THE COMPUTER MAPPING AND ASSESSMENT OF BOREHOLE AND SAMPLING DATA FOR STABLE MINERALS, PARTICULARLY AS APPLIED IN COAL MINING

By R. B. MacGillivray, D. M. Hawkins and M. Berjak

Published in the Journal, January, 1969

Contributions to discussion

V. E. Marting (Associate Member): The authors of this paper should be congratulated in having pioneered the application of the Sharapov predictor to local data, and in having made a valuable contribution to South African experience in the quantitative treatment of geological information using a digital computer.

Although the existing set of three programmes has been described as a 'suite' it does not represent an integrated system in the accepted sense since the communication between runs is made in the form of punched card files to enable the mining engineer to exercise discretion and control over the data entered at each phase. Although the authors are justifiably proud of the successful results obtained through the application of this method, I feel that it is my duty to comment that the computer aspect of the work has not reached full development and is deserving of the financial support needed to increase its widest usefulness. At the present moment a routine to detect 'missing information' has not been commissioned. Due to the lack of documentation the use of the programmes is restricted since the format for the submission of data is known only to the authors. It is very unlikely that a profitable return on the development costs will be enjoyed before the programme is fully documented and made available to wider use.

My criticism of the present incomplete state of the work is being voiced purely with the intent of focussing on the significance of this project and the fact that it is deserving of development on an industry basis.

It should be noted that at the economical cost of R4.00 per sq ft of computer print-out, a mining engineer can hardly afford to miss the opportunity of using the programmes on other (e.g. metalliferous) mineral deposits of tabular formation.

Since the publication of this paper I have had an opportunity to see the results of the programmes developed by Drs M. D. G. Salamon and F. Deist for designing bord and pillar workings. The basic data for these latter programmes are extracted from iso-line plans drawn conventionally by geologists, a task that could be considerably expedited through the use of Mr Hawkins' programmes. It appears that the efforts of these respective teams are complementary and it is in the interests of the industry that these individual components should eventually form part of an integrated system.

E. Margo* (Member): This paper is a good example of how and where the backroom scientist—in this case the mathematical statistician—can aid the mining engineer to develop not only more accurate means for dealing with valuation problems but to develop a tool that can also cope with changed geological specifications.

*Work Study Consultant, Rand Mines Limited.

The mining industry will indeed benefit from this description of a sophisticated mathematical technique and its computer application.

The Rand Mines Group has for some time now been striving to establish acceptable criteria for accurately determining the value of fairly consistent variable thickness tabular deposits from borehole information.

The objectives were twofold. Firstly it was necessary to ascertain the minimum number of boreholes that would give a satisfactory population of values for the purpose of evaluating the deposit within a specified probability of error. It soon became evident that a system of boreholes drilled to a regular pattern would yield the 'best' population. Next, it was necessary to determine the mathematical model most suitable for evaluating the deposit from the available data and which would produce unbiased estimates from which valid confidence limits could be computed. Originally all the analyses were done manually on an electronic desk calculator.

This paper by a team of mathematical statisticians and a mining engineer is the result of the work that has been put into resolving this problem. The success of the team approach is undoubtedly good communication which in turn reinforced inter-disciplinary effort. Each member knew the part he had to play because both the problem and the objectives were clearly defined.

It is interesting to point out that recently Rand Mines had occasion to apply the computer programme to the evaluation of a thick-bedded stratified limestone deposit with very satisfactory results.

In this case the team consisted of a geologist who worked closely with the mathematical statistician. The geologist first defined the size and shape of the deposit, anomalous areas were recognised and borehole cores carefully logged andi nter-preted. This advance geological information assured that the statistical analysis was realistic and pragmatic. Outcrops, cut-offs and geological disturbances were thus assigned meaningful values.

A major benefit derived from the computer programme is to provide management with information for different stated requirements; for example tonnage at a specified grade; grade for specific widths; percentage waste to be handled for various yields and so on.

