

Regression Revisited¹

Isobel Clark²

One traditional method of estimating ore reserves is to allocate the arithmetic mean of local samples to each block or stope within the deposit. Thirty years ago, Krige showed that this method produced a "bias" in the estimates, in that high values were commonly overestimated and low values underestimated. This bias became known as the "regression effect" because it could usually be corrected by an empirical regression relationship derived from previously mined areas. This paper shows how the regression effect can be reconciled with the geostatistical method of kriging and how a simple regression line may be derived before any areas are mined out. The method is similar to kriging with known mean, but acknowledges that the overall mean of the deposit may not be known precisely.

KEY WORDS: Kriging, geostatistics, regression effect.

INTRODUCTION

In many vein and reef deposits it is common practice to estimate the internal value of a block or stoping panel using peripheral samples from development drives or from raises. Typically, two opposite sides of a rectangular stoping panel have been sampled, say at 3- or 5-m intervals, whereas development drives may be 30 or even 50 m apart. In some South African gold reefs it is possible to develop on three or four sides of a panel before estimating the content of the slope. The traditional stope estimate was the arithmetic mean of all the peripheral samples. Providing these are evenly spaced around the panel, there seems much intuitive justification for this approach. In most of these deposits fairly intense slope sampling also takes place – partly to aid in immediate local estimation, and partly to check the "actual" value of the panel after it has been mined out.

Krige (1951) was among the first to notice that this method of estimation produced some discrepancy between the value expected in the stope and that actually found. Stopes for which high values had been predicted generally turned out to be somewhat poorer, while low-grade stopes – if they were mined, which was less common – seemed to have higher values than expected. When a graph is plotted of stope estimator versus actual stope value, it is immediately obvious that the relationship between the two variables is not a 45° line, but one with a shallower slope and a positive intercept on the vertical axis. One could argue for a more sophisticated predictor than a simple straight line, but this leads immediately into the realm of nonlinear estimators such as disjunctive kriging. These more sophisticated estimators are outside the scope of this present paper.

¹ Manuscript received 25 November 1980; revised 1 June 1982.

² Formerly Department of Mineral Resources Engineering, Imperial College of Science and Technology, London SW7 2BP, England.

Krige fitted lines through his data using the normal least-squares regression theory. Because of this, the so-called "bias" on stope estimation became known as the *regression effect*. Without the regression correction payable ground may be left out of the reserves, while seemingly excellent ground leads to disappointing mill feed grades. A regression relationship (linear or otherwise) between the "usual" linear estimator and the "true" stope value is often sufficient to correct this bias and lead to more accurate predictions.

There is only one major disadvantage to the empirical regression approach – it cannot be applied until a substantial amount of stope sample information is available. In the author's experience, at least 25 panels should be available before a *preliminary* regression line may be evaluated. This imposes a severe limitation on the potential use of the regression approach.

KRIGING

The development of kriging is well documented and will not be repeated here. Suffice it to state that

$$T^* = \sum w_i g_i$$

where g_i are the values (e.g., grades) of the members of the sample set S_i ; w_i are the weights allocated to those samples; and T^* is a weighted average used to estimate the true unknown value T . For an unbiased estimator in the absence of trend

$$\sum w_i = 1$$

Kriging produces the Best Linear Unbiased Estimator, that is, that set of $\{w_i\}$ which produce the minimum estimation variance, by setting up a system of simultaneous equations of the form

$$\sum w_j \bar{\gamma}(S_i, S_j) + \lambda = \bar{\gamma}(S_i, A), \quad \text{for } i = 1, 2, 3, \dots, n$$

$$\sum w_j = 1$$

where $\bar{\gamma}(S_i, S)$ is the average value of the semivariogram between sample set i and sample set j and $\bar{\gamma}(S_i, A)$ is the semivariogram between sample set i and the "unknown" area or volume. If this set of equations is solved, it can be shown that the kriging (estimation) variance is

$$\sigma_k^2 = \sum w_i \bar{\gamma}(S_i, A) + \lambda - \bar{\gamma}(A, A)$$

There is a variant on this form if it can be assumed that the overall average of the deposit (or suitable subvolume) is known, which will produce a slightly lower estimation variance.

It has been shown elsewhere (cf. Clark 1979c, Rendu 1979) that kriging and its variants may be applied to "two-dimensional" deposits and will produce results which are consistent with that expected from theory. However, when the estimates produced by strict kriging (i.e., with unknown overall mean) are examined more closely, they turn out simply to be the arithmetic mean of the peripheral development sampling. On the other hand, "kriging with known mean" will produce an estimator which closely resembles a linear regression line.

In most cases, then, we are caught between the two situations. Kriging with unknown mean returns the arithmetic average of the peripheral samples, an estimator that is known to be biased; kriging with known mean would help with the bias, but can rarely be applied in practice. Either of these methods could be applied using only development data, if this is all that is available. The alternative empirical approach, which has been proved over 30 years – the regression method cannot be used without a significant amount of stope sample information.

What seems to be needed is a regression line that can be produced *before* stope sampling is available – or equivalently, kriging with known mean when the overall mean is not *known* but has to be *estimated*.

If the following assumptions are made, the regression approach may be stated in geostatistical terms.

- (i) The overall volume under consideration is very large compared with the size of a single panel.
- (ii) The mean of the whole volume (global mean μ) may be estimated ($\hat{\mu}$) and the standard error of this estimate found, σ_{μ} . This might be done by strict kriging, or any other method as desired.

This approach may be expressed mathematically as

$$T = a + bT^* + \varepsilon$$

giving

$$\hat{T} = a + bT^*$$

that is, T is a simple regression type estimator. It is assumed for simplicity that the relationship between $\hat{T}$ and T^* is linear. Kriging, and hence T^* , is distribution free, so that T should also be unaffected by the distribution of the original values. It is obvious that for T to be unbiased

$$a = (1 - b) \mu$$

It remains to find an estimate of b that will give the "best" regression line. This regression relationship may then be used to produce the improved estimator T . The general form for the estimation variance of $\hat{T}$ is

$$\begin{aligned} \sigma_{\varepsilon}^2 &= E[(T - a - bT^*)^2] \\ &= E\{[T - (1 - b)\mu - bT^*]^2\} \end{aligned}$$

Differentiating this variance with respect to b produces

$$\begin{aligned} \partial \sigma_{\varepsilon}^2 / \partial b &= 2E\{[T - (1 - b)\mu - bT^*](\mu - T^*)\} \\ &= 2E[(T - \mu)(\mu - T^*) + b(T^* - \mu)^2] \\ &= 2bE[(T^* - \mu)^2] - 2E[(T - \mu)(T^* - \mu)] \end{aligned}$$

which should be set to zero to find the minimum, giving

$$\hat{b} = \frac{E[(T - \mu)(T^* - \mu)]}{E[(T^* - \mu)^2]} = \frac{\sigma_{TT^*}}{\sigma_{T^*}^2}$$

where σ_{TT^*} is the covariance between the kriging estimator and the true unknown value, and $\sigma_{T^*}^2$ is the variance between kriging estimates for all like panels over the whole volume.

