



Witwatersrand gold reef evaluation: the ‘variancegram’ tool

by C. Lemmer* and M. Mogilnicki†

Synopsis

The resolution with which the different categories of resources are forecast for Witwatersrand gold reefs should ideally tie in with block sizes that are optimal in terms of the variability structures of the reefs. A tool, called the ‘variancegram’, is proposed as a basis for block size choice. A variancegram is intrinsic to the particular reef and mine concerned. A further requirement is the ability to attach global confidence limits to weighted average estimates built up from combinations of local kriged estimates. Approximations to derive global kriging variances based on variables derived from local kriging deliver hugely inflated results if ordinary kriging is used, and markedly better, but not accurate, values if simple kriging is used. These approximations improve as the number of samples used in kriging each block is increased. It is shown that the behaviour of the different components of the global kriging variance with increasing number of samples, all differs, but they all link to the variancegram for the particular reef. The variancegram can thus be used to correct the different components to the values they would have had if all samples were used in kriging each block, and so deliver the ‘correct’ global kriging variance, even though only a limited number of samples were used in kriging each block. The desirability of having very stable solutions implemented in production systems is taken into account in the proposals. It is anticipated that the same variancegram findings will also hold for other densely sampled deposits, but this remains to be investigated.

Keywords

resources, classification, global kriging variance.

Introduction

The Witwatersrand gold reefs in South Africa are unique in terms of their continuity despite the relatively narrow reef widths. This aspect, together with the wealth of historical channel sample data captured, affords the rare opportunity of studying gold accumulation behaviour over square kilometres in extent. Such studies have in turn provided the backdrop against which large-scale forecasts can be attempted for a phenomenon as inherently variable as gold, and at such extreme mining depths. Due to the latter circumstance, forecasting of values ahead of a mining front is based almost exclusively on extrapolation.

Although the channel sample data itself exhibit very high micro-scale variability (which all but masks any spatial continuity), it is well-known in the industry that, by

averaging these data values into blocks of ever-increasing size, one can uncover spatial structures (variograms) with ever-increasing ranges to underpin the extrapolation. This practice serves the needs of the industry: resources within a short range of the just-mined blocks constitute the ‘annual forecast’ or Measured Resource for the coming year, and need to be estimated with relatively high resolution; resources beyond these constitute ‘medium term’ or Indicated Resources that may be mined within the following five years, and the resolution requirements for these resources are less stringent; finally, the remaining resources constitute the ‘long-term’ or ‘life-of-mine’ or Inferred Resources, and the resolution required for these last estimates is relatively crude.

In choosing the block sizes on which the averaging process will be based, one can be purely practical and choose the block dimension such that the resulting range will cover the area to be estimated. However, one is dealing with a geological phenomenon here, and there is an instinctive feeling that optimum block sizes – in terms of best elucidating the underlying spatial structures – will be peculiar to the reef or deposit studied.

A tool called a ‘variancegram’ is proposed for determining the underlying optimum block sizes. There is a small latitude in the final determination of the sizes, allowing the very desirable feature of having the block sizes as integer multiples of each other and thus multiples of the smallest block size chosen. This feature forms a cornerstone for the development of an overall gold resource system, and the justification for considering the derived block sizes as ‘optimal’ will also be discussed.

* *Geological and Geostatistical Services, Johannesburg, South Africa.*

† *Asfaltowa, Warsaw, Poland.*

© *The Southern African Institute of Mining and Metallurgy, 2014. ISSN 2225-6253.*



Witwatersrand gold reef evaluation: the 'variancegram' tool

Having forecast (kriged) the resources remaining to be mined with different degrees of resolution, the associated practical problem of the need for confidence limits to attach to average values built up from combinations of these varying supports is posed. In essence, this boils down to the intractable problem of the inability to derive a global kriging variance by a simple combination of local kriging variances. A first approximation to the problem, which has been mooted in the literature, is discussed.

The quality of the above-mentioned approximation is very dependent on the number of samples used to krig each of the local estimates. Theoretically, if all available samples were used to krig each local estimate, the resultant global kriging variance should be good. However, for most of the practical numbers of samples routinely used in such kriging, the approximation delivers a hugely inflated global kriging variance if ordinary kriging (OK) was done, and a markedly better, but not accurate, value if simple kriging (SK) was carried out.

The interesting finding is that, with increasing numbers of samples, the behaviour patterns of the different components of the global kriging variance, for OK or SK, differ; but they all link to the behaviour of the variancegram for the particular gold reef under consideration.

A method for correcting the above-mentioned inaccurate global kriging variances is proposed, based on the variancegram.

It is anticipated that the same variancegram findings will also hold for other densely sampled deposits, but this remains to be investigated.

The variancegram

Early work of Krige, published during the 1950s and summarized later (Krige, 1978), found that 'the logarithmic variance of gold values shows a rising linear trend when plotted against the logarithms of the population areas concerned. This remarkable finding is, of course, true in theory and it would certainly be true for infinite deposits, but in practice one finds that the variance of the logarithms of the gold accumulation values, or logvariance for brevity, does not steadily increase without limit, but that it starts to deflect downwards towards an eventual 'sill'.

For the purposes of determining optimum block sizes – to use in the routine evaluation of gold reefs – it was decided to investigate the above deflection points, with the view that they indicate points at which the logvariance decreases abruptly due to correlations between sample values that set in for areas beyond that particular size. The argument is that by averaging point values within an area indicated by, for instance, the first deflection point, one would absorb the first-structure variability within the block averages, and a variogram in terms of these averages would then elucidate longer-range structures.

Channel sample data (gold accumulation values) for a typical reef was available over several square kilometres in area (the 'test data'). The logvariance of this 'point' test data was calculated within grid cells of side 7 m, 8 m, etc., up to cells of dimension 2000 m, and finally 5000 m. For each choice of cell size the logvariance for that cell size, averaged over the entire mine, was determined.

Figure 1 shows the graph resulting when the average logvariance per cell is plotted against the logarithm of the area of the cell.

As is clear from this graph, it was found that the series of deflections does not constitute sharp, *unique* bends. On the contrary, by drawing tangential lines to the graph, several sets of deflection points can be obtained.

A solution to this problem was found by plotting the average logvariance within square cells against the length of the side of the square. Figure 2 shows the resultant graph for the reef under consideration.

In contrast with Figure 1 this plot uses the effective block (cell) side length $l = \sqrt{\text{Area}}$ to characterize the area under consideration. Such a construction will, for the purpose of this discussion, be called a 'variancegram' to distinguish it from the more familiar variogram of classical geostatistics. The distinction is all the more important since (i) the plot has the appearance of a variogram (which it is not), and (ii) in describing its main features one can draw on previous experience of fitting ordinary variograms. Point (ii) will be discussed in more detail below.

One notices from Figure 2 that the *rate* at which the logvariance varies with $l = \sqrt{A}$ changes from a rapid to a much slower one as the side of the block involved is

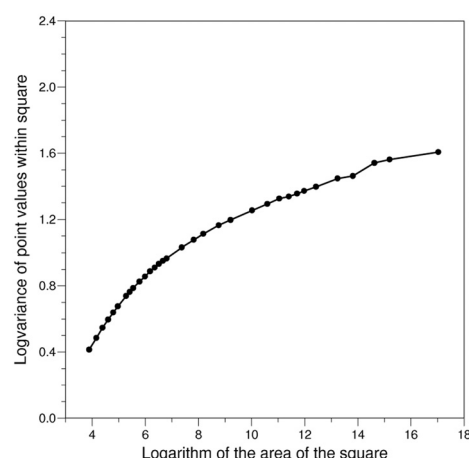


Figure 1 – Logvariance versus log of area of square

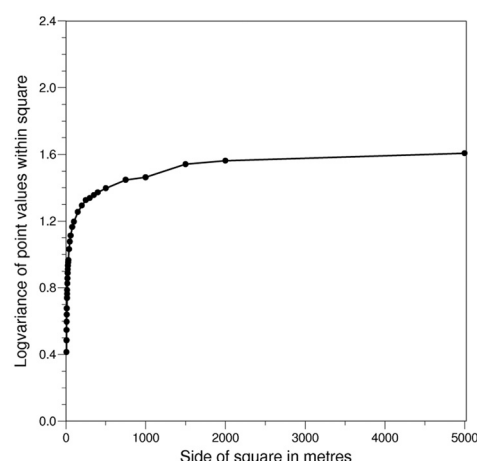


Figure 2 – Logvariance versus side of square

Witwatersrand gold reef evaluation: the ‘variancegram’ tool

increased. If the initial rapid variation in the logvariance is interpreted as a reflection of the microstructure variability in the deposit, then this rapid increase in logvariance with block side would be expected to continue without limit as l increases, if this variability persists at all scales. Obviously this does not happen in the case of the test data. Instead, the logvariance increases less and less rapidly as l is increased, and dramatically so.

