Natural Resources Research, Vol. 9, No. 3, 2000

Comparison Between Kriging Variance and Interpolation Variance as Uncertainty Measurements in the Capanema Iron Mine, State of Minas Gerais—Brazil

Marcelo Monteiro da Rocha[1] and Jorge Kazuo Yamamoto[2]

Received 15 February 2000; *accepted 6 March 2000*

The Capanema Mine, an iron ore deposit, is located in the central portion of the Quadrilátero Ferrífero, State of Minas Gerais, southeastern Brazil. Mine development data from approximately 7000 drillholes were used for a comparative study between kriging variance and interpolation variance as uncertainty measurements associated with ordinary kriging estimates. As known, the traditional kriging variance does not depend on local data and, therefore, does not measure the actual dispersion of data. On the other hand, the interpolation variance measures adequately the local dispersion of data used for an ordinary kriging estimate. This paper presents an application of the concept of interpolation variance for measuring uncertainties associated with ordinary kriging estimates of Fe and silica grades. These data were selected for their distinct statistical characteristics with Fe presenting a negatively skewed distribution and, consequently, a low dispersion, and silica a positively skewed distribution and, therefore, a high variability. Comparative studies between the two uncertainty measurements associated with ordinary kriging estimates of Fe and silica proved the superiority of the interpolation variance as a reliable and precise alternative to the kriging variance.

KEY WORDS: Kriging variance; interpolation variance; ordinary kriging; uncertainty measurements.

INTRODUCTION

Uncertainty has been addressed as a key issue in geostatistics, justifying the development of simulation techniques. The interesting feature of stochastic simulation is that from a set of realizations on a grid node we can derive the mean and the associated variance. Indeed, these statistics are derived from a conditional cumulative distribution function built from a set of stochastic simulation realizations. The number of realizations should be large enough to guarantee that the resulting statistics will be significant as stochastic simulation takes time, even considering available fast processors. Depending on the stochastic simulation algorithm and also on the data set, a typical run can take some hours. On the other hand, there are estimation techniques provided by geostatistics, in which the ordinary kriging has been recognized as a standard one because of its simplicity and easy operation. The great problem of ordinary kriging estimation is that the variance associated with the resulting estimate has no significance, because it is only variogram dependent and not data-value dependent. Journel and Rossi (1989) recognized that the kriging variance does not measure uncertainty, but just the spatial configuration of neighbor data used to make the estimate. Recently Yamamoto (2000) proposed an alternative to the kriging variance, named interpolation variance. As long as the interpolation variance measures local data dispersion

[1] Institute of Geosciences, University of São Paulo, Rua do Lago, 562, CEP 05508-900. São Paulo, SP. Brazil. (e-mail: mmrocha@usp.br)

[2] Laboratory of Applied Geostatistics—LAG, Department of Environmental and Sedimentary Geology, Institute of Geosciences, University of São Paulo, Rua do Lago, 562, CEP 05508-900, São Paulo, SP. Brazil.

it is a reliable measure of uncertainty. In this sense, Yamamoto (1999) proposed to use the interpolation variance for ore-reserve classification. This paper presents a comparative study between kriging and interpolation variance using iron and silica values taken from exploration and development data sets. This study was carried out in the Capanema Iron Mine situated in the Quadrilátero Ferrífero region, a traditional iron ore-producing region in State of Minas Gerais, southeastern Brazil.

LOCATION OF THE CAPANEMA MINE

The Capanema Mine is located in the central portion of State of Minas Gerais. From Belo Horizonte (state capital), it is about 100 km and is reached by interstate BR-356 to Ouro Preto. Figure 1 shows the location map of the Capanema Mine, which belongs to Minas da Serra Geral S.A., and is a joint venture between Vale do Rio Doce and a Japanese consortium headed by Kawazaki Steel.

GEOLOGY AND ORE TYPES

According to Massahud and Viveiros (1983), the Capanema Mine is located in a synclinal structure known as the Ouro Fino Syncline. The stratigraphic section for this area is reproduced in Figure 2 (after Massahud and Viveiros, 1983).

The iron ore of the Capanema Mine is contained in the Cauê Formation located in the core of the Ouro Fino Syncline. The Cauê Formation includes (from bottom to top): basal siliceous itabirite, calcareous itabirite, and siliceous and amphibolitic itabirite. Some intrusive bodies (dikes and sills) of metamorphosed basic rocks cut the orebody. They are altered completely to argillaceous material of varied colors (white, yellow, and locally reddish).

The most accepted hypothesis for the ore genesis of the Capanema Mine is that the relative enrichment of iron resulted from leaching of carbonates by meteoric solutions flowing downward to the watertable. The mineral paragenesis reflects a low-grade metamorphism.

The iron ore exploited in the Capanema Mine is classified into three types, following chemical and grain-size criteria. This classification is based on Fe, SiO_2, Al_2O_3, and P grades as well as visual characteristics such as color, texture, and compaction of the rock:

- hematite A (HA): laminated, banded, friable to pulverulent, soft with rare bands of siliceous beds;
- hematite B (HB): laminated and banded as the previous type but with hard siliceous beds;
- soft itabirite (SI): composed by alternate beds of unconsolidated silica and hematite, probably resulting from a relative enrichment of silica.

Following the same criteria for ore types, the waste is classified into five types:

- weathered hematite (WH): hard, banded and laminated, with some argillaceous and limonitic parts;
- soft itabirite–Pulverulent (SIP): such as SI presents alternate beds of unconsolidated silica and hematite, but more pulverulent;
- soft itabirite–amphibolitic (SIA): composed of alternate beds of silica and hematite with patches of altered amphibole;
- intrusive (IN): intrusive rocks in the form of dikes and sills;
- phyllite (PHI): metamorphic rock occurring below the iron formation.

Apparent density of Capanema's iron ore (Table 1) is relatively low if compared to other iron ores.

Table 2 presents average compositions of ore and waste types as obtained from diamond drill cores during the exploration program.

Figure 3 is a three-dimensional model as viewed from NW for the Capanema Mine according to the ore and waste types. The orebody is elongated in a SE–NW direction as result of a first, brittle-ductile, deformational event, which generate the syncline and the Fundão fault system. The Ouro Fino Syncline then was refolded in its upper portion (NW) by a second deformation event. A third event occurred but it is not represented in the three–dimensional model because it is related only to minor structures. The three-dimensional model shows clearly the iron ore (the brownish unit) in the core of Ouro Fino Syncline.

SAMPLING

When the iron body of the Capanema Iron Mine was discovered in the mid 1960s, development drifts and cross cuts were used to explore it. Then, near the end of 1980s the first diamond drillhole campaign was carried out which resulted in 11 holes. In a reevaluation 59 additional holes were drilled, for a total of 70. Figure 4 shows the location of all the exploration diamond drillholes.

