COURSE ON NONLINEAR GEOSTATISTICS

by

G. Matheron, A. Journel, A. Maréchal, J. Deraisme

———

Summer School

1978

Cours C-74

———

TECHNICAL PARAMETRIZATION OF ORE RESERVES

G. MATHERON

(September 1978)

CONTENTS

0 - INTRODUCTION

Technical parametrization of recoverable ore reserves is intended as a tool for choosing among various technically feasible projects, that which has the greatest chance to lead to the best results under given economical conditions.

Each project being considered as a function of one or two purely technical parameters (e.g. cut-off grades, or tonnage of ore to be processed, or tonnage of waste to be removed), the technical parametrization consists of selecting a family of such projects which are optimal in a purely technical sense(i.e. independently of any economic consideration). This family will have to possess the following property : whatever the economic conditions of the moment (price of metals, costs of extraction and processing etc..), and the formula to be used for the economic value, the economically optimal project (whatever it may be) must necessarily fall within this family of technically optimal projects.

This separation (or preliminary technical parametrization) is not always possible to make rigourously, but it is often possible as a first approximation; In what follows, we will place ourselves in the simple case in which the economic value of a project is a function of the form

$$B(Q,T,V)$$

- Q : Quantity of metal (or of useful material)
- T : Tonnage of ores to be processed
- V : Total tonnage to be extracted (ore + waste)

In order to parametrize the ore reserves, we do not need to explicitly know the function B (which can thus, at this stage, remain indeterminate). We only assume the following (reasonable) hypothesis that B is an increasing function of Q and a decreasing function of T and V.

At that point, technical parametrization becomes possible. Among all of the possible projects involving the same tonnage T of ore

to be processed and the same volume V to be mined, the best will always be the one which maximizes the quantity of recovered metal Q.

As a function of T and V, we will thus define the project which maximizes Q at fixed T and V, and we will set down the corresponding function :

$$Q = Q(T,V)$$

which will produce the parametrization which we are looking for. To choose the economically optimal project, we have only to determine (once the payable function B is known) the values of T and V which produce the maximum of $B[Q(T,V), T, V]$.

In practice, we will often replace T and V by two other parameters z and λ (representing the cut-off grades), and we will set up a parametric representation of the family of technical optima of the form :

$$Q = Q(z,\lambda)$$
$$T = T(z,\lambda)$$
$$V = V(z,\lambda)$$

The economic optimization will then (once B is known) consist of choosing the values of z and λ which maximize the expression :

$$B[Q(z,\lambda), T(z,\lambda), V(z,\lambda)]$$

I - INTRODUCTIVE INTUITION : ZONEOGRAPHY AND HISTOGRAM.

Lasky's Tonnage/Grade curve.

Around 1950, S. Lasky wondered how the recoverable reserves of porphyry copper varied as a function of the usual selection criterion (the cut-off grade in Cu %). He came up with the capital notion of grade/tonnage curve.

Lasky started from a simple representation of his porphyry copper deposit : the grade $Z(x)$ at point x was (implicitly)

considered as a continuous and regular function which, for example, decreased regularly as one moved away from the rich heart of the porphyry (today, we would attribute this behaviour to a _trend_, rather than to the grade itself). He assumed this function to be known (perfectly, or to a very close approximation), and his point of view was that of a _zoneography_ :
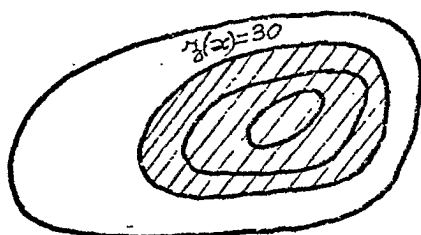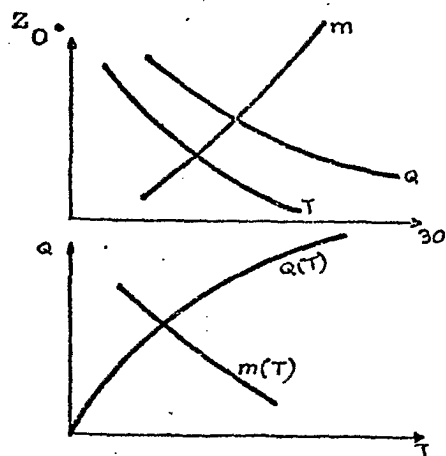

If the cut-off grade of Cu is z_o, we will mine the domain whose limits are the iso-grade z_o, i.e.

$$B_{z_o} = \{x : z(x) \geq z_o\}$$

and with a tonnage and a grade of

$$(1\text{-}1) \qquad T(z_o) = \int_{B_{z_o}} \lambda(dx) \qquad Q(z_o) = \int_{B_{z_o}} z(x)\, \lambda(dx)$$

with : $\lambda(dx) = dx$, or, more precisely; $\rho(x)\, dx$, where ρ is the density. Lasky obtained T, Q, and $m = Q/T$ curves parametrized by z_o.

Q and T are evidently decreasing (and m increasing) with z_o.

We can thus eliminate z_o and set

$$Q = Q(T) \ , \ m = m(T), \ etc...$$

[Ex : the so-called Lasky law of the form $m = \alpha - \beta \log T$].

: $Q(T)$ is increasing and _concave_.

Indeed, if $T(dz)$ is the tonnage which lies between the iso-grades z and $z + dz$:

$$
(1\text{-}2) \qquad
\begin{aligned}
T(z_o) &= \int_{z_o}^{\infty} T(dz) \\[2mm]
Q(z_o) &= \int_{z_o}^{\infty} z\, T(dz)
\end{aligned}
$$

As a consequence, $\frac{dQ}{dT} = z_o > 0$; but $T(z_o)$ is a decreasing function of z_o, hance z_o is also a decreasing function of T : the slope of the tangeant decreases as T increases, and Q is concave.

Notice the relation :

$$Q(z_o) = z_o \, T(z_o) + \int_{z_o}^{\infty} T(z) \, dz$$

Hence :

$$m(T) = Q/T = z_o + \frac{1}{T} \int_{z_o}^{\infty} T(z) \, dz$$

The relation $Q = mT$ than gives $dQ = mdT + Tdm$, hence $z_o \, dT = mdT + Tdm$. From which we deduce that :

$$\frac{dm}{dT} = - \frac{m-z_o}{T} \leq 0$$

Thus, m decreases when T increases.

We can also give another interpretation, in terms of histograms, or of the distribution $F(dz)$ of the grades (point grades $Z(x)$). Indeed, if T_o is the maximum tonnage of ore (the limit of $T(z_o)$ as $z_o \to 0$), we may set

$$F(dz) = \frac{-1}{T_o} \, T(dz)$$

from which we obtain :

$$T(z_o) = T_o \, (1 - F(z_o)) = T_o \int_{z_o}^{\infty} F(dz)$$

$$Q(z_o) = T_o \int_{z_o}^{\infty} z \, F(dz) = z_o \, T(z_o) + T_o \int_{z_o}^{\infty} [1 - F(z)] dz$$

In this very simple case, the above formulae give the technical parametrization of recoverable reserves, as a function of a single parameter : the cut-off grade z_o.

In fact - if we suppose that the point grades are actually

known and that it is possible to perform such a precise selection — the best possible choice of the volume to mine, in order to obtain a given tonnage T, would be that which maximizes Q, that is, the set $\{Z(x) \geq z_0\}$ with a z_0 such that $T(z_0) = T$.

REMARK : From the point of view of geostatistics, Lasky's simplistic model is misleading. Through the (non-realistic) hypotheses made on the regionalized variable (RV) $Z(x)$: a well-known and regular function — three notions — between which one should discriminate very carefully — here become confused :

- ~ Zoneography
- ~ Histogram of grades
- ~ The tonnage/grade curve

With usual RV's, this method would lead to serious overestimations. Indeed, one <u>never</u> has the right to use the histogram of point (or core) grades ; at most, one can get away with using the histogram of grades of blocks, i.e. of the smallest units on which the effective selection will be made (their size depends upon the planned mode of operation). Then arises a first problem (which we know how to solve) :

- how to predict the histogram of blocks knowing that of the cores.

But that is not all. For the identification of the tonnage/grade curve with the distribution of blocks to be legitimate, we would need :
- the real grades to be perfectly known, at very least when the operation starts.
- that no geometric constraints should influence the selection (in other words, that we would be free to select each of the blocks independently of the other blocks).

The first condition is never achieved except (approximately) for Uranium, for which one can measure the radioactivity at the time of mining. In all other cases, the grade of the blocks will <u>never be known</u> exactly at the mining stage, but only estimated : based upon information perhaps richer than that which we have

today, but with a margin of error nonetheless.

As for the second condition, it will rarely (or never) be possible, in practice, to extract a rich block from just any place while leaving its poorer neighbouring blocks in place. In the case of an open pit, such geometric constraints become the most important aspect of the problem.

2 - CRITICAL FACTORS OF THE TECHNICAL PARAMETRIZATION.

These factors are suggested by the preceding critique. Let us point out right away that there will be as many technical parametrizations, hence as many possible definitions of the recoverable reserves, as there are possible modes of mining. For a given mining technique, the following factors must be taken into account :

a) Geometrical constraints.

[Principle : not everything is possible].

The technological characteristics of a mining technique often impose severe constraints upon the geometrical form of the volumes to be mined. This can be expressed with the notion of admissible contours. Example :

- Open pit : The excavation must respect the natural equilibrium slope. If ore is extracted at a point x, all material within the cone T_x, which has its tip at point x (upside down) must also be extracted, and the admissible contours are necessarily a union of such cones.

- In subhorizontal bedded deposits, selection of the footwall and hanging wall limits. We decide to take only one layer of thickness greater than p. meters.

- The selection itself may be simple or it may operate at several levels. For example, we often have to perform a double selection which comprises :

  ~ an extraction selection concerning the choice of the volume V to be extracted (which must be an admissible contour).

~ an <u>ore selection</u> (each of the blocks v contained within the volume V to be extracted can be sent either to the waste or to the mill.

The geometric constraints influence each of these selections in different ways. For example, the selection of ores, at the level of the elementary blocks v, will most often be independent of any constraints (each small block, once it has been extracted, can be sent to the waste or to the mill, independently of the other blocks). On the contrary, the extraction selection will generally fall under more or less severe constraints ("the volume V to be extracted must constitute an admissible contour"). Sometimes (for example, when considering underground mining or in the case of a shallow stratified formation) we may consider (as a first approximation) that the volume V to be extracted is constituted by the union of <u>panels</u> V_j of a given size which we can choose independently of each other (this terminology implies that the blocks v, at whose level the ore is selected, are smaller than the <u>panels</u> V_j to be extracted) : in this case, we can talk about a free selection on two levels (level v for the ore selection, level V_j for the extraction selection).

b) <u>Support</u>.

[Principle : the variables upon which the ore will be mined out are not the same as those which we used in the exploration. <u>We will mine blocks or panels, but never cores</u>].

For each of these selections, we must defined its support, which is the size of the smallest unit which might be selected : <u>blocks</u> v for the ore selection ; <u>panels</u> V_j for the nextraction selection.

c) <u>Information</u>.

The information upon which the selection decisions are (or will be) effectively made at the various levels (ore and extraction)

[Principle : <u>the estimation is never identical to reality</u>].

We must carefully distinguish between :

- The present information (available at the moment of the parametrization of reserves : for example, exploration DDH).

- Future or "ultimate" information, which will be available when the final decision about selection is made, but generally is not available today. For example, for the selection of ores, we will have the results of the blast holes when we will have to decide whather a block v should be sent to the mill or to the waste.

This gap between present information and future information is the source of one of the principal difficulties in parametrization : since the actual future selection is based upon a not-yet-available future information, we cannot predict its effect without anticipating, probabilistically, that future information on the basis of our present information.

This would not be very serious if there are no constraints (if the panels can be selected independently), since a global probabilistic reconstitution would be possible. If there are strong constraints (open pits, for example) the prediction necessarily shifts to the local level. It then requires more powerful tools (transfer functions) and the use of the final information.

d) Criteria.

The criteria used for each of the selections. In all generality a selection S_i (for instance S_1 = extraction selection, S_2 = ore selection) requires a criterion of the form f_i (function of information available at the moment of making the effective decision upon S_i) $\geq z_i$.

These z_i constitute the parameters of the technical parametrization of reserves. Hence, in principle, there will be as many parameters as there are levels of selection. The criteria functions f_i (which we chose ahead of time) will have to respect, in principle, the following condition :

- Whatever the values of the economic parameters in the "pay off"

formula itself (as long as it remains plausible), the economic optimum must belong to the family of projects defined by the criteria f_i .

3 - FREE AND SIMPLE SELECTION.

Here, the deposit is cut up into blocks (or panels) v_i and we can select each of these v_i (i = 1,...,N) independently. Among all of the possible choices of k blocks, the best would be that which maximizes the total recovered metal $v \sum Z(v_i)$ (this is true, regardless of the form of the pay-off formula). But we do not know the actual grades $Z(v_i)$ and we will have only their estimators $Z*(v_i)$.

We will assume, essentially, that these estimators satisfy the conditional non-bias condition.

$$(3-1) \qquad E[Z(v_i)/Z*(v_i)] = Z*(v_i)$$

At the present time, the estimators $Z*(v_i)$ are not yet known, if we have available (for example) only the results of the widely-spaced exploration DDH (which gave Z_α, α = 1,...,n as results).

For a given panel v_i, we should calculate

$$P(Z*(v_i) \geq Z/Z_\alpha)$$

which is the probability (conditional given Z_α) thatbthe panel will be selected in the end, and estimate $T(Z)$ by

$$T(Z) = v \sum_i P(Z*(v_i) \geq Z/Z_\alpha)$$

From this point of view, we estimate the global $T(Z)$ by summing the local estimations of each panel v_i. This is a rather heavy procedure which can be avoided here. It so happens that (if the panels are numerous enough) the summation of these various conditional probabilities approximately reproduces (in the stationary and ergodic cases) the non-conditional probability $P(Z*(V) \geq Z)$ for a

panel of size v, i.e. :

$$T(Z) = N v P(Z*(v) \geq Z)$$

We need only to know the "a priori" distribution law of the future estimator, that is : $F*(dZ)$.

From (3-1), $E(Z(v)/Z*(V)) = Z*(v)$, from which we deduce that

$$E[Z(v)/Z*(v) \geq Z] = E[Z*(v)/Z*(v) \geq Z]$$

This produces our parametrization (with $T_0 = N_V$) :

(3-2)
$$\begin{cases} T(Z_0) = T_0 \int_{Z_0}^{\infty} F*(dZ) \\ Q(Z_0) = T_0 \int_{Z_0}^{\infty} Z F*(dZ) \end{cases}$$

The parametrization is similar to Lasky's (1-2) but $F*$ is now the distribution of the future estimator $Z*(v)$, and not that of the block grades $Z(v)$. What's more, the second relation of (3-2) (but not the first) presupposes the condition of conditional non-bias (3-1).

In practice, we do not know the distribution $F*$ of the $Z*(v)$ (nor that of the $Z(v)$), except if the future information is already available (in this case, we use the experimental histogram of the $Z*(v_i)$. But, with the help of certain "invariance" hypotheses, we can obtain an approximate expression for $F*$ from the histogram of the Z_α (present information).

4 - FREE SELECTION AT TWO LEVELS.

Here, we assume that a first selection S_1 (extraction), based upon the information I_1 (not necessarily available at the time of the estimation) chooses the panels V_j which are to be extracted. Next, a second selection S_2 (of ores) chooses, within each panel,

those blocks v_{ji} which are to be processed, based upon "ultimate" information I_2, which is richer than I_1, but is presently unknown. Let us assume, at the outset, that I_1 (but not I_2) is available right away.

Under these conditions, for a <u>given</u> choice of panels V_j $(j \in J)$ assumed to be selected, the reasoning followed in paragraph 3 can be applied : the economic optimum will belong to the family of projects $\{Z*(v_{j,i}) \geq z\}$ defined by the application of a single cut-off grade z to the (future) estimators $Z* (v_{ji})$ of the various blocks contained within the preselected panels.

Thus, we should perform the same parametrization as in (3-2) but panel by panel, and consequently using the distribution law $F_j(dZ/I_1)$ of the grade $Z*(v_j)$ of a block v_j, which sweeps through the <u>panel</u> V_j : this law is conditional, given the actual information I_1. For each panel V_j, we thus obtain a (local) estimation of recoverable reserves :

$$(4-1) \qquad \begin{cases} T_j(z_0) = V_j \displaystyle\int_{z_0}^{\infty} F_j^* \, (dZ/I_1) \\[2em] Q_j(z_0) = V_j \displaystyle\int_{z_0}^{\infty} F_j^* \, (dz/I_1) \end{cases}$$

The techniques of <u>Disjunctive Kriging</u> are very helpful for rapidly estimating these <u>transfer functions</u> (or local parametrization of reserves), panel by panel.

We now have <u>to optimize the choise J of panels</u> V_j $(j \in J)$ to be extracted. To facilitate the reasoning (although the result which will be obtained will be more generally applicable), let us assume, temporarily, that the "pay off" formula is linear, that is, to within a constant factor :

$$B(Q,T,V) = Q - zT - \lambda V$$

(in this formula, z and λ represent the costs of processing and

extraction. But we attach little importance to this economic interpretation and simply consider z and λ as two parameters which can vary independently).

For a given choice J of panels V_j to be extracted, we have

$$E(B/J) = \sum_{j \in J} [Q_j(z) - z\ T_j(z) - \lambda\ V\ j]$$

With z and λ fixed, a panel j will be selected for extraction if and only if it brings a positive contribution to the sum above. The criterion for the selection of panels is thus :

$$(4\text{-}2) \qquad \frac{Q_j(z) - z\ T_j(z)}{V_j} \geq \lambda$$

If we now consider z and λ as simple variable parameters, we wi will obtain the following parametrization : for given z and λ, we will retain those panels which verify criterion (4-2), and this defines the set $J = J(z,\lambda)$ of selected panels. From this set $J(z,\lambda)$ of panels, we then select the blocks of grade $\geq$ z (ore selection). In this way, we obtain the parametrization

$$(4\text{-}3) \qquad \begin{cases} Q(z,\lambda) = \sum_{j \in J(z,\lambda)} Q_j(z) \\[2mm] T(z,\lambda) = \sum_{j \in J(z,\lambda)} m_j(z) \\[2mm] V(z,\lambda) = \sum_{j \in J(z,\lambda)} V_j \end{cases}$$

REMARKS :

1) Even though, for simplicity's sake, we assumed that the "pay-off" formula is linear, the parametrization (4-3) has a general value. More precisely, the techniques of convex analysis, which we will study, show that the projects defined by (4-3) all correspond to technical optima (for given z and λ, each of these maximizes the quantity of metal Q for fixed $T = T(z,\lambda)$ and $V = V(z,\lambda)$. However, some existing technical optima may escape this parametrization, but : those optima which elude us will always, in practice, be closely bracketed by projects defined by (4-3), and, for the

.usual forms of "pay-off" formula, they would have virtually no chance of corresponding with the economic optimum.

2) We have assumed that the information I_1, for the first selection, is available. If it is not yet available, if we have only the poorer information $I_0 \subset I_1$ the $Q_j(z)$ and $T_j(z)$ must be considered as random variables, and formulae (4-3) must be probabilistically anticipated ; such a procedure is possible, but we will not go into it here.

5 - SIMPLE SELECTION UNDER CONSTRAINT.

The constraint is expressed by stating that any project must correspond to a contour $A \in \beta$ (belonging to the family β of admissible contours). Examples :.

- Open Pit : $A \in \beta$ if A follows the equilibrium slope conditions, hence if A is the union of cones T_x, $X \in A$.

- Footwall and Hanging wall : Here, the deposit is divided into panels P_j. In each panel, the mined portion must be either empty or a horizontal slice of thickness $\geq$ a given p (for example). To simplify things, we assume that the horizontal slices may be independently chosen from the various panels P_j (hence, we neglect the link-up conditions between panels).

Among the possible projects A of volume $V(A) \leq$ a given v, we suppose that the best is that project which maximizes the quantity of metal $Q(A)$. For each v, there are thus one (or several) A_v which verify :

(5-1) $\qquad Sup[Q(A) : A \in \beta , V(A) \leq v] = Q(v)$

If the grade $q(x)$ is known at each point x, then $Q(A) = \int_A q(x)dx$ is known for each A, and the above formula will achieve (theoretically) the technical parametrization of v. When $q(x)$ is not known, nothing essential is changed if the information I, upon which is

based the choice of contour A, is available right away — which is what we assume here. Indeed, we formulate, for each x, the estimator $q^*(x) = E(q(x)/I)$ which satisfies the condition of conditional non-bias. We are back to maximizing $\int_A q^*(x)dx = Q^*(A)$ (since $E(Q(A)/I) = Q^*(A)$ for each $A \in \beta$). We shall therefore drop the asterisk and write $Q(A)$ instead of $Q^*(A)$.

Principle of Convex Analysis

In practice, the problem stated in (5-1) is rarely easy to solve. A tedious combinatorial analysis is often required. Convex analysis techniques consist then in replacing the inaccessible curve $Q = Q(v)$, which is increasing but generally not concave, by its concave hull $Q = \hat{Q}(v)$, the latter curve being in general more easy to get : doing so, we may miss a few technical optima (but they are the most interesting ones in an economical sense, and in any case they are never too far from (known) points of the concave curve $\hat{Q}(v)$
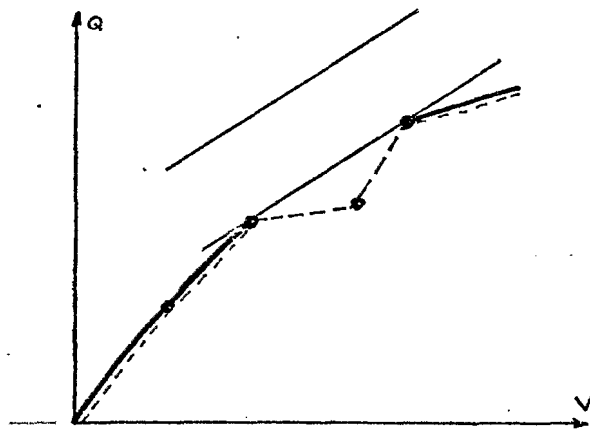
For each $\lambda \geq 0$, let us consider with (Q,v)-coordinates the lines of equation $Q - \lambda v = \gamma$. When γ is great, the curve $Q = Q(v)$ is entirely within the lower semi-plane limited by this line : $Q(v) - \lambda v < \gamma$ for any v. For a limit value $\gamma = \gamma(\lambda)$, the line meets one, or possibly more points of the curve, but this curve is still within the lower semi-plane (i.e. $Q(v) - \lambda v \leq \gamma$). It then appears that $\gamma(\lambda)$ is given by :

$$(5\text{-}2) \qquad \gamma(\lambda) = \sup_{A \in \beta} \{Q(A) - \lambda v(A)\}$$

and that the admissible contour(s) for which this maximum is
met correspond precisely to points where the line $Q - \lambda v = \gamma(\lambda)$
meets the curve $Q = Q(v)$ (such points are called "critical points").

The concave hull $Q = \hat{Q}(v)$, whose explicit expression is

$$\hat{Q}(v) = \underset{\lambda \geq 0}{\mathrm{Inf}} \{\gamma(\lambda) + \lambda v\}$$

is thus defined as the set of critical points of the various lines
obtained when λ varies. "Missing" optimal projects correspond to
convex parts of the curve $Q(v)$ and hence are, in general, of little
interest. Except for those few unimportant missing projects, we
have commuted the tedious problem (5-1) into a much easier problem
(5-2) : find, for each value λ, the project which maximizes $Q-\lambda v$.

The recovered reserves can then be parametrized in terms of λ :
for each $\lambda \geq 0$, the limit line $Q - \lambda V = \gamma(\lambda)$ presents the two ex-
treme critical points $(v_\lambda^-, Q_\lambda^-)$ and $(v_\lambda^+, Q_\lambda^+)$, hence $v_\lambda^- \leq v_\lambda \leq v_\lambda^+$
for any other critical point), and these two extreme points cor-
respond to the two projects A_λ^+ and A_λ^-. We have then come to the
following parametrization :

$$v^+(\lambda) = v(A_\lambda^+) \qquad v^-(\lambda) = v(A_\lambda^-)$$

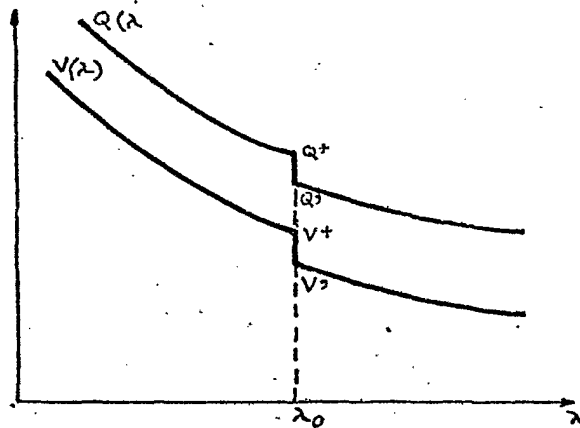$$Q^+(\lambda) = Q(A_\lambda^+) \qquad Q^-(\lambda) = Q(A_\lambda^-)$$

v^+ and Q^+ are decreasing functions and continuous to the left hand
side ; v^- and Q^- are decreasing and continuous to the right hand
side. Moreover :

$$\gamma(\lambda) = Q^+(\lambda) - \lambda \, v^+(\lambda) = Q^-(\lambda) - \lambda \, v^-(\lambda) \text{ is convex, decreasing}$$
and continuous.

This parametrization is unique, whatever the notations v^+, v^-,
etc... because

$$v^-(\lambda) \downarrow v^+(\lambda_0) \quad , \quad Q^-(\lambda) \downarrow Q^+(\lambda_0) \quad \text{for } \lambda \uparrow \lambda_0$$

$$v^+(\lambda) \uparrow v^-(\lambda_0) \quad , \quad Q^+(\lambda) \uparrow Q^-(\lambda_0) \quad \text{for } \lambda \downarrow \lambda_0$$
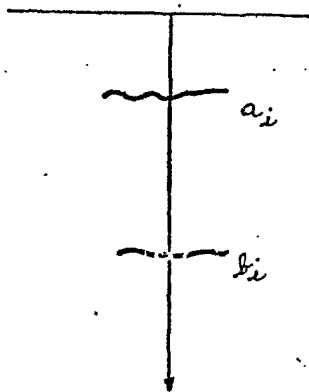
(this results from $v^+(\lambda) \geq v^-(\lambda) \geq v^+(\lambda_0)$ for $\lambda \uparrow \lambda_0$, etc..)

EXAMPLE 1 : Footwall and Hanging wall.

The problem is to define, for each panel P_i, the elevations a_i and b_i of the footwall and hanging wall. For each i, the kriged grade $z_i^*(x)$ of the horizontal slice of elevation x is computed, and it follows that

$$Q - \lambda v = \sum_i \int_{a_i}^{b_i} [z_i^*(x) - \lambda] \, dx$$

Parametrization in terms of λ is obtained by independent optimization of the various terms

$$\int_{a_i}^{b_i} [z_i^*(x) - \lambda] \, dx$$

as we assumed that there was no link-up cons- traints between panels.

For each λ, there are a limited number of pos- sibilities (as $z^*(x) < \lambda$ for any x just above a_i or just below b_i). For each panel :
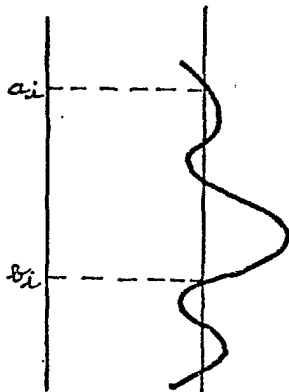
$$T_i(\lambda) = (b_i - a_i)$$

$$Q_i(\lambda) = \int_{a_i}^{b_i} z_i^*(x) \, dx$$

then : $T(\lambda) = \sum T_i(\lambda)$; $Q(\lambda) = \sum_i Q_i(\lambda)$

EXAMPLE 2 : Open Pit design.

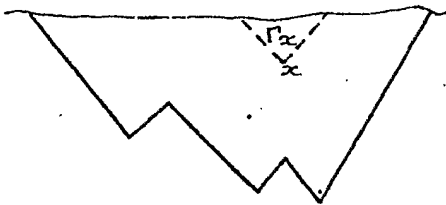Here for equilibrium slope purposes, all cones Γ_x with tip at x must be mined if it is decided to mine out point x, therefore any admissible contour $A \in \beta$ is a union of inside cones Γ_x.

It can be shown that :

1) For a given value $\lambda \geq 0$, there is a minimum size admissible contour B_λ^- and a maximum size B_λ^+ which fulfill the maximum of $Q(A) - \lambda\, v(A)$ for $A \in \beta$, with necessarily $B_\lambda^- \subset B_\lambda^+$ (generally, $B_\lambda^- = B_\lambda^+$, except for a few critical values of parameter λ).

2) These contours are nested within each other, when λ varies :

$$B_\lambda^+ \subset B_{\lambda'}^- , \quad \text{for} \quad \lambda > \lambda'$$

3) For each point x, there exists a maximum value $\Lambda(x)$ for parameter λ, such that x belongs to the optimal contour B_λ^+ as long as $\lambda \leq \Lambda(x)$, and no longer for $\lambda > \Lambda(x)$. This function $\Lambda(x)$, when known, thus allows parametrization of the optimal contours : indeed, we have the following criterium : $x \in B_\lambda^+$ if and only if $\Lambda(x) \geq \lambda$.

4) The relation $y \in \Gamma_x$ can be considered as an ordering relationship (y is inferior to x with respect to this ordering relationship if mining x implies mining y) and a given function f is said to be increasing with respect to the order Γ if $f(y) \geq f(x)$ as soon as $y \in \Gamma_x$. Such a function $\Lambda(x)$ which performs the parametrization of optimal contours must necessarily be increasing with respect to the order Γ.

More precisely, the following result can be stated : Amongst all Γ-increasing functions, the function $\Lambda(x)$ is the one that gives the best approximation of the grade function $Z(x)$, i.e. the function minimizing the integral

$$\int [Z(x) - \Lambda(x)]^2 \, dx$$

In other terms again : $\Lambda(x)$ is the <u>projection</u> of the grade $Z(x)$ on the set of Γ-increasing functions.

On this fundamental result are based procedures which allow practical parametrization of open pit recoverable reserves. The so-called "CG Open-Pit" program allows (under particular approximation) calculation of this function $\Lambda(x)$. <u>Parametrization</u> is then obtained as in the zoneography model of Lasky, replacing the grade $Z(x)$ by this function $\Lambda(x)$.

6 - <u>DOUBLE SELECTION UNDER CONSTRAINT.</u>

Admissible contours are here pairs $(A, A') \in \beta$ where A' is the volume given by extraction selection and $A \subset A'$ is the volume given by more selection. Let us denote $\varpi = (A, A')$; the $\varpi \in \beta$ are then called : "<u>admissible project</u>". For any $\varpi = (A, A')$, the following notations are considered :

$$V(\varpi) = V(A') \quad : \quad \text{Volume to be extracted.}$$

$$T(\varpi) = V(A) \quad : \quad \text{Volume (or tonnage) of recovered ore.}$$

$$Q(\varpi) = Q(A) \quad : \quad \text{Recovered quantity of metal.}$$

with $V(\varpi) \geq T(\varpi)$. The general underlying idea is the duality between the two parametrizations in (V, T) and in (λ, θ).

Parametrization in $(\underline{V}, \underline{T})$.

For any $v \geq T \geq 0$, the function

$$(6-1) \qquad Q(v, T) = \text{Sup} \{Q(\varpi), \varpi \in \beta, T(\varpi) \leq T, V(\varpi) \leq v\}$$

is an increasing function of the two variables v, T. This function provides a possible technical <u>parametrization</u> (under the only hypothesis that, between any two projects leading to the same values of v and T, the best project is the one that gives the greatest quantity of metal.)

But such a technical parametrization is, in practice, unaccessible because of the tedious combinatorial analysis involved. Here again $Q(v,T)$ should be replaced by its concave hull, which amounts to using dual parametrization.

Dual parametrization in λ, θ.

The function $Q(v,T)$ is replaced by its concave hull $\hat{Q}(v,T)$. For this, to each value of the pairs of parameters (λ, θ), we must associate (and it is generally possible to do so) the project or various projects providing the following maximum :

$$\gamma(\lambda, \theta) = \sup_{\omega \in \beta} \{Q(\omega) - \lambda V(\omega) - \theta T(\omega)\}$$

Such projects correspond to critical points of planes $Q - \lambda V - \theta T = \gamma(\lambda, \theta)$, i.e. to the points where these planes meet the surface $Q = Q(V,T)$, this surface being within the semi-space defined by this limiting plane. Those critical points thus define the concave hull of $Q(V,T)$, i.e. :

$$\hat{Q}(V,T) = \inf_{\lambda, \theta \geq 0} \{\gamma(\lambda, \theta) + \lambda V + \theta T\}$$

and they give, as functions of λ and θ, the following technical parametrization :

$$Q = Q(\lambda, \theta) \quad ; \quad T = T(\lambda, \theta) \quad ; \quad V = V(\lambda, \theta)$$

In practice, one parameter is fixed, e.g. $\theta = \theta_0$, and the parametrization in λ is done as was shown in the preceding paragraph. Then θ_0 is made variable, and the required parametrization in (λ, θ) is thus obtained.