It should be appreciated that as exploration and exploitation costs increase and as the discovery of economical ore bodies becomes less frequent, more reliance must be placed on methods that will provide financial and technical management with highly probable information at the least cost. A uni-disciplinary inference may well prove to be inadequate if not perhaps erroneous and therefore so much more financially embarrassing.

We owe the authors our thanks and congratulations for placing on record and so making public a reliable and practical scientific tool for valuating and mapping tabular stable deposits.

H. S. Sichel (Member): The problem with which the authors are concerned$_i$ may be stated as follows:

Given n points in a two-dimensional space with co-ordinates for each point x_i and y_i and a third dimension r_i which is called a 'value', how best can one predict this 'value' r at any other point whose plane co-ordinates are x and y.

The values r_i are observations and as such may be subject to errors of measurements. Even if no such error variations were in evidence, the problem would still be a formidable one as in practice the general surface

$$r = f(x, y)$$

is rather irregular with several maxima, plateaus, saddle points and minima.

The authors make short shrift with a technique well known among mine surveyors which is called the 'triangular' method. Perhaps they may have added the other survey procedure called the 'polygonal' method.

Both survey methods assume linearity of the variable r between two observation points and it is not possible to put a standard error of prediction on estimates derived from such procedures.

Next the authors tried two mathematical approaches, i.e. polynomial surfaces and finite Fourier series. The results obtained on their set of data were not encouraging as one would expect from the variations in the observed values r_i, the irregular spacing of boreholes and the data voids as shown in Fig. 3 of the paper under discussion.

This left the writers with the moving average technique in one or other form.

They start off with a formula due to Sharapov which assigns weights to individual observations r_i, inversely proportional to the linear distance d_i which is raised to a power of k. d_i is the distance between the point (x, y), for which an estimate of the value r is required, and the point (x_i, y_i) for which an observed value r_i is available.

If one wishes to predict a value of such a surface at a point for which one has already an observation, i.e. at a particular borehole site, Sharapov's formula will reproduce this borehole value exactly.

In other words, no smoothing will have taken place and no account is taken of possible errors of measurement—as correctly pointed out by the writers. Furthermore, the multiple correlation coefficient would be a rather meaningless $R = 1$.

To overcome this difficulty, the authors introduce another constant 'a' into Sharapov's weighting formula which constitutes a significant advance in the methodology of trend surface fitting.

The choice of constants 'a' and 'k' is rather crucial and one looks forward to future work which will tell us how to estimate these parameters from a given set of data. As no theory has yet been developed, the authors fell back on that prodigious slave, the computer, to determine heuristically what numerical values for 'a' and 'k' should be taken in their set of data. The trouble with such an empirical approach lies in the fact that it does not teach us anything which we could use in a different set of circumstances, say in a different type of orebody or in another coalfield or with another borehole grid. The whole procedure of empirical fitting will have to be repeated *de novo* and valuable time and money will be lost. The computer cannot think and it is because of its lack of brainpower that we will have to pay a rather expensive bill. The incentive to develop some sound theory is, therefore, apparent.

The authors test the goodness of fit of their iso-line plan with the help of the multiple correlation coefficient and the standard error of prediction. I am not altogether happy with this procedure. Trend surface fitting is different from problems encountered in classical statistics.

Individual observations r_i are self-correlated with the estimates $\hat{r}_i$ at the same point (x_i, y_i). This self-correlation may be made as high as one wishes depending on the choice of parameters 'a' and 'k'. As previously pointed out, if $a = 0$ the multiple correlation coefficient becomes unity. Yet does this mean that the error of prediction becomes zero for any other value r, given the co-ordinates (x, y)? The answer is decisively in the negative and shows up the futility of comparing observed and expected values r_i with $\hat{r}_i$ at the known borehole locations.

As I have stressed before on another occasion there does not exist a mathematical short-cut for the comparison of a value predicted from a trend surface with one subsequently observed when mining has taken place.

We will gain real confidence in trend fitting exercises only if we are honest enough to compare actually observed mining results with what we have predicted *a priori* from our forecasting model.

