HALF A CENTURY OF GEOSTATISTICS FROM A SOUTH AFRICAN PERSPECTIVE

Keynote address, 6th International Geostatistical Congress, South Africa, 2000

D.G. Krige, South Africa

ABSTRACT

The South African background to the original geostatistical contributions is covered, including reference to the detailed statistics, production and reconciliation records kept by the gold mines. These records highlighted the presence of serious conditional biases in orthodox ore block valuations and was the one major factor which, following the statistical explanation of the feature, led to the birth of kriging. Also covered are Sichel’s contributions via the introduction of the lognormal model and, in time, through the development of a flexible set of skew distribution models to cater for the complex diamond and other distributions. The global geostatistical developments in the five decades from the 1950’s to the 1990’s are covered briefly with special reference to the work by Matheron and others. This is done against the more detailed background of South African contributions and practical applications. Of concern from the South African point of view is the fact that, after half a century of phenomenal developments in geostatistics, conditional biases which gave birth to this subject, are still encountered in practical applications and are being justified for reasons which, in the author’s view, are invalid.

1. INTRODUCTION:

Geostatistics can look back on a phenomenal half a century of development, growth and achievements. Geostatistics started off with the publication of papers in various scientific journals such as the Journal of the South African Institute of Mining and Metallurgy, Annales des Mines in France, Geologie en Mijnbouw in the Netherlands, Journal of the Canadian Institute of Mining and Metallurgy, and publications of the SME in the U.S.A. and the I.M.M. in London. However, it soon found a niche in a number of International Meetings. Of these, APCOM provided the first and a regular forum from 1962 onwards, followed by the IAMG congresses, the series of International Geostatistical Congresses of which this one is the sixth, the Pribram meetings in Europe and various other congresses in various parts of the world. In due course publications of monographs and text books followed.

A major contributing factor to the rapid development of geostatistics was the remarkable speed in the spread and use of computers and of suites of geostatistical programs. Without the phenomenal increase in the speed of
computing, by a factor of some one million from 1962 to 1996, the development and routine use of many techniques, such as the kriging of massive ore resources on a routine basis and simulation, would have been impractical. After half a century, the available geostatistical literature, techniques and computer software facilities are so extensive that it is difficult to keep abreast of and retain an objective view of all the latest developments.

2. THE SOUTH AFRICAN BACKGROUND

For more than a century, South Africa has been, and still is, one of the major mining countries in the world, initially following the discovery of diamonds 130 years ago, but essentially based on gold mining in the Witwatersrand basin since the 1880’s. Like most mining operations, these gold mining operations had to be conducted from the start on a selective cut-off basis and thus called for intensive and regular sampling of development ends and of all advancing stope (panel) faces. This was and still is unavoidable because of the impracticability of any intensive core drilling for the valuation of selective mining units. Block valuations were and still are done essentially on an extrapolation basis.

This discipline of proper sampling and production records thus set the scene for the accumulation of very extensive data sets conducive to statistical analyses. This fundamental discipline is still in force today, provides follow-up data from sampling inside ore reserve blocks and an ongoing check and validation of the block valuations. It sets an example to mining enterprises world wide. This practical aspect of readily available follow-up data also accounts for the South African tendency to stress the essential practical side of geostatistics and the need for and benefits to be derived from the proper validation of all applications and techniques.

Studies of frequency distribution patterns for gold grades were initially done some 70 to 80 years ago but were unsuccessful. Real progress was absent until the late 30’s and the 1940’s when Sichel (Sichel,1947) suggested the use of the lognormal model. He was a classical statistician and in the mining field he concentrated his efforts on frequency distribution models. It was thus a natural development for him to also get involved in the more complicated and extremely skew distributions of diamonds where the definition of the upper tail representing the occurrence of small numbers of the larger and most valuable diamonds is of paramount importance. His work in this field culminated in the development of a flexible set of skew distribution models (Sichel,1972;Sichel et al,1992) with useful applications not only for diamonds but also for other minerals.

Apart from Sichel’s work, almost all other South African efforts were concentrated on practical geostatistical models and applications in gold mining with natural extensions in time to other minerals. The basic concept of the selection of ore units of a practical mining size and with grades which would be expected to be above a break-even cutoff grade was and still remains of critical importance. This
concept does not generally apply in non-mining geostatistical applications and is one of the reasons for the differences in emphasis between contributions from South Africa and other countries. The fundamental concept of selective units will, however, also apply to non-mining cases where the pattern of distribution of the variable concerned is of importance, not only globally, but also locally within relatively small parts of the whole project area or volume.

