

DOES GEOSTATISTICS WORK?

Isobel Clark

Lecturer, Dept. of Mineral Resources Engineering,
Imperial College of Science and Technology,
London SW7 2BP, U.K.

INTRODUCTION

The problem of the estimation of ore reserves is essentially a simple one to describe. Given a set of sample data, at specified locations within a deposit, how can we estimate the value (or quantity) of ore in unsampled areas. The question may be posed of a whole deposit (global) or it may be more desirable to estimate smaller volumes within the planned mine area, e.g. stopes or blocks (local). Many methods have been evolved to tackle this problem, from hand-contouring of values by the mine geologist to extremely sophisticated geomathematical analysis.

The main difficulty with any estimation method in this context is that it is almost impossible to check whether the "predicted" value bears a close relationship with the "actual" value. The only reliable way to decide if an estimation method has "worked" is to test the eventual value of ore against the predicted value. This is, more often than not, impossible - even allowing for the time factor involved. To know the "actual" value of a mined block it has to be sampled intensely, and these values with their positions recorded accurately. Alternatively, blocks may be passed through the concentration process individually and these results recorded. In most producing mines neither of these alternatives are practicable or seen to be important. Another problem with comparing in-situ (so-called) mineable reserves with production figures is the introduction of many variables into the concentration process -- dilution, wastage, possible stockpiling, mill recoveries and so on.

This paper describes a study undertaken to answer the question "Does Geostatistics Work?". To remove ambiguity, Geostatistics here refers to the practical applications of the Theory of Regionalised Variables as developed by Georges Matheron in Fontainebleau, France (cf. Journel & Huijbregts, 1978).

Perhaps it would be best to begin by defining the question with a little more precision. What is really meant by "work"? We have (say) a prediction of grade value for some given volume. This will differ from the "true" value over that volume. An error is incurred by the estimation procedure. The main selling point of Geostatistics is that, whatever estimator you use, the "standard error" of that estimate can be calculated using the semi-variogram. Our question is, then, does the "standard error" produced by geostatistical theory really reflect the accuracy of the estimate produced. If the theory matches up with the practical experience, then we can use geostatistical methods with confidence in other applications. For example, if we know that geostatistics measures the standard error of an estimate with

accuracy, then we can be confident that the estimation technique giving the smallest standard error (Kriging) must be the best one.

This paper attempts to answer the question "Does geostatistics genuinely measure the accuracy of predicted values?".

CASE STUDY

The Mine

Geevor Tin Mines Ltd. in its present form was registered in 1911, although there is evidence that mining in the Geevor lease area extends well back into the eighteenth century. The Geevor and Levant mines are situated at Pendeen in Western Cornwall (U.K.). Three shafts serve the mine, although only Victory Shaft is used for ore haulage.

Ore is extracted from the underground lodes by overhead shrinkage stoping, between horizontal development drives (on lode) approximately 100 feet apart. Drilling and blasting is used to break rock in both stope and development heading, utilising compressed-air rock-drills. Trimming cars and electric locomotives provide transport to Victory Shaft for hoisting. Gravity methods are used in the mill to recover the cassiterite concentrate (65% Sn). In the financial year ending March 1976, 106 thousand tonnes of ore were treated yielding just under 900 tonnes of cassiterite concentrate.

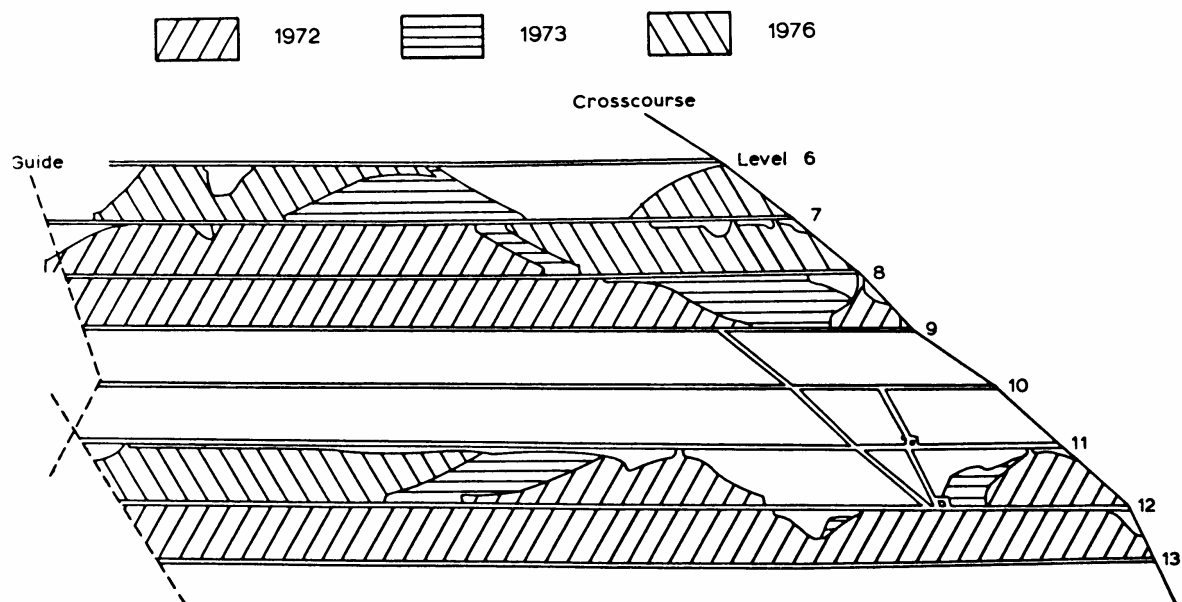


FIGURE 1. Section Diagram of Study Area, Simms Lode, Geevor Tin Mines Ltd.