This (of course) is the usual least squares regression relationship, and is equivalent to kriging with known mean.

However, in practice μ is not known and the estimate $\hat{\mu}$ is used instead. Once a value is defined for b

$$a = (1 - b) \hat{\mu}, \quad \text{so that}$$

$$\sigma_\epsilon^2 = E\{[T - (1 - b) \hat{\mu} - bT^*]^2\}$$

and for the minimum estimation variance

$$\frac{\partial}{\partial b} E\{[T - (1 - b) \hat{\mu} - bT^*]^2\} = 0$$

Performing the differentiation gives

$$2E\{[T - (1 - b) \hat{\mu} - bT^*] (\hat{\mu} - T^*)\}$$

$$= 2E\{[(T - \mu) - b(T^* - \mu) - (1 - b)(\hat{\mu} - \mu)] [(\hat{\mu} - \mu) - (T^* - \mu)]\}$$

Given assumption (i) above and the implicit assumptions of kriging, $\hat{\mu}$ should be independent of T and also of T^* . That is, the estimate of the global mean is independent of the value (or estimated value) of any single panel. Under this condition many of the cross-products in the above equation vanish, leaving

$$-2E[(T - \mu)(T^* - \mu)] + 2bE[(T^* - \mu)^2] - 2(1 - b)E[(\hat{\mu} - \mu)^2]$$

$$= -2\sigma_{TT^*} + 2b\sigma_{T^*}^2 - 2(1 - b)\sigma_\mu^2$$

which should be set to zero to minimize the estimation variance, giving

$$\hat{b} = (\sigma_{TT^*} + \sigma_\mu^2) / (\sigma_{T^*}^2 + \sigma_\mu^2) \quad (1)$$

For a deposit which has a semivariogram with a total sill C_{tot} , the covariance and variance may be expressed in terms of the semivariogram as follows

$$\sigma_{TT^*} = C_{tot} - \sum_i w_i \bar{\gamma}(S_i, A)$$

$$\sigma_{T^*}^2 = C_{tot} - \sum_i \sum_j w_i w_j \bar{\gamma}(S_i, S_j) \quad (2)$$

If T^* is the kriging estimator, found by the normal kriging procedure, then the latter term reduces to

$$\sigma_{T^*}^2 = C_{tot} - \sum w_i \bar{\gamma}(S_i, A) + \lambda$$

In fact, eq. (1) holds for any linear estimator T^* , although it would seem sensible to assume that the kriging estimator is the best one to use.

The calculation of a and b uses only terms that have already been calculated by the kriging system *plus* the estimate of the global mean and its associated standard error. These latter terms, of course, need only be calculated once for any deposit (or subsection of deposit).

The general form of the estimation variance for *any* defined value of b – provided that $a = (1 - b) \hat{\mu}$ – is given by

$$\sigma_r^2 = (1 - b)^2 \sigma_\mu^2 + \sigma_T^2 + b^2 \sigma_{T^*}^2 - 2b\sigma_{TT^*} \quad (3a)$$

or in terms of the semivariogram

$$\sigma_r^2 = (1 - b)^2 (C_{\text{tot}} - \sigma_\mu^2) + 2b \sum w_i \bar{\gamma}(S_i, A) - b^2 \sum \sum w_i w_j \bar{\gamma}(S_i, S_j) - \bar{\gamma}(A, A) \quad (3b)$$

If μ is known, and $\hat{b}$ takes the above form, then the estimation variance becomes

$$\sigma_{ls}^2 = (1 - \hat{b}) C_{\text{tot}} + \hat{b} \sum w_i \bar{\gamma}(S_i, A) - \bar{\gamma}(A, A)$$

where ls has been used to denote the normal least squares estimation. It can also be seen that if $b = 1$, eq. 3 reduces to the strict (linear) kriging case.

It can be shown analytically that σ_{ls} is always less than σ_k . In practice so far, σ_r has always been less than σ_{ls} , but (obviously) greater than σ_{ls} .

To summarize, the technique which is proposed is an attempt to generalize the "kriging with known mean" to the case where the mean is not known, but can be estimated. This gives a more widely applicable method, which seems to take care of the bias problem. The technique is simple and robust in practice, although it obviously does not have the potential power of the nonlinear type of estimator, as proposed in disjunctive kriging and the Multivariate Gaussian approaches of Matheron. The first case study is a short synopsis of the detailed study which lead to the development of the "geostatistical regression" method (georegression).

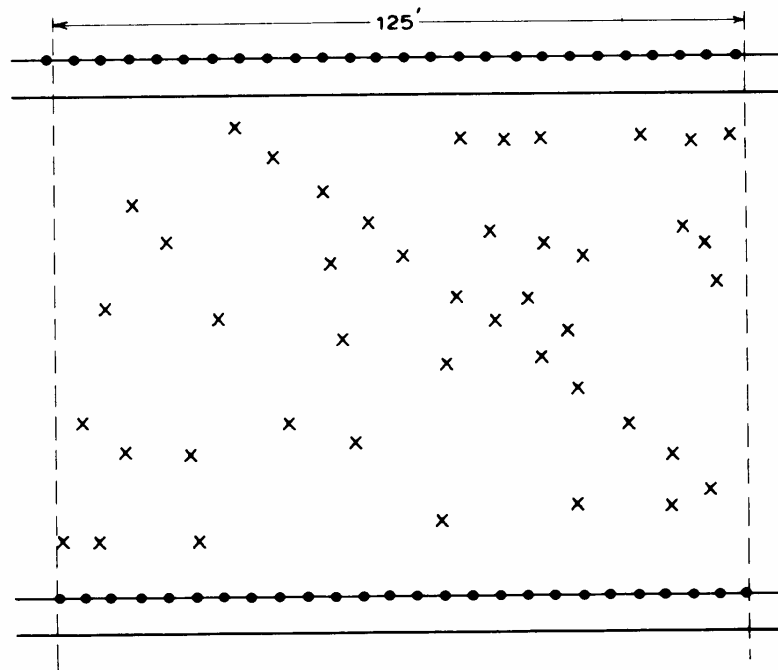


Fig. 1. Typical stope panel on SWF showing drive and stope sampling.

