Mathematical Geology, Vol. 32, No. 4, 2000

An Alternative Measure of the Reliability of Ordinary Kriging Estimates[1]

Jorge Kazuo Yamamoto[2]

This paper presents an interpolation variance as an alternative to the measure of the reliability of ordinary kriging estimates. Contrary to the traditional kriging variance, the interpolation variance is data-values dependent, variogram dependent, and a measure of local accuracy. Natural phenomena are not homogeneous; therefore, local variability as expressed through data values must be recognized for a correct assessment of uncertainty. The interpolation variance is simply the weighted average of the squared differences between data values and the retained estimate. Ordinary kriging or simple kriging variances are the expected values of interpolation variances; therefore, these traditional homoscedastic estimation variances cannot properly measure local data dispersion. More precisely, the interpolation variance is an estimate of the local conditional variance, when the ordinary kriging weights are interpreted as conditional probabilities associated to the n *neighboring data. This interpretation is valid if, and only if, all ordinary kriging weights are positive or constrained to be such. Extensive tests illustrate that the interpolation variance is a useful alternative to the traditional kriging variance.*

KEY WORDS: estimation variance, conditional estimation variance, uncertainty assessment, negative weights, ordinary kriging.

INTRODUCTION

Kriging is a generic name adopted by geostatisticians for a family of minimum-error-variance estimation techniques. Among the various forms of kriging, ordinary kriging (OK) has been used widely as a reliable estimation method. Kriging has specific features (e.g., recognition and modeling of anisotropies, declustering, and unbiasedness) that are not all available with other interpolation methods (polygonal, triangulation, inverse of weighted distance, etc). Another feature of the ordinary kriging technique is the computation of the minimized estimation variance also known as the kriging variance. The kriging variance is used as a quality indicator of the estimator. For instance, when mineral reserves are being

[1]Received 18 February 1998; accepted 14 May 1999.

[2]Department of Environmental and Sedimentary Geology, Institute of Geosciences, University of São Paulo, Caixa Postal 11.348, CEP 05.422-970, São Paulo, SP, Brazil. e-mail: jkyamamo@usp.br

0882-8121/00/0500-0489$18.00/1

estimated, kriging variance is used to assign confidence levels for classification purposes (Diehl and David, 1982; Froidevaux, 1982; Wober and Morgan, 1993).

Because kriging variance is based on a global variogram averaged over the whole estimation domain, it cannot properly measure local data dispersion. It is important to recognize that the nature of the phenomenon under study may vary from one locality to the next (Isaaks and Srivastava, 1989). For instance, if the same data configuration is used, the kriging variance yields the same value, no matter what the local data values are. This is because ordinary kriging weights and variance are data-values independent.

This paper presents an alternative to the measure of the reliability of ordinary kriging estimates, which accounts for both the data configuration and the data values.

KRIGING VARIANCE

An unknown value associated with a point or block can be estimated by ordinary kriging as follows:

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

where $\{\lambda_i, i = 1, n\}$ are the kriging weights computed from a normal system of equations derived by minimization of the error variance.

The error variance is

$$\sigma_{\mathrm{OK}}^2 = \mathrm{Var}\{Z(x_o) - Z^*(x_o)\}$$

which becomes

$$\begin{aligned} \sigma_{\mathrm{OK}}^2 = \mathrm{Var}\{Z(x_o)\} - 2\sum_i \lambda_i \,\mathrm{Cov}\{Z(x_i), Z(x_o)\} \\ + \sum_i \sum_j \lambda_i \lambda_j \,\mathrm{Cov}\{Z(x_i), Z(x_j)\} \end{aligned} \tag{1}$$

Minimization of the error variance constrained by the unbiasedness condition ($\sum_i \lambda_i = 1$) results in the following ordinary kriging system:

$$\begin{cases} \sum_j \lambda_j \,\mathrm{Cov}\{Z(x_i), Z(x_j)\} + \mu = \mathrm{Cov}\{Z(x_i), Z(x_o)\} \quad \text{for } i = 1, n \\ \sum_j \lambda_j = 1 \end{cases} \tag{2}$$

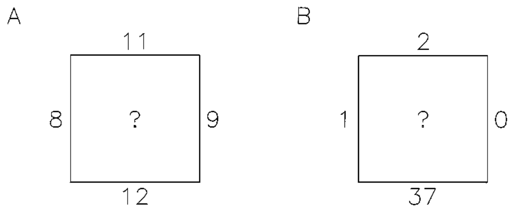

Figure 1. Estimation of blocks from the same data configuration (after Armstrong, 1994). See that in A the estimation variance should be less than in B because of more consistent data values.

The minimized error variance also called kriging variance is then computed as

$$\sigma_{\mathrm{OK}}^2 = C(0) - \sum_i \lambda_i \operatorname{Cov}\{Z(x_i), Z(x_o)\} - \mu \tag{3}$$

Kriging variance is homoscedastic, i.e., it is independent of the data values used to obtain the estimator $Z^*(x_o)$ (Olea, 1991). Figure 1 illustrates why kriging variance is not a full measure of uncertainty.

Because the data layouts are identical in both cases, the kriging variances are identical and if there is no anisotropy so will be the kriged estimates (Armstrong, 1994). The proposed interpolation variance makes use of both the set of weights $\{\lambda_i, i = 1, n\}$ and data values $z(x_i)$, measuring the higher variability prevailing around block B.

According to the "Theorem of combining kriged estimates" (Journel and Huijbregts, 1978), the kriging variance [Eq. (3)] is valid both for point and block estimates.

