Reprinted from
Economic Geology
Vol. 58, 1963, pp. 1246–1266

PRINCIPLES OF GEOSTATISTICS

G. MATHERON

ABSTRACT

Knowledge of ore grades and ore reserves as well as error estimation of these values, is fundamental for mining engineers and mining geologists. Until now no appropriate scientific approach to those estimation problems has existed: geostatistics, the principles of which are summarized in this paper, constitutes a new science leading to such an approach. The author criticizes classical statistical methods still in use, and shows some of the main results given by geostatistics. Any ore deposit evaluation as well as proper decision of starting mining operations should be preceded by a geostatistical investigation which may avoid economic failures.

RESUME

Pour tout mineur et géologue minier, la connaissance des teneurs et du tonnage et l'appréciation des erreurs sur ces grandeurs est fondamentale. Or, jusqu'à présent, il n'existait pas d'approche scientifique correcte de ces problèmes.

La *géostatistique*, dont les principes sont résumés dans cet article, est la nouvelle science qui permet cette approche. L'auteur indique les méthodes statistiques antérieures et encore courantes et donne quelquesuns des résultats principaux de la géostatistique.

Toute évaluation de gisement et toute décision de mise en exploitation devrait être précédée d'une étude géostatistique permettant de limiter le risque d'une déconvenue ultérieure.

INTRODUCTION AND SHORT HISTORICAL STATEMENT

GEOSTATISTICS, in their most general acceptation, are concerned with the study of the distribution in space of useful values for mining engineers and geologists, such as grade, thickness, or accumulation, including a most important practical application to the problems arising in ore-deposit evaluation.

Historically geostatistics are as old as mining itself. As soon as mining men concerned themselves with foreseeing results of future works and, in particular as soon as they started to pick and to analyze samples, and compute mean grade values, weighted by corresponding thicknesses and influence-zones, one may consider that geostatistics were born. In so far as they take into account the space characteristics of mineralization, these traditional methods still keep all their merit. Far from disproving them, modern developments of the theory have adopted them as their starting point and have brought them up to a higher level of scientific expression.

However, assuming they could provide a correct evaluation of mean values, the traditional methods failed to express in any way an important character of mineralizations, which is their variability or their dispersion. Some scores of years ago, classical probability calculus techniques began to be used in order to take into account this characteristic. If an unskillful application of those techniques has sometimes led to absurdities, it remains certain that, on the whole, results have been profitable. In a way this is a paradox, for classical statistical methods, in so far as they are not concerned with the spatial aspect of the studied distributions, actually cannot be applied. As a matter of fact, the South-African school, which has recorded the most remarkable results with Krige, Sichel, used to say, and believed that they were applying classical statistics. But the methods they were developing differed more and more from classical statistics, and adjusted themselves spontaneously to their object.

The second decisive change appeared when the insufficiency of classical probability calculus was clearly understood as well as the necessity of re-introducing the spatial characters of the distributions. It consisted in realizing on a higher level the synthesis between traditional and statistical methods. Hence, geostatistics started elaborating their own methods and their own mathematical formalism, which is nothing else than an abstract formulation and a systematization of secular mining experience. This formalism has inherited from its statistical origin a language in which one still speaks of variance and covariance, including however in those notions a new content. This similarity in vocabulary must not deceive. At the end of a protracted evolution, the geostatistical theory had to admit that it was facing, instead of random occurrences, natural phenomena distributed in space. And, therefore, its methods are approximately these of mathematical physics and more specially those of harmonic analysis.

INSUFFICIENCY OF CLASSICAL STATISTICAL CONCEPTS

To be brief, we shall limit ourselves, in what follows, to the distribution of ore-grades in a deposit. The results that will be obtained will however have a general range and will be applicable to any character owned by a spatial distribution. In an usual statistical approach, the grades of samples picked in a deposit are classified on a histogram. Such a procedure does not take into account the location of samples in the deposit. But it is not enough to know the frequency of a given ore-grade in a deposit. It is also necessary to know in what way the different grades follow each other on the field, and specially what is the size and the position of economic orebodies. At the starting point of the theory we have to face one fact: the inability of common statistics to take into account the spatial aspect of the phenomenon, which is precisely its most important feature.

More precisely, the aim of the classical probability calculus is the study of aleatory variables. The mere example of the heads or tails game shows clearly what is going on. Let us record $+1$ each time the coin falls on tails and -1 in the opposite case. Before throwing the coin, there is no way of forecasting whether $+1$ or -1 will be recorded; we only know that there is one chance out of two for one or the other of these two opportunities. An aleatory variable has classically two essential properties: 1) The possibility, theorically at least, of repeating indefinitely the test that assigns to the variable a numerical value; we can for example, throw the coin as often as we want. 2) The independence of each test from the previous and the next ones; if all the 100 first attempts have given tails, there remains however one chance out of two for the 101st attempt to give heads.

It appears clearly that a given ore-grade within a deposit cannot have those two properties. The content of a block of ore is first of all unique. This block is mined only once and there is no possibility of repeating the test indefinitely. When the grade of a sample is concerned, which may be a groove sample of a given size for example, the result is exactly the same, because the grade of a groove located in a point with coordinates (x, y) is unique and well determined. However it is possible to pick a second sample close to the first, then a third one, etc. . . . which shows an apparent possibility of repeating the test. Actually, it is not exactly the same test but a slightly different one. But even assuming this possibility of repetition, the second property will surely not be respected. Two neighboring samples are certainly not independent. They tend, in average, to be both high-grade if they originate from a high-grade block of ore, and vice-versa. This tendency, more or less stressed, expresses the degree of more or less strong continuity in the variation of grades within the mineralized space.

