SOCIETE
DE
L'INDUSTRIE MINERALE
SECTION
DE MINERALURGIE

Présentation des variables régionalisées

par G. MATHERON

Ingénieur au corps des Mines
Docteur ès Sciences
Ecole Nationale Supérieure des Mines de Paris

I. INTRODUCTION

Variables régionalisées

Un grand nombre de phénomènes naturels se présentent à l'homme sous forme régionalisée : ils se déploient ou se distribuent dans l'espace. De tels phénomènes peuvent se caractériser, localement, par certaines grandeurs qui varient dans l'espace, et constituent, par conséquent, des fonctions numériques (ordinaires). Ce sont de telles fonctions numériques que nous appelons des variables régionalisées : il s'agit là d'un terme neutre, purement descriptif, antérieur en particulier à toute interprétation probabiliste.

Comme exemples simples de variables régionalisées, citons entre autres :

— la teneur, dans un gisement minier,

— le rendement à l'hectare, dans une région céréalière,

— la densité de population, à l'intérieur d'un pays.

Dans ce qui suit, j'emprunterai surtout mes exemples aux sciences de la terre et à l'estimation des gisements miniers, domaines où j'ai surtout travaillé, mais le champ d'application est évidemment plus général.

Caractère mixte aléatoire structuré

Le plus souvent, les variables régionalisées présentent un haut degré d'irrégularité. Elles ne sont généralement pas dérivables, ni même continues. On ne peut les représenter graphiquement que d'une manière grossière et approximative, sous forme de courbes en « dents de scie ». L'étude directe de courbes de ce genre serait prohibitive, et ne présenterait sans doute qu'un intérêt limité, en raison de leur complication même. Il y a dans le comportement des variables régionalisées un *aspect aléatoire*, qui suggère presque irrésistiblement le recours à une interprétation probabiliste.

Cet aspect aléatoire permet de comprendre l'insuffisance des méthodes traditonnelles d'estimation des gisements miniers. Les mineurs admettent que chaque échantillon prélevé représente la teneur réelle de sa zone d'influence. Ils obtiennent bien ainsi une estimation globale utilisable, mais ne peuvent évidemment pas l'assortir d'un calcul d'erreur (d'une variance d'estimation). Il est clair qu'entre l'échantillon et sa zone d'influence, il n'y a pas identité, mais seulement corrélation, d'autant meilleure que la zone d'influence est plus petite et que la minéralisation est elle-même plus continue.

De même encore, les méthodes d'interpolation fonctionnelle surestiment, en général, de façon inadmissible, le degré de continuité du phénomène représenté. Par quatorze points expérimentaux, on peut toujours faire passer un polynome du treizième degré. Mais, en général, ce polynome ne reflète pas le moins du monde l'évolution réelle du phénomène entre les points expérimentaux.

Et cependant, sous la complication et l'irrégularité extrême que présente une régionali-

sation, dans sa variation spatiale se dissimule, en général, *la structure d'un phénomène naturel*. C'est de cette structure spatiale qu'à leur tour ne peuvent pas rendre compte les méthodes purement statistiques. Quand on classe les échantillons sous forme d'histogramme, on fait, par là même, abstraction de l'endroit où ils ont été prélevés : on détruit les structures spatiales, qui constituent justement l'aspect le plus important du phénomène. Dans un gisement minier, par exemple, il ne suffit pas de savoir avec quelle fréquence se répète une teneur donnée. On a besoin, aussi, d'apprécier le degré de continuité de la minéralisation, de connaître la taille et la position des zones exploitables, etc...

Ce n'est pas là seulement une vue de l'esprit. Il est clair que l'erreur que l'on commet lorsque l'on estime les réserves d'un gisement à partir d'un réseau donné de travaux de reconnaissance, dépend au plus haut point du degré de continuité de la minéralisation. Ainsi, la *continuité* d'une minéralisation constitue l'un de ces caractères *structuraux* dont l'histogramme statistique est incapable de rendre compte, et qui doivent cependant impérativement être pris en charge par une théorie réaliste des variables régionaliseés. Il y a bien d'autres caractères analogues, nous les mentionnerons au fur et à mesure que nous les rencontrerons.

Le langage des fonctions aléatoires

Il fallait donc adopter un mode de formulation synthétique, capable de rendre compte à la fois de ces deux caractères contradictoires des variables régionalisées, de leur aspect à la fois aléatoire et structuré, et permettant bien entendu aussi de poser correctement et de résoudre des problèmes essentiellement pratiques, comme celui de l'erreur que l'on commet lorsque l'on estime une variable régionalisée à partir d'un échantillon fragmentaire.

Ce langage adequat — à la fois probabiliste et structuraliste — ce sera celui des *fonctions aléatoires*. Mais il se posera un problème méthodologique assez grave, sur lequel nous reviendrons ci-dessous. Ce langage des fonctions aléatoires implique, en effet, certaines hypothèses fondamentales, de nature probabiliste, et nous devrons procéder à un examen critique de leur sens et de leur valeur. Autrement dit, nous devrons examiner si, en parlant de fonctions aléatoires, nous disons réellement quelque chose de sensé, ou bien si, par hasard, nous ne parlerions pas pour ne rien dire.

Avant, donc, d'introduire et d'examiner ces hypothèses probabilistes, il est sage et utile de regarder au préalable jusqu'où il est possible d'aller sans introduire de telles hypothèses, c'est-à-dire en utilisant uniquement des méthodes de nature purement géométrique (non probabiliste). En fait, nous allons le voir, il est possible d'aller assez loin et, au fond, on pourrait fonder la géostatistique en se passant complètement d'hypothèses probabilistes.

Nous exposerons donc, en premier lieu, ces méthodes géométriques, que l'on appelle *méthodes transitives*, et nous consacrerons une deuxième partie à leur correspondant probabiliste. Il se trouve que — par une convergence remarquable — les résultats ultimes auxquels conduisent ces deux groupes de méthodes sont pratiquement équivalents. C'est là une circonstance très rassurante sur le plan méthodologique. Elle signifie, en effet, que la part d'arbitraire qu'entraîne inévitablement l'introduction d'hypothèses probabilistes est en fait très réduite.

II. LES METHODES TRANSITIVES

Nous examinerons d'abord le cas particulier, de nature purement géométrique, où il s'agit simplement de décrire de manière adéquate la forme d'une surface ou d'un volume. Nous introduirons ensuite la notion générale de *covariogramme transitif*, et nous examinerons sa signification structurale. Nous montrerons ensuite que cet outil minimal permet à lui seul de résoudre le problème de l'estimation.