INTERPOLATION VARIANCE

We propose the interpolation variance defined as follows:

$$s_o^2 = \sum_{i=1}^{n} \lambda_i [z(x_i) - z^*(x_o)]^2 \tag{4}$$

where the λ_i's are the ordinary kriging weights.

This expression is data-value dependent; it is the weighted average of the experimental squared differences between data values and the kriging estimate. Note that this definition requires all weights be positive, since any negative weight can lead to a negative interpolation variance. This variance has the following properties:

- It ensures exactitude: indeed, if a datum location coincides with the point to be estimated, then that datum weight is equal to one with all other weights equal to zero, therefore $s_o^2 = 0$.
- It increases with the dispersion of the retained data values.
- It indirectly uses the variogram structural distance through the ordinary kriging weight λ_i. The more influential datum location receives the greatest weight.

This expression was introduced by Yamamoto (1989) to define an interpolation variance associated with estimated grades using multiquadric equations in mineral deposits. Yamamoto (1991) extended the definition to compute an interpolation variance using the ordinary kriging weights.

Equation (4) is in fact a direct adaptation of the variance computation commonly used in statistics (e.g., Mood, Graybill, and Boes, 1974, p. 67), where the probability density function at x_i is here replaced by the ordinary kriging weight λ_i. Provided that the ordinary kriging weights are all positive and sum up to one, this adaptation seems reasonable. Indeed, kriging weights $\{\lambda_i, i = 1, n\}$ might be interpreted as conditional probabilities attached to the n local data (Journel and Rao, 1996). At any given location x_o sort the n neighboring data in increasing order:

$$z(x_1) \leq z(x_2) \leq \cdots \leq z(x_n)$$

The local conditional cumulative distribution function (ccdf) is then modeled as

$$F(x_o, z_n) = \sum_{i=1}^{n} \lambda_i$$

The variance of this OK-derived ccdf identifies expression (4), it is data-values dependent (heteroscedastic conditional variance) as opposed to the traditional ordinary kriging variance, and as such, it is a better measure of local estimation accuracy (Journel and Rao, 1996). Froidevaux (1993) did also propose calculation of a variance conditional to the n neighboring data as

$$\sigma_o^2 = \sum_{i=1}^{n} \lambda_i \cdot z(x_i)^2 - \left(\sum_{i=1}^{n} \lambda_i \cdot z(x_i) \right)^2$$

which is exactly the same as provided by expression (4).

When $Z(x_o)$ and $\{Z(x_i), i = 1, n\}$ are random variables, the interpolation variance s_o^2 appears as an outcome of the random variables (RV) S_o^2.

Now let us look at some properties of the RV S_o^2 (see Appendix A):

$$S_o^2 = \sum_i \lambda_i [Z(x_i) - m]^2 - [Z^*(x_o) - m]^2 \tag{5}$$

where $m = E[Z(x)]$ is the mathematical expectation of RV $\{Z(x_i), i = 1, n\}$.

Taking the expected value of the RV S_o^2 we have

$$\begin{aligned} E\left\{S_o^2\right\} &= \sum_i \lambda_i \operatorname{Var}\{Z(x_i)\} - \operatorname{Var}\{Z^*(x_o)\} \\ &= \operatorname{Var}\{Z(x)\} - \operatorname{Var}\{Z^*(x_o)\} \geq 0 \end{aligned} \tag{6}$$

which appears as the variance smoothing of the kriging estimator $Z^*(x_o)$. The greater the interpolation variance S_o^2 taken in average over all possible data values for a given data configuration, the greater the variance smoothing of the kriging estimator.

Expression (4) also applies for determining the interpolation variance associated with an OK block estimate (see proof in Appendix B) as follows:

$$S_v^2 = \frac{1}{nV} \sum_l S_l^2 + \frac{1}{nV} \sum_l \left[Z_v^* - Z^*\left(x^{(l)}\right)\right]^2 \tag{7}$$

where S_l^2 is the interpolation variance for the point $x^{(l)}$, $Z^*(x^{(l)})$ is the OK estimate for point $x^{(l)}$, provided that the block V is discretized into nV points and all point kriging estimates use the same n data configuration.

Therefore, the interpolation variance for a block kriging estimate is equal to the average of interpolation variances computed in the nV discretizing points plus a variance between blocks, similar to the Krige's additivity relationship.

CORRECTING FOR NEGATIVE WEIGHTS

Negative weights must be avoided because they can cause both negative interpolation variance [Eq. (4)] and negative estimates. According to Herzfeld (1989), the geologic setting of many problems requires an estimator involving only non-negative weights. For instance, in resource exploration, geologic layers should always have non-negative thicknesses, and hanging and foot walls must not cross each other (Herzfeld, 1989). Moreover, deriving local conditional cumulative distribution functions from ordinary kriging weights is possible only when all weights are positive.

Among the various solutions to the problem of negative kriging weights, consider the direct correction. After solving the ordinary kriging system, negative weights are corrected according to specific algorithms. In this paper, the three algorithms proposed by Froidevaux (1993), Journel and Rao (1996), and Deutsch (1996) will be considered. These algorithms will hereafter be called as Froidevaux's, Journel's, and Deutsch's, respectively.

Froidevaux's correction resets all negative weights to zero; hence, this procedure simply removes the data corresponding to negative weights.