The Data

The study which is described in this paper was carried out on the Simms Lode of Geevor Mine. The cassiterite vein is almost vertical, and its total developed length (at 1976) was just under 7,000 feet in a north-west/southeast direction. The lode lies about a kilometre from Victory Shaft, and comprises nine developed levels from 600 feet below surface to 1400 feet. The cassiterite vein is divided into several major sections by "guides" -- quartz-filled faults -- and minor branchings. The area selected for this detailed study is in the central portion of the lode, and it has been shown by other studies (geological and statistical) that this area may be considered as one continuous unit. Figure 1 shows a section drawing of the study area. The vein is narrow enough and continuous enough to be treated as a reasonably flat plane in two dimensions, and all sections are drawn on this understanding. The different modes of shading merely denote the years at which the data was added to the continuing study. Although some development was available on level 14 (1400 feet below surface) this has been omitted for simplicity.

The lode is sampled by chipping across the vein (but not the gangue) in the hangingwall of the drive or the stope. The vein averages about two feet wide, and the normal stoping width is forty inches. The bagged sample is returned to surface for vanning assay which produces a value for 'content of recoverable cassiterite'. This is expressed in pounds of cassiterite per ton of ore (lb/ton). The width of the lode is measured in inches on site. Prior to March 1971 development drives were sampled at five foot intervals, but this was changed to ten feet, and in 1976 to three metres. In the study area all development except on level 6 and for minor westward extensions on 10, 11 and 12 was completed prior to March 1971. Thus, virtually all drive sampling is available at five foot intervals. Stopes are sampled, usually, at fortnightly intervals. This produces a very approximate grid of samples within the stoped area. Figure 2 shows an enlarged subsection of levels 12 and 13. Each cross denotes a chip

sample. Note how closely the drive is sampled compared with the stope.

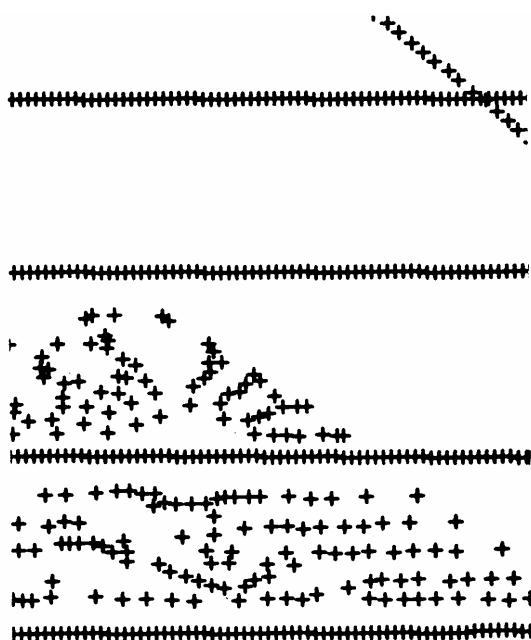


FIGURE 2. Example of sample positions in stope and development.

The sections were originally hand-drawn in the Survey Office at Geevor. For this study each sample was assigned two co-ordinates relative to an arbitrary origin in order to position it in the two-dimensional plane. Figure 2 is actually a computer plot of the sub-area shown. In all there are about 2,700 stope samples and 1,900 drive samples in the study area. It must be assumed that all samples are comparable -- i.e. have the same 'support'.

It is common practice to value vein deposits such as this using the "accumulations" of the samples. That is, the assay value is multiplied by the width. This has not been done in the study described. The variable discussed throughout this paper is "grade" or assay value. Justification for this decision may be

found in Clark (1979). This reference also contains a complete listing of the data used in this investigation.

SEMI-VARIOGRAMS

The semi-variogram is an essential tool in any geostatistical analysis. It provides a graphical and numerical measure of the "continuity" of the mineral values within the deposit. In a normal reserve estimation only development samples would be available to construct the "experimental" semivariograms. This reduces the calculation to a very simple one, since the sampling is done at regular intervals along each drive, either at five or ten foot, except for the occasional cross-cut from the shaft. Experimental values can be calculated for any multiple of five feet. A decision has to be made on how far to carry this calculation. For various reasons, the experimental semi-variograms were produced for distances up to 250 feet on each level in the lode separately. At first glance there seems to be little or no correlation between the assay values along the drives, and vast differences between levels in the lode.