EXAMPLE 1 : Defining footwall and hanging wall with overburden.

With the notations of section 5, it comes here :

$$Q(\omega) - \lambda V(\omega) - \theta T(\omega) = \sum_i \left[\int_{a_i}^{b_i} Z^*(x)dx - \lambda(b_i) - \theta(b_i - a_i) \right]$$

As the limits (a_i, b_i) are determined independently from one panel to another, the following optima :
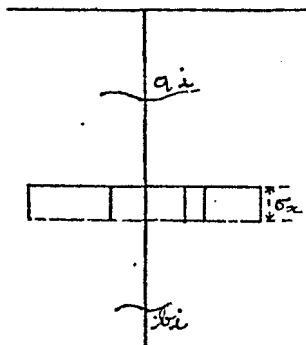
$$\int_{a_i}^{b_i} Z*(x)dx - (\lambda+\theta) b_i + \theta a_i$$

must be determined independently for each panel P_i : there are
but a small number of possibilities, since $Z*(x) < \theta$ for eleva-
tion $x > a$, and $Z*(x) \geq \theta$ for $x < a$; similarly $Z*(x) < \lambda+\theta$ for
$x < b$ and $\tilde{Z}*(x) \geq \lambda+\theta$ for $x > b$. The following parametrization
is thus readily obtained :

$$\begin{cases} Q(\lambda,\theta) = S \sum_i \int_{a_i}^{b_i} Z*(x) \, dx \\ T(\lambda,\theta) = S \sum_i (b_i - a_i) \\ V(\lambda,\theta) = S \sum_i b_i \end{cases}$$
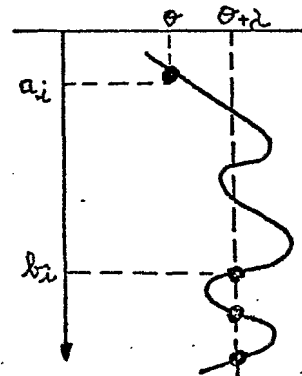
EXAMPLE 2 : <u>Defining footwall and hanging wall with precise se-
lection.</u>

Let us consider an uranium deposit. Within each
panel, an horizontal slice (a_i, b_i) is recovered
(whatever overburden may result). Then at the
mining stage a precise selection of ore is made
on small blocks of thickness δx.

On these blocks v, the criterium $Z*(v) \geq \theta$ al-
lows the technical parametrization of a given
volume (a slice a_i, b_i) already mined out. This
leads to the transfer functions of each hori-
zontal slice $(x, x+dx)$ for each panel P_i (of
surface S_i)

$$\begin{cases} T_i(x;\theta) = S_i \, \delta x \int_{\theta}^{\infty} F_i^*(x;dz) \\ Q_i(x;\theta) = S_i \, \delta x \int_{\theta}^{\infty} z \, F_i^*(x;dz) \end{cases}$$

Then it follows for the slice (a_i, b_i)

$$Q_i - \lambda V_o - \theta T_i = S_i \int_{a_i}^{b_i} (Q_i(x;\theta) - \theta T_i(x;\theta) - \lambda) \, dx$$

For each given value θ, a_i and b_i are given by the procedure detailed in section 5 (replacing $Z*(x)$ with $Q_i(x,\theta) - \theta\, T_i(x,\theta)$). Therefore a_i and b_i appear as functions of λ and θ, and by summing up, the parametrization $Q(\lambda,\theta)$, $T(\lambda,\theta)$, $V(\lambda,\theta)$ is obtained.

EXAMPLE 3 : <u>Open Pit design with free selection of ore</u>.

Each admissible contour being necessarily a union of blocks v, the problem is considered for each particular block v. For each block v_x (centred at x, $x \in E$ where E is finite), the transfer functions $T(x,\theta)$ and $Q(x,\theta)$ are computed. For each given value θ, the problem then reduces to a single λ-parametrization problem with $Z(x) = Q(x,\theta) - \theta\, T(x,\theta)$.

GAUSSIAN ANAMORPHOSIS MODELS

A. MARECHAL

- CONTENTS -

GAUSSIAN ANAMORPHOSIS MODELS

A. MARECHAL

I - INTRODUCTION

In practical problems of estimating recoverable reserves it is necessary to estimate the average density function of the variables within specified volumes (e.g. small blocks or large blocks). This, in its turn, requires the definition of multivariate probabilistic models (i.e. the joint density functions of two or more random variables). In addition to this, since the quantities to be estimated are always the average grades of blocks it is necessary to be able to determine the distribution law of the block grades from the distribution of sample grades. Moreover, the models of these distributions need to be general enough to describe the wide variety of distributions encountered in practice without requiring the estimation of too many parameters.

One possible solution to this kind of non linear problem is to carry out a conditional simulation of the deposit. Since most of the algorithms used for simulations generate data having a normal (i.e. Gaussian) distribution, it is necessary at the outset of the study to transform the Gaussian simulated variates so that they have the same distribution as the experimental data (i.e. so that their histogram is the same as the experimental histogram). For example, for a lognormal random variable, the transformation function (which we shall call the anamorphosis function in the following sections) is simply the exponential function.

As it happens, the models implicitly used in our simulation algorithms satisfy the conditions listed above and also possess a number of interesting properties associated with the properties of the normal distribution. Under these conditions it is natural to choose a Gaussian anamorphosis model as a first attempt of a model, firstly because of the functional simplicity of the normal

distribution, and secondly because it clarifies some of the problems arising in conditional simulations. .

II - REVIEW OF THE NORMAL (GAUSSIAN) DISTRIBUTION.

II-1 - Review of the properties of random functions (r.f.)

A random function can be defined by its spatial distribution function, that is the joint probability density function of N random variables $Z(x_i)$ taken at N sample points x_i. The problems we shall study will always consider only a finite number N of sample points of the random function, so that we shall only have to make use of the multivariate probability laws rather than of the full spatial law. The multivariate distribution law (i.e. the distribution function of a vectorial random variable $(Y_1, Y_2, \ldots, Y_N)$ is

$$F(Y_1, \ldots, Y_N = Prob(Y_1 < y_1, \ldots, Y_N < y_N)$$

Two types of distribution functions can be derived from this multivariate distribution function :

(i) the marginal distribution function

$$F_i(y_i) = F(+\infty, +\infty, \ldots, y_i, \ldots, +\infty)$$

i.e. the distribution of the random variable Y_i considered by itself.

(ii) the conditional distribution functions

$$F(y_i/y_j) = Pr(Y_i < y_i/Y_j = y_j)$$

The random variable defined by this distribution function is the conditional random variable Y_i given that Y_j takes the value y_j and will be noted $(Y_i | Y_j = y_i)$.

A random function $Y(x)$ is said to be totally Gaussian if all of its vectorial components have a multivariate normal distribution. In this case all of its marginal distributions will also be Gaussian.

From now on, we will only consider stationary models with a standard normal distribution, that is, those models where all the marginal distributions $P(Y(x) < y)$ are identically $N(0,1)$ (i.e. normally distributed with mean 0 and variance 1).

II-2 - Multivariate Normal Distribution.

An N-dimensional Gaussian random variable which has been centred and normed (i.e. mean 0, variance 1) is completely defined by a positive definite matrix H which gives the correlations between the components Y_i

$$H = \{\rho_{ij}\} \quad \text{where} \quad \rho_{ij} = \begin{cases} 1 & \text{if } i = j \\ E(Y_i \, Y_j) & \text{if } i \neq j \end{cases}$$

In the particular case where this random variable is associated with a stationary random function, we know that the correlation coefficient ρ_{ij} of any two components Y_i and Y_j is given directly by the covariance function $\rho(h)$ of the random function :

$$\rho_{ij} = \rho(h)$$
$$= \rho(x_i - x_j)$$

For the particular case of a bivariate normal distribution, the distribution will be defined by a single correlation coefficient ρ, the probability density function being

$$g_\rho(y_1, y_2) = \frac{1}{2\pi \sqrt{1-\rho^2}} \quad \exp \left\{ - \frac{y_1^2 + y_2^2 - 2\rho \, y_1 \, y_2}{2(1-\rho^2)} \right\}$$

The marginal density functions would then be the standard normal density functions

$$g(y) = \frac{1}{\sqrt{2\pi}} \exp \left(- \tfrac{1}{2} y^2 \right)$$

The main properties of Gaussian random variables are, for our purposes :

(i) All the marginal distributions are $N(0,1)$.

(ii) Every linear combination $\lambda^i \, Y_i$ of the Y_i is $N(0, \, \lambda^i \, \lambda^j \, \rho_{ij})$.

(iii) Two components Y_i and Y_j are independent if and only if $\rho_{ij} = 0$.

(iv) For a random variable with N+1 components $(Y_0, Y_1 \ldots, Y_N)$ the best estimator of Y_0 given $Y_1, \ldots, Y_N$ is the condition expectation $Y_0^* = E(Y_0/Y_1 = y_i, \ldots, Y_N = y_N)$. If the vector $(Y_0, Y_1, \ldots, Y_N)$ is completely Gaussian then the conditional random variable $E(Y_0/Y_1, \ldots, Y_N)$ is itself a Gaussian random variable with mean value $\lambda^i y_i$ and with variance $1 - \sigma_K^2$ where λ^i is the vector solution of the kriging system, and σ_K^2 is the corresponding kriging variance.

One property of kriging with a known mean and in the Gaussian case (in which case kriging is the conditional expectation) is that the estimator and the error of estimation are 2 independent normally distributed random variables $E(Y_0/Y_1, \ldots, Y_N)$ and $[Y_0 - E(Y_0/Y_1, \ldots, Y_N)]$.

II-3 - Bivariate and Univariate normally distributed random variables.

In the following we will consider the random variables Y_1, Y_2 having standard normal distributions and a correlation coefficient ρ. We let $g(y)$ and $G(y)$ denote the density functions and the distribution functions respectively of the standard normal distribution :

$$g(y) = \frac{1}{\sqrt{2\pi}} \, e^{-\frac{1}{2} y^2}$$

$$G(y) = \int_{-\infty}^{y} g(x) \, dx$$

Let $\Phi(y)$ be a measurable function, and let Y be a normally distributed random variable. Suppose that we wish to study the random variable $\Phi(Y)$ and that it is convenient to approach this random variable by developing a polynomial series of order n :

$$\Phi^*(Y) = a_0 + a_1 Y + a_2 Y^2 + \ldots + a_n Y^n$$

The best polynomial representation for $\Phi^*(Y)$ is the one which

is unbiased and which minimizes the variance of the deviation
$\Phi(Y) - \Phi*(Y)$:

$$I = \int_{-\infty}^{\infty} [\Phi(Y) - a_0 - a_1 Y - \ldots - a_n Y^n]^2 g(y) \, dy$$

The problem is easily resolved by using the Hermite polynomials
$H_1, H_2, \ldots, H_n$ instead of the monomials $Y_1, Y_2, \ldots, Y^n$ where the Hermite polynomials are defined by

$$n > 0 \, , \, H_n(y) \, g(y) = \frac{d^n}{d \, y^n} \, g(y) \, , \, H_0(y) = 1$$

This family of polynomials has the remarkable property of being orthogonal for the Gaussian density $g(y)$

$$\int_{-\infty}^{\infty} H_n(y) \, H_m(y) \, g(y) \, dy = n! \, \delta_{n,m}$$

where $\delta_{n,m}$ = 1   if $n = m$
             0   if $n \neq m$

Consequently they will be particularly convenient in our polynomial representation, $\Phi*(Y)$:

$$\Phi*(Y) = \sum_{n=0}^{N} \frac{\psi_n \, H_n(y)}{n!}$$

where the ψ_n are presently undetermined numerical coefficients.
These coefficients can be obtained directly from

$$\psi_n = \int_{-\infty}^{\infty} \Phi(y) \, H_n(y) \, g(y) \, dy$$

or equivalently from

$$\psi_n = E(\Phi(y) \, H_n(y))$$

The reason for using the first of these two representations for
$\Phi(Y)$ comes precisely from the fundamental property of the family
H_n, i.e. if

$$\Phi(Y) = \sum_{n=0}^{N} \frac{\psi_n \, H_n(Y)}{n!}$$

Since $E[H_n(Y)] = 0$ for $n \neq 0$, we see that

$$E[\Phi(Y)] = \psi_0$$

and
$$V[\Phi(Y)] = \sum_{n=1}^{N} \frac{\psi_n^2}{n!}$$

By taking account of the properties of the family H_n we obtain the classical result for the bivariate normal density

$$g_\rho(y_1, y_2) = \sum_{n=0}^{\infty} \frac{\rho^n}{n!} H_n(y_1) H_n(y_2) g(y_1) g(y_2)$$

The term isofactorial is used to describe a bivariate distribution law whose joint density function $f(y_1, y_2)$ can be expressed in the form :

$$f(y_1, y_2) = \sum T_n \chi_n(y_1) \chi_n(y_2) f(y_1) f(y_2)$$

where y_1 and y_2 have the same marginal density function $f(y)$, and where $\chi_n(y)$ are the orthogonal polynomials associated with the marginal distribution.

The definition given above leads naturally to expressing the density function of the random variable $(Y_1/Y_2 = y_2)$ in terms of Hermite polynomials. The classical expression for $f(y_1/y_2)$ is

$$f(y_1/y_2) = \frac{g_\rho(y_1, y_2)}{g(y_2)}$$

Hence
$$f(y_1/y_2) = \sum_{n=0}^{\infty} \frac{\rho^n}{n!} H_n(y_1) H_n(y_2) g(y_1)$$

A simple representation for the expectation of (Z_1/Z_2) can easily be derived for the case where $Z_1 = \Phi_1(Y_1)$, $Z_2 = \Phi_2(Y_2)$ and where Y_1 and Y_2 have a bivariate normal distribution with a correlation coefficient ρ.

Let
$$\Phi_1(y) = \sum \frac{\psi_{1n}}{n!} H_n(y)$$

and
$$\Phi_2(y) = \sum \frac{\psi_{2n}}{n!} H_n(y)$$

then we have

$$E(Z_1/Z_2) = \sum_n \frac{\psi_{1,n}}{n!} \rho^n H_n(Y_2)$$

or alternatively

$$E(Z_1/Z_2) = \Phi_{1,\rho}(Y_2)$$

where $\Phi_{1,\rho}$ is the function

$$\Phi_{2,\rho}(y) = \sum \rho^n \frac{\psi_{1,n}}{n!} H_n(y)$$

II-4 — The anamorphosis of the Gaussian random variables.

The results from the preceding section on the two random variables Z_1 and Z_2 are more easily seen if it is assumed that the pair (Y_1, Y_2) has a bivariate normal distribution and if Z_1 and Z_2 are written in the following form :

$$Z_1 = \Phi_1(Y_1) \quad \text{and} \quad Z_2 = \Phi_2(Y_2)$$

where

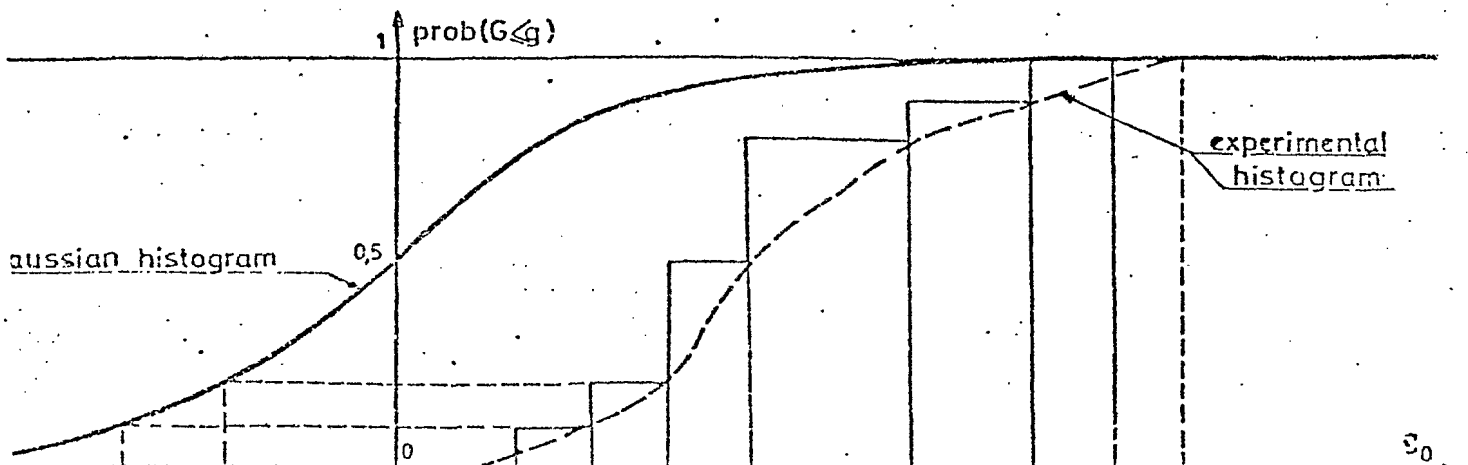$$\Phi_1(y) = \sum_n \frac{\psi_n^1}{n!} H_n(y)$$

$$\Phi_2(y) = \sum_n \frac{\psi_n^2}{n!} H_n(y)$$

and where Y_1 and Y_2 are Gaussian random variables. In particular the problem of the regression of Z_1 on Z_2 is resolved simply as we shall see.

Any continuous random variable can be represented as the function $\Phi_1(Y_1)$ by the operation of anamorphosis (used in Monte-Carlo simulations).

Consequently we can always consider a stationary random function $Z(x)$ as the anamorphosis of a stationary random function $Y(x)$ with a Gaussian marginal distribution where $Z(x) = \Phi[Y(x)]$ but in general the random function will not be completely Gaussian (i.e. in its joint distribution function). To be able to take advantage of properties of the normal distribution , we must, at a minimum, assume that all pairs of random variables $Y(x_1)$ and $Y(x_2)$ have a bivariate normal distribution. In practice we often even assume that $Y(x)$ is completely Gaussian ; for example, in the "discrete Gaussian" model.

III - THE DISCRETE GAUSSIAN MODEL.

III-1 - An example of the lognormal case.

The most commonly used probability model in modelling geological phenomena is without doubt the lognormal. The lognormal random variable Z is already expressed in terms of a Gaussian anamorphosis since by its very definition there exists a normally distributed random variable Y such that $Z = \exp(m + \sigma Y)$ (i.e. $\ln Z \sim N(m,\sigma^2)$). This definition can easily be extended to cover the case of the lognormal random functions encountered in geostatistics : we let $Z(x)$ be a random function such that $Z(x) = \exp[m + \sigma Y(x)]$ where $Y(x)$ is a completely Gaussian random function. Of course a model of this type can only be used to refer to the average grades from constant sized supports or from drill holes of uniform size which, for example, can be considered as point samples of grade $Z(x)$. Continuing with our notation we have

$$Z(x) = \exp[m + \sigma Y(x)]$$

where $Y(x)$ is a random function having a centred and normed, Gaussian distribution with a correlogram $\rho(h)$.

The classical formulae for the lognormal provide us with the main statistical characteristics of $Z(x)$ in terms of the parameters m and σ :

$$M = E(Z) = \exp(m + \frac{\sigma^2}{2})$$

$$\Sigma^2 = \text{Var}(Z) = M^2 (e^{\sigma^2} - 1)$$

$$\Sigma(h) = E[(Z(x+h)-M)(Z(x) - M)] = M^2(e^{\rho(h)} - 1)$$

Moreover, as was shown earlier in the section on conditional distributions, the conditional variable $(Y_0/Y_1 = y_1, Y_N)$ is also completely Gaussian with mean $\lambda^i y_i$ and variance $1 - \sigma_K^2$, and can be written as

$$\lambda^i y_i + \sqrt{1 - \sigma_K^2} \, Y_0^1$$

The transformed variable $(Z_0/Z_1 = e^{m+\sigma y_1}, \ldots, Z_n)$ is then lognormal and can similarly be written

$$(Z_0/Z_1, \ldots, Z_n) = \exp(m + \sigma \lambda^i y_i + \sigma \sqrt{1 - \sigma_K^2} \, Y_0^1)$$

It is evident that provided we are prepared to assume that $Y(x)$ is completely Gaussian, it is possible to determine theoretically the conditional distribution of the point grades from the data given, by using the lognormal theory : their distributions are also lognormal. We may say that this is a kind of "permanency" of the type of distribution.

To establish a lognormal model for a vector of N point grades does not present any particular difficulty, but it does not correspond to the problem most frequently encountered in mining studies, i.e. to know how to define the density function of the average grade Z_V of a panel V and also the joint density function of this grade Z_V and the point grades $Z(x_\alpha)$ of the samples. In the case where the point grades are lognormal, it has been shown experimentally that the average grades of the panels are at least approximately lognormally distributed (with a much smaller variance). This phenomenon is referred to as the "permanency of the lognormal distribution".

Continuing the same notation as before, the parameters of the new lognormal distribution can easily be derived.

Let $$Z(x) = \exp[m + \sigma Y(x)]$$

where $Y(x)$ has a correlogram $\rho(h)$. Then the covariance function for the point grades $Z(x)$ is

$$\sum(h) = M^2 (e^{\rho(h)} - 1).$$

The theoretical variance $\sum_V^2$ of the average panel grade $Z_V = \frac{1}{V} \int_V Z(x) \, dx$ is known from geostatistics to be

$$\sum_V^2 = \overline{\sum(V,V)} = M^2 \left\{ \frac{1}{V^2} \int_V \int_V e^{\rho(x-y)} \, dx \, dy - 1 \right\}$$

This determines $\sum_V^2$ in terms of the point grade parameters m, σ and ρ. If we assume that Z_V is also lognormal, then there exists a Gaussian random variable X such that

$$Z_V = \exp(m^1 + \sigma^1 X)$$

where

$$E(Z_V) = E[Z(x)] \Rightarrow m^1 + \frac{\sigma^{12}}{2} = m + \frac{\sigma^2}{2}$$

$$Var(Z_V) = \sigma_V^2 \Rightarrow \sigma^{12} = \log[e^{\overline{\sigma(V,V)}}]$$

Consequently if we accept the "permanency of the lognormal distribution", the parameter values from the point model allow us to completely determine the distribution of panel grades Z_V. In fact, the same principle can even be applied to the vectorial random variable $(Z_{\alpha 1}, \ldots, Z_{\alpha n}, Z_{V1}, \ldots, Z_{Vn})$ when the point variables Z_{α_i} are considered simultaneously with the regularized variables Z_{V_j}. We are, of course, assuming that the point variables Z_{α_i} are lognormal with $Z_{\alpha_i} = \exp(m + \sigma Y_{\alpha_i})$, that the regularized variables Z_{V_j} are lognormal with $Z_{V_j} = \exp(m^1 + \sigma^1 X_{V_j})$ and that the vectorial random variable $(Y_{\alpha_1}, \ldots, Y_{\alpha_n}, X_{V_1}, \ldots, X_{V_n})$ is completely Gaussian and is centred and normed with correlation coefficients $\{r_{ij}\}$ such that after the lognormal transformations have been applied,

these correlations are equal to the values $\Sigma_{\alpha\beta}$, $\Sigma_{\alpha V}$, $\Sigma_{V_1 V_2}$ of the correlations between the Z_α and the Z_V.

In reality theoretical studies have shown that the "permanency of the lognormal distribution" cannot be strictly true : it is not possible for both the point variables Z_{α_i} and the regularized variables Z_{V_j} to have a lognormal distribution. In fact the discrete Gaussian model has been developed to generalize the lognormal model a little more rigourously.

III-2 - The general case.

The essential principles of this model are the same as for the lognormal model given earlier :

. To define a principle of "permanency" which allows us to determine the distribution of the regularized variable Z_V from the parameters of the distribution of the point random variables $Z(x)$. The $Z(x)$ are considered as the anamorphosis $\Phi(Y(x))$ of a Gaussian random function $Y(x)$, and the Z_V as the anamorphosis $\Phi_V(X_i)$ where X_i is also a Gaussian random function. The principle of "permanency" will be used to link the parameters of the two anamorphoses Φ and Φ_V.

. To assume the variables $Y(x_\alpha)$ and X_i (which are individually normally distributed) are jointly completely Gaussian.

As we have noted earlier, this hypothesis is only possible if we alter slightly the significance of the point variables : in order to do this we are going to select a particular unit volume, the size of which is supposed to have a particular significance (e.g. size of a block for selection, for simulation, etc.). We are going to divide the deposit into a grid of such blocks v_i. The data is assumed to be known at the level of point samples (or is considered to be) on a grid much wider than v so that at most one sample $Z(x_\alpha)$ can be found in a given block v_i. It is then possible to define a model which is completely Gaussian. This is done by considering the sample Z_α to be located at random in the

small block v_α containing it. (By doing this we, of course, lose some of the information available to us concerning the exact location of the sample within v_α). The result of this approximation is that $E(Z_\alpha/Z_{v_\alpha}) = Z_{v_\alpha}$, which allows us to link the parameters of our point sample model to those of the samples of volume V_α.

The deposit is divided into N blocks v_i. The grade Z_{v_i} of each of these blocks is associated by the anamorphosis Φ_V : $Z_{v_i} = \Phi_V(X_i)$ with one of the components of the Gaussian vector $(X_1,\ldots,X_N)$. Each block v_i may also contain a point sample Z_i which is, in turn, associated by a second anamorphosis Φ : $Z_i = \Phi(Y_i)$ with another Gaussian vector $(Y_1,\ldots,Y_N)$.

The following hypotheses are made concerning the Gaussian distributions :

a) For all blocks v_i, the two random variables (X_i,Y_i) have a standard bivariate normal distribution with a correlation coefficient $r > 0$ which is the same for all blocks v_i.

b) As was seen earlier, a consequence of randomizing the location of Z_α within v_α is that $E(Z_\alpha/Z_{v_\alpha}) = Z_{v_\alpha}$. We further assume that for X_i fixed, Y_i is independent of all the other random variables X_j, Y_j in the model.

c) $(X_1,\ldots,X_N)$ is completely Gaussian.

Taken together, the hypotheses a), b) and c) imply that the vector $(Y_1, X_1, Y_2, X_2,\ldots,Y_N, X_N)$ is completely Gaussian, is centred and normed and may be shown to have the following correlations :

- block $\times$ block correlations $R_{ij} = E(X_i X_j)$ are, in fact, given by the covariance function $R(h)$ of the inverse anamorphosis X_i of the block grades, i.e.

$$R_{ij} = R(x_\alpha - x_\beta)$$

- point $\times$ block correlations $r_{i\alpha} = E(X_i Y_\alpha) = r R_{i\alpha}$

- point $\times$ point correlations $\rho_{\alpha\beta} = E(Y_\alpha Y_\beta) = r^2 R_{\alpha\beta} \quad \alpha \neq \beta$

So the resulting set of correlation coefficients is determined by two quantities : $R(h)$, the covariance function of the X, and r, the parameter defined above as $E(Y_\alpha X_i)$ (i.e. the covariance between a block equivalent X_i and its associated point equivalent Y_α).

III-3 - Permanency in the discrete Gaussian model.

The preceding model completely defines the random variables Y_α and X_i which are the inverse anamorphoses of the point grades $Z(x_\alpha)$ and the average grades Z_{v_i}, respectively. As usual the point anamorphosis function Φ will be determined from the experimental histogram of the point grades $Z(x_\alpha)$; the anamorphosis Φ_v of block grades must be derived directly from the preceding model.

Suppose that $\Phi(y)$ can be represented by

$$\Phi(y) = \sum_0^\infty \frac{\phi_n}{n!} H_n(y)$$

Point a) of the model is that within a given block v_i, (X_i, Y_i) is a bivariate Gaussian with correlation r : we thus have

$$E(Z_i / Z_{v_i}) = E(\phi(Y_i)/X_i) = \sum_{n=0}^\infty \frac{\phi_n r^n}{n!} H_n(X_i)$$

Point b) is that $E(Z_i / Z_{v_i}) = Z_{v_i}$ so we have

$$Z_{v_i} = \sum_{n=0}^\infty \frac{\phi_n r^n}{n!} H_n(X_i)$$

in which we recognize precisely the anamorphosis function of block grade ϕ_v

$$\phi_v(y) = \sum_{n=0}^\infty \frac{\phi_n r^n}{n!} H_n(y)$$

This result is interesting because it allows us to adjust the value of r experimentally. In fact the geostatistical study provides us with the variogram $\gamma(h)$ of the point samples and consequently the covariance $S(h) = \gamma(\infty) - \gamma(h)$, and also the covariance of the block grades $S_v(h) = \overline{S(v, v_h)}$. In particular, $S_v(o) = \overline{S(v, v)}$ is the

theoretical variance of the block grades Z_{v_i}. We have seen earlier
that the variance associated with an anamorphosis function can be
calculated directly from the coefficients ψ_n :

Let $\quad\quad \Phi_v = \sum \dfrac{\psi_n \, r^n \, H_n(y)}{n!}$

then

$$Var[\Phi_v(X)] = \sum \dfrac{\psi_n^2 \, r^{2n}}{n!}$$

The parameter r can then be determined from the relation

$$S_v(o) = \sum_1^\infty \dfrac{\psi_n^2 \, r^{2n}}{n!}$$

The solution can easily be found since the series is an increasing function of r. Once r has been determined, the same relation can be used for the covariance function $S_v(h)$ (which is known from a theoretical regularization of the experimental covariance $S(h)$) and $R(h)$ from the block anamorphosis :

$$S_v(h) = \sum_1^\infty \dfrac{\psi_n^2 \, r^{2n} \, R(h)}{n!}$$

The numerical method which was used to determine r, is now used again to provide the values of $R(h_i)$ for a series of values h. A theoretical model for $R(h)$ is then chosen.

III-4 — Conclusion.

The discrete Gaussian anamorphosis model provides a convenient method for carrying out the calculations which arise in connexion with the point distribution of the point grades and the block grades. Moreover, it allows us to do this whatever the distribution of the point samples. However it should be noted that the model suffers from two restrictions :

. Firstly the assumption that the Y_α and the X_i are completely Gaussian. This extremely strong hypothesis is difficult to verify except for the experimental bivariate histograms. It is known that this hypothesis will be false in many cases, but the results coming

from this extremely useful hypothesis are valid in many cases where the hypothesis itself is not strictly true. Besides this, the hypothesis of the independence of the Y_α for fixed X_α has unpleasant theoretical consequences which restrict the generality of the model (for example it can be shown that the local histograms for within a panel are entirely determined when the mean of the panel is fixed).

. The second restriction concerns the theoretical approximations, that is, the assumptions regarding the "permanency" of distributions. In the case of the lognormal distribution this is precisely the "lognormal permanency" which, as was mentioned earlier, is theoretically not possible.

In conclusion, the discrete Gaussian model is not an "all purpose" model which can be applied willy-nilly in all cases. In fact several other models are available (viz the Hermitian Gaussian model, the non-Gaussian isofactorial model) which are better suited for use with some of the variables studied than the discrete Gaussian model.

IV - A PRACTICAL EXAMPLE : THE CALCULATION OF THE OVERALL GRADE/TONNAGE CURVE.

IV-1 - The Problems.

Consider a deposit near the surface which is to be mined by open pit methods. Suppose that the size of the blocks for selection has been fixed (for example, as cubes the size of the bench height). As a first approximation assume that the blocks to be mined may be selected without any constraints, and that all the blocks whose average grade Z_v is above the cut-off grade z_c will be mined. Under these conditions, the exploitable results will be given by the cut-off grade/tonnage curve which is simply the estimator of the proportion of block grades above the cut-off grade z_i. The problems to be resolved are then :

- obtain a good estimator of the distribution of point grades

within the mineralized part of the deposit, given the experimental values of the point grades.

- model the distribution law using an anamorphosis function and to obtain numerically its expansion in terms of Hermite polynomials.

- determine the representation of the distribution of block grades Z_V from the model obtained above by using the "permanency of distributions" as outlined in the previous section.

- use this numerical model to calculate the proportion of blocks v whose grade is above the cut-off grade and to determine the mean grade of these blocks.

IV-2 - Determination of the distribution of point grades.

Let V be the mineralized region. We wish to estimate the function $P_V(z)$ which gives the proportion of the ore on V whose grade $Z(x) < z$. Or, more formally, if $F_x(z)$ is the distribution function of $Z(x)$, the expected value of $P_V(z)$ (to be denoted by $F_V(z)$) is simply the mean value of the point distribution $F_x(z)$, with the average taken over the volume V, i.e.

$$F_V(z) = \frac{1}{V} \int_V F_x(z) \, dx$$

The problem of estimating $P_V(z)$ is similar to that of estimating Z_V, the average grade in V, since for a fixed z, P_V is the mean value over V of the indicator functions $I_z(Z(u))$ where

$$I_z[Z(u)] = 1 \quad \text{if} \quad Z(u) < z$$

$$= 0 \quad \text{if} \quad Z(u) > z$$

One obvious estimator of $P_V(z)$ is, of course, the histogram $F^*(z)$ of the available experimental sample values. This has the same role vis-à-vis $F_V(z)$ as does the arithmetic mean $\bar{Z}_\alpha$ of the samples in the estimation of the average grade Z_V. If the region V is large (or if it covers the whole of the mineralized zone) and if the sample points x_α are distributed sufficiently uniformly throughout V, then we know that $\bar{Z}_\alpha$ is a satisfactory estimator

of Z_V, and similarly that $F^*(z)$ will be a reasonable estimator of $F_V(z)$. On the other hand if V has been chosen in the centre of the mineralization and is surrounded by samples which are consistently of a lower grade than those in V (or if the samples are not evenly distributed) then we know that Z_α is not a good estimator of Z_V because it does not take account of the grade dilution by the lower grades at the boundaries. In the latter case, kriging provides a much better estimator, Z_{KV}, of Z_V. For the same reasons it is advisable to estimate $F_V(z)$ by a suitable weighting procedure using the point samples inside and outside the region under study.