If $\lambda_i < 0 \Rightarrow \lambda_i = 0$, then the new weights are

$$\tau_i = \frac{\lambda_i}{\sum_{j=1}^{n} \lambda_j}$$

Applying Froidevaux's correction, only the remaining samples (with rescaled positive weights) are used for the estimation, thus some information is lost.

The correction proposed by Journel and Rao (1996) only removes the sample with the largest negative weight. After solving the kriging system, a positive constant equal to the modulus of the largest negative weight (if any) is added to all weights:

$$c = -\min(\lambda_i, i = 1, n) \quad \text{if } \min(\lambda_i) \leq 0, c = 0 \text{ otherwise}$$

$$\tau_i = \frac{\lambda_i + c}{\sum_{j=1}^{n} \lambda_j + c}$$

Deutsch (1996) has proposed an approach that zeroes negative weights. Moreover, this procedure also removes samples that have (1) a weight less than the average absolute magnitude of negative weights; (2) a covariance between the location being estimated and the location of the sample that is less than the average covariance between the location being estimated and the locations receiving negative weights.

In the subset of negative weights, the average absolute magnitude of negative weights is

$$\overline{\lambda} = \frac{1}{n'} \sum_{j=1}^{n'} |\lambda_j|$$

and the average covariance between the location being estimated (u) and the locations receiving negative weights is

$$\overline{C} = \frac{1}{n'} \sum_{j=1}^{n'} C(u - u_j)$$

Deutsch's procedure for correction of negative weights is as follows:

$$\text{if } \lambda_i < 0 \Rightarrow \lambda_i = 0$$
$$\text{if } \lambda_i > 0 \quad \text{and} \quad C(u - u_i) < \overline{C} \quad \text{and} \quad \lambda_i < \overline{\lambda} \Rightarrow \lambda_i = 0$$
$$\tau_i = \frac{\lambda_i}{\sum_{j=1}^{n} \lambda_j}$$

After correction of negative weights, the corrected OK estimate is

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

Once we have guaranteed that all weights from ordinary kriging are positive, the interpolation variance [Eq. (4)] is necessarily positive.

TESTING INTERPOLATION VARIANCE

Several tests have been carried out in order to demonstrate the usefulness of the interpolation variance as a measure of uncertainty associated with the estimate. In order to apply the tests, the public domain GSLIB *cluster.dat* was used (Deutsch and Journel, 1992). This data set with 140 values was derived from an exhaustive data set named *true.dat*, with 2500 points distributed on a regular grid of 50×50 nodes. According to Deutsch and Journel (1992), cluster.dat results from a two-step sampling, in which the first step 97 samples were taken on a pseudo-regular grid and in the second step an additional 43 samples were taken preferentially in high-valued regions (as identified from the first 97 samples).

The GSLIB cluster.dat

This data set has clustered data points (Fig. 2) and has a highly skewed distribution (Fig. 3), which are interesting features when computing the interpolation variance.

The variogram model for the exhaustive data set (true.dat) was considered valid for the cluster.dat. The following isotropic model, given by Deutsch and Journel (1992), will be used hereafter:

$$\begin{aligned} \gamma(h) &= 10 + 16\left[1.5\frac{h}{8} - 0.5\left(\frac{h}{8}\right)^3\right] && \text{if } h < 8 \\ \gamma(h) &= 26 && \text{if } h \geq 8 \end{aligned} \tag{8}$$

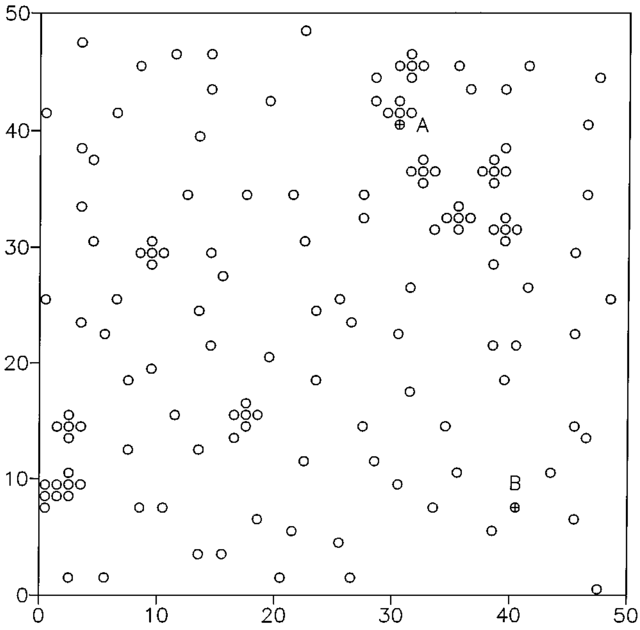

Figure 2. Location map for the data points of cluster.dat (Deutsch and Journel, 1992). Plus (+) signs show the two locations (A and B) to be estimated.

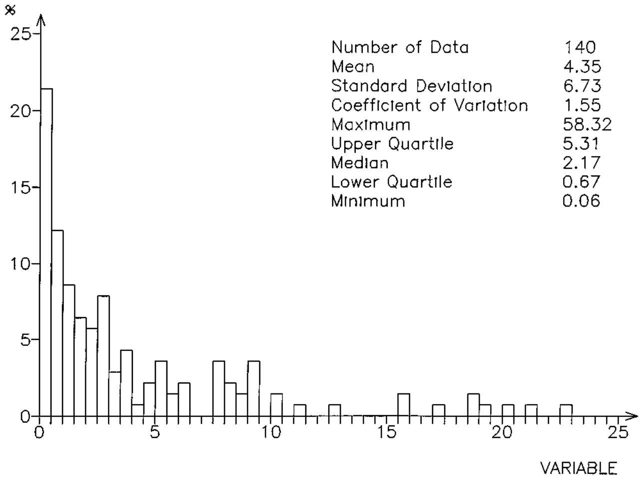

Figure 3. Histogram and statistics of the sample data set, cluster.dat.