Le covariogramme géométrique

Considérons le phénomène de transition le plus simple que l'on puisse imaginer : présence ou absence d'un caractère. Soit, par exemple, une formation géologique S d'extension limitée. Un sondage foré au point x la rencontre, ou ne la rencontre pas. Posons :

$$k\ (x) = \begin{cases} 1 \text{ si } x \in S \\ 0 \text{ si } x \notin S \end{cases}$$

Il s'agit d'un phénomène unique, pour lequel aucune formulation probabiliste n'est possible. Parler de la probabilité pour qu'un point x donné appartienne à S n'aurait pas grand sens. On peut noter que tout l'intérêt se concentre sur la *frontière* de S. En effet, $k\ (x)$ est constante à l'intérieur comme à l'extérieur de S, et c'est au franchissement de cette frontière seulement que $k\ (x)$ varie, passant de 1 à 0, ou de 0 à 1. De là le nom de phénomène de transition et de méthodes transitives.

L'aire S de notre formation est évidemment donnée par l'intégrale

$$S = \int k\ (x)\ dx$$

La valeur de S constitue un renseignement fort intéressant du point de vue pratique. Le plus souvent, c'est elle que l'on cherche à estimer, à partir d'un réseau de sondages. Toutefois, elle n'apporte encore aucune information de nature réellement structurale.

On peut, en effet, définir la structure d'un ensemble comme le système des relations existant entre les éléments ou les parties de cet ensemble. Ainsi, nous n'aurons d'information de nature structurale sur notre surface S qu'à la condition de faire intervenir au moins deux points.

Soient donc x et $x + h$ deux points, c'est-à-dire le plus petit ensemble structurant que nous puissions imaginer. Considérons alors l'expression $k\ (x)\ k\ (x + h)$. Elle vaut 1 si x et $x + h$ appartiennent tous les deux à S, et 0 autrement. Mais dire que $x + h$ appartient à S, c'est dire que x appartient au translaté S_{-h} de S dans la translation $-h$. Ainsi, on a :

$$k\ (x)\ k\ (x + h) = \begin{cases} 1 \text{ si } x \in S \cap S_{-h} \\ 0 \text{ si } x \notin S \cap S_{-h} \end{cases}$$

Si nous intégrons cette expression en x, nous obtenons une fonction de h :

$$(1)\ K\ (h) = \int k\ (x)\ k\ (x + h)\ dx = Mes\ S \cap S_{-h}$$

qui représente la mesure (l'aire) de l'intersection de S et de son translaté par $-h$. Cette fonction est *symétrique*, puisque les deux intersections $S \cap S_{-h}$ et $S \cap S_h$ se déduisent l'une de l'autre par translation. On a donc :

$$\begin{cases} K\ (h) = K\ (-h) \\ K\ (h) \leq K\ (0) \end{cases}$$

Cette fonction $K\ (h)$ est le *covariogramme géométrique* associé à S. Il donne une certaine image de la forme de l'ensemble S. En effet :

1 — Tout d'abord, la mesure de S est $K\ (0)$, comme on le voit en faisant $h = 0$ dans (1). De même, si l'on intègre (1) en h, on obtient le carré S^2 de la mesure de S :

$$\begin{cases} K\ (0) = S \\ \int K\ (h)\ dh = S^2 \end{cases}$$

2 — Si l'on désigne par $r = |\ h\ |$ et par α le module et la direction du vecteur h, $K\ (h)$ est une fonction de la forme $K\ (r\ ;\ \alpha)$. Dans chaque direction α, la fonction $K\ (r\ ;\ \alpha)$ est identiquement nulle lorsque r devient supérieur à une valeur limite $a\ (\alpha)$ que nous appellerons *portée* du covariogramme $K\ (h)$. La portée représente la plus grande dimension de S mesurée dans la direction α. Pour $|\ h\ | > a\ (\alpha)$, l'intersection $S \cap S_{-h}$ est vide. Naturellement, la portée $a\ (\alpha)$ dépend de la direction α, et la manière dont elle se modifie lorsqu' α varie donne une image des propriétés d'anisotropie de la surface S.

3 — Le comportement du covariogramme au voisinage de l'origine présente lui aussi un grand intérêt. Il est toujours *linéaire*. En effet, si $\delta\ r$ est très petit, l'aire balayée par le petit vecteur de module $\delta\ r$ et de direction α dont l'origine décrit le contour de S vaut :

$$2\ [K\ (0) - K\ (\delta\ r\ ;\ \alpha)]$$

Mais, si nous désignons par $2\ D_\alpha$ la somme des projections sur un axe perpendiculaire à α des éléments d'arc du contour de S, cette aire est aussi égale à $2\ D_\alpha\ \delta\ r$. On a donc :

$$K\ (0) - K\ (\delta\ r\ ;\ \alpha) = D_\alpha\ \delta\ r$$

c'est-à-dire :

$$-K'\ (0\ ;\ \alpha) = D_\alpha$$

Dans le cas où S est convexe, D_α est la longueur du contour apparent de S en projection sur l'axe perpendiculaire à α (*figure* 1) et cette importante caractéristique détermine la pente à l'origine du covariogramme.

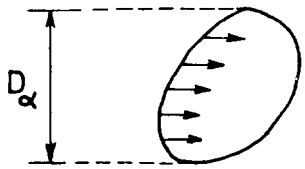

Fig. 1

Le covariogramme transitif

Plus généralement, au lieu d'une fonction $k\ (x)$ susceptible de prendre les seules valeurs 0 et 1, on peut considérer une fonction $f\ (x)$ représentant une variable régionalisée quelconque. Nous supposerons seulement que $f\ (x)$ n'est différente de 0 que dans un domaine borné. A cette fonction $f\ (x)$, nous associerons son *covariogramme* $g\ (h)$ défini par :

$$(2) \qquad g\ (h) = \int f\ (x)\ f\ (x + h)\ dx$$

Cette fonction $g\ (h)$ est symétrique et vérifie les relations :

$$\begin{cases} g\ (-h) = g\ (h) \\ g\ (h) \leq g\ (0) \end{cases}$$

Exactement comme son homologue $K\ (h)$, $g\ (h)$ reflète certaines propriété structurales de la régionalisation :

1 — Désignons par $Q = \int f\ (x)\ dx$ l'intégrale de $f\ (x)$ — la « quantité de métal ». En intégrant (2) en h, on obtient :

$$\int g\ (h)\ dh = Q^2$$

Par contre, la valeur $g\ (0)$ de $g\ (h)$ en $h = 0$ n'est pas égale à Q, mais à l'intégrale du carré de $f\ (x)$:

$$g\ (0) = \int [f\ (x)]^2\ dx$$

2 — A chaque direction correspond une portée $a\ (\alpha)$, distance au-delà de laquelle $g\ (h)$ est identiquement nul. La manière dont la portée se modifie en fonction de α reflète les anisotropies de la régionalisation.