An immediate question arises as to whether these differences between the levels are simply due to "sampling variation" or whether they are due to some inherent structure in the lode. Some authorities (David, Guarascio) have suggested that such differences may be due to the samples following a log-normal type of distribution. In the log-normal distribution the standard deviation of a set of samples is directly proportional to the arithmetic mean of those samples. Thus if the "local mean" shifts slightly, the standard deviation will also shift accordingly. Since the sill of a semivariogram (if it has one) is equal to the variance of the sample values, a shift in the 'local mean' will have an even greater effect on the semivariogram sill.

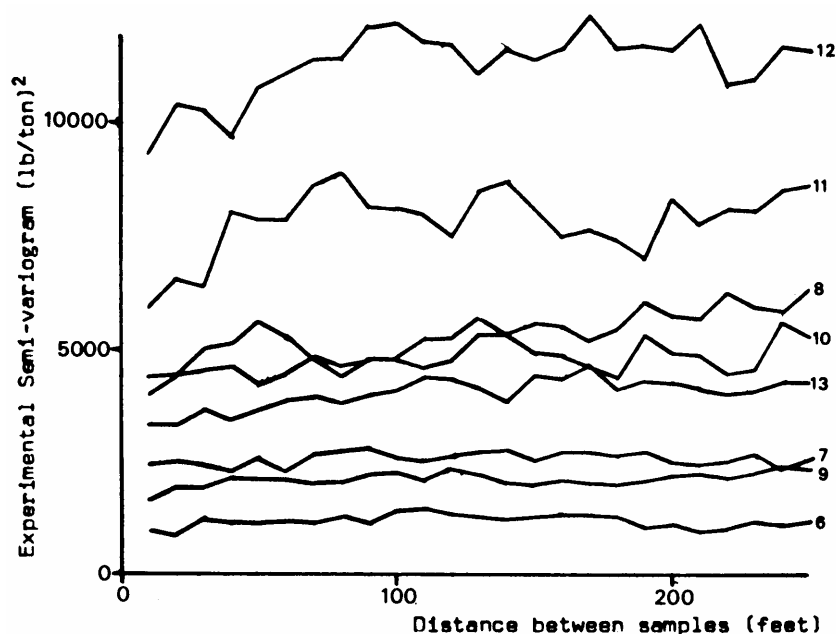


FIGURE 3. Experimental semi-variograms by level.

Table 1: Sample means and standard deviations by level.

level	arithmetic mean	standard deviation
6	34	34
7	40	51
8	56	78
9	46	48
10	57	73
11	77	89
12	79	105
13	53	63

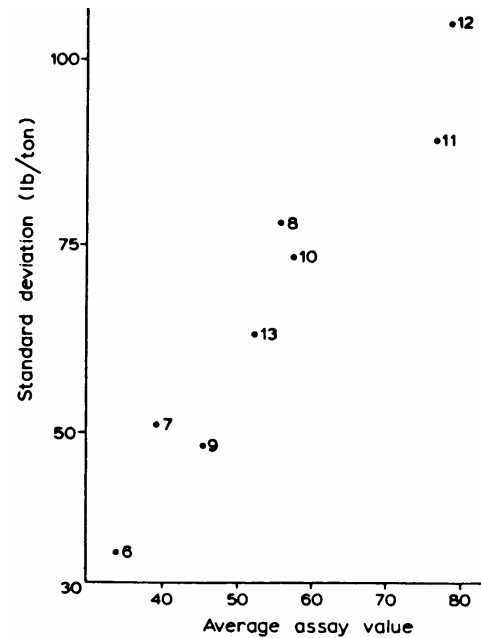


FIGURE 4. Graph of sample mean versus standard deviation for each level in the lode.

Table 1 shows the mean and standard deviation of the sample grades for each level separately. It can easily be seen that the standard deviations rank in the same order as the experimental semi-variograms in Figure 3. A graph of mean versus standard deviation illustrates the proportional effect very well. The correlation coefficient is, in fact, 0.95.

It would seem from this evidence that the difference between the level semi-variograms is due to the proportional effect - and presumably the log-normality of the grade distributions. In actual fact, the samples are not log-normally distributed but may be characterised by a mixture of two log-normals - so we must presume that proportionality also holds under this circumstance.

According to the authorities (op cit) the next stage in the variogram analysis is to "standardise" each level semi-variogram by the square of the local mean. That is, each calculated semi-variogram is divided by the square of the arithmetic mean of the samples which were used in its calculation. This was done, and the resulting semi-variograms are shown in Figure 5. The picture is now much more homogeneous. The highest and lowest semivariograms now differ only by a factor of two, whereas in the "unstandardised" graph the highest and lowest differed by a factor of ten. Since the levels now appear to have comparable structures we can combine them to form one overall experimental semi-variogram for the whole section, and to fit a model for use in the later estimation procedures.