The definition of the theoretical estimator of $F_V(z)$ requires the techniques of disjunctive kriging which will be considered else. In practice, we take advantage of the fact that the Z_V have usually been kriged (either in an overall (global) kriging of V or alternatively in a panel by panel kriging of V) to define the weighted cumulative histogram of the experimental data Z_V as the estimator of $F_V(z)$.

As part of the kriging process, each of the samples available was assigned its own weighting factor λ_α (obviously 0 if x_α was too far away from V). The weighted histogram is obtained by cumulating to the class in which a value belong to a relative frequency λ^α rather than the usual absolute frequency (1). Each class of the histogram then has a relative frequency proportional to the sum of the λ_α for that class. The drawbacks of this procedure are :

- the necessity of storing all the λ_α used in the kriging.

- the slight possibility of obtaining a negative frequency which would then be set arbitrarily to zero.

On the other hand, the advantages are that

- it takes into account the position of the sample points (because it uses the variogram).

- the experimental histogram thus obtained has a mean of precisely Z_{KV}.

IV-3 - Hermitian anamorphosis model for the distribution of point samples.

Once an estimate $F*(z)$ has been obtained for the average distribution of point samples, we can find the Gaussian anamorphosis $Z* = \Phi(Y)$ of the associated variable $Z*$. If $G(y)$ denotes the distribution function of a standard normal $N(0,1)$, we have

$$F*(\Phi(y)) = G(y)$$

$$\Phi(y) = F*^{-1}[G(y)]$$

Since the experimentally graded function $F*(z)$ is defined over the classes z_i, the relation $F*^{-1}$ is therefore defined only over these classes. We can therefore

- either model $F*(z)$ by an appropriate distribution function F^m (or by a combination of such functions) so that the representation $F^m(z)$ is continuous and monotonic increasing. The inverse $F^{m-1}[G(y)]$ then provides a continuous model for $\Phi(y)$.

- or alternatively define $\Phi(y)$ in a discontinuous manner for the values of y corresponding to the experimental classes z_i and then modify this model to be continuous.

In fact, the model of the function $\Phi(y)$ is not particularly important since, as we have seen earlier, all the calculation can be done using the Hermitian expansion of $\Phi(y)$. Consequently we use the alternative method and try to adjust the graph of the points (y_i, z_i) using a polynomial of the type

$$\sum_{0}^{N} \frac{\psi_n}{n!} H_n(y)$$

We know that the coefficients ψ_n are given by the relations

$$\psi_n = \int_{-\infty}^{\infty} \Phi(y) H_n(y) g(y) dy$$

The problem is to estimate these integrals numerically, given the values of $Z = \Phi(y)$ for the sample points (z_i, y_i).

We use one or the other of the following two methods :

(i) Gauss's method for integrating : the integral giving ψ_n is estimated from the values of $\Phi(y_\alpha)$, $\alpha = 1,\ldots,k$ where the points y_α are defined theoretically as the solutions of the equation $H_k(y) = 0$, so that

$$\psi_n^* = \sum_{j=0}^{k} \varphi(y_k) \, H_n(y_k) \, W_k \quad (W_k \text{ appropriate weights})$$

(see reference A. Maréchal : "The practice of Transfer functions : Numerical methods and their applications").

(ii) Exact integration of a simplified representation of $\Phi(y)$ (i.e. linear or quadratic) in the classes (z_i, z_{i+1}) of the histogram.

In both cases it is necessary to assume the form taken by the experimental histogram, especially for its tails. Both methods are interactive by nature : they allow us to try different numerical values for the tails of the histogram.

Lastly in this section we consider representing $F_V(z)$ as an anamorphosis function expanded in terms of Hermite polynomials, i.e.

$$\Phi(y) = \sum_{n=0}^{N} \frac{\psi_n \, H_n(y)}{n!}$$

It should be noted that $\Phi(y)$ completely defines the model of the distribution function, $F_V(z)$, and in particular

$$\psi_0 = \text{Mean } [F_V(z)] = z_V$$

$$\sum_{1}^{N} \frac{\psi_n^2}{n!} = \text{Variance } [F_V(z)] = \Sigma^2 - \Sigma_V^2$$

Consequently, since the structural analysis of the sample values $Z(x_\alpha)$ has a point variogram $\gamma(x)$, the variance $\overline{\gamma(V,V)}$ must be equal to $\sum \frac{\psi_n^2}{n!}$. In fact the region V usually covers practically all the region under study and so the variogram model is chosen

so that the sill $\gamma(\infty)$ equals $\sum \dfrac{\phi_n^2}{n!}$.

IV-4 - Modifying the anamorphosis for a change of support size.

The anamorphosis function $\Phi_V(y)$ for the variable Z_V defined in the region V can be deduced from the function Φ using one of the properties of the discrete Gaussian model.

$$\Phi_V(y) = \sum_o^N \frac{\phi_n}{n!} r^n H_n(y)$$

where the coefficient r satisfies the equation

$$\Sigma_V^2 = \sum_{n=1}^N \frac{\phi_n^2}{n!} r^{2n} \qquad n > 0$$

Σ_V^2 can be obtained by reading the charts (if the variogram $\gamma(n)$ is a linear combination of standard variograms) or otherwise by numerical integration of $\gamma(h)$ over the block v. The equation given above for r can easily be solved numerically because the function

$$S(r) = \sum_1^N \frac{\phi_n^2}{n!} r^{2n}$$

is a monotonic increasing function of r ; it suffices to increase r from 0 to 1 until the value of Σ_V^2 is reached (this is necessarily less than

$$\sum_1^N \frac{\phi^n}{n!}$$

IV-5 - Calculation of the overall (i.e. global) cut-off grade/ tonnage curve.

If the block size for selection is v, the recoverable reserves in the region V are found by considering

$T_V(z_c)$ - the recoverable tonnage for cut-off grade z_c.

$Q_V(z_c)$ - the total quantity of metal recovered.

Let T_o denote the total tonnage contained in V. Since $F_V(z)$ is

the distribution function corresponding to the anamorphosis function Φ_V , we have

$$T_V(z_c) = T_0 \left[1 - F_V(z_c) \right]$$

$$Q_V(z_c) = T_0 \int_{z_c}^{\infty} z \, d \, F_V(z)$$

Since $z = \Phi_V(y)$, $F_V(z) = G(y)$.

To begin with, the value of y_c corresponding to z_c

$$z_c = \sum_{n=0}^{N} \frac{\psi_n \, r^n}{n!} \, H_n(y_c)$$

is determined numerically. The value of $T_V(z_c)$ can then be determined directly from the integral $G(y_c)$. (Excellent numerical approximations are available for this function in all mathematical tables).

$$T_V(z_c) = T_0 \left[1 - G(y_c) \right]$$

The calculation of Q_V is slightly more complicated :

$$Q_V(z_c) = T_0 \int_{z_c}^{\infty} \Phi_V(y) \, z(y) \, dv$$

$$= T_0 \sum_{n=0}^{N} \frac{\psi_n \, r^n}{n!} \int_{y_c}^{\infty} H_n(y) \, g(y) \, dy$$

A direct consequence of the relation defining the Hermite polynomials is :

$$\int H_n(y) \, g(y) \, dy = H_{n-1}(y)$$

Hence

$$Q_V(z_c) = T_0 \left\{ \psi_0 \left[1 - G(y_c) - g(y_c) \sum_{1}^{N} \frac{\psi_n \, r^n}{n!} \, H_n(y_c) \right] \right\}$$

The average grade recovered $m_V(z_i)$ is then

$$m_V(z_c) = \frac{Q_V(z_c)}{T_V(z_c)}$$

IV-6 - Concluding remarks on the calculation of the grade/tonnage curve.

In the preceding sections, it has been shown that the discrete Gaussian model allows us to estimate the recoverable reserves for blocks of size v provided it is assumed that the blocks can be selected freely. It is clear that this will never be the case in reality. However these techniques provide a method for estimating the recoverable reserves very soon after the initial geological studies have been carried out. A more precise estimation of these reserves can be carried out later using transfer functions.
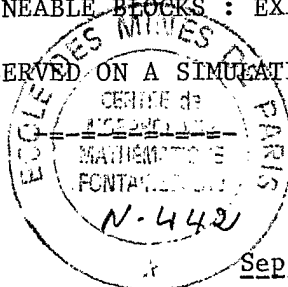
Rajouter
dans original
Après la page 22 :

+ N-442 (G. Matheron)

+ ~~[illisible]~~

Fontainebleau   N-442

SELECTING MINEABLE BLOCKS : EXPERIMENTAL

RESULTS OBSERVED ON A SIMULATED OREBODY

A. MARECHAL   Septembre 1975

NATO A.S.I., Rome, Italie, 13-25 Octobre 1975

SELECTING MINEABLE BLOCKS : EXPERIMENTAL RESULTS OBSERVED ON A
SIMULATED OREBODY

A. Maréchal

Centre de Morphologie Mathématique, Ecole Nationale
Supérieure des Mines de Paris, Fontainebleau, France.

ABSTRACT. The problem of optimal selection from estimates is stu-
died theoretically in a simple case : it is shown that the solu-
tion is to apply a cut-off criterion to the conditional expectation
of each panel's grade ; but if it is impossible to compute such an
estimator, then no simple, nearly optimal substitute can theoreti-
cally be found. However, we know that kriging is the best linear
least-square approximation of conditional expectation, so we may
hope that it would be a good practical substitute : this is expe-
rimentally shown on a two-dimensional simulated orebody on which
selection is performed with various estimators.

0. INTRODUCTION

When assuming that the true block's grades are known, it is
easy to demonstrate that maximizing a linear benefit by selecting
mineable blocks is achieved by using a cut-off grade criterion on
the true block's grades. When reality is unknown, the problem is
set in probabilistic terms : if the objective function is the ex-
pected value of a linear benefit, the optimal procedure consists
in using a cut-off grade criterion on the conditional expectation
of the block's grades. However, in practice the conditional expec-
tation cannot be determined and will be replaced by a linear esti-
mator of the block's grade. Hence it can be expected that the result
obtained with this procedure will not be so good as the optimum
selection. However, the actual loss in expected profit is diffi-
cult to determine theoretically. Therefore it is interesting to
compare the results of the true and approximate selections made on
a simulated orebody. Moreover, we shall take profit of the knowledge

of both true and estimated values to point out some experimental properties of the estimators used in geostatistics.

1. A VERY SIMPLE CASE OF SELECTION

In order to allow an easy theoretical exposition, let us imagine a simple situation : we mine a mineralized layer whose width is considered as constant, with a mining method based on elementary mineable panels of tonnage T_ℓ. We essentially assume that no technical constraints of accessibility are imposed on mining, so that any subset of elementary panels P_i is technically mineable.

We assume that the total profit B yielded by mining a total tonnage T of mean ore grade m is :

$$B = p\, m\, T - p_1 T - p_2$$

p : sale price of metal/ton metal

p_1 : mining and processing cost/ton of ore

p_2 : fixed costs

To simplify the notation, we shall write the profit as

$$B = p\, \left[T(m-m_c) - r\right] \quad \text{with } m_c = p_1/p \; , \; r = p_2/p$$

There is no loss of generality in taking p = 1.

1.1 Selection when the true block grades are known

We know each m_i, mean grade of the panel P_i. Selection will consist in choosing a subset $\{P_i, i \in I\}$ of the sets of panels P_i. Mining this subset will yield the total profit :

$$B = \sum_{i \in I} T_\ell (m_i - m_c) - r$$

Maximizing B consists in choosing I so that each term $(m_i - m_c)$ ≥ 0, that is keeping each panel whose ore grade m_i is greater or equal to the cut-off m_c. If the corresponding maximized profit B is positive, mining the orebody is profitable for these precise economical conditions.

1.2 Selection when the true block grades are unknown

We must abandon the deterministic representation of grades and replace it by a probabilistic one. Let us consider the quantities, either known or unknown, which arise in the problem :

- M_i, $i = 1, N$: Panel mean grades (unknown).

- $Z_{\alpha j}$, $j = 1, n$: Available sample grades (known).

According to the commonly used probabilistic representation, we define the vectorial random variable (v.R.V.) n+N dimensional $(M_i, i = 1, N, Z_{\alpha j}, j = 1, n)$.

The total profit yielded by mining the subset $\{P_i, i \in I\}$ is then a random variable $B = \sum_{i \in I} T_\ell (M_i - m_c) - r$. Optimization may then apply to various quantities, such as :

Max $E(B)$, or Max $(\text{Prob } B > b_o)$ or Min $\{\text{Prob } B < b_m\}$

The selection of one of those criteria depends of the company policy : the most commonly used among large companies seems to be maximizing the expected profit $E(B)$. Moreover, it is the only criterion which allows an easy theoretical computation.

In order to ease the presentation, we may, without any loss of generality, take $r = 0$ and $T_\ell = 1$. Furthermore, we note that B_i being the profit yielded by each panel P_i, Max $E(B) = \sum_{i \in I} E(B_i)$ with $\{i \in I \Leftrightarrow E(B_i) \geq 0\}$. This relation is due to the linearity of the operator E and to our hypothesis of no technical constraints in the choice of mineable panels : hence the problem comes down to examining, for a given panel P_i, whether $E(B_i) \geq 0$ or not.

In the practical situation, selection will involve the known values of sample grades $Z_{\alpha j} = z_{\alpha j}$, $j = 1, n$. The a priori R.V. M_i is then replaced by the conditionalized variable

$(M_i / Z_{\alpha j} = z_{\alpha j}, j = 1, n)$ and the profit B_i by :

$(B_i / Z_{\alpha j} = z_{\alpha j} \ldots)$. Defining a selection criterion of panel P_i according to the information $(Z_{\alpha j} = z_{\alpha j}, j = 1, n)$ consists in defining a function $C_i(z_{\alpha 1}, \ldots z_{\alpha n})$ taking only two values : 1, when P_i

is selected, and 0 when P_i is abandoned. So, after selection, profit is

$$(B_i^o \mid Z_{\alpha j} = z_{\alpha j}, \ldots) = C_i(Z_{\alpha 1}, \ldots z_{\alpha n})(M_i \mid Z_{\alpha j} = z_{\alpha j} \ldots - m_c)$$

and noting $m_i^*(z_{\alpha 1}, \ldots z_{\alpha n}) = E(M_i \mid Z_{\alpha j} = z_{\alpha j} \ldots)$, it follows :

$$E(B_i^o \mid Z_{\alpha j} = z_{\alpha j} \ldots) = C_i(z_{\alpha i}, \ldots z_{\alpha n})(m_i^* - m_c)$$

The condition $E(B_i^o \mid Z_{\alpha j} = z_{\alpha j} \ldots) \geq 0$ defines our selection criterium.

$$C_i(z_{\alpha i}, \ldots z_{\alpha n}) = 1 \text{ when } m_i^*(z_{\alpha 1} \ldots z_{\alpha n}) \geq m_c$$

$$= 0 \text{ when } m_i^*(z_{\alpha 1} \ldots z_{\alpha n}) < m_c$$

In another way, $H(u)$ being the step function, $H(u) = 1$ $u \geq 0$, $H(u) = 0$, $u < 0$, we have

$$C_i(z_\alpha \ldots z_{\alpha u}) = H(m_i^* - m_c)$$

Provided that the estimate of each panel grade M_i is

$$m_i^*(z_{\alpha 1}, \ldots z_{\alpha n}) = E(M_i \mid Z_{\alpha j} = z_{\alpha j} \ldots)$$

the selection criterium is identical to the one found in the deterministic case, i.e. selecting each panel, the estimated grade of which is higher than the cut-off grade. Furthermore, as a consequence of the preceding reasoning, we shall estimate the unknown optimal real profit $B_i^o \mid Z_{\alpha j} = z_{\alpha j} \ldots$ with the estimator

$$B_i^* = C_i(z_{\alpha j} \ldots z_{\alpha n})(m_i^* - m_c)$$

which is precisely $E(B_i^o \mid Z_{\alpha j} = z_{\alpha j} \ldots)$.

Conclusion : If we are able to determine the function $m_i^*(z_{\alpha 1}, \ldots z_{\alpha n})$, conditional expectation of M_i, not only can we select optimally the reserves, but we can estimate the optimal recovered profit without bias.

1.3 Practical near-optimal selection

In practical situations, we cannot determine $E(M_i \mid Z_{\alpha j} = z_{\alpha j} \ldots)$

because it requires knowledge of the whole distribution of the multidimensional v.R.V. $(M_i, Z_{\alpha 1}, \ldots Z_{\alpha n})$ which is inaccessible to statistical inference. We shall usually be restricted to a linear estimator m_i^* of the unknown panel grade M_i : $m_i^* = \lambda_i^\alpha Z_\alpha$, so that our selection criterion will only be a function $C_i(m_i^*)$. This means that we reduce our use of the initial set of information $\{z_{\alpha 1}, \ldots z_{\alpha n}\}$ to the single value $m_i^* = \lambda_i^\alpha Z_\alpha$, so that our initial a priori R.V. M_i is replaced by the conditionalized variable $M_i | m_i^*$, and the post-selection profit $B_i^o | m_i^* = C_i(m_i^*)(M_i | m_i^* - m_c)$.

The rest follows as before, noting $\mathcal{M}_i(m_i^*) = E(M_i | m_i^*)$

$$E(B_i^o | m_i^*) = C_i(m_i^*)(\mathcal{M}_i(m_i^*) - m_c)$$

Just as in section 2, the condition $E(B_i^o | m_i^*) \geq 0$ defines our selection criterium as $C_i(m_i^*) = H(\mathcal{M}_i(m_i^*) - m_c)$, together with an unbiased estimator of $B_i^o | m_i^*$:

$$B_i^* = H(\mathcal{M}_i(m_i^*) - m_c).(\mathcal{M}_i(m_i^*) - m_c)$$

Let us now consider the main difference from the results of section 2 :

- The optimum derived from this new criterion is necessarily smaller, in expected value, than the optimum achieved by cutting on conditional expectation : the reason is that we actually have optimized the selection on the R.V. $M_i | m_i^*$ which is only an "approximation" of the R.V. that we have optimized before, i.e. $M_i | Z_{\alpha 1} = z_{\alpha i} \ldots$

- The cut-off procedure is to be applied, not on the estimate m_i^*, but on the function $\mathcal{M}_i(m_i^*)$, which must also be used to estimate the recovered profit without bias.

From these remarks, we conclude that the "estimate" m_i^* arises only as an element of conditionalization. The true estimator used for the cut-off procedure and for the estimator B_i^* is actually $\mathcal{M}_i(m_i^*)$, and an optimal selection will be performed only if we are able to determine it. The amount of statistical information required for the determination of $\mathcal{M}(m_i^*)$ is a little smaller than for the determination of $E(M_i Z_{\alpha j} \ldots)$. Actually we need a

the use of a scattered estimate will improve only the "apparent" estimated profit B_i^*, while probably reducing the true profit B_i : we finally obtain a lower, and, even worse, a grossly overestimated profit.

1.4 Influence of the quantity of information

Recalling the formulation of the optimal selection seen in section 2, we may say that B_i^o being the post-selection profit, we find a function $m_i^*(z_{\alpha 1}, \ldots z_{\alpha n})$ in the set of n variable, measurable functions such that by following the selection procedure, we optimize $E(B_i^o)$. Now if we suppose that the available information is only a subset of $\{Z_{\alpha j}, j = 1, n\}$, we shall find our optimizing function $m_i^{*'}(z_{\alpha k}, k = 1, n')$ in the set of n'-variable functions, a subset of the preceding one, hence obtaining a lower expected profit. This is a well-known property of any optimizing procedure.

Though this result is only exact for the true optimizing procedure, we may expect to find it also when optimizing with linear estimators, for instance with kriging estimators : we shall see an experimental example of this property in part II.

2. THE SIMULATED OREBODY AND THE ESTIMATORS

2.1 Simulated orebody.

We have simulated a two-dimensional distribution of point grades, by the classical "Turning band method". The simulated orebody is divided into 500 square panels of 22 x 22 m., on 10 lines of 50 panels. In each panel, we have simulated a 11 x 11 points grid, the mean value of which was taken as the panel's mean ore grade. Consider the central value of each panel as sampled : so the exploration campaign consists of a regular 22 x 22 m. grid with 500 samples. In Fig. 1 and 2 we show the experimental variograms and histograms of samples. The experiment consisted in comparing, for different estimators, the profit yielded by selecting panels with a cut-off criterion applied to the estimated panel's mean ore grade. To avoid the computation of special estimators at the border of the orebody, we made the comparison on the 48 x 8 interior blocks of the simulated orebody.
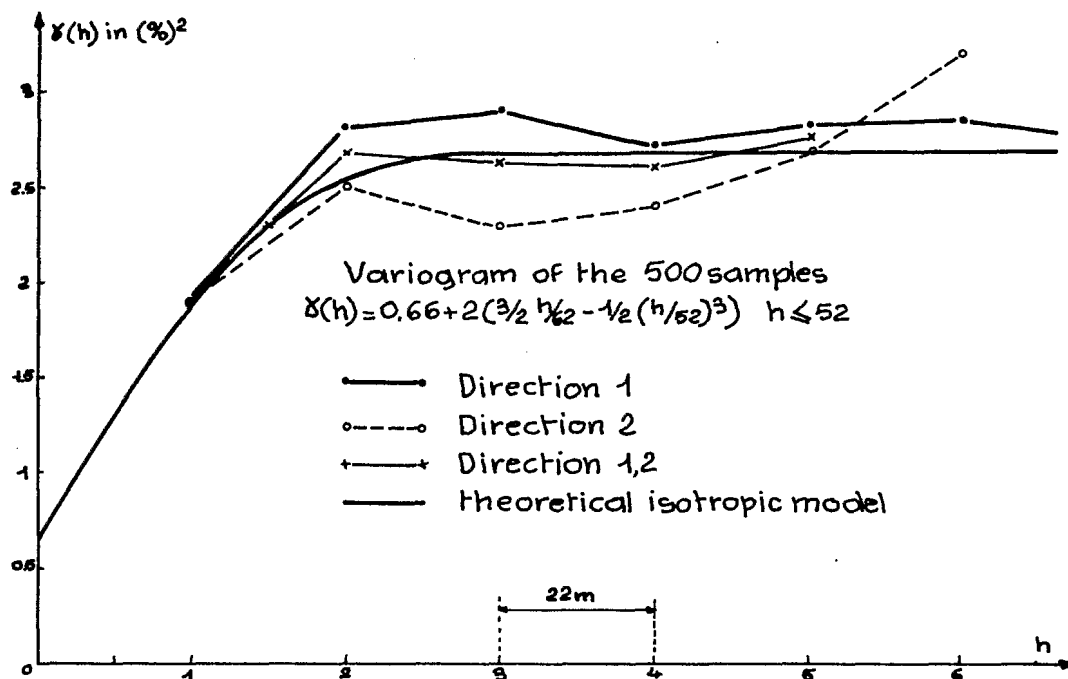
$\gamma(h)$ in $(\%)^2$

Variogram of the 500 samples
$\gamma(h)=0.66+2(\frac{3}{2} \frac{h}{52}-\frac{1}{2}(\frac{h}{52})^3)$ $h \leqslant 52$

•——• Direction 1
○----○ Direction 2
+——+ Direction 1,2
—— theoretical isotropic model

22m

h

Fig. 1

Frequency in %

Histogram of samples
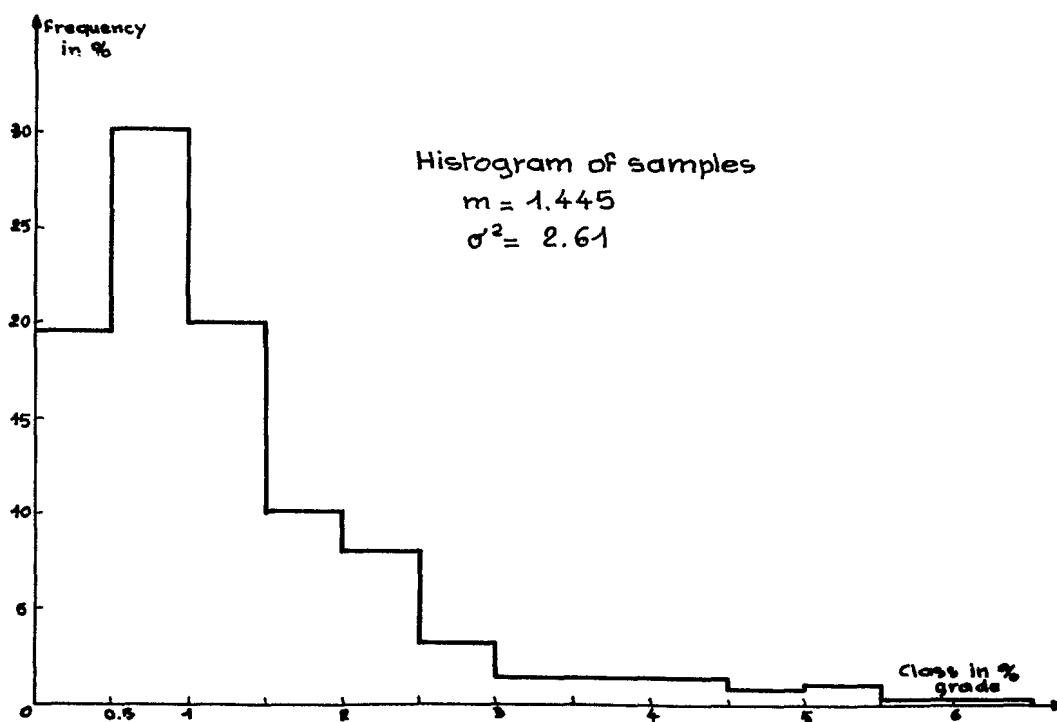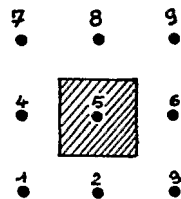m = 1.445
$\sigma^2 = 2.61$

Class in % grade

Fig. 2

2.2 The estimators

For each 22 x 22 m. panel's grade, we build four estimators :

- a disjunctive kriging estimator
- a kriging estimator
- a zone of influence estimator
- a kriging estimator using only the information of a 44 x 44 m. grid.

The first three estimators using the 22 x 22 m. information are computed, for each panel P_i, with the 9 nearest samples and the overall mean of the 500 samples.

With the last kriging estimate, we estimate the same panels as before, with the four nearest samples and the overall mean of the 125 samples. We note that this compels us to distinguish three different position of panels.
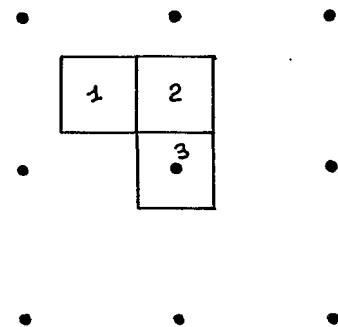
2.2.1. Kriging with 22 x 22m. grid.

$$Z_V^* = m^* + \lambda^1 (\overline{Z}_1 - m^*) + \lambda^2 (\overline{Z}_2 - m^*) + \lambda^3 (\overline{Z}_3 - m^*)$$

m^* = arithmetic mean of the 500 samples

$\overline{Z}_1$ = value of sample number 5

$\overline{Z}_2$ = mean value of samples n° 2, 4, 6, 8

$\overline{Z}_3$ = mean value of samples n° 1, 3, 7, 9

Weights λ^1, λ^2, λ^3 are calculated with the kriging system of residuals, with the model of variogram appearing in Fig. 1 :

$$\lambda^1 = 0.416 \; ; \; \lambda^2 = 0.110 \; ; \; \lambda^3 = 0.033 \quad \sigma_K^2 = 0.327$$

2.2.2 Disjunctive kriging with the 22 x 22 m. grid. The point-based variable $Z(x)$ is considered as an anamorphosis of a Gaussian R.F. $Y(x)$, i.e. $Z = \varphi(Y)$, and this function φ is expressed as a development in term of normalized Hermite polynomial η_n. $Z = \sum_{p=o}^{n} f_o^p \eta_{\bar{p}}(Y)$. In our study, we tried $\varphi(Y) = e^Y$, so that we made a structural study of variable $Y = \log Z$ on the 500 samples (see Fig. 3 and 4). We recall that the D.K. estimator is $Z_V^{DK} = \sum_{\alpha} f_\alpha(Z_\alpha)$ where functions f_α are developed, which gives $Z_V^{DK} = f_o^o + \sum_{\alpha} (\sum_{p=1}^{n} f_\alpha^p \eta_p(Z_\alpha))$ where coefficients f_α^p are computed in the fol-lowing systems :

for $p = 1,\ldots n$ $\qquad \sum_{\alpha} f_\alpha^p \rho_{\alpha\beta}^p = f_o^p \overline{\rho_{\beta v}^p}$ $\qquad \alpha,\beta = 1,N$

$$\overline{\rho_{\beta v}^p} = \frac{1}{v} \int \rho^p (x-x\beta)\, dx \quad , \quad \rho_{\alpha\beta}^p = \left[\rho(x\alpha - x\beta) \right]^p$$

In our problem, the function e^Y was developed up to degree 5, the samples $x\alpha$ were the 9 points surrounding the panel and covariance $\rho(h)$ is the normalized one corresponding to the variogram of logarithms. (See fig. 3) . In Z_V^{DK}, experimental values intervene in terms such as $\eta_p(Y_\alpha)$, so that we must use each sample in an individual manner : however, it remains clear that for geometrical reasons, we shall have the following identity :

$$p = 1,5 \qquad f_2^p = f_4^p = f_6^p = f_8^p$$

$$p = 1,5 \qquad f_1^p = f_3^p = f_5^p = f_9^p$$

p/α	α = 5	α = 2	α = 1
1	- 0.3725	- 0.0992	- 0.0265
2	0.1977	0.0456	0.0111
3	- 0.0745	- 0.0144	- 0.00346
4	0.0223	0.00357	0.00077
5	- 0.0062	0.00075	- 0.00014

In conclusion, the whole D.K. estimator is determined by the resolution of 5 3x3 linear systems. We also note that our D.K. estimator does not use any permanence of law hypothesis.
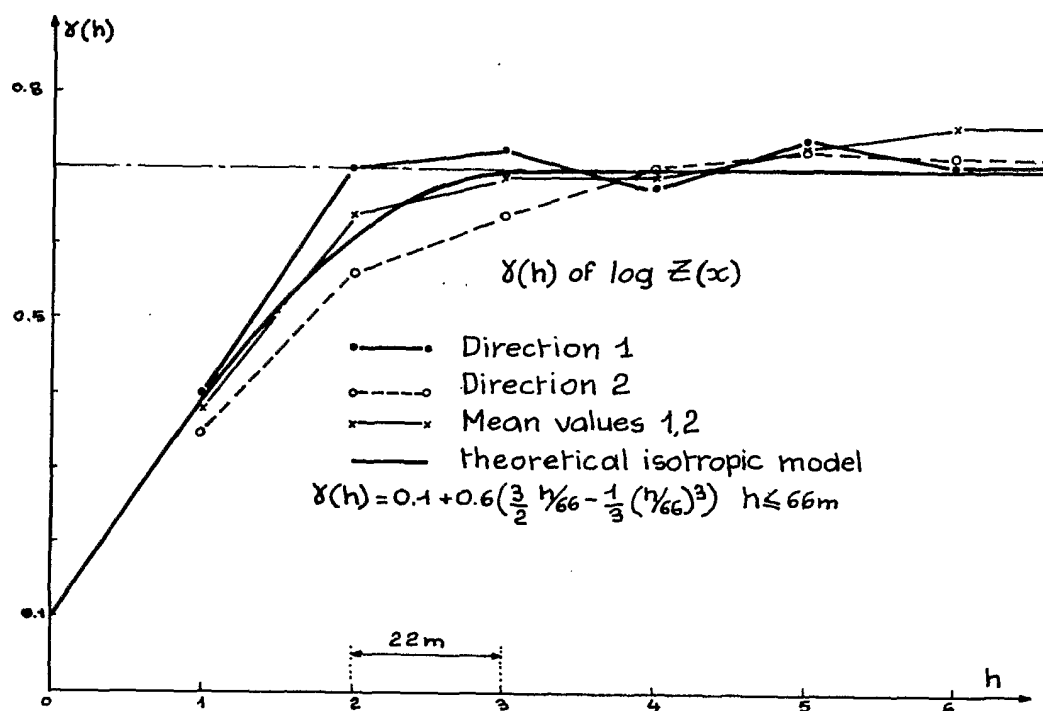
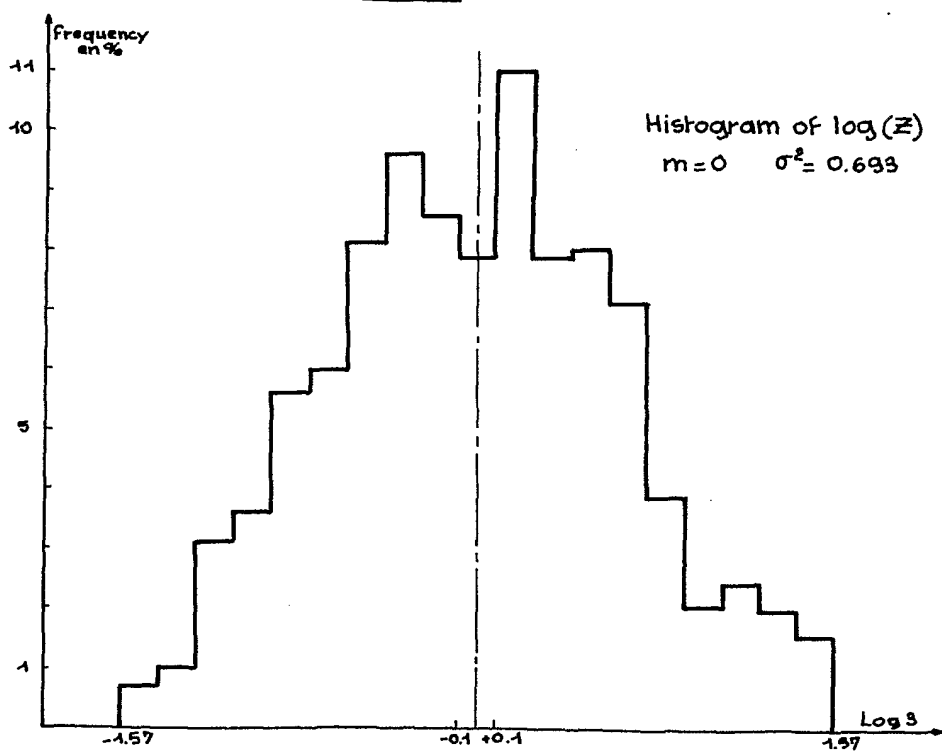$\gamma(h)$ of $\log \bar{Z}(x)$

•———•   Direction 1

o- - - -o   Direction 2

×———×   Mean values 1,2

————   Theoretical isotropic model

$\gamma(h) = 0.1 + 0.6 \left(\frac{3}{2} \frac{h}{66} - \frac{1}{3} \left(\frac{h}{66} \right)^3 \right)$ $h \leq 66m$

22m

Fig. 3

Histogram of $\log(\bar{Z})$

$m = 0 \qquad \sigma^2 = 0.693$

Fig. 4

2.3 Polygonial zone-of-influence estimator

According to the method of influence zone, each panel's estimator is the central sample value : in our case, the direct consequence is that, for estimating our ·48 x 8 = 384 panels, we shall use only 384 of our available 500 samples.