This empirical finding is to be interpreted as reflecting the presence of a set of superimposed micro-to-macro variability structures in the deposit, each exhibiting a characteristic range a over which it is operative, and beyond which it levels off to a constant value. For block sizes of dimension larger than a , it would lead to a smaller logvariance than would otherwise have been found had the particular structure continued to increase at the same rate for all distances.

The preceding qualitative observation can be made quantitative by drawing on the understanding of the meaning of the spherical model in fitting *variograms*. As is well-known, the spherical model formula is equivalent to surrounding each point grade by a ‘sphere of influence’ with radius equal to one-half the variogram range a , and arguing that the covariance of two grades at positions separated by $h < a$ is proportional to the overlap volume of their respective spheres of influence. This assumption immediately gives

$$C(h) = C(0)\left[1 - \frac{3}{2}\left(\frac{h}{a}\right) + \frac{1}{2}\left(\frac{h}{a}\right)^3\right], \quad h \leq a \quad [1]$$

$$= 0, \quad h > a$$

for the covariance, from which the standard spherical variogram formula follows,

$$\gamma(h) = C(0)\left[\frac{3}{2}\left(\frac{h}{a}\right) - \frac{1}{2}\left(\frac{h}{a}\right)^3\right], \quad h \leq a \quad [2]$$

$$= C(0), \quad h > a$$

i.e. the variogram is proportional to the non-overlapping volume, and when the two spheres are separated by a distance of more than a , the non-overlapping volume has reached a constant value.

These ideas can now be taken over to provide the basis for a *quantitative* method for describing the variancegram as well. For, by analogy with the variogram case, one can argue that the variability associated with a structure of range a in the deposit increases with square side l within a ‘sphere of influence’ of radius a , beyond which the variability levels off once l exceeds a . Then it immediately follows that the logvariance (due to all structures present), as a function of increasing square size, will gradually start to saturate as it approaches a final value, instead of increasing indefinitely. Calling the variancegram $\Gamma(l)$ to carefully distinguish it from the usual variogram, one can write

$$\Gamma_i(l) = C_i\left[\frac{3}{2}\left(\frac{l}{a_i}\right) - \frac{1}{2}\left(\frac{l}{a_i}\right)^3\right], \quad l \leq a_i \quad [3]$$

$$= C_i, \quad l > a_i$$

for the influence of the i th structure in the variancegram. The variancegram itself is then constructed as follows:

$$\Gamma(l) = \sum_{i=1}^p \Gamma_i(l) \quad [4]$$

It is to be emphasized again that this is a *parametrization* of the variancegram that has been motivated by ideas suggested by the meaning of correlation in the case of the

ordinary variogram. The big advantage of the approach is that one draws on existing expertise and programs to fit the curve.

In fitting the test data, it turned out that the spherical model gives an extremely good representation of the empirical variancegram. Figure 3 shows the experimental data points, with the model fitted.

The parameters of the fitted model are the following (using the standard spherical model notation):

$$C_0 = 0$$

$$C_1 = 0.710, \quad a_1 = 20 \text{ m}$$

$$C_2 = 0.280, \quad a_2 = 60 \text{ m}$$

$$C_3 = 0.190, \quad a_3 = 180 \text{ m}$$

$$C_4 = 0.118, \quad a_4 = 540 \text{ m}$$

$$C_5 = 0.205, \quad a_5 = 1620 \text{ m}$$

$$C_6 = 0.097, \quad a_6 = 4860 \text{ m}$$

The advantage of fitting a nested spherical model is that it fits a continuous curve, even though there are six structures. There is thus no need to try to divine successive bends in the experimental curve.

In establishing a good fit, there is a small amount of latitude in choosing the ‘sills’ C and the ranges a . However, the systematic increase in the ranges is no doubt an intrinsic property of the variability structures of the deposit, and arbitrarily chosen ranges will not deliver a fit.

As is clear from the above parameters, the small amount of latitude that does exist is used to stagger the ranges in such a way that they form integer multiples of each other. The order of magnitude of this latitude is such that one could have fitted any range between, say, 56 m and 64 m, instead of 60 m, and still have obtained a reasonable fit. In this particular case of the test data the scheme – of multiplying each range by three to obtain the next one – worked well. A similar set of ranges a , but not sills C , were observed to fit the variancegrams of the same Witwatersrand reef on different mines.

The fact that the spatial variability of the phenomenon can so successfully be decomposed in individual structures with finite ranges provides experimental justification for using these ranges as the optimum block sizes on which to base the different resolution levels of the resource

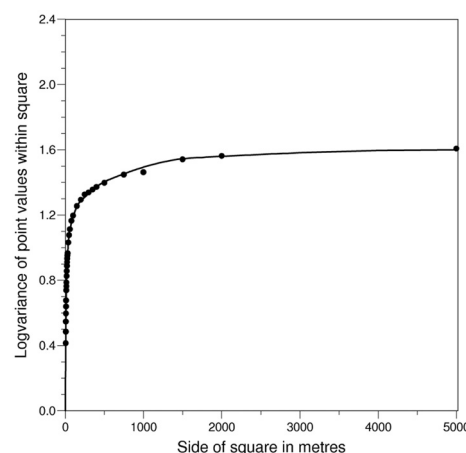


Figure 3—Model fitted to experimental variancegram points

Witwatersrand gold reef evaluation: the ‘variancegram’ tool

calculations. It certainly provides a natural, coherent framework for an integrated system of increasing block sizes, as will become clear.

In the case of the test data, the block sizes implicated above turn out to be very convenient and appropriate. The 20 m block size has been used as the basis for annual resource forecasts; the 60 m block size provides variogram structures that are very appropriate for medium-term forecasts, and the 180 m block size could be useful for life-of-mine resources.

A note in passing: the variancegram was fitted here up to a range of about 5000 m, which corresponds to the limit of experimental calculations. We will demonstrate that the variancegram can be used as a tool to correct the global kriging variance approximations, as stated above.

What is very important, though, is that for the latter purpose the variancegram has to be fitted over a very specific distance. Such a procedure does alter the parameters of the fit somewhat, i.e. they are not the same parameters taken over a shorter distance, as will be shown.

The problem of calculating global kriging variances

Having derived optimal block sizes on which to base forecasts for the different categories of resources, the kriging of these resources is straightforward. Additional considerations will determine whether OK or SK is used.

The main theme of the approach is thus not to use the original channel samples values as data to forecast different size blocks ahead, but to first *average* the point data into different size blocks to create data of different support. The new support now has a longer, more useful, range and, since the blocks to be forecast have the same support, point kriging is performed. The result is a resource forecast defined on blocks of varying support, where the blocks are integer multiples of each other in area. Naturally, each block has an associated local kriging variance as measure of the confidence that can be placed in its estimate.

