Geostatistical Methods of Ore Reserve Estimation

Jean-Michel Rendu

1. Introduction

The economic value of a mineral deposit is first a function of its ore reserves. During the exploration and feasibility phases of a mining project, the need for timely and reliable ore reserve estimates is of critical importance particularly as a basis for mine planning in the early production years. This is a first step in assessing the economic merit of a deposit, and must be done to permit monitoring of the appropriateness of continued capital investment in the project.

Underestimation of reserves may result in rejection of a viable project, significant opportunity losses, and possible embarrassment if the deposit is later successfully mined by a Competing Company. Overestimation may result in the inappropriate development of a subeconomic deposit, disappointing results and significant financial losses. Poor evaluation may also lead to an undesirable mine development sequence, improper mill design, and inappropriate mining equipment selection.

Once in production, reserves and local block grades are re-evaluated periodically for medium and long-range planning perhaps annually, down to weekly or even daily for short-term planning and grade control. Poor reserve estimation will result in waste blocks being predicted as ore, while ore blocks will be treated as waste, with direct impact on the grade of anticipated mill feed and cash flow. The precision which can be attained in each evaluation will vary significantly with the information available, including sampling and mapping data as well as the current understanding of the deposit geology. Similarly the method to be used for reserve evaluation must be adapted to take into account the quantity and quality of information to be processed, as well as the purpose for which the reserve estimate is to be made.

Whatever the amount of effort put into the calculation of the reserves, some degree of uncertainty will remain concerning the true characteristics of the deposit. This is a direct consequence of the need to estimate the properties of thousands of tonnes of material from samples whose total weight may only be a few kilograms. The overall precision with which the reserves are estimated is a function of numerous factors, including the type of mineralization, the geologic complexity of the deposit, the quality and quantity of sample information available, as well as the methodology used to process this information.

Considering the complexity of the parameters which must be taken into account, and the interdisciplinary nature of the studies required, reserve estimation must be a team effort, involving geologists, mining engineers, mineral process engineers, statisticians and, usually, computer specialists. Furthermore, in all instances where the quality of the reserve estimates is likely to have a significant influence on the financial and economic health of a mining company, high level executive
attention should be given to the study to ensure not only that the correct answer is obtained, but that its meaning and its basis is also fully understood.

The purpose of this paper is to present the principal statistical and geostatistical techniques of ore reserve estimation. To fully appreciate both the power and the limitations of these techniques, the preceding remarks concerning ore reserve estimation should be kept in mind throughout.

Traditional statistical methods of reserve evaluation do not take into account the physical position of the samples with respect to each other. They are extremely useful in the early analysis of a deposit, and can be used to verify and interpret the sample values available, as well as to obtain preliminary reserve estimates. Geostatistical methods, on the other hand, are based on the common observation in mineral deposits that samples taken close to each other are more likely to have similar values than if they are located far apart. For this reason they are particularly well adapted to detailed reserve evaluations, when relatively dense sample information is available.

Statistical and geostatistical methods, take into account the variability in an orebody's mineralization, and can be used to quantify the uncertainty associated with the reserve estimates. As with any other reserve estimation procedure, considerable judgment is needed in the use of geostatistics but the methods are easily adapted to take into account the specific properties of each deposit, the nature of the sample information available, the degree of continuity present in the mineralization, as well as the geologic controls which may influence the ore distribution.

In this paper the emphasis is on the practical aspects of applied geostatistics. Mathematical equations have been kept to a minimum. For more information on theoretical geostatistics, the reader is referred to the publications given as references, and more specifically, to the test books published by RENDU (1981), DAVID (1977), JOURNEL and HUIJBREGTS (1978), and MATHERON (1971).

2. A Historical Overview

Early attempts to apply statistical methods to the estimation of mineral deposits can be traced back to the turn of the century. Mathematical statistics was still in its infancy, computing tools were not readily available to process large amounts of data, and these early studies had no significant impact on the industry. Maybe the first significant breakthrough occurred in the late 1940's and early 1950's with work by H. J. Wijs, H. S. Sichel, and D. G. KRIEGE on the South African gold fields. The lognormal distribution of gold values was recognized at this time and statistical methods were developed accordingly to improve estimation of new gold deposits.

Practical grade control problems, including the well documented overestimation of high grade zones and underestimation of low grade zones, were statistically analyzed. In the 1950s, multivariate linear regression was successfully proposed by D. G. KRIEGE as a solution to this problem, a solution which led to the development of what is now known as the kriging method of reserve evaluation.

The 1960's saw the development by G. MATHERON and his disciples of a formal mathematical theory to statistically analyze spatially distributed variables. The theory was developed to explain or justify results and observations made in the South African gold fields. The semivariogram was proposed as the main tool to be used to quantify the statistical properties of mineral deposits, and kriging became the accepted geostatistical method for reserve evaluation. The theory was successfully applied to a number of deposits.

The ground was thus prepared for the explosive development in both applied and theoretical geostatistics which occurred in the late 1960's and throughout the 1970's. Significant progresses were made in the theory of geostatistics, with the development of new and increasingly complex mathematical models by the Centre de Morphologie Mathematique in Fontainebleau, France. Numerous books and articles were published on the subject. Courses in geostatistics were offered in many mining
schools worldwide. The public image of geostatistics changed, from a very complex tool to be used only by mathematicians and computer specialists, to an acceptable means of analyzing and evaluating deposits. All of these changes were greatly facilitated by the growing availability of the electronic computers needed to process large amounts of information, and by a rapidly expanding mining industry.

The 1980's seem to herald new challenges for geostatisticians. Even though progresses will continue to be made in theoretical geostatistics, it is the author’s opinion that the main developments will be in this as yet poorly understood zone where geology, mining engineering, and geostatistics overlap. Increasingly careful attention will be given to the interdisciplinary nature of applied geostatistics. The ability of statistical models to fully take into account geologic controls will be increasingly scrutinized. Once again, progresses made in the field of computing with the development of extremely powerful microcomputers will have a significant impact on geostatistics. Mine site geostatistics will become commonly accepted, as an integrated tool for exploration, grade control, reserve evaluation, and short term mine planning. The emphasis will be on making geostatistics easily useable by geologists and mining engineers. Geostatistics originated 40 years ago as a practical solution to practical mining problems. Huge progresses have been made since then in the development of theoretical geostatistics. It seems that the next decade will be characterized by a return to the practical aspect of geostatistics.

3. Statistical Methods

3.1 Introduction

Traditional statistical methods are based on the assumption that all sample values are equally representative of the properties of the deposit under study. The physical position of the samples with respect to each other is not taken into account. For this reason, application of these methods to reserve evaluation is usually limited to the early stages of orebody analysis, when a global estimate is sufficient; or to the study of highly variable mineralizations in which the similarity between sample values is negligible, even at short distances. Statistical methods are also extremely useful to test the reasonableness of estimates obtained using more sophisticated approaches, such as kriging. Many of the visual and mathematical tools described here can also assist in the analysis, verification and interpretation of sample data and in this respect they may be found useful throughout the life of a mining project.

Perhaps the easiest and fastest method of estimating the average grade of a deposit, or part of a deposit, consists of calculating the average value of the samples in the area of interest, ignoring their relative position. As an example, if a coal seam is intersected by 10 drill holes and the seam thickness is measured in each hole, the average thickness of the seam can be estimated by the mean value of the 10 drill-hole intercepts. If the drill holes are approximately uniformly distributed, this estimate is likely to be perfectly acceptable. Simple statistical formulae can then be used which give reasonable preliminary estimates of the coal tonnage in the deposit and the precision with which it is known.

If on the other hand, we consider an uranium vein which has also been intersected by 10 drill holes, and want to estimate the average grade times thickness (uranium accumulation measured in meter x%U₃O₈), the average value of the 10 drill holes is likely to give a very poor estimate of the vein properties. Indeed, the uranium accumulation may vary from a minimum of say 0.01%U₃O₈ over 0.5 meters (0.005 meter x%U₃O₈) to a maximum of say 1.00%U₃O₈ over 1.5 meters (1.5 meter x%U₃O₈), and the mean value of the 10 drill holes will be extremely sensitive to the presence of extreme values, especially extremely high values. In such instances where a very high level of variability is observed, more complex statistical methods may be required.