2.4 Kriging with the 44 x 44 m. grid

As shown in the preceding figure, we suppose that we know only 125 samples on a 44 x 44 m. grid, with which we estimate the same 384 22 x 22 m. panels as before. As previously noted, we have to determine 3 different kriging configurations, according to the relative position of the samples and the panel to be estimated.

3. EXPERIMENTAL RESULTS

We shall present the experimental results in order to point out the properties of the different estimators with respect to three problems :
- for a given amount of information, quality of the estimators with respect to the magnitude of the level of estimation.
- for a given amount of information, quality of the selection obtained by applying a cut-off grade to the estimators.
- quality of the selection of the kriging estimator when the amount of information changes.

3.1 Errors of estimation

On the following table, we see the results of the three estimators, computed with the 22 x 22 m. grid information :

D.K. : Disjunctive Kriging.

K.1. : Kriging

POLY : Central value of panels.

	$m(\varepsilon)$	$D^2(\varepsilon)$	σ_E^2	m^*	$D^2(v^*\|V^*)$	$D^2(v^*\|V^*)+D^2(\varepsilon)$
D.K.	0.040	0.323	0.29	1.424	0.879	1.201
K.1.	0.030	0.380	0.327	1.434	0.930	1.310
POLY	0.039	1.128	0.98	1.425	2.585	3.713

TABLE 1

$m(\varepsilon)$, $D^2(\varepsilon)$: experimental mean and variance of errors

σ_E^2 : estimation variance of each panel's grade

m^*, $D^2(v^*|V^*)$: experimental mean and variance of estimated values.

We note that :

- with respect to the global estimation, the three estimators are nearly equivalent. The variance of estimation of the mean value of the 384 panels (exact value 1.464) computed with the variogram model of Fig. 1 is $\sigma_m^2 = 0.002$, the corresponding standard deviation of which is $\sigma_m = 0.045$. The error in the global estimation is thus smaller than the theoretical standard deviation of error.

- with respect to the variance of error in estimating each panel, the estimators are ranking as forecast, though the observed variance of errors is slightly higher than the theoretical variances of estimation.

- with respect to the variance of dispersion of estimates $D^2(v^*|V^*)$, we see that, although being better, D.K. estimates are less scattered than K.1 and POLY.

- The last value $D^2(v^*|V^*) + D^2(\varepsilon)$ was computed to recall that the theoretical smoothing relation, $D^2(v|V) = D^2(v^*|V^*) + D^2(\varepsilon)$, is only valid when v^* is calculated by kriging. The theoretical value $D^2(v|V)$ being 1.34, we actually see that only the kriging estimates follow the relation.

- the histogram of kriging errors is fairly different from a Gaussian one (Fig. 5) : it is not truly symmetric and, above all, largely more skewed in the central classes. The histogram of errors from other estimators have a similar shape.

3.2 Optimization of the selection.

For each estimate, selection is achieved by cut-off on the estimated values. With the selected values, we estimate profit B^*, which is $B^* = \sum_{i \in I} (m_i^* - m_c)$ or, defining as $T(m_c)$ the tonnage after selection and as $m^*(m_c)$ its estimated mean ore grade, we find :

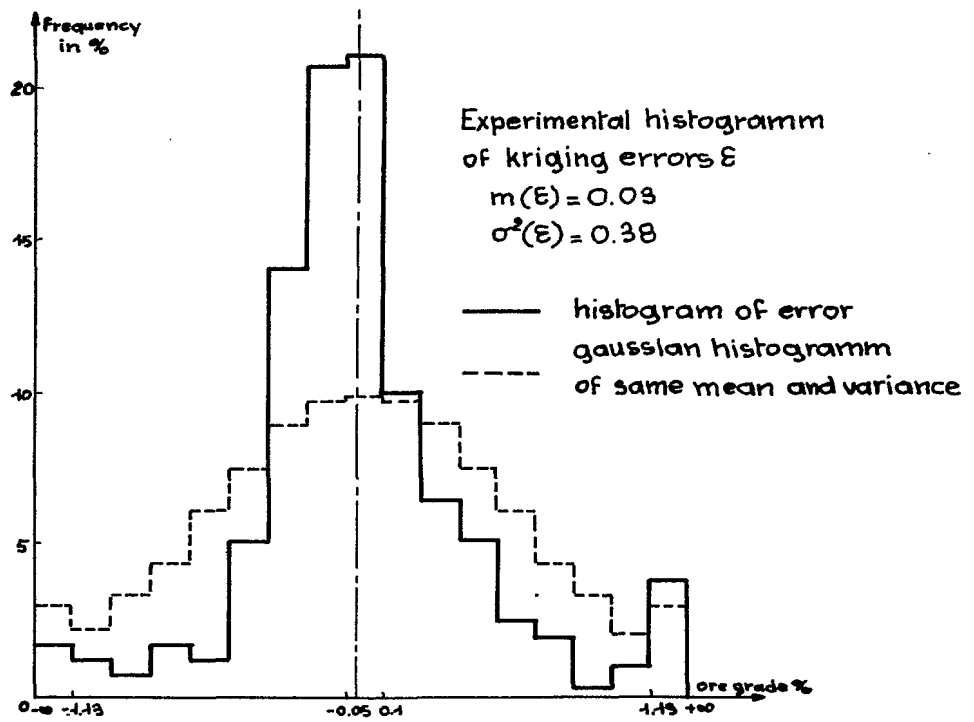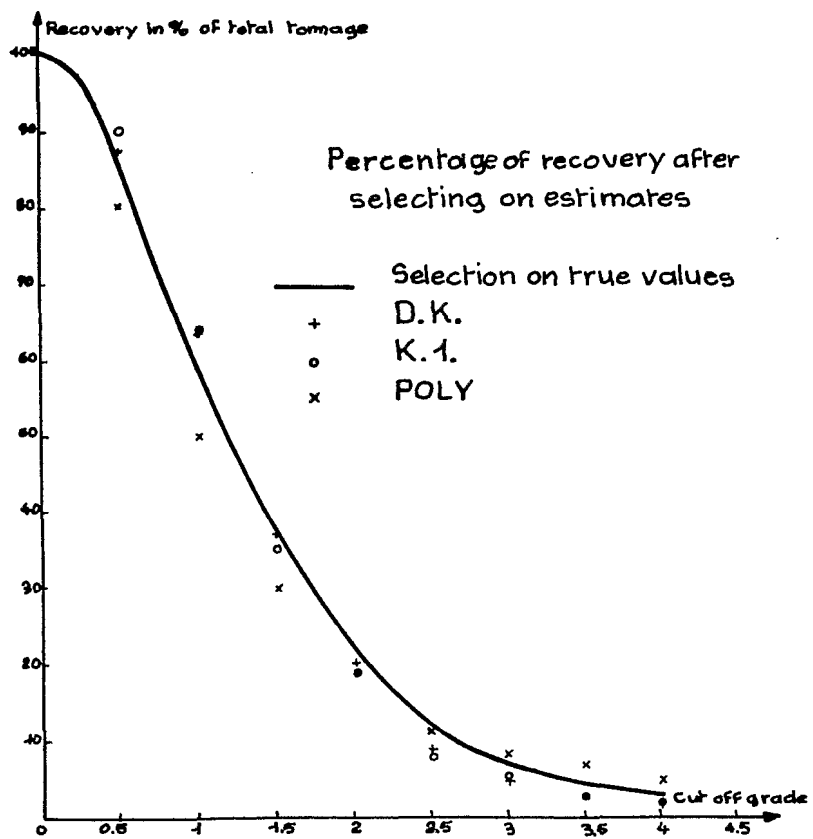$$B^* = T(m_c) \left[m^*(m_c) - m_c \right]$$

Fig. 5

Fig. 6

For each selection, we know the actual mean ore grade $m_v(m_c)$ corresponding to the selected tonnage $T(m_c)$, so that we can compare the estimated profit B^* to the actual recovered profit $B = T(m_c) \left[m_v(m_c) - m_c \right]$. I remind the reader that the aim of this comparison is to check whether or not the selectionwith a given estimate is close to the optimal selection performed with the conditional expectation estimator : as this estimator is actually complicated, we shall compare our results with the D.K. estimator which is very close to the conditional expectation. We can see on Fig. 7 that the D.K. estimator verifies very closely the fundamental property of conditional expectation : the mean of the true values of a block having a given estimation is equal to this estimation. (All the numerical results on Figures 10 to 14 appear on Table 2).
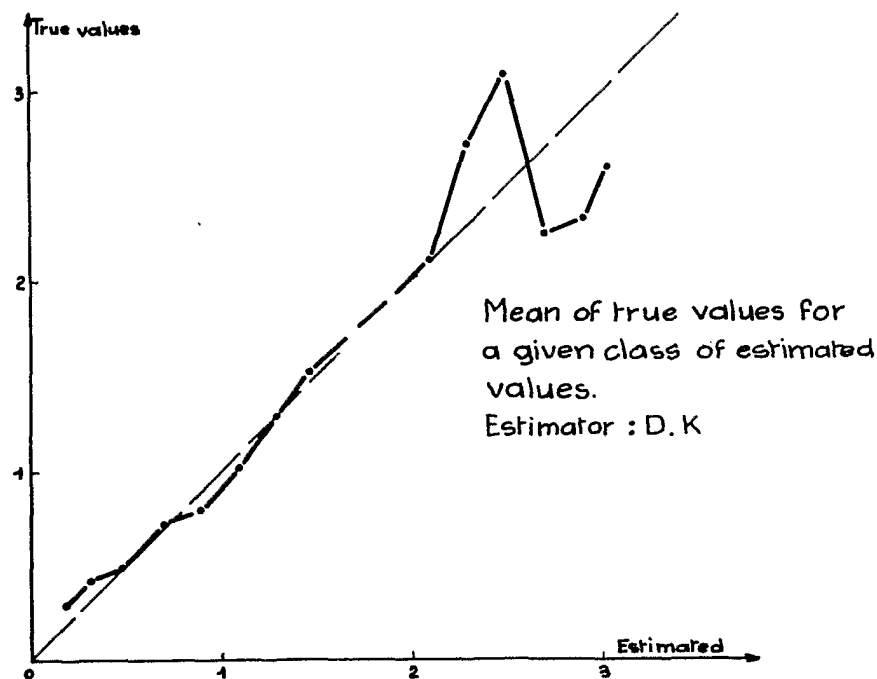

Mean of true values for a given class of estimated values.
Estimator : D.K

Fig. 7

3.2.1 <u>Actual profit recovered</u>. The largest profit will be re-covered by selection on estimates having the same property as con-ditional expectation, i.e. the same tonnage-cut-off curve and the same property of conditional unbiasedness. We see on Fig. 6 that the D.K. and K.1 tonnage curves are very close, while the POLY one is quite distinct. In the same way, Fig. 8 shows that kriging is "nearly" conditionally unbiased, while in Fig. 9 we see a notice-able discrepancy between the experimental curve and the unit slope line.

The consequence, appearing in Fig. 10, is that selection on kriging estimates gives a near optimal result, whatever the cut-off grade may be, while selection on zone-of-influence estimates induces a loss of profit which is especially important for high cut-off.

Therefore we may conclude that kriging is a near-optimal tool for any cut-off grade. Conversely, the polygon method is systema-tically non-optimal and so leads to waste of reserves.

3.2.2 <u>Estimation of recovered profit</u>. We have seen that the pro-perty of conditional unbiasedness implies that an estimator gives an unbiased estimation of the <u>selected</u> reserves.

$$B^* = T(m_c) \left[m^*(m_c) - m_c \right] \text{ while } B = T(m_c) \left[m_V(m_c) - m_c \right]$$

Thus - if it exists - the bias will appear to be proportional to $m_V - m^*$, difference between the true and estimated mean grade of the selected reserve. We see in Fig. 11, 12 and 13 the experi-mental curves $m_V(m^*)$ corresponding to different cut-off grades and for estimators D.K., K.1 and POLY respectively. For each esti-mator, the bias will be proportional to the deviation between $m_V(m^*)$ and the unit slope line.. We see that the bias is small and non systematic for D.K. and K.1, corresponding to statistical fluc-tuations, while the bias of estimation from POLY is systematic : not only will a selection from POLY give a loss of profit (compared with the optimum one) but furthermore the estimation of the profit will be dangerously overvalued.

3.2.3 <u>Influence of the quantity of information</u>. In the first section, we have recalled the fact that the expected optimum profit was a non decreasing function of the quantity of information : we have just seen experimentally that kriging was a fairly good approxi-mation of conditional expectation, so that in practice it will re-place it. We shall see on the following results (Fig. 14 and Table 2) the true benefit resulting from a selection on kriging estimates,
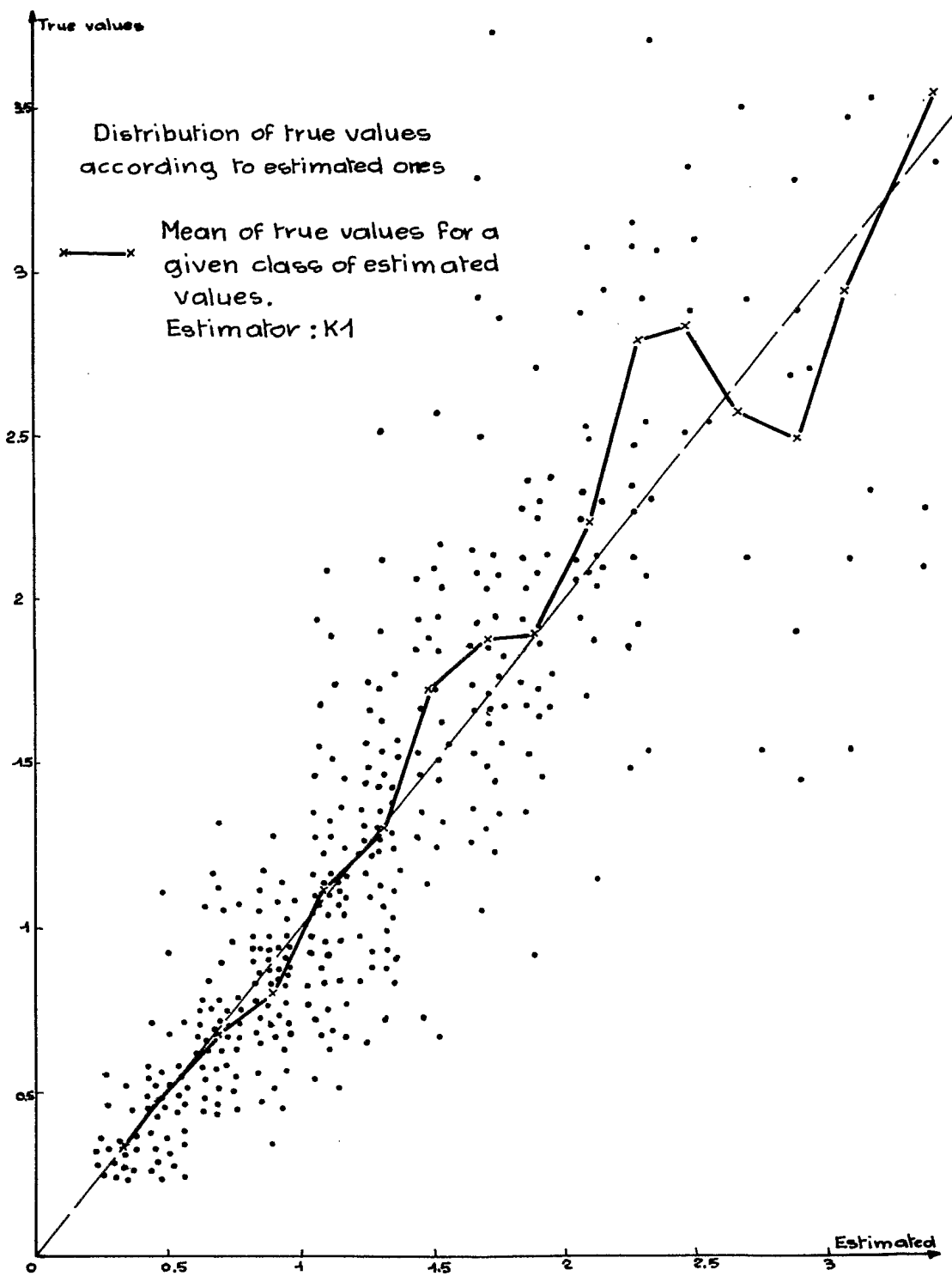
Fig. 8

K.1 being built with the 22 x 22 m. information, K.2 with the 44
x 44 m. information : we note the degrading of the quality of the
selection and by comparison with Fig. 13, we see that this degra-
ding is greater between K.1 and K.2 than between K.1 and POLY.
Although it is the best linear estimator,kriging will give a low
quality selection if the information is not sufficient (of course,
using polygon estimation would be even worse), and this property
gives a direct "monetary" value to mining information. As a secon-
dary effect of the deficiency of information, we also note in Table
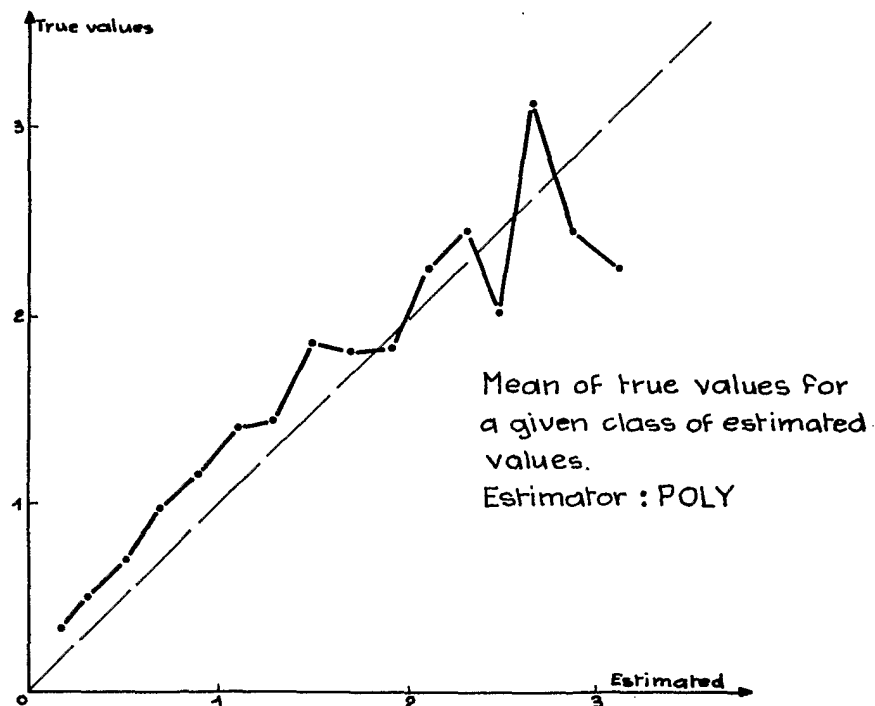2, that the lower profit is, moreover, ill-estimated.

Mean of true values for
a given class of estimated
values.
Estimator : POLY

Fig. 9

4. CONCLUSION

    The preceding experimental results were obtained from a simu-
lated orebody of lognormal type and fairly large relative disper-
sion : $\frac{\sigma^2}{M^2}$ = 1.25. Such a random function is very different from
a Gaussian one, for which kriging is theoretically the optimal so-
lution. The first important experimental conclusion is that kriging
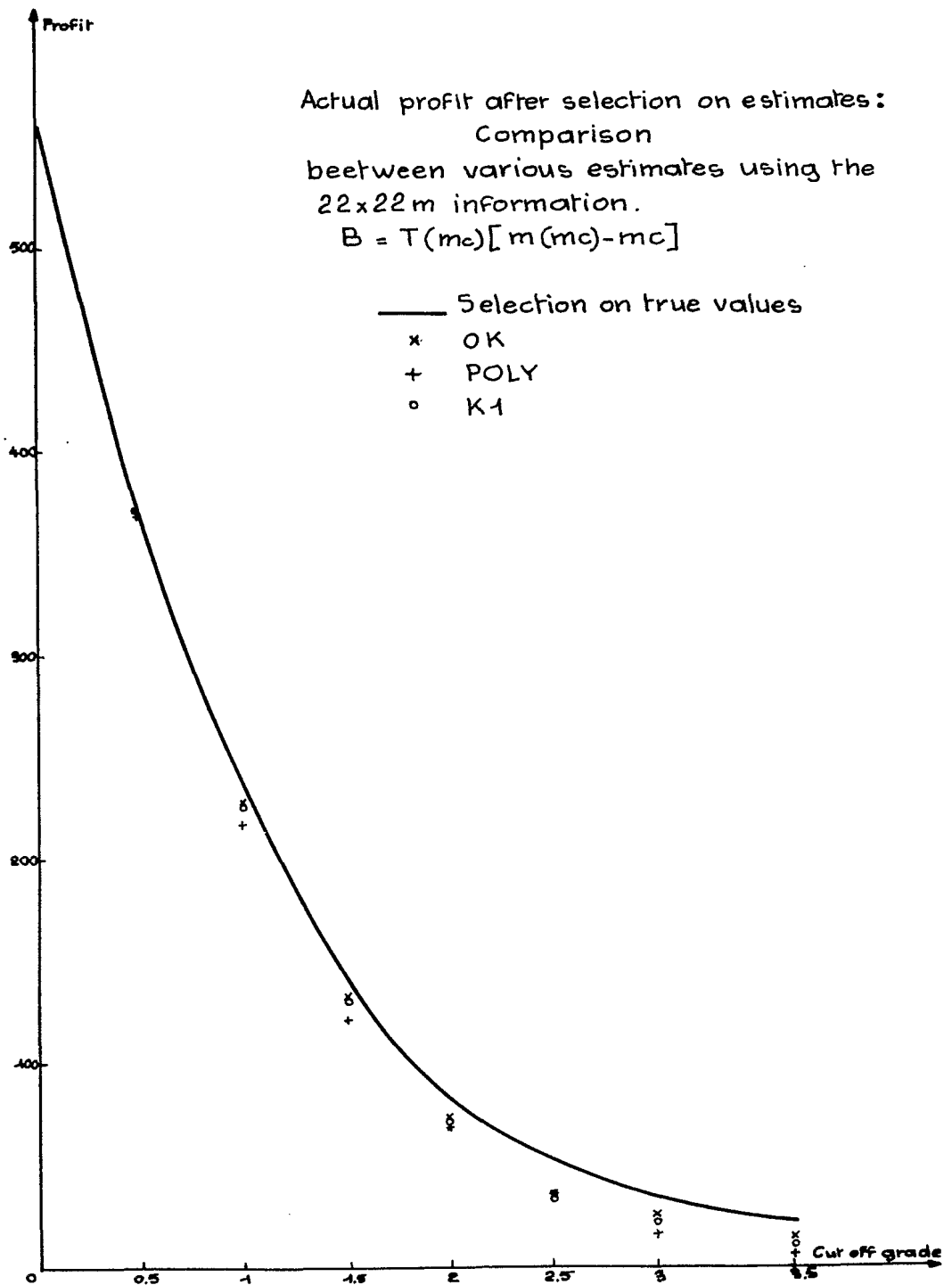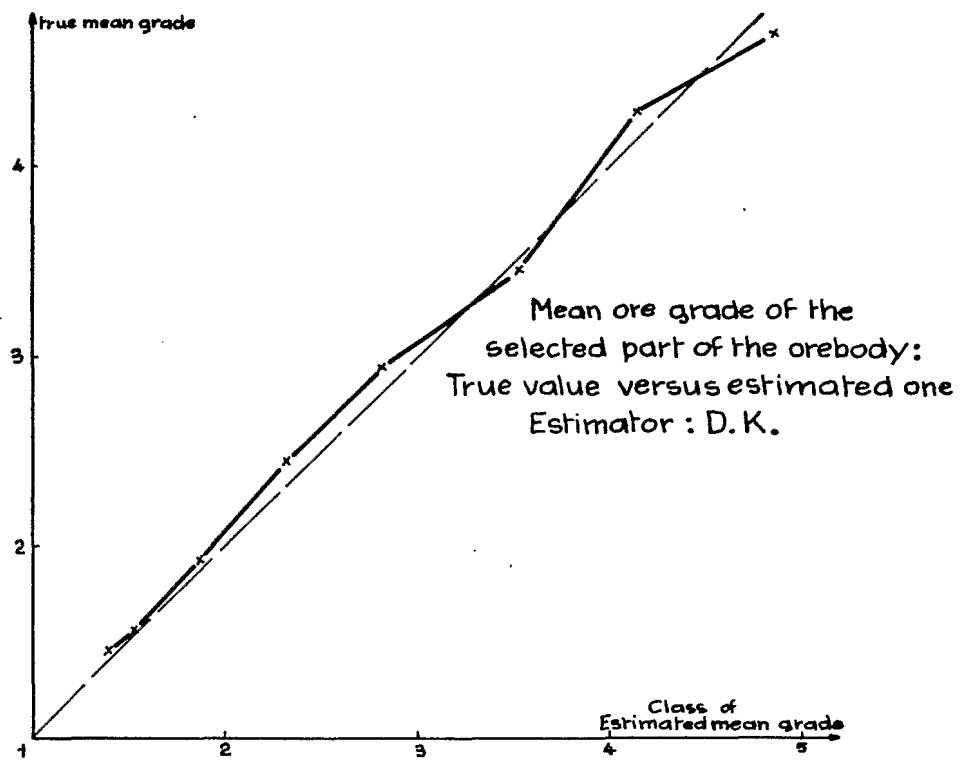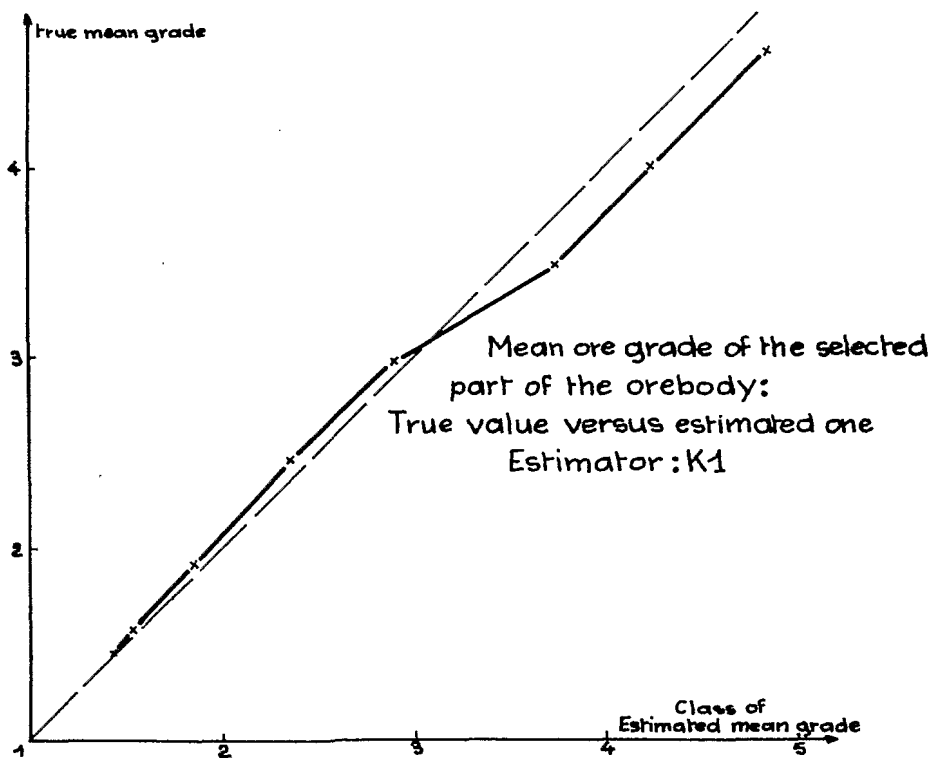
Fig. 10

Fig. 11

Fig. 12

is a robust approximation of conditional expectation and keeps its important property (conditional unbiasedness) even for a very particular R.F. The second conclusion is drawn from the comparison between D.K., K.1, POLY and K.2 concerning the actual recovered profit : we have already noted that the results of K.2 were worse than the results of POLY, although the experimental estimation variance of the panels were 1.13 and 0.65 respectively. This emphasizes the fact that the quality of selection is not entirely related to the estimation variance of each panel, but is largely function of the overall estimation variance of the whole orebody. Lastly, the drawbacks of an imprecise estimator seem to be the same, whatever the reasons of the lack of precision (i.e. wrong estimator or lack of information) :

- loss of actual recovered profit.