The practical problem that formed the main thrust of this investigation stemmed from the requirement to be able to place confidence limits on the weighted average value calculated for a number of block estimates of varying support. These blocks could be specified by a single polygonal boundary that encloses them, or by several polygonal boundaries that are disjunct.

In the discussion that follows, the assumption is made in the first place that the forecast blocks are of different support, but that the data used in the forecast is all of the same support; say, the smallest data block used in the forecasts. The generalization to varying data supports will be made subsequently.

Assume that the total area(s) enclosed within the specified polygon(s) is denoted by V , and the supports of the M enclosed local estimates are referred to as ‘blocks’ v_j . A kriged estimate Z_j^* is available for each of these blocks. Then the ‘global’ kriged estimate that pertains to the entire area V is given by the weighted average (see for example, Journel and Huijbregts, 1978) of the Z_j , i.e.

$$Z^* = \sum_{j=1}^M \frac{v_j}{V} Z_j^* \quad [5]$$

The proviso on this formula is that the *same set of all n data samples* be used to krig the Z_j . Similar relations then

hold for the ‘global’ weight λ_i pertaining to the contribution of the i th data sample to the estimate Z^* , and also to the Lagrange multiplier μ in the event that ordinary kriging is employed:

$$\lambda_i = \sum_{j=1}^M \frac{v_j}{V} \lambda_{i(j)}, \quad i = 1, 2, \dots, n \quad [6]$$

$$\mu = \sum_{j=1}^M \frac{v_j}{V} \mu_j \quad [7]$$

Recall again that M counts the number of blocks making up V and n the number of data points used to calculate the n weights $\lambda_{i(j)}$, $i = 1, 2, \dots, n$ that give the kriged estimate $Z_j^* = \sum_{i=1}^n \lambda_{i(j)} Z_i$ of the j th block. From this, the above equation for the global weights follows,

$$\begin{aligned} Z^* &= \sum_{j=1}^M \sum_{i=1}^n \frac{v_j}{V} \lambda_{i(j)} Z_i \\ &= \sum_{i=1}^n \left(\sum_{j=1}^M \frac{v_j}{V} \lambda_{i(j)} \right) Z_i \\ &= \sum_{i=1}^n \lambda_i Z_i, \quad \lambda_i = \sum_{j=1}^M \frac{v_j}{V} \lambda_{i(j)} \end{aligned} \quad [8]$$

A similar argument produces the ‘global’ Lagrange multiplier μ .

If it were simple to krig each block using all n available data values there would be no problem in producing the required global kriging variance associated with Z^* , since this error variance is given by

$$\sigma_V^2 = \bar{C}(V, V) - \mu - \sum_{i=1}^n \lambda_i \bar{C}(V, v_i) \quad [9]$$

where the Lagrange multiplier $\mu < 0$. In the case of SK, $\mu = 0$. The global weights λ_i and the global Lagrange multiplier μ would be known from the above equations, and the covariances could be calculated straightforwardly from the variogram based on the smallest sample support. However, this is not the case.

As it stands, the implementation of the kriging equations to determine the $\lambda_{i(j)}$ becomes impractical if the number of data points n is large (e.g. several hundreds or thousands, as would be the case here). Some of the weights become so small that rounding errors in the computer pose a serious problem.

The reduced data-set method (RDM)

An approximation to the above stringent requirements is clearly needed. One knows that only the ‘nearest data neighbours’ to the centre (or centre of gravity) of each block v_j are influential in determining Z_j^* . Therefore an obvious approximation suggests itself: pretend that all the data values were used in kriging, but set artificially to zero the weights of those data samples that were not actually used to krig a specific block (and which are way out of variogram range anyway). Call this the ‘reduced data method’ or RDM for short. Although they used a different approach to calculating global kriging variances, Kim and Baafi (1984) alluded to this possibility, at least in theory. This dramatically reduces the number of data values required in kriging each block, and hence reduces the dimension of the corresponding matrix to be inverted.

Witwatersrand gold reef evaluation: the ‘variancegram’ tool

This approximation yields results that are not too unreasonable for the global kriging variance in the case of SK, provided the number of data values n used per block is not too small.

For OK, the fact that the unbiasedness condition must be satisfied requires that the global weights sum to unity. This too is satisfied, provided the same requirement is met for the weights $n^l < n$, say, *not* set formally to zero:

$$\sum_{i=1}^{n^l} \lambda_{i(j)} = 1, \quad j = 1, 2, \dots, M \quad [10]$$

$$\begin{aligned} \sum_{i=1}^{n^l} \lambda_i &= \sum_{i=1}^{n^l} \sum_{j=1}^M \frac{v_j}{V} \lambda_{i(j)} = \sum_{j=1}^M \left(\sum_{i=1}^{n^l} \frac{v_j}{V} \lambda_{i(j)} \right) \\ &= \sum_{j=1}^M \frac{v_j}{V} (1) = 1 \end{aligned} \quad [11]$$

As with most approximations, there is a price for introducing these simplifications: in artificially setting a selection of otherwise small but finite weights to zero, and implicitly leaving these data values out of the kriging equations, one finds that the global kriging variance is hugely overestimated.

Examples of the results obtained with this approximation for the test data are given next.

Results obtained with the RDM

For ease of reference, the first term in Equation [9] (the global kriging variance formula) will be called the *polygon covariance* term, the second term the *global Lagrange multiplier*, or μ , and the last term the *sample/polygon* term, or sp .

Channel sample data was averaged into 239 60 m × 60 m blocks, and variograms based on the raw average gold accumulation values showed good structure (the histograms of block averages are only slightly skewed, allowing direct calculation on raw values, instead of on logarithms). For the purposes of this illustration, both the data and the unknown blocks to be estimated were taken to be of the same support, i.e. 60 m square blocks.

A rectangle of 18 × 5 (60 m) blocks, adjacent to the data blocks, was kriged (point kriging) using various schemes. Included in the latter was an exhaustive run for both OK and SK using all 239 blocks to kriging every unknown block, with the idea that this would render the ‘true’ value for the global kriging variance in each case that one is striving for. This forms the last entry in the tables that follow. The other runs were based on using respectively $n = 1, 2, 4, 8, 16, 32, 64$ data samples to kriging each block, for both OK and SK.

The polygon, for which the global kriging variance was estimated, included all 90 kriged blocks. The value obtained for the polygon covariance is, of course, a constant. The results obtained for SK are the following (global kriging variance in the last column):

n	$\sum_{j=1}^n \lambda_j$	$\sum_{i=1}^n \lambda_i \bar{C}(V, v_i)$	σ_V^2
1	0.223	34 214	161 354
2	0.328	50 660	144 908
4	0.421	64 044	131 524
8	0.479	71 037	124 531
16	0.507	75 996	119 572

32	0.475	81 561	114 007
64	0.425	81 078	114 490
239	0.454	81 195	114 373

For values of n upwards of 16, a large number of negative weights are obtained and their impact is clear. Using a limited number of samples with which to kriging each block does overestimate the global kriging variance, but not too dramatically.

The results obtained for OK, under otherwise exactly the same circumstances, are the following:

n	$-\mu$	$\sum_{i=1}^n \lambda_i \bar{C}(V, v_i)$	σ_V^2
1	166 8273	153 594	1 710 247
2	979 379	154 284	1 020 663
4	576 688	152 143	620 113
8	380 837	144 328	432 077
16	257 711	137 838	315 441
32	192 633	132 324	255 877
64	138 827	123 468	210 927
239	44 929	97 512	142 985

The amount by which the global kriging variance is overestimated climbs drastically as the number of samples n used in OK is reduced.

Faced with the above problem, a study was made of how sp (the second column in the above tables) in the case of OK and SK, and $-\mu$ in the case of OK, decrease/increase with n , the number of samples used in kriging each block.

