59389

Developments in Geomathematics 2

M. DAVID

Elsevier Scientific Publishing Company

10494

ION IN REMOTE

Developments in Geomathematics 2

GEOSTATISTICAL ORE RESERVE ESTIMATION

by

MICHEL DAVID

Ecole Polytechnique, Université de Montréal, Montréal, Québec, Canada

To my dear friend Jeff
Michel

Elsevier Scientific Publishing Company
Amsterdam Oxford New York 1977

ELSEVIER SCIENTIFIC PUBLISHING COMPANY
Molenwerf 1,
P.O. Box 211, 1000 AE Amsterdam, The Netherlands

Distributors for the United States and Canada:

ELSEVIER SCIENCE PUBLISHING COMPANY INC.
52, Vanderbilt Avenue
New York, N.Y. 10017

First edition 1977
Second impression 1978
Third impression 1979
Fourth impression 1982

Library of Congress Cataloging in Publication Data
David, Michel, 1945-
Geostatistical ore reserve estimation.

(Developments in geomathematics ; 2)
Bibliography: p.
Includes index.
1. Mine valuation--Statistical methods. 2. Mines and mineral resources--Statistical methods. I. Title. II. Series.
TN272.D37 622'.13 76-39608
ISBN 0-444-41532-7

ISBN 0-444-41609-9 (series)
ISBN 0-444-41532-7 (vol. 2)

© Elsevier Scientific Publishing Company, 1977
All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of the publishers, Elsevier Scientific Publishing Company, P.O. Box 330, 1000 AH Amsterdam, The Netherlands

Printed in The Netherlands

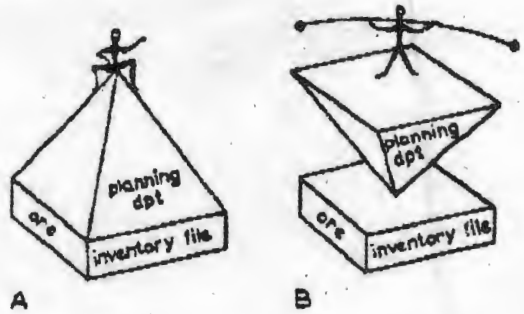


Fig. 34. Mine manager and planning department.

our planning department built? If the situation is the one of Fig. 34A, then chances are that the planning will be good. If it looks like the one in Fig. 34B, action has to be taken immediately! Now the problem is that one quite often doesn't know in which case one is! It will be one of the main tasks of geostatistical computations to answer that question.

We will concentrate on the problem of obtaining an ore inventory file and this will at the same time generate all the concepts which will be sufficient and necessary to solve almost any other estimation problem in a mine-life time and appraise any estimation method.

3.2 WHAT IS AN ORE RESERVE ESTIMATION?

What we intend to do here is to briefly review the actual practices in this field in order to isolate the underlying concepts and the questions which arise.

We can state the general problem as follows: one knows the actual grade of some localized samples and one wants to know the grade of a larger block of ground or even of the whole mine. The natural solution is thus knowing a few values ($t_1, t_2, \ldots, t_n$), to manipulate these values in order to come out with a combination of these grades which will be considered as the real grade of the block under study. Doing these manipulations, whatever they are, implies two basic concepts:

— *the concept of extension,*
— *the concept of error of estimation.*

3.2.1 *The concept of extension*

We can group the ore reserve estimation methods under two headings: Geometrical and Distance-weighting methods. They are all, however, based on the same principle.

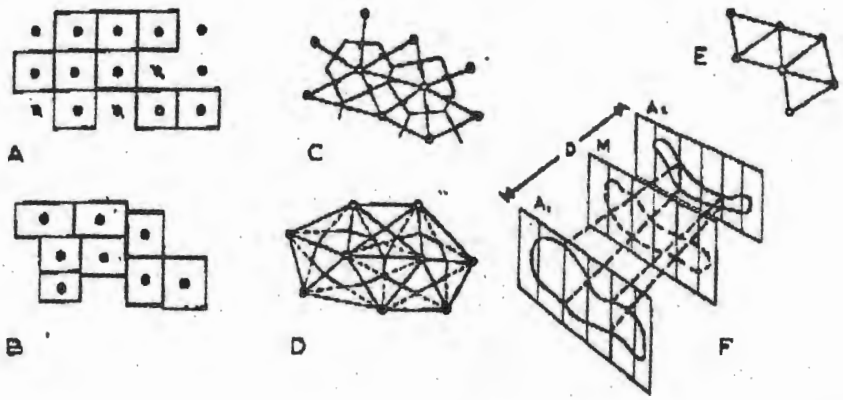


Fig. 35. Geometrical methods of ore reserve estimation. A. Regular blocks. B. Irregular blocks. C. Polygonal. D. Angular bissection. E. Triangular. F. Sections. From Patterson (1959).

3.2.1.1 *Geometrical methods*

These were the most employed until the advent of computers and are still favoured in many companies. They include polygonal, triangular and cross-sectional methods and have been summarized in Fig. 35. They were also the first to be computerized (Hewlett and Faick, 1961). In any case, what one does is always to assign to a block the grade of a portion of a diamond drill hole. We extend to the block the grade of a portion of a diamond drill hole. We extend to the block the grade of something else. We will call this an extension. This is the basic principle.