CASE STUDY 1 - CORNISH TIN LODGE

An intensive study of various local estimation methods has been carried out by the author (Clark 1979a) using sample data from the Simms Lode of Geevor Tin Mines, Cornwall. This illustration will consider a geologically homogeneous subsection of the lode, which will be referred to as "Simms West of the Fault" (SWF). Development drives are available at approximately 100-ft

levels within this almost vertical lode, and these have been chip sampled at 5-ft intervals. The one exception to this in the study area is the topmost level (6) which was sampled at 10-ft intervals. Four of the levels were extensively mined by the 31st of March, 1976, and a total of about 2700 chip samples were available in a fairly regular pattern throughout these stopes. For this investigation it was decided that panels 125 ft along the strike and approximately 100-ft high (between drives) would be estimated. This gave 49 non-overlapping panels in the study area.

Each panel has up to 52 peripheral samples from drives on two parallel sides, and between 27 and 65 interior samples. Most panels contained about 45 stope samples. It is generally assumed that the interior sampling adequately represents the "true" stope panel average. A typical panel is shown in Fig. 1, with different symbols for drive and stope sampling.

Two variables are measured at each sample site: the assay value of the ore (at that time in pounds per ton of SnO_2 ; and the width of the lode (in inches). For this study, the lode widths have been selected as the variable of interest. For each panel we have peripheral drive samples – on which local estimation would normally be made – and stope samples, which should give an accurate reflection of the stope panel value. It can be shown that kriging produces an estimator that is effectively identical to the arithmetic mean of the peripheral development samples. The one exception to this is in the case of the topmost stope, where level 6 was sampled at a different interval. In that case, the kriging estimator was equal to the average of the two individual drive means.

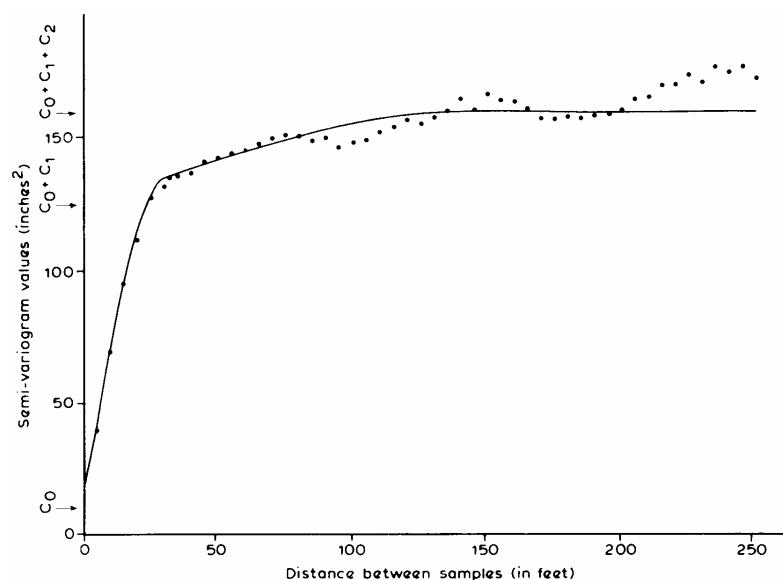


Fig. 2. Experimental semivariogram for lode widths on SWF and model.

The semivariogram of the lode widths was calculated from the drive samples only, and is shown in Fig. 2. The model comprises two intermeshed spherical schemes with ranges of 30 and 150 ft, and a small nugget effect. From this we find that using drive samples to estimate the stope panels produces an estimator equal to the arithmetic mean of the samples with a standard error (σ_k) of about 3.65 in. Over the whole area (SWF) the average lode width is 23 in.

When estimating the value of the stope panel using the stope samples, the kriging system *does* improve over the arithmetic mean. This is especially true where the spatial distribution of the

samples is erratic. Slope estimation standard errors over the 49 panels range between 1.03 and 2.25 in.

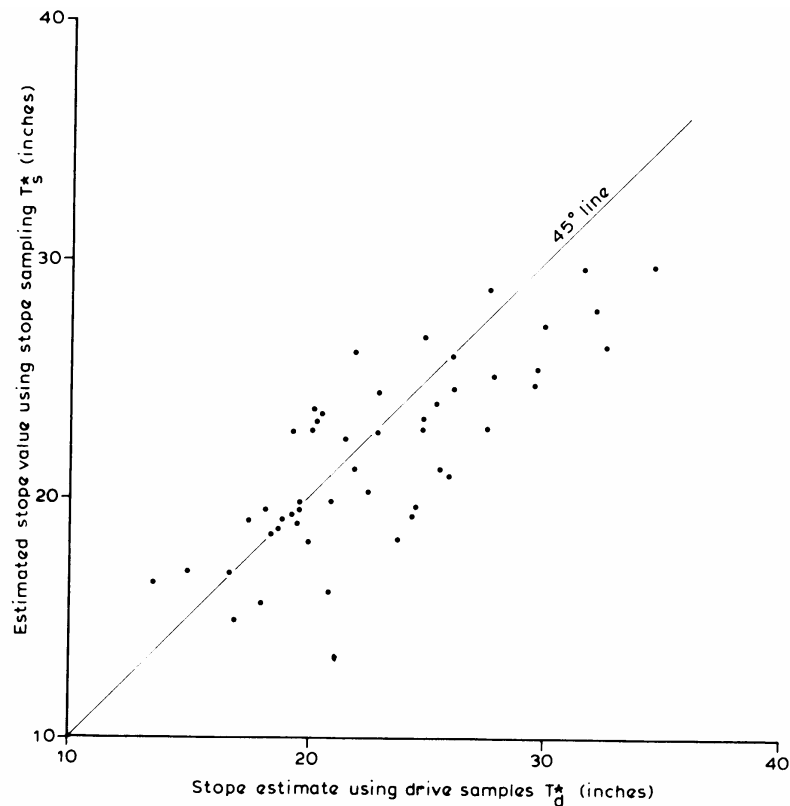


Fig. 3. Comparison of stope values with drive values for lode widths on SWF.