The authors claim that their weighting formula does not depend on the type of borehole grid employed, yet, in another part of the paper, they concede that a regular square or triangular grid will give optimum results.

The fact remains that *ex nihil, nihilo*. Data voids will play havoc with any trend surface including the one proposed. The more boreholes we have at our disposal, and the more regular the drilling pattern, the better will be our power of prediction as obtained from trend surface models.

In conclusion I wish to congratulate Messrs MacGillivray, Hawkins and Berjak for a most stimulating paper. Their modification of Sharapov's formula has certainly opened up new vistas in the important problem of predicting ore characteristics *in situ* from a limited number of sample values.

Dr D. G. Krige (Member): The authors are to be congratulated on a very interesting and stimulating paper on the relatively new approach to evaluation of ore values, i.e. the representation of ore values by two-dimensional contour plans. This approach has received considerable attention in recent years in our gold mines and also overseas, particularly in the U.S.A. Experience to date on gold values confirm the authors' objections to mathematically defined surfaces and the preference for a weighting system which decays with increasing distance. The system in use for gold values is in principle the same as that used by the authors except that:

(i) because very large numbers of values are available practical considerations place a limit on the distance up to which a value(s) can be used, and requires the use of *averages* of values grouped over areas; it would clearly be quite impractical to use all the thousands of individual gold values to estimate the grade at every point on the valuation grid;

(ii) whereas the authors use an exponentially decaying function for weights and experimentally determined parameters for rate of decay and 'errors' of observation, the weighting system and level of 'error' for gold values are determined from detailed data available in areas already mined out; furthermore, the weights need not necessarily decay exponentially;

(iii) there is the possibility of a sufficient number of follow-ups on the predictions for gold values to establish statistically the efficiency of the system; in the case of borehole values the follow-up is usually long deferred and only a single or limited number of follow-ups are possible.

The main obstacle to using the system established for gold values in a borehole analysis such as that of the authors, is that the limited number of borehole values available will in most cases not allow for the necessary convariance and multiple correlation analyses to establish a calculated weighting system based directly on the actual observations. This is the essential difference to be noted and not the 'regularity' of the data on the gold mines and the more 'haphazard' nature of the borehole data as mentioned by the authors. Where as many as a hundred borehole values are available such as used by the authors it might, however, be worthwhile to attempt such a correlation analysis.

A further point of difference is that the authors do not provide for possible directional trends in the borehole values which would necessitate the use of lower rates of decay for the weights in the direction(s) of such trends. Recently this aspect has been receiving more detailed attention on the gold mines with interesting and promising results. Where adequate borehole data is available for estimating any such trends it should be possible to show a substantial improvement in the efficiency of interpolation between and extrapolation beyond the borehole values.

I would like to ask the authors to provide more specific information about their trail fits of polynomial and Fourier surfaces and also as to whether they found the 'best' values for the parameters '*a*' and '*k*' to be the same for all the different measurements taken such as seam width, ash, calorific value, etc.

In conclusion I would like to express the hope that the interest in and research into these new approaches to ore valuation as evidenced by tonight's paper will continue to spread over an ever increasing circle of research workers to the economic advantage of our mining industry.

A. H. Munro (Associate Member): Over the last few years we have seen the flow of papers to this Institute, dealing with the application of modern computing machinery to the solution of the real problems of mine planning, increase from a mere trickle to a steady stream. The paper, which was so ably presented, is a welcome addition to this flow. One of its distinctions is that it is so very readable.

I am sure I am not alone in wishing to congratulate the authors on their ability to describe their work in such a way that it may be readily appreciated by, not only a few specialists but also, the many who may be interested in applying their results. This latter comment is not intended to denigrate the efforts of previous authors, most of whose work has been less empirical and therefore less easily described in common langauge.

There are a few points about which I feel bound to quibble and two major objections which I think must be raised. But for these points, I found the paper most stimulating and am certain that the computer programmes will prove their usefulness through continued use.