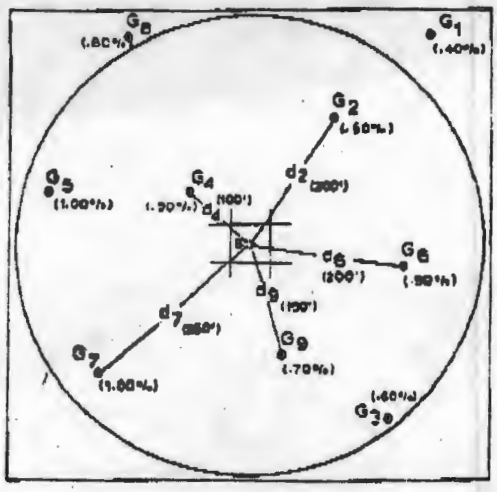
3.2.1.2 *Distance-weighting methods*

The same extension prevails in these methods which became popular only after computers were available, due to the very large number of repetitive calculations required. These methods assign to a block, or to a point, a linear combination of the grades of surrounding samples. Whatever the coefficients of that combination, in other words whatever the weighting method used, this is again an extension of the grade of some samples to a larger volume. Fig. 36 gives an example of that method.

3.2.2 *The concept of error of estimation*

3.2.2.1 *The error*

Due to the simple fact that samples and blocks are not strictly equivalent there is an error involved in any estimation procedure. The validity of an estimation method will be assessed by the magnitude of the errors involved. Note that it is not possible to forecast an error, since if an error is known,



$$B = \frac{G_2 \times \frac{1}{(d_2)^2} + G_5 \times \frac{1}{(d_5)^2} + G_9 \times \frac{1}{(d_9)^2} + G_7 \times \frac{1}{(d_7)^2} + G_4 \times \frac{1}{(d_4)^2}}{\frac{1}{(d_2)^2} + \frac{1}{(d_5)^2} + \frac{1}{(d_9)^2} + \frac{1}{(d_7)^2} + \frac{1}{(d_4)^2}}$$

$$B = \frac{0.5 \times \frac{1}{(200)^2} + 0.8 \times \frac{1}{(200)^2} + 0.7 \times \frac{1}{(150)^2} + 1.0 \times \frac{1}{(250)^2} + 0.9 \times \frac{1}{(100)^2}}{\frac{1}{(200)^2} + \frac{1}{(200)^2} + \frac{1}{(150)^2} + \frac{1}{(250)^2} + \frac{1}{(100)^2}} = 0.77\%$$

Fig. 36. Hypothetical block calculation from neighbouring samples using an inverse squared distance method. After O'Brien and Weiss (1968).
Legend: G = assay composite value, d = distance, B = block assay.

there is no error left! Now what we want to do is to select, before the actual mining of the blocks, the best estimation method. What does "best" mean? Everybody will agree that the best method is the one which yields the smallest errors in general, considering all the blocks in the deposit. We know that an estimation procedure will not yield a constant error, sometimes in a porphyry copper deposit for instances, it will be ± 0.1% Cu, sometimes, hopefully less often, it will be 0.4% Cu. We start to see the concept of probability of occurrence of an error. When there is a cut-off involved we also want the classification to be correct, ore and waste tonnage should be correctly evaluated.

3.2.2.2 *The distribution of the estimation error*

Suppose that in one deposit we have a number, n, of identical blocks which we want to estimate from a similar setting of samples. Let the true

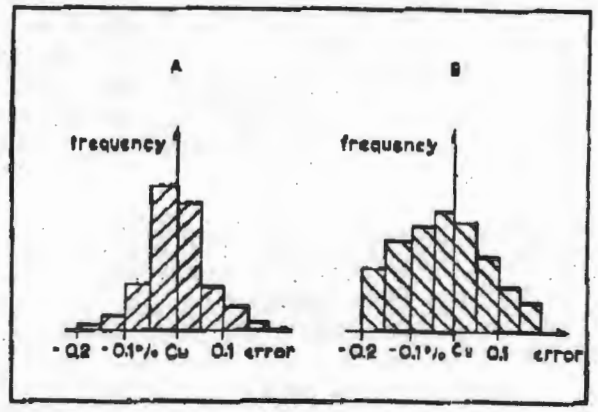


Fig. 37. Distributions of errors for two methods of estimation of block grades in a porphyry copper deposit.

real grade of blocks be $Z_1, Z_2, \ldots, Z_n$. Let the estimated grades — using one particular method — be $Z_1^*, Z_2^*, \ldots, Z_n^*$. The resulting errors are thus $e_1 = Z_1 - Z_1^*$; $e_2 = Z_2 - Z_2^* \ldots$; $e_n = Z_n - Z_n^*$. We will be satisfied if the average of the estimates is the same as the average of the real values; in other words: we want no systematic error. This is the non-bias condition. Then we will be happy if most of the errors have a low absolute value, knowing that some of them might be larger and hoping that a very small number of them will be large.

In statistical words we say: "We want the distribution of the errors to have a low dispersion: a low variance".

This is illustrated in Fig. 37. Suppose again that we are dealing with a porphyry copper averaging say 0.8% Cu. Then for blocks of 50′ x 50′ x 50′, a distribution of errors like the one of Fig. 37A is acceptable while a distribution like the one of Fig. 37B is unacceptable.

3.2.2.3 *The variance of the error on a single value*

In some instance we will estimate only one value, for instance the average grade of a deposit. This value is not error-free. Also it is obvious that we cannot know the error incurred, otherwise, there would be no error. We have already discussed this in Chapter 2 under the heading "Standard error of the mean". Let us see it from a more intuitive point of view.