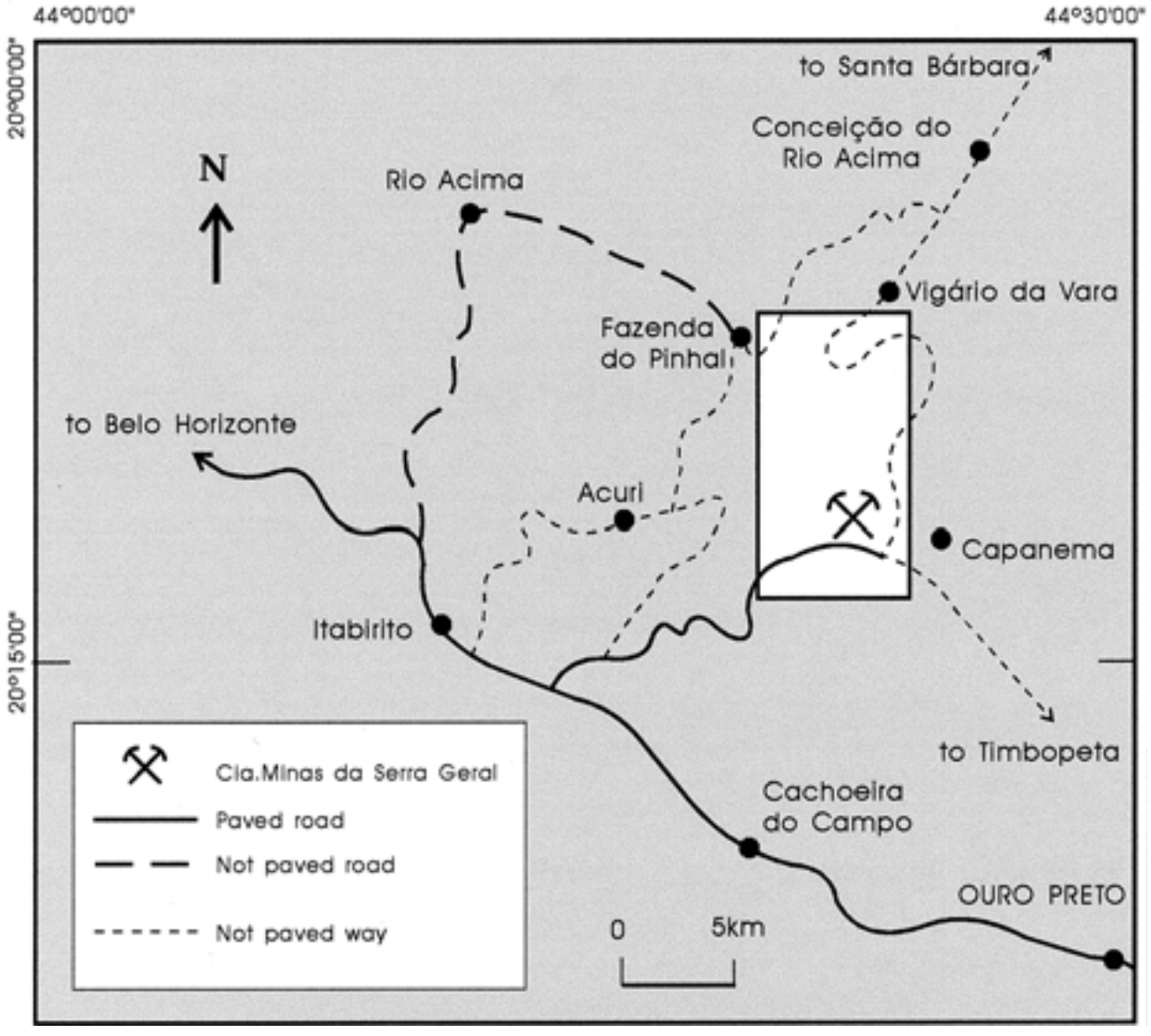

Figure 1. Location map of Capanema Mine.

In this mine, the ore-grade control is made using the dust of rock drillholes that has been collected systematically and then analyzed. For each block of a bench, a rock drillhole is made in its center. As the blasting block is equal to 28.8 × 28.8 × 13 m, rock drillholes are arranged in a regular grid spaced 28.8 m. Figure 4 also presents the location of all rockdrill holes, totaling 6988 holes.

As seen in Figure 4, the diamond drillholes sampled the eastern part of the Capanema orebody. On the other hand, the rock drillholes have sampled the entire orebody because they were drilled as the mine developed. Thus, rock drillholes reproduce better the orebody in relation to its characteristics such as composition, grain size, etc. Moreover, the high sampling density provides precise statistics of the orebody, which authorizes a comparative study between uncertainty measurements associated with ordinary kriging estimates.

STATISTICAL AND GEOSTATISTICAL ANALYSES

Using the two available databases, a statistical analysis has been done in order to characterize statis-

Super group	Group	Formation	Main lithologies
Minas	Itabira	Cauê	itabirites, amphibolitic itabirites and amphibolitic iron formation
	Caraça	Batatal	sericitic phyllites
		Moeda	quartzite and kyanite quartzite
Rio das Velhas	Maquiné	undivided	chlorite schist with interbedded quartzite

Figure 2. Stratigraphic section for Capanema Mine region (after Massahud and Viveiros, 1983).

tical properties of variables. For this study, just two variables were selected from databases, that is Fe and SiO_2. Because this paper compares uncertainty measurements derived from ordinary kriging estimates, we selected these two variables for reasons to be explained later.

Figure 5 presents histograms and their statistics for Fe and SiO_2. Fe and silica illustrate two different behaviors, with the former presenting a negative asymmetry and the silica a positive one. Consequently, the coefficient of variation is low for Fe (0.086) and higher for silica (1.013). Fe presents a negative asymmetrical distribution because it has an upper limit equal to 69.9%, which is the stoichiometric limit of Fe in Fe_2O_3 (hematite).

Variogram models fitted to experimental data revealed the models presented in Figure 6.

ORDINARY KRIGING ESTIMATION AND ASSOCIATED ERRORS

Ordinary kriging (OK) has been used extensively as an estimation technique because of its simplicity and for its reliable estimates. This technique allows estimation of an unsampled location based on neighbor data values:

$$Z^*(x_o) = \sum_{i=1}^{n} \lambda_i Z(x_i) \tag{1}$$

where $\{\lambda_i, i = 1, n\}$ are the ordinary kriging weights and $\{Z(x_i), i = 1, n\}$ are the n neighbor data close to the unsampled location (x_o).

The kriging variance associated to an OK estimate is:

$$\sigma^2_{OK} = \sum_{i=1}^{n} \lambda_i \gamma(x_i - x_o) - \mu \tag{2}$$

where $\gamma(x_i - x_o)$ is the variogram value for a distance between the *ith* sample and the location x_o to be estimated and μ is the Lagrange multiplier resulting from solution of the ordinary kriging system. Details for deriving ordinary kriging system are given in textbooks such as Journel and Huijbregts (1978), Isaacks and Srivastava (1989), Wackernagel (1995), Goovaerts (1997), and, more recently, Olea (1999).