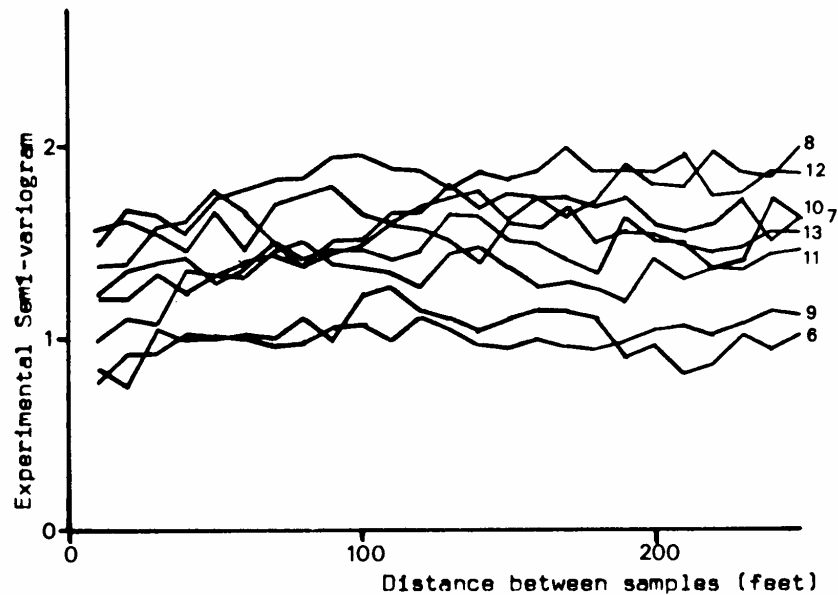


FIGURE 5. Experimental semi-variograms by level
- standardised by local mean

The overall "standardised" experimental semi-variogram is shown as the dots in Figure 6. The dashed line is the compound value of the variance divided by the square of the arithmetic mean. This would be a first estimate at the value of the sill of the model (if it exists). Attempts to fit one of the simple models for semi-variograms, i.e. linear, spherical, de Wijsian etc. fail completely.

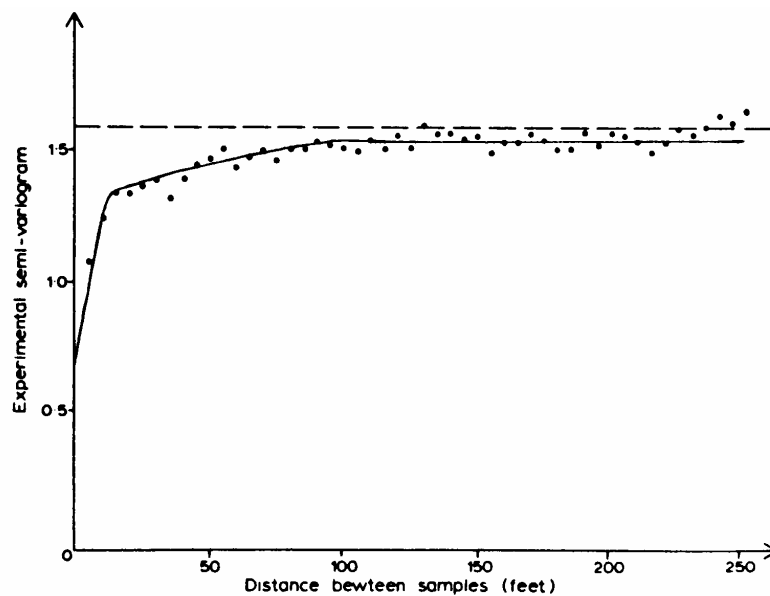


FIGURE 6. Overall "standardised" semi-variogram

It is clear that a large nugget effect is present, at a level of about 0.70. Note that this semi-variogram is dimensionless, since all the values have been divided by the square of the mean grade. Thus this nugget variance must be read as "0.70 multiplied by the square of the local mean". The sill clearly exists and takes the value 1.55. However, the curve between the two (intercept and sill) cannot be approximated by a spherical model. It is possible to fit a

model which is a mixture of two spherical models. A fitted model is shown as the solid line in Figure 6. It has two spherical components with ranges of 15 feet and 125 feet respectively, and with sills of 0.60 and 0.25. This implies that samples within 125 feet of one another are correlated, but to a fairly weak extent. Samples within 15 feet are much more highly correlated, but about 45% of the total variation in the sample values seems to be due purely to random behaviour - perhaps a reflection of the type of mineralisation. This tends to imply that no matter how many samples are taken, there will always be a large unpredictable component in whatever we try to estimate.

The semi-variogram model to be used in the estimation procedure will be:

for distances (h) less than 15 feet

$$\gamma(h) = 0.70 + 0.60 \left\{ \frac{3h}{30} - \frac{h^3}{2(15)^3} \right\} + 0.25 \left\{ \frac{3h}{250} - \frac{h^3}{2(125)^3} \right\}$$

for distances between 15 and 125 feet

$$\gamma(h) = 1.3 + 0.25 \left\{ \frac{3h}{250} - \frac{h^3}{2(125)^3} \right\}$$

and for distances greater than 125 ft.

$$\gamma(h) = 1.55$$

TESTING THE ESTIMATION PROCEDURE

Having produced a model for the semi-variogram of this section of the lode, we can now embark upon an estimation procedure. Suppose we wish to estimate a volume which possesses an actual (unknown) grade, say T.