Let X be the value which we have obtained. There is only one value for the error and we cannot draw a distribution like in Fig. 37. We can feel, however, that this error has a given probability to be between such and such values, and a small probability to be above another large number.

Suppose we are dealing with a 1.6% Cu vein. We will feel, for instance, that there is a 50% chance that the error will be less than 0.1% and maybe

80% chance to be smaller than 0.2%. These feelings are nothing more th' the probability distribution of the error, and thus we can again associate our estimate a variance, which gives the spread of the distribution of the error. We could sum up by saying that we have an estimated mean value X and we want to know the standard error of the mean value.

The exact type of the distribution of that error is and will remain unknown. However, for usual situations, the distribution of that error is irrelevant and whatever the model we select, either normal or lognormal, for instance, we will obtain the same results in terms of confidence intervals.

At this point we should make the difference between academic statistical problems and actual engineering problems. Concerning the previous statement about the exact distribution of the error being irrelevant we can give the following example.

Example. Suppose that we have an estimated value of 0.65% Cu with a standard deviation of 0.10% Cu. If we choose a normal model for the distribution of the error, the value which has 84% chance of being surpassed is $0.65 - 0.1 = 0.55\%$ Cu. If we choose a lognormal model then this value is $0.65 \exp(-0.1/0.65) = 0.557$, which for practical purposes is equivalent to the first one. Now if we have an estimated value of 0.65% with a standard deviation of 0.30% Cu, then the two models respectively yield 0.35% Cu and 0.41% Cu, which are different. But in the second case, whatever the true unknown answer (academic), the only information to remember (engineering) is that we have no idea of what might be in the block; it is irrelevant to differentiate between something very bad and something worse!

3.2.2.4 *Exercise: assessing the quality of an ore reserve estimation procedure*

The purpose of this exercise is to find a way to select an estimation procedure. We have a series of block values (real values) and two series of estimated values for these same blocks, using two different methods (see Table 3.I). The first method is by "inverse squared distance" (I.S.D.) and the second is by kriging.

The blocks are from a very regular porphyry copper deposit where the mean grade is 1.19% Cu and the variance of point samples is 0.20 $(\%)^2$.

The blocks are 50′ x 50′ x 50′ and the estimates have been made from samples on a 200′ grid.

A common way to assess the quality of the estimation is to examine the frequency distribution of the errors. The two corresponding distributions are given in Fig. 38.

We see that the kriging procedure is in this case much better; the spread, or variance of the distribution is much smaller. We should add that this criterion is not the only one to select a method of estimation. In many cases the two distributions will be closer to one another than here. Then other

TABLE 3.I

Block values, estimated and real for Exercise 3.2.2.4

Real	Kriged	I.S.D.	R − K	R − I.S.D.
0.99	1.07	1.21	−0.08	−0.22
1.67	1.48	1.29	0.19	0.38
1.63	1.48	1.31	0.15	0.32
1.43	1.24	1.25	0.19	0.18
0.98	0.99	1.17	−0.01	−0.19
0.88	1.04	1.20	−0.16	−0.32
1.67	1.63	1.27	0.04	0.40
1.31	1.36	1.26	−0.05	0.05
1.15	1.16	1.21	0.01	−0.06
1.01	0.96	1.17	0.05	−0.16
0.81	0.99	1.20	−0.18	−0.39
1.72	1.53	1.26	0.19	0.46
1.10	1.26	1.26	−0.16	−0.16
1.23	1.26	1.23	−0.03	0.00
1.06	1.06	1.19	0.00	−0.13
0.90	0.92	1.17	−0.02	−0.27
1.58	1.32	1.21	0.16	0.37
1.21	1.28	1.26	−0.02	−0.05
1.65	1.59	1.33	0.06	0.32
1.39	1.20	1.29	0.19	0.10
0.97	0.92	1.16	0.05	−0.19
1.33	1.24	1.21	0.09	0.12
1.20	1.23	1.27	−0.03	−0.07
1.94	1.75	1.36	0.19	0.58
1.58	1.45	1.33	0.13	0.25
0.99	1.01	1.37	−0.02	−0.38
1.13	1.10	1.22	0.03	−0.09
1.32	1.30	1.30	0.02	0.02
2.06	1.99	1.39	0.07	0.67
1.54	1.66	1.35	−0.12	0.19
0.97	1.02	1.16	−0.05	−0.19
0.97	1.12	1.21	−0.15	−0.24
1.27	1.30	1.29	−0.03	−0.02
1.55	1.80	1.36	−0.25	0.19
1.21	1.54	1.23	−0.33	−0.02
0.99	0.96	1.13	0.03	0.14
0.96	1.07	1.18	−0.11	−0.22
1.27	1.22	1.21	0.05	0.06

qualities of the estimation method will have to be investigated as the next paragraph will show.

For kriging: The average error is 0.0005% Cu, or 0.0 as it should be and the variance of the error is $\sigma_2^2 = 0.015$, i.e., the standard deviation is 0.122% Cu.

[illegible] average error is 0.037
$\sigma^2 = 0.069$
$\sigma = 0.26$

Fig. 38. Distribution of errors for two methods of estimation (I.S.D. and kriging) of block grades. Data of Table 3.I.

For I.S.D.: The average error is 0.037, so that the estimates are systematically biased, and the variance of the error is 0.069 giving a standard deviation of 0.26% Cu.