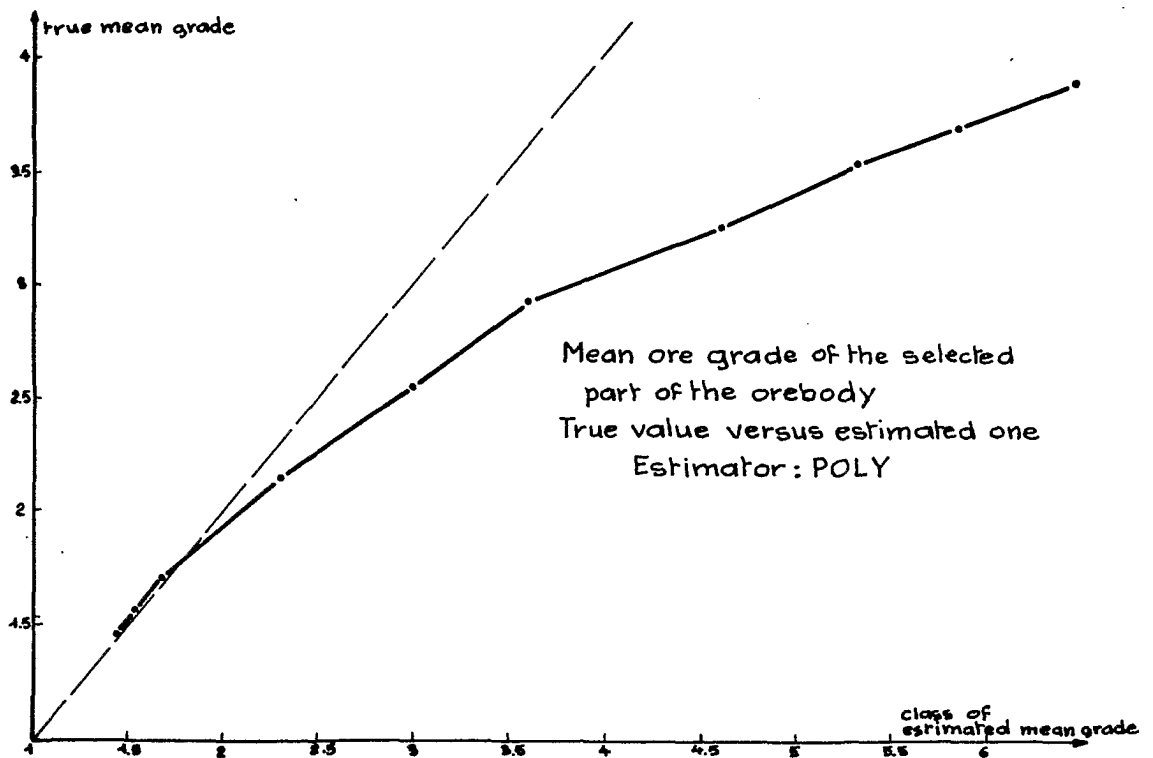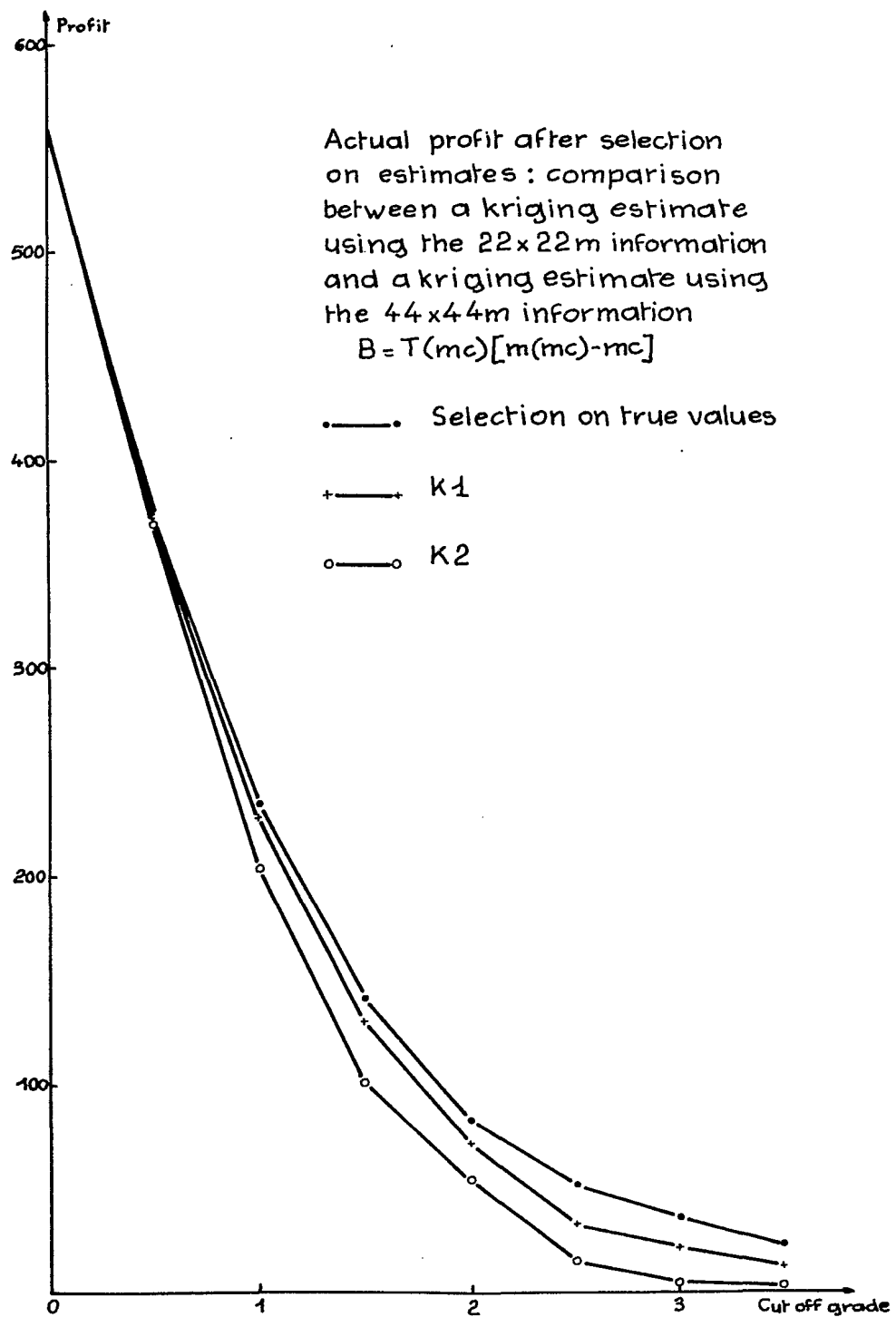- poor estimation of recovered profit.

Mean ore grade of the selected part of the orebody
True value versus estimated one
Estimator : POLY

Fig. 13

Fig. 14

TRUE VALUES

m_c	T	m_V	B
.0	384	1.464	562.2
0.5	338	1.611	375.5
1	225	2.049	236.
1.5	148	2.467	143.1
2	90	2.948	85.3
2.5	44	3.726	53.9
3	29	4.211	35
3.5	17	4.915	24.1
4	11	5.507	16.6

T : Tonnage of Selected panels

m_c : Cut-off grade

m_V^* : estimated mean grade of selected panels

m_V : true mean grade of selected panels

B^* : estimated profit of selected orebody

B : true profit of selected orebody

D.K.

m_c	T	m_V^*	m_V	B^*	B
0	384	1.424	1.464	546.8	562.2
0.5	337	1.577	1.608	362.9	373.4
1	246	1.875	1.927	215.3	228
1.5	142	2.329	2.430	117.7	132.1
2	76	2.817	2.949	62.1	72.1
2.5	34	3.539	3.459	35.3	32.6
3	19	4.143	4.293	21.7	24.6
3.5	11	4.782	4.701	14.1	13.2
4	6	5.67	5.56	10	9.4

POLY

m_c	T	m_V^*	m_V	B^*	B
0	384	1.425	1.464	547	562.2
0.5	309	1.688	1.69	367.1	367.7
1	192	2.287	2.135	247.1	216
1.5	115	2.983	2.549	170.5	120.5
2	73	3.624	2.934	118.6	68.2
2.5	45	4.583	3.261	93.7	34.2
3	32	5.329	3.535	74.5	17.1
3.5	26	5.831	3.668	60.6	4.4
4	20	6.479	3.919	49.6	-1.6

K.1

m_c	T	m_V^*	m_V	B^*	B
0	384	1.433	1.464	550.3	562.2
0.5	348	1.54	1.574	361.9	373.8
1	248	1.852	1.916	211.3	227.2
1.5	136	2.358	2.455	116.7	129.9
2	72	2.901	2.983	64.9	70.8
2.5	32	3.746	3.492	39.9	31.7
3	21	4.245	3.999	26.1	21
3.5	12	5.026	4.59	18.3	13.1
4	8	5.646	4.375	13.2	3

K.2

m_c	T	m_V^*	m_V	B^*	B
0	384	1.418	1.464	544.5	562.2
0.5	384	1.418	1.464	352.5	370.2
1	296	1.582	1.694	172.3	205
1.5	142	1.949	2.234	63.8	104
2	48	2.379	3.154	18.2	55.4
2.5	15	2.726	3.596	3.4	16.4
3.	2	3.40	5.32	0.8	4.6
3.5	1	3.63	7.93	0.1	4.4
4	0				

GEOSTATISTICS FOR

CONDITIONAL SIMULATION OF OREBODIES

by A.C. Journel

A B S T R A C T

Simulation techniques are frequently used to solve various problems of Operational Research in the fields of mining industry and more generally of earth sciences (hydrogeology, gravimetry, meteorology....). First, the model to be simulated is characterized, for example the spatial dispersion of grades in an orebody. Then a simulation technique is elaborated, which must be operational particularly in terms of computer time. The efficiency of the simulation produced is obviously linked to the capacity of the model to fit the main characteristics of the revealed reality. Now, one of the most important of these characteristics, namely the spatial auto-correlation of variables, is often ignored by the models commonly presented in classical literature.

The originality of the Conditional Simulation derives :

- First, from the fact that these simulations meet the particular spatial auto-correlation function (covariance or variogram) which characterizes the reality observed.

- Second, from the conditionalization of the experimental data : the simulated values at data locations are equal to the experimental values.

- Third, from the possibility to work in the real 3-dimensional space. The simulation technique proposed (turning bands method) consists in simulating on lines (1-dimensional space), then in turning these lines in the 3-dimensional space. This procedure avoids the well-known explosion of computer time and memories involved by classical procedures when extended to several dimensions spaces. This last point represents the very originality of Conditional Simulation techniques with regard to Spectral Analysis techniques

I - AIMS OF A SIMULATION

Before a mine starts, it is often important to predict the fluctuations at various scales (day, month, year) of the various characteristics of the recovered ore. The daily fluctuations of the overburden ratio or the mineralized thickness may condition the extraction procedure or the choice of the mining tool. The fluctuations at various scales of the grades may condition the existence of a blending station or the flexibility of the mill.

A perfect knowledge of the recovered ore would be required to study these dispersions, and more precisely :

- if $z_0(x_1, x_2, x_3) = z_0(x)$ denote the true grade value at any location $x \in R^3$ of the 3-dimension space with coordinates x_1, x_2, x_3, one would require the knowledge of the real function $z_0(x)$. For lack of such knowledge, a simulation $z_c(x)$ of that reality is needed.

Information about the real surface consists of :

1 - local knowledge of values $z_0(x_i)$ at data locations $x_i \in I$.

2 - structural knowledge of the spatial auto-correlations between data $z_0(x_i)$, $z_0(x_j)$. These spatial auto-correlations characterize the "fluctuations pattern", i.e. the degree of regulatiry of the real surface $z_0(x)$.

Hence, the simulated surface $z_c(x)$ must meet this information, i.e. :

1 - at data locations $x_i \in I$, simulated values and experimental values must be equal : $z_c(x_i) = z_0(x_i)$.

2 - the simulated surface must present the same "fluctuation pattern" that the real surface.

When the real surface $z_0(x)$ is known only at a limited number of locations $x_i \in I$, the simulated surface can be known on almost every point x of the deposit. It is then possible to apply to that

simulation the various processes of extraction , hauling, stock-
piling, etc..., to study their technical or economical consequences
and by feed-back correct these processes.

This idea of simulating orebodies is not new. Disposing of
high performance computers, many authors have proposed various deter-
ministic or probabilistic simulations of the mine reality (see in
summary [7], J.W. Harbaugh, 1970 - [14], Y.C. KIM, 1973). But most
of these "classical" simulation procedures fail to reproduce the
most important characteristic of reality, namely the spatial auto-
correlations of variables. Variables are auto-correlated in all de-
posits ; these auto-correlations can be characterized by covariances,
correlograms, or, still better, by variograms. Geostatistics shows
that fluctuations of grades, thicknesses or any other variables are
directly linked to the particular auto-correlation function of the
variable considered. Hence, it is of prime importance that the simu-
lation $z_c(x)$ meets this very particular auto-correlation function.

II - THE PROBABILISTIC REPRESENTATION OF REALITY

Before exposing the Conditional Simulation procedure, let's
recall how Geostatistics characterizes the spatial display of a va-
riable $z_0(x)$ (see also [1], G. Matheron, p. 50).

Consider a 3-dimensional deposit D surveyed by drill-holes and
cores of equal section and length. Let $x = (x_1, x_2, x_3) \in R^3$ be a point
of the 3-dimensional space and x_1, x_2, x_3 be its coordinates. Let
$z_0(x) = z_0(x_1, x_2, x_3)$ be the grade of the core which gravity centre
is located at point x. The dimensions of the core are generally
small enough with regard to the extension of deposit D ; we will
then consider the grade $z_0(x)$ as a punctual grade.

The deposit D is constituted of an infinity of punctual grades
$\{z_0(x), x \in D\}$, some of them being known at data locations
$\{z_0(x_i), x_i \in I\}$. Consider any line within the deposit, the true

curve $z_0(x)$ is a saw-tooth like curve which however presents structural characteristics (see Fig. 1) :

- There exist poor zones and rich zones.

- 2 grades $z_0(x)$ and $z_0(x+h)$ distant of h are on average more similar as their distance h decreases.

The probabilistic representation of $z_0(x)$ will take these structural characteristics into account :

Locally at each point x_1 the true value $z_0(x_1)$ is interpreted as a particular realization of a random variable $Z(x_1)$. The two random variables $Z(x_1)$ and $Z(x_1+h)$ are correlated, that is, their two realizations $z_0(x_1)$ and $z_0(x_1+h)$ are not independent. Now, the whole reality $\{z_0(x), x \in D\}$ is interpreted as a particular realization of a random function $Z(x)$, this random function can be seen as the set of an infinity of correlated random variables located at each point x : $\{Z(x), x \in D\}$.

Consider the first two moments of the random function $Z(x)$:

- moment of order 1 or expected value : $m(x) = E\{Z(x)\}$

- moment of order 2 or covariance :

$$C(x,h) = E\{Z(x).Z(x+h)\} - m(x)\ m(x+h)$$

All throughout this article, we will suppose the hypothesis of local stationarity of order 2, that is, for distances h inferior to a certain limit, the two previous moments are independent of the location x.

$E\{Z(x)\} = m$ constant whatever x may be. For mere reasons of writing simplicity, we will consider the constant m as nought. Hence :
$E\{Z(x)\ Z(x+h)\} = C(h)$ whatever x may be.

In practice, there is always a scale within which reality meets that hypothesis of local stationarity.

Geostatistics uses another moment of order 2, the variogram :

$2\gamma(h) = E\{[Z(x) - Z(x+h)]^2\}$ which requires a weaker hypothesis : local

stationarity of increments $[Z(x)-Z(x+h)]$. When the covariance exists, both tools - covariance and variogram - are identical and correspond to each other by the relation

(1)
$$\boxed{\gamma(h) = C(o) - C(h)}$$

The covariance or the variogram appears as the sought after auto-correlation function that characterizes the "fluctuations pattern" or degree of regularity of the true surface $\{z_o(x), x \in D\}$. $2\gamma(h)$, for example, is estimated from the experimental data :

$$2 \gamma^*(h) = \frac{1}{N} \sum_{i=1}^{N} \{z_o(x_i) - z_o(x_i+h)\}^2$$

if there are N couples of data distant of h.

Let's recall now how that simple variogram tool can characterize the "fluctuations pattern" of reality.

Variances of Dispersion - (see also [1], G. Matheron, p.66).

The dispersion of true punctual values $z_o(x)$ when x describes an area $V \subset D$ included in the deposit is characterized by a dispersion-variance which depends only on the variogram :

$$D^2(0/V) = \frac{1}{V^2} \int_V dx \int_V \gamma(x-y)dy = \overline{\gamma}(V,V)$$

$\overline{\gamma}(V,V)$ denoting the mean value of $\gamma(h)$ when the two extremities M and M' of vector $\vec{h} = \overrightarrow{MM'}$ describe independently the volume V. An estimator $D^{2*}(0/V)$ of that variance is the a priori variance of data $\{z_o(x_i), x_i \in V\}$.

$$D^{2*}(0/V) = \frac{1}{n(n-1)} \sum_{i<j} [z_o(x_i)-z_o(x_j)]^2$$

if there are n data $z_o(x_i)$, $z_o(x_j)$,... within area V.

Similarly, the dispersion variance of mean grade $Z_p(x)$ of blocks of volume P within the area V depends only on the variogram :

(2) $$\boxed{D^2(P/V) = \overline{\gamma}(V,V) - \overline{\gamma}(P,P)}$$ with $P \subset V \subset D$

$\overline{\gamma}(P,P)$ denoting the mean value of $\gamma(h)$ when the two extremities of vector $\vec{h}$ describe independently the volume P.

Hence, Geostatistics interpret the reality $\{z_o(x), x \in D\}$ as a particular realization of a random function $Z(x)$ characterized by a particular variogram $2\,\gamma(h)$, this variogram being estimated from the real data $\{z_o(x_i), x_i \in I\}$. Now, simulation merely consists in drawing another realization $\{z_s(x), x \in D\}$; $z_s(x)$ is not the reality $z_o(x)$, but both $z_s(x)$ and $z_o(x)$ are drawn from the same random function $Z(x)$ characterized by the auto-correlation function variogram $2\,\gamma(h)$. Conditionalization is a second operation that consists in selecting among all the possible simulations $z_s(x)$, the $z_c(x)$ ones that meet the experimental values at data locations, i.e. :

$$z_c(x_i) = z_o(x_i) \quad \forall\ x_i \in \text{Information} \quad (\text{see Fig. 1})$$

fig:1 Reality and simulations
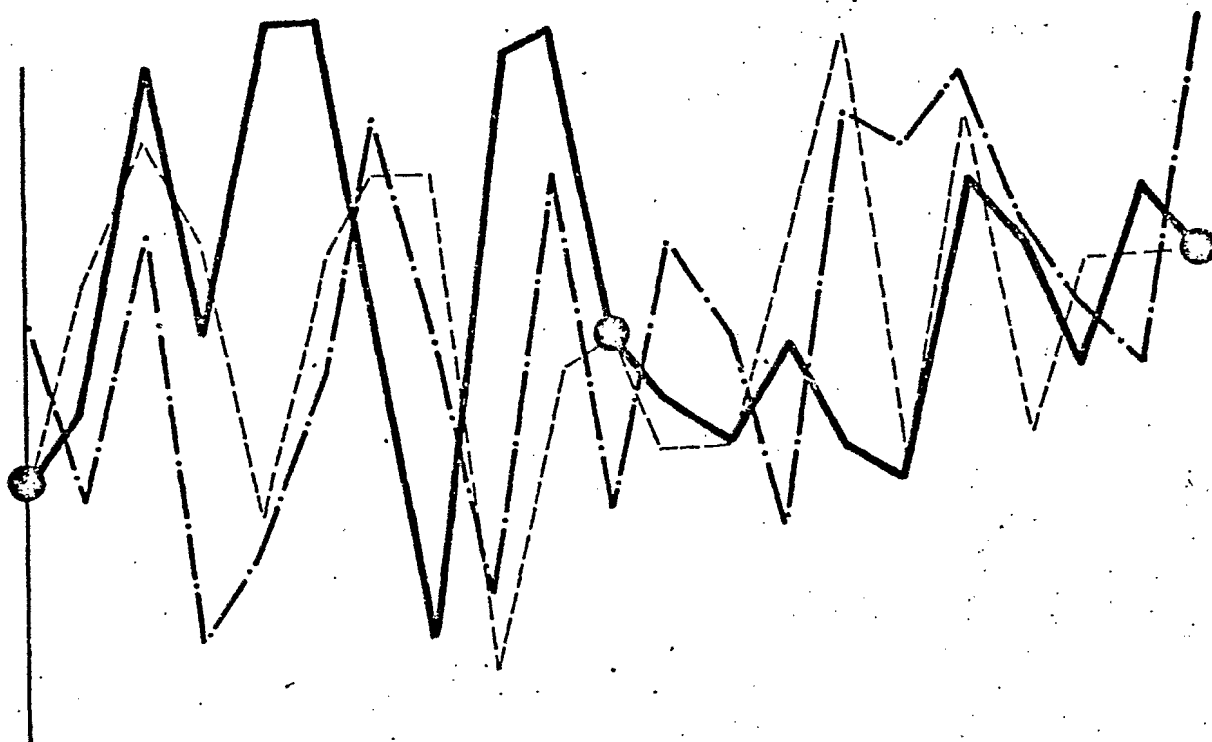
⊙ Data locations

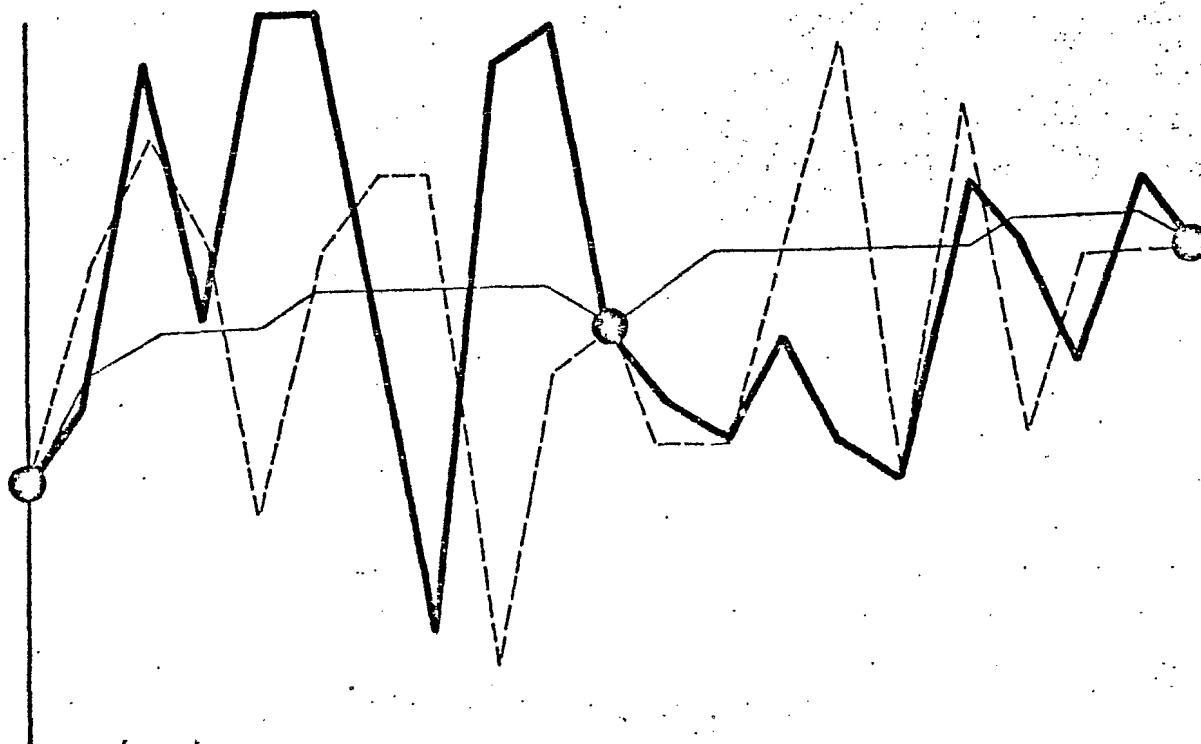|	
▬▬▬▬	$Z_0(x)$ Reality
▬·▬·▬	$Z_s(x)$ Non conditional simulation
▬ ▬ ▬	$Z_t(x)$ Conditional simulation
▬▬▬	$Z_k^*(x)$ Kriging

fig:2 Reality_Simulation_Kriging

Simulation or Estimation -

A question arises : couldn't the simulated surface $\{z_c(x),\ x \in D\}$ be considered as an estimation of the real unknown surface $\{z_0(x),\ x \in D\}$? Obviously $z_c(x)$ is an estimator of reality $z_0(x)$, but a very poor one ; more precisely :

The quality of an estimator $z*(x)$ of a reality $z_0(x)$ is characterized by the estimation variance or mean quadratic error, the error $Z(x) - Z*(x)$ being considered as a random function :

(3)
$$\sigma_E^2 = E\{[Z(x) - Z*(x)]^2\}$$

- The best linear estimator $z_K^*(x)$ is given by the formalism of Kriging (see also [i], G. Matheron, p. 115) that precisely minimizes the previous estimation variance, whereas the conditional simulation $z_c(x)$ considered as an estimator provides an estimation variance twice greater than the kriging variance (see demonstration further in III-1) :

$$E\{[Z_c(x) - Z(x)]^2\} = 2\ E\{[Z_K^*(x) - Z(x)]^2\} = 2\ \sigma_K^2$$

- The kriged value is on average closer to reality, but the kriged surface $\{z_K^*(x),\ x \in D\}$ has no reason to reflect the same "fluctuations pattern" as the real surface $\{z_0(x),\ x \in D\}$. In particular, the kriged surface will smooth the real fluctuations ; see Figure 2 : the dispersion variance $D_K^{2*}(O/V)$ of the kriged value considered as a random function $Z_K^*(x)$ is inferior to the dispersion variance $D^2(O/V)$ of the true values :

$$D_K^{2*}(O/V) < D^2(O/V)$$

In summary, simulation and estimation have two different purposes :

- the <u>simulation</u> $\{z_c(x),\ x \in D\}$ reflects the structural characteristics of the revealed reality. For problems dealing with ore characteristics fluctuations, simulation is required.

- the <u>estimation</u>, in particular the kriging $\{z_K^*(x),\ x \in D\}$

minimizes the estimation variance but smoothes the true spatial fluctuations. For problems dealing with resources or recovered reserves estimation, kriging is required.

III - CONDITIONAL SIMULATION - THEORY

III-1 - Conditionalization.

At each point $x \in$ Deposit, consider the real value $z_0(x)$ and the kriged value $z_K^*(x)$. They differ of an error :

$$z_0(x) = z_K^*(x) + [z_0(x) - z_K^*(x)]$$

This error is unknown, just like the true value $z_0(x)$ itself, and will be simulated. The kriging error $[z_0(x) - z_K^*(x)]$ has a very particular property which is the key to the Theory of Conditional simulations : this error considered as a random function $[Z(x) - z_K^*(x)]$ is orthogonal to $Z_K^*(x)$, i.e. in terms of covariance :

$$E\{Z_K^*(x)[Z(x') - Z_K^*(x')]\} = 0 \qquad \forall \; x, \; x' \in D \qquad \text{which entails :}$$

$$E\{Z_K^*(x) \; Z_K^*(x')\} + E\{[Z(x) - Z_K^*(x)][Z(x') - Z_K^*(x')]\} =$$

$$= E\{Z(x) \; Z(x')\} = C(x-x') = C(h)$$

This property derives directly from the minimization of the estimation variance of Kriging (see [1], G. Matheron, p. 123).

Now, if we consider a random function $Z_S(x)$ independent of $Z(x)$ but isomorphic to it, i.e. $Z_S(x)$ admits the same covariance $C(h)$ as $Z(x)$, the same kriging procedure leads to the analogous decomposition :

$$Z_S(x) = Z_{S,K}^*(x) + [Z_S(x) - Z_{S,K}^*(x)]$$

Moreover, $[Z_S(x) - Z_{S,K}^*(x)]$ is isomorphic to $[Z(x) - Z_K^*(x)]$. Hence the random function :

$$Z_C(x) = Z_K^*(x) + [Z_S(x) - Z_{S,k}^*(x)]$$

constructed by superimposing the residuals $[Z_s(x)-Z^*_{s,K}(x)]$ of the second random function $Z_s(x)$ to the kriging $z^*_K(x)$ of the initial random function $Z(x)$ is isomorphic to $Z(x)$, i.e. $\underline{Z_c(x)}$ admits the same covariance $C(h)$ as $Z(x)$. Furthermore, the corresponding realizations :

(4)
$$z_c(x) = z^*_K(x) + [z_s(x)-z^*_{s,K}(x)]$$

are conditional to the real experimental data : at any data location $x_i \in I$, by definition of kriging, the two residuals $z(x_i) - z^*_K(x_i)$ and $z_s(x_i) - z^*_{s,K}(x_i)$ vanish, it comes then :

$$z_c(x_i) = z(x_i) \qquad \forall\ x_i \in I$$

Hence the random function $Z_c(x)$ and its realizations $z_c(x)$ give a complete solution to the problem of Conditional Simulation :

- the simulated surface $\{z_c(x),\ x \in D\}$ does pass by the experimental data and reflects the same "fluctuations pattern" (i.e. same covariance or variogram) than the real surface.

NOTES : $-z^*_K(x)$, kriged value of $z_0(x)$, is a linear combination of neighboring data : $z^*_K(x) = \sum\limits_i \lambda_i\ z_0(x_i)$

-The weighting factors λ_i are given by the kriging system which is a system of linear equations (see [1], G. Matheron, p. 124).

$-z^*_{s,K}(x)$ is the kriged value of $z_s(x)$ based on the same data pattern as $Z_K(x)$. It is a linear combination of the data $z_s(x_i)$ located at the experimental points $x_i \in I$:

$$z^*_{s,K}(x) = \sum\limits_i \lambda_i\ z_s(x_i)$$

Obviously the weighting factors λ_i are identical for $z^*_{s,K}(x)$ and $z^*_K(x)$.

This technique of conditionalization provides at every point $x \in$ Deposit, both the simulated value $z_c(x)$ and the kriged value $z^*_K(x)$. The estimation variance of the simulation $Z_c(x)$ considered as an estimator of $Z(x)$ is derived from formula 4 :

$$E\{[Z_C(x)-Z(x)]^2\} = 2 \; E\{[Z_K^*(x)-Z(x)]^2\}$$

In terms of estimation variance, the kriged value is twice better than the simulated value. Estimation is not the purpose of simulation.

III-2 - Non-Conditional Simulation.

The following problem remains : to generate a realization $z_S(x)$ of a random function $Z(x)$ with imposed covariance $C(h)$. $z_S(x)$ does not have to be conditional to the real data, i.e. at data locations x_i : $z_S(x_i)$ can be different from $z_O(x_i)$.

- Let us recall that the writings $z_S(x)$ or $z_O(x)$ denote variables spread in the 3-dimensional space $(x \in R^3)$ and that the covariance $C(h)$ is a function of the modulus h of the vector $\vec{h}$.

In the literature, there exist numerous probabilistic procedures to generate one-dimensional realizations $\{z_S(x), \; x \in R^1\}$ of stochastic process $Z(x)$ with imposed covariance. But when extended to 3-dimension problems, these procedures reveal unefficient, if not unreasonable in terms of computer time. They generally consist (see [5], G.M. Jenkins and D.G. Watts, [4], D. Guibal) :

- either in simulating orthogonal random measures the variance density of which meets the spectrum measure of the imposed covariance.

- or in considering moving averages (on circles, spheres, or any figure) which describe a field with a Poisson dispersion.

The originality of the "turning-bands" method initiated by G. Matheron and the Centre de Morphologie de Fontainebleau (see ref. [2], [3] and [4]) consists in reducing any 3-dimensional simulation to simple one-dimensional simulations on lines that turn in the 3-dimensional space. This "turning bands" method provides 3-dimensional simulations with acceptable computer costs of the order of the costs required by classical 1-dimensional procedures.

The turning bands method.

In the 3-dimensional space, consider the line D_1 defined by the

unit vector $\vec{u}_1$ (see Figure 3). On line D_1, consider a one-dimensional random function $Y(u_1)$ stationary of order 2, i.e. with

- a nought expected value $E\{Y\} = 0$

- a 1-dimensional stationary covariance $c^{(1)}(h_{u_1})$

$(\vec{u}_1, \vec{v}_1, \vec{w}_1)$ represents a unit system of orthogonal axis, and $(h_{u_1}, h_{v_1}, h_{w_1})$ the projections of vector $\vec{h}$ on these axes.