Table 1. Apparent Density for Ore and Waste Types of Capanema Mine

Ore and waste types	Apparent density (t/m^3)
HA	2.89
HB	2.57
SI	2.29
WH	2.76

Table 2. Average Composition of Ore and Waste Types of Capanema Mine

Type	Fe (%)	Al_2O_3 (%)	SiO_2 (%)	P (%)	LOI[a] (%)
WH	61.61	2.85	1.20	0.091	7.39
SIP	54.01	0.59	20.3	0.025	1.68
SIA	62.68	0.90	3.03	0.082	6.02
SI	57.49	0.38	15.12	0.042	1.97
HA	64.40	1.45	2.25	0.060	3.88
HB	61.57	0.87	7.47	0.072	3.12

[a] LOI, loss on ignition.

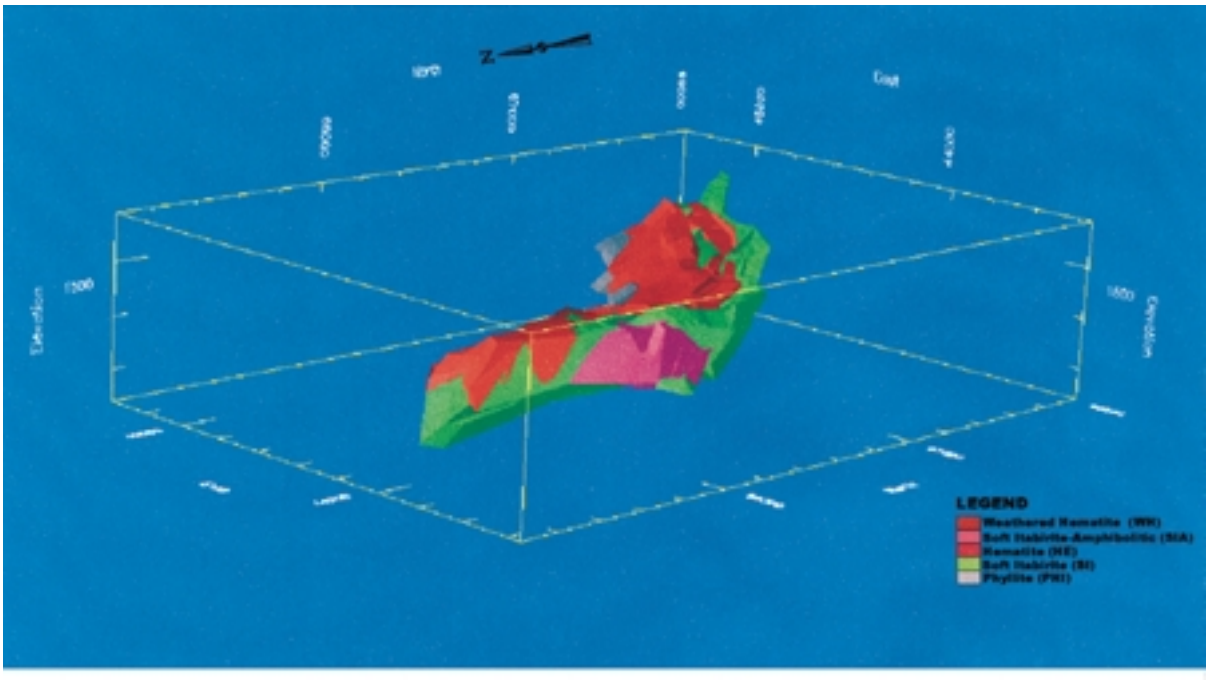

Figure 3. Three-dimensional model as viewed from NW to SE for Capanema Mine according to ore and waste types.

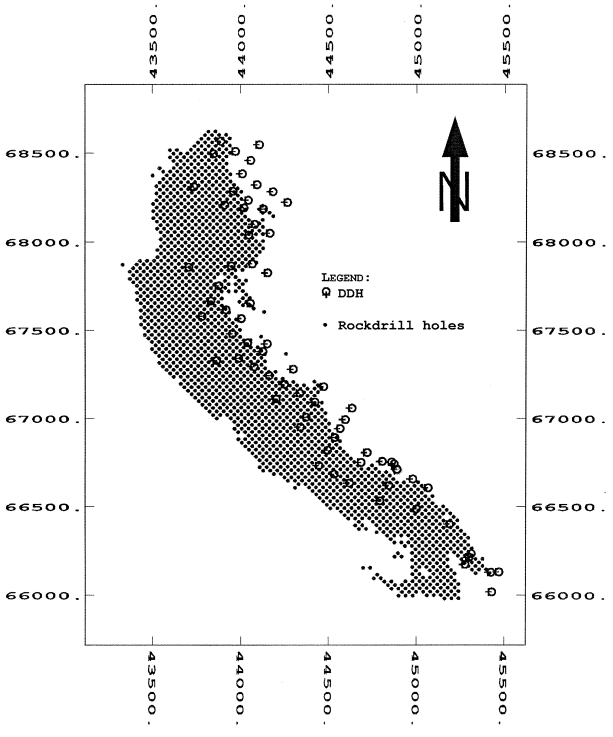

Figure 4. Location map of diamond drillholes and rock drillholes.

Ordinary kriging was the first estimation technique that provided computation of an estimation error by the kriging or estimation variance. The kriging variance was used extensively as an uncertainty measure associated with OK estimate until the end of the 1980s. In fact, Journel and Rossi (1989) stated that the kriging variance does not measure uncertainty, but just the spatial configuration of neighbor data used to make the OK estimate. Yamamoto (2000) proposed an alternative measurement of uncertainty to compute an estimation variance from ordinary kriging weights. This proposal is based on the interpolation variance that is simply the weighted average of the squared differences between data values and the OK estimate.

$$S_o^2 = \sum_{i=1}^{n} \lambda_i [Z(x_i) - Z^*(x_o)]^2 \qquad (3)$$

where S_o^2 is the interpolation variance and λ_i are the ordinary kriging weights.

Positiveness of the interpolation variance depends on all ordinary kriging weights being positive. If some negative weights occur, then the interpolation variance can be negative, which is unacceptable, as well as a negative grade, which is also unacceptable. Users of ordinary kriging estimation techniques are used to seeing negative grades, especially if they are

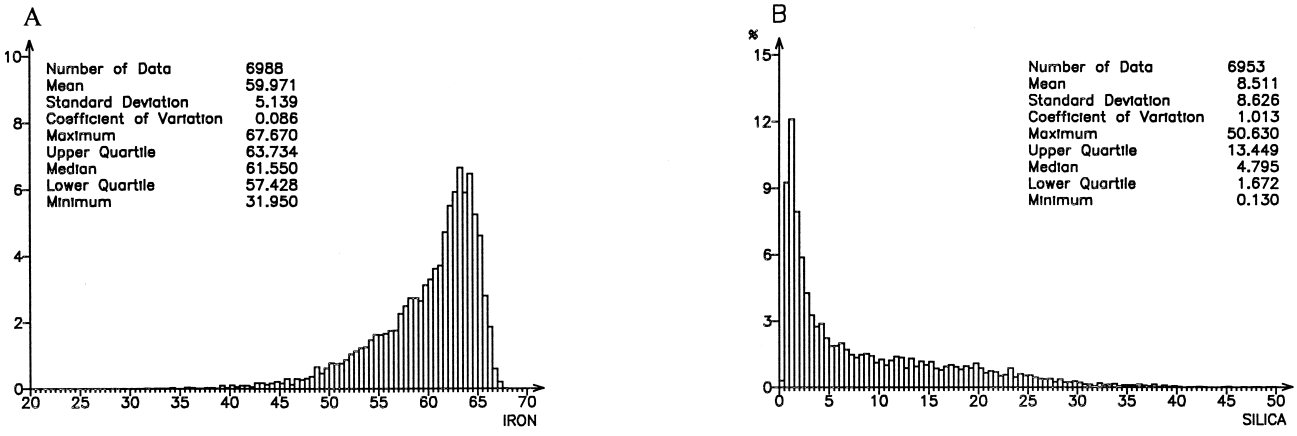

Figure 5. Histogram and statistics for Fe(A) and for silica (B).