The link to the variancegram in the case of OK

The link to the variancegram will be illustrated graphically in the first place. The link will be made even though the variancegram is based on logvariance calculations, and the kriging was done on raw values. In fact, raw variances can be used as basis for the variancegram too, but in the case of Witwatersrand gold reefs the logvariances probably result in a more stable graph, which is very important.

We require the (log)variance as a function of the block side l . As was demonstrated (Figure 3), this function $\Gamma(l)$ can be faithfully reproduced by a set of nested spherical variogram structures, as given in Equations [3] and [4].

An important requirement for the link is that the variancegram has to be fitted specifically up to a square side that is indicated by the total area defined by all the data blocks. In the case of the test exercise the total number of data blocks is 239 and they cover 239 × 3600 m², i.e. 860 400 m². The side of an equivalent square is 928 m. Thus the variancegram has to be refitted only up to 928 m, which changes the parameters somewhat.

The required fit to the variancegram was accomplished using five spherical variogram structures with parameters:

$C_0 = 0$	
$C_1 = 0.710,$	$a_1 = 20 \text{ m}$
$C_2 = 0.280,$	$a_2 = 60 \text{ m}$
$C_3 = 0.190,$	$a_3 = 180 \text{ m}$
$C_4 = 0.090,$	$a_4 = 540 \text{ m}$
$C_5 = 0.180,$	$a_5 = 928 \text{ m}$

It is useful for the illustration to introduce the logvariance difference,

$$f(l) = \text{Sill} - \Gamma(l) \quad [12]$$

where the ‘sill’ in this equation can be any convenient value, since it will fall away in the final calculations.

Witwatersrand gold reef evaluation: the ‘variancegram’ tool

We chose a sill of $f(0) = 2.2276$, the overall logvariance of the data area.

Inserting the above values for the fit of the variancegram up to a square side l of 928 m, one obtains the graph in Figure 4 for $f(l)$ versus l .

The global Lagrange parameter $-\mu$ in the case of OK

The remarkable finding is that the same behaviour as the above variancegram fit is displayed when $-\mu(n)$ is plotted against $(n-1)$, where n is the number of samples used in kriging. Figure 5 gives the experimental points obtained for the test data.

It is immediately clear from Figures 4 and 5 that these two graphs can be scaled to coincide, and the scaling constants will give proportionality criteria. Furthermore, when $-\mu$ is plotted versus $f(l)$ (see Figure 11) the relationship is shown to be a linear one.

The variancegram is known beforehand, and therefore, three known points on the $-\mu$ versus $(n-1)$ curve will determine the scaling constants for the entire curve. This in turn enables one to determine $-\mu$ for a value of n of, say, 239 blocks, *i.e.* all the data blocks, which means that the correct value of the $-\mu$ component of the global kriging variance can be computed.

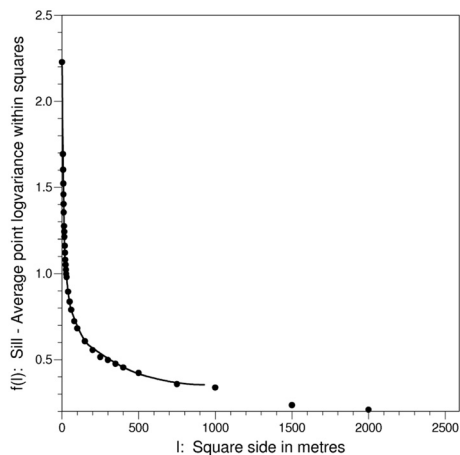


Figure 4—Exp. variancegram points with fit of $f(l)$ vs l up to $l = 928$ m

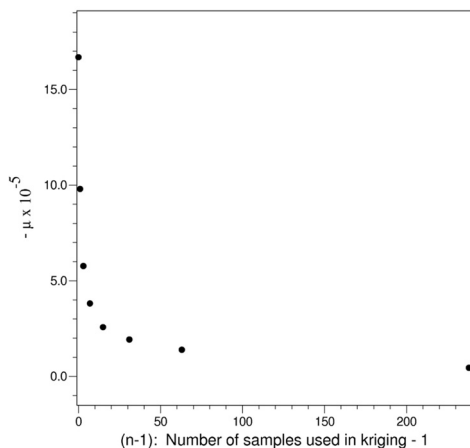


Figure 5—Experimental behaviour of $-\mu(n)$ versus $(n-1)$

The actual determination of the scaling constants will be discussed below.

The sample/polygon term sp in the case of OK

$sp(n)$ also decreases with $(n-1)$, where n is the number of samples used in kriging. However, its behaviour is different from that of $-\mu(n)$ versus $(n-1)$, shown in Figure 5.

Figure 6 gives the experimental points obtained for the test data when $sp(n)$ is plotted versus $(n-1)$. There is not the immediate resemblance to the way in which $f(l)$ decreases with l .

The scaling behaviour of sp is clearly more complicated than was the case for $-\mu$. When sp is plotted versus $f(l)$ the result is a curve resembling $\Gamma(l)$ versus l . Figure 7 gives the former curve, and Figure 8 shows the latter for comparison.

The two curves can again be scaled to coincide, identifying the necessary scaling constants and allowing extrapolation to $n = 239$ for the correct value of $sp(n)$.

The procedures are discussed below.

The link to the variancegram in the case of SK

In the case of SK there is of course only the sample/polygon term to correct in the equation for the global kriging variance,

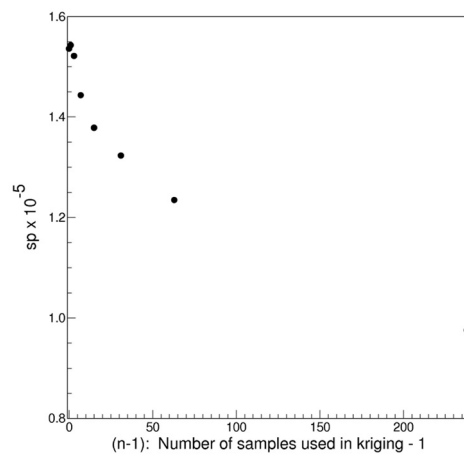


Figure 6—Experimental behaviour of $sp(n)$ versus $(n-1)$ for OK

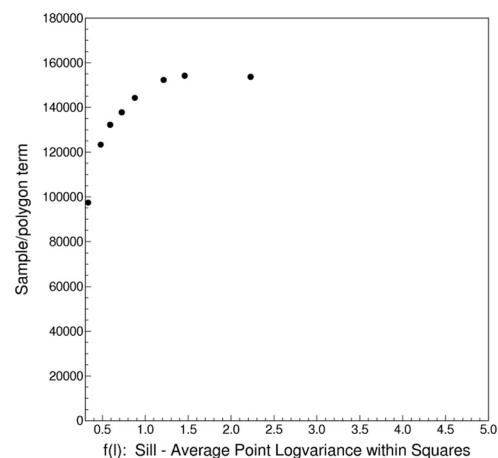


Figure 7—Sample/polygon term versus $f(l)$ for OK

Witwatersrand gold reef evaluation: the ‘variancegram’ tool

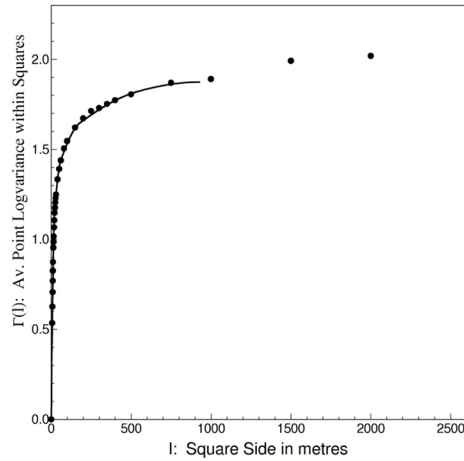


Figure 8—The variancegram $\Gamma(l)$ versus l

since the Lagrange multiplier is absent. The procedure for correcting sp here is very similar to that used for correcting $-\mu$ in the case of OK, with only one difference: sp increases with n (see Figure 9) in a manner reminiscent of the way in which $\Gamma(l)$ increases with l .