The misunderstanding of this fact and the rough transposition of classical statistics has sometimes led to surprising misjudgments. Around the fifties, in mining exploration, it was advised to draw lots to locate each drilling (i.e., to locate them exactly anywhere). Miners of course went on still using traditional regular grid pattern sampling, and geostatistics could later prove they were right. Or else again, it was urged that the accuracy of ore evaluation of a deposit depended only on the number of samples (and not on their location) and varied as the square root of this number. This unskillful transposition of the theory of errors led to absurdities. For example, if a given deposit is explored by drilling, it would suffice to cut the cores in 5 mm pieces instead of 50 cm pieces to obtain 100 times more samples, and therefore 10 times higher accuracy. This, of course, is wrong. The multiplicity of samples thus obtained is a fallacy, and does nothing more than repeat indefinitely the same information, without yielding anything else. Geostatistics actually show that accuracy is the same with pieces of 5 mm and 50 cms, as every miner understands instinctively.

NOTION OF REGIONALIZED VARIABLE

Thus a grade cannot in any way be assimilated to an aleatory variable. We speak of regionalized variables precisely in order to stress the spatial aspect of the phenomena. A regionalized variable is, *sensu stricto*, an actual function, taking a definite value in each point of space.

In general such a function has properties too complex to be studied easily through common methods of mathematical analysis. From the point of view of physics or geology, a given number of qualitative characteristics are linked to the notion of regionalized variable.

a) In the first place, a regionalized variable is *localized*. Its variations occur in the mineralized space (volume of the deposit or of the strata), which is called *geometrical field* of the regionalization. Moreover such a variable is in general defined on a *geometrical support* (*holder*). In the case of an ore-grade, this support is nothing but the volume of the sample, with its geometrical shape, its size and orientation. If, in the same deposit, the geometrical support is changed, a new regionalized variable is obtained, which shows analogies with the first one, but does not coincide with it.

For instance, samples of 10 Kg corresponding to drill cores are not distributed in the same way as samples of 10 tons corresponding to blasts. Often the case of a punctual support will be considered. A punctual grade, for example, will take value 0 or value +1 according to whether its support will fall into a barren or mineralized grain.

b) Secondly, the variable may show a more or less steady continuity in its spatial variation, which may be expressed through a more or less important deviation between the grades of the two neighboring samples mentioned above. Some variables with a geometrical character (thickness or dip of a geological formation) are endowed with the strict continuity of mathematicians. Fairly often (for grades or accumulations) only a more lax continuity will exist or, in other words, a continuity "in average." In some circumstances, even this "in average" continuity will not be confirmed, and then we shall speak of a *nugget effect*.

c) Lastly the variable may show different kinds of *anisotropies*. There may exist a preferential direction along which grades do not vary significantly, while they vary rapidly along a cross-direction. Those phenomena are well known under the names of runs, or zonalities.

To those general characters, common to any regionalized variable, specific features can be superimposed. For example, in the case of a sedimentary deposit, a *stratification effect*, will be noted. Large-scale stratification provides individualizable and separately minable strata. Inside each strata it may appear by the existence of beds following one another vertically, and separated by discontinuity surfaces. The grade, almost constant or barely varying inside a given bed, will vary abruptly from one bed to another; however common and familiar this phenomenon appears to be, it is still fundamental, and a theoretical formulation of the problem that would not take it into account would miss the point. It will happen as well that to those vertical discontinuities, stressed by jointing, will be added lateral discontinuities, owing to the lenticular endings of beds. This *bed-relaying phenomenon*, when it does exist, shows up at each stratigraphic level a partitioning of the sedimentation area into micro-basins with almost autonomous evolution, and may appear during operation through *grade-limit effect*.

In the same way in stockwerk types of deposits, high-grade veinlets or granules individualized in a more or less impregnated mass will be observed. This *stockwerk effect*, just as the stratification and bed relaying effects, expresses the appearance of a discontinuity net-work within a homogeneous geometrical field. On a very different scale, that of granularity, the nugget effect appears as a phenomenon of the same nature, the net-work of discontinuities being here that one separating barren from mineralized grains.

Those different specific aspects of spatial distribution of regionalized variables—far apart from classical probability calculus—must compulsorily be taken into account by geostatistics. This is made possible owing to a simple mathematical tool: the *variogram.*

THE VARIOGRAM

The variogram is a curve representing the degree of continuity of mineralization. Experimentally, one plots a distance d in abcissa and, in ordinate, the mean value of the square of the difference between the grades of samples picked at a distance d one from the other. Theoretically, let $f(M)$ be the value taken in a point M of the geometrical field V by a regionalized variable defined on a given geometrical support v (in general support v will be small and the limit may be considered as punctual). The *semi-variogram* $\gamma(h)$, or law of dispersion, is defined, for a vectorial argument h, by the expression:

$$\gamma(h) = \frac{1}{2V} \iiint_V [f(M+h) - f(M)]^2 dV. \qquad (1)$$

In general, the variogram is an increasing function of distance h, since, in average, the farther both samples are one from the other, the more their grades are different. It gives a precise content to the traditional concept of the *influence zone* of a sample. The more or less rapid increase of the variogram represents, indeed, the more or less rapid deterioration of the influence of a given sample over more and more remote zones of the deposit. The qualitative characteristics of regionalization are very well expressed through the variogram:

a) The greater or lesser regularity of mineralization is represented by the more or less regular behavior of $\gamma(h)$, near the origin. It is possible to distinguish roughly four types (Fig. 1). In the first type the variogram has

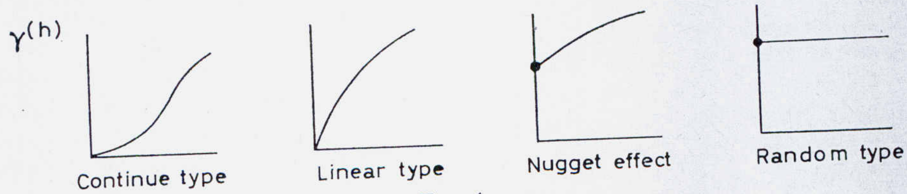

FIG. 1.

a parabolic trend at the origin, and represents a regionalized variable with high continuity, such as a bed-thickness.

The second type, or linear type, is characterized by an oblique tangent at the origin, and represents a variable which has an "in average" continuity. This type is the most common for grades in metalliferous deposits.

The third type reveals a discontinuity at the origin and corresponds to a variable presenting not even an "in average" continuity, but a nugget effect.