Suppose we use a linear estimator, that is one which is a simple linear combination of the grades of the surrounding samples. We may denote this by T^* . In the prediction an error will be made which will be equal to $T - T^*$. Using the semi-variogram we can produce a measure of the reliability of the estimator in the form of the "estimation variance", σ_e^2 .

Since we are using a "standardised" semivariogram, the estimation variance produced will be "relative to the local mean". That is, the calculated variance must be multiplied by the square of the local mean before it becomes a true measure of the spread of the errors. The calculated variance can be denoted by σ_i^2 , so that:

$$\sigma_e = \sigma_i \times \text{"local mean"}$$

The only problem now is to define what exactly is meant by the local mean. The published authorities are not very explicit on this point, but presumably this should be interpreted as the "expected value at this point". This must mean the actual value at that point (or over that volume), i.e. T. Thus, the standard error of the estimate T^* should be given by

$$\sigma_e = \sigma_1 \times T$$

In practice, of course, we would not know the value of T so that the standard error of T* would be estimated by

$$\sigma_e = \sigma_1 \times T^*$$

To test the estimation three quantities are needed — T, T* and σ_e . The estimator T* will (according to the theory) be drawn from some kind of probability distribution with mean T* and standard deviation σ_e . That is, for any estimation we have one sample from such a distribution. Unless it is possible to carry out several estimations which result in the same standard error, it is difficult to test whether the procedure is "working" or not. Suppose a new variable is proposed:

$$z = \frac{T - T^*}{\sigma_e}$$

No matter what kind of distribution T* follows, the z variable will have a distribution with a zero mean and a standard deviation of 1. Now, if several estimations are carried out, each will yield a z value. If the theory and method of application are correct then all of these z values will come from a distribution with mean zero and standard deviation 1. This produces a situation which can easily be tested -- carry out several estimations. calculate the z values and see if they have a mean of zero and a standard deviation of 1.

The Case Study

In the study area under consideration the only "actual" values which we know with reasonable certainty are the measured values at the sample sites within the stopes and development drives. Since the drive samples were used to calculate the experimental semi-variograms, they will be excluded from the testing of the estimation technique. This should ensure an objective and unbiased test, and leaves upwards of 2,700 "known" T values to use in the testing. The next problem is to decide on which estimator to use. Any linear estimator may be chosen, and in this case the choice was the Kriging estimator. The procedure followed was:

- 1) Consider each stope sample in turn;
- 2) Find all other stope samples within a given distance (δx) in the horizontal direction, and within the same stope;
- 3) Use these stope samples to produce a Kriging estimate and Kriging variance for the "point" under considerations
- 4) Correct the Kriging variance for the local mean;
- 5) Calculate the corresponding z value;
- 6) Include this z value in the calculation of mean and standard deviation, and in the histogram.

In the computer program written for this analysis the histogram was centred on zero and scaled so that the total range ran from -3 to $+3$. The individual z values were not stored at all. A decision has to be made on the size of δx . The larger it is, the longer the program takes to run. Mini-tests using level 13 only (726 samples) produce virtually the same output for $6x$ ranging between 25 feet and 65 feet. Distances over 65 feet were considered impracticable -- especially since the dominant range of influence is only 15 feet. The final results given in this paper were all calculated for $6x$ equal to 25 feet. In correcting the Kriging variance for the local mean, the "true" value of the stope sample was used, so that

$$z = \frac{T - T^*}{\sigma_1 T}$$

for each stope sample.

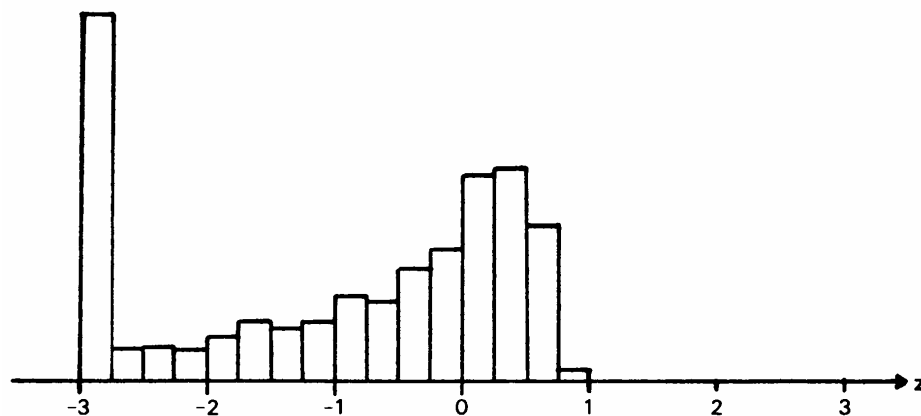


FIGURE 7. Histogram of transformed errors using the "standardised" semi-variogram

Table 2: Parameters of transformed errors by level.

level	arithmetic mean	standard deviation
7	-1.28	3.09
8	-1.86	2.94
9	-1.29	2.10
12	-1.76	3.31
13	-1.63	3.14

The histogram of the resulting z values is shown as Figure 7. Negative values of z imply that T' has over-estimated T , and positive values imply under-estimation. For clarity, all samples below -2.75 have been included in the group -3 to -2.75 . Similarly, all samples greater than $+3$ are included in the group $+3$ to $+3.25$. The mean of the z values is -1.60 and the standard deviation is 2.96 . Table 2 shows the means and standard deviations by level so that it may be seen that these values are consistent over the whole lode.