working with a data set with some clustered points and a highly skewed distribution. Thus, in order to avoid negative estimates some algorithms for correcting negative weights have been proposed. Examples of correction algorithms for negative weights are given in Froidevaux (1993), Journel and Rao (1996), and Deutsch (1996). We have adopted the procedure proposed by Journel and Rao (1996) that consists on adding a value to all weights equal to the modulus of the largest negative weight and rescaling them to sum up to one.

The ordinary kriging allows estimation of an unsampled point where the true value is not known. On the other hand, the cross-validation procedure makes use of an estimation technique (e.g., ordinary kriging) for a sampling point, but without considering its value as neighboring data points. Thus, cross validation provides, for each point, the true and estimated values, from which actual errors can be derived. For this reason, we adopted the cross-validation procedure to make ordinary kriging estimates and to compare uncertainty measurements associated with them.

Figure 7 presents histograms and statistics of OK estimates for Fe and silica. Distributions of estimated values are more smoothed than their respective data sets, that is:

$$Var\lfloor Z^*(x_o)\rfloor < Var[Z(x)] \qquad (4)$$

As long as $Z^*(x_o)$ is a linear combination of n nearest neighbor data points $\{Z(x_i),\ i = 1,\ n\}$, $Var\lfloor Z^*(x_o)\rfloor$ will be less than $Var[Z(x)]$. Any other estimation method based on a linear combination of data points (weighted average) will result in smoothing for resulting estimates.

Figure 8 presents cross-validation scattergrams for Fe and silica. In Figure 8A note that iron values are limited to the stoichiometric limit of 69.9%. Correlation coefficients for Fe and silica are good (0.715 and 0.801, respectively) considering a total of 6884 points. Observed dispersions around minimum square regression lines are indeed the result of both the smoothing effect and the natural variability of the orebody.

Figures 9 and 10 present histograms and statistics for estimation errors associated with OK estimates, respectively, for iron and silica.

Histograms for the interpolation standard deviation present more classes than those for the kriging standard deviation. Actually, the kriging standard deviation values are distributed through three–five classes, which reflect configurations of neighboring points used to make estimates. Moreover, observe that the kriging standard deviation values are limited to the square root of the sill of fitted variogram models (Fig. 6). On the other hand, the interpolation standard deviation values do not present this limitation because they depend on retained estimate and n neighbor data values.

ANALYZING THE PROPORTIONAL EFFECT

The proportional effect is a heteroscedastic condition in which the variance of the error is proportional to some function of the local mean of the data values (Olea, 1991). In this sense, as long as the interpolation variance depends on the local data (n neighbor data values) it recognizes the proportional effect. Yamamoto (2000) analyzed a highly skewed distribu-

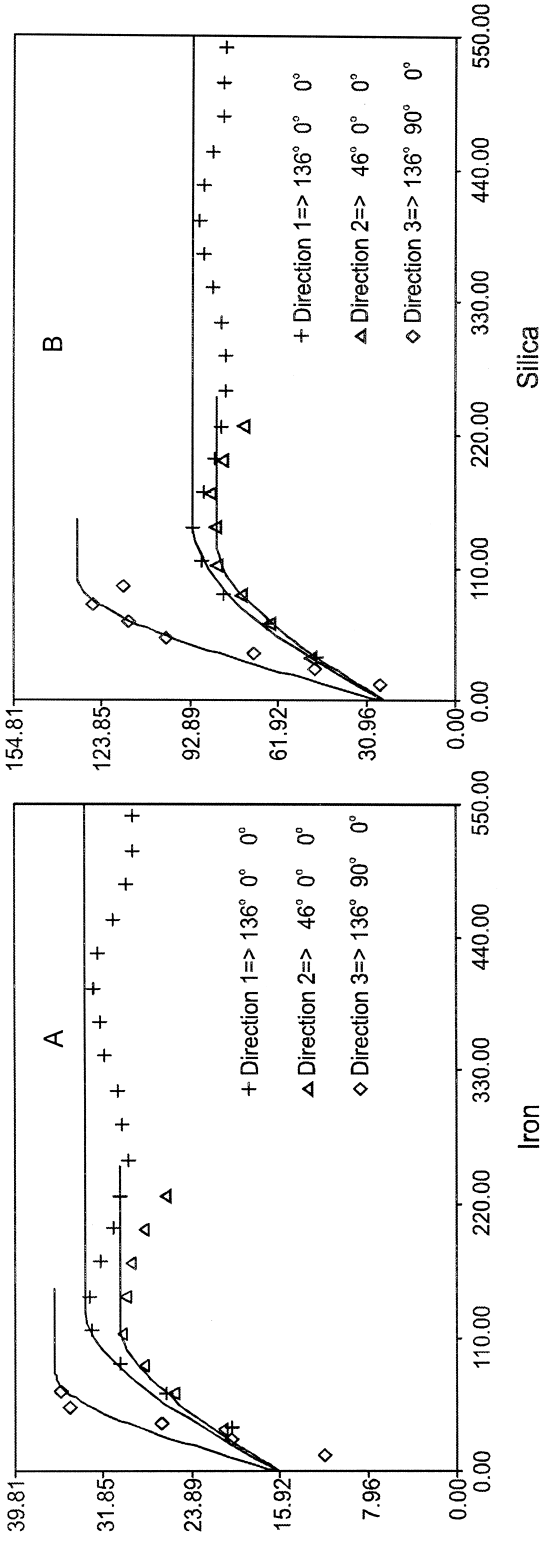

Figure 6. Variogram models computed for Fe (A) and silica (B).

