

## LES PRINCIPES DE LA GEOSTATISTIQUE

---

### INTRODUCTION ET BREF HISTORIQUE

Au sens le plus large, la géostatistique est l'étude de la manière dont se répartissent dans l'espace les grandeurs utiles au mineur ou au géologue, telles qu'une teneur, une puissance, une accumulation, etc ..., avec une application pratique d'importance majeure aux problèmes posés par l'estimation des gisements.

Historiquement, on peut dire que la géostatistique est aussi vieille que l'art des mines lui-même. Dès que les mineurs se sont préoccupés de prévoir à l'avance les résultats de leurs travaux à venir, et, plus particulièrement, dès qu'ils ont commencé à prélever et à analyser des échantillons, et à calculer des teneurs moyennes pondérées par des puissances et des zones d'influence, on peut dire que la géostatistique était née. En tant qu'elles tiennent le plus grand compte des caractéristiques spatiales de la minéralisation, ces méthodes traditionnelles conservent toute leur valeur. Loin de les réfuter, les développements modernes de la théorie les ont prises comme point de départ, et les ont portées à un niveau d'expression scientifique plus élevé.

Cependant, si elles permettaient une estimation correcte des valeurs moyennes, ces méthodes traditionnelles n'exprimaient en aucune façon une importante caractéristique des minéralisations, qui est leur variabilité ou leur dispersion. C'est pour tenir compte de cette caractéristique que l'on a commencé, il y a quelques dizaines d'années, à utiliser les techniques du calcul des probabilités classique. Si une application maladroite de ces techniques a parfois conduit à des absurdités, il est indéniable que, dans l'ensemble, les résultats obtenus ont été fructueux. C'est d'ailleurs un peu un paradoxe, car les méthodes statistiques classiques, en tant qu'elles font abstraction de l'aspect spatial des répartitions étudiées, ne sont pas réellement applicables. En fait, l'Ecole d'Afrique du Sud, qui a obtenu les résultats les plus notables avec KRIGE, SICHEL, le hollandais de WIJS, etc ... disait et croyait qu'elle appliquait la statistique classique. Mais les méthodes qu'elle mettait au point en différaient de plus en plus, et s'adaptèrent spontanément à leur objet.

Le deuxième changement décisif s'est manifesté quand on a clairement réalisé l'insuffisance du calcul des probabilités classiques, et la nécessité de réintroduire les caractères spatiaux des répartitions. Il s'agissait de faire, à un niveau supérieur, la synthèse entre les méthodes traditionnelles et statistiques. Dès lors, la géostatistique a commencé à élaborer ses propres méthodes, et son propre formalisme mathématique, qui n'est pas autre chose qu'une mise en forme abstraite et une systématisation de l'expérience minière séculaire. Ce formalisme a hérité, de ses origines statistiques, un langage où l'on continue, par exemple, à parler de variances ou de covariances, tout en donnant à ces notions un contenu nouveau. Cette similitude de vocabulaire ne doit pas faire illusion. Au terme d'une longue évolution, la géostatistique a reconnu qu'elle avait, en face d'elle, non pas des événements aléatoires, mais des phénomènes naturels répartis dans l'espace. Et, en conséquence, ses méthodes se sont rapprochées de celles de la physique mathématique, et, en particulier, de celles de l'analyse harmonique.

## 2/- Insuffisance des concepts statistiques classiques.

Dans ce qui suit, nous parlerons, pour abrégé, de la répartition des teneurs à l'intérieur d'un gisement. Les résultats auxquels nous parviendront auront, cependant, une portée générale, et pourront s'appliquer à n'importe quelle grandeur à répartition spatiale. Dans un traitement statistique ordinaire, on prend les teneurs des échantillons prélevés dans le gisement, et on les classe sous forme d'histogramme. Par là même, on fait abstraction de l'endroit où ils ont été prélevés. Or, il ne suffit pas de savoir avec quelle fréquence une teneur se répète dans un gisement. Il faut aussi savoir de quelle manière les teneurs se succèdent sur le terrain, et, en particulier, quelle est la taille et la position des zones payantes. Au point de départ de la théorie, on trouve ainsi la reconnaissance d'un fait : l'impuissance de la statistique ordinaire à rendre compte de l'aspect spatial du phénomène, qui est, justement, son aspect le plus important.