The true average lode width value over a stope panel may be denoted by T ; the estimate produced from the drive samples by T_d^* , and that from the stope samples by T_s^* . Figure 3 shows a scattergram of T_s^* versus T_d^* for the 49 panels of this study. The 45° line is also shown, and it can clearly be seen that the development average (a linear estimator) produces a consistent bias in the estimation of the stope value. High development values are associated with high stope values, but the value in the stope is generally lower than that in the drives. The reverse is true at the lower end of the graph – the drives systematically underestimate the stopes. Figure 4 shows the relationship of the kriging error to both T_s^* and T_d^* . The estimation errors ($T_s^* - T_d^*$) are highly correlated with the estimator (-0.6) but not with the stope values (-0.05). Figure 4a especially illustrates the bias in the estimation. High drive values will give considerable overestimation and vice versa.

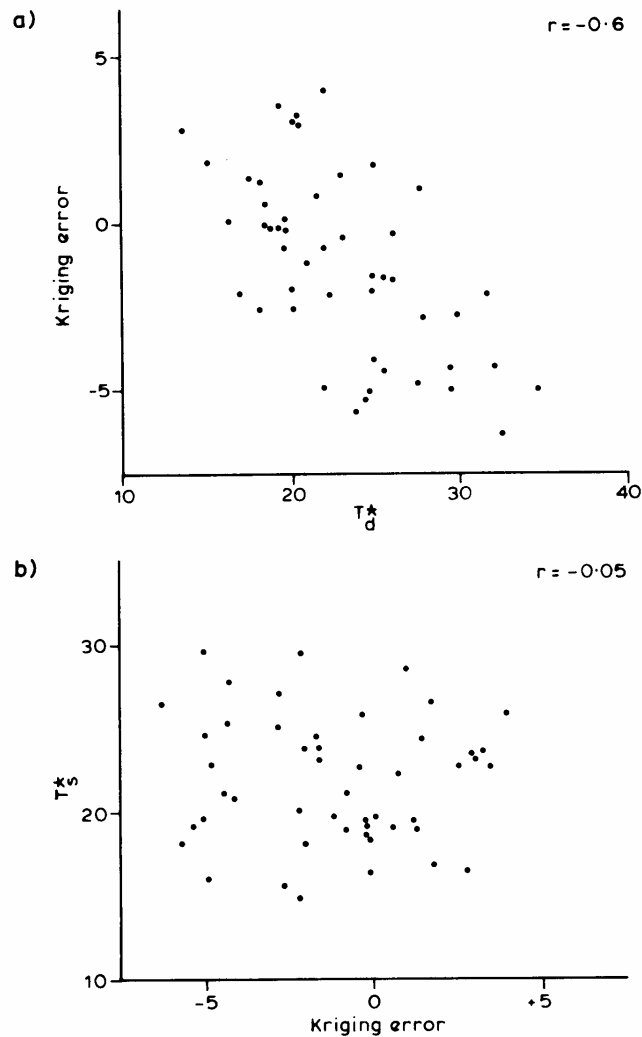


Fig. 4. Kriging error ($T_s^* - T_d^*$) versus estimates.

With this amount of data it is possible to fit an empirical regression line to the drive/stope relationship. The least squares regression line is shown in Fig. 5, and the discrepancy between it and the 45° line can easily be seen. An envelope has been drawn outlining the data points from the known stopes. The line shown is

$$T_s^* = 7.102 + 0.643 T_d^*$$

Classical statistical theory leads to a confidence limit on this line, such as that shown in Fig. 5. This boundary is narrowest close to the mean stope value and widest at the edges of the graph. In theory, a lower $\alpha\%$ confidence level is given by

$$C_\alpha = a + bT_d^* - t_{(1/2)\alpha, n-2} s_{y/x} \left[1 + (1/n) + \left\{ (T_d^* - \bar{x})^2 / \sum (x_i - \bar{x})^2 \right\}^{1/2} \right], \text{ e.g.}$$

$$C_{90} = 7.102 + 0.643 T_d^* - 1.300 \times 4.950 \left\{ 1 + \frac{1}{49} + [(T_d^* - 23.00)^2 / 1120] \right\}^{1/2}$$

For example, when T_d^* equals 40 in., the predicted value in the stope is 32.8 in. with a lower 90% confidence limit of 25.6 in.

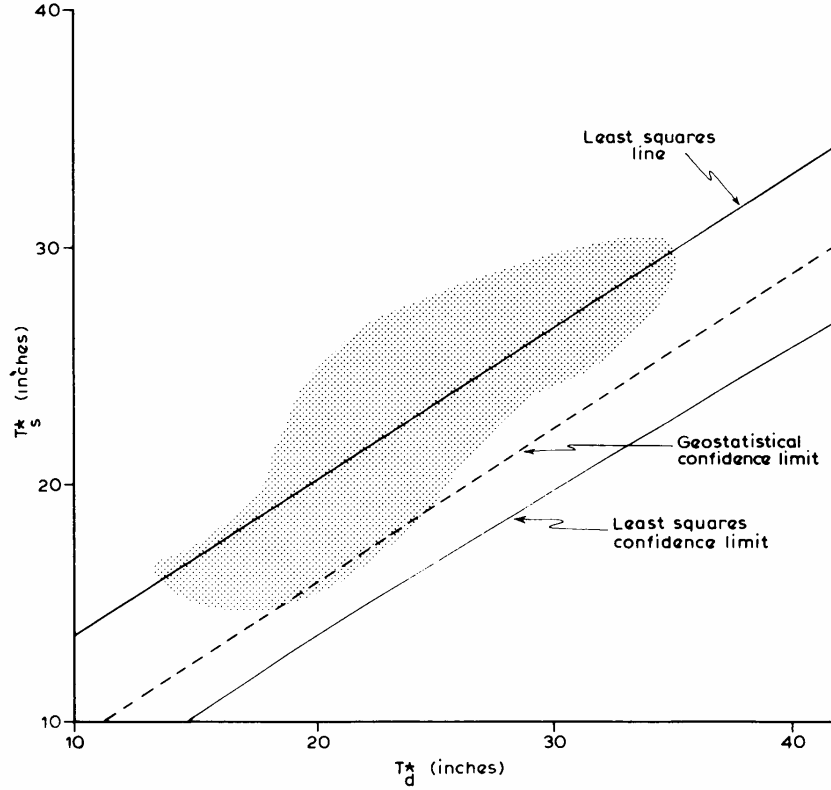


Fig. 5. Empirical regression line and supposed confidence limits.

It is interesting to note that this limit is lower than the corresponding kriging limit, which would be constant at the predicted value minus $1.28 \times 3.65 (= 4.67 \text{ in.})$. On the other hand, these confidence limits are based on the assumption that all of the errors are independent and identically distributed, and also that the observations (T_s^* and T_d^*) are randomly and independently collected. These assumptions would not seem realistic in the present case.