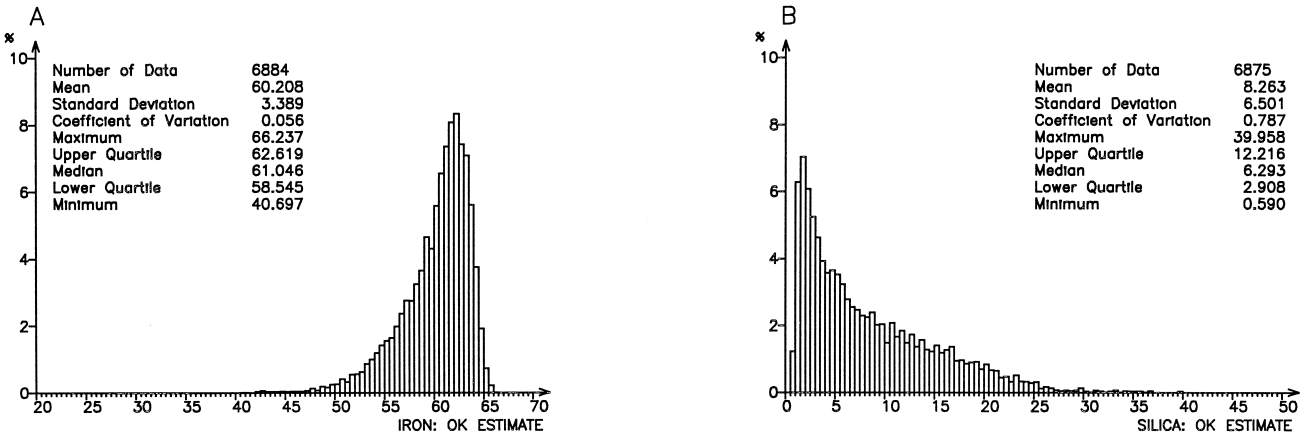

Figure 7. Histogram and statistics for OK estimates for Fe (A) and for silica (B).

tion (positive asymmetry) and concluded that the interpolation variance is proportional to the OK estimate. In this paper two distinct distribution are analyzed, both asymmetric, but with iron presenting a negative asymmetry and silica a positive one.

Thus, for analyzing the proportional effect scattergrams of the estimation errors were plotted against OK estimates as presented in Figures 11 and 12, respectively, for Fe and silica.

Figures 11 and 12 illustrate that the kriging variance does not present any correlation with the OK estimates for Fe and silica. Interpolation variance presents a good correlation with the OK estimates for both variables, but the correlation coefficient is negative for Fe and positive for silica. This is an interesting feature and shows why the interpolation variance is a reliable measurement of uncertainty. As iron presents a negatively skewed distribution, in which the upper limit is given by the stoichiometric limit, samples close to this limit, that is iron ores, will present similar iron contents indicating that estimation error should be small. On the other hand, toward the low-grade ores, the estimation error presents more dispersion because iron contents after with more freedom. It explains why the correlation coefficient for the interpolation

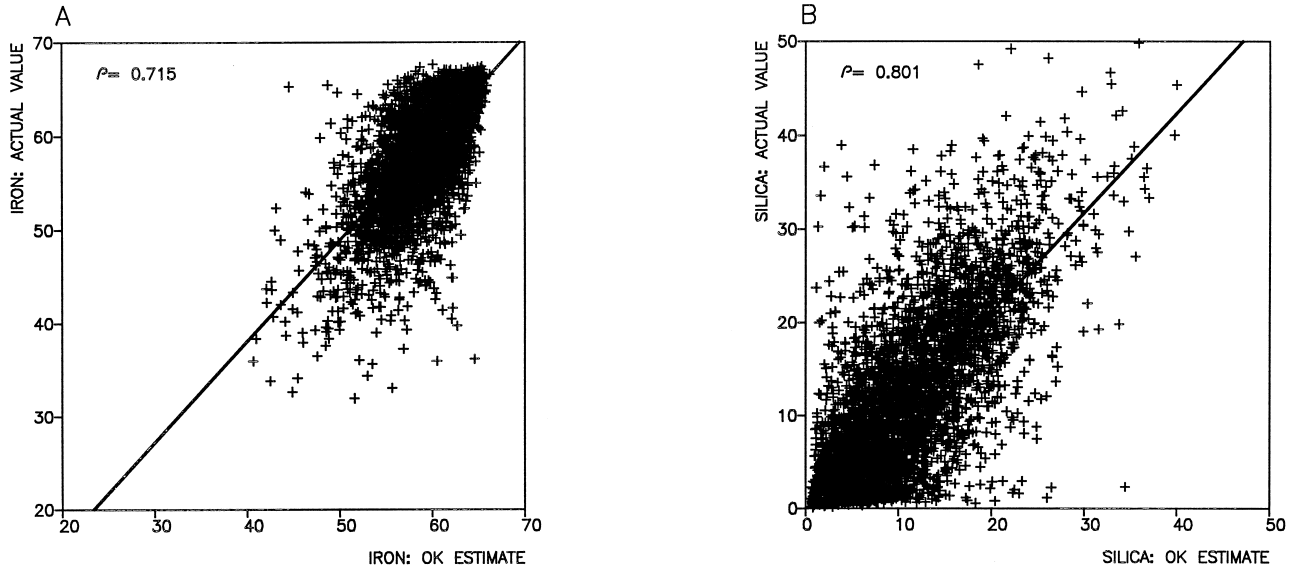

Figure 8. Cross-validation scattergrams for Fe (A) and for silica (B).

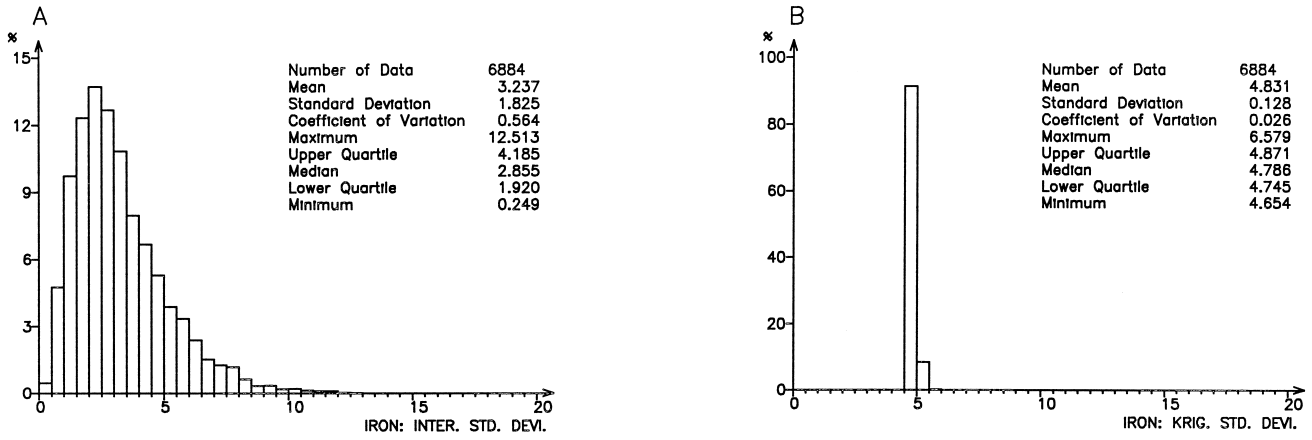

Figure 9. Estimation errors calculated for iron OK estimates: interpolation standard deviation (A) and kriging standard deviation (B).

standard deviation plotted against iron OK estimates is negative.