D'une manière plus précise, le calcul des probabilités classiques a pour objet l'étude des variables aléatoires. L'exemple simple du jeu de pile ou face montre très bien de quoi il s'agit. Convenons de marquer + 1 chaque fois que la

pièce tombe sur pile, et  $-1$  dans le cas contraire. Avant de lancer la pièce, je n'ai aucun moyen de savoir si je marquerai  $+1$  ou  $-1$  : je sais seulement qu'il y a une chance sur deux pour que l'un ou l'autre se produise. Une variable aléatoire présente, classiquement, deux caractères essentiels :

- la possibilité, au moins théorique, de répéter indéfiniment l'épreuve qui attribue à la variable une valeur numérique : je peux, par exemple, lancer la pièce de monnaie aussi souvent que je le veux.

- l'indépendance de chacune des épreuves vis-à-vis des précédentes et des suivantes : si les 100 premiers coups ont tous donné pile, il n'en reste pas moins une chance sur deux pour que le 101ème donne face.

Il est clair qu'une teneur, dans un gisement, ne peut pas posséder ces deux caractères. La teneur d'un panneau, tout d'abord, est unique. On n'exploite qu'une fois ce panneau là, et la possibilité de répéter indéfiniment l'épreuve n'existe pas. S'il s'agit de la teneur d'un échantillon, qui peut être par exemple un rainurage de dimensions données, il en est exactement de même, puisque la teneur du rainurage implanté au point de coordonnées  $(x,y)$  est unique, et bien déterminée. Cependant, on peut prélever un deuxième échantillon à côté du premier, puis un troisième, etc ..., ce qui donne la possibilité apparente de répéter l'épreuve. En réalité, il ne s'agit pas exactement de la même épreuve, mais d'une épreuve légèrement différente. Mais, même en admettant cette possibilité de répétition, le deuxième caractère des variables aléatoires ne sera sûrement pas respecté. Deux échantillons voisins ne sont sûrement pas indépendants. Ils ont tendance - en moyenne - à être riches tous les deux s'ils proviennent d'un panneau riche, et réciproquement. Cette tendance, plus ou moins marquée, exprime le degré de plus ou moins grande continuité de la variation des teneurs dans l'espace minéralisé.

La méconnaissance de ce point et la transposition brutale de la statistique classique a parfois conduit à de singuliers errements. Vers les années 50, on recommandait, pour la reconnaissance d'un gisement, de tirer au sort l'implantation de chacun des sondages (c'est-à-dire, exactement, de les planter n'importe où). Les mineurs ont, naturellement, continué à utiliser les mailles régulières traditionnelles, et la géostatistique a pu, par la suite, démontrer qu'ils avaient raison. Ou bien encore, on soutenait que la précision, sur l'évaluation d'un

... / ...

gisement, ne dépendait que du nombre d'échantillons prélevés (et non de leur implantation), et variait comme la racine carrée de ce nombre. Cette transposition maladroitement de la théorie des erreurs menait à des absurdités. Par exemple, si un gisement est reconnu par sondages, il suffirait de découper les carottes en tronçons de 5 millimètres au lieu de 50 centimètres pour avoir 100 fois plus d'échantillons, donc une précision 10 fois meilleure. Naturellement, cela est faux. La multiplicité des échantillons ainsi obtenus est illusoire, et ne fait que répéter indéfiniment la même information, sans rien apporter de plus. La géostatistique montre effectivement, que la précision est la même pour les tronçons de 5mm et 50 cm, comme n'importe quel mineur le comprend d'instinct.