The conclusion is left to the reader. . . the drawing of Fig. 38 also!

3.2.3 *The correct assignment of blocks to ore and waste*

Maintaining a small estimation error is not the only criterion to assess the validity of an ore estimation technique. Tonnage is also a very important (perhaps the most important) factor. When in a given overall envelope there exists ore and waste one wants the small blocks to be correctly classified.

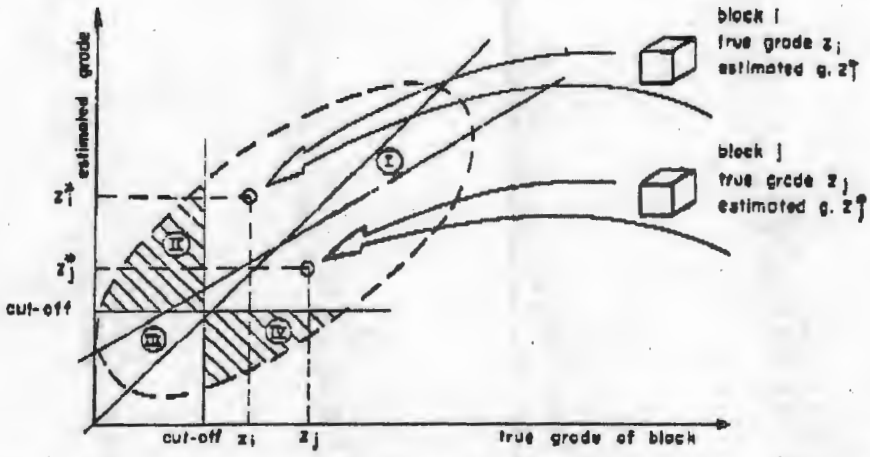


Fig. 39. Correlation between predicted and actual grades of blocks. Shaded areas show groups of blocks wrongly allocated.

Chapter 4

WHAT IS A VARIOGRAM?

SYNOPSIS — *In this chapter we will try to show that the idea of a variogram introduces itself naturally when one starts to think about estimation and variability problems. We will first define what the variogram function is, then show that it incorporates several geological features which are important in mineral deposit valuation, such as the continuity, size and shape of a zone of influence. Then a little arithmetic will show us that the variogram function intervenes in all estimation variances and block variances computations. This will be illustrated by two exercises in order to show what kind of real results one can expect from the variogram, and finally before going through the theoretical development we will show how one can compute a variogram in a simple case, and introduce a most frequently used variable in the reserve computation, the accumulation.*

4.1 SPATIAL CORRELATION

All the geological features which we have mentioned in the previous chapter could in fact be summarized under one name. They represent the similarity or rather the dissimilarity which exists between the grade at one point and the grade some distance away. The error of estimation is thus a function of the similarity which can be expected between samples and a block. This expected variation can be called Spatial Similarity or rather Spatial Correlation. We now see that what we have to do is to find a way to characterize that similarity of grades within a deposit or, rather, within a homogeneous area of the deposit where we can suspect that the geological relationships are the same and where consequently the precision of an estimation procedure for a given geometrical pattern depends only on the pattern and not on the particular location in the homogeneous area.

4.2 DEFINITION OF THE VARIOGRAM

The most natural way to compare two values, say two grades $Z(x)$ and $Z(x+h)$ at two points, x and $x+h$, h feet away from x, is to consider their difference. Since we are not really interested in the sign of the difference but rather in the absolute value, we should consider the value $|Z(x) - Z(x+h)|$. This value expressing the dissimilarity between two particular points is of very little interest. We rather want the average difference between two points h feet apart. Thus we should consider the value {Average $|Z(x) - Z(x+h)|$} for all possible points x and $x+h$. Since absolute values

are difficult in calculus, we will rather consider the squared differences and thus select as our dissimilarity function:

$$2\gamma(\vec{h}) = \text{AVE}\,[Z(x) - Z(x + \vec{h})]^2 \tag{4.1}$$

which should be properly expressed as:

$$2\gamma(\vec{h}) = \frac{1}{V}\int_{\underline{V}} [Z(x) - Z(x + \vec{h})]^2\,dx \tag{4.2}$$

where V is the volume of the deposit $\underline{V}$.

This is the variogram, $2\gamma(\vec{h})$. It is a function of a vector, in other words, a distance and the orientation of that distance, and it expresses how grades differ in average according to the distance in that direction.

This definition involves an integral, in fact, a triple integral calculated over the whole deposit. In practice we will have to estimate that integral through the available data. Suppose we have a total of $N(\vec{h})$ pairs of samples separated by a vector $\vec{h}$, then we will estimate the variogram by:

$$2\gamma(\vec{h}) = \frac{1}{N(\vec{h})}\sum_{i=1}^{N(\vec{h})} [Z(x_i) - Z(x_i + \vec{h})]^2 \tag{4.3}$$

As a simple case, take samples regularly distributed along a line as in Fig. 52; we have n samples at an interval of d feet, thus we will have $(n - 1)$ pairs to compute $\gamma(d)$, $(n - 2)$ to compute $\gamma(2d)$ and so on.