Using the geostatistical approach and knowledge of panel geometry, we find that eq. (3b) may be used to produce a more realistic standard error for the regression relationship. Since we are using the overall mean of the stope panels in the calculation of the a coefficient, we may assume that the global mean is known, so that σ_μ is zero. The semivariogram terms are different for each of the panels, but an average value may be used here, for illustration purposes. $\bar{\gamma}(S,A)$ gives a typical value of 146.3464; $\bar{\gamma}(S,S)$ is 139.0662; and $\bar{\gamma}(A,A)$ is 140.3155. Since the estimator being used is the arithmetic mean of the development samples, we can write the variance formula without weightings, so that

$$\begin{aligned}\sigma_e^2 &= (1 - \hat{b})^2 C_{\text{tot}} + 2\hat{b}\bar{\gamma}(S, A) - \hat{b}^2\bar{\gamma}(S, S) - \bar{\gamma}(A, A) \\ &= 0.357^2 \times 160 + 1.286 \times 146.3464 - 0.643^2 \times 139.0662 - 140.3155 \\ &= 10.78 \text{ in.}^2\end{aligned}$$

This gives a standard error for the regression line of 3.28 in. and a lower 90% confidence level 4.2 in. below the regression line. Note that this is a major difference between the geostatistical and the classical confidence levels. Both of these lines are shown in Fig. 5. The difference between these two confidence levels may be explained by the difference in approach between the two theories. The classical statistical approach assumes that the line produced is an *estimate* of the true line and hence has an error associated with it *in addition* to the so-called random error. This is why the confidence bands widen toward the edges of the graph. The geostatistical approach assumes that the line is the true one since it is built up from a model of the deposit, and not the local sample values. Perhaps a truly realistic bound would be produced by a compromise between these approaches – where the semivariogram model would be regarded as an estimate of the true model of the deposit.

This empirical regression line has been constructed from 49 stope panels which were estimated using over 2500 stope samples. It can be seen that the standard error of the panel estimates has been reduced from the kriging value of 3.65 in. to one of 3.28 in. using the regression line. This is an improvement in estimation of over 10%. However, by this stage the lode has been nearly half mined. An estimator is required at a much earlier mining stage.

Using eq. (1) and (2) above, a regression relationship can be produced using only the following information

- (a) the semivariogram constructed on the drive samples
- (b) the size and shape of the stope panel
- (c) the drive samples peripheral to the panel
- (d) the overall mean of the study area and (if necessary) its standard error

Assuming that no stope samples are available, the procedure is then as follows

- (i) estimate the overall mean and its standard error
- (ii) for each panel the kriging estimator T_d is produced. In this process the terms

$$\sum w_i \bar{\gamma}(S_i, A), \quad \sum \sum w_i w_j \bar{\gamma}(S_i, S_j), \quad \bar{\gamma}(A, A)$$
 have been evaluated.
- (iii) correct T_d^* by calculating $\hat{b}$ and $\hat{a}$ as given in eq. (1), to produce $\hat{T}$, an improved estimate.
- (iv) calculate the standard error of this estimate using eq. 3.

In practice each panel has a slightly different configuration, and hence slightly different $\hat{a}$ and $\hat{b}$ values. To provide an overall "georegression" line for comparison with the empirical line, the various terms were again averaged over the 49 panels. The following typical georegression line was produced

$$\hat{T} = 8.004 + 0.652 T_d^*$$

which has a constant standard error of 3.28 in.

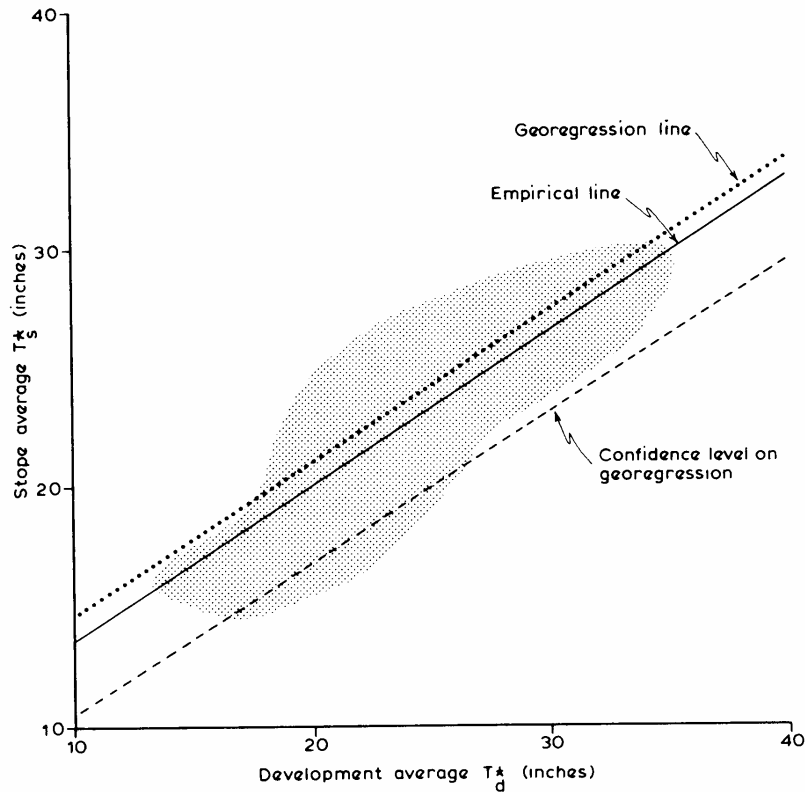


Fig. 6. Comparison between georegression and empirical regression lines.

This georegression line, constructed on development data only, is shown alongside the empirical least squares regression line which used over 2500 stope samples, in Fig. 6. The lower 90% confidence limit on the georegression is shown, in addition to the outer envelope of the stope panel data values.

CASE STUDY 2-SIMULATED IRON ORE EXAMPLE

To illustrate the effect of georegression when used as a contouring or point estimation method, a simple set of previously published data (Clark 1979b) has been used. A simplified low-grade iron ore deposit has been sampled by means of 50 randomly positioned boreholes perpendicular to the dip of the orebody (see Fig. 7). This problem is simplified by considering only the average grade over the borehole intersection as a point in two-dimensional space. An area 400 m² is selected for contouring, and the data is known to have a spherical semivariogram with a range of 100 m and a sill of 25 % Fe². That is, the standard deviation of the iron grades is 5 % Fe. The average grade of the selected area is estimated to be 34.04 % Fe with a standard error of 0.56 % Fe.

Three different approaches are illustrated here for comparison purposes. These, of course, only represent a selection of the possible approaches that could be used for this problem.