The fourth type is a limit case corresponding to the classical notion of random variable. Between type I (continuous functional) and type 4 (purely random) appears a range of intermediates, the study of which is the proper object of geostatistics.

b) The variogram is not the same along different directions of the space. Function $\gamma(h)$ defined in (1) does not only depend upon the length, but also upon the direction of vector h. Preferential trends, runs, and shoots are revealed through the study of the distortion of variogram when this direction is altered. Geological interpretation of such anisotropies is often instructive.

c) Structural characters are also reflected in the variogram. For instance, the bed-relaying phenomenon appears in the experimental curve as a level stretch of the variogram beyond a distance, i.e., a range equal to the mean diameter of the autonomous micro-basins of sedimentation. And the

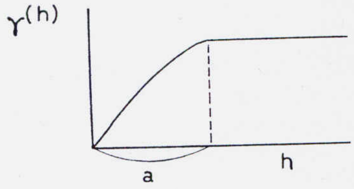

FIG. 2.

fact that these ranges are not the same along different directions makes it possible to determine the directions of elongation, and the average shape of the micro-basins.

This tool, the variogram, does not represent the totality nor the local details of the mineralizing phenomenon, but it expresses in a synthetic form their essential characters. The harmonic analysis of a vibratory phenomenon assigns for each harmonic a phase and an amplitude. The local outline of the phenomenon depends mostly upon phases, but energy depends only upon the square of amplitudes. The spectral curve giving the squares of the amplitudes does not describe the whole phenomenon but gives an account of the essential, i.e., the energetic characteristics. The variogram (or more precisely its Fourier's transformed curve) plays exactly the part of such a spectral curve.

In the following paragraphs, a few of the possible applications of variograms will be run over. It is obviously out of the question to give here a systematic study. I will merely mention some examples and several characteristic formulae. For more details I kindly ask the reader to refer to my "Treatise of Applied Geostatistics." [1]

[1] Editions *Technip*, Paris, Tome I (1962)—Tome II (Le Krigeage) (in press). Tome III (l'effet de pépite et les phénomènes de transition) to be published.

ABSOLUTE DISPERSION (OR INTRINSIC) LAW

The semi-variogram defined in (1) is bound to the geometrical field V of the regionalized variable. If, instead of the total field V, only a portion V' of it would have been considered, a function $\gamma'(h)$ possibly different from $\gamma(h)$ would have been obtained. However, we have the intuitive notion that in a geologically homogeneous geometrical field there might be something intrinsic, independent from location, in the characteristics representing the variabilities of regionalized variable.

Formulated in an accurate way, that intuition leads to the hypothesis of an absolute dispersion or intrinsic law expressed through the equation:

$$\gamma'(h) = \gamma(h)$$

which means that the variogram is independent from the portion V' of the deposit V selected for its calculation. It may be said at once, that this hypothesis is not really essential to the development of the theory and it is possible to eliminate it because of some mathematical complications.[2]

Nevertheless it makes the statement of the theory much easier, and for that reason it will be followed here. A slow deviation of the variogram in space is generally ascertained through experience and if this drift does not take too much importance, the results yielded by the hypothesis of an absolute dispersion law provide an excellent approximation of reality (on the condition that $\gamma(h)$ actually employed has been calculated from the actual portions of the considered deposit).

When this hypothesis is verified, the semivariogram $\gamma(h)$ itself acquires an intrinsic significance. It is often designated under the name of intrinsic (or absolute) *dispersion law* or, more shortly, intrinsic function of regionalized variable.

VARIANCES AND COVARIANCES

Let us consider in the first place, a regionalized variable (which will be called grade in order to simplify) defined in a field V, on a punctual support and submitted to an intrinsic dispersion law $\gamma(h)$. Let $f(M)$ be the value taken by the grade in a point M of the field V. Instead of the punctual grade $f(M)$ we usually are concerned with the grade $y(M)$ of a sample v, of a given size, shape and orientation, picked at point (M).[3] This new variable is deducted from the previous one through an integration performed within the volume v centered in M.

$$y(M) = \frac{1}{v}\int_v f(M+h)dv. \tag{2}$$

To this variable will be bound a parameter measuring its dispersion inside V, called variance, as in classical probability calculus. The mean

[2] Loc. cit., Tome III.

[3] This means that the center of gravity of v is located at point M.

value of the punctual variable inside V being m,

$$m = \frac{1}{V}\int_V f(M)dV$$

the variance of $y(M)$ inside V is defined as the average value within V of the square of the expression $[y(M) - m]$, let:

$$\sigma^2 = \frac{1}{V}\int_V [y(M) - m]^2 dv. \tag{3}$$

It will be noted that this notion has, at the outset, a geometrical and not a probabilistic meaning. It will not deter us from calculating these variances, in the applications, from experimental data with common statistical methods. Should they be taken according to their spatial order, as in integral (3), or previously rearranged in histograms, the same expressions $(y - m)$ are appearing, with the same weights in both the calculation procedures. But, on a conceptual ground, definition (3) has a physical content that the statistical motion has not. From expression (1) of the variogram, of (2), and the definition (3) of the variance, one may deduce, reversing the order of the integrations.

$$\sigma^2 = \frac{1}{V^2}\int_V dV\int_V \gamma(h)dV' - \frac{1}{v^2}\int_v dv\int_v \gamma(h)dv'. \tag{4}$$

Each one of these sextuple integrals has a very clear meaning: it represents the average value of the $\gamma(h)$ inside V(or v) when both the extremities of vector h sweep, each one for its own account, the volume V(or v).

If we write:

$$F(V) = \frac{1}{V^2}\int_V dV\int_V \gamma(h)dV,$$

i.e., $F(V)$ = average value of $\gamma(h)$ inside V.

One gets:

$$\sigma^2 = F(V) - F(v). \tag{5}$$

Thus knowledge of the variogram of punctual grades allows the "a priori" calculation of the variance of any sample v within any portion V of a deposit. It will be noted that this variance does not depend only upon the sizes of volumes v and V, but also upon their shapes and orientation.