3 — Le comportement de $g\ (h)$ au voisinage de l'origine, de son côté, donne une image du degré de continuité ou de régularité de la variable régionalisée. En effet, on a toujours :

$$g\ (0) - g\ (h) = \frac{1}{2} \int [f\ (x + h) - f\ (x)]^2\ dx$$

Donc, tout d'abord, si $f\ (x)$ est partout dérivable, $g\ (h)$ va présenter un comportement parabolique au voisinage de l'origine.

Par contre, si $f\ (x)$ est continue « par morceaux » — c'est-à-dire continue et dérivable partout sauf en certains points constituant un ensemble de mesure nulle où elle présente des discontinuités de première espèce, ou sauts — $g\ (h)$ présentera un comportement linéaire au voisinage de l'origine : tel était le cas du covariogramme $K\ (h)$ associé à la variable $k\ (x)$ passant de $+\ 1$ à 0 à la traversée du contour de S.

Plus généralement, on caractérisera un $g\ (h)$ par son développement limité au voisinage de $h = 0$:

$$g\ (h) = g\ (0) + a_2\ h^2 + a_i\ h^i + \ldots + \sum c_\lambda\ h^\lambda$$

Dans ce développement, on distinguera deux parties : une partie *régulière*, comprenant uniquement des termes entiers pairs h^{2n}, et une partie *irrégulière* comprenant des termes en h^λ (λ différent d'un entier pair) ou éventuellement des termes du type $h^{2n}\ log\ h$.

Si la partie régulière existe seule, $g\ (h)$ est indéfiniment dérivable en $h = 0$, et on peut alors montrer qu'il en est de même de $f\ (x)$. C'est donc la partie *irrégulière* qui caractérise le degré de plus ou moins grande continuité de $f\ (x)$, et avant tout son terme de plus bas degré. Si donc λ est l'ordre du terme irrégulier de plus bas degré, λ est un paramètre structural donnant une mesure de la régularité de la régionalisation.

L'estimation

Nous venons de passer rapidement en revue les propriétés qui font du covariogramme transitif un outil pour l'analyse structurale. Ce n'est pas le seul, loin de là, mais c'est celui qui apparaît le premier. Il correspond, en effet, à la plus petite figure structurante que l'on puisse envisager — celle qui est constituée de deux points seulement.

Or — c'est une circonstance remarquable — cet outil structural minimal suffit, à lui tout

seul, pour résoudre le problème de l'estimation des variables régionalisées : problème d'importance pratique considérable.

Soit, en effet, une variable régionalisée $f(x)$, et un réseau de prélèvements implantés selon une maille régulière a (nous utilisons des notations unidimensionnelles, mais la transposition aux espaces à *2* ou *3* dimensions est immédiate). Si y est l'un des points de prélèvements, on connaît les valeurs numériques de $f(y + pa)$ pour tous les entiers p, positifs et négatifs.

On peut estimer l'intégrale :

$$K = \int f(x)\, dx$$

à partir de l'estimateur :

$$(3) \qquad Q^*(y) = |a| \sum_p f(y + pa)$$

Cet estimateur $Q^*(y)$ est une fonction périodique, $Q(y + pa) = Q^*(y)$, puisqu'il est manifestement indifférent de prendre comme origine y du réseau l'un ou l'autre des points de prélèvement.

Si l'on admet que l'origine y du réseau est implantée au hasard à l'intérieur du segment (pour l'espace à une dimension) ou du rectangle (pour l'espace à deux dimensions) de maille a, cet estimateur $Q^*(y)$ devient une variable aléatoire de moyenne Q et de variance:

$$\sigma^2(a) = \frac{1}{a} \int [Q^*(y) - Q]^2\, dy$$

Telle est la *variance d'estimation* associée à la maille a. Pour expliciter son expression, on remplace $Q^*(y)$ par son expression (3), et on obtient facilement :

$$(4) \qquad \sigma^2(a) = |a| \sum_p g(pa) - \int g(h)\, dh$$

On conçoit, intuitivement, qu'une estimation doit être d'autant plus précise que le réseau de prélèvements est plus serré, et que la régionalisation est elle-même plus régulière dans sa variation spatiale. L'influence de ces deux facteurs est visible sur la formule (4). En effet, la variance d'estimation apparaît comme la différence entre la valeur approchée et la valeur exacte d'une même intégrale $\int g(h)\, dh$, c'est-à-dire comme l'erreur commise dans le calcul approché de cette intégrale à partir des valeurs numériques des $g(pa)$. Elle est d'autant plus faible que les valeurs numériques disponibles sont plus nombreuses (que le réseau de prélèvement est plus serré) ; et aussi, comme le montre la théorie du calcul numérique approché des intégrales, d'autant plus faible que la fonction $g(h)$ est plus régulière, donc que la variable régionalisée $f(x)$ est elle-même plus régulière, dans sa variation spatiale. Mais naturellement, la formule (4) donne une expression mathématique précise à cette vue intuitive des choses.

Dans la pratique, cependant, la formule (4) conduirait à des calculs longs, puisqu'il s'agit de mettre en évidence la différence de deux quantités voisines l'une de l'autre, et pénibles à calculer. On est donc conduit à mettre au point des formules d'approximation permettant un calcul plus rapide.

On peut distinguer deux cas extrêmes, selon que la maille a est petite ou grande devant la portée.

1 — Si la maille a est plus grande que la portée, les $g(pa)$ de la formule (4) sont tous nuls, à l'exception du terme $g(0)$ correspondant à $p = 0$; il vient donc :

$$(5) \qquad \sigma^2(a) = |a|\, g(0) - \int g(h)\, dh$$

Cette expression possède une signification aléatoire évidente. La maille étant supérieure à la portée, un prélèvement au plus prend une valeur différente de *0*, et tout se passe comme si ce prélèvement unique était implanté au hasard dans un rectangle égal au rectangle de maille et contenant la minéralisation. La variance de la teneur $f(y)$ de ce prélèvement unique est alors :

$$\frac{1}{|a|} \int [f(x)]^2\, dx - \frac{1}{|a|^2}\, Q^2$$

et celle de l'estimateur $|a|\, f(y)$ est bien donnée par (5).