- The 3-dimensional random function $Z_1(u_1, v_1, w_1)$ defined by

$$z_1(u_1, v_1, w_1) = y(u_1)$$

is stationary of order 2 with a nought expected value and a 3-dimensional covariance :

$$C_1(h) = C_1(h_{u_1}, h_{v_1}, h_{w_1}) = c^{(1)}(h_{u_1})$$

To generate a realization z_1 of Z_1, the value $y(u_1)$ of the 1-dimensional realization at point u_1 of line D_1 is assigned in practice to all the points interior to the band centred on the place $\{u_1 = $ Constant$\}$ (see Fig. 3). The thickness of the band is the spacing of the values $y(u_1)$ on line D_1.
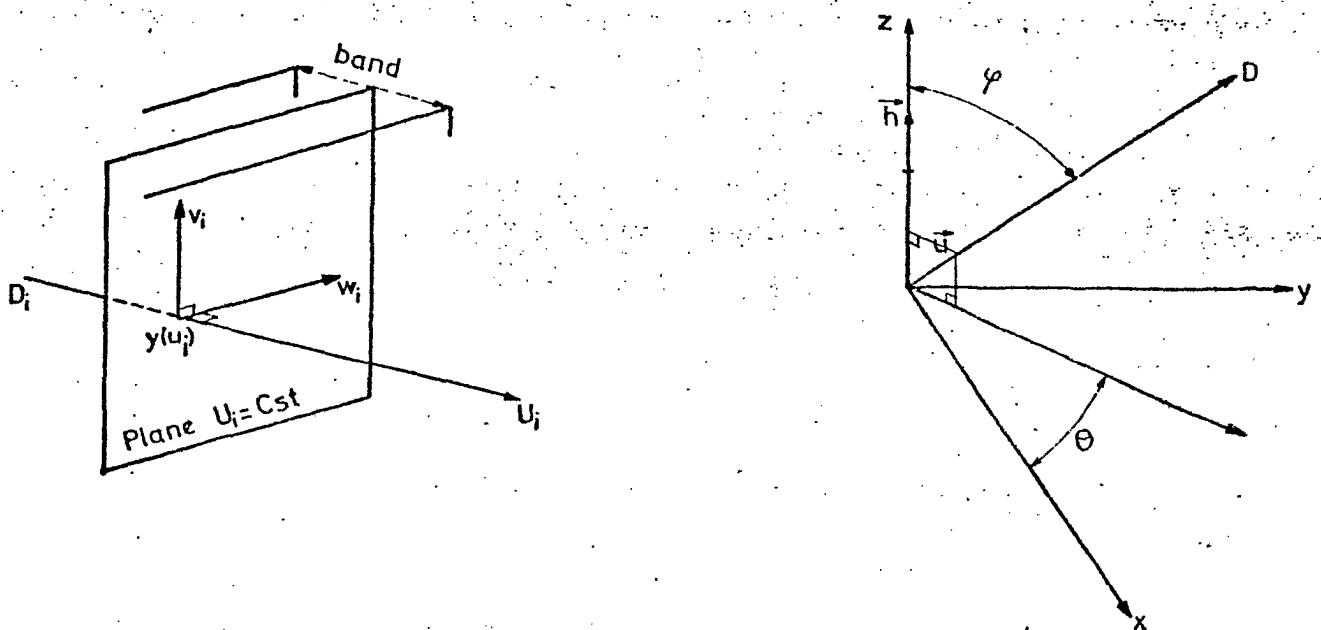

fig:3 The turning bands

- We then consider N directions $u_1, u_2, \ldots u_i, \ldots u_N$ uniformly distributed on the unit sphere. A new realization $y(u_i)$ of a random function Y_i equivalent to Y is generated on every line D_i, the N random functions $\{Y_i, i = 1, N\}$ being independent. To each 1-dimensional realization $y(u_i)$ corresponds the 3-dimensional realization $z_i(u_i, v_i, w_i) = y(u_i)$

- We then consider the sum $z = \sum_{i=1}^{N} z_i$ which is a realization of a 3-dimensional random function $Z(x_1, x_2, x_3)$, stationary of order 2, with a nought expected value and a covariance that tends towards the following isotropic covariance when the number of lines $N \to +\infty$:

$$C(h_1, h_2, h_3) = C(h) = \int_{\substack{1/2 \text{ unit} \\ \text{sphere}}} c^{(1)}(< \vec{h}, \vec{u} >) \, du$$

$< \vec{h}, \vec{u} > = h_u$ denoting the projection of vector $\vec{h}$ on axis $\vec{u}$.

$h = \sqrt{h_1^2 + h_2^2 + h_3^2}$ is the modulus of vector $\vec{h}$. The preceding integral is written with spherical coordinates (See Fig. 3) :

$$C(h) = \int_0^{2\pi} d\theta \int_0^{\pi/2} c^{(1)}(|h \cos \varphi|) \sin \varphi \, d\varphi = \frac{2\pi}{h} \int_0^h c^{(1)}(s) \, ds$$

The 3-dimensional covariance $C(h)$ is imposed, hence the 1-dimensional covariance $c^{(1)}(s)$ derives from formula :

(5) $\boxed{c^{(1)}(s) = \frac{1}{2\pi} \frac{\partial}{\partial s} s \, C(s)}$

It is shown that the function $c^{(1)}(s)$ given by that formula is a covariance function (see [2], G. Matheron). Hence, there always exists a solution to the problem : what is the covariance $c^{(1)}(s)$ of the 1-dimensional random function Y(u) to be simulated on the lines D_i? By turning these lines in the space, the sought after realization of the 3-dimensional random function Z(x) of imposed covariance C(h) is obtained.

NOTES - In practice, N is never infinite. We may consider for example the N = 15 lines that join the middles of opposite edges of a regular

isocahedron. The sum $z = \sum\limits_{i=1}^{15} z_i$ will then reflect a covariance

$$C_d(h) = \sum_{i=1}^{15} c^{(1)}(< \vec{n}, \vec{u}_i >)$$

which differs from the limit covariance $C(h) = \dfrac{2\pi}{h} \displaystyle\int_0^h c^{(1)}(s)ds$. The

bias between $C_d(h)$ and $C(h)$ can be theoretically calculated using the geometrical properties of the regular isocahedron. It has been shown in practice (see ref. [4], D. Guibal, p. 24) that it is sufficient to introduce a multiplicative corrective factor.:

- if $c^{(1)}(s)$ was a constant k, it would come :

.for the limit covariance : $C(h) = 2\pi k$

for the practical covariance : $C_d(h) = 15\ k$

Hence we could consider that the covariance $C_d(h)$ provided by the isocahedron approximation is :

$$C_d(h) \neq \frac{15}{2\pi}\ C(h)$$

- The identification of an expected value $E\{Z(x)\} = m \neq 0$ is immediate. As the previous 3-dimensional realization corresponds to a nought expected value, it is sufficient to add a constant equal to the imposed expected value m to each simulated value $z(x)$

III-3 - One-dimension simulations.

The very last problem left is to generate realizations of the 1-dimensional random function Y(u) with the known imposed covariance

$$c^{(1)}(s) = E\{Y(u)\ Y(u+s)\}$$

the expected value of Y(u) being nought : $E\{Y(u)\} = 0$. There exist several methods to generate these 1-dimensional realizations. The most classical method is based on spectral analysis ; its advantage is to be absolutely general, whatever the imposed covariance $c^{(1)}(s)$ may be, but in practice, this spectral analysis procedure may require heavy Fourier transforms calculations.

Particular cases occur quite frequently in practice, when the

covariance $c^{(1)}(s)$ can be written as a convolution product of a known function $f(u)$ by its transposed $\check{f}(u) = f(-u)$:

(6)
$$c^{(1)}(s) = f * \check{f} = \int_{-\infty}^{+\infty} f(u).f(u+s)du$$

Then the procedure of moving average with the imposed weighting function $f(u)$ reveals to be much easier and economical than the general spectral analysis procedure. Precisely :

- We start from a stationary 1-dimensional random measure $T(dv)$ with a Dirac covariance measure, for example the differential forms of a Poisson process or a Brownian movement. Then with each measure $T(dv)$ is associated the 1-dimensional random function $Y(u)$ defined by :

$$Y(u) = \int_{-\infty}^{+\infty} f(u+v).T(dv)$$

$Y(u)$. appears as a stationary random function the covariance of which is precisely : $c^{(1)}(s) = f * \check{f}$. (This is a classical probabilistic result - see also [1], G. Matheron p. 14 and Exercise 10, p. 104).

- In practice, A discrete approximation of the previous random measure $T(dv)$ is used :.

- First, we start drawing independent realizations of a random variable T, and affect these values to each mesh-point of a regular grid defined on line D - see Fig. 4. The random variable T presents a nought expected value and an imposed variance $E\{T^2\} = \sigma^2$. Several subroutine packages provide such realizations of a random variable, for example the IBM Sub-package RANDU.

Let $t_{i-k},\dots t_i,\dots t_{i+k},\dots$ be these independent realizations at locations $i-k,\dots i,\dots i+k\dots$ on line D. The constant spacing is b.

- Second, we consider the moving average y_i defined at every location i by. the imposed weighting function $f(u)$, i.e. :

$$y_i = \sum_{k=-\infty}^{+\infty} t_{i+k} .f(kb)$$

The values y_i can be interpreted as a realization of a 1-dimensional

random function $Y_i = Y(u)$ with a nought expected value and the following covariance :

$$c_d^{(1)}(s) = E\{Y_i \ Y_{i+s}\} = E\left\{\sum_{k=-\infty}^{+\infty} T_{i+k} \cdot f(kb) \sum_{k'=-\infty}^{+\infty} T_{i+s+k'} \ f(k'b)\right\}$$

With the change : $\ell = s+k'$, it comes :

(7)
$$c_d^{(1)}(s) = E\left\{\sum_{k=-\infty}^{+\infty} \sum_{\ell=-\infty}^{+\infty} T_{i+k} \cdot T_{i+\ell} \cdot f(kb) \cdot f[(\ell-s)b]\right\}$$

This covariance (7) is the discrete approximation of the imposed covariance $c^{(1)}(s)$ of formula (6).

Formula (7) must be developed according to each particular weighting function f ; the small bias between the discreet covariance $c_d^{(1)}(s)$ and the imposed covariance $c^{(1)}(s)$ is then easily corrected by acting on two parameters : variance σ^2 and spacing b.
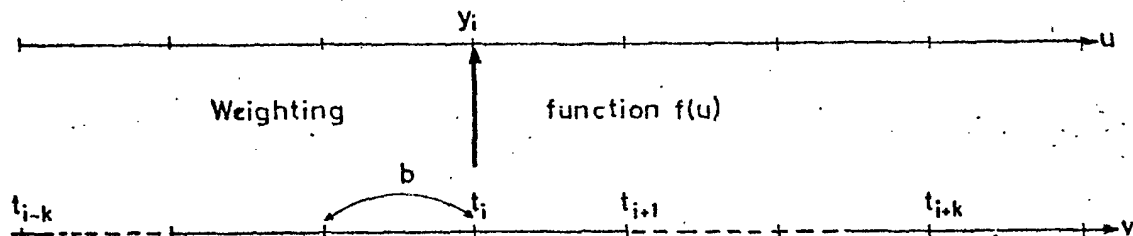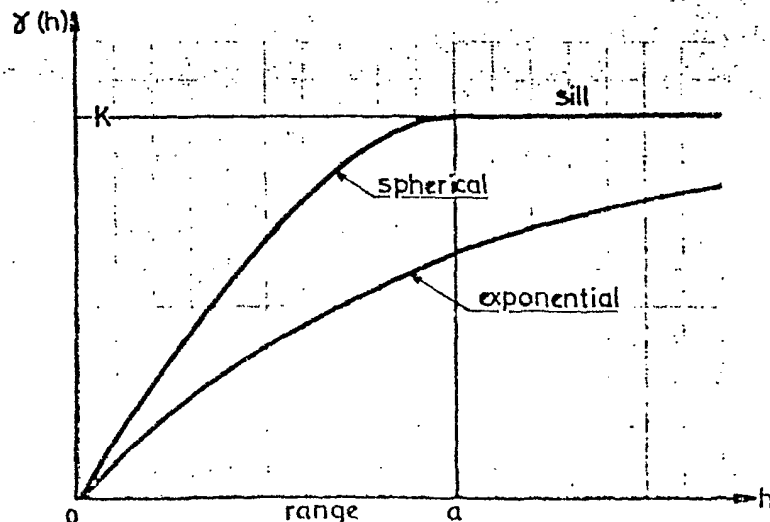

fig:4 1 dim. simulation

fig:5 Spherical and exponential
schemes

Some practical examples.

The preceding moving average procedure provides 1-dimensional realizations reflecting covariances which can be written as a convolution product (see formula (6)) : $C^{(1)}(s) = f * \overset{\vee}{f}$.

Now, the two 3-dimensional covariances $C(h)$ most frequently used in Geostatistics, namely the spherical covariance and the exponential covariance, correspond precisely to the 1-dimensional imposed covariances $C^{(1)}(s)$ that can be written as convolution products.

More precisely, with the previous notations, and $C(h)$ being the covariance of the 3-dimensional random function $\{Z(x), x \in \mathbb{R}^3\}$:

- 1 - The spherical covariance is written

$$(7) \, C(h) = \begin{cases} K[1 - \dfrac{3h}{2a} + \dfrac{h^3}{2a^3}] & \forall \ 0 \le h \le a \\ \\ 0 & \forall \ h \ge a \end{cases} \quad \text{with } K, \, a > 0$$

The corresponding variogram is deduced from formula (1) :

$$\gamma(h) = C(o) - C(h) \qquad \text{(See Fig. 5)}$$

The constant K is called "sill" of the variogram. Independence between the 2 variables $Z(x)$ and $Z(x+h)$ is reached for distances h superior to a. The distance a, which characterizes the zone of influence of the variable $Z(x)$, is called "range" of the variogram.

- 2 - The exponential covariance is written :

$$(8) \qquad C(h) = K \, e^{-ah} \qquad \forall \ h \ge 0 \qquad \text{with } K, \, a > 0$$

The corresponding semi-variogram is (see Fig. 5) :

$$\gamma(h) = K[1 - e^{-ah}] \qquad \forall \ h \ge 0$$

If we want to simulate 3-dimensional realizations that reflect the 2 previous covariances $C(h)$, we must first simulate 1-dimensional realizations on lines that reflect the covariance $C^{(1)}(s)$ given by formula (5) :

1 - For the spherical covariance, it comes :

$$c^{(1)}(s) = \begin{cases} \frac{K}{2\pi} \left[1 - \frac{3s}{a} + \frac{2s^3}{a^3} \right] & \forall \quad 0 \leq s \leq a \\ \\ 0 & \forall \quad s \geq a \end{cases}$$

This covariance is nothing but the convolution product : $c^{(1)}(s) = f * \overset{v}{f}$ with the following weighting function :

$$f(u) = \begin{cases} \sqrt{\frac{6K}{\pi a^3}} \cdot u & \text{if} \quad -\frac{a}{2} \leq u \leq \frac{a}{2} \\ \\ 0 & \text{if not} \end{cases}$$

The demonstration is immediate.

2 - For the exponential covariance it comes :

$$c^{(1)}(s) = \frac{K}{2\pi} \ (1-as) \ e^{-as} \qquad \forall \ s \geq 0$$

This covariance is nothing but the convolution product : $c^{(1)}(s) = f * \overset{v}{f}$ with the following weighting function :

$$f(u) = \begin{cases} \sqrt{\frac{2Ka}{\pi}} \ (1-au) e^{-au} & \text{if} \ u \geq 0 \\ \\ 0 & \text{if not} : u < 0 \end{cases}$$

It is recalled that the practical discrete approximations $c_d^{(1)}(s)$ differ from the previous expressions $c^{(1)}(s)$ (see formula (7)), and that these differences must be corrected (see precise calculations in ref. [4], D. Guibal, p. 19 and p. 30).

NOTE - The two examples of covariance $C(h)$ (spherical and exponential) given here have no absolute character. They are frequently used in practice

- first, for the simplicity of their mathematical expressions. This fact is of considerable help, because Geostatistics constantly considers integrals of the covariance or variogram function.

- second, because they depend on two parameters only, sill K and range a, that can be easily deduced from the available experimental covariance or variogram.

III-4 - Gaussian Anamorphosis.

The regionalization of the true values $\{z_o(x),\ x \in \text{Deposit}\}$ has been interpreted as a particular realization of a random function $\{Z(x),\ x \in D\}$. This random function is defined by its spatial law, and not only by the first 2 moments - mean and covariance - imposed to the simulation. Hence, the simulated value $\{z_c(x),\ x \in D\}$ does not reflect the whole space law of $Z(x)$, but only the first two moments imposed. In practice, this is not too embarassing for we have shown (see Chapter II and formula (2)) that all dispersion variances depend only on these imposed first two moments. Moreover, experimental data seldom allows recognition of moments of order superior to 2.

The turning bands method, that consists in summing up in space a great number N of independent realizations defined on lines, generates realizations $z_c(x)$ of a random function $Z_c(x)$ with a <u>Gaussian space law</u> (central limit theorem). Hence, under the stationary hypothesis, the dispersion law of $Z_c(x)$ - x being fixed - is Gaussian and independent of x. Therefore, the histogram of the simulated values $\{z_c(x),\ x \in D\}$, which is an estimation of that dispersion law of $Z_c(x)$ - x being fixed - will be of a Gaussian type.

In practice, real values $\{z_o(x),\ x \in D\}$ are very often positive or nought (thicknesses, grades, etc..). Hence, the experimental histogram of real data $\{z_o(x_i),\ i \in \text{Information}\}$ will reflect a lognormal, Gamma or other dispersion law that ensures $z_o(x) \geq 0$, $\forall\ x \in D$. Whereas the simulation, being Gaussian, may present a fair proportion of negative simulated values : $z_c(x) < 0$.

Hence, it is of prime importance that the histogram of simulated values reflects the same dispersion law as the experimental histogram of real data. This is the purpose of the Gaussian anamorphosis :

$\{z_o(x),\ x \in D\}$ being the real values, let $g_o(x) = F\{z_o(x)\}$ be the transformed of $z_o(x)$ by function F, such that $\{g_o(x),\ x \in D\}$ presents

a Gaussian histogram. Instead of simulating $z_0(x)$, we shall simulate its transform $g_0(x)$. This simulation $\{g_c(x),\ x \in D\}$ will reflect a Gaussian histogram. Hence, the inverse transformation F^{-1} will provide the sought after simulation $\{z_c(x) = F^{-1}\{g_c(x)\},\ x \in D\}$ that presents the good histogram, i.e. a histogram similar to the histogram of real data $\{z_0(x_i),\ i \in I\}$. Obviously the covariance $C(h)$ imposed to the simulation of $g_0(x)$ is deduced from the available data $\{g_0(x_i) = F\{z_0(x_i)\},\ i \in I\}$.

Figure 6 shows the experimental histogram of real values $\{z_0(x_i),\ i \in I\}$. A lognormal dispersion law could be fitted to that experimental histogram. Hence, the transformed variable $g_0(x) =$ Log $\{z_0(x)\}$ will present a Gaussian histogram. The spatial regionalization $\{g_0(x),\ x \in D\}$ is then considered as a particular realization of the random function $\{G(x),\ x \in D\}$ with imposed expected value and covariance. These two imposed moments are deduced from the available data $\{g_0(x_i) = \text{Log } \{z_0(x_i)\},\ i \in I\}$. The turning bands method then provides a simulation $\{g_c(x),\ x \in D\}$ that is conditional to the data, i.e. at each location x_i of real data

$$g_c(x_i) = g_0(x_i) = \text{Log } \{z_0(x_i)\} \qquad \forall\ i \in I$$

The histogram of the simulated values $\{g_c(x),\ x \in D\}$ being Gaussian, the inverse transformation $z_c(x) = \exp \{g_c(x)\}$ will provide a simulation $\{z_c(x),\ x \in D\}$ that :

- presents a lognormal histogram

- is conditional to the real data $z_c(x_i) = z_0(x_i) \qquad \forall\ i \in I$

- reflects the same first 2 moments, expected value and covariance, as the revealed reality $\{z_0(x_i),\ i \in I\}$.
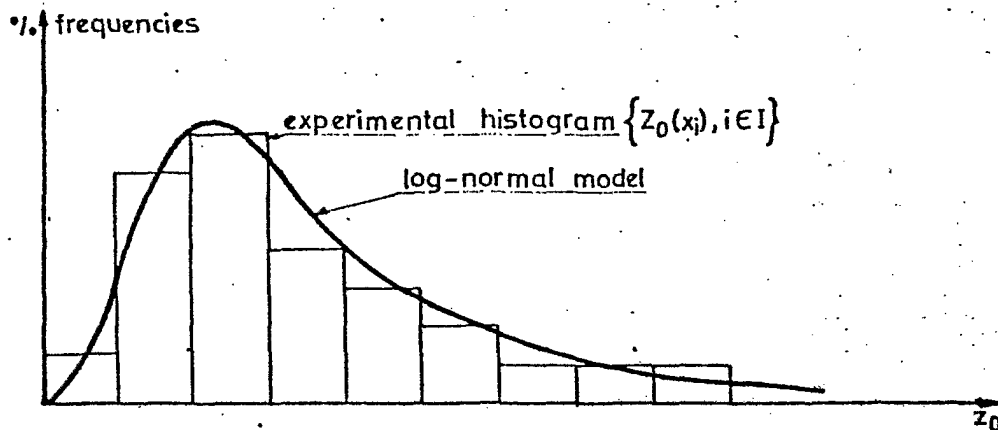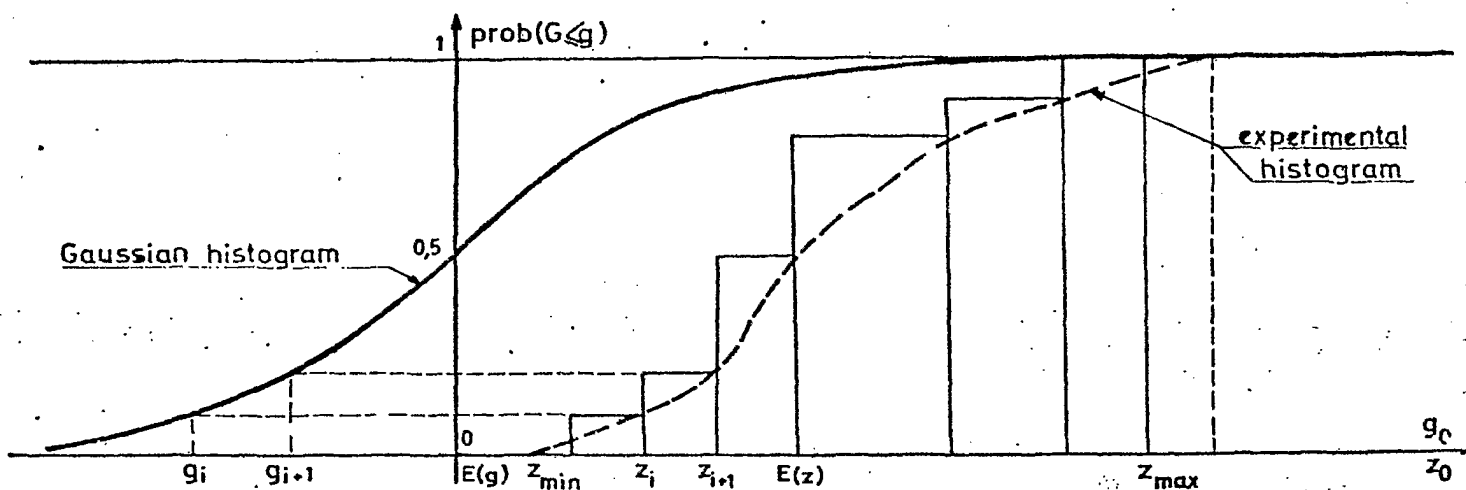
fig:6 Gaussian anamorphosis

fig:7 Graphic anamorphosis

Graphical Anamorphosis.

In practice, if a precise and simple dispersion law - for example lognormal or Gamma - cannot be fitted to the experimental histogram of real data $\{z_0(x_i), i \in I\}$, the anamorphosis F will be done graphically :

The two following cumulative histograms (see Fig. 7) are considered :

- the theoretical cumulative histogram of the reduced Gaussian variable G, i.e. the curve that gives the probability law $\text{Prob}\{G \leq g\}$.

- the experimental cumulative histogram of the real data
$\{z_0(x_i), i \in I\}$. Each class (z_i, z_{i+1}) of the experimental curve
corresponds graphically to a class (g_i, g_{i+1}) of the theoretical
Gaussian curve. We then proceed to a linear interpolation within
each class, $z \in [z_i, z_{i+1}]$ or $g \in [g_i, g_{i+1}]$. Hence, the bijection
$g \overset{\leftarrow}{\underset{\rightarrow}{}} z$ is graphically defined, i.e. the anamorphosis :

$$g_0(x) = F\{z_0(x)\}$$

This graphical procedure is nothing but the well-known Monte-
Carlo technique (see [14], Y.C. Kim, [7], J.W. Harbaugh, and many
others).

III-5 - Simultaneous simulations of several variables.

In practice, we often have to deal with several intercor-
related variables, for example : the simulation of an orebody usually
consists in simulating simultaneously various variables such as :

- the two thicknesses of overburden and ore

- the various corresponding metal accumulations, a metal accu-
mulation being defined as the product of a thickness by the average
metal grade on this thickness.

Structural Analysis.

Let $\{z_k(x), x \in D, k = 1 \text{ to } n\}$ be the n variables defined on
the deposit D.

Before simulating them, we have to characterize their spatial
"fluctuation pattern" :

1 - Each of these variables $z_k(x)$ is considered as a particular
realization of a random function $Z_k(x)$, stationary of order 2, charac-
terized by the first 2 moments :

Expected value : $E\{Z_k(x)\} = m_k = $ Constant

Covariance : $C_{kk}(h) = E\{Z_k(x) Z_k(x+h)\} - m_k^2$

these two moments being estimated from the available data
$\{z_k(x_i), x_i \in I\}$. The covariance $C_{kk}(h)$ characterizes the spatial

<u>auto-</u> correlations of $Z_k(x)$.

2 - Now the n variables are <u>inter-</u>correlated in space, i.e. there exist cross-covariances :

$$C_{kk'}(h) = E\{Z_k(x)\ Z_{k'}(x+h)\} - m_k\ m_{k'}, \qquad \forall\ k,\ k' = 1\ \text{to}\ n$$

these cross covariances being estimated from the available data $\{z_k(x_i),\ z_{k'}(x_j)\ ;\ x_i, x_j \in I\}$.

<u>NOTE</u> - The classical correlation coefficient between $Z_k(x)$ and $Z_{k'}(x)$ is nothing but :

$$\rho_{ij} = \frac{C_{kk'}(o)}{\sqrt{C_{kk}(o) \cdot C_{k'k'}(o)}}$$

This correlation coefficient characterizes only local inter-correlation between two variables located at the same point x (i.e. for h = 0), when the cross-covariance $C_{kk'}(h)$ characterizes the spatial inter-correlation (i.e. for any h).

Hence, the "fluctuation pattern" of the set of the n variables $\{z_k(x),\ x \in D,\ k = 1\ \text{to}\ n\}$ is characterized by the covariance matrix $\{C_{kk'}(h)\}$, just as the "fluctuation pattern" of the single variable $\{z_k(x),\ x \in D\}$ was characterized by the single covariance

$$C(h) = C_{kk}(h)$$

<u>Conditional Simulations.</u>

Simulating n variables with an imposed covariance matrix presents no difference from simulating one variable with an imposed covariance function. The turning bands method applies :

- First, we simulate on lines 1-dimensional realizations of the n random functions $\{Y_k(u),\ k = 1\ \text{to}\ n\}$ with the imposed 1-dimensional covariance matrix $\{C_{kk'}^{(1)}(s)\}$. Each covariance $C_{kk'}^{(1)}(s)$ of this matrix is given by formula (5) :

$$C_{kk'}^{(1)}(s) = \frac{1}{2\pi}\ \frac{\partial}{\partial s}\ s\ C_{kk'}(s)$$

- Second, these lines are turned in the 3-dimensional space,

thus generating the desired 3-dimensional realizations of the n random functions $\{Z_k(x), \; k = 1 \text{ to } n\}$ with the imposed 3-dimensional covariance matrix $\{C_{kk'}(h)\}$.

In practice, difficulties arise mainly from 2 facts :

1 - Experimental data $\{z_k(x_i), \; z_{k'}(x_j) \; ; \; x_i, x_j \in I\}$ may not allow correct estimation of all the terms of the covariance matrix $\{C_{kk'}(h)\}$. Usually, the direct covariances $C_{kk}(h)$ are well estimated, whereas some cross-covariances $C_{kk'}(h)$ are less easily estimated.

2 - Spectral analysis always provides a solution to the problem of simulating 1-dimensional realizations reflecting the imposed co-variance matrix $\{C_{kk'}^{(1)}(s)\}$, but may require heavy calculations (see ref. [5], G.M. Jenkins, p. 321). The simple method of moving average exposed in III-3 for simulation of one single variable can be extended to simultaneous simulation of n variables if all the covariances of matrix $\{C_{kk'}^{(1)}(s)\}$ are proportional to a unique covariance $C^{(1)}(s)$, which is a convolution product :

(9)

$$
\begin{array}{|ll|}
\hline
C_{kk'}^{(1)}(s) = \lambda_{kk'} \cdot C^{(1)}(s) & \\
& \\
C^{(1)}(s) = f * \overset{\vee}{f} & \\
\hline
\end{array}
$$

$\forall \; k, \; k' = 1 \text{ to } n$

$\lambda_{kk'}$ being a constant

In III-3, we have seen how to generate a 1-dimensional realization $p_\ell(u)$ of a random function $P(u)$ with the imposed covariance $C^{(1)}(s) = f * \overset{\vee}{f}$. Consider n such realizations $p_1, \ldots p_2, \ldots p_\ell, \ldots p_n$ independent of each other, and the n following linear combinations :

$$
y_k(u) = \sum_{\ell=1}^{n} a_{k\ell} \; p_\ell(u) \qquad \forall \; k = 1 \text{ to } n
$$

The n realizations $\{y_k(u), \; k = 1 \text{ to } n\}$ thus obtained correspond to n random functions $\{Y_k(u), \; k = 1 \text{ to } n\}$, the covariance matrix of which is

$$
C_{kk'}^{(1)}(s) = \left[\sum_{\ell=1}^{n} a_{k\ell} \; a_{k'\ell} \right] \cdot C^{(1)}(s) \qquad \forall \; k, \; k' = 1 \text{ to } n
$$

with the following constants identification :

$$\sum_{\ell=1}^{n} a_{k\ell} \, a_{k'\ell} = \lambda_{kk'}$$

The covariance matrix of $\{Y_k(u), \ k = 1 \text{ to } n\}$ is precisely the imposed covariance matrix of formula (9).

IV - FIRST CASE STUDY : THE PRONY DEPOSIT

(Readers can refer to [9], [10], [12], Ch. Huijbregts and A.G.
Journel for more detailed accounts on the geostatistical sequential
exploration of the Prony orebody).

Presentation of the Deposit.

The Prony deposit (New Caledonia) is a subhorizontal formation
of nickel-laterites which are products of weathering of basic rocks,
largely peridotites. The deposit is surveyed by systematic grids of
vertical boreholes, each borehole is analyzed meter by meter for
Ni grade and impurity Mg grade (see Fig. 8).

Sequential exploration of the deposits provided successively :
- global and local estimation of the resources in situ
- global and local estimation of the recovered reserves, with a
limit mineable ore thickness criterion.

Then, there was the problem of predicting the fluctuations of
overburden ratios and of grades (Ni and Mg) that exploitation and
milling will meet at various scales (day - month - year).

A solution to that problem could have been the drilling of a
systematic square grid of 10m. spacing : 17,600 boreholes on the
first production zone, which is unreasonable. With the help of the
available grid of 100m. spacing (176 boreholes), conditional simu-
lation provides the simulated data of these 17600 boreholes with
10m. spacing. Hence, that key problem of predicting ore charac-
teristics fluctuations receives an immediate and inexpensive so-
lution.

From a practical point of view, it was proceeded as follows
to realize these conditional simulations :
Vertical structural analysis.

Vertical profiles of Ni grades show a common form. From the
top downwards a progressive enrichment is observed, followed by a
fairly rapid decrease in grade. In geostatistical terms, this re-
presents a "vertical drift" of Ni grades ; residual fluctuations
are superimposed on this drift or mean tendency. Universal Kriging

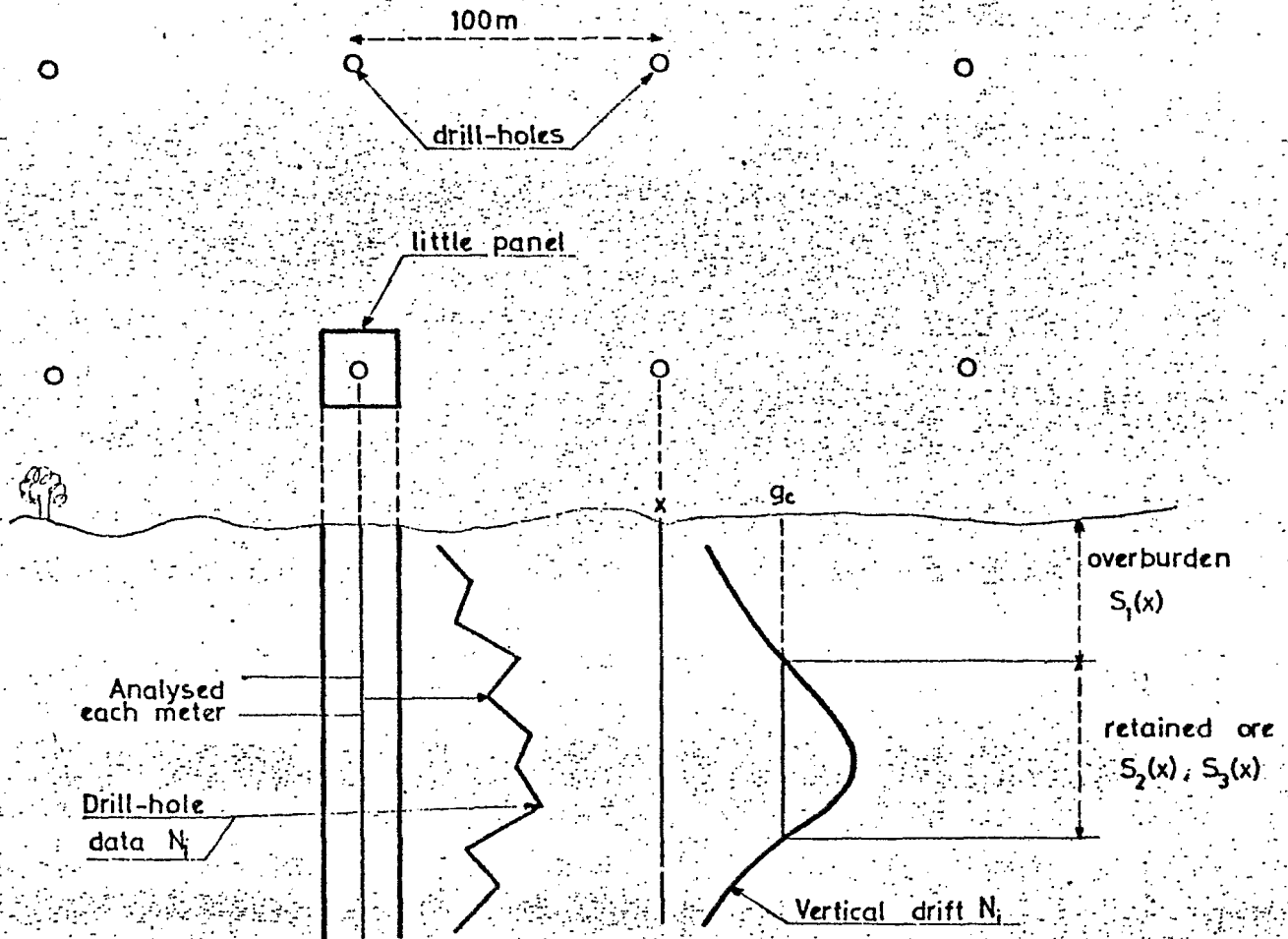permits estimation of this drift curve from the data of each bore-
hole (see Fig. 8).

fig:8   Definition of the service-variables

The vertical series of Ni grades observed along a borehole
(saw-tooth curve) is not representative of the series of mean Ni
grades in a small panel centered on the borehole. It has been shown
(ref. [9]) that an estimator of that vertical series of mean Ni
grades is precisely the previous vertical drift curve.