### 3/- La Notion de Variable Régionalisée

Ainsi, une teneur ne peut en aucune façon être assimilée à une variable aléatoire. Nous parlerons de variables régionalisées, précisément pour mettre l'accent sur l'aspect spatial du phénomène. Une variable régionalisée est, au sens strict, une véritable fonction, prenant en chaque point de l'espace une valeur définie. En général, une telle fonction a des propriétés trop complexes pour que l'on puisse l'étudier commodément par les méthodes ordinaires de l'analyse mathématique. Du point de vue de la physique, ou de la géologie, un certain nombre de caractères qualitatifs sont attachés à la notion de variable régionalisée.

a/- En premier lieu, une variable régionalisée est localisée. Ses variations ont pour théâtre l'espace minéralisé (le volume du gisement, ou de la couche) que l'on appelle champ géométrique de la régionalisation. De plus, une telle variable est, en général, définie sur un support géométrique. Dans le cas d'une teneur, ce support n'est autre que le volume de l'échantillon, avec sa forme géométrique, ses dimensions et son orientation. Si, dans un même gisement, on change le support géométrique on obtient une nouvelle variation régionalisée, qui présente des analogies avec la première mais ne coïncide pas avec elle. Par exemple, des échantillons de 10 kgs, correspondant à des carottes de sondage, ne se distribuent pas de la même manière que des échantillons de 10 tonnes, correspondant à des volées. On considérera souvent le cas d'un support ponctuel. Une teneur ponctuelle, par exemple, prendra la valeur 0 ou la valeur + 1 selon que son support ponctuel tombera dans un grain stérile ou minéralisé.

b/- En deuxième lieu, la variable peut présenter une plus ou moins grande continuité dans sa variation spatiale, qui se traduira, par exemple, par un écart

plus ou moins grand entre les teneurs des deux échantillons voisins envisagés ci-dessus. Certaines variables à caractère géométrique (la puissance, ou le pendage d'une formation) posséderont la continuité stricte des mathématiciens. Le plus souvent (pour une teneur, ou une accumulation) il ne s'agira que d'une continuité plus lâche, ou continuité "en moyenne". Dans certains cas, même cette continuité en moyenne ne sera pas vérifiée, et nous dirons alors qu'il y a effet de pépite.

c/- Enfin, la variable peut présenter diverses variétés d'anisotropies. Il peut exister une direction préférentielle, le long de laquelle les teneurs varient peu, tandis qu'elles se modifient rapidement le long d'une direction transversale. Ces phénomènes sont bien connus sous le nom de runs, de zonalités, etc ...

A ces caractères généraux, communs à toutes les variables régionalisées, peuvent se superposer des traits spécifiques. Par exemple, dans le cas d'un gisement sédimentaire, on observera un effet de stratification. La stratification en grand donnera des couches, individualisables et exploitables séparément. A l'intérieur de chaque couche, elle pourra se traduire par l'existence de bancs, qui se succèdent verticalement, séparés par des surfaces de discontinuités. La teneur, quasi-constante ou peu variable, à l'intérieur d'un banc, variera brusquement au passage d'un banc à l'autre : si banal et familier que paraisse ce phénomène, il n'en est pas moins fondamental, et une formulation théorique du problème qui n'en tiendrait pas compte passerait à côté du réel. Il arrivera aussi qu'à ces discontinuités verticales, matérialisées par les joints stratigraphiques, s'ajoutent des discontinuités latérales, dues à la terminaison lenticulaire des bancs. Ce phénomène de relai des bancs, quand il existe, met en évidence, à chaque niveau stratigraphique, un compartimentage de l'aire de sédimentation en micro-bassins évoluant de façon quasi-autonomes, et peut se manifester, à l'exploitation, par des effets de seuil de teneur.