2 — Si la maille, au contraire, est petite vis-à-vis de la portée, on est conduit à rechercher un développement limité de $\sigma^2(a)$ au voisinage de $a = 0$. Sous réserve d'un terme fluctuant, ou « Zitterbewegung », à peu près imprévisible, mais de valeur moyenne nulle, on peut montrer que ce développement limité corres-

pond, terme pour terme, à la partie irrégulière du covariogramme (la contribution à ce développement de la partie régulière est identiquement nulle).

Ainsi, dans l'espace à *une dimension*, on trouve qu'au terme r^λ ($\lambda \neq 2\ n$) correspond le terme :

$$\sigma^2(a) = A_\lambda\ a^{1+\lambda}$$

A_λ étant une constante (dépendant de λ). Si L désigne la longueur du champ minéralisé total à 1 dimension, et $n = L/a$ le nombre des sondages positifs (ou plutôt la valeur probable de ce nombre), on trouve donc pour ce terme en r^λ :

$$\sigma^2(a) = A_\lambda L^{1+\lambda}\ \frac{1}{n^{1+\lambda}}$$

La variance est en $1/n^{1+\lambda}$ et non pas en $1/n$, comme l'aurait suggéré une application mécanique de la statistique usuelle.

Dans l'espace à *deux dimensions*, si a_1 et a_2 désignent les deux côté du rectangle de maille, avec $a_1 \leq a_2$, et pour un covariogramme isotrope $g(r)$ (ne dépendant que du rayon vecteur r et non de la direction du vecteur h), on obtient une expression de la forme :

$$(6)\quad \sigma^2(a_1, a_2) = B_\lambda\ a_2^{2+\lambda} + C_\lambda\ a_2\ a_1^{1+\lambda}$$

Lorsque a_1 est nul, cette variance se réduit à son premier terme, qui ne dépend que de a_2 et représente l'erreur que l'on commet en étendant au gisement entier la teneur moyenne des lignes de plus grande densité de sondages, lorsque cette teneur moyenne de lignes est parfaitement connue. Le deuxième terme, qui dépend de a_1, représente l'erreur que l'on commet en estimant la teneur moyenne des lignes à partir des prélèvements espacés de a_1 sur ces lignes. La variance totale est la somme de ces deux termes, de sorte que la formule (6) exprime un principe de composition des termes de lignes et de tranche : tout se passe comme si les erreurs que l'on commet, en estimant les lignes de plus grande densité à partir des prélèvements ponctuels, et en estimant le gisement lui-même à partir des lignes supposées parfaitement connues, pouvaient être regardées comme deux variables indépendantes, au sens du calcul des probabilités, et comme si, par suite, les variances correspondantes pouvaient être purement et simplement additionnées.

Si S désigne la surface minéralisée totale, et $n = S/a_1\ a_2$ le nombre des sondages positifs, et si l'on pose $\mu = a_1/a_2 \leq 1$, la formule (6) peut aussi bien s'écrire :

$$\sigma^2(a_1, a_2) = \left(\frac{S}{n}\right)^{1+\frac{\lambda}{2}} \left[B_\lambda \frac{1}{\mu^{1+\frac{\lambda}{2}}} + C_\lambda\ \mu^{\lambda/2} \right]$$

La variance est cette fois en $1/n^{1+\lambda/2}$.

Application à l'estimation d'une surface minéralisée

Nous allons maintenant appliquer la formule (6) dans le cas où la fonction $f(x) = k(x)$ ne peut prendre que les valeurs 0 ou 1, et représente ainsi une surface S : le problème posé est donc celui de l'estimation d'une surface minéralisée S à partir d'un réseau de sondages à maille restangulaire $a_1\ a_2$ ($a_1 \leq a_2$). Le covariogramme transitif $K(h)$ associé à l'aire S est linéaire au voisinage de l'origine :

$$K(h) = K(0) - |\ h\ |\ D_\alpha$$

D_α représentant, comme nous l'avons vu, la demi-projection des éléments d'arc du contour de S sur un axe perpendiculaire à la direction α du vecteur h.

1 — Plaçons-nous d'abord dans le cas isotrope, c'est-à-dire dans le cas où $D_\alpha = D$ est à peu près indépendant de la direction α. La formule (6) est directement applicable, et donne :

$$\frac{\sigma^2 s}{S^2} = \frac{1}{n^{3/2}}\ \frac{D}{\sqrt{S}} \left[\frac{1}{6}\sqrt{\mu} + 0{,}061\ \frac{1}{\mu^{3/2}} \right]$$

$$\left(\mu = \frac{a_1}{a_2} \leq 1 \right)$$

La variance est en $1/n^{3/2}$, n désignant le nombre des sondages positifs. Pour utiliser effectivement cette formule, à partir des données disponibles expérimentalement, on peut estimer S en attribuant à chaque sondage son rectangle d'influence, ce qui donne :

$$S = n a_1\ a_2$$

et D à partir du contour de la réunion de ces zones d'influence (cf. *figure* 2). On comptera

donc les nombre N_1 et N_2 des éléments parallèles à a_1 et a^2 respectivement, qui constituent ce périmètre, et on aura $D = N_1\, a_1 = N_2\, a_2$ (puisqu'il y a isotropie). Par suite, il vient :

$$(7)\quad \frac{\sigma^2_S}{S_2} = \frac{l}{n^2}\left[\frac{l}{6} N_2 + 0{,}061\, \frac{(N_1)^2}{N_2}\right] \quad (N_2 \leqq N_1)$$

2 — En général, cependant, le contour ne sera pas suffisamment isotrope pour que D_α puisse être regardée comme constante. Il présentera, par exemple, une direction principale d'allongement. Si l'un des côtés de la maille est parallèle à cette direction principale, ce qui sera souvent le cas, la formule (7) précédente reste applicable : en effet, prenant cette direction comme axe des x, si nous multiplions les ordonnées par un module convenable, nous obtenons une nouvelle figure, isotrope celle-là, pour laquelle (7) est valable ; mais cette transformation linéaire n'a modifié, ni n, ni N_1, ni N_2, ni la variance relative σ^2_S/S^2, de sorte que (7) est valable aussi dans le cas de la figure anisotrope initiale.