The entire mineralized thickness cut by each borehole cannot
be exploited economically. On each unit of exploitation (approxima-
tely a 20×20 m^2 panel), only a certain useful thickness is retained

the thickness defined by a specific cut-off grade g_c in nickel.
This cut-off grade ought to be applied to the drift curve repre-
sentative of the panel ; hence, for each cut-off value g_c, a cer-
tain number of "service variables" which are spread in the 2-dimen-
sional horizontal space (coordinates $(x_1,x_2) = x \in \mathbb{R}^2$) are defined :

$S_1(x)$: overburden thickness (see Fig. 8)

$S_2(x)$: thickness of the retained ore

$S_3(x)$: Ni-metal accumulation corresponding to the retained
thickness $S_2(x)$. $S_3(x)$ is the product of $S_2(x)$ by the corresponding
average Ni grade read on the drift curve.

Horizontal structural analysis.

The first production zone has been surveyed with a systematic
square grid of 100 m. spacing. Moreover, two cross-shaped grids
with 20 m. spacing have been drilled to recognize the short-dis-
tances structures. That information allowed the characterization
of the horizontal structures of the service-variables $S_1(x)$, $S_2(x)$,
$S_3(x)$. Figure 10 shows the following examples :

- The experimental histogram of data $\{S_2(x_i), x_i \in I = 100$ m.
spacing grid} for the cut-off $g_c = 0\%$ Ni.

- The experimental points and the theoretical semi-variograms
models $\gamma_3(h)$ adjusted. The two variables considered are the prece-
ding ore thickness $S_2(x)$ for $g_c = 0\%$ Ni and the corresponding Ni-
metal accumulation $S_3(x)$.

- On table 9 : the local correlation coefficients $\rho_{S_1-S_2}$ and
$\rho_{S_2-S_3}$ estimated from the data $\{S(x_i), x_i \in I\}$.

Conditional Simulations

The conditioning real data are located on the 176 mesh-points
of the available grid of 100m. spacing : $\{S_1(x_i), S_2(x_i), S_3(x_i),
x_i \in I\}$. The two thicknesses of overburden $S_1(x)$ and retained ore
$S_2(x)$ were simulated independently with the only constraint :

$$S_1(x) + S_2(x) \le 40 \text{ meters}$$

The two variables $S_2(x)$ and $S_3(x)$ were simulated simultaneously with an imposed covariance matrix of the form (9) (see Chapter III-5).

Five cut-off grades in nickel, $g_c \in [0 \text{ to } 1,5\%]$ were considered. The corresponding 5 groups of service-variables $\{S_1(x), S_2(x), S_3(x)$ were simulated independently.

The various preceding service-variables were simulated on a systematic 10 m. spacing grid corresponding to 17600 boreholes. Simulated data are equal to real data at each one of the 176 locations of real boreholes.

Histograms, variograms, cross-variograms, and correlation coefficients between simulated values were systematically compared with the corresponding features deduced from real experimental data. All these verifications were positive. Figure 10 shows the following examples :

- The histogram of simulated data $\{S_2(x)\}$ for the cut-off grade $g_c = 0\% \text{ Ni.}$
- The two semi-variograms γ_{S_2} and γ_{S_3} deduced from the simulated data $\{S_2(x), S_3(x)\}$ for $g_c = 0\% \text{ Ni.}$
- On table 9 : the local correlation coefficients $\rho_{S_1-S_2}$ and $\rho_{S_2-S_3}$ deduced from the simulated data for $g_c = 0\% \text{ Ni.}$

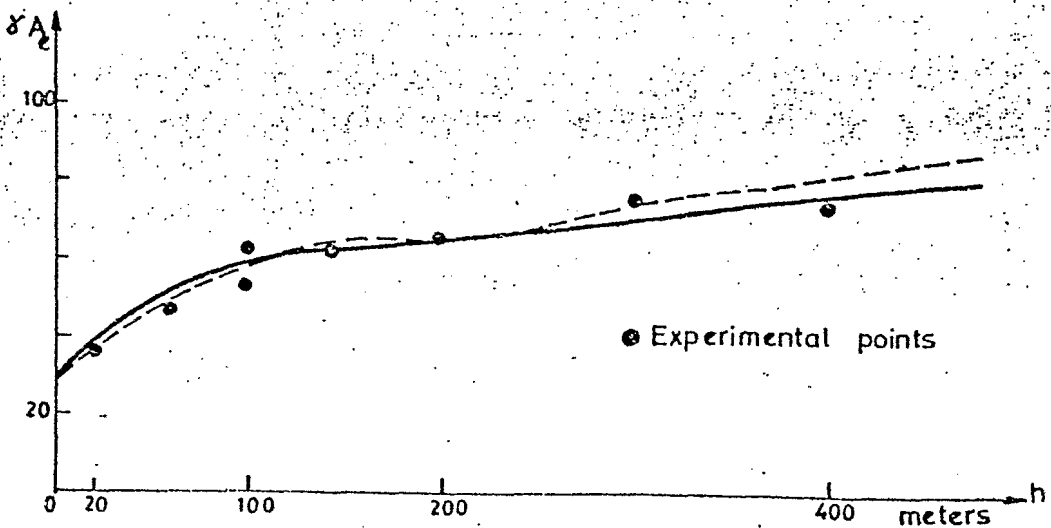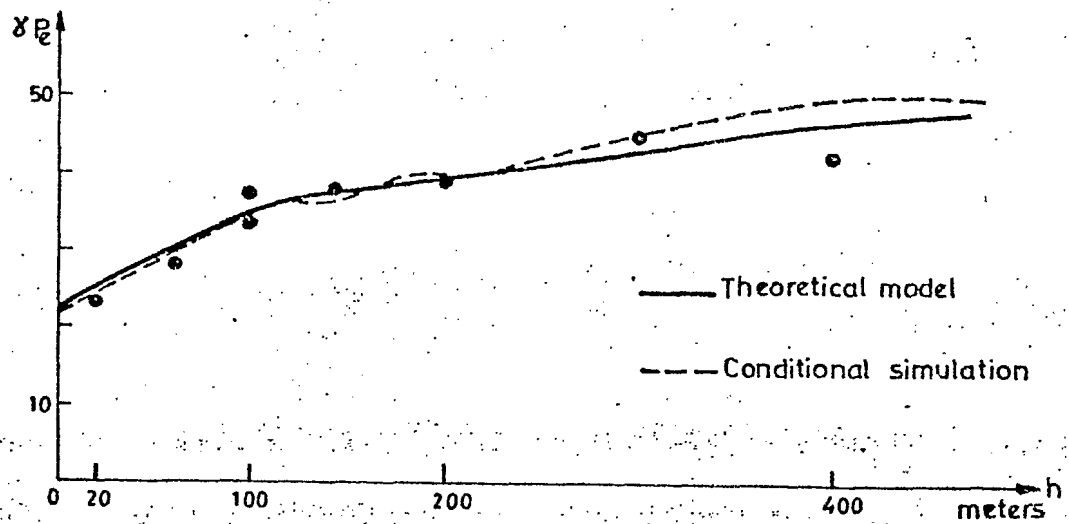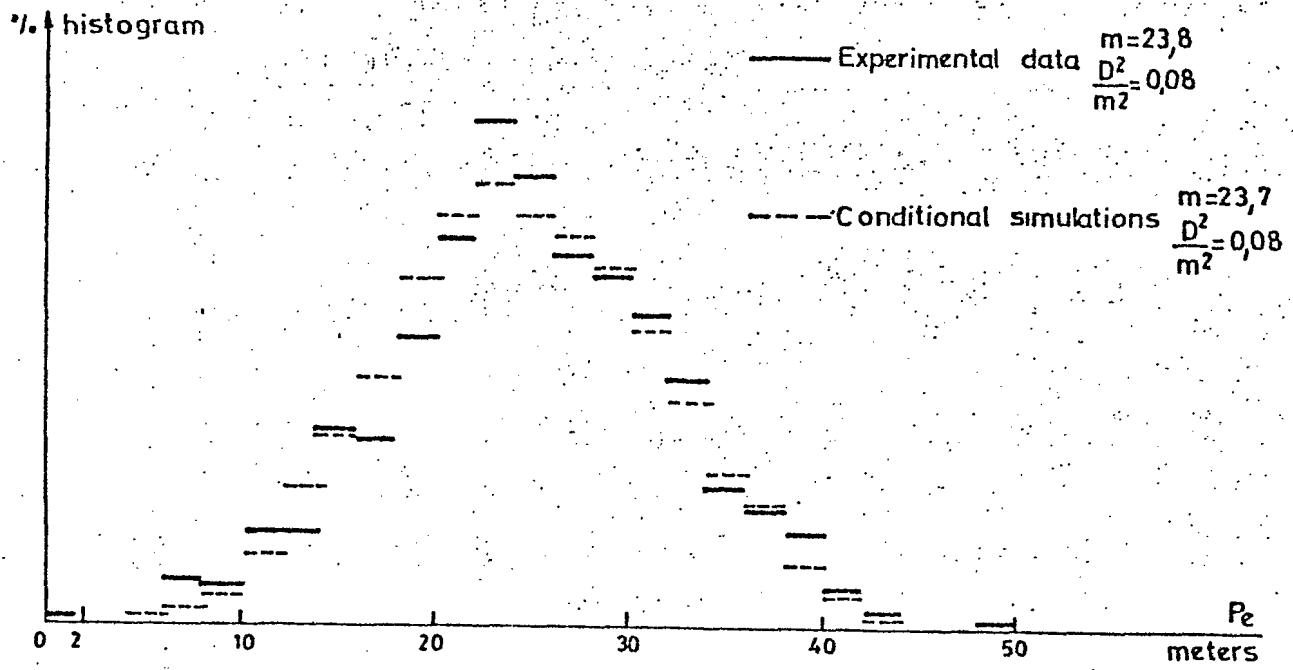| cut/off	Experimental		Simulated	
g_c	$\rho_{S_1-S_2}$	$\rho_{S_2-S_3}$	$\rho_{S_1-S_2}$	$\rho_{S_2-S_3}$
0.0		0.87		0.83
1.0	-0.35	0.99	-0.25	0.98

Table:9 Correlations coefficients

fig:10 Prony simulations, cut-off $g_c = 0\%\, N_i$

V – SECOND CASE STUDY – A SIMULATION DATA BASE

It has already been indicated that the two models of 3-dimensional covariances or variograms most frequently used in practice were the spherical and the exponential scheme (see Chapter III-3, formulae (7) and (8)). These two schemes depend on two parameters only :

- the sill K which is a multiplicative factor

- the range a which characterizes the zone of influence (see Fig. 5). Now, we can record on magnetic tapes or disks a simulation data base constituted of various 3-dimensional realizations of random functions, the isotropic covariances of which are spherical or exponential with a unique sill K = 1 and various ranges a. In practice, if we look for a particular realization with an imposed exponential or spherical covariance (K and a being fixed), it is sufficient to draw the adequate realization (spherical or exponential with a range close to the fixed value a) from the simulation data base, and to multiply each simulated variable by the fixed square-root $\sqrt{K}$.

As an example, consider a subhorizontal sedimentary deposit with a preferential horizontal structure due for instance to horizontal stratifications. The 3-dimensional experimental variogram characteristic of the spatial display of grade in such deposit may be adjusted by the following anisotropic model constituted of three intermeshed structures :

$$(10) \quad \gamma(h_x, h_y, h_z) = K_0 + K_1 \, \gamma_1\left(\sqrt{h_x^2, \; h_y^2, \; h_z^2}\right) + K_2 \, \gamma_2\left(\sqrt{h_x^2 + h_y^2}\right)$$

(h_x, h_y, h_z) being the three coordinates of vector $\vec{h}$, and

K_0, K_1, K_2 being three positive constants,

γ_1 and γ_2 being two spherical schemes with ranges a_1, a_2 and sills equal to 1. Hence, the spatial display of the variable grade $g(x,y,z)$ is interpreted as the sum of three independent realizations of three random functions G_0, G_1, G_2 :

$$g(x,y,z) = g_0(x,y,z) + g_1(x,y,z) + g_2(x,y,z)$$

We want to simulate a 3-dimensional realization $g_s(x,y,z)$ that will reflect the previous anisotropic variogram $\gamma(h_x, h_y, h_z)$.

1) $G_0(x,y,z)$ characterizes a white noise (a nugget effect in geostatistical term) i.e. a microstructure of very small range, that is to say with practically no spatial auto-correlation, or an error of measure. All the values $\{g_0(x,y,z),\ x,\ y,\ z \in$ Deposit$\}$ are independent of each other ; their simulation is very easily done by drawing random numbers with the imposed dispersion variance K_0.

2) $G_1(x,y,z)$ characterizes an isotropic 3-dimensional structure with range a_1. The simulation data base provides a simulated realization $g_{1,s}(x,y,z)$ that reflects the imposed spherical scheme $K_1\ \gamma_1$.

3) $G_2(x,y,z)$ characterizes the horizontal stratification phenomenon : the variable g_2 varies only in the horizontal plane, i.e. : $g_2(x,y,z) = g_2(x,y,z')$ whatever $z \neq z'$ may be.

The semi-variogram $K_2\ \gamma_2$ depends on the horizontal distance $\sqrt{h_x^2 + h_y^2}$ only, and will usually present a greater range $a_2 > a_1$. On any horizontal plane $(z = z_0)$, the simulation data base provides a realization $g_{2,s}(x,y,z_0)$ that reflects the imposed spherical scheme $K_2\ \gamma_2$. The 3-dimensional realization is hence :

$$g_{2,s}(x,y,z) = g_{2,s}(x,y,z_0) \quad \text{whatever } z \text{ may be}$$

4) By summing up the three independent components, we obtain the final desired simulation of the variable grade :

$$g_s(x,y,z) = g_{0,s}(x,y,z) + g_{1,s}(x,y,z) + g_{2,s}(x,y,z)$$

5) Now, if we want a conditional simulation, i.e. a simulation $g_c(x,y,z)$ that meets the real experimental values at data locations $\{x_i, y_i, z_i \in I\}$, we proceed as indicated in Chapter III-1, formula (4) :

$$g_c(x,y,z) = g_K^*(x,y,z) + [g_s(x,y,z) - g_{s,K}^*(x,y,z)]$$

The Centre de Morphologie Mathématique of Fontainebleau has realized that simulation data base over a parallelepipedic volume of 550 × 110 × 20 cubic units (see Figure 11). The imposed 3-dimensional isotropic scheme is the spherical scheme with a sill K = 1 and various ranges a varying from 10 to 200 units. Hence, for each range a, 1,210.000 values were simulated. By grouping the punctual values over panels of dimensions 11 × 11 × 20 cubic units, we obtain the corresponding simulations of 500 panels and if required, the simulations of the 500 vertical centrally located boreholes (see Fig. 11).
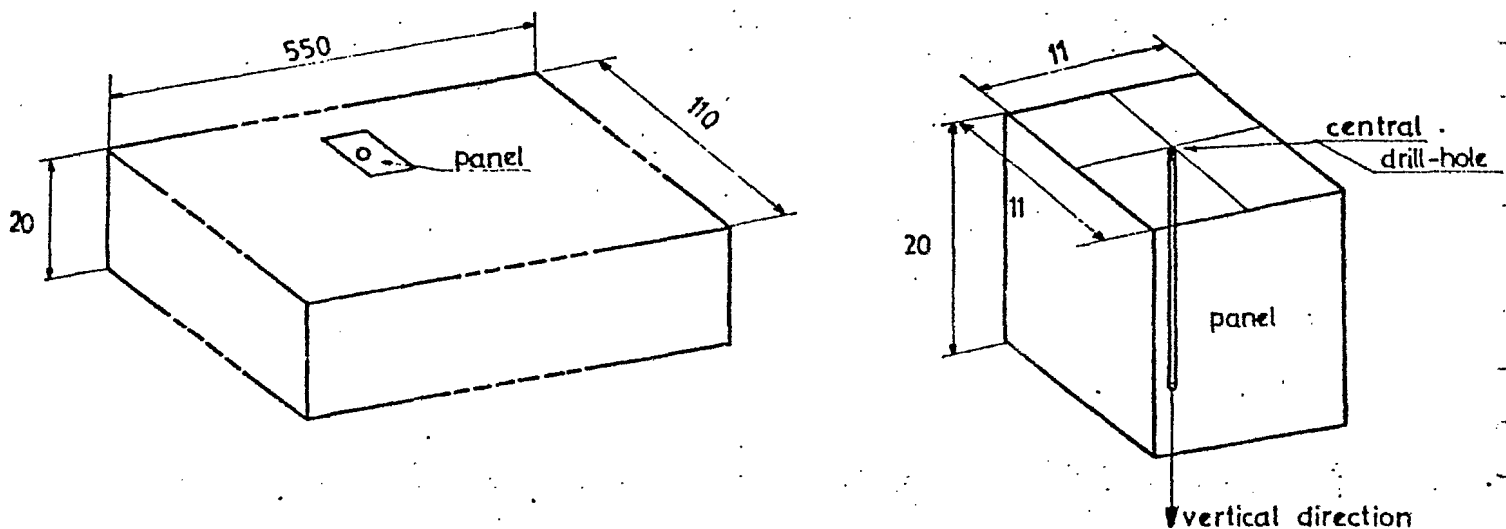

fig:11 The simulation data base

From that simulation data base, a 3-dimensional anisotropic orebody has been simulated. On that orebody, the Centre de Morphologie Mathématique intends to :

- test various estimation procedures : geostatistics - adjustment by mean squares - polygons of influence - inverse and inverse square distance weightings

- verify some practical approximations of Geostatistics

- hence, bring out an essential contribution to answer the old question : Geostatistics or more classical procedure ?

On Figure 12, we can see as examples :

- the vertical profile of simulated mean grades of a panel of section 11 × 11 square units and 20 units height.

- the vertical profiles of simulated grades of the centrally located borehole and of another borehole distant of 4 units, therefore still interior to the preceding panel.

A random vertical drift has been simulated in that orebody ; the fluctuations around that vertical drift are characterized by a variogram of the type (1o) previously discussed.

We can note on Figure 12 that the fluctuations (saw tooth aspect) of the borehole grades are not representative of the fluctuations of the panel mean grades.
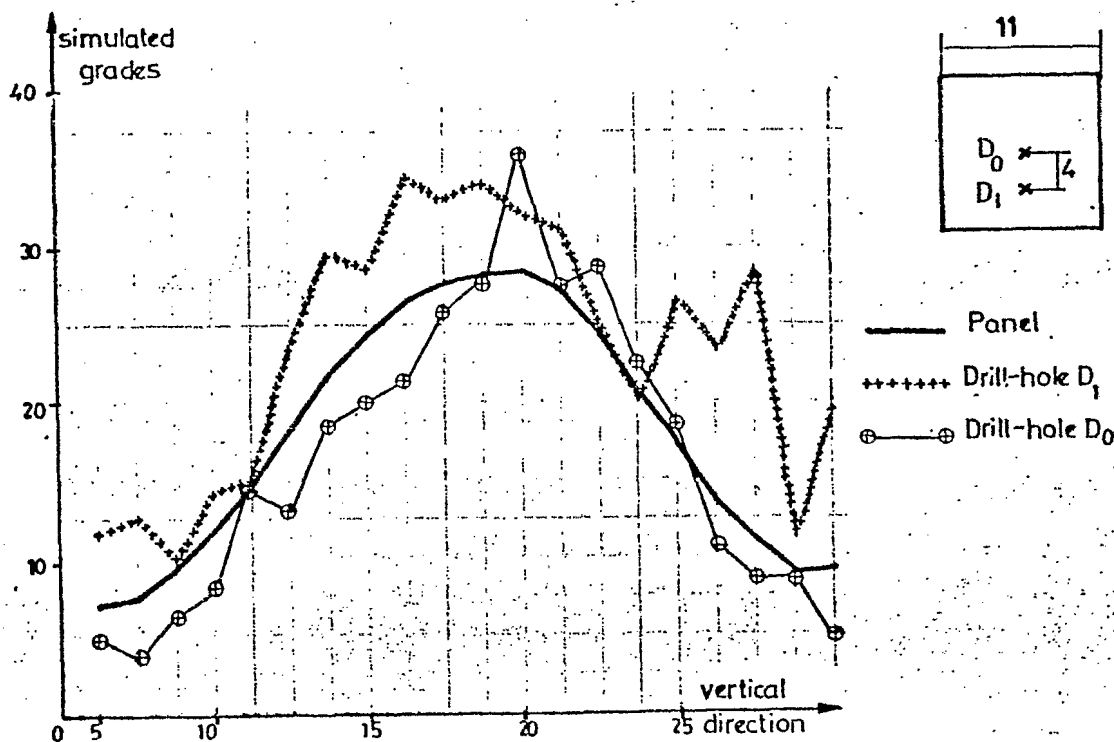

fig:12 Panel grades and drill-holes grades

VI - THIRD CASE STUDY - SIMULATION OF RAIN

This practical example is derived from the prominent study (ref. [6], 1973) of J.P. DELHOMME and P. DELFINER.

The desertic province of Enedi (Tchad, Central Africa) is being surveyed for water supply. The problem is to evaluate the quantity of water that a single rainfall can pour down on a limited area. Over the Kadjemeur area, a certain number of rain gauges (see their location on Figure 13) have recorded the water height of one rainfall. As rainfalls are not frequent in Enedi, it was necessary to simulate over the area various rainfalls similar to the real one recorded by the rain gauges. Thus, hydrogeologists would be able to evaluate the true dispersion of water heights over the area and at the same time critically reconsider the number and the location of their rain gauges. The study proceeded as follows :

1) An experimental semi-variogram was drawn from the rain-gauges data (water heights), and revealed to be more or less iso-tropic and linear : $\gamma(h) = a|h|$. Figure 13 shows the experimental points of that semi-variogram and the adjusted linear model. Figure 13 also shows the estimated map of the true rainfall kriged on the basis of the rain gauges data, as well as the corresponding map of the estimation standard deviations. It can be noted that the highest standard deviations are located on the right hand-side of the map, where rain gauges are rare. The shaded zones correspond to water heights superior to 40 millimeters.

2) Three conditional simulations of the variable water-height have been realized on a regular 1 km. spacing grid. These three si-mulated rainfalls reflect the previous auto-correlation function, i.e. the linear variogram, and meet the true experimental water heights at rain-gauges locations (see Figure 14).

It appears that :

- The true dispersion of water heights is much greater than the dispersion that could have been estimated from the kriged map of Figure 13. Studies of water runoff and recuperation must be done on the simulated maps and not on the estimated map.

- the three simulated rainfalls can differ considerably on high standard deviation zones, i.e. on the zones with few rain-gauges.
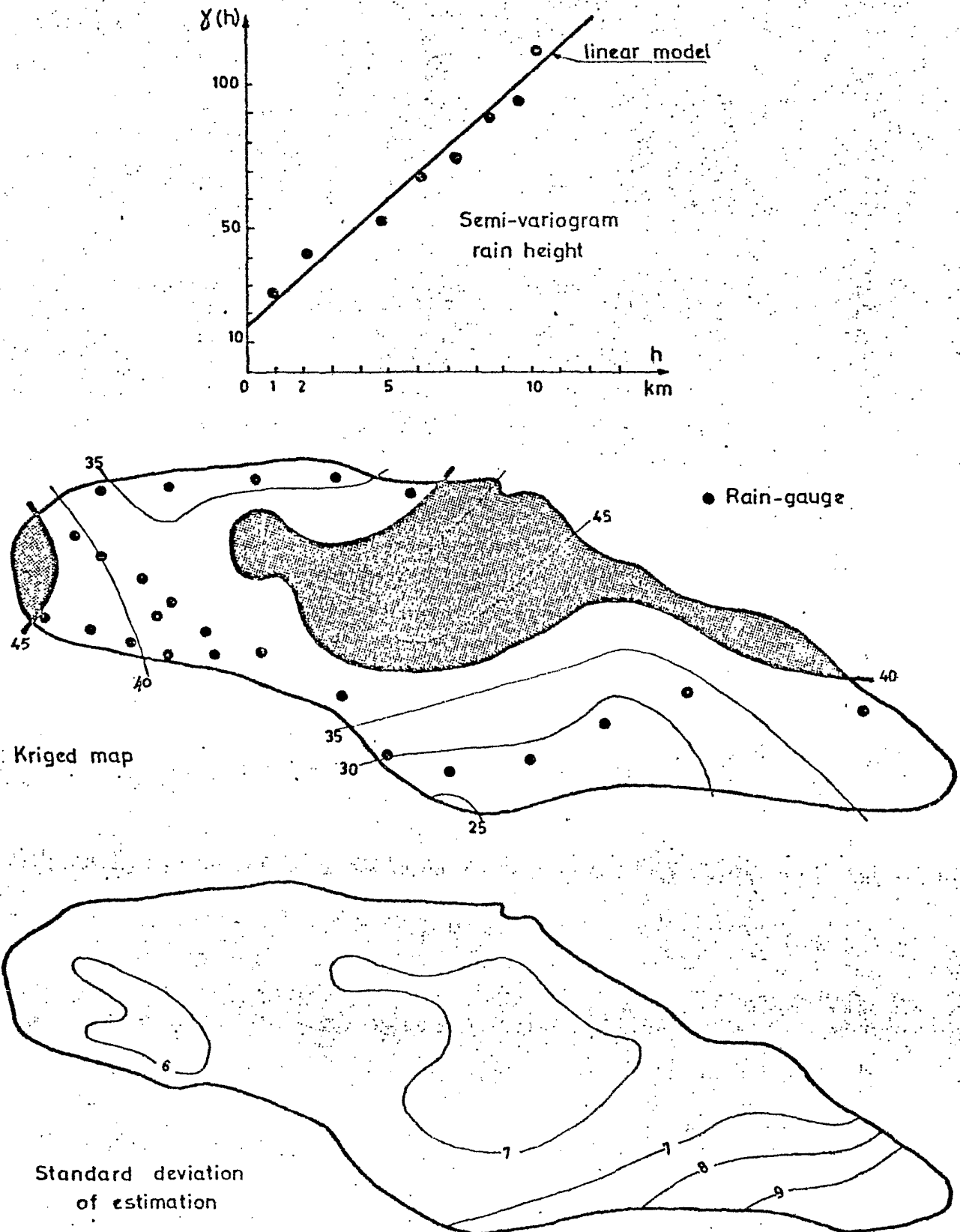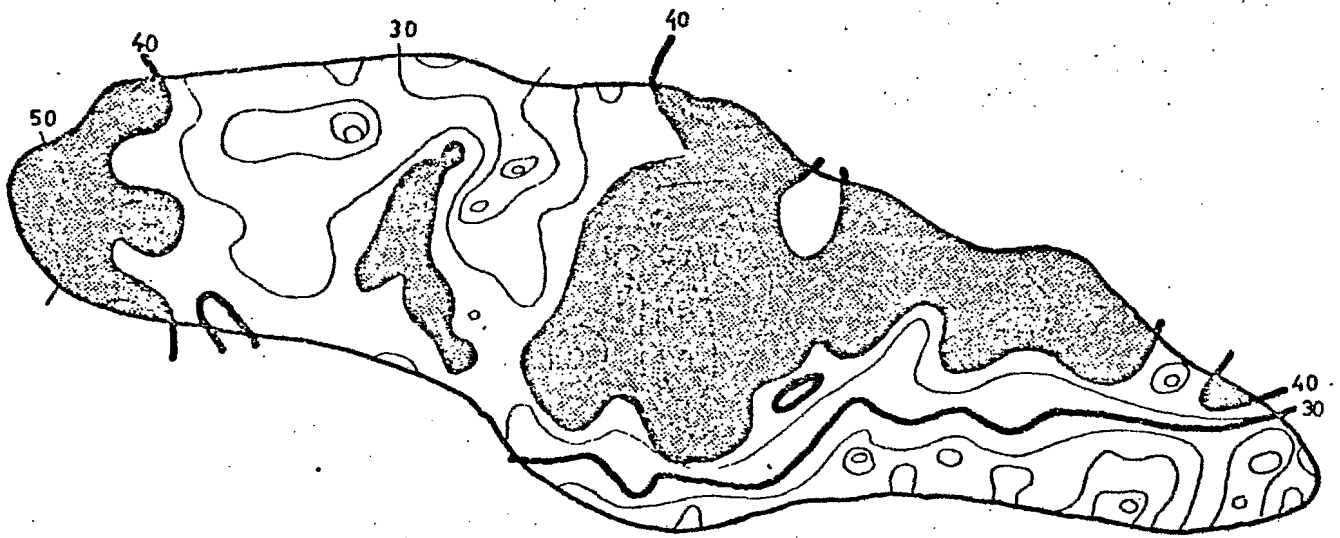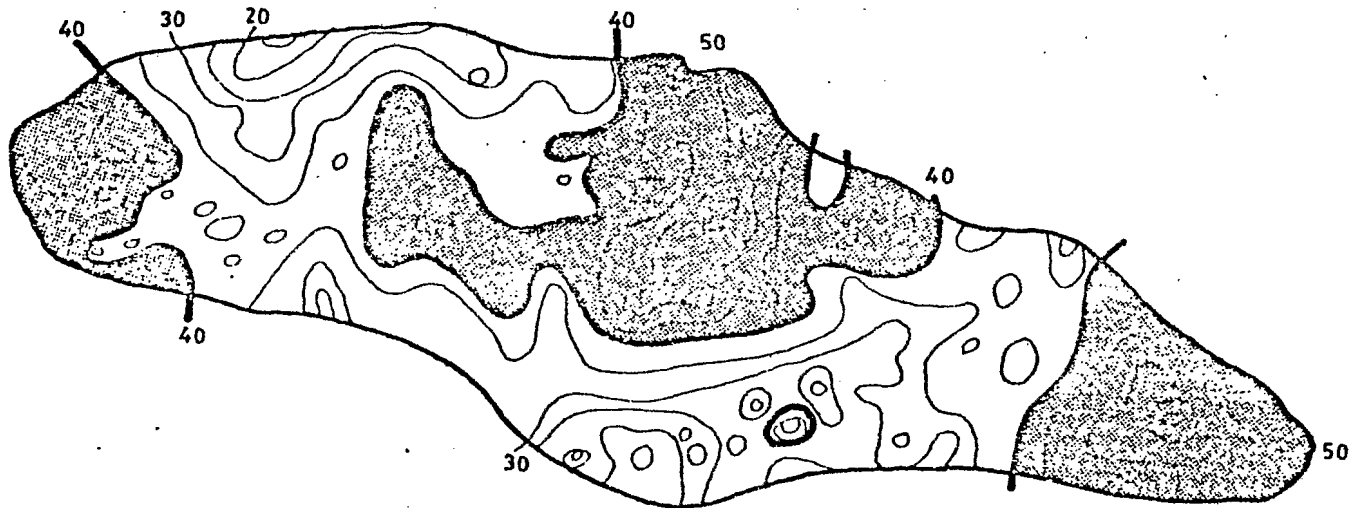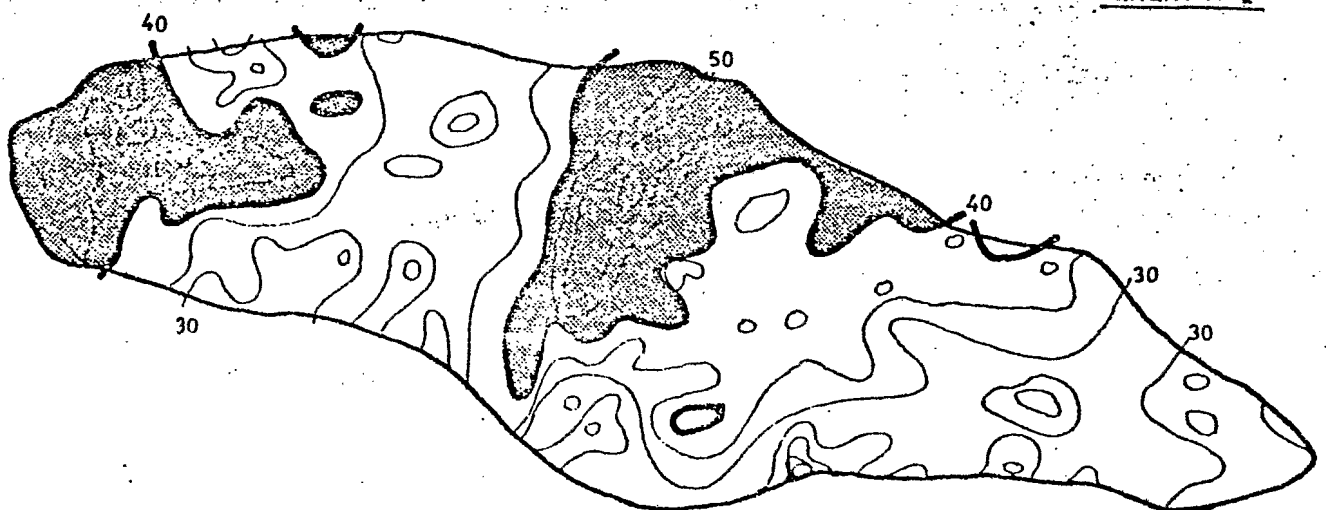
Semi-variogram rain height

Kriged map

Rain-gauge

Standard deviation of estimation

fig:13 Estimation of the showers

Variant Nº1

Variant Nº2

Variant Nº2

fig:14   3 showers simulated

CONCLUSIONS

For various problems, simulation of orebodies, and more gene-
rally simulations of regionalized phenomena, appear as the only
approach that does not require the expensive investments of very
precise systematic surveys or important bulk samples. The availa-
bility of high performance computers, and above all the originality
of the turning bands method that reduces any 3-dimensional simula-
tion to several 1-dimensional simulations, now allow the realiza-
tion of such 3-dimensional multi-variable simulations within reason-
able computer costs. The practical experience already gained by the
Centre de Morphologie Mathématique of Fontainebleau predicts consi-
derable developments of these Conditional Simulations in fields as
differents as : Mining industry : simulation of extraction, blending,
milling procedures etc...