Computing Kriging and Interpolation Variances

Figure 4 shows the estimated values for points A and B and their respective kriging (σ_{OK}) and interpolation standard deviations (S_o). Note that the data set for estimating point A presents larger variability than that used for estimating point B. Both points were estimated using the 8 nearest neighbor data points searched by the quadrant procedure. The kriging standard deviations computed for both points were equal to 4.09 and 4.85, showing that the data configuration for estimating point A is slightly better than for point B. However, the interpolation standard deviations (20.20 and 0.96, respectively for points A and B recognize the difference in data variability. In this regard, the interpolation standard deviation is a better measure of uncertainty. Because error variance is affected by the magnitude of the local variogram model, fluctuations in the local variability as apparent from the local data values must be taken into account when assessing uncertainty (Isaaks and Srivastava, 1989).

Cross Validation

Cross validation was used to test the relevance of the interpolation variance. This procedure provides true and estimated values for each sample location, so that actual estimation errors can be computed and compared with the measures of local accuracy proposed.

Ordinary kriging estimation was carried out at each sample location of cluster.dat using the eight nearest neighboring points (two nearest points per quadrant). In this case study no negative weight was detected in the cross validation procedure due to the high nugget effect of the variogram model considered [Eq. (8)]. Therefore, the procedure for correcting for negative weights was not applied. For each estimated point, kriging [Eq. (3)] and interpolation [Eq. (4)] variances were calculated.

The distributions of kriging and interpolation standard deviations are shown in histograms of Figure 5.

The kriging standard deviation values are concentrated in a few classes (just five) while the values for interpolation standard deviation are distributed over a wider range, showing a bimodal distribution separating estimations in regions with low and high dispersion of data values.

Define the true error as the difference between estimated and true values:

$$\text{True Error} = [z^*(x_o) - z(x_o)]$$

It can be either negative or positive; thus consider also the absolute true error:

$$\text{Abs True Error} = |z^*(x_o) - z(x_o)|$$

Figure 6 presents histograms of the true error and absolute true error.

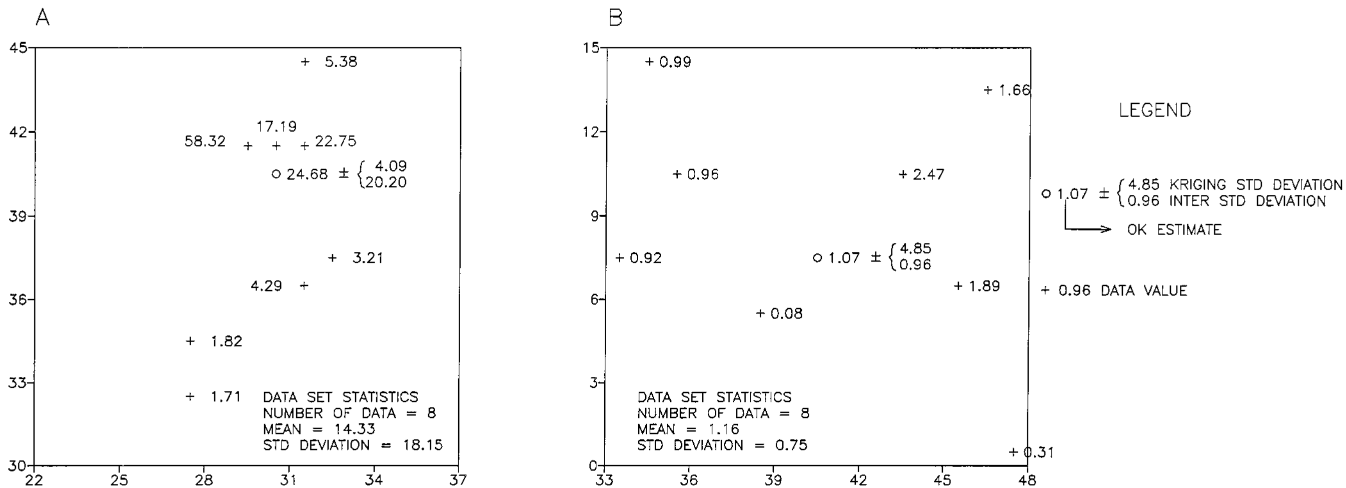

Figure 4. Resulting kriging and interpolation standard deviations for estimating locations A and B.

