for cokriging are not clearly stated in literature. However, a correlation coefficient that exceeds 0.7 favors method 4, while one below 0.2 does not produce good results. Between

0.2 and 0.7, the results may not be of high quality. (MINNITT; DEUTSCH, 2014)

The cross correlograms indicated by Equation 5 show that approximately 78% of the total cross-covariance of the phenomenon (sum of the nugget effect and first structure of the correlogram) deteriorates quickly at the first lags, i.e.

has low spatial continuity of approximately 2 m. This leads to low soft data weights, even when the data is similar to that of the points being estimated. This high variability is related to the presence of sampling and preparation errors in the soft data (greater inaccuracy), which increases the nugget effect.

This higher imprecision is clear when the variograms are compared: the nugget effect from the soft data (Equation 6) is 20% higher than that from the hard data (Equation 5), thus affecting the weights received by the soft data. In addition, soft data is irregularly spaced along the area studied.

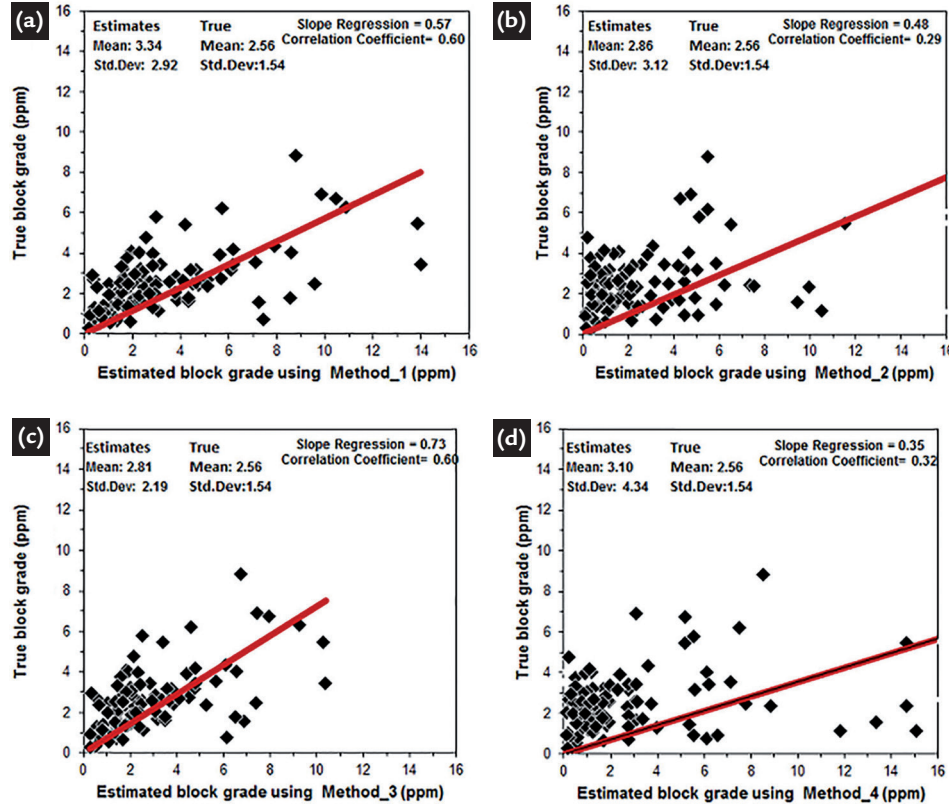


Figure 1
Scatter-plot of estimates vs. the reference block model for: a) Method 1_ Ordinary kriging with hard and soft data, without consideration of data quality; b) Method 2_ Ordinary kriging with only hard data; c) Method 3_ Standarized ordinary kriging with hard and soft data; and d) Method 4_ Standarized ordinary cokriging.

Figure 2 shows cut-off grade x tonnage and cut-off x average grade curves for the reference block grade (red lines), as well as for estimated models. The estimates made using Method 1 (ordinary kriging with hard and soft data combined – light blue lines) produce a poorer grade

tonnage curve. Ordinary kriging overestimates the grades above cut-off. Also, the largest deviations from the tonnage predicted by the true model occur with the ordinary kriging block model. For all cutoffs, the best results are obtained from the grade tonnage curve produced using

Method 3 (standardized ordinary kriging with hard and soft data pooled – black lines). The result produced is closest to the reference curve. These results show that soft data may improve short-term geological mine planning when an appropriate methodology is used to integrate it.