De même, dans un gisement du type amas-stockwerk, on observera des filonnets et des amandes riches s'individualisant au sein d'une masse plus faiblement imprégnée. Cet effet de stockwerk, tout comme l'effet de stratification et de relai de bancs, traduit l'apparition d'un réseau de discontinuités au sein d'un champ géométrique homogène. A une échelle bien différente, qui est celle de la granulométrie, l'effet de pépite apparaît comme un phénomène de même nature, le réseau de discontinuités étant ici celui qui sépare les grains de stérile et de minéral.

Ces différents aspects spécifiques de la distribution spatiale des variables régionalisées, profondément étrangers au calcul des probabilités classiques - doivent être obligatoirement pris en compte par la Géostatistique. Cela est rendu possible par un outil mathématique simple, le variogramme.

4/- Le Variogramme.

Le variogramme est une courbe qui représente le degré de continuité de la minéralisation. Expérimentalement, on porte en abscisse une distance  $d$  et, en ordonnée, la valeur moyenne du carré de la différence des teneurs des échantillons prélevés à la distance  $d$  l'un de l'autre. Théoriquement, on désigne par  $f(M)$  la valeur prise en un point  $M$  du champ géométrique  $V$  par une variable régionalisée définie sur un support donné  $v$  (en général, le support  $v$  sera petit et, à la limite, pourra être considéré comme ponctuel). Le demi-variogramme  $\gamma(h)$ , ou loi de dispersion, est défini pour un argument vectoriel  $h$  par l'expression :

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

Le variogramme est, en général, une fonction croissante de la distance  $h$ , puisque deux échantillons ont, en moyenne, des teneurs d'autant plus différentes qu'ils sont plus éloignés. Il donne un contenu précis à la notion traditionnelle de zone d'influence d'un échantillon. La croissance plus ou moins rapide du variogramme représente, en effet, la manière plus ou moins rapide dont se détériore l'influence d'un échantillon donné sur des zones de plus en plus lointaines du gisement. Les caractéristiques qualitatives de la régionalisation s'expriment très bien à l'aide du variogramme :

a/- La plus ou moins grande régularité de la minéralisation est représentée par le comportement plus ou moins régulier du  $\gamma(h)$  au voisinage de l'origine. On peut, en gros, distinguer quatre types (Figure 1). Dans le premier type, le

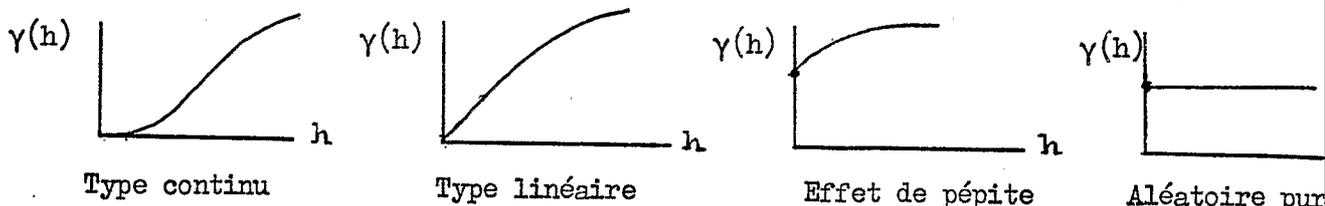
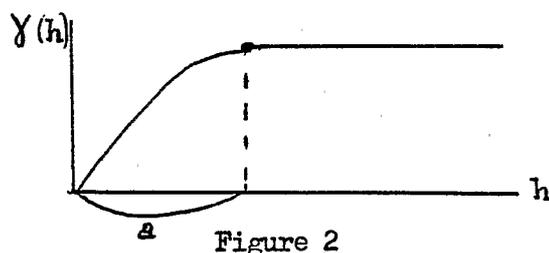