The authors reject the 'triangle method' of constructing a valuation surface, mainly on the grounds that its use involves 'the implicit assumption that the value parameters are measured with perfect accuracy and that the cores obtained from boreholes are representative of the areas nearby'. I agree whole heartedly with both their rejection and their grounds for it. However, they subsequently quote what they call a multiple correlation, and what I imagine to be a simple correlation, coefficient for their Fourier and the polynomial surfaces fitted by least squares. This coefficient is also given for their moving average surface. As the authors quite rightly state 'this co-efficient is a measure of the accuracy of the valuation surface in representing the data', that is, a measure of the extent to which the surface reproduces the actual observations. Surely the use of the measure as a means of comparing the different surfaces also involves, implicitly, the assumption which was regarded as untenable in the case of the 'triangle method'?

No alternative measure is suggested in the paper. Further, no mention is made of the criterion by which 'spurious detail' was judged to have been introduced, or 'broad features' were judged to have been smoothed out, with particular choices of the parameters '*a*' and '*k*' of their moving average surface.

A measure, which is by no means the most satisfactory, is the coefficient of correlation between the calculated surface and the underlying valuation surface which it is intended to estimate. This brings me to the first of my major objections—just what is this underlying surface? In other words what is the target? Until this target has been specified, and preferably estimated by a relatively much more precise method in at least a few test cases, surely no conclusions can be reached regarding the efficiency of the technique. A considerable amount of discussion at the 1966 Symposium on ore valuation[1] centered on this point. This discussion has continued elsewhere[4].

My second objection is to the statement that the weighted moving average technique of Krige[2], developed for gold ore valuation, 'breaks down' in the case of minerals which have more stable distribution properties than gold and therefore 'the regular grid of samples will be replaced by a much more haphazard set of borehole readings'.

The objection is multifold. In the first place, as Krige has stated, gold mine sampling is very irregular indeed in areas of interest, from the valuation point of view, despite the fact that samples are taken at regular intervals in development ends and along stope faces. The point is simply that gold mine workings are seldom regular when viewed in plan. In the second place, and more important, by suitable choice of the weighting system used by Krige in his formulation of a moving average predictor, a valuation surface practically indistinguishable from that produced with the present author's formulation could be obtained. Thirdly, the inverse square, or higher even power, law used by the authors is incapable of admitting directional differences in the patterns of variation of the value parameters of the ore body. Consequently the claim of the authors, that their moving average technique is more general than that of Krige, must be disputed.

As I see the situation, the essential difference between the present formulation of the moving average predictor and that of Krige, is that Krige has available a large number of not very informative samples whereas the number of samples available to the present authors is much smaller and each sample is assumed to throw at least some light on the value parameter at each point in the entire ore body. It is quite reasonable therefore that formulations which are little different in principle should be adopted, but that the structure of the computational routines, by which an estimated valuation surface is obtained, should differ to a considerable extent.

The costs of running the programmes developed by the authors might appear relatively low compared with those of running the programmes developed on a co-operative basis by the gold mining industry[3]. It is my opinion that this is due to a difference in scale. It would be of interest for the Bureau, which undertakes to run these programmes for gold mines, and has almost completed re-designing them, to estimate the costs of runs comparable with that which produced Fig. 4 of the paper.

In conclusion I wish to thank the authors for an interesting paper which will do much to stimulate further development of the range of ore value predictors which is needed to deal with the many different types of ore which are now being exploited locally.

REFERENCES

1. Proceedings of 'Symposium on mathematical statistics and computer application in ore valuation', *S. Afr. Int. Min. Metall.* 1966, Johannesburg.
2. KRIGE, D. G. 'Two dimensional weighted moving average trend surfaces for ore valuation', *ibid* pp 13-38.
3. PIROW, P. C. *et al.* 'Computer programmes for the estimation of a two dimensional trend surface using a weighted moving average', *ibid* pp 80-103.
4. MATHERON, G. 'Kriging or polynomial interpolation procedures', *C.I.M. Bulletin*, **60**, (1967) pp 1041-1045.