Exemple : Sur la *figure* 2 ci-dessous, l'aire minéralisée S est estimée à *10* fois le rectangle de maille $a_1\, a_2$. Elle comporte un trou (une lacune). Dans le décompte de N_1 et N_2 doivent figurer aussi bien les éléments du contour intérieur que ceux du contour extérieur. On lit donc sur la figure :

$$2\, D_1 = 12\, a_1 \qquad \text{soit } N_1 = 6$$
$$2\, D_2 = 8\, a_2 \qquad \text{soit } N_2 = 4$$

D'où, par conséquent :

$$\frac{\sigma^2_S}{S^2} = \frac{1}{100}\left[\frac{4}{6} + 0{,}061\, \frac{36}{4}\right] \simeq \frac{1{,}21}{100}$$

soit un écart type relatif $\sigma_S/S = 11/100$, et une fourchette d'erreur relative de $\pm$ 22 %.

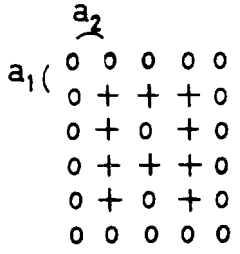

Fig. 2

III. LES SCHEMAS INTRINSEQUES

Dans la première partie de cet exposé, nous avons mentionné l'existence des problèmes méthodologiques graves que pose l'interprétation d'une variable régionalisée comme réalition d'une fonction aléatoire. Ces difficultés sont dues essentiellement à l'unicité des phénomènes naturels, et à l'impossibilité de l'inférence statistique.

Reprenons, en effet, les termes du problème. Une variable aléatoire ordinaire Y est définie par sa loi de probabilité $G(y) = P(Y < y)$. Si l'on effectue un tirage au sort selon la loi $G(y)$, on obtient une valeur numérique particulière, par exemple $y = 98$. Naturellement, à partir de cette valeur numérique unique $y = 98$, il n'est pas possible de reconstituer la loi $G(y)$. Il faut pour cela effectuer un nombre assez grand d'expériences indépendantes donnant chacune des valeurs y différentes.

Une fonction aléatoire F est une variable aléatoire vectorielle à une infinité de composantes : chacune de ces composantes représente la valeur (aléatoire) $F(x)$ prise par F en chacun des points x de l'espace. Si l'on effectue un tirage au sort selon la loi de F, on obtient pour chaque point x une valeur particulière $f(x)$: $f(x)$ constitue ce que l'on appelle une *réalisation* de F.

Ainsi, entre la fonction aléatoire F et sa réalisation f, il y a le même rapport qu'entre la variable aléatoire ordinaire Y et la valeur numérique particulière $y = 98$ obtenue à l'issue d'un tirage au sort effectué selon la loi de Y. Si donc nous nous donnons une variable régionalisée $f(x)$, nous ne pouvons par dire que $f(x)$ est une fonction aléatoire. Une telle expression n'aurait pas plus de sens que si l'on disait : le nombre *98* est une variable aléatoire. En termes corrects, l'hypothèse probabiliste que nous désirons introduire doit se formuler comme suit : « la variable régionalisée $f(x)$ peut être considérée comme une réalisation d'une fonction aléatoire F ».

Or, il y a dans telle hypothèse, et dans les termes mêmes où nous l'avons énoncée, un aspect presque *platonicien :* la seule réalité digne de ce nom devient alors, en effet, la

fonction aléatoire elle-même, définie idéalement par une loi de probabilité donnée sur un espace abstrait. Le phénomène réel, régionalisé, dont nous étions partis, n'est plus considéré que comme une réalisation, c'est-à-dire comme un reflet lointain et déformé, une imitation maladroite de ce modèle idéal.

Malheureusement, nous n'avons pas accès directement au ciel des Idées, et nous devons examiner soigneusement si une telle hypothèse a un sens. Il faut au minimum, pour cela, que l'inférence statistique soit possible. Or, il n'est jamais possible de reconstituer la loi de la variable aléatoire ordinaire Y à partir d'une seule valeur numérique. De la même manière, on peut craindre qu'il ne soit pas possible de reconstruire la loi de F à partir d'une réalisation unique f. Dans les sciences de la Terre, contrairement à ce qui se passe en physique, où l'on peut répéter une expérience aussi souvent que l'on veut, les phénomènes auxquels nous avons affaire sont uniques. Et c'est justement parce que l'inférence statistique n'est pas possible, en général, qu'il a fallu mettre au point les méthodes transitives. Mais, heureusement, il y a toute une catégorie de cas où l'inférence statistique redevient possible, même à partir d'une réalisation unique : c'est le cas des fonctions aléatoires stationnaires.

Le cas stationnaire : Pour définir une fonction alétoire F, il faut connaître sa loi spatiale, c'est-à-dire toutes les lois :

$$G(f_1, f_2 \ldots ; x_1, x_2 \ldots) = \\ = P[F(x_1) < f_1, F(x_2) < f_2, \ldots]$$

pour tous les points d'appui x_1, x_2, ... x_k en nombre k quelconque, mais fini. Ainsi, pour k points d'appui particuliers, on doit en général estimer au minimum k moments d'ordre *1* et $\frac{k(k+1)}{2}$ moments d'ordre *2*, de sorte que le problème reste indéterminé. La fonction aléatoire F est alors dite stationnaire si sa loi spatiale est invariante par translation, c'est-à-dire si l'on a, pour tout vecteur h :

$$G(f_1, f_2 \ldots ; x_1, x_2 \ldots) = \\ = G(f_1 f_2 \ldots ; x_1 + h, x_2 + h, \ldots)$$

Cette notion de stationnarité a le sens d'une sorte d'homogénéité statistique dans l'espace : quel que soit le lieu où l'on se place, le phénomène présente les mêmes caractéristiques moyennes. On peut dire, dans un langage plus imagé, que le phénomène se répète lui-même dans l'espace. Grâce à cette répétition, on conçoit que l'inférence statistique redevienne possible : du fait que la loi spatiale est invariante par translation, le nombre des paramètres dont elle dépend est, en effet, substantiellement réduit.

Nous ferons donc l'hypothèse que notre variable régionalisée peut être considérée comme une réalisation d'une fonction aléatoire stationnaire : cette hypothèse n'est d'ailleurs pas toujours acceptable ; c'est ainsi qu'un phénomène manifestant une décroissance zonée à partir d'un cœur riche ne peut pas être considéré comme stationnaire, et ne peut être décrit que par les méthodes transitives. Il y a cependant de très nombreux cas où cette hypothèse stationnaire est parfaitement admissible.