This means that the only change necessary in the approach used for $-\mu$ is a change in what $f(l)$ in Equation [12] now stands for:

$$f(l) \rightarrow \Gamma(l) \quad [13]$$

By empirically scaling the axes in the respective graphs of sp versus $(n-1)$ and the new $f(l)$ versus l , the two curves can again be made to coincide. The relationship is linear.

Note that for $n = 64$ and $n = 239$ in Figure 9, the large number of negative weights obtained have an impact on the experimental value of the sample/polygon term. The star in the graph indicates the calculated ‘true’ sp value for $n = 239$, which is unaffected by negative weights. Aside from that, it is not difficult to see that the above graph can be scaled to coincide with $f(l) = \Gamma(l)$.

An approximate analytical model

In this section a simple analytical model kriging system is examined to obtain some insight into the behaviour of μ and sp for OK, for example, with increasing n . Analogous results can also be obtained for SK. The crux of the problem is to see if one can understand their behaviour from the structure and solutions of the kriging equations themselves. The OK kriging system to be solved is, as usual,

$$\sum_{k=1}^n C(v_i, v_k) \lambda_k + \mu = C(v_i, v) \quad [14]$$

$$\sum_{k=1}^n \lambda_k = 1 \quad [15]$$

where n data points have been used to kriging block v . To obtain approximate analytic solutions for the values of λ and μ for arbitrary n replace the off-diagonal elements $C(v_i, v_k)$, for $i \neq k$, by their common average, and the right hand terms $C(v_i, v)$ by their average taken over all data points. Making the replacements

$$C(v_i, v_i) = C, \quad (\text{no approximation}) \quad [16]$$

$$C(v_i, v_k) \rightarrow \overline{C(v_i, v_k)} = C_{dd}, \quad i \neq k$$

$$C(v_i, v) \rightarrow \overline{C(v_i, v)} = C_{d,v}$$

in Equation [14], the resulting system of equations can be solved exactly. One finds that $\lambda_i = 1/n$ for all i ,

$$\lambda_1 = \lambda_2 = \dots = \lambda_n = \frac{1}{n} \quad [17]$$

while μ is given by

$$-\mu(n) = \frac{C}{n} + C_{dd}(1 - \frac{1}{n}) - C_{d,v} \quad [18]$$

This exhibits the same qualitative behaviour as was seen empirically: for $n = 1$ the value is $\mu(1) = C - C_{d,v}$, dropping down to a lower limit of $\mu(\infty) = C_{dd} - C_{d,v}$. A graph of the solution for $-\mu(n)$ is shown in Figure 10 (compare to Figure 5).

In generating this illustrative figure, the typical estimates $C = 1.245 \times 10^6$, $C_{dd} = 0.187 \times 10^6$, and $C_{d,v} = 0.019 \times 10^6$ have been used.

In this connection it is therefore important to realize that the strong n -dependence of μ arises solely as a consequence of the restriction $\sum_{k=1}^n \lambda_k = 1$. This condition forces the common λ values to equal $1/n$, and thus the μ in Equation [18] to become a function of $1/n$.

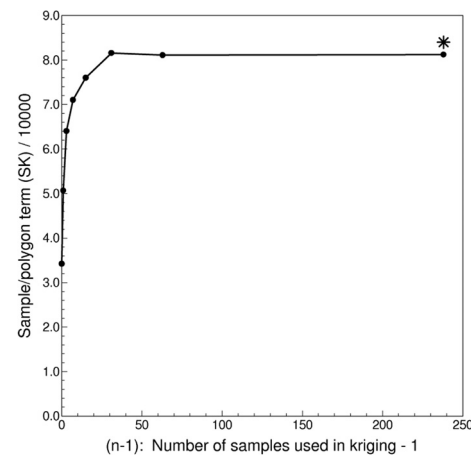


Figure 9—Sample/polygon term sp (SK) versus $(n-1)$

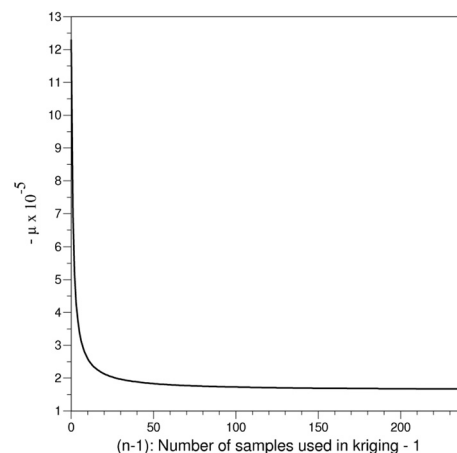


Figure 10—Behaviour of modelled $-\mu$ with $(n-1)$

Witwatersrand gold reef evaluation: the ‘variancegram’ tool

Note that the order of magnitude of the model values of $sp(n)$ and $-\mu(n)$ are not expected to reproduce the empirical values, due to the extremely crude assumptions made. However, the model faithfully reproduces the *qualitative* behaviour of the empirical observations to a striking degree.

The application of the findings to the computing of correct global kriging variances from the variables obtained by local kriging is set out next.

Correcting global kriging variances in the case of OK

The correction procedures referred to above will be illustrated using the same exercise that was used previously: the data-set consisted of average gold accumulation values for 239 60 m × 60 m blocks, and it was used to krig an adjacent rectangle of 18 × 5 (60 m) blocks. The relevant distance over which the variancegram had to be fitted was 928 m.

As before:

$$f(l) = \text{Sill} - \Gamma(l) \quad [19]$$

where an arbitrary ‘sill’ value of 2.2276 (in this case the overall logvariance of the data area) was chosen for convenience, and $\Gamma(l)$ reproduced the variancegram by means of a model fitted with five nested spherical structures:

$$\begin{aligned} \Gamma(l) &= \sum_{i=1}^{p=5} \Gamma_i(l) \\ \Gamma_i(l) &= C_i \left[\frac{3}{2} \left(\frac{l}{a_i} \right) - \frac{1}{2} \left(\frac{l^3}{a_i^3} \right) \right] \quad \text{for } l \leq a_i \\ &= C_i \quad l > a_i \end{aligned} \quad [20]$$

l denotes the side length of the square for which each logvariance is calculated in the experimental variancegram.

Methodology for correcting $-\mu$, the global Lagrange parameter

It was found that $-\mu$ decreases with n , the number of samples used to krig each block, in a manner reminiscent of the way in which $f(l)$ decreases with l (see Figures 4 and 5). By empirically scaling the axes in the respective graphs, these two curves can be made to essentially coincide.

Mathematically this means that the variable pairs $[-\mu, f(l)]$ and $[n, l]$ are linearly related. Hence we may write

$$l = \alpha(n - 1) \quad [21]$$

$$-\mu = \beta f[l] + \gamma = \beta f[\alpha(n - 1)] + \gamma \quad [22]$$

where the $(n - 1)$ dependence on n is necessary since the minimum value of n is 1, and this has to be associated with the minimum value of l of the variancegram, which is zero. The function f was defined in Equation [19]. The constants α , β , and γ perform the necessary scaling to cause the two curves in question to overlap. Since there are three of these unknowns, we need three sets of data input to determine them. From a practical viewpoint, one would expect to know values for small sample numbers n_1, n_2, n_3 , say. Then, by manipulating Equation [22] one can solve in turn for α , β , and γ from

$$\frac{-\mu_1 + \mu_2}{-\mu_1 + \mu_3} = \frac{f[\alpha(n_1 - 1)] - f[\alpha(n_2 - 1)]}{f[\alpha(n_1 - 1)] - f[\alpha(n_3 - 1)]} \quad [23]$$

$$\beta = \frac{-\mu_1 + \mu_2}{f[\alpha(n_1 - 1)] - f[\alpha(n_2 - 1)]} \quad [24]$$