The only conclusion which we can draw from this histogram and these parameters is that the Kriging variances do not reflect the true accuracy of the estimation procedure. A standard deviation of 3 for the z values implies that (in general) the supposed Kriging standard

deviations are too small by a factor of 3. That is, the Kriging variances are a factor of 9 too small. Closer inspection of the histogram allows us to break this conclusion down into more detail. The computer program does not record the extreme values of z , so we do not know how long the negative tail is. However, we can see that the positive side of the graph is truncated at a value of +1. This implies that for samples which have been underestimated the Kriging variance given by theory is far too large. For samples which have been over-estimated the supposed variance is far too small. In other words, we will be vastly overoptimistic about over-estimated samples and rather pessimistic about underestimated ones. This is not only undesirable but could prove disastrous!

A Retrial

In practice we would not know the value of T which we were trying to estimate, but would only know T^* . Suppose in (4) above the Kriging variance was corrected by the estimate of the local mean, i.e. T^* , so that

$$z = \frac{T - T^*}{\sigma_1 T^*}$$

The above procedure was repeated on this basis, and the resulting histogram is shown in Figure 8. The overall mean of the z values is 0.09 and the standard deviation is 1.45. Table 3 gives the level by level parameters, again showing that these parameters are consistent throughout.

Table 3: Parameters of 'estimated' transformed errors by level.

level	arithmetic mean	standard deviation
7	0.22	2.46
8	0.11	1.47
9	0.08	1.29
12	0.06	1.28
13	0.06	1.17

The picture is a lot more encouraging than the previous one. The mean 'transformed' error is very close to zero, although consistently positive on all levels. The standard deviation is still somewhat larger than the expected 1.00. There is also a truncation, but this time at the lower end, and there is still rather a long tail -- but not nearly so remarkable as that in Figure 7. We now seem to be over-optimistic about the accuracy of under-estimated samples and vice-versa. On the whole the Kriging standard deviations seem to be under-estimated by about 45%.

This latter analysis taken in isolation would seem to support the "working" of the application of geostatistics. However, it is somewhat disturbing when taken in context with the previous analysis. If the theory is correct then surely the theoretical analysis - with T known - should give a much better picture than the practical analysis - with T unknown.

The case study so far has resulted in entirely the opposite situation. Should we therefore conclude that geostatistics does not work as according to theory?

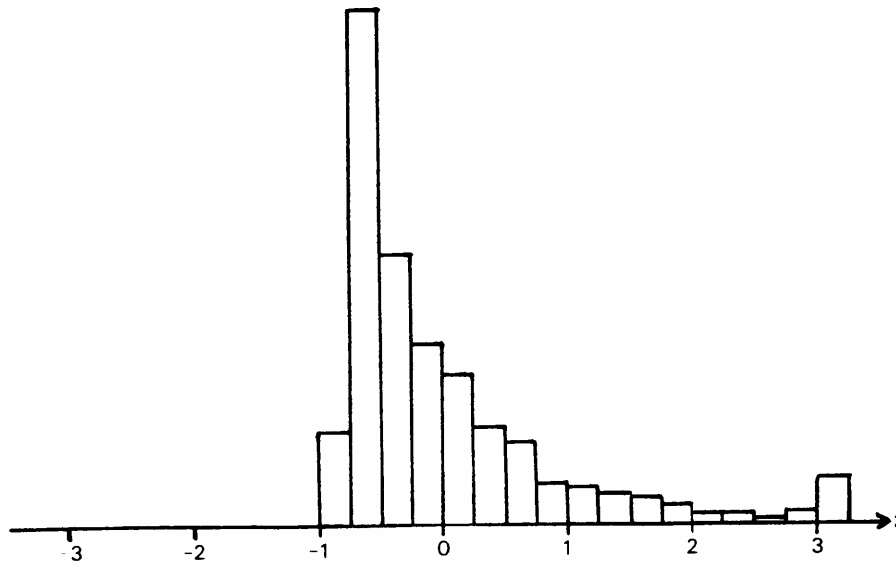


FIGURE 8. Histogram of 'estimated' transformed errors using the "standardised" semi-variogram.