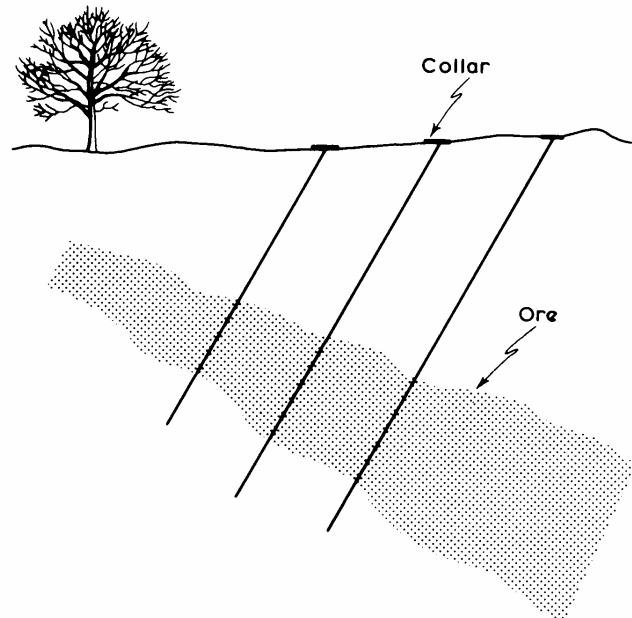
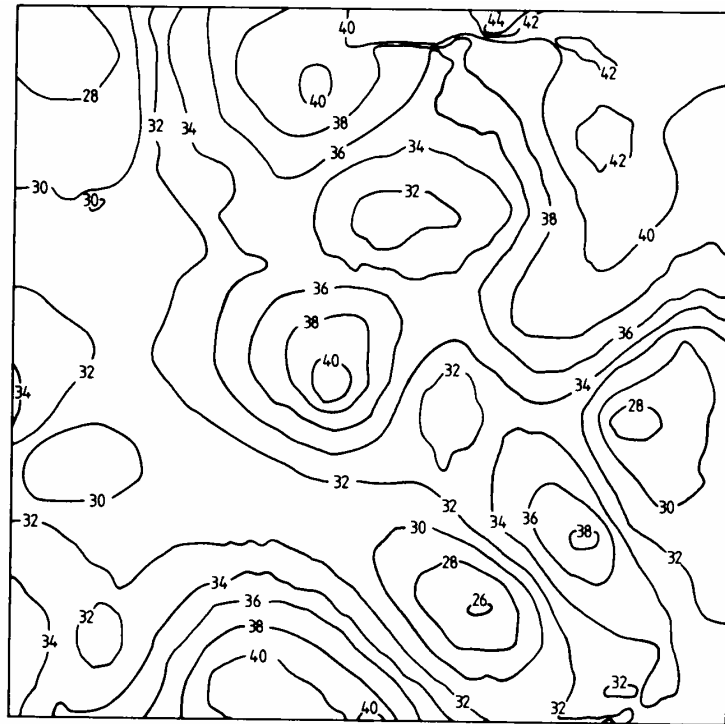


Fig. 7. Hypothetical iron ore deposit.

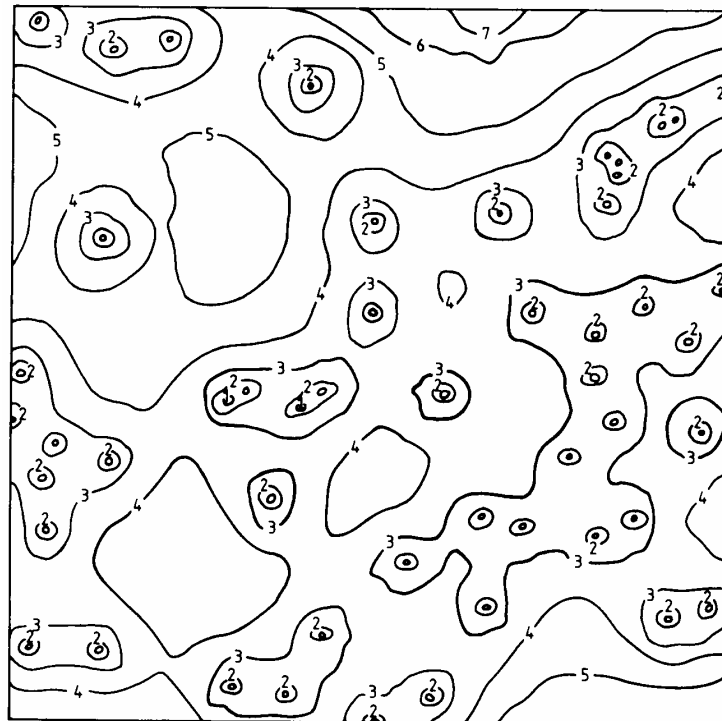
First, "strict" kriging was applied to the data set. That is, only samples which lie within the range of influence (100 m) of a location are used in the estimation of the value at that position. The interpolated map and its corresponding standard error map are shown in Figs. 8a, b, respectively.

It can be seen that the standard errors vary between zero — at a datum point — and $5\sqrt{2}$, where there are actually no samples within the range of influence of the point. For the interpolated map, such points have been allocated the value of the nearest sample. These maps were constructed by estimating the value at points on a 5-m grid over the area, that is, a total of 6561 points. On the Imperial College Cyber 174 computer, strict kriging required 45 sec in Central Processor (CP) time. It is useful to note from Fig. 8b the poorly estimated areas in the north and in the southeast corner. The interpolated values in these areas can be seen to be relatively unstable, giving erratic contours as opposed to the smooth variation in the "well" estimated areas.

The second set of maps, in Fig. 9, was produced using *all 50* samples in each estimation. For each of the 5-m grid points all samples were included in a kriging system. This is still kriging with unknown mean, but using all available information in every situation. It can easily be seen that in the well-sampled areas ($\sigma k < 4$) there is no difference in standard error, and little in the predicted values. The standard error contour at 4 % Fe is marginally affected, but the real improvement comes in the really badly estimated areas where the standard error is greater than 4 % Fe. The maximum σk in Fig. 9 is 5.1 1 % Fe as opposed to 7.07 % Fe in Fig. 8. The new stability can also be seen in the interpolated map, where the unknown area is now assigned a value close to that of the overall average — 34.20 % Fe. Note that this estimator is *not* equal to the arithmetic mean of all the samples. Table 1 shows the weights accorded to each sample (samples ordered as in Clark 1979b) when estimating these isolated points. Instead of a constant weighting of 0.02, the weights vary between -0.004 and 0.043. This is, of course, because the kriging system takes into account the relative intersample positions as well as the relationship between sample and point to be estimated. Thus, two sample points close together will share weight, while a relatively isolated point will gain more importance in the estimation.