$$\gamma = -\mu_1 - \beta f[\alpha(n_1 - 1)] \quad [25]$$

The notation is $\mu_i = \mu(n_i)$. The first of these three relations is an equation for α . Once this has been solved, β and γ follow. To solve this equation we cross-multiply and set

$$R(x) = \frac{(-\mu_1 + \mu_2)(f[x(n_1 - 1)] - f[x(n_3 - 1)]) - (-\mu_1 + \mu_3)(f[x(n_1 - 1)] - f[x(n_2 - 1)])}{f[\alpha(n_1 - 1)] - f[\alpha(n_2 - 1)]} \quad [26]$$

where x is a running variable. One sees that $R(x)$ always passes through zero due to the form of f . Thus $R(x)$ has one real root at $x = 0$; if $R(x)$ were strictly a cubic polynomial in x one could argue for three real roots: one at 0, plus one negative and one positive root (we ignore the unphysical, but mathematically possible, outcome of one real root at zero, plus two complex conjugate roots). This is unfortunately not the case due to the discontinuities in the $\Gamma_i(l)$ that enter the form of $f(l)$. However, if the real, positive root is not so large as to exceed the shortest range in Equation [20], one can perhaps expect $R(x)$ to ‘act’ like a cubic. From the scaling hypothesis it is clear that α must always correspond to the positive, non-zero root. A specific example of how $R(x)$ behaves when it has three real roots (the only case of practical interest) is shown in Figure 11.

A word of caution is in order here. If the coincidence of the $-\mu$ versus n and f versus l curves were exact, it would not matter which known values were used to obtain α , β and γ . However, since we are dealing with a so far empirically established relation which need not be exact, it will matter to some extent what input values are used. As it turns out, $n_1 = 1, n_2 = 2$, and $n_3 = 8$ or 16 or 32, etc. give very good results.

Methodology for correcting sp , the sample/polygon term

As was shown above, the sample/polygon term, $\lambda_\alpha \bar{C}(V, v_\alpha)$, also exhibits a scaling behaviour, but one which is different from that of $-\mu$. We first exploit the known $l = \alpha(n - 1)$ relation of Equation [21] to convert

$$sp(n) = \sum_{\alpha=1}^n \lambda_\alpha \bar{C}(V, v_\alpha) \quad [27]$$

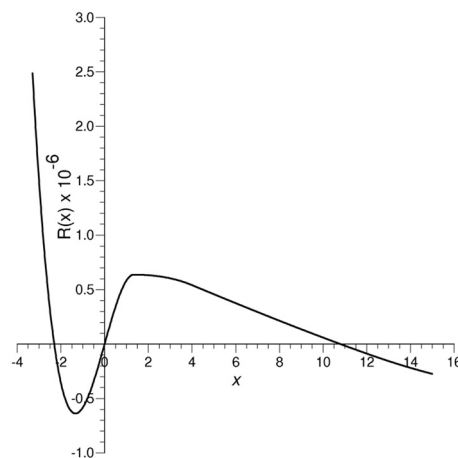


Figure 11 – The function $R(x)$ versus x

Witwatersrand gold reef evaluation: the 'variancegram' tool

to a function of $f = f[\alpha(n-1)]$ instead of n . (This is what gave a straight line for $-\mu$). In the case of sp , however, one finds a curve resembling $\Gamma(l)$ versus l (see Figures. 7 and 8). The two curves, sp versus f and Γ versus l can again be scaled empirically to coincide. This means that

$$sp = b\Gamma[af] + c = b\Gamma[af[\alpha(n-1)]] + c \quad [28]$$

where a, b, c are three new scaling constants. The equations determining these have the same structure as Equations [23]–[25] with $sp_i = sp(n_i)$ substituted for $-\mu_i$ and $\Gamma[af[\alpha(n_i-1)]]$ for $f[\alpha(n_i-1)]$. Thus

$$\begin{aligned} \frac{sp_1 - sp_2}{sp_1 - sp_3} &= \frac{\Gamma[af[\alpha(n_1-1)]] - \Gamma[af[\alpha(n_2-1)]]}{\Gamma[af[\alpha(n_1-1)]] - \Gamma[af[\alpha(n_3-1)]]} \\ b &= \frac{sp_1 - sp_2}{\Gamma[af[\alpha(n_1-1)]] - \Gamma[af[\alpha(n_2-1)]]} \\ c &= sp_1 - b\Gamma[af[\alpha(n_1-1)]] \end{aligned} \quad [29]$$

The least squares fit modification

Instead of relying on the assumption that the straight line of $-\mu$ versus f in Equation [22] can be forced to pass through three of the data (experimental) points, which may lead to no real root for α in Equation [26], one can try a 'milder' approach of only requiring a fit in the least squares sense. This more robust procedure should always produce a solution, which is important for a production system. To do so one requires that the sum over the differences squared of S sample points that are going to be used to determine the scaling constants,

$$F(\alpha, \beta, \gamma) = \sum_i^S [\beta f[\alpha(n_i-1)] + \gamma - (-\mu_i)]^2 \quad [30]$$

be minimized as a function of the parameters γ, β , and α . This means that

$$\frac{\partial F}{\partial \gamma} = 2\beta \sum_i f_i + 2S\gamma + 2\sum_i \mu_i = 0 \quad [31]$$

$$\frac{\partial F}{\partial \beta} = 2\sum_i f_i^2 + 2\gamma \sum_i f_i + 2\sum_i \mu_i f_i = 0 \quad [32]$$

$$\frac{\partial F}{\partial \alpha} = 2\sum_i (n_i - 1)f'_i[\beta f_i + \gamma + \mu_i] = 0 \quad [33]$$

where $f_i = f[\alpha(n_i-1)]$ and f'_i is the derivative of $f(x)$ with respect to its argument $x = \alpha(n_i-1)$. The first two equations are the usual least square conditions, arrived at by differentiating w.r.t. β and γ . The last equation is new. It arises upon differentiating w.r.t. α and requiring that this too should vanish. Then, upon defining

$$D = S \left(\sum_i f_i^2 \right) - \left(\sum_i f_i \right)^2 \quad [34]$$

one finds, as usual, that

$$\begin{aligned} \beta &= \left[-S \left(\sum_i \mu_i f_i \right) + \left(\sum_i \mu_i \right) \left(\sum_i f_i \right) \right] \frac{1}{D} \\ \gamma &= \left[\left(-\sum_i \mu_i \right) \left(\sum_i f_i^2 \right) + \left(\sum_i f_i \right) \left(\sum_i f_i \mu_i \right) \right] \frac{1}{D} \end{aligned} \quad [35]$$

Substituting these two values into Equation [33], one finds an equation that determines the as-yet unknown α as a root of an algebraic equation, very much as in Equation [26].