3.2 Visual Methods of Statistical Analysis

The first step in a statistical analysis of a mineral deposit involves plotting the histogram of the sample values. Examples of histograms are given on Figure 1 for a coal
Fig. 1 Examples of histograms from a coal seam. These curves show the range of sample values which can be expected. They can be used to detect extreme, possibly erroneous values. The values which are occurring most often can be visually determined. Histograms can also help to detect mixtures of geologic environments (Fig. 2).

From the histogram, the frequency distribution can be calculated where the number of samples falling in a given class interval is replaced by a proportion given as a percentage of the total number of samples. Consider the heat content of coal samples listed in Table 1. The corresponding frequency distribution is given in Table 2. This table shows, for example, that 18.37% of the samples have a value between 9600 and 9700 BTU/lb. If the samples are representative of the actual coal seam, this distribution also indicates that 18.37% of the entire deposit has a value between 9600 and 9700 BTU/lb. Extreme care must, however, be applied when using sample statistics to determine the economic properties of deposits which can only be mined using panels or blocks of size significantly larger than the samples. The influence of block size and selectivity on the grade-tonnage relationship is discussed later in this paper. Sample statistics will give reasonable estimates when applied to deposits with relatively low variability such as coal or industrial minerals. At best, an order of magnitude estimate will be obtained if highly variable values are studied, such as gold or uranium grades.

Another tool which is useful in the statistical study of sample values is the cumulative frequency distribution, which indicates the proportion of the samples with value less than a given upper limit. Looking again at Table 2, one sees that 34.69% of the samples have a

<table>
<thead>
<tr>
<th>Heat Content (BTU/lb)</th>
<th>Frequency Distribution (Proportion of Samples in Class, %)</th>
<th>Cumulative Frequency Distribution (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>9000</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>9100</td>
<td>9100 1</td>
<td>2.04</td>
</tr>
<tr>
<td>9200</td>
<td>9200 2</td>
<td>4.08</td>
</tr>
<tr>
<td>9300</td>
<td>9300 3</td>
<td>6.12</td>
</tr>
<tr>
<td>9400</td>
<td>9400 4</td>
<td>12.24</td>
</tr>
<tr>
<td>9500</td>
<td>9500 5</td>
<td>18.37</td>
</tr>
<tr>
<td>9600</td>
<td>9600 6</td>
<td>24.69</td>
</tr>
<tr>
<td>9700</td>
<td>9700 8</td>
<td>34.69</td>
</tr>
<tr>
<td>9800</td>
<td>9800 8</td>
<td>53.06</td>
</tr>
<tr>
<td>9900</td>
<td>9900 8</td>
<td>69.29</td>
</tr>
<tr>
<td>10000</td>
<td>10000 2</td>
<td>85.71</td>
</tr>
<tr>
<td>10100</td>
<td>10100 2</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 1 List of coal sample heat contents (BTU/lb).

Table 2 Distribution of heat contents in a coal seam.
Fig. 3 Cumulative frequency distribution of heat content in coal seam.

Fig. 4 Examples of scatter diagrams

Fig. 5 Scatter diagrams indicating errors and mixture of geologic environments.

heat content of less than 9600 BTU/lb. This also means that 65.31% of the sample values exceed 9600 BTU/lb. Again, if the samples are representative of the deposit, one may be able to conclude that 65.31% of the entire deposit will exceed 9600 BTU/lb.

Cumulative frequency distributions are usually plotted using a probability scale along the horizontal axis, as illustrated on Figure 3. The probability scale is such that if the sample values are approximately normally distributed, the cumulative frequency distribution will be well represented by a straight line. The cumulative frequency distribution calculated in Table 2 is plotted on Figure 3.

Another graphical representation of the distribution of sample values which is extremely helpful in understanding the properties of a deposit is the scatter diagram. A scatter diagram can only be plotted if at least two values are available for each sample. Scatter diagrams for a coal deposit are shown on Figure 4. These diagrams show the relationship between the different coal qualities. Scatter diagrams can also be used to detect anomalous values, as well as mixtures of geologic populations (Fig 5).

3.3 Precautions to be Taken in Interpreting the Results

The use of graphical methods for the preliminary analysis of sample values is usually extremely instructive. One should, however, remember that such methods can give meaningless results unless the geology of the deposit is first taken into consideration to classify the sample values. The spatial distribution of the samples and drill holes must also be kept in mind when interpreting the results. For example, frequency distributions may be meaningless if the drill holes analyzed are located on a 50-meter grid in the high grade part of the deposit and a 200-meter grid in the lower grade parts of the deposit. Finally, all statistical analyses require that the values analyzed correspond to samples, composites or...
blocks of constant size. Using geostatistical terminology, it will be said that the values must correspond to supports of constant size.

The concept of support size is extremely important in all statistical studies of geological data. This is clearly seen on Figure 6 which shows the histogram of sample values in a tin deposit. The sample length varied from 3 centimeters in high grade cassiterite veins to 9 meters in poorly mineralized areas. The grade of these samples is shown on the right-hand side, while the average grade of 3-meter long composites is on the left-hand side. The highest sample assay is 2.0% Sn over 3 cm, while none of the 3-meter composites exceeds 1.0% Sn.

A numerical example of the influence of support size is shown on Figure 7. Channel samples of one meter length have been taken along a drift in a high-grade copper deposit. The values of the one-meter samples vary from

![Fig. 6 Comparison of samples of variable size, with fixed length composites.](image)

![Fig. 7 Influence of support size on the frequency distribution](image)
1% Cu to 15% Cu, while 8 m composites remain between 5.25% Cu and 8.25% Cu. This figure shows the effect of the support size on the frequency distribution. Typically, when the support size increases, the values get closer to each other and, therefore, closer to the mean of the orebody. Similarly, if longer or larger samples are used to analyze a deposit, the variability of the corresponding values is likely to be reduced, unless an increase in assaying errors counters the effect of the decreased geologic variability.

### 3.4 Normal Distributions

The methods described in this section apply to normally distributed sample values, i.e. values whose frequency distribution is symmetrical around the mean and can be plotted as a straight line on probability paper (Fig. 3). These methods are also commonly used when the sample values are only approximately symmetrical, but not necessarily normally distributed (Fig. 1). Other conditions of application are discussed later. Examples of deposits to which these methods can be applied include iron, coal, potash, phosphate and industrial minerals as well as some high grade base metal orebodies. Seam or vein thicknesses can also often be treated in the same fashion.

If \( n \) samples have been taken from a mineral deposit, with values \( x_1, x_2, \ldots, x_n \), an estimate of the average value of the deposit is the sample mean \( m \):

\[
m = \frac{x_1 + x_2 + \ldots + x_n}{n} \quad (1)
\]

For example, if 5 drill holes intersect a coal seam indicating coal thicknesses of 5, 2, 7, 4 and 3 meters respectively, the average thickness of the deposit can be estimated at 4.20 meters. This number is, however, only an estimate of the true average coal thickness which could be thicker (say 5 meters) or thinner (say 3 meters). The degree of certainty with which the deposit characteristics are known is a function of the number of samples available, the variability in the sample values, and of whether the sample values are truly representative of the deposit.

If \( \mu \) is the unknown average thickness of the deposit, \( m \) is an estimate of \( \mu \). The difference between \( m \) and \( \mu \) represents the error made in estimating the deposit parameter in question. This difference is unknown, but statistical methods can be used to quantify its likely magnitude. The expected value of the squared difference between \( m \) and \( \mu \) is known as the **error variance** \( \sigma^2_E \):

\[
\sigma^2_E = \text{expected value of } (m - \mu)^2 \quad (2)
\]

The square root \( \sigma_E \) is the (standard) error of estimation of \( \mu \). The magnitude of \( \sigma_E \) gives an indication of the precision with which the mean is estimated. In many circumstances it is reasonable to state that there are two chances in three (68% probability) that the true value \( \mu \) is greater than \( m - \sigma_E \) and less than \( m + \sigma_E \). There are also approximately 19 chances in 20 (or 95% probability) that the true value of the deposit exceeds \( m - 2\sigma_E \) and is less than \( m + 2\sigma_E \).