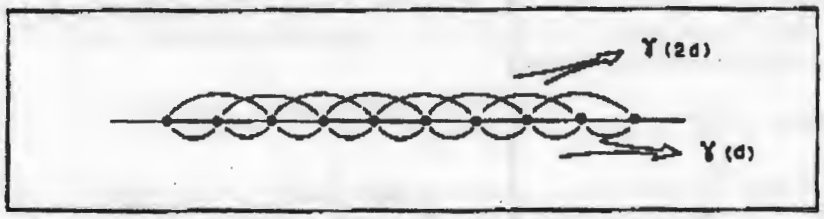


Fig. 52. Schematic computation of a variogram, using pairs of samples a given distance apart.

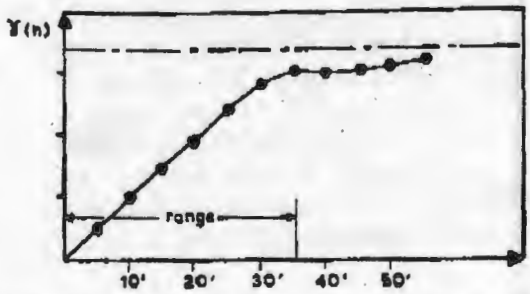


Fig. 53. Typical variogram from a lead—zinc deposit showing a good continuity. The mineralization is disseminated.

4.3 THE VARIOGRAM AS A GEOLOGICAL FEATURES DESCRIPTOR

All the geological characteristics which we have mentioned so far have to appear quantitatively in the variogram. This should help to make precise the "feeling" of a geologist, specially at the beginning of an operation.

4.3.1 *The continuity*

The continuity is reflected by the rate of growth of $\gamma(h)$ for small values of h. In a sedimentary deposit, changes usually occur very slowly. We obtain variograms like the one of Fig. 53 (a lead—zinc deposit), with a gentle, regular growth from zero. In other cases the mineralization may occur as nuggets, or blobs, often concentrated in veinlets, then rapid changes may happen over a very short distance. This will be called later a Nugget Effect. We see such a variogram (from a moly deposit) in Fig. 54. The continuity may also be completely non-existent. Then we have a variogram like those of Fig. 55: whatever the distance between two samples, they are simply independent of each other. This is common in gold deposits and this is why that type is [illegible]d the Pure Nugget Effect type.

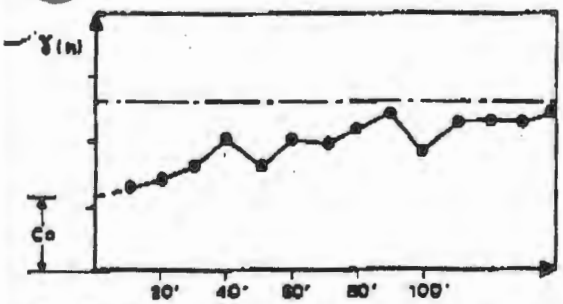


Fig. 54. Variogram from a molybdenum deposit showing continuity curtailed by a high low-scale variability (C_0). The mineralization occurs in thin (a few mm) veinlets.

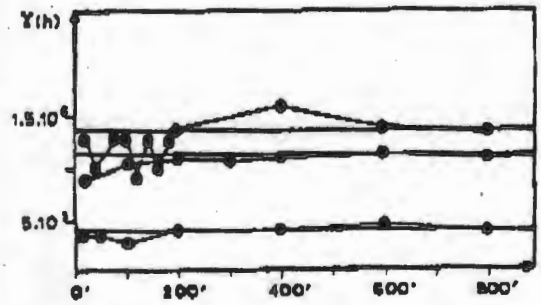


Fig. 55. Variograms from another molybdenum deposit showing no continuity. Mineral[illegible] occurs in thick (1") veins.

4.3.2 *The zone of influence*

The zone of influence is the zone beyond which the influence of a sample disappears. It is normal to characterize that zone of influence in a given direction by the distance at which the variogram eventually reaches a plateau. Typical zones of influence can be seen in Figs. 56 and 53. In Fig. 56 we have a 200′ zone of influence in an iron deposit. In Fig. 53 the range of influence (we will say in short, range) was only 35′. In Fig. 55 one can say that it is less than one sampling interval, i.e. virtually zero.

Also note that in some instances the further apart the samples, the more different they are. This happens frequently in hydrothermal deposits: the variogram shown in Fig. 57 is from a Cu–Ni deposit. When a range exists, the variogram is said to be of a transition type.

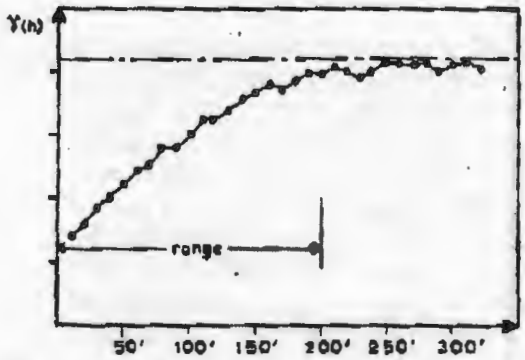


Fig. 56. A very "good" variogram from an iron ore deposit showing a 200′ (vertical) zone of influence.

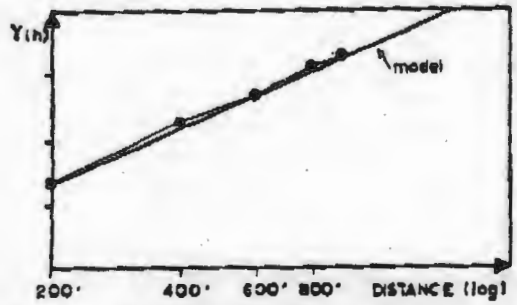


Fig. 57. Variogram of nickel grade in a sulfide vein.