P. L. Bezuidenhout (Visitor): I had the opportunity of attending the very interesting lecture on computer mapping by means of a moving average technique, as developed by Messrs MacGillivray, Hawkins and Berjak. This method is new to me and I will be grateful if I may forward a few questions.

In the two series developed it is clear that r_i values very far from (x, y) have but little effect and there must be a criterion by which the pair of series converges as well as the predictor r. As we have a moving average, cannot moving sets of (r_i, x_i, y_i) within a specified radius from (x, y) be sufficient to determine r. This would lessen computer time especially when several hundreds of data are to be handled.

Coalfield variables can be investigated in a two-dimensional plane with a variable being the third dimension. Do the authors see any objection in using the formula proposed in three-dimensional space and having the variable as a fourth dimension. Thus trends not available in a two-dimensional plane can also contribute to the value of the predictor r.

Authors' reply to discussion

Miss Marting

We would like to thank Miss Marting for her comments. It is conceded that the computer implementation is far from complete.

Apart from the programmatic requirements mentioned by Miss Marting, it is necessary that the variables processed, source documents and input formats be standardized, and that a manual be written to explain the use and facilities of the computer programmes. This project has been constrained by the fact that it has only had the part-time attention of three investigators, and by the necessity of keeping computer usage within budgeted limits. The development mentioned above has top priority, and should be completed shortly.

Rand Mines, Limited is not committed to complete restriction on use of the programmes. The University's Economic Geology Research Unit has been given permission to use them in some of its current research projects, and the surface fitting routine has been applied to a lime deposit.

The possibility of an interface with the Salaman/Deist suite is being investigated.

Mr Margo

We are grateful to Mr Margo for his adumbration of the background to and practical significance to this research project.

Dr Sichel

We would like to thank Dr Sichel for his contribution to the discussion.

Dr Sichel points out the need for a mathematical model which will provide optimum values for a and k. As work on this model has been done by Professor Sharapov, we shall not duplicate his efforts until we have had an opportunity to study his original publication.

The authors share Dr Sichel's distaste for the multiple correlation coefficient as a measure of concordance between true and fitted surface. It should be emphasized that T is used as *a* measure, and not *the* measure of fit. As such, it is felt to be not without value. R as defined for the moving average is, in fact, an artificial quantity defined by analogy with multiple regression surfaces. If used with due care, it can provide some sort of idea of how well the surface fits the data, and give some indication of the relative merits of various fitting techniques. Of course this assessment must be regarded as a poor second-best to a correlation of the fitted and the true surfaces.

Dr Krige

We thank Dr Krige for his contribution—it elucidates many points not discussed in detail in our paper, and raises others concerning his moving average technique of which we were not aware.

Dr Krige points out that the parameters in the moving average technique used in gold mines are obtained by regression techniques. At this stage, follow-up analyses of coal mines are not available to us, and it is not possible to use regression to suggest values of a and k, or to check the moving average prediction against actual performance. This paucity of data also precludes an analysis of directional trend. If this trend were known in any detail, then one could take account of it by varying a and k according to the direction of each borehole from the point predicted.

On the question of whether the 'optimum' values of a and k are the same for all value parameters, we ourselves are very much in the dark. This might be expected *a priori* on the basis of Professor Sharapov's model. A more detailed investigation would require a large amount of computer time, and cannot be undertaken at present.

Mr Munro

The authors thank Mr Munro for his comments on the paper.

Mr Munro, too, has queried our use of R, and we refer again to the qualifications in our use of it mentioned in the reply to Dr Sichel's comments. As regards the analogy between this technique and that of Dr Krige, we must emphasize that use of this implementation of the Sharapov predictor is envisaged only in the case where the number of data is small, and the sampling irregular. It is readily conceded that if this is not the case, the predictor may well be quite unsatisfactory.

The comments on Dr Krige's technique are not intended pejoratively, but are merely to indicate why his moving average is unsuitable for the problem at hand.