Contrary to the behavior of iron, the interpolation standard deviation for silica presents positive correlation coefficients as long as its histogram is positively skewed. This is the behavior expected for variables presenting the proportional effect. The kriging standard deviations do not present any significant correlation with the silica OK estimates.

COMPUTING CONFIDENCE INTERVALS FROM UNCERTAINTY MEASURES

Confidence intervals around estimates can be derived from estimation errors using the well-known Central Limit Theorem:

$$CI_{CL(\%)} = \frac{S \cdot t}{\sqrt{n}} \tag{5}$$

where S is the standard deviation, n is the number of samples, t is the score of t distribution for a given confidence level (CL), and degrees of freedom ($n - 1$).

By replacing the standard deviation by interpolation standard deviation or kriging standard deviation, two-sided confidence intervals are derived that can be used to predict if the actual values fall in a given confidence level. For this study, the test was performed using the following confidence levels: 90, 95, and 99%. Table 3 presents proportion of matches of actual iron grades within two-sided confidence intervals; Table 4 provides the same for silica.

Apparently the kriging standard deviation seems

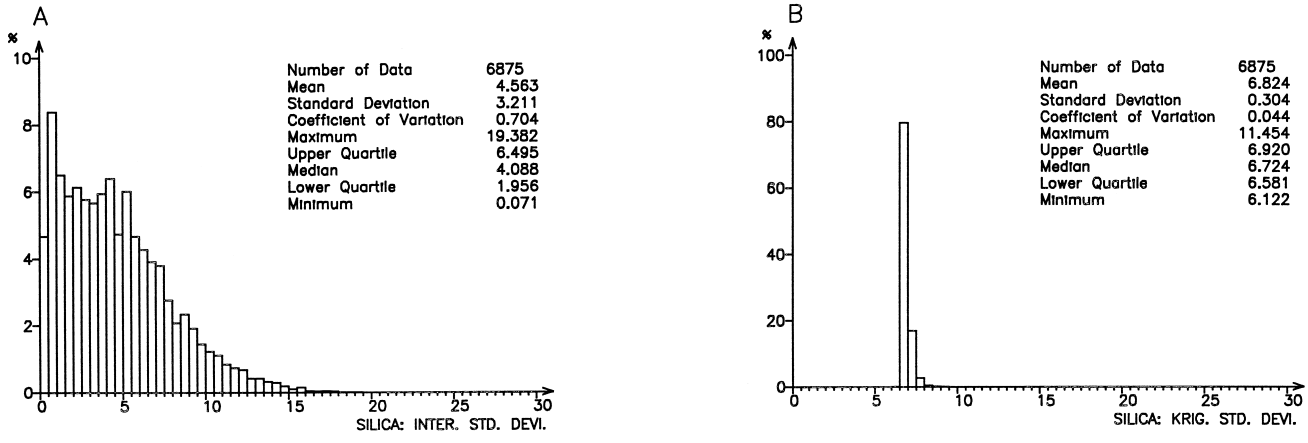

Figure 10. Estimation errors calculated for silica OK estimates: interpolation standard deviation (A) and kriging standard deviation (B).

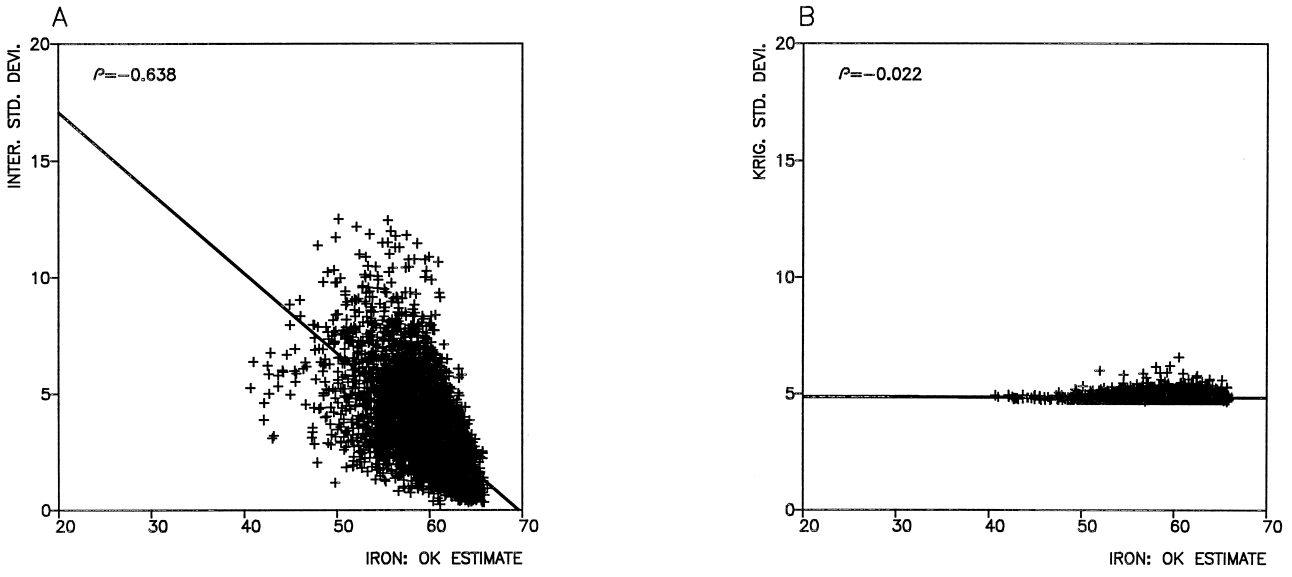

Figure 11. Estimation errors plotted against iron OK estimates according to interpolation standard deviation expression (A) and to kriging standard deviation expression (B).

superior to the interpolation standard deviation. However, this is not necessarily true because the kriging standard deviation is always greater than the interpolation standard deviation as can be seen on histograms of ratios of kriging standard deviation to interpolation standard deviation (Fig. 13).

It is clear that the standard kriging deviation is greater than the interpolation standard deviation. For iron, we observe that 82.6% of kriging standard deviation values are greater than the interpolation standard deviation ones, and for silica this proportion is equal to 77.4%. Thus, we can conclude that the