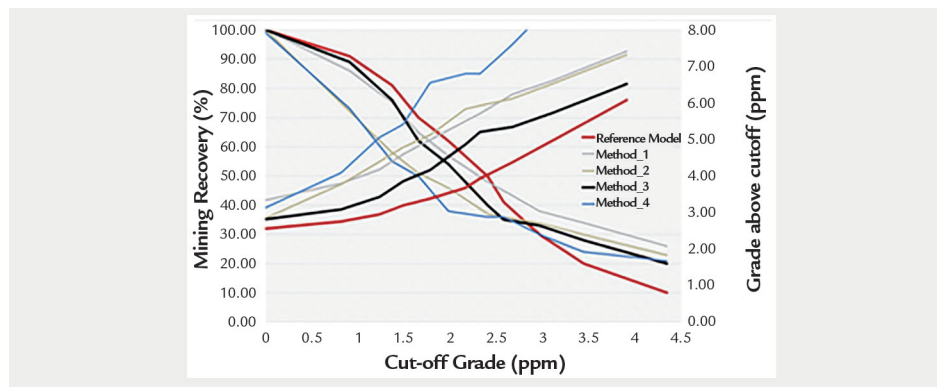


Figure 2
Cut-off grade x tonnage and cut-off x average grade curves for the reference block grade model (red lines) and the estimates.

Figure 3 shows a histogram of the errors of each estimation method considered. The error median was chosen for bias assessment, as it tends to be statistically less sensitive to extreme

values. The median error for ordinary kriging using only hard data is -0,52 (Figure 3a), while using ordinary kriging with hard and corrected soft data produces an error of -0,17 (Figure 3b),

and standardized ordinary cokriging produces an error of -0,65 (Figure 3c). A smaller bias was obtained using ordinary kriging with hard and bias-corrected soft data.

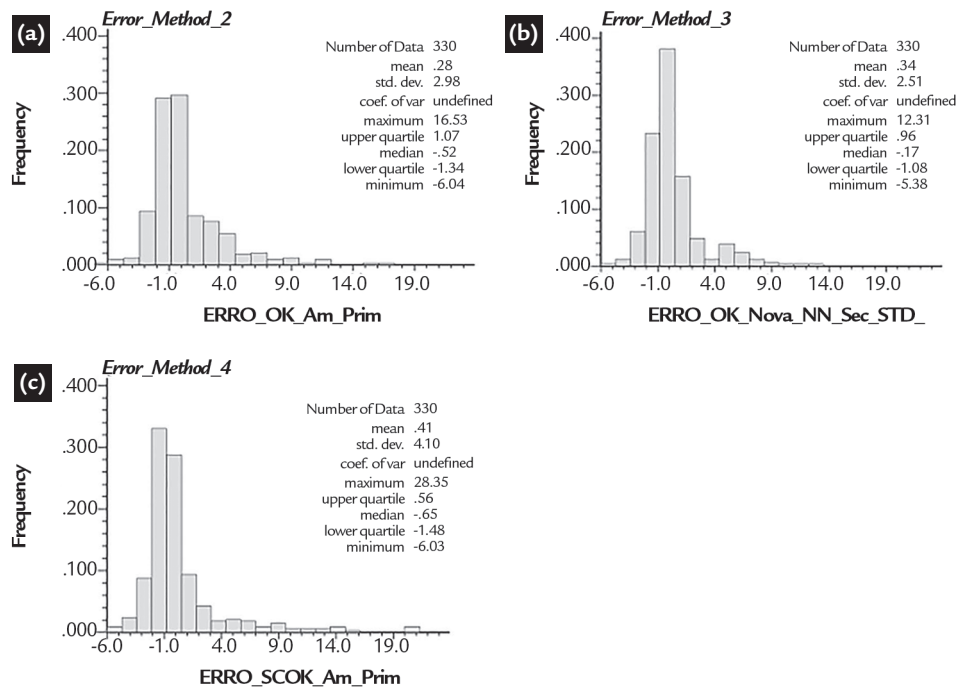


Figure 3 Histogram showing the error of each estimation method considered a) Method 2_Ordinary kriging with only hard data; b) Method 3_Standarized ordinary kriging with hard and soft data; and c) Method_4 Standardized ordinary cokriging.

Figure 4 shows the total number of misclassified blocks, the number of ore blocks classified as waste, and the number of waste blocks classified as ore. Four cutoffs were considered: 1.18 g/t (Q20), 1.69 g/t (Q40), 2.43 g/t (Q60), and 3.16 g/t (Q80). The standardized ordinary

kriging model using hard and soft data (black line) minimized block misclassification. This difference can be seen with the cut-off grade 1.69 g/t in Figure 4b, where Method 4 erroneously classifies 25 more ore blocks as waste than Method 3. Use of an adequate estimation workflow

reduces the number of ore blocks disposed of, and consequently provides more ore blocks for processing, increasing mine metal recovery. Thus, use of a more precise methodology leads to better decision-making when choosing destinations for the mined blocks.