Physical meaning of relation (4) is highly instructive. The variance of a macroscopic sample v, considered as the juxtaposition of a great number of microsamples dv, does not depend in any way on the number of those micro-samples nor on their variances, but only on the average value of intrinsic function $\gamma(h)$ inside the geometrical volume v. Classical statistics, considering these micro-samples as independent, should lead to a variance in terms of $1/v$. There does not actually exist any deposit in which 10 ton blasts would have a variance a thousand times lower than that of 10 kg cores. Formula (4) shows why. The grades of micro-samples are not independent at all. They are inserted into a spatial correlation lattice, the nature of which is bound to the more or less steady continuity of mineralization, and which is expressed precisely through the intrinsic dispersion law $\gamma(h)$. The grades of the micro-samples are much less different, on the average, than classical statistics would indicate, and in consequence 10 ton blasts have a much higher variance than the thousandth of the variance of 10 kg cores.

The expression of the variance in form (5) shows a law of additivity. If we consider panel V' and samples v within a field V and if $\sigma^2(V',V)$, $\sigma^2(v,V)$ and $\sigma^2(v,V')$ designate the variances of V' inside V, of v inside V and v inside V', we get:

$$\sigma^2(v,V) = \sigma^2(v,V') + \sigma^2(V',V).$$

This formula is known as *Krige's Formula*. It has been established by D. G. Krige in the case when the grades are distributed according to a (statistical) lognormal law. Its validity is actually not linked to a special statistical distribution law, but only to the existence of an intrinsic dispersion law.

Besides the variance, geostatistics introduce the notion of *covariance*. If $y(M)$ and $z(M+h)$ are the grades of two samples v and v' centered in two points M and $M+h$, covariance (inside V) of y and z is the function of h defined by:

$$\sigma_{yz} = \frac{1}{V^2}\int_V [y(M) - m][z(M+h) - m]dv.$$

It can be expressed through the variogram with a relation similar to (4):

$$\sigma_{yz} = F(V) - \frac{1}{vv'}\int_v dv \int_v \gamma(k)dv'. \tag{6}$$

The second integral represents the average value of $\gamma(k)$, when both extremities of vector k sweep, respectively, volume v and volume v', at a distance h one from the other.

Let us consider, as a particular case, the isotropic de Wijs's [4] scheme. It is defined by an intrinsic isotropic function of the form:

$$\gamma(r) = 3\alpha \ln r \tag{7}$$

in which $r = |h|$ represents the modulus of the vectorial argument h, or otherwise the distance between the two points M and $M+h$. When symbol ln represents the natural logarithm, parameter α is called *absolute dispersion*. It characterizes indeed the dispersion of grades independently from the shape and the volume of the samples and of the deposit. In the

[4] The starting point of development of the present theory is the original De Wijs's reasoning which is a remarkable example of transition from classical statistics to geostatistics. Reference to "Traite de Geostatistique Appliquee," where bibliographical references will be found.

particular case where the volume of the samples is geometrically similar to the volume V of the deposit, formulae (4) and (7) give:

$$\sigma^2 = \alpha \ln \frac{V}{v}. \tag{8}$$

This formula, which is the *Wijs's formula*, does express a principle of similitude. It ceases to be appliable generally as soon as the deposit is not geometrically similar to the samples. It is however possible to associate to any geometrical volume v its *linear equivalent* d devined by relation:

$$\ln d - \frac{3}{2} = \frac{1}{v^2} \int_v dv \int_v \ln r dv'. \tag{9}$$

Formula (4) entails that sample v has the same variance in any deposit as the linear sample of length d. If D and d are the linear equivalents of the deposit and of the samples respectively, the variance may be set into the form:

$$\sigma^2 = 3\alpha \ln \frac{D}{d}.$$

The linear equivalents have been calculated and tabulated for a certain amount of geometrical figures, and, in addition, we have at our disposal some simple approximation formulae. For example for a rectangle with sides a and b we have:

$$d = a + b.$$

For a parallelogram, with sides a, b, and surface S:

$$d = \sqrt{a^2 + b^2 + 2S}.$$

For a triangle with sides a, b, c, and Surface S:

$$d = \sqrt{\frac{a^2 + b^2 + c^2}{3} + 2S}.$$

For a trapezium with basis $a = \dfrac{L + l}{2}$,

$$b = \frac{L - l}{2}$$

Median: m

Surface: S

$$d = \sqrt{L^2 + l^2 + m^2 - \frac{l^2 m^2}{3L^2} + 2S}$$

For a rectangular parallelepiped with sides $a > b > c$,

$$d = a + b + \frac{c}{2}.$$

For an oblique parallelepiped with edges r_1, r_2, r_3, faces S_1, S_2, S_3 and volume V, we put up:

$$\begin{cases} R^2 = r_1^2 + r_2^2 + r_3^2 \\ S^2 = S_1^2 + S_2^2 + S_3^2, \end{cases}$$

and we obtain the following approximate equivalent:

$$d = \sqrt{R^2 + 2S + \frac{V^2R^2}{S^3}}.$$

This notion of linear equivalent allows an easy comparison between samplings of different natures, at least in the case, common in metalliferous deposits, where the law of dispersion has the form (7).

ESTIMATION VARIANCE AND EXTENSION VARIANCE

One of the most practical problems geostatistics are supposed to resolve is the size of the possible error in the evaluation of a deposit. The general characteristics of regionalized variables indicate that this error does not only depend upon the amount of picked samples, but first of all upon their shapes, their sizes and their respective locations, in other words, on the whole, upon *the geometry of achieved mining workings*. These indications get a precise meaning through the geostatistical notion of the estimation variance. Let us suppose that, in order to estimate the real unknown grade z of a deposit or of a panel V, we know the grade x of a given net-work of mining workings Mw. The estimation error $(z - x)$ has a simple, well determined value, although unknown, for a given panel V as for the net-work Mw located preferentially. In order to make out of this error a regionalized variable, geostatistics consider the panel or the deposit to be estimated as a panel extracted from a very large fictive deposit K. This deposit is supposed to be ruled by the intrinsic dispersion law $\gamma(h)$ defined by the experimental variogram controlled in mining works Mw. We shall see that the shape and the sizes assigned to K do not actually intervene. Let us imagine that panel V which is being estimated travels across the large deposit K, drawing with its attached mining works, the error $(z - x)$ then appears as a regionalized variable with an average value equal to zero and a variance:

$$\sigma^2 = \sigma_z^2 + \sigma_x^2 - 2\sigma_{zx}. \tag{10}$$

This variance called *estimation variance* is calculated after variances σ_z^2, σ_x^2 and covariance σ_{zx} of the variables z and x inside the field K, which are themselves given by formulae of type (4) or (6). Field K interferes in the expression of σ_z^2; σ_x^2 and σ_{xz} by the simple constant $F(K)$ which is eliminated in equation (10), so that the estimation variance σ^2 is independent from the choice of K, and is calculated after the formula:

$$\sigma^2 = \frac{2}{VV'}\int_V dV \int_{V'} \gamma(h)dV' - \frac{1}{V^2}\int_V dV \int_V \gamma(h)dV' - \frac{1}{V^2}\int_{V'} dv \int_{V'} \gamma(h)dv'. \quad (11)$$

In (11) V is the volume of the deposit being estimated and V' that of mining works Mw. The estimation variance σ^2 is calculated after integration of the intrinsic function $\gamma(h)$ inside the geometrical volumes of the deposit and of the samples. In the same way, as the variogram could give to the concept of the influence zone of a sample a precise content, one may say that the estimation variance (11) can give a precise meaning to the "influence" of mining works over the whole deposit.

In practical calculations, formula (11) should be difficult to use. Mining works usually frame a discontinuous net-work in which the samples themselves may be picked discontinuously (for instance, groove samples cut off according to a regular grid pattern in the drifts on a vein developed at different levels). Volume V' interfering in (11) is the discontinuous volume set up by the lattice of samples actually cut off and analyzed. An influence zone is traditionally assigned to each individual sample, in the center of which it is located and supposed to represent the grade. The error usually performed in extending the grade of such an individual sample to its influence zone can be represented by a type (11) variance, where V is the volume of influence zone and V' that of the sample.

Such a variance is called *elementary extension variance* and can be calculated for a given $\gamma(h)$ in terms of geometrical parameters of the sample and its influence-zone. On condition of certain approximation hypothesis, it is possible to prove that an estimation variance of type (11) can be calculated by composing the elementary extension variances.

In practice, two cases are to be distinguished essentially. The elementary samples network, for an isotropic function $\gamma(h)$ may be isotropic[5] or not. Let us mention, as an easy example of isotropic network, the square grid pattern drilling. The errors made for an isotropic network by extending to each influence zone the grade of its central sample may be considered as independent (in other words having a geostatistical covariance equal to zero). In this case *estimation variance is obtained by dividing the extension variance σ_E^2 of each sample within its influence zone by the number N of these influence zones.*

$$\sigma^2 = \frac{1}{n}\sigma_E^2. \quad (12)$$

[5] More generally, for any given function $\gamma(h)$, the lattice may or not be adjusted to the anisotropy of function $\gamma(h)$. For questions concerning the different types of anisotropy, one should refer to the *Treatise of Applied Geostatistics*.

If, on the contrary the network is not isotropic, we are led to rearrange the samples along lines or planes of maximum density, and to compose extension variances of different natures. For example let us suppose a vein-type deposit developed by drifts and channel sampled. In the first place, we have to consider the extension variance $\sigma_{E_1}{}^2$ of a channel within the length of a drift from which it has been cut off. If N is the total number of channels, one can see that the estimation variance $(1/N)\sigma_{E_1}{}^2$ represents the error obtained by extending the grade deduced from channel samples over the mining works themselves. We consider afterwards the extension variance $\sigma_{E_2}{}^2$ of the grade (supposed to be perfectly well known) of a drift inside its influence zone. The influence zone is here the panel composed by joining both the half-levels located above and below the drift. If n is the number of developed levels, the estimation variance $(1/n)\sigma_{E_2}{}^2$ represents the error obtained by extending the average grade supposed to be perfectly well known of the mining works to the whole deposit. The resulting estimation variance becomes:

$$\sigma^2 = \frac{1}{N}\sigma_{E_1}{}^2 + \frac{1}{n}\sigma_{E_2}{}^2. \tag{13}$$

It is usually necessary to add an additional variance to this expression, representing the sampling and analyses errors. The second term in such an expression is usually broadly predominating. The greater part of the error proceeds from the extension of data from the mining works to the deposit. In particular, it would be no use to increase indefinitely the number N of samples without carrying out supplementary mining works. In fact, the estimation variance coincides very soon with the $(1/n)\sigma_{E_2}{}^2$ limit below which it cannot decrease.

Tables and graphs giving the numerical values of elementary extension variances have been established [6] for a given number of intrinsic functions (especially for type (7) of de Wijs's function). They allow a fast computation of estimation variances assigned to different drilling and underground exploration schemes.

We offer for example a vein deposit conformable to a type (7) isotropic de Wijs's scheme and developed by drifts. Let us also assume that drifts have been sufficiently well sampled as to reduce the first term of equation (13) to zero.

Let h be the raise between two consecutive levels (measured inside the plane of the vein). The extension variance of a drift of length l within an influence panel lh is proved to be:

$$\sigma_E{}^2 = \alpha \frac{\pi}{2} \frac{h}{l}.$$

This formula is valid only if h is small compared to l, but it may be used until $h = l$. When $h > l$, it must be replaced by a different formula. Let

[6] *Treatise of Applied Geostatistics*, Vol. I, for the de Wijs's functions. Vol. III for the case of a nugget effect.

us assume that lengths $l_1, l_2 \cdots l_n$ are all superior to h. The estimation variance is obtained by weighting the extension variance of each drift within its influence panel by the square of the surface of this panel:

$$\sigma^2 = \frac{l_1^2\sigma_{L_1}^2 + l_2^2\sigma_{E_2}^2 + \cdots}{(l_1 + l_2 + \cdots)} = \alpha \frac{\pi}{2} h \frac{l_1 + l_2 + \cdots}{(l_1 + l_2 + \cdots)^2}.$$

The explored mineralized surface being $S = h(l_1 + l_2 \cdots + l_n)$ and the total developed length being $L = l_1 + l_2 + \cdots l_n$, we obtain the following remarkable formula:

$$\sigma^2 = \alpha \frac{\pi}{2} \frac{S}{L^2}.$$

Once the estimation variance has been calculated, one has still to interpret it for practical uses under the form of *conventional error spread*. This aim is reached by allocating to this variance a probabilistic meaning. By implicit reference to a gaussian model, we shall take it that the actual average grade of the deposit is included within a 95% probability in the range $m \pm 2\sigma$, m being the estimated grade. In other cases, particularly if 2σ is not small towards m, we shall take the spread $m \exp(\pm 2\sigma/m)$, by reference to a lognormal model.