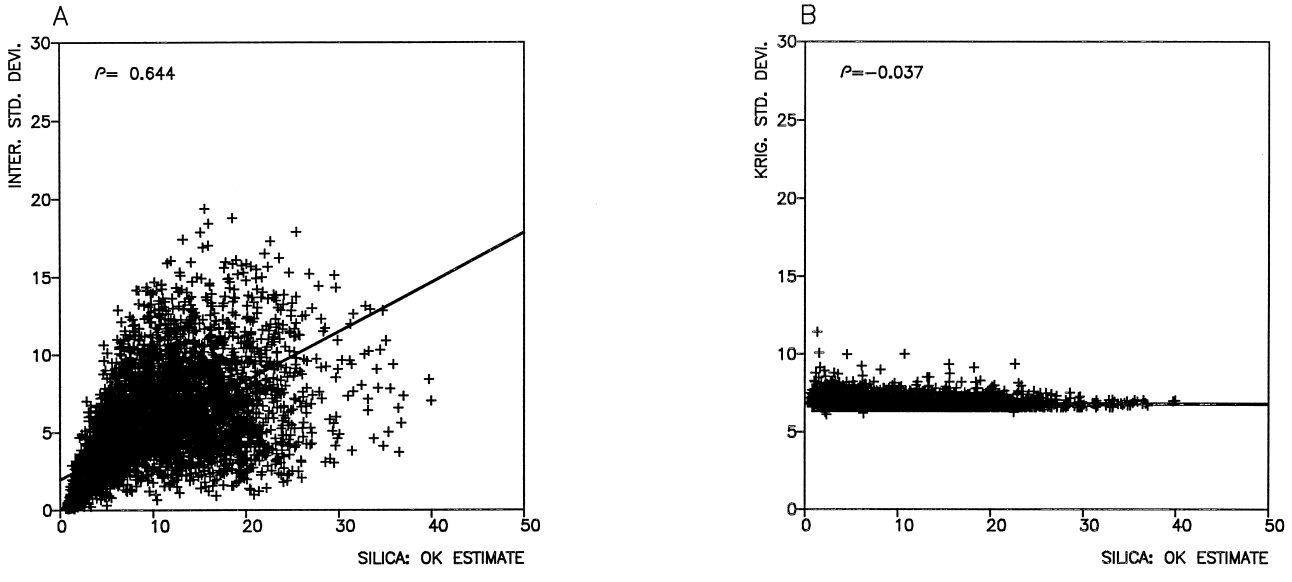

Figure 12. Estimation errors plotted against silica OK estimates according to interpolation standard deviation expression (A) and to kriging standard deviation expression (B).

Table 3. Proportion of Matches of Actual Iron Grades within Two-Sided Confidence Intervals around Estimated Value

Confidence level (%)	S_o based two-sided confidence interval	σ_{OK} based two-sided confidence interval
90	50.5	73.9
95	61.3	81.8
99	79.5	91.9

Table 4. Proportion of Matches of Actual Silica Grades within Two-Sided Confidence Intervals Around Estimated Value

Confidence level (%)	S_o based two-sided confidence interval	σ_{OK} based two-sided confidence interval
90	53.1	74.2
95	65.0	81.7
99	82.1	91.1

confidence intervals based on interpolation standard deviation can predict reasonably well (65% for a confidence level equal to 95%) the true values. Obviously, this proportion of matches will depend on variability of data under study and, consequently, on sampling density. Because we have used the cross-validation procedure, it is available for each sampling point of the estimated and true values from which we can derive the absolute true error:

$$ABS(\text{true error}) = |z^*(x_o) - z(x_o)| \qquad (6)$$

The absolute true error should be comparable to the confidence interval for a given confidence level. In other words, the confidence interval based on interpolation standard deviation or kriging standard deviation should present a linear relationship with the true absolute error. Figures 14 and 15 show scatter diagrams of the confidence intervals (confidence level of 95%) against true errors for iron and silica, respectively.

Once more, it is possible to verify that the confidence intervals, based on kriging standard deviation, do not present any correlation with the absolute true errors. However, those based on interpolation standard deviation do present a reasonable correlation with the absolute true errors illustrating that the interpolation standard deviation provides a reliable measurement of uncertainty associated to ordinary kriging estimate.

CONCLUSION

This paper illustrated the usefulness in an application of the interpolation variance as an uncertainty measure using a real database from development work at Capanema Iron Mine. Fe and silica were selected for their distinct statistical characteristics with the former presenting a negative skewed distribution and the second variable presenting a positively skewed distribution. Although the kriging standard deviation does not

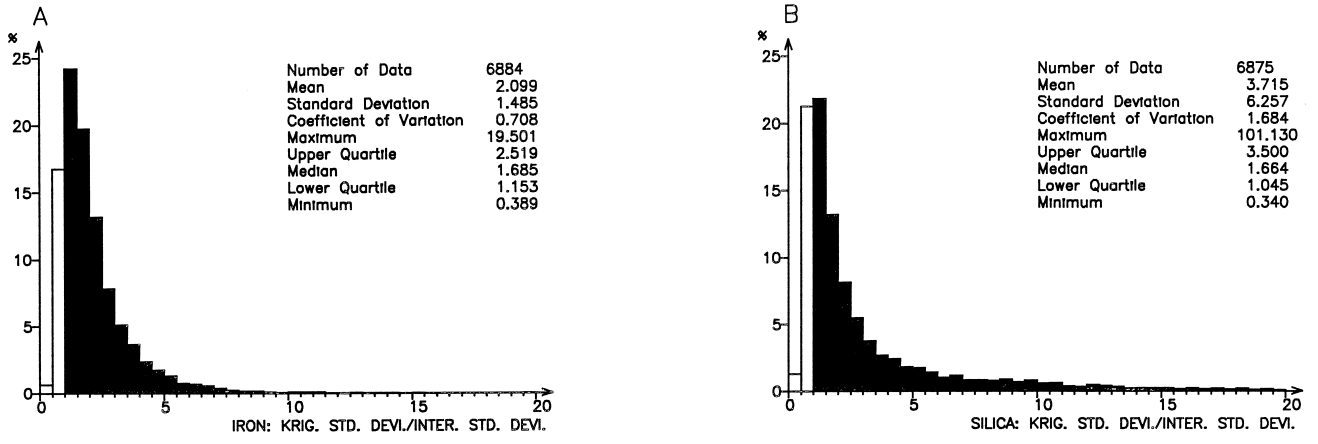

Figure 13. Distribution of ratios of kriging standard deviation to interpolation standard deviation for Fe (A) and for silica (B). Black bars represent ratios greater than 1.