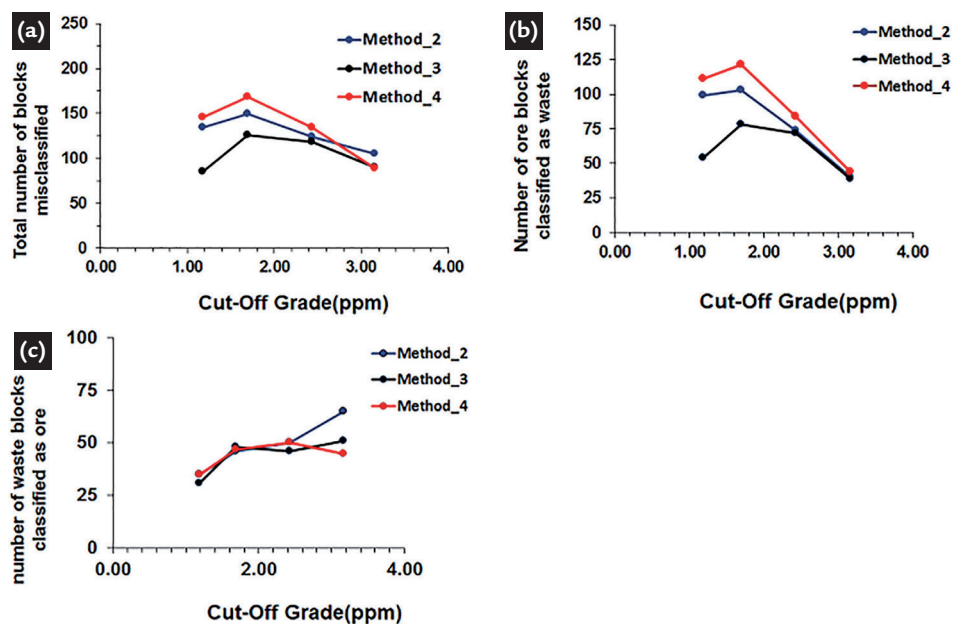


Figure 4 Block misclassification: a) total number of misclassified blocks; b) number of ore blocks classified as waste; and c) number of waste blocks classified as ore.

5. Conclusion

Geostatistical workflows with differing treatments of hard and soft data were investigated in order to integrate samples with known bias into the estimation process. Four methodologies were tested: ordinary kriging with only hard data, ordinary kriging with hard and soft data combined, standardized ordinary kriging with hard and soft data, and

standardized ordinary cokriging.

When cokriging was used, a moderate correlation between data types was noted when modeling the cross-covariance with short spatial continuity. Consequently, low weights were assigned to secondary soft data. Cokriging produced a lower coefficient of correlation with the reference model than standardized kriging

with hard and soft data.

Ordinary kriging with sparse hard data samples (a small number of accurate data points) produced global statistics that were similar to those of standardized kriging with hard and soft data. However, the first exhibited local bias as indicated by the lower coefficient of correlation with the reference model.

A correction factor was applied to biased and imprecise samples in order to mitigate global and conditional biases. The results show that the soft data can be standardized using the declustered mean and standard deviation when the samples are in the same domain. Next, the soft data is rescaled using the mean and standard deviation of the hard data.

After these corrections, two types of samples (hard and soft data) were

given equal weight in the standardized ordinary kriging workflow. The results produced by this methodology showed that the bias presented in the samples was not reproduced in the estimates.

This case study shows that soft data may improve short-term geological modeling when an appropriate method is used to integrate it. The best option is Method 3- standardized ordinary kriging with hard and

soft data, assuming that hard and soft data are moderately correlated and exhibit short-range cross variograms. The scatter-plot of estimates versus reference values indicates one of the highest available coefficients of correlation, a slope of regression closer to one, an absolute error close to zero, and better block-classifying efficiency than other methods for most of the cut-off grades tested in this study.

Acknowledgments

The authors would like to thank the Brazilian National Research Agency

(CNPq), Vale S.A., and Petrobras for supporting the Mineral Exploration and Min-

ing Planning Research Unit (LPM) at the Federal University of Rio Grande do Sul.

References

- COSTA, J.F. *Developments in recoverable reserves estimation and ore body modelling*. Queensland, Australia: University of Queensland, 1997. 666p. (PhD Thesis.)
- GOOVAERTS, P. *Geostatistics for natural resources evaluation*. New York: Oxford University Press, 1997. 483p.
- JOURNEL, A. G. Geostatistics for conditional simulation of ore bodies. *Economic Geology. Mining Geostatistics*, v. 69, n. 5, p. 673-687, 1974.
- MATHERON, G. Principles of geostatistics. *Economic Geology*, v. 58, p. 11246-11266, 1963.
- MINNITT, R. C. A., DEUTSCH, C. V. Cokriging for optimal mineral resource. *Journal of the Southern African Institute of Mining and Metallurgy*. v. 114, n. 3, p. 189-203, 2014.

Received: 12 February 2016 - Accepted: 23 October 2017.