These implicit references to probabilistic models are mainly arbitrary. Actually, the notion itself of statistical distribution of an estimation error is doubtlessly meaningless. The only thing which has an objective physical meaning is the variance. This is why we speak about conventional spreads. Their practical interest resides in the fact that they draw a more intuitive picture of the possible errors than variances themselves.

KRIGING

A second application of major importance is provided by a geostatistical procedure called "kriging." It consists in estimating the grade of a panel by computing the weighted average of available samples, some being located inside others outside the panel. The grads of these samples being x_1, $x_2, \cdots x_n$, we attempt to evaluate the unknown grade z of the panel with a linear estimator z^* of the form:

$$z^* = \sum a_i x_i.$$

The suitable weights a_i assigned to each sample are determined by two conditions. The first one expresses that z^* and z must have the same average value within the whole large field V and is written as:

$$\sum a_i = 1.$$

The second condition expresses that the a_i have such values that estimation variance of z by z^*, in other words the kriging variance, should take the smallest possible value.

This is formulated with a linear equation system related to a_i, the coefficients of which are expressed with the help of the variances and covariances of the samples and of the panel. It is thus possible to tabulate, for each intrinsic function, the coefficients and the kriging variance in terms of geometrical parameters, appropriately for different configurations. Numerous drilling and underground work configurations have thus been tabulated in

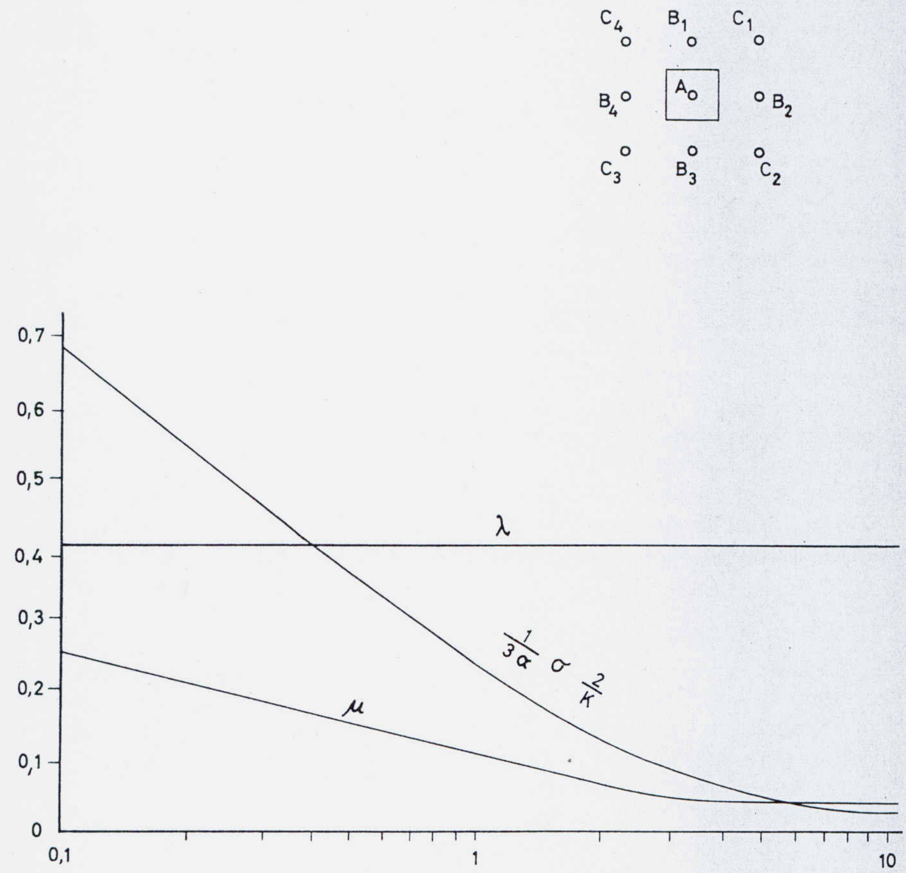

FIG. 3.

the case of an isotropic scheme of de Wijs. For information we show an example in Figure 3. The studied configuration is useful for the appraisal of a deposit explored by drilling, or for open cast selective mining. It consists, in the case of a square grid pattern drilling, in the kriging of the influence blocks of a drilling A with help of the grade of this central drilling A, and those of the 8 nearest drillings rearranged into two "aureolae" $B_1B_2B_3B_4$ and $C_1C_2C_3C_4$. Let u be the grade of A, v and w the average grades of drillings

B and C, the estimator to be used is:

$$z^* = (1 - \lambda - \mu)u + \lambda v + \mu w.$$

In Figure 3, is plotted in abscissa the ratio h/a between the width h of the formation and the size a of the mesh of the drilling grid and the numerical values of λ and μ are read on the curves as well as these of the expression $(1/3\alpha)\sigma_K^2$. The multiplication of this last expression by three times the value of absolute dispersion, 3α, yields the kriging variance.

Theoretically it is advantageous to "krige" each panel by all the samples located in the deposit, inside and outside this panel. In addition to the great complexity of computation which grows very fast, it appears in numerical examples that it is usually unnecessary to take into account remote samples. In general the one or two proximate aureolae of external samples are enough to remove practically the whole effect of remaining external samples. This is, in particular, the case of the configuration studied in Figure 3 where both aureole B and C form an almost perfect screen towards all other external drillings.