In the mining scenario the emphasis was as follows:

i) at the exploration stage with limited data available, global valuations were required of the ore to be selected above a range of cutoffs and the corresponding average grades;
ii) where more data were available at the exploration stage but still limited, some block valuations were possible but of doubtful use for specific selection purposes; and
iii) at the mining stage ore blocks had to be valued on the data then available. The efficient selection of practical mining units above the cutoff was and still has to be based on the ‘best’ use of the available data for the valuation of such individual units.

On the South African gold mines ore reserve valuations of ore blocks under scenario iii) was done regularly, at least once a year. Due to mining restrictions, the data used were confined to those parts of the periphery of each block which were exposed and thus available for sampling. However, the discipline of the subsequent regular sampling of the stope (panel) faces advancing into each block and the recording and comparison of these results with the original block valuations prevailed and thus provided a continuous and progressive set of ‘follow-up’ records amenable to statistical analyses and measures of the standard of the block valuations. Thus, block valuations were, and still are, under constant surveillance; this applies equally today to any set of sophisticated geostatistical block valuations and ensures that neither the responsible geostatistician nor management can ignore a poor standard of ore reserve estimation shown up by these records. This also provides the reason for the author being ‘obsessed’, but very necessarily so, with conditional biases where ever these still show up in ore valuations.

Scenario i) became particularly important after the 2nd world war when new mines in the Orange Free State, Klerksdorp and for West Rand fields were opened up on the basis of relatively small numbers of deep, expensive and time consuming core holes drilled from surface (Krige, 1952). It is still important today in the assessment of new projects. Scenario ii) was not directly applicable on the gold mines, but is, at present, of interest in various types of exploration projects.
3. OVERALL GEOSTATISTICAL DEVELOPMENTS OVER HALF A CENTURY

The 1950’s: Following the initial work by Sichel (Sichel, 1947, 1952) on the lognormal model this was the period of the pioneering developments in South Africa and the publication of the first two papers (Krige 1951, 1952), the implications of which were far-reaching. This work immediately aroused interest worldwide, particularly in France where, under Prof. Allais, the papers were republished in French (Krige, 1955). One of his students, later to become world renowned as Prof. Matheron, started the development of the theory of Regionalised Variables. His early works were all in French, e.g. Matheron (1957), and it was not until the 1960’s when some publications by him and his students became available in English and contributed to the spread of geostatistics worldwide (Matheron, 1963).

In South Africa, the 1950’s was a period of the use on a routine basis, for the first time, of regression estimates for ore reserve blocks (Krige, 1951, 1962a). This technique which, effectively, was the first use of what became known in the 1960’s as kriging, can properly be called Simple Elementary Kriging. The 1950’s was also the period of the establishment of the first global geostatistical model for the estimation of recoverable reserves above a cutoff grade from limited drill hole data and of a number of fundamental geostatistical concepts still valid today (Krige, 1952).

The 1960’s: During this period routine ore reserve valuations on a simplified but proper kriging basis was introduced in South Africa (a world wide first), applications were published locally, including a practical elementary introduction to geostatistics (Krige, 1962), and internationally mainly at APCOM symposia (Krige 1962b, 1963, 1964a, 1964b, 1966, 1969), the 3-parameter lognormal was introduced as an improvement on the 2-parameter model (Krige, 1960), the variogram was substituted for earlier covariance models (Krige, 1964b), anisotropic spatial structures were introduced (Krige, 1969, 1970), and other practical improvements to the various techniques were developed.

In the meantime Matheron’s work aroused interest worldwide and papers were published in English and presented at international meetings. He also introduced the term ‘kriging’ for BLUE estimates (Best Linear Unbiased Estimates) and this term was generally accepted towards the end of the 1960’s (Matheron, 1967). In the late 1960’s geostatistical knowledge was disseminated worldwide following the formation of the Fontainebleau School in France and the work of four outstanding students from there, i.e. Huijbrechts, David, Journel and Marechal, in Europe and later in Canada and the U.S.A.

The 1970’s: In this period South African applications spread from gold to copper (Krige, 1973, 1975), uranium as a byproduct from gold mining (Krige, 1979a) and to diamonds via important classical statistical developments, culminating in the Compound Poisson Model (Sichel, 1972). The basis for the
geostatistical covering of uncertainty in overall risk analyses for new mining projects was also recorded (Krige, 1972, 1979b – also 1984).

Outside South Africa, the 70’s was a period of phenomenal growth, not only of the geostatistical community as such, but of new models, e.g. disjunctive kriging (Matheron, 1975a), transfer functions (Matheron, 1975b), universal kriging (Huijbrechts and Matheron, 1977) and conditional simulation (Journel, 1974). Reference works followed (Royle, 1971; Agterberg, 1974; David, 1977; Krige, 1978; Rendu, 1978; Journel, 1978), and the first International Geostatistical Congress was held in Rome in 1975.