4.3.3 *The anisotropies*

The anisotropies are easily depicted by computing the variogram in different directions. On Fig. 58 for instance one sees four variograms from four directions in a porphyry molybdenum. They can be considered as identical,

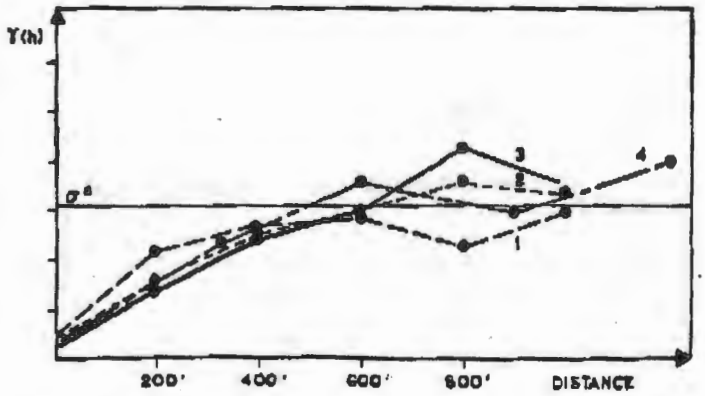


Fig. 58. Four variograms corresponding to four different directions in a porphyry molybdenum deposit. (*1* = E–W, *2* = N–S, *3* = vertical, *4* = cube diagonal) showing no change of behaviour with direction.

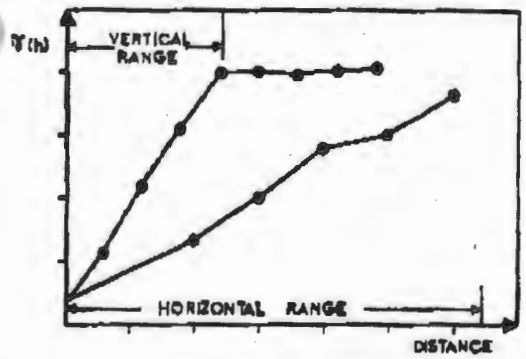


Fig. 59. Horizontal and vertical variograms in an iron mine, showing difference of behaviour with direction.

the continuity of the ore is thus the same in all directions. This is the simplest case but, unfortunately, the least frequent! A more common situation is seen on Fig. 59. The vertical and horizontal range of influence in an iron deposit differ widely. This can be interpreted by saying that a given vertical distance is equivalent to the same horizontal distance multiplied by an anisotropy factor k (greater than one in this case). For instance if the anisotropy factor is 3, then ten vertical feet are equivalent to thirty horizontal feet as far as grade variations are concerned.

Other types of anisotropies are encountered. For instance, when a layering of the mineralization exists, then the global variation in any direction can be thought of as the sum of two independent sources of variation, one isotropic component plus a zonal component, which depends only on the number of layers crossed. The difference between the grades of two points depends

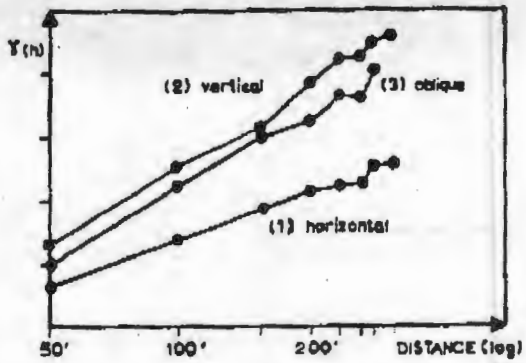


Fig. 60. Difference of continuity in three directions for the thickness of a coal seam, as shown by the three variograms in these directions.

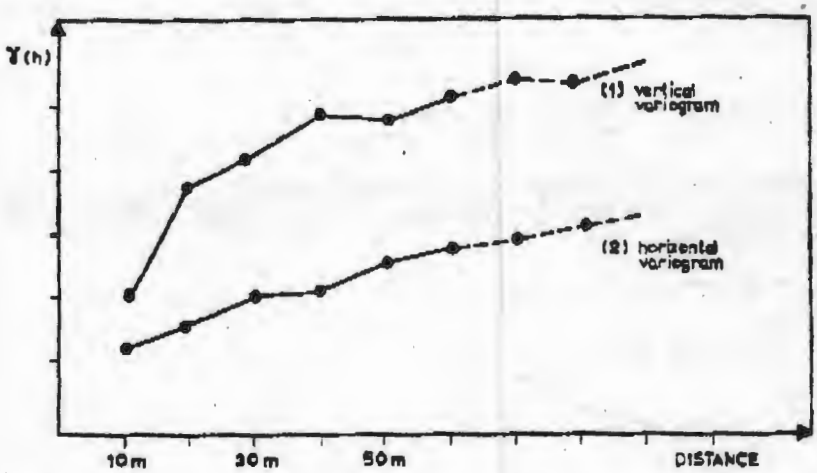


Fig. 61. Difference of continuity in the horizontal and vertical direction at the Exotica Mine, as shown by the variograms.

thus not only on their relative distance, but also on their difference of elevation in the layered system. Note that the layering may not be visible. Such variograms are more complex to interpret. Two different types can be seen in Fig. 60 and 61.

4.3.4 *Conclusion*

We see that the variogram takes into account all the geological features which we mentioned earlier. Thus we can expect that an estimation procedure based on it can be geologically reliable. We will now show how the estimation variance is a function of the variogram.