One can even notice that for high values of the h/a ratio, the weight μ of the second aureole becomes slight, so that the aureole made out of the four B drillings constitutes a screen by itself alone. This *screen effect* is a general phenomenon and plays an important part in the kriging theory.

From a practical point of view, the advantage of kriging is double. First of all, as a result of the definition itself of this procedure, it leads to achieve the best possible estimation for a given panel, that is to say the estimation with minimal variance. It can pay most appreciable services by improving, for example, the monthly output forecast for different mine-sections, and especially in the case where the mine operator is compelled to supply ores with characteristics as constant as possible.

However appreciable they are, the improvements of accuracy provided by the kriging would not always justify the amount of calculations it requires. In most cases, the major interest of the procedure does not come for the reduction of estimation variances but from its being able to eliminate the cause of systematical error. A deposit seldom happens indeed to be payable in the whole. Only some panels chosen as payable according to the grades of the samples cut off within them, are considered as payable. D. G. Krige [7] has proved that the results based only on inside samples inevitably led to over-estimating rich panels and underestimating poor ones. The geostatistical notion of kriging allows to expound this phenomenon easily and to rectify its effects. The selected panel being a rich one, the aureola of outside samples has, in general, a lower grade than that of inside samples. Yet its influence on the panel to be estimated, is not negligible, since it is allocated a weight different from zero by the kriging. Not to take in account this external aureole inevitably introduce therefore a *cause of systematical error by over-estimation* which can be eliminated by kriging.

[7] D. G. Krige's original reasoning constitutes a second example of an implicit passage from classical statistics to geostatistics. It is essentially based on the fact that the variance of a panel is always lower than that of its inside sampling. For references, see "Treatise of Applied Statistics."

THE NUGGET EFFECT

In the presence of a strong nugget effect the general rules outlined in the above paragraphs may suffer some apparent objections. The nugget effect has been defined in Figure 1 by a variogram characterized by a discontinuity at the origin, and corresponding to a regionalized variable that does not have the "in average" continuity. Its nature may be purely granulometrical, as in gold or diamond deposits, or, more generally, it may reveal the existence of discontinuous micro-structures. The presence of veinlets or microfractures with high-grade fillings in a stockwerk may promote such an effect. In gold deposits, the grades of two very close or even adjoining samples may be different if, by chance, one of them contains a large nugget. The smaller the samples, the more important this effect is, and it may reach a considerable magnitude for samples of several liters in volume. A translation of some millimeters only of the geometrical support of a sample is enough for it to contain or not a large nugget able to modify its grade in a proportion of 1 to 10 or 1 to 100. The possibility for a marginal nugget to be embodied, or not, inside a sample appears as an entirely random event. Actually, however, the behavior of the grade can be considered as random locally only. If it were not so, the panels of several thousand tons, on which marginal nuggets have no more detectable effect, would present almost constant grades (their variance being then a million times lower than that of samples of several kilograms). It is well known that actually, even in gold deposits, there are rich panels and poor panels. But this random effect may locally be so strong that it entirely hides the underlaying regionalization. The frequency of some expressions such as "erratic," "monstrous," or "mammoth grades" etc. . . . alluding to an hypothetical anomalous behavior of mineralization in the literature devoted to these deposits is striking. Certainly the classical statisticians were right when they noted that there was no actual anomaly, and that those monster grades, actually existing in the deposit, appeared from time to time in the sampling, with frequency determined by random laws. Historically, a clear distinction between the notions of regionalized and aleatory variables was doubtlessly hampered for a long while by the fascination aroused by this nugget effect. It appears, from the geostatistical point of view, that in fact, the ingenious terminology was not wrong while suggesting the existence of some anomaly; but the aberrant fact is not the presence of some "anomalously" high grades, but rather the locally aleatory behavior of all the grades, high or low, as well as in the deterioration of the spatial correlations grid. Those mammoth grades of the ingenious terminology are not aberrant by themselves, but the fact that they are not assorted with influence zones is so. And, on the other hand, classical statisticians were right stressing the fact that the apparitions of these aberrant grades are ruled by random laws. But they failed to note that the phenomenon can be considered as aleatory locally only.

Without trying to make a systematical statement,[8] let us show briefly how geostatistics allow us to represent a nugget effect. Let us examine the

[8] See *Treatise of Applied Geostatistics*, Vol. III.

$\gamma(r)$ semi-variogram representing the third type of Figure 1. We shall stick here to the case where $\gamma(r)$ is an isotropic function (in other words depending only upon the r modulus of the h vectorial argument). The C discontinuity, or jump, noticed at the origin on the $\gamma(r)$ of a variable with punctual support, is called *nugget constant.* $H(r)$ being Heaviside's function, thus defined:

$$\begin{cases} H(r) = 1 & r > 0 \\ H(r) = 0 & r = 0 \end{cases}$$

the semi-variogram may be divided into 2 components:

$$\gamma(r) = CH(r) + \gamma_1(r). \tag{15}$$

The first component $CH(r)$ represents the pure nugget effect. The second one $\gamma_1(r)$, continuous at the origin, represents the underlaying regionalization. All the variances and the covariances that have to be introduced, may then be calculated as if the variable $x(M)$ with punctual support was the sum:

$$x = x_0 + \epsilon \tag{16}$$

of a theoretical regionalized variable x_0 following the $\gamma(r)$ dispersion law continuous at the origin, and of an aleatory ϵ variable with a zero average and C variance.

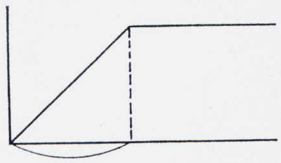

FIG. 4.

The x_0 and the ϵ are independent, and the ϵ assigned to two distinct points even very close, are independent as well. If we limit our study to the variation of the punctual grade x in the proximity of a given point, or, in other words, we consider only the small values of the distance r, $\gamma_1(r)$ will vary so slightly that it might be taken for a constant equal to C. The locally detectable variations are to be assigned almost solely to ϵ. That is what we mean when we say that the regionalized variable behaves locally as an aleatory variable. But, on a larger scale, i.e., for higher values of r, the increase of the continuous component $\gamma_1(r)$ can no longer be neglected and the regionalization of x_0 becomes perceptibly apparent.