The error variance can be estimated as follows:

\[
\sigma^2_E = \frac{s^2}{n} \quad (2')
\]

where:

\[
s^2 = \frac{(x_1 - m)^2 + \ldots + (x_n - m)^2}{n-1}
\]

Using the coal seam example given above, we calculate \( n = 5 \), \( s^2 = 3.70 \) and \( \sigma^2_E = 0.74 \) or \( \sigma_E = 0.86 \) meter. Therefore if \( \mu \) is the unknown average thickness of the deposit,

- Prob \( (3.34 \text{ m} < \mu < 5.06 \text{ m}) = 68\% \)
- Prob \( (2.48 \text{ m} < \mu < 5.92 \text{ m}) = 95\% \)

If the sample values reasonably well approximate a normal distribution, the confidence intervals can be calculated with increased precision, using the Student’s-\( t \) statistics, a method described in detail in most texts on statistics or geostatistics (David, 1977; Rendu, 1981).

The above formulae can generally be applied to most deposits, independent of whether or not the sample values are normally distributed provided the number of samples is large enough, such that the error of estimation \( \sigma_E \) does not exceed 20% of the mean.

### 3.5 Lognormal Distributions

In many more deposits the sample distribution is not symmetrical, but is characterized by a longer tail towards the high values (Fig. 8): this is known as a positively skewed distribution. This type of distribution is extremely common among lower grade deposits, characterized by a high variability in sample values. When sample values have a very skewed distribution, the simple statistical
methods discussed earlier in the context of the normal theory may give meaningless results.

In many instances, as illustrated on Figure 9, the frequency distribution of the sample values will plot as a straight line on logarithmic probability paper. The distribution is then said to be lognormal and the logarithm of the sample values is normally distributed. Logarithmic probability paper is characterized by a probability scale on the horizontal axis and a logarithmic scale on the vertical axis. The data plotted on Figure 9 were those listed on Figure 8. Also shown on Figure 9 are graphical methods of analysis of lognormal distributions which are discussed by KRIGE (1978) and RENDU (1981). The lognormal distribution of sample values was first recognized by Dr. SICHEL in the South African gold fields (SICHEL, 1952).

In some instances, the lognormal assumption is not exactly satisfied and the cumulative frequency distribution shows a downward curvature when plotted on lognormal probability paper (Fig. 10). Dr. D. G. KRIGE showed (KRIGE, 1960) that a lognormal distribution can often be obtained if a constant is added to the sample values. The distribution is then said to be three parameter lognormal. The three parameters are the mean, the logarithmic variance, and the additive constant.

Statistical methods have been developed by Dr. SICHEL (1966) to estimate deposits within which the sample values have a lognormal distribution. The mean value of a deposit is estimated using what is commonly known as the SICHEL's t-estimator. This estimator, and how to calculate the corresponding confidence limits, is discussed in numerous publications,
Geostatistical methods

(a) PORPHYRY COPPER GRADE DISTRIBUTION

Fig. 11 Examples of cumulative frequency distributions. Including Sichel (1966), Rendu (1981), David (1977), and Wainstein (1975).

Lognormal theory was first successfully used in the 1950's in the evaluation of South African gold mines, where it was commonly necessary to achieve accurate estimation of new deposits from a small number of drill hole intercepts. It has been used extensively since then in many types of deposits, including disseminated and vein gold, molybdenum, tin, sulfur in coal deposits, bismuth in zinc deposits, lead, copper, zinc and uranium (Fig. 11).

Eventhough the distribution of the sample values often well approximates the two or three parameter lognormal model, significant departure from these relatively simple models is not unusual. Such departures may be indicative of a number of factors, but in all instances, simple explanations should first be sought for complex value distributions. An irregular sampling density will often result in an excessive representation of the high-grade portions of the deposit. A mixture of geologic environments may produce apparently abnormal distributions. The presence of clay zones in the middle of a disseminated gold deposit may result in an excess of low-grade values. Secondary enrichment may, on the other hand, result in an excess of high-grade values. If more than one mineralizing event occured, it may be possible to see it in the shape of the frequency distribution. Changes in the sampling density can be taken into account by using declustering techniques, which consist in weighting the sample values in inverse proportion to the drill hole density. If mixed geologic zones are suspected, additional geologic analysis and partitioning of the deposit may be required before meaningful statistical results can be obtained.
4. The Foundations of Geostatistics

4.1 The Geostatistical Approach

The statistical methods described earlier permit the estimation of the average value of a deposit parameter and the calculation of confidence limits for this value. They are based on the very important assumption that sample values are independent of one another and are equally representative. As soon as a reasonable number of samples has been obtained from the deposit the assumption of the independence of these samples becomes invalid and represents in fact an ignoring of an important portion of the knowledge gained during the sampling program i.e. the relationships between the sample values and their relative positions. Whatever method is used for the valuation of a block of ore, the following assumptions are always made:

$\frac{1}{n} \sum_{i=1}^{n} (x_i - x_i')^2$

where:
- $n$ = number of pairs of samples a distance $h$ apart
- $x_i$ = value of first sample in $i$-th pair
- $x_i'$ = value of second sample in $i$-th pair

These assumptions will hold true only if the following condition is also satisfied, as it is in virtually every known type of mineral deposit:

§ there exists a relationship between sample values which is a function of their distance apart.

The geostatistical method of ore reserve estimation is distinguished from other methods available by its inclusion of this factor. At the outset, it involves estimating the spatial relationship that exists between the known sample values and from that, deriving a model of this relationship for the orebody or some zone within it. This model is then used to calculate the relationships between the known sample values and the unknown block value and thus to produce the best possible estimate of this block value.

4.2 The Semivariogram

4.2.1 Definition

Once a geologic understanding of the deposit has been obtained and preliminary non-spatial statistics have been completed, more complex statistical methods, which take into account the relative position of the sample values, must be used.

It is obvious that samples located a few feet apart are likely to indicate similar properties of the mineralization, and that this similarity would be expected to decrease when the distance between samples increases. In some deposits, it may be possible to define a distance of influence beyond which the similarity between sample values becomes negligible. However, this distance may vary significantly between deposits, between different parts of the same deposit or even between different directions at the same point in the deposit. This degree of similarity or dissimilarity between values can be quantified and can be of great assistance in understanding the properties of the deposit and optimizing its evaluation.

The simplest method of comparing two sample values is to calculate the difference between them: the smaller the difference the more similar the values. Generally, when the distance between samples increases, the difference between values increases. Rather than working with differences between sample values which can be positive or negative, the geostatistician uses the average squared difference, also known as the variogram. The value of the semivariogram $\gamma (h)$ for a distance $h$ between samples is equal to one-half the average squared difference between sample values. The one-half factor is used such that the values of the semivariogram are equivalent to variances.

The general formula for calculation of the semivariogram is:

$$\gamma (h) = \frac{1}{2n} \sum_{i=1}^{n} (x_i - x_i')^2$$

where:
- $n$ = number of pairs of samples a distance $h$ apart
- $x_i$ = value of first sample in $i$-th pair
- $x_i'$ = value of second sample in $i$-th pair

This definition is best explained through an example:

A bedded deposit has been intersected by 10 drill holes located on a straight line as shown in Figure 12. Most drill holes are 1,000 feet a-
part and the thickness of the deposit has been measured at each point. The semivariogram value for a distance of 1,000 feet between holes is calculated as follows:

\[
\gamma(1000) = \frac{1}{2} \times \left[ (5-5)^2 + (5-7)^2 + (12-11)^2 + (11-8)^2 + (8-7)^2 + (7-2)^2 \right] / 7 = 2.86 \quad (4)
\]

For a distance of 2,000 feet between holes the following results are obtained:

\[
\gamma(2000) = \frac{1}{2} \times \left[ (5-7)^2 + (7-12)^2 + (12-8)^2 + (11-7)^2 + (8-2)^2 + (2-3)^2 \right] / 6 = 8.17 \quad (5)
\]

A similar calculation for 3,000 feet gives a result of 15.67. These results are plotted on Figure 12.

4.2.2 Calculations

One dimensional

The example given in the previous section illustrates how the semivariogram can be calculated in a one-dimensional environment. If channel samples were taken at regular intervals along parallel drifts, the semivariogram will be calculated within each drift as indicated in the above example and then averaged between drifts. If a number of vertical drill holes were drilled and samples of equal length and diameter were taken and assayed using identical methods, an average downhole semivariogram will be calculated in the same fashion. If the samples are not of equal length, they should be first grouped or divided to form composites of equal length. As indicated earlier, statistics calculated on values corresponding to supports of variable size are usually meaningless. If a number of drill hole sizes, sampling methods or assaying methods were used, each data set should be statistically analyzed separately. If the statistics indicate no significant difference, it may be possible to combine different types of data in the geostatistical analysis.