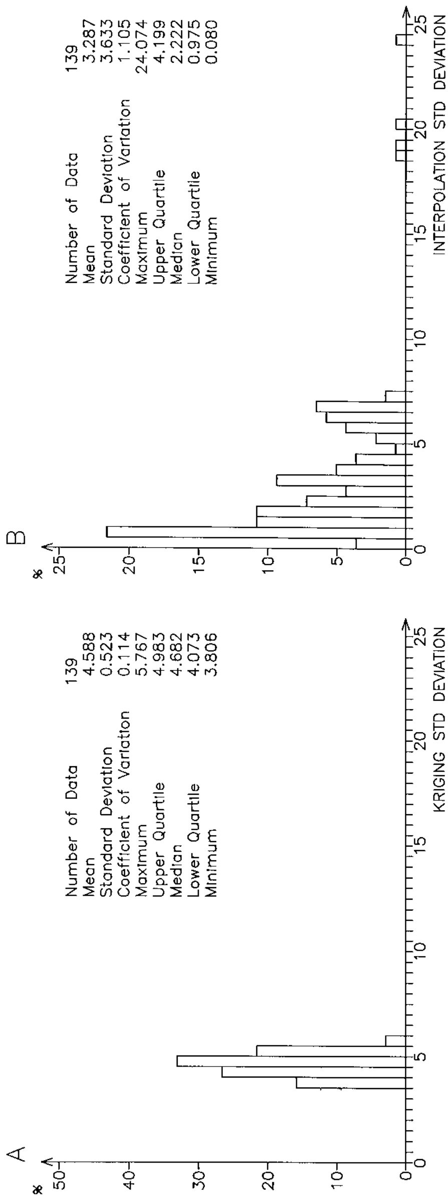

Figure 5. Histograms and statistics of the kriging standard deviation, A, and the interpolation standard deviation, B.

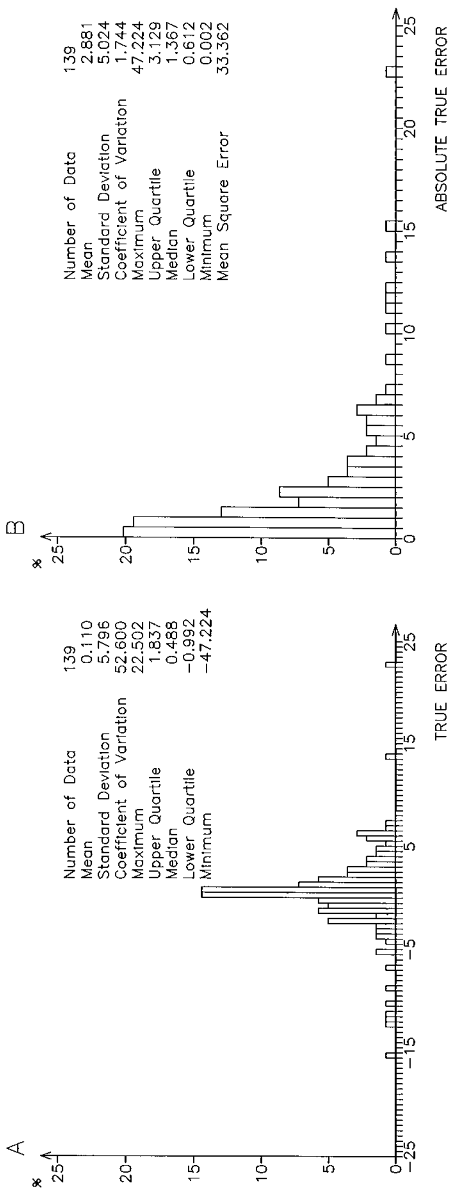

Figure 6. Histograms of true error, A, and absolute true error, B.

Because we are working with positive kriging and interpolation standard deviations, it makes sense using the absolute true error for comparison purposes. Actually, the comparison is done using both raw and ranked values. As done by Journel and Rossi (1989), the kriging and interpolation standard deviations as well as the absolute true errors were ranked by increasing values.

Figure 7 presents the correlation between the absolute true error and the kriging standard deviation. The kriging standard deviation has essentially no correlation with the local absolute true error; actually the correlation is negative: $\rho = -0.357$ and $\rho = -0.460$, respectively, for raw and ranked values. This result confirms that the kriging standard deviation is not a measure of local accuracy as stated by Journel and Rossi (1989).

Figure 8 shows the correlation between the absolute true error and the interpolation standard deviation. Both scattergrams show reasonable linear relationships with correlation coefficients equal to 0.462 (for raw data) and 0.670 (for ranked data), confirming that the interpolation standard deviation can be used as a measure of local uncertainty. It is important to remember that we are working with a difficult data set—cluster.dat–which presents some difficult features associated to clustering and high data dispersion (coefficient of variation $= 1.55$).

Recognizing Proportional Effect

Natural phenomena with skewed distribution of variables may exhibit proportional effect. According to Olea (1991), the proportional effect is a heteroscedastic condition in which the variance of the error is proportional to some function of the local mean of data values. Thus it is important to check that the interpolation standard deviation, as a local measure of uncertainty, recognizes the proportional effect when present. Previous cross validation kriging and interpolation standard deviations are plotted against OK estimates in Figure 9. The kriging standard deviation plotted against the OK estimate presents a negative linear correlation ($\rho = -0.646$), indicating that the kriging variance does not recognize local proportional effect. On the other hand, the correlation between the interpolation standard deviation and the OK estimate ($\rho = 0.932$) is significantly high, proving that this measure for uncertainty does recognize proportional effect.