Variances a priori infinies

Dans la théorie des fonctions aléatoires stationnaires d'ordre deux, l'outil de travail est la fonction

$$K(h) = E[f(x) \cdot f(x + h)]$$

qui représente la covariance des valeurs prises par la fonction aléatoire $f(x)$ en deux points distants de h. Cette covariance n'existe que si la variance a priori existe également, c'est-à-dire si l'on a

$$K(0) < \infty$$

Or, d'assez nombreux phénomènes présentent une capacité de dispersion illimitée et ne peuvent pas être décrits correctement si on leur attribue une variance a priori finie.

Ici d'ailleurs, la nature nous tend une sorte de piège. Lorsque l'on prélève des échantillons v dans un champ géométrique V, on obtient un histogramme à partir duquel on peut toujours calculer numériquement une variance qui prend ainsi une valeur parfaitement définie. Mais cette variance est, en réalité, une fonction $\sigma^2(v \mid V)$ du support v et du champ V. Elle augmente, en particulier, lorsque le champ V augmente. Si les échantillons de taille v pos-

sèdent une variance a priori, celle-ci doit apparaître comme la limite pour V infini de la variance expérimentale $\sigma^2 (v \mid V)$. Il y a une dizaine d'années, à propos du grand gisement d'or du Rand, des auteurs d'Afrique du Sud (D.G. KRIGE, H.S. SICHEL, DE WIJS, etc...), ont calculé, — à partir de centaine de milliers d'échantillons — la variance de ces échantillons dans des panneaux de plus en plus grands, puis dans une concession entière, puis dans le gisement du Rand dans son ensemble. Ils ont ainsi obtenu, expérimentalement, une relation du type :

$$\sigma^2 (v \mid V) = \alpha \log \frac{V}{v}$$

La croissance de la variance observée se poursuit sans défaillance selon cette loi logarithmique jusqu'au dernier point expérimental pour lequel V/v est de l'ordre de plusieurs milliards. On peut conclure en toute certitude qu'il n'existe pas ici de variance a priori finie.

Les schémas intrinsèques

Cependant, même lorsque la variance a priori est infinie, il arrive souvent que les accroissements $f(x + h) - f(x)$ conservent une variance finie. On appellera schéma intrinsèque une fonction aléatoire à accroissements stationnaires d'ordre deux. Etudier une fonction aléatoire en tant que schéma intrinsèque revient à l'étudier par l'intermédiaire de ses accroissements, c'est-à-dire à une constante près (c'est ainsi d'ailleurs que l'on étudie le processus de Wiener-Levy ou les processus poissoniens). L'outil de base qui remplace la covariance $K(h)$ est le *demi-variogramme*, défini comme l'espérance du carré des accroissements :

$$\gamma (h) = \frac{1}{2} E [(f (x + h) - f (x))^2]$$

Si la covariance existe, on a

$$\begin{cases} \gamma (h) = K (0) - K (h) \\ K (h) = \gamma (\infty) - \gamma (h) \end{cases}$$

de sorte que le variogramme et la covariance constituent un seul et même outil. Il n'en est ainsi que si $\gamma (\infty)$ existe. Si au contraire $\gamma (h)$ tend vers l'infini avec $\mid h \mid$, la variance a priori $K (0)$ est infinie et $K (h)$ n'existe pas. Mais le variogramme $\gamma (h)$ reste défini et permet d'étudier la fonction aléatoire (ainsi, les processus classiques de Poisson ou de Wiener-Levy ont un variogramme linéaire $\gamma (h) = a \mid h \mid$ qui tend vers l'infini).

La variance $\sigma^2 (v \mid V)$ d'un échantillon v dans un champ V peut s'exprimer à l'aide du variogramme. On trouve

$$(8) \quad \sigma^2 (v \mid V) = \frac{1}{V^2} \int_V \int_V \gamma (x - x') \, dx dx' - \frac{1}{v^2} \int_v \int_v \gamma (x - x') \, dx dx'$$

ou, si l'on veut :

$$(9) \qquad \sigma^2 (v \mid V) = F (V) - F (v)$$

la fonctionnelle $F (V)$ représentant la valeur moyenne de $\gamma (h)$ lorsque les deux extrémités du vecteur h décrivent le volume V.

On voit que, si $\gamma (h)$ tend vers l'infini avec h, $F (V)$ tendra vers l'infini avec V et par conséquent aussi la variance de v dans V. Avec un variogramme du type

$$\gamma (h) = 3 \alpha \log \mid h \mid$$

(et en supposant v et V géométriquement semblables), on obtient

$$\sigma^2 (v \mid V) = \alpha \log \frac{V}{v}$$

c'est-à-dire justement la formule obtenue expérimentalement par les auteurs d'Afrique du Sud.

De l'expression (9) résulte une *relation d'additivité*

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

dont le sens intuitif est évident : la variance d'un échantillon v dans le gisement V est la somme de la variance de v dans un panneau V' et de la variance du panneau V' dans le gisement V.

On peut définir également la covariance de deux échantillons v et v' (situés à une distance fixe l'un de l'autre). On obtient une formule analogue à (9), $F (v)$ étant remplacée par la valeur moyenne de $\gamma (h)$ lorsque les deux extrémités de h décrivent respectivement les volumes v et v'. Si $\gamma (h)$ tend vers l'infini avec

h, variances et covariances sont des infiniment grands équivalents à $F(V)$, de sorte que le *coefficient de corrélation* expérimental tend toujours vers l'unité : ce coefficient n'est plus un instrument de travail adéquat, et c'est en définitive le variogramme lui-même qui exprime au mieux le réseau des corrélations et la structure spatiale du phénomène :

1) le variogramme donne un contenu précis à la notion traditionnelle de *zone d'influence* d'un échantillon. Sa croissance plus ou moins rapide reflète la manière plus ou moins rapide dont se détériore l'influence d'un échantillon sur des zones de plus en plus lointaines du gisement.

2) *les anisotropies* se manifestent par le comportement différentiel du variogramme dans les différentes directions de l'espace.

3) les phénomènes *de transition* se traduisent sur le variogramme par des paliers et des seuils dont l'analyse permet souvent de mettre en évidence la superposition de plusieurs structures d'échelles différentes.

4) *la continuité* et la régularité d'une régionalisation, enfin, sont très bien exprimées par le comportement du variogramme au voisinage de l'origine. Par ordre de régularité décroissante, on peut distinguer quatre types :

— *Comportement parabolique au voisinage de l'origine* (*figure* 3)

$\gamma(h)$ est deux fois dérivable en $h = 0$.