In many instances, it may be necessary to use sample values which are not strictly comparable. This is commonly the case when a deposit has been analyzed over a number of periods, by different operating companies using different types of equipment. Minor differences such as in drill hole diameters can often be ignored. But chip samples taken along a drift should not be combined with core samples taken ahead of drifting.

Two Dimensional

When a vein or bedded deposit is analyzed, the values of interest can often be treated two-dimensionally. The information supplied by each drill hole intercept is reduced to an ore thickness and an average grade between hangingwall and footwall. The grade values can be analyzed directly or alternatively the product of grade times thickness (the accumulation) can be calculated. The thickness analyzed may be either the true thickness or the thickness measured normal to a reference plane, this being either a horizontal or vertical plane or a dipping plane parallel to the deposit.

The method commonly used to calculate the semivariogram in a twodimensional environment is illustrated on Figure 13 and uses an
Fig. 14 Drill hole location map and drill hole values—test case orebody.

Fig. 15 Cumulative frequency distribution of test case data.

Table 3 Experimental semivariogram of test case data.

<table>
<thead>
<tr>
<th>Direction (Azimuth)</th>
<th>Number of Pairs</th>
<th>Ave. Distance (meter)</th>
<th>Semivariogram</th>
</tr>
</thead>
<tbody>
<tr>
<td>North-South</td>
<td>11</td>
<td>11</td>
<td>14.59</td>
</tr>
<tr>
<td>(0°±22.5°)</td>
<td></td>
<td>117.39</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>175.19</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>231.65</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>20.77</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.75</td>
<td></td>
</tr>
<tr>
<td>Northeast-Southwest</td>
<td>3</td>
<td>60.84</td>
<td>14.59</td>
</tr>
<tr>
<td>(45°±22.5°)</td>
<td></td>
<td>116.58</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>188.96</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>262.52</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>332.38</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>430.50</td>
<td></td>
</tr>
<tr>
<td>East-West</td>
<td>8</td>
<td>52.77</td>
<td>14.59</td>
</tr>
<tr>
<td>(90°±22.5°)</td>
<td></td>
<td>100.55</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>188.43</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>257.40</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>324.14</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>404.56</td>
<td></td>
</tr>
<tr>
<td>Northwest-Southwest</td>
<td>2</td>
<td>58.31</td>
<td>14.59</td>
</tr>
<tr>
<td>(135°±22.5°)</td>
<td></td>
<td>114.29</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>180.47</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>264.92</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>320.48</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>399.05</td>
<td></td>
</tr>
</tbody>
</table>

**angular search.** The value of the semivariogram is to be estimated in a direction α for a distance h between samples. If the reference plane is horizontal, α will represent the azimuth. Two samples with value x₁ and x₂ will be considered in this estimation, provided the direction defined by these two samples is greater than α-δα and less than α+δα, and the distance between samples is greater than h-dh and less than h+dh. δα and dh are the tolerance angle and distance which should be kept small whenever possible.

Consider the 25 sample values given in Figure 14. Their cumulative frequency distribution is shown on Figure 15. It differs only slightly from a normal distribution. The semivariogram has been calculated in the following directions: north-south, northeast-southwest, east-west, northwest-southeast. An angular tolerance of 22.5° was used. The distances considered were grouped in classes of 75 meters. The results obtained are given in Table 3. For example, in the north-south direction, 11 pairs of samples were found, whose distance apart is between 0 and 75 meters. The average distance between these samples was 59.32 meters. The calculated value of the semivariogram was 14.59. These results are plotted on Figure 16 and 17. Figure 16 is a plot of γ (h) as a function of the distance h. It shows that the lowest level of variability is in the north direction, while the highest level of variability is in the east or northwest direction. In all directions, the semivario-
Fig. 16 Directional semivariograms in test case orebody - one-dimensional representation.

Fig. 17 Directional semivariogram in test case orebody - two-dimensional representation.

Fig. 18 Calculated semivariogram and fitted model in principal directions - test case orebody.

Geostatistical methods

Three dimensional

Consider now a large porphyry copper orebody sampled by 200 vertical drill holes. Typically, if the deposit is to be mined using 40-foot benches, the samples will be grouped in 40-foot long composites. Within reasonably homogeneous geological environments, an average downhole semivariogram will be calculated using the one-dimensional approach discussed earlier. Also, horizontal semivariograms will be calculated within each bench and averaged.
between benches. This preliminary semivariogram study will be used to determine the direction of best continuity in the horizontal plane and to compare the horizontal and vertical directions. The results obtained may be useful in the development of exploration drilling, mine design, and grade control strategies. Unfortunately, a semivariogram study which considers only horizontal and vertical directions is not based on any geological foundation and can lead to incorrect conclusions. Except for a small number of significant exceptions, geological controls are not preferentially directed in a horizontal fashion.

The method most commonly used to calculate semivariograms in a three-dimensional environment is based on a conical search (Fig. 19-a). The direction of interest is defined by an azimuth and dip. A tolerance angle is specified around this direction and the pairs of samples are grouped according to distance. The directions to be considered are chosen taking known geologic controls on the mineralization into account, wherever possible. The conical search is particularly helpful in large disseminated deposits, such as porphyries. Its main drawback, however, is that it usually requires considerable computer resources. It is also poorly adapted to the analysis of stratified, bedded or fault related deposits, or any type of deposit in which the controls of mineralization are essentially planar. For example, consider an orebody whose economic mineralization is fracture controlled and the major fracture zones are striking east and dipping 60° north. The directions to be considered should include:

Fig. 19 Three dimensional search

Fig. 20 Three dimensional anisotropy in massive sulfide deposit.

§ a number of directions within or parallel to the fracture plane
§ one direction normal to the fracture plane orientation, i.e., dipping 30° south

The anisotropy of the semivariogram will then be defined by three principal directions:
§ the direction of best continuity which would normally correspond to the plunge direction within the fracture plane orientation.
§ the direction normal to plunge within the fracture plane.
§ the direction of least continuity which is typically normal to the fracture plane.

A method of semivariogram calculation particularly well adapted to the analysis of mineralization with planar controls is the "slice" window search illustrated on Figure 19-b. The method initially involves preparation of the data so as to assign the bench composites to parallel slices through the deposit, whose strike and dip is user-controlled and is chosen as a function of the known geology. The semivariogram is then calculated within these inclined slices. This approach is only applicable to mineralization with planar control. If there is only a limited geologic understanding of the deposit the method can be used to determine the principal directions of continuity thus assisting in the geologic interpretation. The method has the additional advantage of involving minimal computer time compared with the conical
4.3 Semivariogram Modeling

Once a semivariogram has been calculated, it must be interpreted by fitting to it a mathematical formula or "model" which will help to identify the characteristics of the deposit and yield numerical parameters which describe the deposit's continuity. Examples of such models are shown in Figure 21. To the practiced eye, each model has a particular significance.

The "general model" in Figure 21 shows the main parameters derived from a semivariogram model, namely:

- The **sill**, which shows the highest level of variability measured by the semivariogram. Some semivariograms do not have a sill. If a sill exists, its value is usually close to but higher than the sample variance.

- The **range** is the distance at which the semivariogram plateaus or reaches the sill value and represents a measure of the maximum distance of influence of a drill hole in the direction concerned. Beyond this distance, sample values are independent of one another. Some semivariograms do not have a range.

- The **nugget effect** is the value of the semivariogram at zero distance. It represents the sample variability at small distance caused by small scale geologic controls. It also gives an important indication of the presence and magnitude of sampling and assaying errors.

The most common single model is the **spherical model** which has the equation:

\[
\gamma(h) = \begin{cases} 
N + C & \text{for } h < a \\
N + C \left(\frac{3h}{a} - \frac{1}{2} h^3/a^3\right) & \text{for } h \geq a
\end{cases}
\]

where:
- \(C\) is the sill
- \(N\) is the nugget effect
- \(a\) is the range, and
- \(h\) is the distance lag.

Other models which are commonly found include (Fig. 21):

- The exponential model, which does not have a range but reaches a sill asymptotically.
- The linear model, which indicates that the variability is directly proportional to the distance.
- The parabolic model, which indicates a linear drift or trend and a high level of continuity in the sample values. It will be observed if there is a systematic linear increase or decrease in values.
- The Gaussian model, which behaves like the parabolic model for short distances but plateaus at large distances.
- The "hole effect" mode, which is indicative of periodicities in the mineralization.