As a matter of fact, the Heaviside function does not represent with entire satisfaction the random aspect of the behavior of a punctual variable. Unless we suppose the constant C to be infinite, the term $CH(r)$ will lose all influence over the variance of a sample of a size different from zero. It is automatically eliminated in formula (4). It means that the mean value of the ϵ independent aleatory variables, located in infinite number inside an unpunctual support, has compulsorily a zero variance.

The notion of a random variable ϵ with a punctual support has actually no physical meaning. The actual physical phenomenon will never involve a true discontinuity at the origin but a narrow transition zone in the proximity of $r = 0$. $H(r)$ must be replaced by the transition function $T(r, a)$ defined by:

$$\begin{cases} T(r, a) = \dfrac{r}{a} & \text{if} \quad r \leq a \\ T(r, a) = 1 & \text{if} \quad r > a. \end{cases}$$

The a constant, or *range*, gives the scale of the transition zone, that is to say the size of the nuggets. In the case of homogranular nuggets of same volume u it is shown that:

$$u = \frac{\pi}{3} a^3.$$

The intrinsic function $\gamma(r)$ of a punctual grade is decomposed in the following way:

$$\gamma(r) = CT(r, a) + \gamma_1(r).$$

C is still the nugget constant, and $\gamma_1(r)$ the continuous component.

The punctual grade x can be given by a sum similar to (16) in which ϵ is a regionalized variable admitting $CT(r, a)$ as its intrinsic function. Now the ϵ are only independent for distances superior to the range a. For smaller distances they are bound by a linear variogram. The nugget effect will therefore reflect itself on samples of size v different from zero. If v is large in regard to the grain size a^3 the transition zone will be diluted in the integration volume v, and the nugget effect will yield an additional variance of the type a^3/v. Indeed, let σ_P^2 (nugget variance) be the share of $CT(r, a)$ for the variance of sample v. According to (4) we have to compute integrals of the type:

$$\frac{C}{v^2} \iiint_v dv_1 \iiint_v T(r, a) dv_2.$$

If all sizes of v are supposed to be large in regard to a, each point inside v brings to the sextuple integral the following part:

$$C(v - \tfrac{4}{3}\pi a^3) + \frac{C}{a} \int_0^a 4\pi r^3 dr = C\left(v - \frac{\pi}{3} a^3\right).$$

This is valid only for points located at a distance superior to (a) from the boundary of v; but, when v is large, the boundary points only interfere with superior order terms. With such an approximation, the sextuple integral is equal to $C(1 - (\pi/3)(a^3/v))$.

As the integral inside V is computed in the same way, we finally have

$$\sigma_P^2 = C\frac{\pi}{3}\left[\frac{a^3}{v} - \frac{a^3}{V}\right]. \tag{17}$$

Practically a^3/V is negligible and the nugget variance is in terms of a^3/v, i.e., in an inverse ratio to the number of grains contained inside the sample.

Any time a nugget effect does exist, i.e., anytime a regionalized variable shows a locally aleatory behavior, an additional variance is assigned to macroscopic samples, called nugget variance, inversely proportional to their size.

The variance of those samples appears as the sum:

$$\sigma^2 = \sigma_P{}^2 + \sigma_\theta{}^2$$

of the nugget variance and of the theoretical variance $\sigma_\theta{}^2$ calculated with the continuous component $\gamma_1(r)$ of the intrinsic function.

When v is increasing, the theoretical variance is decreasing much slower than the nugget variance. In the presence of a very strong nugget effect $\sigma_p{}^2$ may happen to be widely predominating for samples of several kilograms. The underlying regionalization is almost completely hidden at the scale of these samples. If we limit the variation of the volume v in the interval of a few liters up to tens of liters, the experimentally observable variations of the variance will be those of the nugget variance effect only, and we may take the risk to conclude that the variance varies in inverse ratio of the volume.

Whereas if we consider samples of several tens of tons, the term $\sigma_p{}^2$ decreases and disappears, and the theoretical variance $\sigma_\theta{}^2$ becomes prominent. The effect of the underlying regionalization appears again and the variance is steadily decreasing as v is increasing, but much slower than $1/v$.

We have somewhat insisted upon the nugget effect in order to show, through a crucial example, how geostatistical concepts allow us to rediscover the local results that are fluently obtained from common statistical reasoning (nugget variance inversely proportional to volume) but inserting them in the general prospect of an underlying regionalization. As for the practical use of this theory, let us succcinctly mention the two following points:

In the presence of a nugget effect, the extension and the estimation variances are both increased by a term $C(\pi/3)(a^3/v)$ inversely proportional to the total volume of available samples and, therefore, in particular to the number n of those samples. In this regard, the additional estimation variance due to the nugget effect behaves itself as the sampling and analyses variances, and may be rearranged with them.

As for the kriging, the nugget effect results in partly removing all the screens. Practically, we are led to use the special forms of kriging called "aleatory kriging" which are not different from those proposed formerly by D. G. Krige himself, in connection with the gold deposit of the Rand, in which the nugget effect is probably very strong.

SEARCH FOR OPTIMUM IN MINING EXPLORATION

Geostatistics are able, through estimation variances, to provide an accurate measurement of the information yielded, by a given amount of underground workings on a deposit. Generally, these workings are expensive, and their cost must be weighed against the economic value of the provided information. Thus appears the possibility to determine the optimum amount of credits to be allocated for the exploration of a deposit, and particularly the possibility to choose the suitable moment for stopping the exploration, as well as for taking a positive or negative decision towards starting the exploitation of the deposit. These methods, permit one to solve, at least partly, one of the main problems raised by mining exploration, will be published in another connection, and cannot be treated here. Let us only, as a conclusion, stress the fact that they appear as the natural extension of geostatistics. The possibility of their adjustment was bound to the preliminary elucidation and to the thorough scientific study of the different ideas which have been summarized in this paper.

Bureau de Recherches Geologique et Minieres,
Paris, France,
June 10, 1963