Ce type correspond à une variable régionalisée dérivable en moyenne quadratique, donc à haut degré de régularité.

— *Tangente oblique à l'origine* (*figure* 4)

$\gamma(h)$ est continu à l'origine, mais n'est pas deux fois dérivable.

Ce type correspond à une variable régionalisée continue mais non dérivable en moyenne quadratique, donc déjà moins régulière.

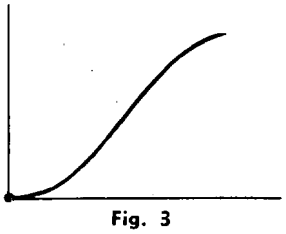

Fig. 3

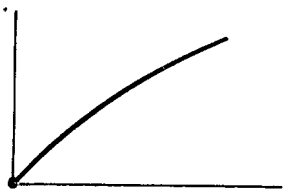

Fig. 4

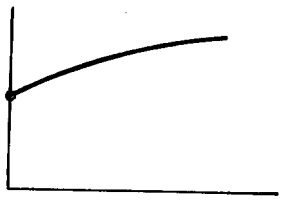

Fig. 5

— *Effet de pépite* (*figure* 5)

Il se caractérise par une discontinuité à l'origine, $\gamma(h)$ ne tendant pas vers 0 avec h. Il correspond à une variable régionalisée qui n'est pas continue en moyenne quadratique, donc extrêmement irrégulière.

— *Variogramme plat, ou effet de pépite à l'état pur* (*figure* 6)

Ce cas limite correspond à une variable régionalisée purement aléatoire (mesure aléatoire à accroissements indépendants).

Partie irrégulière

Lorsque $\gamma(h)$ est de la forme $\gamma(r)$, c'est-à-dire ne dépend que du rayon vecteur $r = |h|$, on peut caractériser son comportement au voisinage de l'origine par un développement limité de la forme :

$$\gamma(r) = \sum a_{2n} r^{2n} + \sum C_\lambda r^\lambda$$

Sur un tel développement, on distingue une partie régulière et une partie irrégulière. La

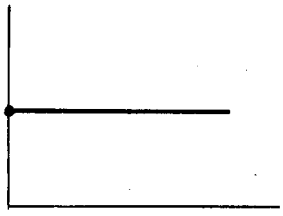

Fig. 6

première est constituée de termes de degrés entiers pairs. Si elle existait seule, $\gamma(r)$ serait indéfiniment dérivable, donc aussi la variable régionalisée elle-même qui présenterait ainsi le plus haut degré de régularité. C'est donc la partie irrégulière, constituée de termes du type r^λ, où λ est différent d'un entier pair (éventuellement aussi des termes logarithmiques en $r^{2n}\ log\ r$) qui définit le type d'irrégularité de la variable régionalisée et, dans cette partie irrégulière, c'est le terme du plus bas degré qui joue le rôle principal.

On est ainsi amené à définir le degré de régularité d'une variable régionalisée comme l'ordre λ du terme irrégulier r^λ du plus bas degré de son variogramme.

Régularisation et montée

Lorsque l'on remplace une variable régionalisée $f(x)$ par la valeur moyenne de $f(x)$ dans un échantillon v prélevé au point x (ou par une moyenne pondérée plus générale) on obtient une nouvelle variable régionalisée dont le variogramme se déduit du $\gamma(h)$ initial par des opérations de convolution, sur lesquelles nous n'insisterons pas ici, mais dont l'effet est toujours régularisant. Un exemple particulièrement instructif, et important dans les applications, de cette régularisation est la *montée* : la montée, opération permettant de passer d'une variable régionalisée définie dans l'espace à n dimensions à des variables régionalisées définies dans les espaces à $n-1$, $n-2$ ... dimensions, est la transposition de la technique minière qui consiste à forer des sondages verticaux, tracer des niveaux horizontaux, etc. ... Par exemple, si $f(x, y, z)$ est la teneur au point de coordonnées $(x\ y\ z)$ à l'intérieur d'une formation stratiforme horizontale, un sondage implanté au point de coordonnées x, y de la surface topographique contient la teneur (exprimée en quantité de métal au mètre carré)

$$\int f(x\ y\ z)\ dz$$

Cette intégrale représente une nouvelle variable régionalisée, définie dans l'espace à 2 dimensions (la surface topographique), déduite de l'ancienne par une montée d'ordre 1.

La régularisation d'une variable régionalisée à la montée se manifeste par l'effet suivant : si le terme irrégulier de plus bas degré du variogramme de la variable initiale est en r^λ, celui de la variable qui s'en déduit par montée d'ordre 1 est en $r^{\lambda+1}$. Le degré de régularité augmente d'une unité (lorsque λ est un entier impair, on a la séquence logarithmique :

$$log\ r \rightarrow r \rightarrow r^2\ log\ r \rightarrow \ldots$$

où alternent les termes impairs r^{2k-1} et les termes logarithmiques $r^{2k}\ log\ r$).

Variation d'estimation

Montrons maintenant comment le variogramme permet de résoudre un problème essentiellement pratique, comme celui de l'estimation d'une variable régionalisée à partir d'un échantillonnage fragmentaire. Soit par exemple à estimer la teneur moyenne Z d'un domaine V :

$$Z = \frac{1}{V}\int_V f(x)\ dx$$

à partir de prélèvements (ici supposés ponctuels) effectués en n points $x_1 \ldots x_n$. On forme l'estimateur

$$Y = \frac{1}{n}\sum f(x_i)$$

et, pour apprécier l'ordre de grandeur de l'erreur possible, on introduit la variance $D^2(Y-Z)$ de la différence $Y-Z$. Si les variances a priori existent, on doit avoir :

$$(10) \qquad D^2(Y-Z) = \sigma^2_Z + \sigma^2_Y - 2\,\sigma_{ZY}$$

Par contre, si $\gamma(h)$ devient infini avec h, ces variances et cette covariance a priori n'existent plus. Mais la variance d'estimation $D^2(Y-Z)$ reste définie, car la différence $Y-Z$ est une combinaison linéaire d'accroissement de $f(x)$. On montre que cette variance est :

$$(11)\ D^2(Y-Z) = \frac{2}{nV}\sum_i \int_V \gamma(x-x_i)\ dx - \frac{1}{n^2}\sum_{i,j}\gamma(x_i-x_j) - \frac{1}{V^2}\int_V\int_V \gamma(x-x')\ dx dx'$$

Cette formule a la même signification que (10). A la covariance σ_{YZ} correspond le terme mixte, où figurent une intégrale et une sommation discrète, tandis que σ^2_Y et σ^2_Z ont pour homologues la somme double et l'intégrale dou-

ble. Mais (11) est plus générale que (10), et subsiste même si les variances a priori n'existent pas.