In addition, the **de Wijsian model** should be mentioned, as it has great historical significance as the first semivariogram model recognized (MATHERON, 1962) (KRIGE, 1978).

Most semivariograms can be represented by the sum of a number of submodels, including a nugget effect, a short-range submodel and a long-range submodel. With the exception of
the parabolic, gaussian and hole effect models, a combination of spherical submodels can be used to represent most other calculated or so called experimental semivariograms. Spherical submodels were used on Figure 18 and 20.

Semivariograms which are best represented by a parabolic model are often obtained when studying the thickness of highly continuous strata bound deposits. If the continuity is only over limited distances the gaussian model may be applicable. A gaussian model with nugget effect was used to represent the semivariogram of a coal seam shown on Figure 22. The model indicates high continuity in the east-west and northeast-southwest directions, and a high level of variability in the north-south and northwest-southeast directions.

The hole effect semivariogram can be used to represent and quantify periodicities in sample values. Figure 23 shows the semivariogram obtained in a fracture controlled deposit with a number of vertical fracture zones oriented in a north-south direction. A hole effect is observed in the east-west direction, normal to the ore zones. The fitted model shows that the average distance between fracture zones is 240 meters. If drilling is intended to discover new fracture zones for exploration purposes, a possible drill hole spacing in the east-west direction is 240 meters. An equal or larger spacing can be used in the north-south direction. On the other hand, if drilling is intended for grade control and mine planning, the spacing between holes should reflect the short-range variability, with a possible spacing of 40 meters in the direction of the ore zones and 15 meters in the perpendicular direction.

As illustrated in the previous examples in two or three-dimensional environments, different semivariogram models may be applicable depending on the direction considered, thus reflecting anisotropies or changes in continuity in the mineralization. Two types of anisotropies can be found, depending on the geologic environment considered. Zonal anisotropies are observed in bedded or stratified deposits. The variability within a stratum may be low and remain low for very large distances, while the variability between strata may be much larger, even at short distances. Zonal anisotropies can also be present in deposits where strongly directional geologic controls have influenced the mineralization. If a zonal anisotropy is present, the semivariograms will have different sills, depending on the directions considered. Figure 22 shows a zonal anisotropy which resulted from the presence of erosion channels in a coal seam.

A more common form of anisotropy is the geometric anisotropy. If such an anisotropy is present, the sill remains the same in all directions, but the range is variable. A typical example was shown on Figure 20. Often the range in any direction can be represented by the radius of an ellipse or ellipsoid. If nested structures or structures "en echelon" are present, neither the direction nor the ratio of anisotropy need be the same for each structure. An example is illustrated on Figure 24: the ore
Fig. 24 Two-dimensional semivariogram showing structures en-echelon.

Fig. 25 Logarithmic semivariograms in two zones of a precious metal deposit.

Pods are oriented in a roughly north 20° east direction while the mineralized areas containing these pods are elongated in a north-south direction.

4.4 Practical Considerations

As for all statistical studies, the semivariogram analysis will give meaningful results only if a number of so-called stationarity constraints are satisfied. Stationarity first means geological homogeneity. A geostatistical study should first start with a thorough review of all geological information concerning the deposit, including rock types, folding, faulting, zones of primary and secondary enrichment, and a model of the genesis of the deposit if available.

In most situations, mixing of geologic environments will result in meaningless or erroneous geostatistical results.

In a porphyry copper deposit for example, a sulfide zone should not be mixed with an oxide zone. The mineralization may differ with rock type as well as with the intensity of that mineralization. The significance of geologic discontinuities such as faults on the semivariogram should be obvious. Figure 25 illustrates the influence of changes in geologic environment on the semivariogram in different sections of a precious metal deposit.

The influence of the sample or support size on the semivariogram is similar to the influence of sample size on the sample variance: the larger the sample size, the lower the sample variability. This is illustrated on Figure 26 where the semivariograms of 2-meter and 4-meter composites are compared. The decrease in variability is known under the name of regularization and can be calculated from the semivariogram model (RENDU, 1981).

Differences in sampling methods will also result in differences in the semivariogram. On Figure 27 the results obtained in a tin deposit are illustrated. Two types of samples were analyzed: core samples and bulk samples taken from successive rounds along horizontal drifts. The semivariograms shown on Figure 28
Fig. 27 Comparison of semivariograms in a tin deposit.

Fig. 28 Comparison of diamond drill hole and channel sample semivariograms in a silver vein deposit.

A number of methods are used to take into account the proportional effect, the applicability of which will vary depending on the deposit under study. These methods include:

§ Division of the deposit into "grade zones" or zones of fairly uniform grade, within which the proportional effect can be ignored.

§ Calculation of a relative semivariogram. The relative semivariogram is \( \gamma(h)/m^2 \) where \( m \) is the mean value of the samples used to calculate each point of the semivariogram \( \gamma(h) \).

§ Calculation of the logarithmic semivariogram. If the sample values are approxi-
approximately lognormally distributed, the semivariogram of the logarithms of the sample values will be independent of the sample grade.

Figure 30 shows the average downhole semivariogram calculated in a precious metal deposit; first using untransformed sample values, then using the logarithms of the sample values. This figure dramatically illustrates the impact of the logarithmic transformation of the data prior to semivariogram estimation.

4.5 Some Applications of the Semivariogram

The semivariogram study is one of the most, if not the most, important parts of any geostatistical study. The semivariogram is a prerequisite to the use of the kriging method of ore reserve evaluation. In addition, the semivariogram may supply information concerning:

§ The orientation and size of high-grade pods
§ The average spacing between ore zones and periodicities in the mineralization.
§ The strike and dip of fracture zones and the plunge of the mineralization within these zones.
§ The distance of influence of a sample or the distance beyond which the similarity between values is negligible.
§ The precision with which the mineralization is known at the point where it is sampled.

Using this information, answers can be given to some critical questions such as:

§ What is the optimal drill hole orientation and drill hole spacing required to complete a drilling program?
§ What will be the effect of a highly selective mining method, as opposed to a large scale bulk mining method, on the tonnages and grades mined?
§ What sampling method, including sample spacing, should be used for grade control during production?
§ What method of reserve estimation should be used to optimize the economic evaluation of the deposit?

Hypotheses concerning the genesis of the orebody can be tested and conclusions can be drawn concerning future drilling to find new deposits in the same district or new ore zones in the same deposit.

In summary, there are very few deposits where the process of exploration and evaluation cannot be greatly enhanced through semivariogram analysis. In addition to an improved understanding of the geology of the deposit such analysis provides a quantitative model of the variability of the mineralization which can be directly used as a basis for subsequent reserve calculations.

5. Geostatistical Reserve Evaluation

5.1 Error of Estimation

5.1.1 Definition

When estimating an orebody or part thereof from a number of sample values, an error is inevitably made. It can be very valuable to
have some measure of the potential size of this error. Consider a panel in a vein deposit, with size 20 meters along strike and 15 meters down dip (Fig. 31). The true average thickness of the vein in this panel and therefore the corresponding tonnage of ore are unknown. Let \( Z \) be the unknown average thickness which can be estimated by drilling a hole in the center of the panel. Let \( Z^* \) be the thickness measured in this hole, say 10 meters. By definition, the variance of estimation of the panel thickness is the expected mean squared difference between \( Z \) and \( Z^* \):

\[
\sigma_E^2 = \text{expected value of } (Z - Z^*)^2
\]

The square root \( \sigma_E \) of the estimation variance is the standard error of estimation. It is often assumed that the error of estimation is approximately normally distributed. Under this condition there is approximately a 68% probability (two chances in three) that the true average thickness of the vein in the panel is greater than \( Z^* - \sigma_E \) and less than \( Z^* + \sigma_E \).

There is approximately 95% probability (19 chances in 20) that this thickness is greater than \( Z^* - 2\sigma_E \) and less than \( Z^* + 2\sigma_E \).

The precision which can be attached to any estimate of the average thickness of a vein in a panel will vary depending not only on the sample information available to estimate it, but also on the method of estimation and the variability of the vein thickness. This variability is quantified by the semivariogram. Although the true thickness of the panel is unknown, the variance of estimation for a particular estimation method can be calculated as a function only of the semivariogram, the location of the samples with respect to the panel, the weight given to each one of the samples, and the size and orientation of the panel.