4.4 THE VARIOGRAM AS THE FUNDAMENTAL FUNCTION IN ERROR COMPUTATIONS

Although we have not yet given the theoretical probabilistic models describing the grade of a deposit as a random function, we can write a few lines of statistics in order to appreciate what is exactly required in error estimation.

4.4.1 *The variance of the error of estimation*

Let us try to appreciate the variance of the error of estimation ϵ_i which we mentioned before. The $Z(V_i)$ are true unknown grades of the blocks V_i, the $Z^*(V_i)$ are the linear combinations $Z^*(V_i) = \Sigma_{j=1}^{n} a_j Z(X_j)$ of the known sample grades at locations $X_j (j = 1, \ldots, n)$. Then:

$$\mathrm{VAR}(\epsilon_i) = \mathrm{VAR}[Z(V_i) - Z^*(V_i)]$$

$$\mathrm{VAR}(\epsilon_i) = \mathrm{VAR}[Z(V_i)] - 2\,\mathrm{COV}[Z(V_i), Z^*(V_i)] + \mathrm{VAR}[Z^*(V_i)]$$

$$\mathrm{VAR}(\epsilon_i) = \mathrm{VAR}[Z(V_i)] - 2 \sum_j a_j \,\mathrm{COV}[(Z(V_i), Z(X_j)] + \sum_j \sum_k a_j a_k \,\mathrm{COV}[Z(X_j), Z(X_k)] \tag{4.4}$$

If we consider each of these VAR, and COV separately, we will see that they should be computable from the variogram.

4.4.2 *The variance of the grade of blocks: VAR [Z(V$_i$)]*

The variability of the grade of blocks should depend on the average difference in grade which exists between any two points inside the block. If this average difference is low, then the variance of the block should be low. If it is high, the variance will be high. Since the difference between the grades at two points is expressed by the variogram, we can say that we should be able to compute the variance of the blocks from the variogram. This will be strictly checked and proved later.

4.4.3 *The covariance of the grade of a block and the grade of a sample: COV [Z(V$_i$), Z(X$_j$)]*

This value expresses the correlation which exists between the grade of a block and the grade of a given sample. It should again obviously be equal to the average correlation between any point in the block and a point in the sample. This correlation being expressed by the variogram, we thus can state that once we know the variogram we will be able to know the covariances.

4.4.4 *The covariance of the grades of two samples: $COV[Z(X_j), Z(X_k)]$*

After what we have just said we see that this is just a particular case of $COV[Z(V_i), Z(X_j)]$. This time we know that the correlation or rather, the lack of correlation between two samples is directly expressed by the variogram function. Thus $COV[Z(X_j), Z(X_k)]$ should be some constant minus $\gamma(h_{jk})$ if h_{jk} is the distance between samples X_j and X_k.

4.5 CONCLUSION

We have now completed our intuitive examination of the problem of ore reserves estimation. We have reached the stage where we know theoretically, how to compute the magnitude of the error involved in one procedure or the other. Thus we should be able to select the best method, the one which will give the minimum variance. What is left now is to firmly establish the theory which is called the Theory of Regionalized Variables and, secondly, to examine which practical computational steps and interpretations are involved in practical cases. The actual optimization procedure — kriging — will be seen later in Chapter 9. Before we proceed, we will show in two exercises how the variogram helps to compute all what we need and how to compute it in a simple case. Those who are not concerned with the theoretical basis or practical aspects of variograms and variance computations can directly proceed to Chapter 9.

4.6 EXERCISES

4.6.1 *Variances and the variogram*

This exercise is the follow-up of Exercise 3.2.5. In that previous exercise, we had shown which physical realities our concepts were covering, namely what block variances and estimation variances are. At that time, in order to compute these values, we have had to use block grades which are normally only available after the material has been mined out, i.e. too late. In fact we never know the real grade of blocks, so that in daily mining practice it is not possible to compute the variances which we have just obtained, in the way we did.

Geostatistics has shown that this is, however, possible. From one drilling campaign one can estimate the variogram of the deposit and then fit a model to it.

This defines a few parameters like the zone of influence in all directions (which we call the range, a), the sampling error (which we call the nugget effect, C_0) and the variance of the samples (which we call the sill, $C + C_0$). Using these parameters and charts we can obtain the variances which we had before. How these charts have been computed is irrelevant so far. All what we need to know is that they can be deduced from samples only.

apter 9

TIMIZATION OF THE GRADE ESTIMATION: KRIGING

NOPSIS — *This chapter gives the theoretical answer to the problem of optimum estimation of block grades. This is kriging. After a theoretical roduction and a complete example, some time will be spent to review ige's original work and the correction factors presently in use in the lustry. Capital properties unique to kriging and essential for planning will pointed out and a simple kriging program, sufficient for 2-D problems will proposed. A last section will deal with universal kriging and non-station- phenomena, and the practical ways to cope with these.*

Foreword. This chapter could have been placed just after Chapter 4 and it uld be reduced to one paragraph. This is probably the feeling a statistician ll have after reading it. This would, however, mean not recognizing all the rti r problems which arise due to the fact that we are only concerned ing problems and that we need economic solutions for any situ- n. Kriging is in fact a word which has been coined to cover both "best ear unbiased estimator" (B.L.U.E.) of a point and the best linear weighted oving average of a block. The particular nature of estimation problems in ne planning is such that it most probably deserves the use of a special me. This name appeared around 1960 to designate a technique created in ance by Matheron, after the name of D.G. Krige who was probably the st to make use of spatial correlation and B.L.U.E. in the field of mineral ources evaluation (1951).