CONCLUSION

This paper presents a simple expression to compute a measure of local uncertainty from the ordinary kriging weights as an alternative to the kriging variance. The proposed interpolation variance applies to OK for both point and block estimates. As opposed to the traditional kriging variance, which is homoscedastic,

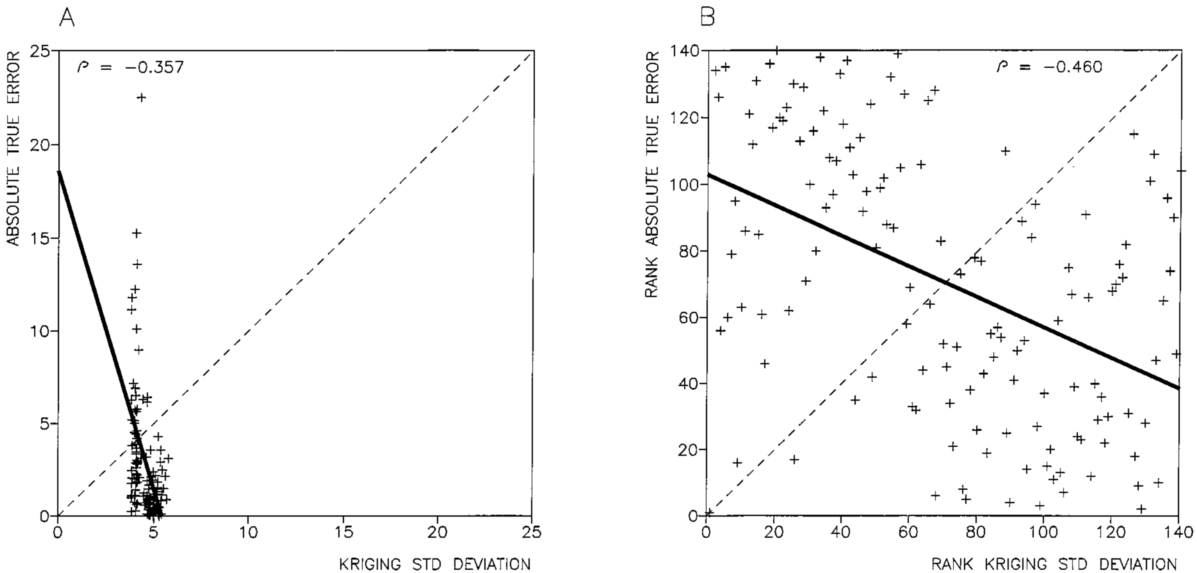

Figure 7. Correlation between absolute true error and kriging standard deviation, A, and the same after ranking, B. Dark lines represent minimum square regression lines.

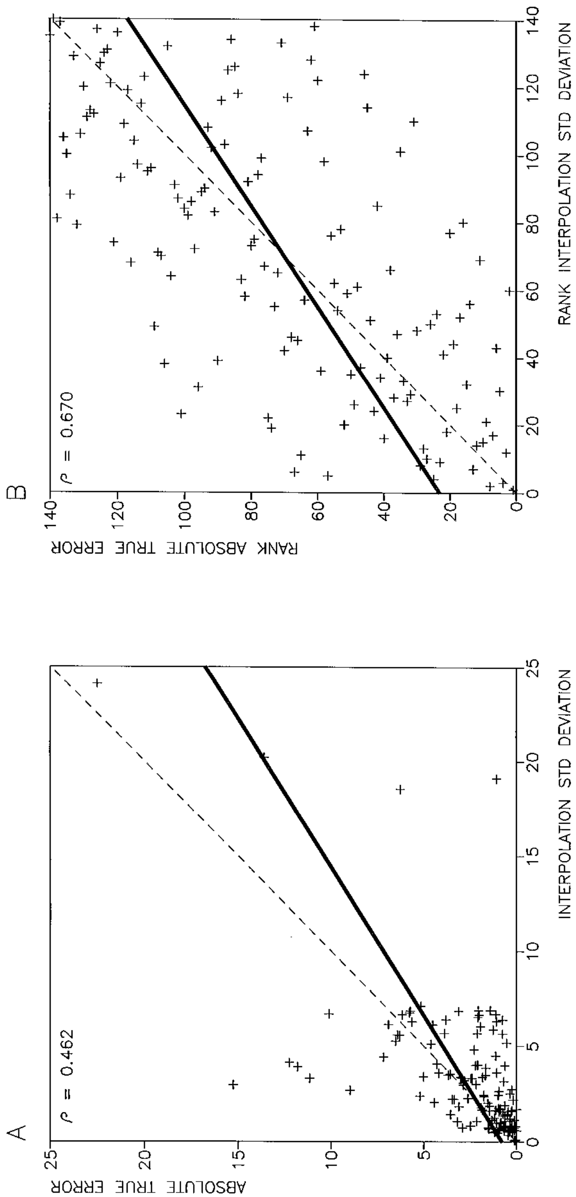

Figure 8. Correlation between absolute true error and interpolation standard deviation before, A, and the same after ranking, B. Dark lines represent minimum square regression lines.