Given the isotropic semivariogram shown on Figure 31, the precision with which a panel of size 20 by 15 meters would be estimated from a single central sample can be calculated. A standard error of 3.5 meters is obtained. If the central sample indicated a 10 meter thickness, there is 68% probability (two chances in three) that the average thickness of the panel is less than 13.5 meters and greater than 6.5 meters.

In the same vein deposit, consider a panel of the same size, 20 by 15 meters, estimated by the arithmetic average of four corner samples. The corresponding standard error is 2.3 meters. If the thicknesses observed at each corner are 6, 12, 14, and 8 meters respectively, the average thickness of the vein will be estimated at 10 meters with 68% central confidence limits 12.3 and 7.7 meters.

5.1.2 Error of estimation and risk analysis

The error of estimation is extremely useful for estimating the risk involved in opening a new mine. The need for additional drilling can be quantified by comparing the cost of drilling with the expected resulting decrease in the estimation error (RENDU, 1980b). The optimal locations of additional drill holes can also be determined as they must be chosen to minimize the resulting error of estimation.

The error of estimation can also be used to determine, on a monthly or yearly basis, the precision with which the properties of the mineral to be produced are known. Confidence limits for critical parameters such as pounds of sulfur per million BTU in a coal seam or bismuth content in a massive sulfide zinc deposit can be obtained. The potential for
unacceptable variations in these parameters can be detected and might lead to a decision for additional drilling, a change in mining strategy, and/or an increase in storage and blending capacity (Fig. 32).

5.1.3 Error of estimation and selective mining

In most mining operations, some degree of selectivity is applied during mining. Blocks or panels whose estimated value exceeds a specified cutoff grade are processed as ore blocks while blocks estimated below cutoff grade are treated as waste. The decision to classify a block as an ore or waste block is made on the basis of the detailed but incomplete information available at the time mining takes place.

This information which may include channel samples, exploration drill holes, or assayed blast holes is usually not sufficient to allow exact calculation of the block values. Consequently some blocks will be estimated as being below cutoff grade whose true grade exceeds the cutoff grade. They will be incorrectly treated as waste blocks. Other blocks will be estimated as being above cutoff grade whose true grade is below cutoff grade. They will be incorrectly treated as ore blocks. This situation is illustrated on Figure 33. A scatter diagram is shown on which each block is represented by a point, with horizontal coordinate the (known) block estimated value, \( Z^* \), and vertical coordinate the (unknown) block value, \( Z \). Ideally, one would mine only the ore blocks with true value \( Z \) exceeding the cutoff grade. In fact, one will mine the blocks whose estimated value \( Z^* \) exceeds the cutoff grade.

The scatter diagram is divided in four sectors. The first sector contains ore blocks which will be treated as waste. The second sector contains ore blocks which will be treated as ore. The third sector contains waste blocks which will be treated as ore. The fourth sector contains waste blocks which will be treated as waste. The number of blocks falling in the first and third sectors should be minimized (RENDU, 1980b).

It has long been observed in the South African gold mines and other mineral deposits that the traditional methods of reserve estimations often introduce systematic local biases (KRIGE, 1962):

§ low grade blocks tend to be underestimated. On Figure 33, this is translated by most points being above the 45° line for low grade blocks

§ high grade blocks tend to be overestimated. On Figure 33, this is translated by most points being below the 45° line for high grade blocks.

These systematic biases are often recognized and corrected by an empirical dilution factor which can only be accurately determined using historical data. The higher than expected grade of low grade blocks is explained by the addition of surrounding high grade material. Conversely low grade dilution is blamed from down grading high grade blocks. However, this explanation cannot be used when mining is done under very selective and controlled conditions as it is in most large South African gold mines. Dr. D. G. KRIGE was first to recognize that the observed biases could be explained if
not totally, at least in part by the method used to estimate the block values.

In a selective mining environment, two types of dilutions should be taken into account, sometimes known as external and internal dilutions. The first **external dilution** is caused by limitations in the mining method. It is usually impossible to mine a 15 by 15 by 20 meter block without some dilution from surrounding material. The second, **internal dilution**, is of a completely different origin, as it reflects a weakness in the method of reserve evaluation. Dr. D. G. Krige showed that a smoothing method of block estimation such as the one now known as kriging, must be used to eliminate the systematic overestimation of high grade blocks and underestimation of low grade blocks. If such a method is used, the need for an internal dilution factor may disappear, but external dilution still remains to be accounted for.

### 5.2 Calculation of Estimation Variance

#### 5.2.1 General formula

All traditional methods of reserve evaluation consist in estimating the value of a panel or a block $W$ as a linear combination of surrounding sample values. If $n$ samples are used to estimate the block, its estimated value will be:

$$ Z^* = \sum_{i=1}^{n} b_i X_i $$

where $b_i$ is the weight given to the $i$-th sample. The sum of the weights is equal to 1.0. If the inverse square of the distance method is used for example, and $d_i$ is the distance between the $i$-th sample and the center of the block being estimated, then

$$ b_i = K/d_i^2 $$

where $K = \sum_{i=1}^{n} 1/d_i^2$ (7)

The error of estimation of the true block value $Z$ is given in the following equation:

$$ \sigma^2 = \sigma_w^2 + \sum_{i=1}^{n} \sum_{i=1}^{n} b_i b_j \sigma_{ij} - 2 \sum_{i=1}^{n} b_i \sigma_{iw} $$

(8)

where:

- $\sigma_w^2$ = block variance
- $\sigma_{ii}$ = sample variance
- $\sigma_{ij}$ = covariance between the sample $i$ and the sample $j$.
- $\sigma_{iw}$ = covariance between the sample $i$ and the block $W$.

Fig. 34 Estimation of variances and covariances by discrete representation.

#### 5.2.2 Calculation of variances and covariances

The variances and covariances whose values determine the error of estimation can be calculated from the semivariogram $\gamma (h)$, using the autocovariance function $\sigma (h)$ defined as follows:

$$ \sigma (h) = \sigma^2 - \gamma (h) $$

(9)

where $\sigma^2$ is the sample variance ($\sigma^2 = \sigma_{ii}$) and is usually taken equal to the sill of the semivariogram. The block variance $\sigma_w^2$ is equal to the average value of the autocovariance function when all possible pairs of points in the block $W$ are taken into consideration. For simple cases, graphs, and exact mathematical equations, known as "auxiliary functions", are available to estimate the block variance. In most cases, however, a numerical method is used. Since there is an infinite number of points in the block, an exact calculation would be prohibitively time consuming even with the fastest computers. A simple approximation method is commonly used, known as discrete representation. The block is divided into a number of small volumes of equal size and a single point is taken within each volume. The value of $\sigma_w^2$ is then estimated as the average value of the autocovariance function between these points. As an example, consider a block of size 40 by 40 by 15 meters (Fig. 34). The semivariogram has been calculated using 15-meter composites. The model accepted is the sum of a nugget effect $N=0.2$ and an isotropic spherical model with sill $C=0.8$ and range $a=150$ meters. The sample
The variance is estimated as follows:

\[ \sigma^2 = C + N = 1.0 \]

The mathematical equation of the spherical semivariogram was given earlier. If we divide the block \( W \) into four equal cubes of size 20 by 20 by 15 meters, the block variance is estimated as follows (Fig. 34a):

\[
\sigma_{w^2} = \frac{[4\sigma(0) + 8\sigma(20) + 4\sigma(20\sqrt{2})]}{16} = \frac{1}{16} (4 \times 0.8 + 8 \times 0.651 + 4 \times 0.590) = 0.673
\]

(10)

This approximation is poor and in practice it is considered that a 16-point discrete representation is needed to obtain acceptable results (Fig. 34a). The corresponding value would be 0.641.

\( \sigma_{ij} \) is equal to the sample variance \( \sigma^2 \) for \( i = j \), and to the covariance \( \sigma(h_{ij}) \) for \( i \neq j \), where \( h_{ij} \) is the distance between the two samples.

As an example, given the semivariogram on Figure 34, the covariance between two samples at distance 100 meters is \( \sigma(100) = 0.119 \).

\( \sigma_{iw} \) is equal to the average value of the covariance \( \sigma(h) \) when \( h \) is the distance between the sample \( i \) and all possible points in the block \( W \). Auxilliary functions can be used in simple cases. More commonly, numerical integration will be used, which is based on discrete representation of the block \( W \).