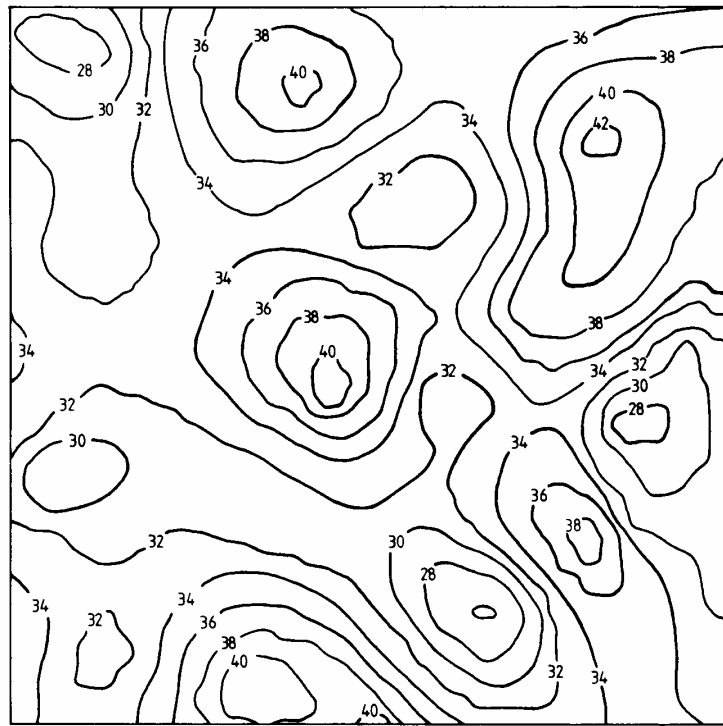


(a)

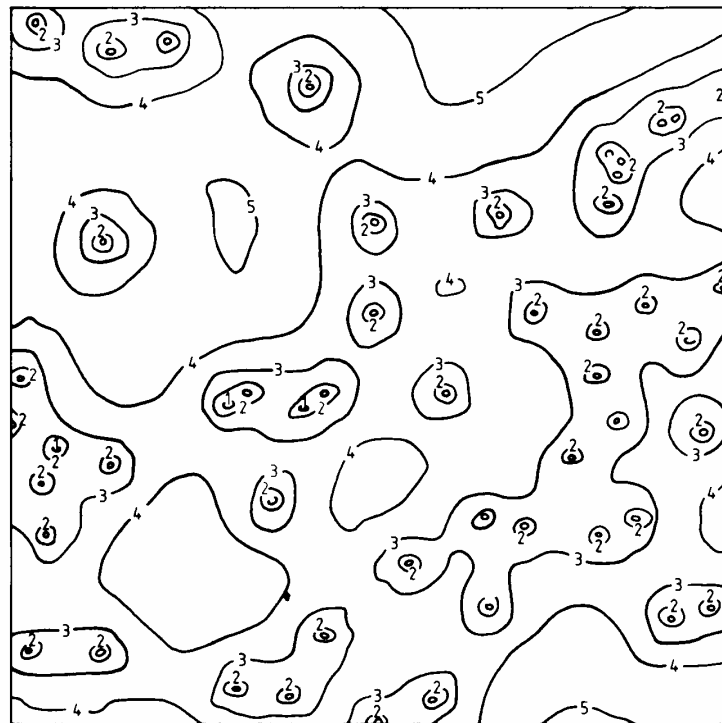


(b)

Fig. 8. (a) Interpolated map using "strict" kriging, (b) standard error map using "strict" kriging.



(a)



(b)

Fig. 9. (a) Interpolated map using kriging on all 50 samples, (b) standard error map from kriging using all 50 samples.

It is interesting to compare these weights with those accorded to the samples when estimating the overall average of the rectangular area. Table 2 shows the corresponding weights for estimating μ . Once again, the weights vary markedly from the arithmetic mean, ranging from 0.004 to 0.048. The distribution of the weights is quite different from that in Table 1, although the final overall average is similar.

Table 1. Kriging Weights for Isolated Points³

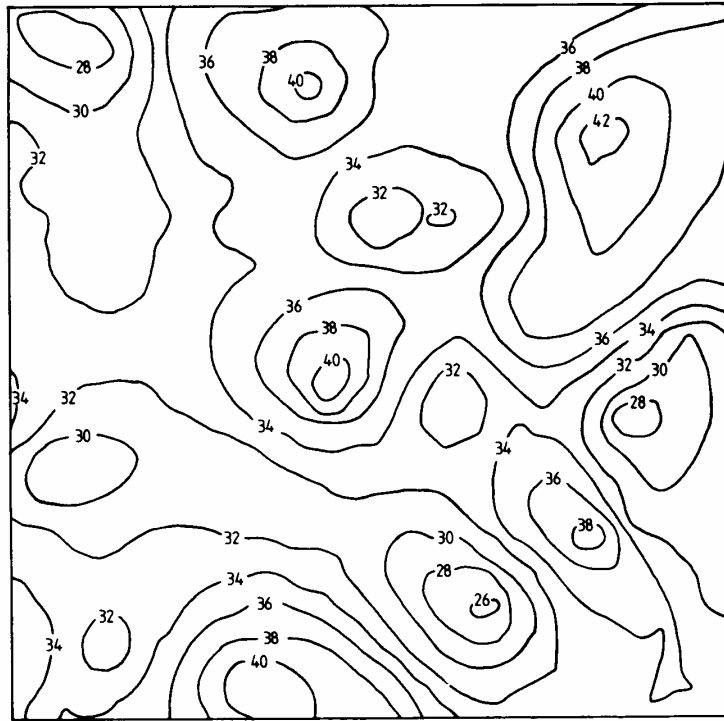
0.014	0.029	0.010	0.027	0.031
0.017	0.018	0.004	0.031	0.014
0.022	0.021	0.010	0.011	0.006
0.010	0.009	0.020	0.037	0.032
0.043	0.032	0.036	0.007	0.005
0.030	0.022	-0.004	0.027	0.012
0.032	0.013	0.016	0.025	0.023
0.022	0.028	0.014	0.016	0.033
0.011	0.004	0.021	0.014	0.035
0.041	0.029	0.008	0.020	0.012