This equation can be recast as

$$Q(y) = \beta \left(\sum_i (n_i - 1) f_i f'_i \right) + \gamma \left(\sum_i (n_i - 1) f'_i \right) + \left(\sum_i (n_i - 1) \mu_i f'_i \right) = 0 \quad [36]$$

i.e. the value of y at which $Q(y)$ vanishes is the preferred value of α . Once this has been determined, the preferred values of β and γ follow from Equation [35]. Next, one has to check that the solution so obtained indeed leads to a minimum value of $F(\alpha, \beta, \gamma)$, evaluated at these values of α, β , and γ . The condition for this to be the case is that all three second partial derivatives are positive there. Doing the calculations, one finds the three conditions

$$\frac{\partial^2 F}{\partial \gamma^2} = 2S > 0 \quad [37]$$

$$\frac{\partial^2 F}{\partial \beta^2} = 2 \sum_i f_i^2 > 0 \quad [38]$$

$$\frac{\partial^2 F}{\partial \alpha^2} = 2\beta \sum_i (n_i - 1)^2 [\beta f_i'^2 + (\beta f_i + \gamma + \mu_i) f_i''] > 0 \quad [39]$$

Of these, the first two are obviously always met: the third should be used as a test to ensure that one is indeed at a minimum of F as envisaged by the least squares problem. In the above development, pathological cases where F is neither at a maximum or at a minimum have been excluded. One can check for such cases by verifying in addition that

$$\left(\frac{\partial^2 F}{\partial \alpha \partial \beta} \right)^2 < \frac{\partial^2 F}{\partial \alpha^2} \frac{\partial^2 F}{\partial \beta^2} \quad [40]$$

$$\left(\frac{\partial^2 F}{\partial \beta \partial \gamma} \right)^2 < \frac{\partial^2 F}{\partial \beta^2} \frac{\partial^2 F}{\partial \gamma^2} \quad [41]$$

$$\left(\frac{\partial^2 F}{\partial \gamma \partial \alpha} \right)^2 < \frac{\partial^2 F}{\partial \gamma^2} \frac{\partial^2 F}{\partial \alpha^2} \quad [42]$$

which, in combination with the vanishing first derivatives, guarantee a minimum or a maximum.

The derivatives of f that are required come from Equation [19] with Equation [20] inserted for $\Gamma(l)$. Note that this consists of a sum of individual spherical variogram forms that have zero derivatives once the argument exceeds the range. This should be borne in mind in programming the derivative. One has

$$\frac{d\Gamma_j}{dl} = \frac{3C_j}{2a_j} \left(1 - \frac{l^2}{a_j^2} \right) \quad l \leq a_j \quad [43]$$

for the j th structure. Hence

$$f' = \frac{df}{dl} = - \sum_{j=1}^p \frac{3C_j}{2a_j} \left(1 - \frac{l^2}{a_j^2} \right) \theta(a_j - l) \quad [44]$$

where the step function θ is merely a reminder that each term in the sum will contribute only provided l does not exceed its particular range a_j . The expressions for the required derivatives are

$$f'_i = - \sum_{j=1}^p \frac{3C_j}{2a_j} \left[1 - (n_i - 1) \frac{\alpha^2}{a_j^2} \right] \theta(a_j - l_i) \quad [45]$$

Witwatersrand gold reef evaluation: the ‘variancegram’ tool

$$f_i'' = \sum_{j=1}^p \frac{3C_j}{a_j^2} \left[(n_i - 1) \frac{\alpha}{a_j} \right] \theta(a_j - l_i) \quad [46]$$

Notice that both derivatives are a sum over the p structures making up f^i and f^j , evaluated at the particular value $l = l_i = (n_i - 1)\alpha$.

It is interesting to observe that in all cases where low values of n_1, n_2, n_3 were considered to obtain the scaling constants, the least squares method gave exactly the same values for the scaling constants as were found from the (largest) root of $R(x)$ previously.

Calculation of a corrected $-\mu$

To calculate the scaling constants for the test data the values for $-\mu(n_i)$ of $[n_1, n_2, n_3] = [1, 2, 16]$ are used. Figure 11 shows the plot of $R(x)$ versus x , the positive root of which will determine α .

The value of the positive root in this case is $x = 10.7968$. This means that

$$\alpha = 10.7968 \quad [47]$$

$$\beta = 867276 \quad [48]$$

$$\gamma = -256987 \quad [49]$$

Knowing these constants, we can write the way $-\mu$ is expected to vary with the sample number $(n - 1)$:

$$-\mu(n) = 867276 f[10.7968(n - 1)] - 256987 \quad [50]$$

Extrapolating with this formula to $n = 239$, one finds

$$-\mu(239) = 48988 \text{ versus } 44929 \quad [51]$$

by direct calculation.

Verification of the scaling hypothesis for $-\mu$

In order to verify that these scaling factors are indeed correct, we show the values of $f[10.7968(n - 1)]$, and the corresponding values of $-\mu(n)$ found in the kriging run:

Experimental $-\mu \times 10^{-6}$	Samples used n	Calculated f $f[10.7968(n-1)]$
1.6683	1	2.2276
0.9794	2	1.4241
0.5767	4	0.9544
0.3808	8	0.7484
0.2517	16	0.5395
0.1926	32	0.4865
0.1388	64	0.3755
0.0489	239	0.3528

The number pairs $[f, -\mu]$ are plotted as closed circles in Figure 12.

The straight line is given by Equation [22], plotted as a function of the variable f . The agreement is truly remarkable!

The predicted point at $n = 239$ has also been included to illustrate that increasing values of n ‘pile-up’ on each other along the straight line of $-\mu$ versus f .

It is useful to point out again that the inclusion of a sill to define the logvariance difference $f(l) = Sill - \Gamma(l)$ is merely one of convenience. From the previous development it is clear that one could equally well have considered $-\mu$ as a function of the variancegram $\Gamma(l) = [\alpha(n - 1)]$. Then

$$\begin{aligned} -\mu &= \beta(Sill - \Gamma(l)) + \gamma = -\beta\Gamma[\alpha(n - 1)] + \\ &\underbrace{(\beta Sill + \gamma)}_{\gamma'} = -\beta\Gamma[\alpha(n - 1)] + \gamma' \end{aligned} \quad [52]$$

Here $-\mu$ is a linear function of the variancegram with gradient $-\beta$ and a revised constant $\gamma' = Sill \times \beta + \gamma$. Thus the entire exercise of determining the scaling constants could equally well have been implemented by using the variancegram: the choice of a sill is irrelevant.

Calculation of a corrected sp

By cross-multiplying in Equation [29] one finds a function that is the exact analogue of $R(x)$ in Equation [26]. The root of the former equation now gives the value of a in Equation [28] for fixed α . Using the input values of sp_i for $[n_1, n_2, n_3] = [2, 4, 16]$ one determines the scaling constants as

$$a = 71.03 \quad [53]$$

$$b = 233614 \quad [54]$$

$$c = -283409 \quad [55]$$

Knowing these constants we can similarly write down an equation for sp versus n :

$$\begin{aligned} sp(n) &= 233614 \Gamma \\ &(71.03[f[11.914(n - 1)]] - 283409 \end{aligned} \quad [56]$$

The extrapolation to $n = 239$ gives

$$sp(239) = 116883 \text{ versus } 97512 \quad [57]$$

Notice a small but important point: the value of $\alpha = 11.914$ appears here instead of $\alpha = 10.7968$. This happens because the samples for $[n_1, n_2, n_3] = [2, 4, 16]$ were used for the sp , since sp_1 and sp_2 lie too close to give a useful determination of a, b, c . One must thus use the same n 's for the coefficient α to remain consistent.

Verification of the scaling hypothesis for sp in the case of OK

We verify the linear behaviour again by evaluating $\Gamma(n) = \Gamma(f)$ for the values at which the sp are known from the kriging runs

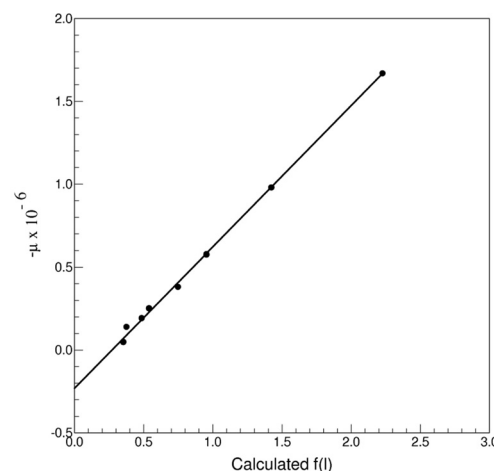


Figure 12— $-\mu$ versus $f(l)$

Witwatersrand gold reef evaluation: the 'variancegram' tool

$sp \times 10^{-5}$	n	$\Gamma(71.03f[11.914(n-1)])$
1.536	1	1.874
1.543	2	1.874
1.521	4	1.864
1.443	8	1.837
1.378	16	1.803
1.323	32	1.765
1.235	64	1.71
1.169	239	1.713

The straight line character of the plot sp versus Γ is verified in Figure 13.