On notera la structure remarquable de la formule (11), où alternent des expressions exactes et approchées des mêmes intégrales. Le deuxième terme apparaît comme une valeur approchée du premier, lui-même approximation du dernier. La théorie du calcul numérique approché des intégrales nous fait donc pressentir que la variance d'estimation sera d'autant plus faible :

— que le réseau de prélèvement sera plus serré et plus représentatif de la géométrie du domaine V à estimer,

— que le variogramme $\gamma(h)$ sera lui-même plus régulier analytiquement, et donc que la variable régionalisée sera elle-même plus régulière et continue dans sa variation spatiale.

Ces conclusions sont bien conformes à ce que suggère l'intuition. Toutefois, la formule (11) conduirait à des calculs assez longs, lorsque les prélèvements sont un peu nombreux, et on est amené à utiliser des principes d'approximation :

1) Dans l'espace *à une seule dimension,* et pour des prélèvements à maille régulière a, l'estimation du champ de longueur $L = na$ constitué de la juxtaposition des zones d'influence de longueur a au centre desquelles sont implantés chacun des n échantillons, admet la variance

$$D^2(Y - Z) = \frac{1}{n}\,\sigma_E^2$$

σ_E^2, *variance d'extension* élémentaire, est fonction de a et du variogramme $\gamma(h)$. Si r^λ est le terme irrégulier de plus bas degré, la partie principale de la variance d'extension est proportionnelle à a^λ :

$$\sigma_E^2 = A(\lambda)\, a^\lambda$$

Comme $a = \frac{L}{n}$, on voit que la variance d'estimation est de la forme :

$$D^2(Y - Z) = \frac{C}{n^{1+\lambda}}$$

Elle est en $\frac{1}{n^{1+\lambda}}$ et non plus en $\frac{1}{n}$ comme dans la statistique ordinaire. Le cas $\lambda = 0$, qui correspond à un effet de pépite, redonne bien la formule usuelle en $\frac{1}{n}$; cela est naturel, puisque l'effet de pépite à l'état pur correspond à des échantillons indépendants. Mais pour une valeur usuelle comme $\lambda = 1$, la variance est en $\frac{1}{n^2}$ et décroît beaucoup plus vite : dans les applications, ce phénomène permet des économies substantielles.

2) Dans l'espace *à plusieurs dimensions,* on applique un principe d'approximation connu sous le nom de composition des termes de tranches et des termes de lignes. Exposons-le sur un exemple. Soit un gisement filonien reconnu par traçages (galeries horizontales tracées dans le plan du filon) à des niveaux équidistants. Chaque traçage est lui-même échantillonné par des prélèvements équidistants (on suppose que la maille a de ces prélèvements est inférieure à la relevée h entre traçages). L'erreur d'estimation apparaît comme la somme de deux erreurs :

— celle que l'on commet en étendant aux traçages la teneur moyenne des prélèvements,

— celle que l'on commet en étendant au gisement lui-même la teneur des traçages supposée parfaitement connue.

La première erreur admet une variance

$$\frac{1}{n}\,\sigma_1^2(a)$$

où n est le nombre total de prélèvements et $\sigma_1^2(a)$ la variance d'extension d'un échantillon dans sa zone d'influence, à une dimension, de longueur égale à la maille a. On la calcule comme ci-dessus.

La deuxième erreur admet la variance

$$\frac{1}{N}\,\sigma_2^2(h),$$

N étant le nombre des niveaux tracés, et $\sigma_2^2(h)$ la variance d'extension d'un traçage dans sa tranche d'influence ($\sigma_2^2(h)$ se calcule à partir du variogramme déduit de $\gamma(h)$ par montée d'ordre 1).

Si a est inférieur à h, on peut montrer que ces deux erreurs sont à peu près indépendantes, de sorte que la variance d'estimation est :

$$D^2 (Y - Z) = \frac{1}{n} \sigma_1^2 (a) + \frac{1}{N} \sigma_2^2 (h)$$

En général, les prélèvements reviennent moins chers que le mètre de galerie. Néanmoins, il est inutile de les multiplier indéfiniment : la variance d'estimation ne descendra pas en-dessous de $\frac{1}{N} \sigma_2^2 (h)$, limite assez vite atteinte dès que n est grand. Il arrive un moment où on ne peut plus améliorer l'estimation sans travaux miniers supplémentaires.

Connaissant le prix de revient du prélèvement et du mètre de traçage, et l'expression de $\sigma_1^2 (a)$ et $\sigma_2^2 (h)$ en fonction de a et de h, un calcul d'optimisation élémentaire permet de déterminer le nombre d'échantillons que l'on doit prélever au mètre de traçage pour minimiser la variance d'estimation à dépense donnée.

D'une manière générale, la théorie des variables régionalisées, en donnant, sous la forme d'une variance d'estimation, une mesure précise de l'information disponible, fournit une base solide pour la recherche d'optima économiques en matière de reconnaissance minière. Je ne peux malheureusement pas développer ce point ici, non plus que bien d'autres applications intéressantes, telles que : le krigeage (ou recherche de l'estimateur optimal), la simulation numérique de gisements miniers, la théorie des milieux poreux, etc... pour lesquelles on me pardonnera de renvoyer à la bibliographie sommaire suivante :

LISTE BIBLIOGRAPHIQUE

— Traité de géostatique appliquée. Tome I (1962), Tome II (1963). Ed. Technip. Paris.

— Les variables régionalisées et leur estimation (thèse). Masson, Paris, 1965.

— Recherche d'optimum économique dans la reconnaissance et la mise en exploitation des gisements miniers. En collaboration avec Ph. FORMERY. *Annales des Mines*, Mai et Juin 1963.

— Principles of Geostatistics. *Economic Geology*. Décembre 1963.

— Structure et composition des perméabilités. *Revue de l'I. F. P.* Avril 1966.

— Comparaison entre les échantillonnages à poids constant et à effectif constant. *Revue de l'Industrie Minérale*. Août 1966.

— Eléments pour une théorie des milieux poreux. Paris. Masson, 1967.