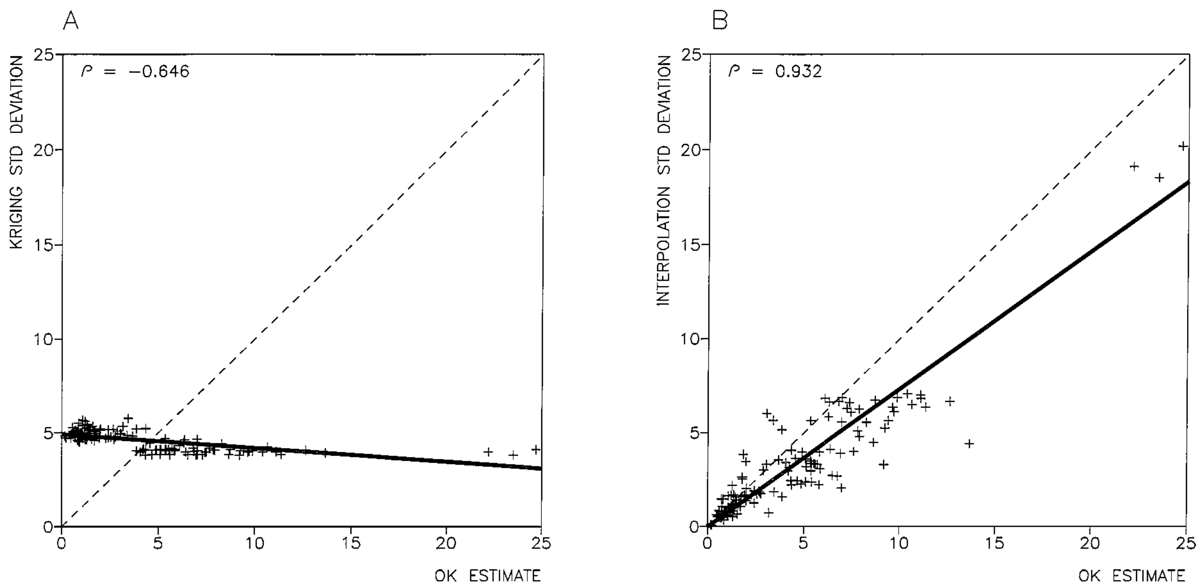

Figure 9. Scatterplots of kriging standard deviation vs. ordinary kriging estimate, A, and interpolation standard deviation vs. ordinary kriging estimate, B. Dark lines represent minimum square regression lines.

interpolation variance is both data-geometry and data-value dependent and therefore provides a more complete measure of local accuracy.

Tests carried out using the GSLIB's cluster.dat illustrate the usefulness of the interpolation variance as a reasonable measure of local uncertainty, recognizing local data variability as well as any proportional effect when present.

ACKNOWLEDGMENTS

I am very grateful to the National Council for Scientific and Technological Development-CNPq (Process 304612/89-8) as well as to the Foundation for Research Sponsorship of State of São Paulo-FAPESP (Processes 95/9964-2, 95/8812-4, and 96/7156-9) for the financial support to conduct this research. I sincerely wish to thank Prof. Andre G. Journel and Prof. Donald Myers for their invaluable suggestions and recommendations, which significantly improved this paper. Prof. Journel greatly contributed with constructive criticism and patience in reviewing previous submitted draft versions.

REFERENCES

Armstrong, M., 1994, Is research in mining geostats as dead as a dodo?: *in* Dimitrakopoulos, R., ed., Geostatistics for the next century: Kluwer Academic, Dordrecht, Netherlands, p. 303–312.

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

Deutsch, C. V., and Journel, A. G., 1992, GSLIB: Geostatistical software library and user's guide: Oxford University Press, New York, 340 p.

Diehl, P., and David, M., 1982, Classification of ore reserves/resources based on geostatistical methods: CIM Bulletin, v. 75, no. 838, p. 127–136.

Froidevaux, R., 1982, Geostatistics and ore reserve classification: CIM Bulletin, v. 75, no. 843, p. 77–83.

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

Herzfeld, U. C., 1989, A note on programs performing kriging with nonnegative weights. Math. Geology, v. 21, no. 3, p. 391–393.

Isaaks, E. H., and Srivastava, R. M., 1989, An introduction to applied geostatistics: Oxford University 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 (Report No. 9), Stanford, 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.

Mood, A. M., Graybill, F. A., and Boes, D. C., 1974, Introduction to the theory of statistics (3rd edition): McGraw-Hill, New York, 564 p.

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

Wober, H. H., and Morgan, P. J., 1993, Classification of ore reserves based on geostatistical and economical parameters: CIM Bulletin, v. 86, no. 966, p. 73–76.
Yamamoto, J. K., 1989, A new method for ore reserve estimation and modeling: Brasil Mineral, v. 68, p. 52–56. (In Portuguese.)
Yamamoto, J. K., 1991, Comparison of computational methods for ore reserve estimation: A case study in the Chapada Copper Deposit, State of Goiás, Brazil: unpubl. Ph.D. thesis, University of São Paulo, 175 p. (In Portuguese.)

APPENDIX A: PROPERTIES OF THE RANDOM VARIABLE S_o^2

In this appendix we develop expressions (5) and (6).

$$\begin{aligned} S_o^2 &= \sum_i \lambda_i [Z(x_i) - Z^*(x_o)]^2 = \sum_i \lambda_i [[Z(x_i) - m] - [Z^*(x_o) - m]]^2 \\ &= \sum_i \lambda_i [Z(x_i) - m]^2 + [Z^*(x_o) - m]^2 - 2[Z^*(x_o) - m] \sum_i \lambda_i [Z(x_i) - m] \end{aligned} \tag{A1}$$

Since

$$Z^*(x_o) = \sum_i \lambda_i Z(x_i), \quad \text{and} \quad \sum_i \lambda_i = 1, \text{ per ordinary kriging,}$$

then

$$\sum_i \lambda_i [Z(x_i) - m] = Z^*(x_o) - m$$

Hence,

$$S_o^2 = \sum_i \lambda_i [Z(x_i) - m]^2 - [Z^*(x_o) - m]^2, \text{ that is, expression (5).}$$