AN ALTERNATIVE APPROACH

In the analyses above we have followed the stages of the accepted form of geostatistical analysis. Having calculated "local" experimental semivariograms and demonstrated that the "proportional effect" was present, we standardised each level by the local mean, combined the individual levels and emerged with an overall experimental standardised semi-variogram. To this we fitted a model which was 'relative to the local mean'. We then followed the usual Kriging procedure to estimate each stope sample in turn and compared the errors in prediction by means of the z transformation. The conclusion reached was that if we take the actual value of the stope sample (T) as the "local mean", the resulting estimation variances bear very little relationship to reality. If we take the estimated value (T^*) in place of the local mean the resulting estimation variances give an acceptable picture of the reliability of the errors, although the standard deviations are under-estimated by about 45% and the shape of the histogram is decidedly odd.

Suppose we take a rather different approach and return to the original 'raw' experimental semi-variograms. Each level is calculated from several hundred inter-related sample values and we have illustrated that in each level the sill is directly proportional to the square of the sample mean. Suppose an alternative hypothesis is put forward. Suppose that the variation of mean and standard deviation of the sample grades between each level is due simply to them being different sets of samples from the same log-normal type distribution. Under this hypothesis the semi-variograms are not standardised but combined to form one overall semi-variogram -- this might be referred to as the 'raw' experimental semi-variogram. This combined semivariogram is shown as the dots in Figure 9, with the overall sample variance drawn in as the dashed line.

This graph exhibits a strange zigzag behaviour which is presumably the proportional effect differentiating between those levels with five foot sampling and those with ten foot. A sill appears to be present at about 5,675 (lb/ton)², and a nugget variance of 2,575 (lb/ton)². This nugget effect represents 45% of the total sill value - exactly the same proportion as in the standardised model. Bearing this in mind, a model was proposed with two ranges of 15 and 125 feet, and with sills of 2,200 (39% of total sill) and 900 (16%) respectively. That is, exactly the same model as before but with a total sill of 5,675 (lb/ton)² instead of 1.55. This is shown as the solid line in Figure 9. and appears to be a perfectly adequate fit. This, perhaps, supports the alternative hypothesis, since the fitting of essentially the same model would tend to suggest that 'standardisation' is superfluous. However, it can be seen that deriving such a complex model from the zigzag semi-variogram in Figure 9 would have been rather more difficult if we had not had the standardised model to refer to.

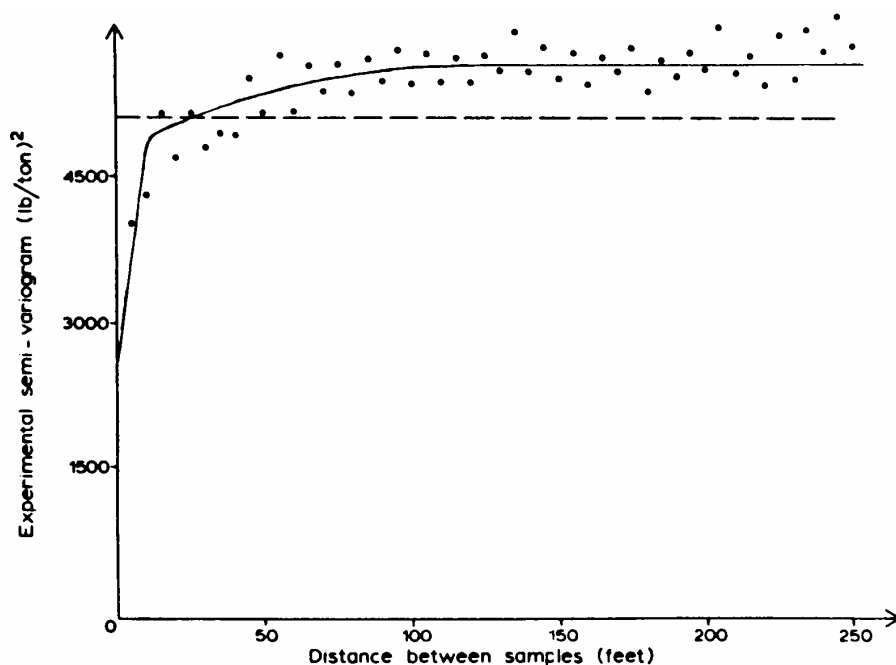


FIGURE 9. Overall "raw" experimental semi-variogram

Table 4: Parameters of errors by level

level	arithmetic mean	standard deviation
7	0.01	0.60
8	0.00	1.10
9	0.00	0.63
12	0.00	0.84
13	0.00	0.94

Since the models are the same except for a multiplying factor, the estimates produced will be identical. The only difference will be in the estimation variances. Whereas the standardised model gave standard errors directly proportional to the value being estimated, the "raw" model gives standard errors which are independent of the size of either true value or estimator. In other words, the former approach assumed that low grades could be estimated more accurately than high ones. The present approach assumes that the accuracy of the estimation depends only on the sample configuration -- not the sample grades.