As an example, consider again a block of size 40 by 40 meters, and a sample at a distance of 50 meters, as shown on Figure 34b. If a four point discrete representation is used, the covariance between the sample and the block is estimated as follows:

\[
\sigma_{1W} = \frac{1}{4} [2\sigma(\sqrt{40^2+10^2}) + 2\sigma(\sqrt{60^2+10^2})] = 0.408
\]

(11)

Typically, a 16 point discrete representation is preferred. In this example the result would not be changed.

If the block value was estimated from one sample only, at distance 50 meters (Fig. 35a), the estimation variance would be:

\[
\sigma_E^2 = \sigma^2 + \sigma_{i1} - 2\sigma_{1W} = 0.641 + 1.0 - 2 \times 0.408 = 0.825
\]

(12)

If the block value was estimated from two samples located on both sides of the block, as shown in Figure 35b, each sample being given an equal weight \( b_1 = b_2 = 1/2 \), the estimation variance would be:

\[
\sigma_E^2 = \sigma_w^2 + \sigma_{ii} - 2\sigma_{1W} = 0.641 + 1.0 + 2 \times 0.119 - 2 \times 0.408 = 0.385
\]

(13)

Note that, provided the stationarity conditions are satisfied, this error of estimation is dependent on the location of the samples \( w_1 \) and \( w_2 \) but not on their value.

5.3 Kriging

5.3.1 General

Intuitively, when estimating a panel or block of ore, the closer samples should be given more weight than the ones farther away. The directions showing the best continuity should also be considered in assigning a weight to a sample. In addition, the relative position of the drill holes with respect to each other should be taken into account. The size of the block to be estimated is also an important factor. Most important is the fact that the optimal method of reserve estimation should be a function of the specific properties of the deposit or part of the deposit within which the block is located. The geostatistical method of reserve estimation takes all these factors into account.

As indicated earlier, the precision with which a block is estimated can be quantified using the semivariogram. The error of estimation is a function of the estimation method or, more specifically, of the weight given to each sample in estimating the block. Instead of calculating the error of estimation which corresponds to a given set of sample weights, it is possible to calculate the weights which will result in the minimum error. The optimal weights, known as kriging weights, are obtained by solving a system of linear equations, known as the kriging system.

An illustration of a simple kriging problem is given in Figure 36. Nine drill hole intercepts are available within a bench to estimate a block of size 4,000 by 4,000 feet. The average grades
measured at each hole are shown on Figure 36 as well as the semivariogram calculated from all the drill holes in the deposit. If the polygon method is used, the block is estimated by the central sample, with value 4% zinc and an error of 1.20% zinc is made. If the mean of the nine samples is used, an average of 6.11% zinc is calculated with an error of 0.72% zinc. If kriging is used, the central sample is given most of the weight and the outside samples are given weights which reflect the anisotropy shown by the semivariogram. The block estimated average grade is 5.10% zinc with an error of estimation of 0.52% zinc.

Kriging can be used to estimate the reserves within an entire deposit, to estimate the reserves scheduled for production on a yearly basis, or to estimate a regular grid of blocks. If a regular grid is estimated, the corresponding errors of estimation can be contoured, as well as the estimated block values. Areas with relatively high errors can be studied with respect to the proposed mining sequence to determine whether additional drilling is required. The block values estimated by kriging can also be used for computerized or manual mine design.

The kriging method of evaluation is the only method which uses to its full potential the statistical and geologic properties of the deposit as represented by the semivariogram. The kriging weights and the error of estimation depend on the block size, block location and sample locations, and on the semivariogram. They are independent of the sample values. This is true only if the necessary stationarity conditions are satisfied.

The kriging method of block estimation has been applied to the test data set given in Figure 14, using the semivariogram model shown in Figure 18. Blocks of size 100 by 100 meters have been estimated using the ten samples closest to each block. The results are shown on Figure 37. All calculations were made on a programmable pocket calculator.

5.3.2 Kriging calculations

If a block \( W \) is estimated from the sample values \( x_i \) using a formula of the form:

\[
Z^* = \sum_{i=1}^{n} b_i x_i
\]

the error made in estimating the true value \( Z \) is a quadratic function of the sample weights \( b_i \). Kriging consists of calculating the weights which should be given to the sample values, so as to minimize the error of estimation. It can be shown (RENDU, 1981) that these weights are obtained by solving the following system of linear equations:

\[
\begin{align*}
\sigma_{11} b_1 + \sigma_{12} b_2 + \ldots + \sigma_{1n} b_n + \lambda &= \sigma_{1W} \\
\sigma_{12} b_1 + \sigma_{22} b_2 + \ldots + \sigma_{2n} b_n + \lambda &= \sigma_{2W} \\
\vdots \quad & \quad \vdots \\
\sigma_{1n} b_1 + \sigma_{2n} b_2 + \ldots + \sigma_{nn} b_n + \lambda &= \sigma_{nW} \\
b_1 + b_2 + \ldots + b_n &= 1
\end{align*}
\]

There are \( n+1 \) equations with \( n+1 \) unknowns, the sample weights \( b_i \) and the lagrange multiplier \( \lambda \), which must be introduced to impose
the constraint that the sum of the weights adds up to 1.0.

This system of equations is known as the kriging system. We saw earlier how all the variances and covariances in this system can be calculated from the semivariogram. The optimal weights which are solutions to this kriging system of equations are the kriging weights. The corresponding minimum error of estimation is the kriging error and can be written in the following form:

\[ \sigma_K^2 = \sigma_y^2 - \sum_{i=1}^{n} b_i \sigma_{yw} - \lambda \]  \hspace{1cm} (15)

As an example, consider again a block of size 40 meters by 40 meters, located in a deposit whose semivariogram is as shown on Figure 34. Three samples are available to estimate the block, a central sample and two samples located at a distance of 50 meters from the center of the block (Fig. 38). Using a 16 point discrete representation of the block, the following variances and covariances are obtained:

\[ \sigma_W^2 = 0.641 \]
\[ \sigma_{1W} = \sigma_{2W} = 0.408 \]
\[ \sigma_{3W} = 0.681 \]
\[ \sigma_{11} = \sigma_{22} = \sigma_{33} = 1.0 \]
\[ \sigma_{12} = \sigma_{13} = \sigma_{23} = 0.415 \]

Hence the kriging system:

\[ 1.0b_1 + 0.119b_2 + 0.415b_3 + \lambda = 0.408 \]
\[ 0.119b_1 + 1.0b_2 + 0.415b_3 + \lambda = 0.408 \]
\[ 0.415b_1 + 0.415b_2 + 1.0b_3 + \lambda = 0.681 \]
\[ b_1 + b_2 + b_3 = 1 \]

The solution to this system is \[ b_1 = b_2 = 0.214, \]
\[ b_3 = 0.572 \text{ and } \lambda = -0.069. \]

The corresponding estimation variance is \[ \sigma_K^2 = 0.146. \]

If the polygon method was used, only the central sample would be considered and the error of estimation would be:

\[ \delta_K^2 = 0.641 + 1.0 - 2 \times 0.681 = 0.279 \]

5.4 Additional Remarks about Kriging

According to geostatistical theory and practical experience, if a deposit is mined selectively, the optimal method of evaluation of individual blocks is by kriging. This method results in the smallest estimation error. The systematic overestimation of high grade zones and underestimation of low grade zones, common with most methods of reserve evaluation, is also minimized. However even if kriging is used, an error remains in the estimation of the blocks. Consequently the tonnage mined and the corresponding average grade will differ from the tonnages and grades which would be mined if we had exact information concerning the true grade distribution.

As for any method of reserve evaluation, application of the kriging method will give meaningful results only if detailed geologic studies of the deposit have been completed. Statistical analyses as discussed earlier, including the semivariogram, can help develop a better understanding of the deposit. A geologic model must however be built before kriging can be completed. Since kriging usually requires the use of a computer, all relevant geologic controls must be computerized, and when this is not possible, the need may exist to carefully verify and manually override meaningless results.

Strictly speaking, the linear kriging method of reserve estimation is truly optimal only if minimum stationarity conditions are satisfied and the sample values are normally distributed. If departure from this ideal but unfortunately rather uncommon model occurs, other geostatistical methods may be applicable. Logarithmic kriging (Jouzel and Huijbrecht, 1978; Rendu, 1979) was developed to estimate deposits in which the sample values are lognormally distributed. Disjunctive kriging (Matheron, 1976; Rendu 1980a) can be used to estimate deposits within which the sample frequency distribution is well known but cannot be represented by simple models such as the normal or lognormal mode. Universal kriging (Jouzel and Huijbrecht, 1978) was developed to estimate block values when drifts are present and the necessary stationarity conditions are no longer satisfied.