The 1980’s: Two more International Geostatistical Congresses were held (Tahoe and Avignon). It was also the period of the rapid expansion of the availability and speed of computers and of suites of programs for kriging, variogram estimation, and various new techniques. Of these the most notable were indicator kriging (Journel, 1982), various further simulation techniques, and also new techniques for the correction of the smoothing effect of block kriging in the estimation of recoverable reserves, e.g. uniform conditioning. Isofactorial disjunctive kriging was also introduced to handle the very skew distribution of diamonds and the problem of discrete particles. Emphasis was directed to various practical sampling problems in this regard (Kleingeld et al, 1996).

Activities in the late 1980’s were covered by Dowd (1992) and included a detailed discussion of simulation developments, indicator kriging, interval estimation and applications to hydrocarbon reservoirs and hydrology.

The 1990’s: In this period various ideas were introduced in South Africa, such as the use of ‘soft’ data in a Bayesian approach (Krige et al., 1990), co-kriging of virgin areas using limited drill hole data plus adjacent regularised data from mined out areas (Krige, 1992a), the use of macro kriging (Krige, 1992b), a detailed analysis of the massive data bases available from the gold mines, the use of the Principal Components technique for estimating anisotropical spatial structures (Krige, 1999c), direct conditioning (Assibey-Bonsu, 1999b), and a practical series of studies of the effects and implications of conditional biases in block estimates when using a limited search routine or simulation techniques (Krige, 1994, 94a, 96a, 96b, 97, 99a, 99d).

Outside South Africa there was a very significant expansion of geostatistical techniques into the fields of petroleum and environmental studies. This is evident from the papers read at the last two International Congresses (Troia and Wollongong) and also from the programme for this Congress. Geological phenomena via categorical and pluro-gaussian technique also received attention (Le Loc’h and Galli, 1996) as well as neural networks (APCOM XXIV, 1993; Dowd, 1993).
The problems of Conditional Biases have not yet been fully laid to rest and will be reviewed in the balance of this address as part of the fundamental geostatistical concepts which, from a South African point of view, have not and still do not receive the attention they deserve.

4. ORIGINAL FUNDAMENTAL CONCEPTS

The early South African contributions (Krige 1951, 1952), covered the following fundamental concepts which have not lost their relevance after half a century

i) The skew frequency distribution pattern for grade data, suitably covered in most cases by the 2- or 3-parameter lognormal model and where necessary, by the compound lognormal. The logical parallel is that confidence limits for grade estimates will also be skew.

ii) The need for a proper geological model of the mineralisation to subdivide the ore body into homogenous sub-populations.

iii) The concept of supports and the effect of changes in support sizes on the variance parameter of the distribution, leading directly to the first model for the spatial structure of data, i.e. the ‘variance – size of area’ graph, usually linear on log-log basis. This model was fitted to the Basal Reef data for the main sector of the Orange Free State field in 1952 (Krige,1952).

iv) The use of this model to estimate the frequency distribution model for the selective mining units on which the recoverable resources expected for the field, could be based. It did not take into account the information effect still present in all final block estimates which are based, invariably, on incomplete data. Nevertheless, when compared with the ‘follow-up’ underground sampling results up to 1980, was shown to be acceptable (Krige 1998).

v) The correlation between ore reserve block estimates and follow-up sampling results of advancing stope (panel) faces inside the blocks (Krige,1951). This provided the explanation for the long known but unexplained serious under-valuation of blocks valued as low grade and the serious over-valuation of blocks valued as high grade, i.e. for the conditional biases inherent in all block valuations based on limited data. It also showed that where such valuations are expanded to supplement the limited peripheral data by the incorporation of the surrounding data via the local mean grade, i.e. by using a regression or simple kriging approach, conditional biases were eliminated and the error variances of the block estimates were reduced significantly.

vi) The development under v) led directly to the fundamental concept that all block valuations which are not conditionally unbiased are sub-optimum and cannot be used for proper selective mining decisions, for detailed grade control purposes or for feasibility studies. In fact they do not meet the original requirements for BLUE estimates.

vii) Subsequent developments such as the introduction of anisotropical spatial structures and more sophisticated kriging techniques provided further improvements in the standard of block estimates. However, these
tended to supplement, but never replaced, the fundamental and major improvement gained from the elimination of conditional biases.

5. CONDITIONAL BIASES -- GENERAL

South Africa’s contributions to the birth of geostatistics originated from a practical study of conditional biases observed in the routine valuation of ore blocks for ore reserve purposes.