- Meteorology - Hydrology : simulation of rainfalls, ground
waters, atmospheric disturbances etc...

- Gravimetry - Bathymetry : simulation of submarine surfaces,
marine survey lines, etc...

- Oil Industry : simulation of reservoirs

REFERENCES

Basic Theoretical Works :

[1] MATHERON, G.- The theory of regionalized variables and its applications (211 p).
  Ed. : Ecole Nationale Supérieure des Mines, Paris, 1971.

[2] MATHERON, G.- The intrinsic random functions and their applications.- Advances in Applied Probability - Dec. 1973.

[3] MATHERON, G.- Quelques aspects de la montée.
  Centre de Morphologie Mathématique, Unpublished Note N° 271, Jan. 1972.

[4] GUIBAL, D.- Simulation de schémas intrinsèques.
  Centre de Morphologie Mathématique, Unpublished Note N° 291, June 1972.

[5] JENKINS, G.M. and WATTS, D.G.- Spectral Analysis and its applications. 525 p.
  Ed. : Holden-Day, San Francisco, 1968.
  (A practical exposé designed for engineers).

Other References :

[6] DELFINER, P. et DELHOMME, J.P. - Optimum interpolation by kriging.
  Proceedings of NATO Advanced Study Institute for Display and Analysis of Spatial Data.- July 1973.
  Ed. : Wiley and Sons, London.

[7] HARBAUGH, J.W.- Computer simulation in Geology.
  Ed. : Wiley and Sons - New York, 1970.

[8] HARBAUGH, J.W. and MERRIAM, D.F.- Computer applications in stratigraphic analysis.
  Ed. : Wiley and Sons, New York, 1968.

[9] HUIJBREGTS, Ch. and JOURNEL, A.- Estimation of lateritic type deposit.
  10th APCOM Symp. pp. 207-212 - Ed. : South African Inst. of Min. and Metall., Johannesburg, 1972.

[10] HUIJBREGTS, Ch. and JOURNEL, A.- Simulation du gisement de nickel latéritique de Prony.
  Research Report, So. PENARAX and Centre de Morphologie Mathématique, Fontainebleau, 1972.

[11] JACOD, J. and JOATHON, P.- Conditional simulation of sedimentary cycles in 3 dimensions.
  in "Mathematical Models of Sedimentary Processes"
  Ed. : D.F. Merriam, Plenum Press, New York, 1972.

[12] JOURNEL, A.- Geostatistics and Sequential Exploration.
          in "Mining Engineering", October 1973.

[13] JOURNEL, A.- Le Formalisme des relations Ressources-Réserves -
          Simulations de gisements miniers.
          in "Revue de l'Industrie Minérale" N° 4, 1973.

[14] KIM, Y.C. and others - A short course on mine system simu-
          lation using a digital computer.
          Unpublished course - College of Mines, University
          of Arizona, Tucson, 1973.

[15] MARECHAL, A.- El problema de la curva Tonelaje - Ley de Corte
          y su estimacion.
          in Boletin de Geoestadistica, Vol. 1, 1972.
          Ed. : Universidad de Santiago de Chile.

SIMULATION OF MINING PROCESSES ON A MODEL DEPOSIT

Mining geostatistics has for a long time been concerned with the problems of evaluating in situ ore reserves. In the last few years geostatistics has also been applied to the field of ore processing. The problem is to study the effects of various proposed mining methods or mineralurgical processes on the ore under treatment. The central idea is to simulate the process in question on a model of the deposit. After the statistical characteristics of the ore in situ (i.e. its mean and its variability) have been reproduced in the model, the real ore and the ore in the model will undergo exactly the same changes as the process under study is carried out. Consequently the simulation of mineral processes on a model enables us to predict the effect of the process on the ore. These simulations therefore provide an additional method for studying mining projects which can assist the decision makers in choosing the method best suited to the deposit under study. Several projects can easily be evaluated without actually having to put them into operation. The methods developed by " downstream geostatistics" provide a new approach to the problems of ore selection, homogenization and mine planning, which take account of the technical constraints inherent in a particular mining method, of the type of production equipment used, of the way in which the ore is being processed, and so on.

Before the simulation itself can be commenced, 4 preliminary operations must be carried out :

- firstly, to determine the parameters of the problem ; that is, the technical constraints imposed by the type of mineralization, the factors known to influence the results, the technical and economic criteria used to judge the quality of the results and to compare different processes.

- secondly, to take account of the amount of information which is likely to be available to the decision makers at the time when the final selection and mining planning will be carried

out. In the case where this information will not all be available
when the processing commences, this information will be replaced
by simulated values in order to duplicate the real conditions.

- thirdly, to build a numerical model of the deposit. Conditional
  simulations provide just such a model.

- finally, to model the process which one wishes to simulate ; that
  is, to provide a realistic representation which can be programmed
  and can give the characteristics of the ore at the end of the pro-
  cessing for specified initial conditions.

The simulation is carried out from the point of view of the en-
gineer, who must make the day-to-day decisions concerning the pro-
duction. The proposed method of simulation does not attempt to use
a computer program to optimize the output, firstly because these
programs tend to be cumbersome and time-consuming, and secondly
because it is not always possible to unite such a program given the
number of factors which have to be taken into account. Rather than
attempt this, a mini-computer with a conversational mode is used to
simulate the decisions made by the production engineer. Several fea-
sible decisions can be tested with the computer, providing the end
results of these decisions. The process which best satisfies the
management's objectives can then be chosen.

At this point an example of a simulation of an open pit mine
is presented. The text may be found in "Mining Geostatistics", by
A. Journel, pp. 688-698, to be published by Academic Press.

Case study : Choice of a mining method from a numerical* model
of a deposit. after Deraisme, J. (1977).

* To avoid confusion, the term "simulation" will be used
here only for the simulation of the various methods of mining
and blending ; the simulated deposit will be referred to as
the numerical model.

The numerical model.

The deposit model used in this study does not represent an
actual deposit, but rather a type of deposit having characteris-
tics similar to the homogeneous zone in a porphyry copper depo-
sit (e.g. the sulphide zone).

The zone G of homogeneous mineralization considered here
has horizontal dimensions 450 x 450 m^2 and consists of 4 levels
of parallelepipedic blocks v with a square horizontal section
18 x 18 m^2 and a 5 m. height. The numerical model consists of
the true grades of the blocks and the grades of vertical drill
intersections centred on these blocks. The following characteris-
tics were imposed on the model :

~ mean grade : m = 1% Cu.

~ dispersion variance of the true block grades : $D^2(v/G)$ =
$0.235(\%)^2$, i.e. a relative standard deviation of 48% : D(v/G) =
0.48.

~ a strongly asymmetrical histogram of the true block grades;
as shown on Figure VII-19.

~ an isotropic regionalization of the grades characterized by
a spherical model with nugget effect, and a range of approxima-
tely 70 m.

Figure VII-19 also shows the histogram of the block values
kriged from the central drill core intersections on a horizontal
18 x 18 m. grid. This histogram of kriged values has a mean
m* = 1% Cu and a variance $D_K^{2*}(v/G) = 0.185(\%)^2$; the kriging
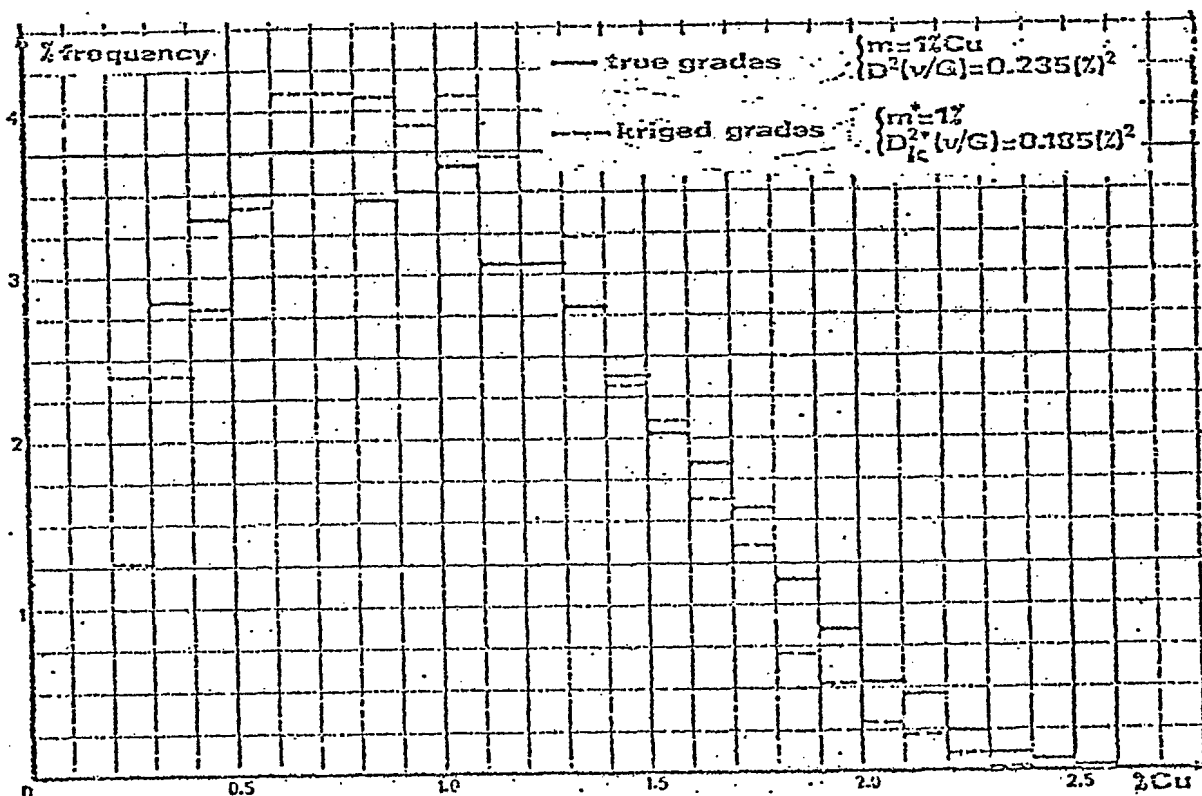variance is : $\sigma_K^2 = 0.032(\%)^2$.

Fig. VII-19 — Histograms of true and kriged block grades.

Outline of problems.

The minimum mining units are blocks $18 \times 18 \times 5 \text{ m}^3$, each containing 4730 tons of ore. The mining rate is fixed at 35000 tons/day which is equivalent to 8 blocks. A year's production consists of 280 days. The mill is planned to treat this daily production. The metal recovery is a function of the mill capacity to absorb fluctuations in the feed grade. Automatic process control of the milling operation is a possible solution, but it is expensive and not adapted to all scales of variability, since it is usually designed to respond to small scale (hourly) and not to large scale fluctuations (daily – fortnightly).

In this study, the period of overall feed grade regulation was considered to be two weeks. The miner predicts the quality of the ore to be sent to the mill during the next fortnight and during this period, it is the miner's responsibility to ensure that the fluctuations in the daily mean grade are not significantly different from his prediction. The short term (e.g. hourly) fluctuations of the mill feed grade are the responsibility of the metallurgist and the process controls.

It is assumed here that grade is the primary variable affecting mill operation. Any other regionalized parameter (impurity grades, crushing indexes, etc...) can also be simulated and used in an analogous manner.

Projected mining methods.

Three different mining methods were considered to meet the required degree of homogeneization of daily mean grades over a fortnight production. These methods are all non-selective : every block is eventually sent to the mill.

Method 1 : Zone G is mined by two 10 m. high benches, cf. Figure VII-20-a. From each bench, a big shovel of limited manoeuvrability can extract two contiguous $18 \times 18 \times 10$ m^3 blocks per day. The two shovels advance parallel to the mining front and cannot reverse : they must extract all the ore along any given front ($18 \times 25 = 450$ m. long) before beginning on the next front. All extracted ore is sent directly to the mill. Since the motion of the shovels is pre-determined in this method, there is no choice for the mining engineer, who must relinquish his part in the homogeneization process.

Method 1'. Mining is identical to the previous method, but some of the mined ore can be sent to a stockpile where homogeneization can take place, cf. Figure VII-20-b.

The stockpile consists of two sub-piles, one rich, the other poor. The maximum capacity of the total stockpile is a 2-day's production (70000 tons), which is the equivalent of 8 large blocks. A block will be sent to the stockpile when there is a risk that its mean grade affects the day's mean feed grade. When a block

is sent to the rich subpile, the equivalent of a block is taken away from the poor subpile to meet the daily required tonnage (35000 tons).

In addition, the existence of the stockpile provides some degree of flexibility to the mining operation, since on certain days, three blocks may be extracted from one bench and one on the other, provided that equilibrium is restored within 8 days ; alternatively, 5 blocks may be extracted within any given day, provided that this over-production is balanced by an equivalent under-production within the next 8 days. But mining is expected to adhere to the norm as often as possible, i.e. daily production of 2 blocks per bench.
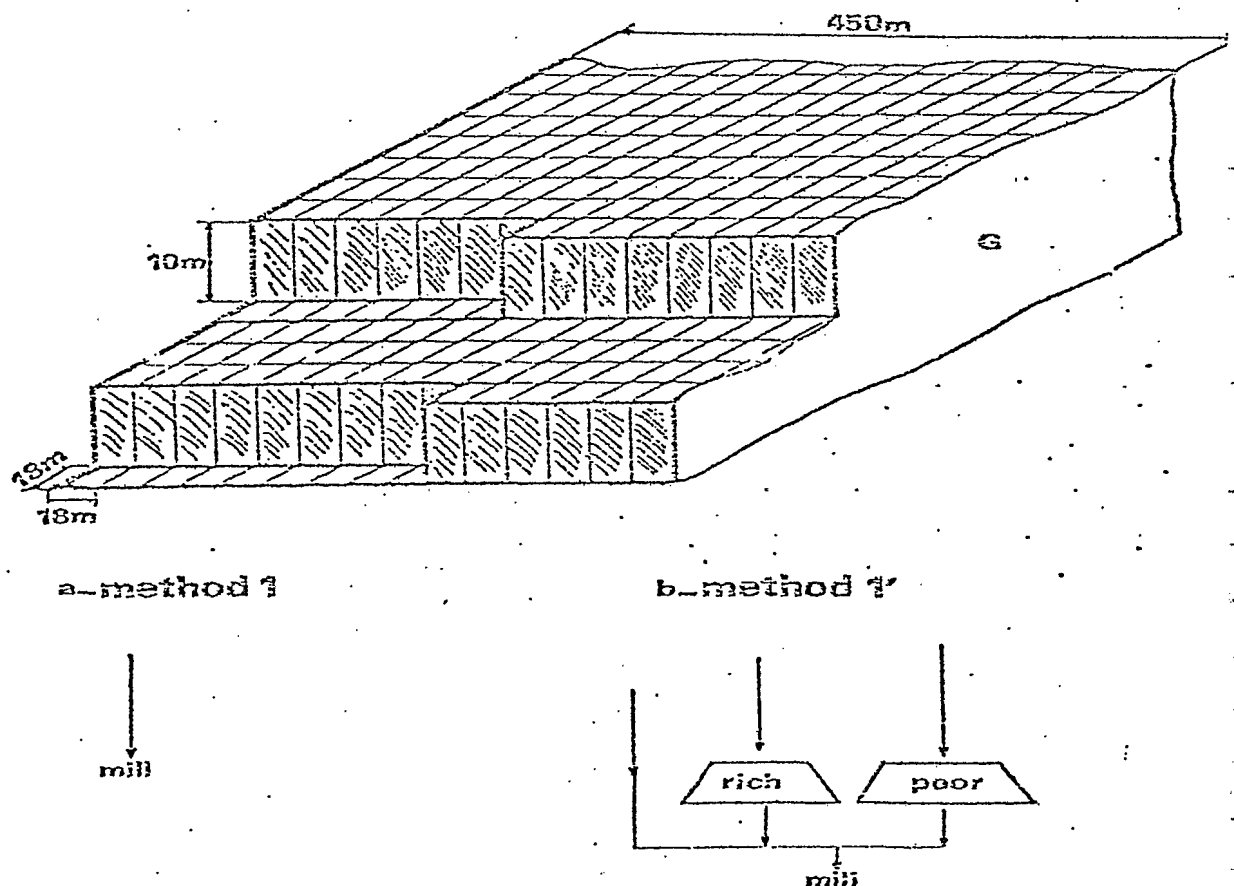

a—method 1

b—method 1'

Fig. VII-20 – View of the mining fronts (blocks 10 m. high)

Method 2 : The same zone G is mined by four 5 m. high benches, cf. Figure VII-21. From each of these benches, a mobile shovel can extract two small 18 x 18 x 5 m³ blocks per day. The resulting daily production of 35000 tons is sent directly to the mill.

The mobility of the shovels allows the choice of the 8 blocks to be extracted, to best stabilize the daily mean grades over the fortnight. All the ore of any given mining front (450 m. long) must be extracted before beginning on the next, and the total moving distance of each shovel is to be kept to a minimum.

The homogenization obtained by blending the production coming from the four working levels will result in higher mining costs.
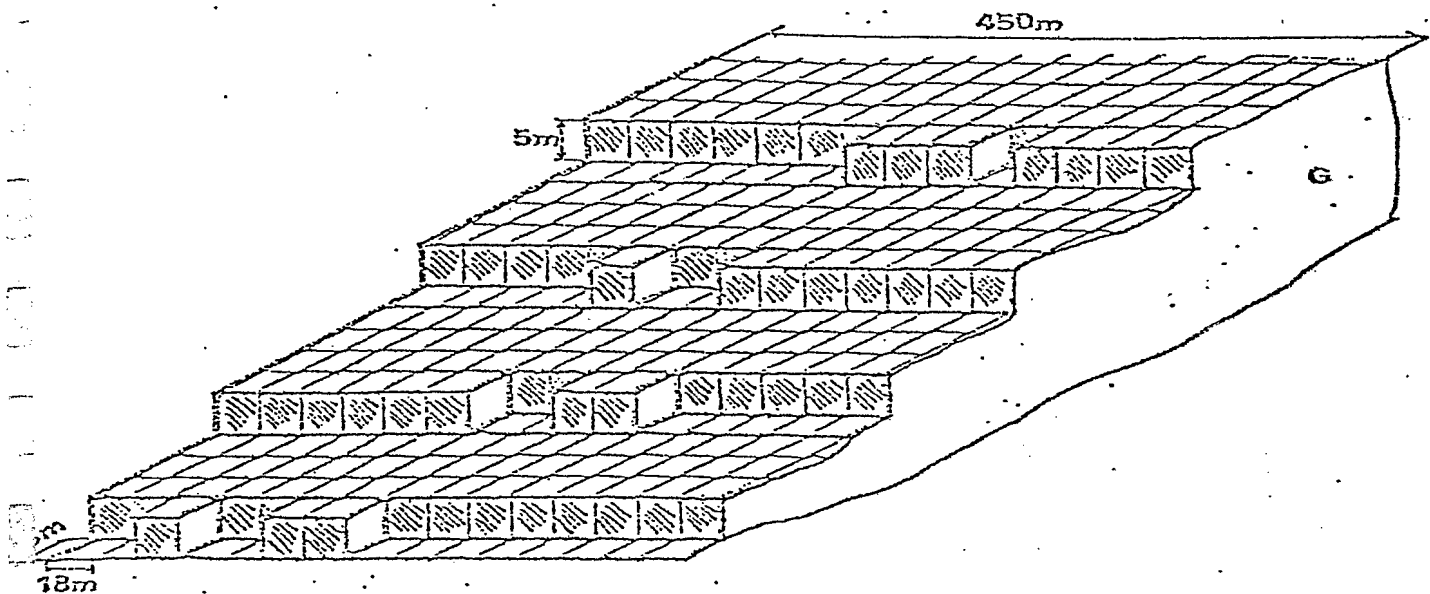

Fig. VII-21 — View of the mining fronts (5 m. high blocks)

Simulation of the mining processes.

Only method 1, for which the path of the big shovels is entirely determined, is completely simulated by computer. For the methods, 1' and 2, the daily decisions such as which block to mine, and whether to send it to the mill or to the stockpile, are made by an operating mining engineer, just as they would be in a real situation. The decisions were made through visualization of the state of the fronts and the stockpiles, taking the past and immediate future mill feed grades into account. The decision of the mining engineer is certainly not always optimal, but the object of this study is precisely to reproduce reality (including, in particular, human errors) and not an absolute optimum which is often inaccessible in practice.

The engineer makes all his decisions on the basis of kriged estimates of the mining blocks, the blocks being kriged from the central drill core intersections. However, it is the true block grades that are delivered to the mill, cf. section VI-A-3.

As for the type of blending performed within the subpiles considered in method 1', it is assumed that the true mean grade of the unit taken from a subpile is a normally distributed random variable, with :

~ an expectation m_S equal to the true mean grade of the subpile at the time considered. This instant mean grade varies around 1.6% Cu for the rich subpile, and 0.6% Cu for the poor subpile.

~ a fixed relative standard deviation σ_S/m_S equal to $0.25(\%)^2$ for the rich subpile and $0.3(\%)^2$ for the poor. These values were chosen so that the confidence interval $[m_S \pm 2 \sigma_S]$ for the rich subpile for example, includes all individual true block grades sent to this particular subpile during the course of one year.

This model of the blending effect of the stockpile is very simplistic and pessimistic, but is enough to give an order of magnitude of the influence of the stockpile on the homogenization of the mill feed grades. In a real case, it would be a

simple matter to replace this model either by a deterministic
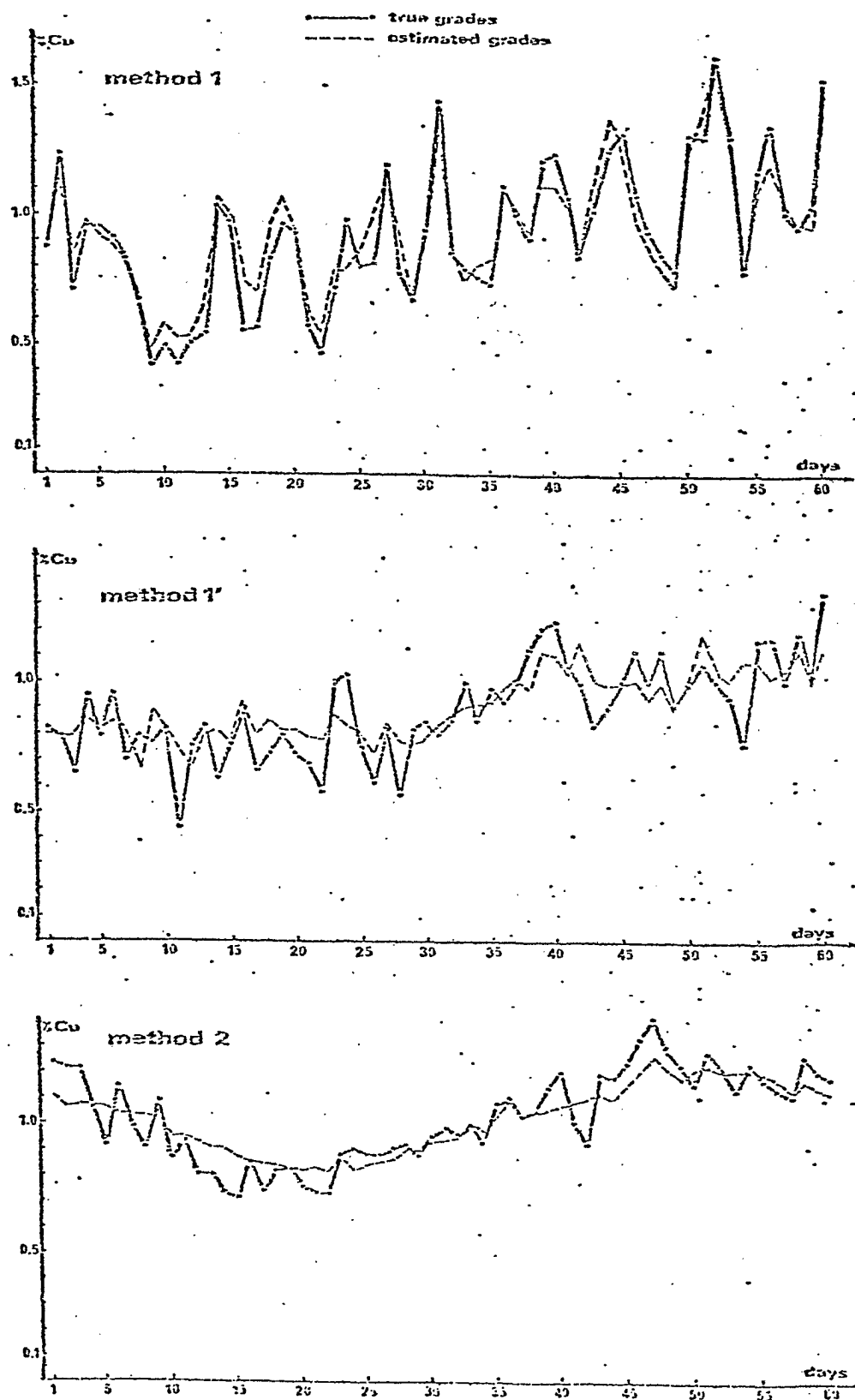system or a more precise probabilistic model.

Fig. VII-22 — Fluctuations of the daily mean grades
at the mill-feed

Results.

For the 3 simulated methods of mining zone G, Figure VII-22 shows the variation in daily mean grades during the first two months (60 production days). J. Deraisme's study was carried out for a year's production (280 days) but there is not enough space here to show all the results.

The full line curves show the variations of the true daily grades (i.e. the homogenization actually achieved), while the broken line curves represent the variation of the estimated daily grades (i.e. the homogenization predicted on the basis of the kriged block grades).

The mill operator will obviously prefer methods 1' and 2 to method 1, which the mining engineer may prefer because of its lower mining cost. Thus there is a real problem involved, which can only be solved by assigning costs to the mill recovery losses caused by method 1, and to balance these costs with the ones caused by stockpiling (method 1') or by the more flexible mining operation of method 2.

Table VII-23 gives the different dispersion variances observed at different scales, and thus gives an overall picture of the effect of each mining method on the fluctuations of the mill-feed grades.

	Method 1	Method 1'	Method 2
Dispersion variance 1 day within 7 days	0.0452 0.0362	0.00115 0.0040	0.0074 0.0018
Dispersion variance 1 day within 15 days	0.0501 0.0422	0.0136 0.0052	0.0082 0.0016
Dispersion variance 1 day within 1 month	0.0525 0.0474	0.0156 0.0070	0.0106 0.0038
Dispersion variance 1 day within 1 year	0.0816	0.0405	0.0310
Dispersion variance 15 days within 1 year	0.0260	0.0249	0.0227

Table VII-23 - Dispersion variances for the 3 mining methods.
(when there are two values, the upper one re-
fers to the true grades while the lower refers
to estimates).

The following remarks can be made about these results :
- There is a considerable variation in the fluctuations of the
daily grades sent to the mill, depending on the mining method
used. The dispersion variance of the daily grades in fifteen
days for method 1 is six times greater than for method 2. These
orders of magnitude of the mill-feed grades variabilities at
various scales can be predicted only by simulation. Standard
estimation procedures (including kriging) are irrelevant for
predicting these variabilities, for they take no account of
the destructuring effect of mining, trucking, stockpiling, etc..

- The deviation between the dispersions resulting from the different mining methods increases as the time interval decreases. Thus, whatever method is used, the amount and location of the ore extracted during a given fortnight will be approximately constant, and the dispersion variances of a fortnight's mean grade over one year will also be approximately the same. On the other hand, methods 1' and 2 are particularly efficient in stabilizing the daily grades within any week or fortnight.

- The dispersions of estimated grades are always less than those of true grades. This is a result of the classical smoothing property of the kriging estimator, cf. relation (VI-1 or 2) : the dispersion of estimators is not a good estimate of the real dispersion. For method 2, even though the homogenization of the estimated grades is nearly perfect, there is still an irreducible fluctuation of the true grades, due to the estimation variance of the daily grades, cf. Figure VII-22.

The experimental estimation variances of the daily grades for each method are :

method 1 : $\sigma_E^2 = 0.0089\ (\%)^2$

method 1' : $\sigma_E^2 = 0.0131\ (\%)^2$

Method 2 : $\sigma_E^2 = 0.0061\ (\%)^2$

Note that for methods 1 and 2, the dispersion variance D^2 of the true grades can be deduced from that of the estimated grades D_K^{2*} by the approximate smoothing relation (VI-2) :

$$D^2 \simeq D_K^{2*} + \sigma_E^2$$

This relation is not applicable for method 1' because daily production does not only come from the mine but also from the stockpile.

- For one year of simulated production, 17% of the production from method 1' passed through the stockpile. This represents a

considerable proportion for a stockpile with a maximum capacity of 2 days production. In fact, to obtain the homogenization shown on Figure VII-22, almost 50% of production days involved departure from normal production (2 blocks per bench). One solution would be to increase the capacity of the stockpile by, for example, adding a third subpile of average quality ore. In any case, it would appear that the capacity of the stockpile should not be greater than 3 or 4 days' production. Method 1' produces dispersions of the same order of magnitude as those of method 2 and since its mining cost is lower (10 m. benches instead of 5 m.) a feasibility study would probably result in method 1' being chosen.

- Globally, there is no selection involved in any of the methods considered in this study. However, in his more complete study, Deraisme (1977) studied various selection criteria for method 2 and compared the results of applying these criteria on the simulations to the results that could be predicted on the basis of grade histograms alone. The differences were significant, and illustrate the well-known effects of concentration and dilution of reserves : a rich block with a grade above the cut-off will not be selected when it is located in an overall poor zone ; conversely, a poor block with a grade below the cut-off will be selected when located in an overall rich zone. The prediction made on the basis of the grade histogram (cf. the various techniques outlined in section VI-A) are unable to take account of this spatial localization of rich and poor grades ; more generally, they cannot take the technological conditions of the mining operation into account. Once again, a solution is afforded by the simulation of the mining operation on a numerical model of the deposit : the simulations can be used to evaluate the effect of these technological conditions on the actual recovery of the reserves.

Any mining or mineralurgical process can be simulated in this way ; for example, extraction of the ore, haulage, stockpiling, processing... Several other studies (see J. Deraisme : "Simulations sur modèles de processus miniers et minéralurgiques". Thèse de Docteur-Ingénieur, 1978) have been carried out. The principal conclusions of the studies were :

- the optimal size of a stockpile can be determined after only a few iterations by these simulations.

- by dividing the model of the deposit into sufficiently fine blocks it is possible to simulate the operation of an ore stockpile used for homogenizing plant feed.

- day-to-day mine planning decisions can be tested by these simulations and then put into operation by using a minicomputer in a conversational mode.

At present the work in progress to make "geostatistical mineral processing" a practical and useful tool in the decision-making process for mining projects, is being carried in 4 main areas :

- a diversification of the type of deposits studied.

- more parameters are being taken into account (in particular, the granulometric properties of the ore at the end of the crushing process).

- the simulation of the complete production cycle from extraction up to the concentration of the metal.

- the improvement of the simulations, particularly from the point of view of the visual presentation of the results.