In Chapter 4, the problem of the accuracy of the estimation of the grade a block of ore, from a series of samples $x_i (i = 1, \ldots, n)$ has been for- ulated. As it has been possible to express the precision of a linear com- nation as a function of the variogram and the weights chosen, it should be ssible to find a set of weights which minimizes the estimation variance. will be shown how this can be done, then how it was originally done by ige and how this can be linked to various correction factors traditionally ed in the industry. Finally the problem of dealing with trends by means universal kriging will be discussed. Doing it this way rather than following e historical development will make discussions easier, and probably the nk between all earlier attempts will become obvious once the terminology kriging is understood.

1 THE GENERAL PROBLEM AND ITS SOLUTION

Considering a block V having a true unknown grade $Z(V)$, and a series of samples of known grade $Z(x_i) (i = 1, \ldots, n)$ we want to find that set of

weights $a_i (i = 1, \ldots, n)$ which will make the weighted average

$$Z^* = \sum_{i=1}^{n} a_i Z(x_i)$$

the best estimator of $Z(V)$.

Best can be qualified in many ways. In Chapter 4, it was found that a natural way to measure the quality of an estimation procedure is to compute the variance of the error it involves. The estimation variance of $Z(V)$ by Z^* was expressed as:

$$\sigma_e^2 = \sigma_V^2 - 2 \sum_{i=1}^{n} a_i \sigma_{Vx_i} + \sum_i \sum_j a_i a_j \sigma_{x_i x_j} \tag{9.1}$$

where:

σ_V^2 stands for VAR$[Z(V)]$, the variance of the grade of blocks like the block V studied.

σ_{Vx_i} stands for COV$[Z(V), Z(X_i)]$, the covariance of the grade of block V and the grade of sample X_i.

$\sigma_{x_i x_j}$ stands for COV$[Z(X_i), Z(X_j)]$, the covariance of the grades of sample X_i and sample X_j.

Now as all these σ coefficients can be computed from the variogram it is possible to minimize σ_K^2 with respect to the a_i's in order to obtain that weighted average which will have the smallest estimation variance. This is just a matter of calculus.

So we will take as a definition of best, the minimum variance linear estimator. This is not sufficient. It is also required that on average, the value which is computed, should be equal to the real value, rather than systematically higher or lower. This is the unbias condition.

Unbiased means that one should have:

$$E(Z^*) = m \text{ which requires that } E\left[\sum_i a_i Z(X_i)\right] = m$$

$$\sum_i a_i E[Z(X_i)] = m \text{ and as } E[Z(X_i)] = m \text{ then } \sum_i a_i = 1$$

Altogether what is desired is the best linear unbiased estimator; the B.L.U.E. of statisticians.

The problem. To find those weights a_i which minimize the estimation variance σ_e^2 under the condition that $\Sigma_i a_i = 1$.

The solution. To minimize a function Q of a_i's one knows that the derivatives of the function with respect to the a_i's should be made equal to zero. Now when there is a constraint $(C = 0)$ the Lagrange principle tells us that $F = Q + 2\mu C$ should be minimized, where μ is a new unknown, the Lagrange multiplier. In our case, we should thus take the derivative of:

$$F = \sigma_e^2 + 2\mu\left(\sum_i a_i - 1\right)$$

with respect to all unknowns (a_i's and μ) and make them equal to zero. Writing the detail of computations yields:

$$F = \sigma_v^2 - 2\sum_i a_i\sigma_{vx_i} + \sum_i\sum_j a_i a_j \sigma_{x_i x_j} + 2\mu\left(\sum_i a_i - 1\right)$$

and the derivatives are:

$$\begin{cases} \dfrac{\partial F}{\partial a_i} = -2\sigma_{vx_i} + 2\sum_j a_j\sigma_{x_i x_j} + 2\mu = 0 \quad (i = 1, \ldots, n) \\ \dfrac{\partial F}{\partial \mu} = \sum_i a_i - 1 = 0 \end{cases}$$

This is a linear system of $n + 1$ equations with $n + 1$ unknowns, the a_i's and μ. It can be written in the usual form:

$$\begin{cases} \sum_j a_j\sigma_{x_i x_j} + \mu = \sigma_{vx_i} \quad (i = 1, \ldots, n) \\ \sum_i a_i = 1 \end{cases} \tag{9.2}$$

or in matrix form:

$$[\Sigma]\,[A] = [D]$$

where Σ, A and D are, respectively (using σ_{ij} as a short for $\sigma_{x_i x_j}$):

$$[\Sigma] = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} & 1 \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} & 1 \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \qquad [A] = \begin{bmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_n \\ \mu \end{bmatrix} \qquad [D] = \begin{bmatrix} \sigma_{vx_1} \\ \sigma_{vx_2} \\ \cdot \\ \cdot \\ \cdot \\ \sigma_{vx_n} \\ 1 \end{bmatrix}$$

Note that $[\Sigma]$ is a symetric matrix which depends only on the known samples while [D] depends on both the unknown block and samples.

The solution is thus $[A] = [\Sigma]^{-1}[D]$. The problem is solved as all σ's are derived from the variogram. Let us see quickly how this is done in simple cases.