As mentioned before, on the gold mines current production records continually monitor the follow-up results when ore reserve blocks are mined, and thus it is impossible for a geostatistician involved in the mine’s ore valuation, to lose sight of the presence of any conditional biases. This situation also applies directly to mine management even where management is not fully acquainted with the more technical aspects of geostatistics. The reaction from the South African point of view was, thus, natural when in recent years, serious conditional biases in block valuations for mines outside South Africa were observed. Initially these were related to a practice of limiting the search routine for ordinary kriging in exploration projects with the object of avoiding the smoothing effect of kriging when a proper search routine is used. More recently serious conditional biases were also observed in routine final ore reserve estimations when the remaining smoothing is unavoidable and has to be accepted.

Conditional biases also became obvious with the advent of simulation techniques and the proposal to use one of a series of repeated simulations for block valuations, grade control, and feasibility studies. Also, the variation between repeated simulations was proposed as a direct estimate of the actual uncertainty attaching to the relevant single simulation.

6. LIMITED SEARCH ROUTINE

The two scenarios where this problem has arisen are:

i) At the selective mining stage blocks are valued for ore reserve purposes and no further data will become available for the selection of blocks above or below the cut-off. There is obviously no excuse at that stage for limiting the search routine to such an extent that conditional biases arise. Yet the author has encountered this in practice for an open cast base metal mine, where the follow-up blast hole grades showed a level of under-valuation of some 50% for blocks valued in the low grade categories, and the over-valuation of blocks valued as high grade where the actual follow-up grades were only some 50% of the block estimates. This situation prevailed throughout the property for all geologically homogenous sections and for the grades of all the impurities as well. This practice cannot be justified on any basis. This was stressed at the
panel discussion on limited search routines held during the ‘99APCOM symposium at the Colorado School of Mines (Oct.1999).

The implementation of such a practice effectively ignores all the benefits gained from the proper kriging techniques developed over half a century and places block estimations back in the position which prevailed before the advent of geostatistics. The overall effects of such a practice on the efficiency of the selection of blocks above cut-off, on grade control and on the net profitability of a mine was covered fully at the last International Geostatistical Congress (Krige,1996) and at the ‘99APCOM (Krige,1999b).

ii) At the exploration stage when the level of available data is much lower than that on which the final valuation and selection of ore blocks will be based, a limited search routine has also been advocated in order to arrive at a set of block estimates with a dispersion variance close to that which is expected for selective mining units (SMU’s) at the mining stage. This is aimed at avoiding any ‘smoothing’ and to provide a realistic grade-tonnage profile for block estimates and feasibility studies. Whereas this objective could be achieved for global estimates, the valuations of individual blocks will be conditionally biased and will provide a poor basis for the provisional classification of blocks into grade categories, and also for any initial mine planning. This will also apply to the profiles of production and profits to be achieved as mining proceeds from section to section in the ore body over time and are thus also not suitable for feasibility studies. This practice cannot be condoned and, in any case, the global estimates mentioned above can be achieved via various available techniques without any individual block estimates.


7. SIMULATION ESTIMATES

The overall effects of using simulation for estimating recoverable resources via a series of repeated simulations, the selection of one of these, and the acceptance of the spread of the repeated simulations to reflect the level of uncertainty, has been covered by a series of papers (Krige 1998, 1999a, 1999b) and by the author’s paper to be presented at this congress. From these studies it is clear that although the average of such repeated simulations can provide a reasonable profile of likely results and could be conditionally unbiased, it will approach the results of proper kriging and will, therefore, also be subject to ‘smoothing’. It, therefore, presents no real solution. Furthermore, any single simulation can be expected to be conditionally biased not only for individual blocks, but also for sections of the ore body likely to be mined, say, annually. Hence it cannot be expected to provide a profile of results suitable for planning or feasibility studies.
Also the spread of results of repeated simulations can be misleading as an indication of the uncertainty involved in using an individual simulation.

8. CONCLUSIONS:

The record of geostatistics over half a century has been phenomenal and covers a very impressive set of publications and ranges of techniques and computer programs. From a practical point of view, however, some fundamental concepts are not always observed and this should receive serious attention. The main concern is that this record will be tarnished by the all too ready acceptance of estimates which are still conditionally biased.

For the future, I would like to see geostatistics continue to grow from strength to strength with new models, techniques and applications, but where these are all validated properly by way of checks to confirm the absence of biases and the practical advantages to be gained when they are applied in practice.

9. ACKNOWLEDGEMENT:

The assistance from Drs.W.J.Kleingeld and W.Assibey-Bonsu in covering the overall developments over the half century, is appreciated.

10. REFERENCES