Now taking the expected value of the random variable S_o^2, rewritten as (A1), we have

$$\begin{aligned} E\left\{S_o^2\right\} &= \sum_i \lambda_i E\{[Z(x_i) - m]^2\} + E\{[Z^*(x_o) - m]^2\} \\ &\quad - 2 \sum_i \lambda_i E\{[Z^*(x_o) - m][Z(x_i) - m]\} \\ &= \sum_i \lambda_i \operatorname{Var}\{Z(x_i)\} + \operatorname{Var}\{Z^*(x_o)\} - 2 \sum_i \lambda_i \operatorname{Cov}\{Z(x_i), Z^*(x_o)\} \end{aligned}$$

Per stationarity:

$$\mathrm{Var}\{Z(x_i)\} = C(0), \quad \forall i \Rightarrow \sum_i \lambda_i \ \mathrm{Var}\{Z(x_i)\} = C(0)$$

Per the OK system (2) and the kriging variance (3):
it becomes:

- $$\begin{aligned} \mathrm{Var}\{Z^*(x_o)\} &= \sum_i \sum_j \lambda_i \lambda_j \ \mathrm{Cov}\{Z(x_i), Z(x_j)\} \\ &= \sum_i \lambda_i [\mathrm{Cov}\{Z(x_i), Z(x_o)\} - \mu] \\ &= \left[C(0) - \sigma^2_{\mathrm{OK}} - \mu\right] - \mu \\ &= C(0) - \sigma^2_{\mathrm{OK}} - 2\mu \end{aligned}$$

- $$\begin{aligned} \mathrm{Cov}\{Z(x_i), Z^*(x_o)\} &= E\left\{[Z(x_i) - m]\left[\sum_j \lambda_j Z(x_j) - m\right]\right\} \\ &= \sum_j \lambda_j E\{[Z(x_i) - m][Z(x_j) - m] \\ &= \sum_j \lambda_j \ \mathrm{Cov}\{Z(x_i), Z(x_j)\} \\ &= \mathrm{Cov}\{Z(x_i), Z(x_o)\} - \mu \end{aligned}$$

then

$$\begin{aligned} \sum_i \lambda_i \ \mathrm{Cov}\{Z(x_i), Z^*(x_o)\} &= \sum_i \lambda_i \ \mathrm{Cov}\{Z(x_i), Z(x_o)\} - \mu \\ &= C(0) - \sigma^2_{\mathrm{OK}} - 2\mu \equiv \mathrm{Var}\left\{Z^*_{\mathrm{OK}}\right\} \end{aligned}$$

Finally,

$E\{S_o^2\} = C(0) - \mathrm{Var}\{Z^*(x_o)\}$, i.e., expression (6), which is necessarily greater than or equal to zero, even if some of the weights λ_i are negative.

APPENDIX B: INTERPOLATION VARIANCE FOR AN OK BLOCK ESTIMATE

This appendix shows the development of expression (7). In the case of block kriging, provided that all point kriging estimates within the block V use the same

n data configuration:

$$Z_V^* = \sum_{i=1}^{n} \lambda_i^{(V)} Z(x_i) \equiv \frac{1}{nV} \sum_{l=1}^{nV} Z^*\big(x^{(l)}\big)$$

where $\lambda_i^{(V)}$ are the block kriging weights, and the $x^{(l)}$'s are the nV points discretizing the block V.

Per consistency of block kriging with the nV point krigings:

$$Z^*\big(x^{(l)}\big) = \sum_{i=1}^{n} \lambda_i^{(l)} Z(x_i)$$

$$\lambda_i^{(V)} = \frac{1}{nV} \sum_{i=1}^{nV} \lambda_i^{(l)}, \quad \forall i = 1, \ldots, n$$

where the $\lambda_i^{(l)}$ are the kriging weights for point $x^{(l)}$.

The interpolation variance of the block estimator Z_V^* is written as

$$\begin{aligned}
S_V^2 &= \sum_{i=1}^{n} \lambda_i^{(V)} \big[Z(x_i) - Z_V^*\big]^2 \\
&= \sum_i \frac{1}{nV} \sum_l \lambda_i^{(l)} \big[\big[Z(x_i) - Z^*\big(x^{(l)}\big)\big] - \big[Z_V^* - Z^*\big(x^{(l)}\big)\big]\big]^2 \\
&= \frac{1}{nV} \sum_l \Bigg[\sum_i \lambda_i^{(l)} \big[Z(x_i) - Z^*\big(x^{(l)}\big)\big]^2 + \big[Z_V^* - Z^*\big(x^{(l)}\big)\big]^2 \\
&\quad \times \sum_i \lambda_i^{(l)} - 2\big[Z_V^* - Z^*\big(x^{(l)}\big)\big] \sum_i \lambda_i^{(l)} \big[Z(x_i) - Z^*\big(x^{(l)}\big)\big]\Bigg]
\end{aligned}$$

with

- $S_l^2 = \sum_i \lambda_i^{(l)} \big[Z(x_i) - Z^*\big(x^{(l)}\big)\big]^2$, being the interpolation variance of point $x^{(l)}$;
- $\sum_i \lambda_i^{(l)} = 1, \quad \forall l;$
- $\sum_i \lambda_i^{(l)} \big[Z(x_i) - Z^*\big(x^{(l)}\big)\big] = Z^*\big(x^{(l)}\big) - Z^*\big(x^{(l)}\big) = 0, \quad \forall l$

it becomes

$$S_V^2 = \frac{1}{nV} \sum_l S_l^2 + \frac{1}{nV} \sum_l \left[Z_V^* - Z^*\left(x^{(l)}\right) \right]^2$$

which is the interpolation variance for an OK block estimate [expression (7)].