Resulting estimated global kriging variance in the case of OK

The directly calculated ('true') OK global kriging variance for the test data is

$$\sigma_{OK}^2 = 195568 + 44929 - 97512 = 142985 \quad [58]$$

The estimated OK global kriging variance based on the proposed correcting method is

$$\sigma_{OK}^{*2} = 195568 + 48988 - 116883 = 127673 \quad [59]$$

As is clear from comparison of the above two equations, the estimated $-\mu$ and sp values are both higher than the true values, and since they are subtracted from each other, the resultant error in the global kriging variance should be reduced. However, the main source of discrepancy is the difference in the sp values. This could be ascribed to the fact that the true value of 97 512 is anomalously low due to the impact of negative weights, whereas the impact is absent in the estimated value.

Nevertheless, the estimated value exhibits an error of less than 11% of the true value, whereas the uncorrected RDM value based on kriging with 16 samples showed an error of 121%.

Correcting global kriging variances in the Case of SK

The correction procedure here focuses only on sp , since the Lagrange multiplier is absent. It was illustrated in Figure 9, that for SK, sp increases with n in a manner reminiscent of the way in which $\Gamma(l)$ increases with l . This again means that the variable pairs $[sp, f(l)]$ and $[n, l]$ are linearly related.

Hence the rest of the formulism is exactly the same as was the case for $-\mu$, except that $-\mu$ is everywhere replaced by sp , and $f(l)$ is reinterpreted as being $\Gamma(l)$.

The least squares fit modification of the method to ensure a solution in a production environment follows immediately with the same replacements in the formulae of the relevant subsection above.

The experimental values obtained for sp in the case of SK based on kriging with $n = 1, 2, 4, 8$, etc. are given above. Using $n_1 = 1$, $n_2 = 2$ and $n_3 = 16$, the following scaling constants were obtained when solving the equivalent of Equations [23]–[25], replacing $-\mu$ by sp :

$$\alpha = 8.058 \quad [60]$$

$$\beta = 26564 \quad [61]$$

$$\gamma = 34214 \quad [62]$$

Knowing these constants, one can again write down the way sp is expected to vary with the sample number $(n-1)$:

$$sp(n) = 26564 f[8.058(n-1)] + 34214 \quad [63]$$

Extrapolating with this formula to $n = 239$, one finds

$$sp(239) = 83983 \text{ versus } 81195 \quad [64]$$

by direct calculation.

This estimated value for sp (239) is indicated by a star in Figure 9. It would indeed seem that this would have been the value of sp if the influence of negative weights were absent.

The scaling hypothesis for sp in the case of SK is verified next by evaluating $f(n) = \Gamma[8.058(n-1)]$ for the values at which the sp are known from the kriging runs.

$sp \times 10^{-4}$	n	$\Gamma[8.058(n-1)]$
3.4214	1	0.0000
5.0660	2	0.6191
6.4044	4	1.1905
7.1037	8	1.4284
7.5996	16	1.5729
8.1561	32	1.6912
8.1078	64	1.8122
8.1195	239	1.8736

The straight line character of the plot sp versus Γ is again verified in Figure 14.

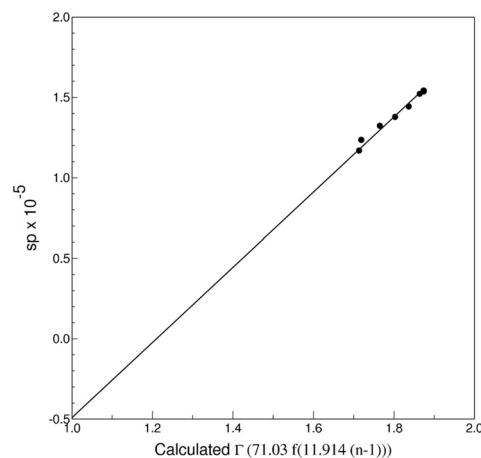


Figure 13— sp versus $\Gamma(71.03 f[11.914(n-1)])$

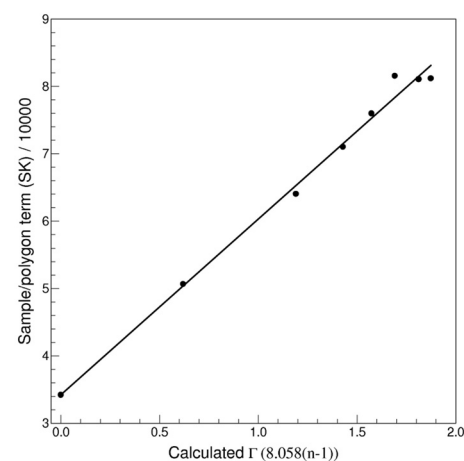


Figure 14— sp versus $\Gamma(8.058(n-1))$

Witwatersrand gold reef evaluation: the 'variancegram' tool

The deviations from the straight line at high n values are once more a manifestation of the impact of negative weights on the experimental sp values when kriging is actually performed with such large numbers of samples.

The directly calculated ('true') SK global kriging variance for the test data is

$$\sigma_{SK}^2 = 195568 - 81195 = 114373 \quad [65]$$

The *estimated* SK global kriging variance based on the proposed correcting method is

$$\sigma_{SK}^{*2} = 195568 - 83983 = 111585 \quad [66]$$

The estimated SK global kriging variance is within 2% of the true value, whereas the uncorrected RDM SK global kriging variance based on kriging with 16 samples is within 4.5% of the true value.

Conclusion

With research into new mining methods opening up the possibility of going to unprecedented depths in the Witwatersrand gold mines, the legacy of Krige's work remains as relevant as ever.

References

- JOURNEL, A.G., and HUIJBREGTS, Ch. J. 1978, Mining Geostatistics. Academic Press, New York. 600 pp.
- KIM, Y.C., and BAAFI, E.Y. 1984. Combining local kriging variances for short-term mine planning. *Geostatistics for Natural Resources Characterization*, Part 1. Verly, G., David, M., Journel, A.G., and Marechal, A. (eds.). D. Reidel, Dordrecht, The Netherlands. pp. 187.
- KRIGE, D.G. 1978. Lognormal-de Wijsian geostatistics for ore evaluation. *Monograph Series M1*. South African Institute of Mining and Metallurgy, Johannesburg. p. 26. ◆

Celebrating Kimberley's Firsts



E-H 19717



DE BEERS
GROUP OF COMPANIES

1871: First diamond mining begins in what becomes the City of Kimberley

1873: Not always first! Kimberley was once the second largest town in South Africa after Cape Town

1881: Kimberley Royal Stock Exchange, in the country's first industrial centre

1882: First city in the Southern Hemisphere to install electric street lighting

1887: First scientific identification of diamond bearing rock type - Kimberlite - named in honour of Kimberley

1896: South Africa's first school of mines opens giving a foundation to three other major South African Universities

1904: First electric tram in South Africa

1912: First woman trained as a pilot

1913: South Africa's first aviation school opens

1914: The Big Hole closes after 34 years reopening as the first South African mining heritage museum

1931: First airport to install lighting equipment and first night landing

1969: First female judge: Miss Justice Leonora van den Heever

1992: First City Council to amalgamate all group areas

2006: De Beers Consolidated Mines empowered with Ponahalo Holdings

2013: First woman head of Kimberley based De Beers Sightholder Sales South Africa: Mpumi Zikalala

2014:
De Beers becomes the first major South African mining company to empower its trading business. Empowerment of De Beers Sightholder Sales South Africa with Ponahalo Holdings.

Our diamond trading business is an empowered entity in Kimberley, creating skilled employment and economic activity; sorting and valuing most of South Africa's diamond production. Trading diamonds in Kimberley to all our South African clients' factories.