The six-point analysis described previously was respected with the "raw" model, omitting step (4). The histogram of transformed z values is shown in Figure 10, and possesses a mean of 0.00 and an overall standard deviation of 0.76. Table 4 shows the parameters by level. In this case the mean error at all levels is zero, but the standard deviations vary between 0.60 and 1.10 in no recognisable pattern. This is an encouraging sign since it suggests that the variation is stochastic and not due to some consistent bias. The histogram of the transformed errors is much more intuitively pleasing than either of its predecessors. Whilst still slightly skewed, there are no truncation points or excessive concentration of values. This evidence leads us to conclude that in this case the standardising of the experimental semi-variograms has led us into an erroneous picture of the standard errors of the estimation procedure.

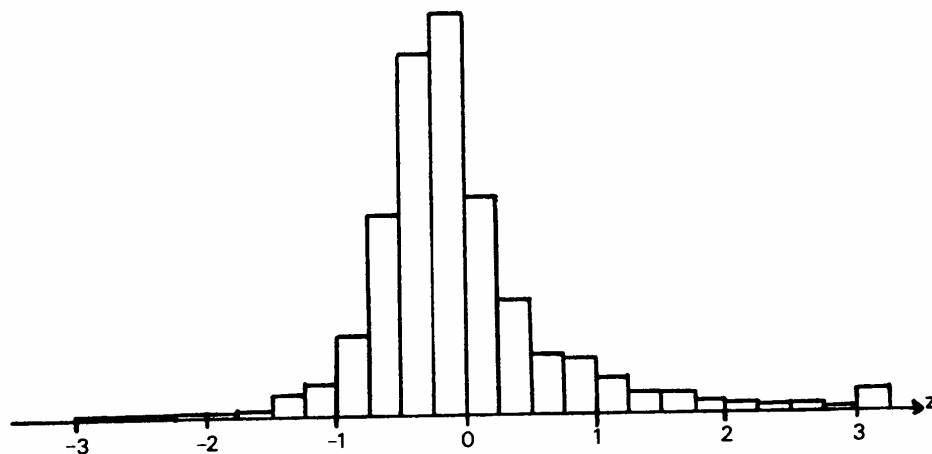


FIGURE 10. Histogram of transformed errors using the "raw" semi-variogram

CONFIDENCE INTERVALS

It is common in quoting standard errors for estimates to attempt to convert these into confidence intervals. The main difficulty in this process is in what distribution to assume for the errors. Most workers happily assume that their prediction errors follow an approximately Normal distribution, so that a 95% Confidence Interval for the estimate would be given by:

$$\tau^* \pm 2\sigma_e$$

An empirical 95% Confidence Interval can be found from the histogram of the transformed errors in our case study. This is done by taking the centre 95% of the histogram, and using

the z values at either end of this interval instead of the ± 2 above. Taking the three approaches in turn produces:

- ◆ Standard errors calculated by using the standardised semi-variogram multiplied by the 'true' mean (T) produces an empirical 95% interval of

$$\begin{aligned} T^* &= 5.3\sigma_e \\ \text{to} \\ T^* &+ 0.70\sigma_e \end{aligned}$$

- ◆ Multiplying the standard error by the estimated mean T' produces an interval of

$$\begin{aligned} T^* &= 0.89\sigma_e \\ \text{to} \\ T^* &+ 3.12\sigma_e \end{aligned}$$

- ◆ Using the 'raw' semi-variogram gives an empirical 95% confidence interval

$$\begin{aligned} T^* &= 1.19\sigma_e \\ \text{to} \\ T^* &+ 2.33\sigma_e \end{aligned}$$

This lends weight to the decision that the 'raw' semi-variogram leads to a better evaluation of the actual errors incurred in the estimation. Assuming a 95% Confidence Interval of $\pm 2\sigma_e$ in either of the standardised cases could lead to disastrous consequences, whereas in the "raw" case it will give a fairly accurate representation of the actual Confidence Interval.

CONCLUSION

The question posed by this paper was "Does Geostatistics Work?". We have considered one case study only and the conclusion which must be drawn is "Yes provided the so-called proportional effect is ignored". If the "raw" semi-variogram is used, then the standard error predicted by geostatistical theory accurately measures the reliability of the estimation method. We have also seen that the usual assumption of the Central Limit Theorem -- i.e. that a 95% Confidence Interval may be produced by taking $\pm 2\sigma_e$ -- is also fairly safe if the "raw" semivariogram is utilised.

ACKNOWLEDGEMENTS

The author wishes to thank the management of Geevor Tin Mines Ltd. for the use of the Simms Lode data in this and the rest of the research for the degree of Ph.D obtained in January 1979. The computer programs were developed and implemented on the CDC Cyber 175 at Imperial College Computer Centre.

REFERENCES

Clark, I "Geostatistical Reserve Estimation in Cornish Tin Mining", Ph.D Thesis, University of London, 1979

David, M Geostatistical Ore Reserve Estimation, Elsevier, Amsterdam. 1977

Guarascio, M. "Improving the uranium deposit estimations, the Novazza case" Advanced Geostatistics in the Mining Industry, Reidel, Netherlands, 1976

Journel, A.G. & Huijbregts, Ch. Mining Geostatistics, Academic Press, 1978 .