A number of geostatistical methods are also available to estimate recoverable reserves, when...
only wide spaced drill hole information is available at the time the estimate is made, but detailed blasthole samples will be used for selective mining. Finally, geostatistical methods have been developed to assist in the evaluation of possible grade control, blending, and short term mining and processing problems. These methods, known as "orebody simulation" (Journel and Huijbregts, 1978) are aimed at creating a computer model of the orebody which shows the same short range variability as is present in the deposit itself.

All geostatistical methods other than the linear kriging discussed in this paper, should be used with extreme caution, as the necessary conditions of application may be extremely restrictive, and significant errors may result if these conditions are not satisfied.

Acknowledgements: The author wishes to acknowledge the constant encouragements received from Dr. Tetsuya Shoji, Tokyo University, who made publication of this paper possible.

The author is also indebted to Golder Associates for the time and resources made available, and to Graham Mathieson and Tony Edey who helped in the editing of this paper.

References


礦量推定のための地質統計学的方法

鉱量計算に対する統計学的方法の応用は、1940年代後半から1950年代前半にかけて、南アメリカの銅鉱床を対象に始まった。このとき、品位が変動正規分布することが確認され、新しい銅鉱床の鉱量推定のための統計学的方法が開発された。1950年代に、Krigingが品位空間トーローの問題を、多変量統計の回帰の方法を応用して、解決した。これは現在のクライギング法と呼ばれているものである。1960年代に入って、Matheronとその一派によって、空間的に分布する斑状を統計的に解析するための数論的理論が開発された。このとき、セバリオグラムが提案され、クライギング法が鉱床推定のための地質統計学的方法として受け入れられ、1960年代後半から1970年代を通じて、地質統計学の理論が発展するとともに、世界中の銅鉱床の教育がそれを取り入れられ、鉱床評価に利用されるようになった。

鉱床を統計的に解析する手始めは、ヒストグラムである（Fig.1）。ヒストグラムを描くことにより、サンプル値の分布状況が分かるとともに、異常値も検出できる。また、地質環境の異なるデータが観察していることを発見する場合がある（Fig.2）。ヒストグラムから頻度分布を計算することにより、与えられた範囲内におけるサンプル値推定値を数値で示すことができる（Table 2）。これによると、9600 BU/ lbの間の発熱量も少ない場合、サンプル値を解析する他の方法として、累積頻度分布がある（Table 2）。これによると、9600 BU/lbより発熱量の小さなサンプルは34.69%であることが示される。したがって、頻度分布が正規分布していれば、累積頻度分布は正規確率纸上で直線となる（Fig.3）。グラフを用いた他の方法として、スキャッターダイアグラムが考えられる。これにより、異常値や異なる母集団の混合などの検出ができる（Fig.4, Fig.5）。

グラフによる解析は非常に有効であるが、(1)サンプル値を分類するとときに地質を考慮していないと、無意味な結果が得られる。(2)サンプルの空間の位置と推定値についていない。なお統計学的処理としてはサンプルの大きさが等しい必要があるという点を忘れてはならない。Fig.6の右側には、サンプル値が3cmから9mまで変化しているときの証の品位が、左側にはこれらを3cm幅に統一したときの品位が示されている。これにより、品位の変動高値は3cm幅では2.0%であるが、3m幅では1.0%である。サンプルの大きさが大きくなるにつれて、値はお互いに近くなり、鉱床の平均値に近くとも考え、対応する値の変動は小さくなる。

鉱、鉱石、鉱床、工業原料鉱物、高品位の鉱金属鉱床の品位は正規分布する。これに対し、金、モリブデン、銅、鉱石中の硫黄、亜鉛鉱床中のスズ、鉛、銅、亜鉛、ウランの品位のヒストグラムは通常高い方へ傾きを示した分布を示す（Fig.8）。対数正規確率紙上で直線とする（Fig.9）。なお、対数正規確率紙上で線図を示すデータに対して、適当な値を加えたり加減したりすることをする(M10, Fig.10)。正規分布を示す変数の平均値をm（式1）, 標準偏差をa（式2）とした場合、真の平均μがmーaと、m＋2aの間にある確率は68%、mー2aと、m＋2aの間にある確率は95%である。

以上に述べた統計学的方法により、鉱床を特微づけるパラメーターの平均値を推定し、その値の信頼限界を計算することができる。この場合、サンプル値がお互い独立で、それぞれ等しい表現力を持っていることを前提としている。それにより、鉱床から多数のサンプルを取ることで、サンプルの独立性という前提は成り立たない。サンプル値は距離の関数で、近接するサンプルの値は遠くのものに較べて一般に近い。このことを考慮し、鉱床推定法が地質統計学である。近接するサンプルは鉱化作用に関連して似た性質を示し、類似性は距離が大きくなるにつれて減少する。類似性は非無視できるような距離を影響範囲と定義でき、その大きさは鉱床ごとに、あるいは同一鉱床の異なる部分、あるいは同一部分の異なる方向によって変化する。距離がhだけ離れた2点のサンプル値の平均の平均値の半分をセバリオグラムδ(i, j)という（式3）。

Fig.12に示すデータについての計算例が式(i)と式(i)である。このような計算をする場合、通常、サンプル間距離が等しくないもので、ほぼ等しい距離のグループに分けて計算する必要がある。

二次元の場合には、Fig.13に示すように、hーdhより大きくhーdhより小さく、aーdaより大きく、aーdaより小さい矩形でグループ分けをする。Fig.14に示すように、具体的に計算を実行すると、正規分布（Fig.15）および距離と方位で分類したセバリオグラム（Table 3, Fig.16）を得ることがある。このためにおけるγ(i, j)は距離の関数であり、その最適値は北の方位で最も低く、東または西北東の方位で最大である。すべての方向で、セバリオグラムは最適値上昇し、その後急激に減少している。これでサンプル値の方位で周期性をもっているか、対称性をもっているためである。等セバリオグラム図（Fig.17）によると、N25°Eの方位が最も連続性があり、65°Wの方位が最も悪い。Fig.18は、これら両方向についてセバリオグラムを計算した結果である。以上の結果から、この鉱床の鉱床はN30°Nの方位に偏っており、65°Wの方位で円筒形をしていることから、鉱床帯の幅は約200mであることが分る。

三次元のセバリオグラムの計算には、Fig.19に示すような円錐台あるいは円形のある矩形を用いる。Fig.20は三次元の異方性のある塊状酸化物鉱床について、セバリオグラムを計算した例である。

セバリオグラムの一般的モデルをFig.21に示す。セバリオグラムで制限される変動の最高値をシル、セバリオグラムが一定になるか、シルに達する距離をレンジ、距離が0におけるセバリオグラムの値をナゲット効果という。最も一般的に使われる球形モデルのセバリオグラムは、式6で表現される。これ以外のモデルとして、指数モデル、直線モデル、放射線モデル、誤差関数モデル、標準差モデル、ホワイト効果モデルがある（Fig.21）。連続性の良い層状規制鉱床の厚さは放射線モデルが、連続性があまり良くない場合には誤差関数モデルがよく合う。ナゲット効果のある誤差関数モデルは炭層の厚さに（Fig.22）。ホワイト効果のあるセバリオグラムは周期性のある鉱床に用いられる。Fig.23は多数のN-S方向の破砕帯がある割合に規制された鉱床のセバリオグラムである。
 экономическ的にも、これらの観点から、セパレートグラフによる解析の必要性を示すことができる。

Fig. 31に示すような鉱物を考え、この鉱物の中心において、鉱物の厚さの推定値Zと測定値Zとの差の利用可能な期待値を推定する方法である。もし推定誤差が正規分布していると、鉱物の厚さの平均値がZ±0.674である。平均値を推定する場合の精度は推定に使用する情報を信頼性に応じて、事前推定方法と厚さの変動をも考慮する。この場合、変動性はセパレートグラフによって定量化できる。鉱物の厚さの平均値は分からないが、使用する推定可能な推定精度は、セパレートグラフモデルの推定値、推定による精度、鉱物に対する精度、鉱物の大きさと方向の関数を計算することができる。推定誤差は、新しい鉱物を対象する場合の精度を推定するのに必要であり、試料数とそれから得られる結果を比較して、さらに試験を実施する。精度を最大にできるための鉱物の試験位置はどこかを決めることが必要になる。