Table 2. Kriging Weights for Estimating Overall Average ($\hat{\mu}$)³

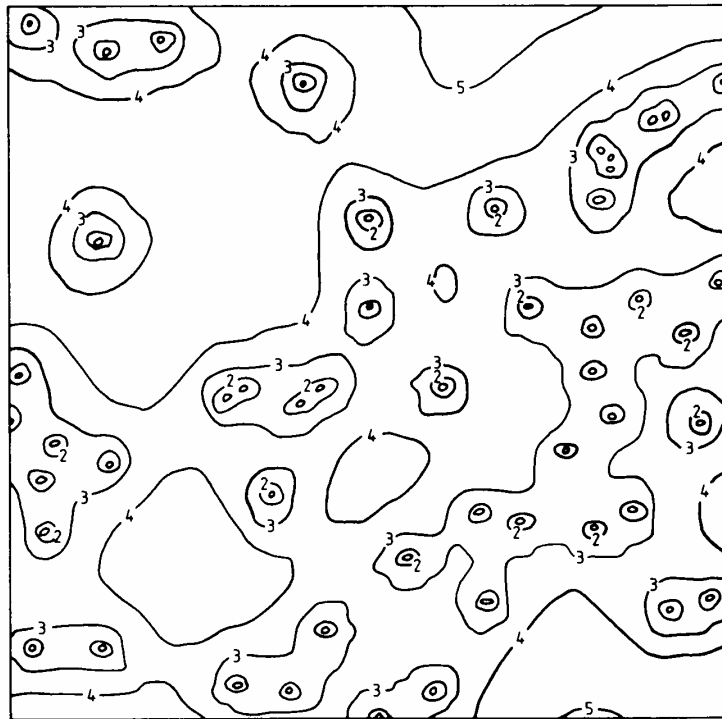
0.002	0.025	0.016	0.018	0.004
0.021	0.005	0.048	0.033	0.034
0.020	0.026	0.021	0.028	0.016
0.037	0.007	0.004	0.045	0.026
0.018	0.038	0.024	0.011	0.028
0.021	0.036	0.016	0.027	0.035
0.011	0.023	0.018	0.011	0.015
0.011	0.022	0.014	0.011	0.005
0.015	0.015	0.010	0.029	0.012
0.007	0.025	0.015	0.022	0.019

It is difficult to quote a relative timing for the production of these maps. Using a generalized kriging routine – that is, estimating each point separately requires an astronomical amount of CP time. To kriging just one point using all 50 samples takes 0.80 sec on the Cyber 174. This implies a run-time of something like 5250 sec for the 6561 points on the 5-m grid. On the other hand, writing a tailor-made computer program cuts down the CP requirement. Even then the 5-m grid took approximately 135 sec in CP, a three-fold increase over strict kriging. However, the improvement in prediction would seem to justify the extra expenditure.

³ Order as in Clark, 1979b



(a)



(b)

Fig. 10. (a) Interpolated map using strict georegression, (b) standard error map using strict georegression.

The final pair of maps, shown in Figs. 10a, b, were produced using the georegression correction on a strict kriging estimator. Effectively, the maps in Fig. 8 were calculated and then each point corrected by its appropriate geostatistical regression factors. The regression coefficient $\hat{b}$ was calculated as described in eq. (1), with a corresponding $\hat{a}$ coefficient. As with Fig. 9 the areas most affected are the poorly estimated ones. The highest standard error now is 5.03 % Fe, which is not significantly different from that of Fig. 9. However, it may be noticed in Fig. 10b that the areas with standard errors greater than 5 % Fe have been reduced-one contour in the northwest vanishes entirely. Comparison of the two interpolated maps, Figs. 9a, and 10a, show minor differences mainly in the improved areas around the regions with highest standard error.

Georegression estimation of the 5-m grid required 46 sec of CP time, plus almost 1 sec to estimate the overall average of the area. This represents an increase in computer time of less than 5% over the strict kriging routine. It is fairly obvious that the more points (or panels) estimated the smaller will be the comparative increase in calculation time.

SUMMARY

The principle of the regression relationship in mining is fairly well established in practice. The ability to quantify the continuity of a deposit by means of a semivariogram and to produce optimal estimators is the main advantage of a geostatistical analysis. These two factors have been combined to produce a simple, robust (possibly nonlinear) estimator to improve the prediction given by linear kriging. Comparative timings on computer runs show that georegression produces a (slightly) better result than the best possible kriging system, in approximately one-third of the CP time. Of course, this factor will vary from one application to another. The more samples available, the greater will be the advantages of using the georegression approach. Programming the georegression correction into an existing kriging system requires the addition of only four or five lines of code.

The two case studies which have been discussed illustrate the application of this method. Another two case studies with somewhat less ideal characteristics are presented elsewhere (Clark and Clausen, 1981).

ACKNOWLEDGMENTS

The georegression method was developed as part of the research undertaken for the degree of Ph.D. from the University of London. The author particularly wishes to thank the staff at Geevor Tin Mines Ltd. for their help and encouragement during that study. The simulated iron ore example was produced for Clark (1979b) by R. P. B. Durham at Imperial College.

REFERENCES

- Clark, I., 1979a, Geostatistical reserve estimation in Cornish tin mining: Unpublished Ph.D. thesis, University of London, 360 p.
- Clark, I., 1979b, Practical geostatistics: Applied Science, Barking, Essex, 129 p.
- Clark, I., 1979c, Does geostatistics work?: Proceedings of the 16th APCOM Symposium, Tucson, Arizona (SME-AIME Publ), pp. 213-225.

- Clark, I. and Clausen, F. L., 1981, Simple alternative to disjunctive kriging: Trans. Inst. Min. Metall., v. 90, pp. A13-A24.
- Jackson, M. and Marechal, A., 1979, Recoverable reserves estimated by disjunctive kriging: a case study: Proceedings of the 16th APCOM symposium, Tucson, Arizona, p. 240-249.
- Krige, D. G., 1951, A statistical approach to some basic mine valuation problems in the Witwatersrand: Jour. Chem. MetaR. Min. Soc. S. Afr., v. 52, p. 119-139.
- Matheron, G., 1975, A simple substitute for conditional expectation: disjunctive kriging, in Guarascio, M., Huijbregts, Ch. J., and David, M., (Eds.), Advanced geostatistics in the mining industry: D. Reidel (Dordrecht-Holland), p. 221-236.
- Rendu, J. M., 1978, An introduction to geostatistical methods of mineral evaluation: S. Afr. Inst. Min. Metall., Monograph Series, 61 p.
- Rendu, J. M., 1979, Kriging, logarithmic kriging and conditional expectation: comparison of theory with actual results: 16th APCOM symposium, Tucson, Arizona, p. 199-212.