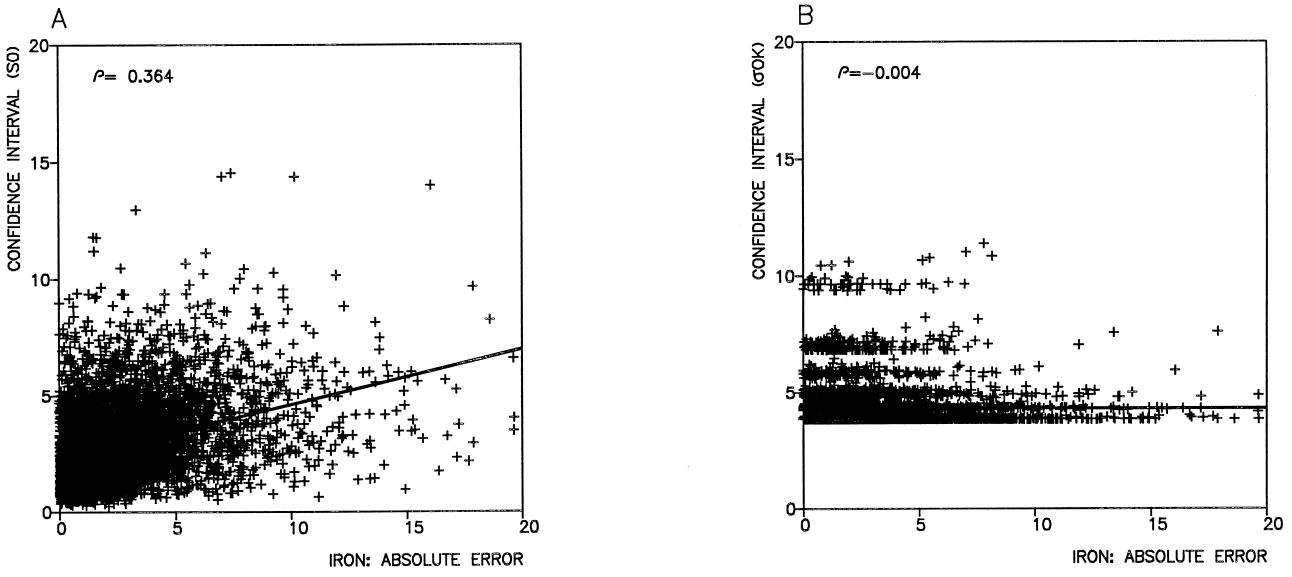

Figure 14. Scatter diagrams of confidence intervals against absolute true errors for iron. Confidence interval based on interpolation standard deviation (A) and on kriging standard deviation (B).

recognize the proportional effect existing in both distributions, the interpolation standard deviation showed clearly that there is a proportional effect for silica, that is, in a scatter diagram of the interpolation error against estimated value, there is a positive correlation. On the other hand, because Fe presents a negatively skewed distribution the linear relationship in the scatter diagram for analyzing the proportional effect also is negative. Confidence intervals based on interpolation standard deviation always are less than those based on kriging standard deviation and they can predict the true values reasonably well. Finally, all results confirm the usefulness of the interpolation variance as a precise and reliable measure of uncertainty.

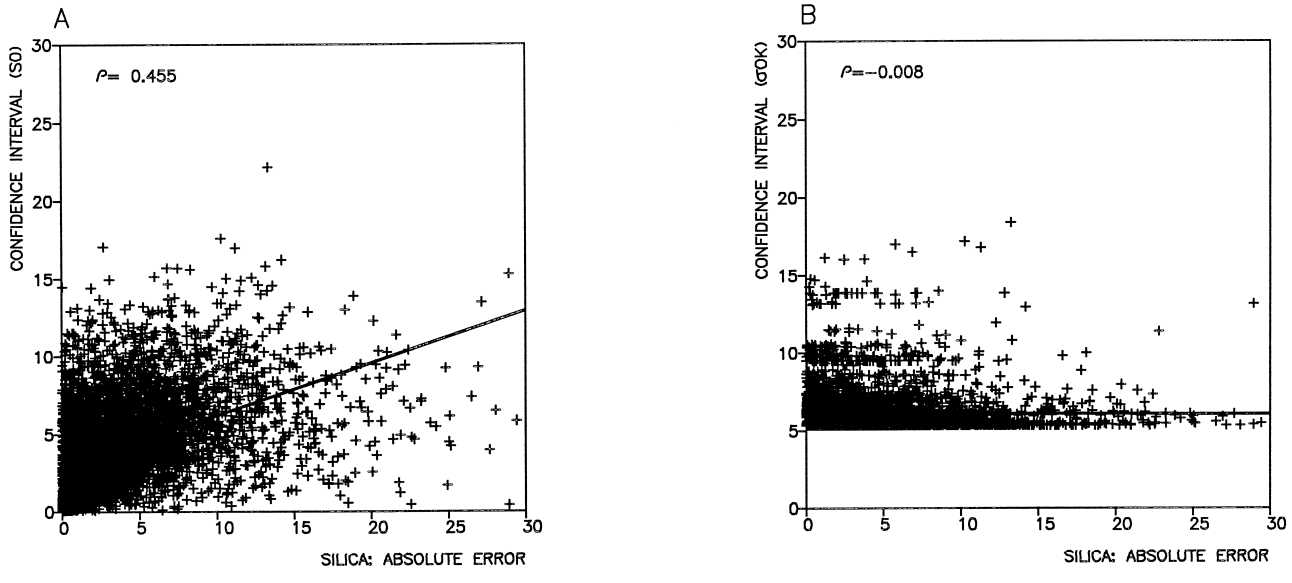

Figure 15. Scatter diagrams of confidence intervals against absolute true errors for silica. Confidence interval based on interpolation standard deviation (A) and kriging standard deviation (B).

ACKNOWLEDGMENTS

The authors wish to thank the Foundation for Research Sponsorship of State of São Paulo—FAPESP (Processes 96/7156-9 and 97/08628-4), as well as to the National Council for Scientific and Technological Development—CNPq (Process 304612189-8) for the financial support to conduct this research. The authors also wish to thank the Empresa Minas da Serra Geral S/A and its staff and especially its director, Juarez César da Fonseca (*in memoriam*).

REFERENCES

Deutsch, C. V., 1996, Correcting for negative weights in ordinary kriging: Computers & Geosciences, v. 22, no. 7, p. 765–773.

Froidevaux, R., 1993, Constrained kriging as an estimator of local distribution functions, *in* Capasso, V., Girone, G., and Posa, D., eds., Proc. Intern. Workshop on Statistics of Spatial Processes: Theory and Applications (Bari, Italy), p. 106–118.

Goovaerts, P., 1997, Geostatistics for natural resources evaluation: Applied Geostatistics Series, Oxford Univ. Press, New York, 483 p.

Isaaks, E. H. and Srivastava, R. M., 1989, An introduction to applied geostatistics: Oxford Univ. Press, New York, 561 p.

Journel, A. G., and Huijbregts, C. J., 1978, Mining geostatistics: Academic Press, London, 600 p.

Journel, A. G., and Rao, S. E., 1996, Deriving conditional distributions from ordinary kriging: Stanford Center for Reservoir Forecasting, Rept. No. 9, Stanford Univ., 25 p.

Journel, A. G., and Rossi, M. E., 1989, When do we need a trend model in kriging?: Math. Geology, v. 21, no. 7, p. 715–739.

Massahud, J. S., and Viveiros, J. F. M., 1983, Geology of Capanema iron ore deposit: a review, *in* Simpósio sobre minério de ferro da Fundação Carl Duisberg. Belo Horizonte, 453 p.

Olea, R. A., ed., 1991, Geostatistical glossary and multilingual dictionary: Oxford Univ. Press, New York, 175 p.

Olea, R. A., 1999, Geostatistics for engineers and earth scientists: Kluwer Acad. Publ., Norwell, Massachusetts, 303 p.

Wackernagel, H., 1995, Multivariate geostatistics: an introduction with applications: Springer-Verlag, Berlin, 256 p.

Yamamoto, J. K., 1999, Quantification of uncertainty in ore-reserve estimation: applications to Chapada Copper Deposit, State of Goiás, Brazil: Natural Resources Research, v. 8, no. 2, p. 153–163.

Yamamoto, J. K., 2000, An alternative measure of the reliability of ordinary kriging estimates: Math. Geology, v. 32, no. 4, p. 489–509.