Figure 1

variogramme a un comportement parabolique à l'origine, et représente une variable régionalisée à haute continuité, comme une puissance. Le deuxième type, ou type linéaire, est caractérisé par une tangente oblique à l'origine, et représente une variable qui ne possède que la continuité " en moyenne ". C'est le type le plus courant pour les teneurs dans les gisements métalliques. Le troisième type manifeste une discontinuité à l'origine, et correspond à une variable ne présentant même pas la continuité en moyenne, c'est-à-dire à un effet de pépite. Le quatrième type est un cas limite, qui correspond à la notion classique de variable aléatoire. Entre le type 1 (fonctionnel continu), et le type 4 (aléatoire pur), on voit apparaître une gamme d'états intermédiaires, dont l'étude est l'objet propre de la géostatistique.

b/- Le variogramme n'est pas le même dans les différentes directions de l'espace. La fonction  $\gamma(h)$  définie en (1) ne dépend pas seulement de la longueur, mais aussi de la direction du vecteur  $h$ . En étudiant comment se déforme le variogramme lorsque l'on change cette direction, on met en évidence des directions préférentielles, des runs, des zonalités, etc ... L'interprétation géologique de telles anisotropies est souvent instructive.

c/- Les caractères structuraux se reflètent aussi sur le variogramme. Par exemple, un phénomène de relais des bancs se manifeste, sur la courbe expérimentale,



par un palier du variogramme, au delà d'une distance  $a$ , ou portée (Fig.2) égale au diamètre moyen des micro-bassins de sédimentation autonomes. Et le fait que ces portées ne soient pas les mêmes dans les différentes directions

permet de déterminer les directions d'allongement et la forme moyenne des micro-bassins.

Cet outil, le variogramme, ne représente pas la totalité, ni les détails locaux du phénomène minéralisateur, mais il en exprime, sous forme synthétique, les caractères essentiels. Dans l'analyse harmonique d'un phénomène vibratoire, on met en évidence, pour chaque harmonique, une phase et une amplitude. La configuration locale du phénomène dépend au plus haut point des phases, mais l'énergie ne dépend que du carré des amplitudes. La courbe spectrale, qui donne le carré des amplitudes

ne décrit pas tout le phénomène, mais rend compte de l'essentiel, c'est-à-dire des caractéristiques énergétiques. Le variogramme (ou, pour être précis, sa transformée de Fourier) joue exactement le rôle d'une telle courbe spectrale.

Dans les paragraphes qui suivent, nous allons passer en revue quelques unes des applications possibles de la notion de variogramme. Il n'est naturellement pas question d'en donner ici un exposé systématique. Nous nous contenterons de citer quelques exemples, et quelques formules caractéristiques. Pour plus de détail, nous demandons au lecteur - et nous nous en excusons - de se reporter à notre *Traité de Géostatistique Appliquée* <sup>(1)\*</sup>.

5/- Loi de dispersion absolue, ou intrinsèque.

Le demi-variogramme  $\gamma(h)$ , défini en (1), est lié au champ géométrique  $V$  de la variable régionalisée. Si, au lieu du champ total  $V$ , on n'avait considéré qu'une portion  $V'$  de celui-ci, on aurait obtenu une fonction  $\gamma'(h)$  peut être différente de  $\gamma(h)$ . Cependant, on a la notion intuitive que, dans un champ géométrique géologiquement homogène, il doit y avoir quelque chose d'intrinsèque, ou d'indépendant du lieu, dans les caractéristiques représentant la variabilité d'une variable régionalisée. Formulée d'une manière précise, cette intuition conduit à l'hypothèse d'une loi de dispersion absolue, ou intrinsèque, qui s'exprime par l'égalité :

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

et signifie que le variogramme est indépendant de la portion  $V'$  du gisement  $V$  que l'on a retenue pour son calcul. Disons tout de suite que cette hypothèse n'est pas réellement indispensable au développement de la théorie, et que l'on peut s'en affranchir, moyennant certaines complications mathématiques. <sup>(2)\*</sup> Néanmoins, elle facilite beaucoup l'exposé de la théorie, et, pour cette raison, nous l'adopterons ici. Expérimentalement, on constate, en général, une lente dérive du variogramme dans l'espace, et, si cette dérive ne prend pas trop d'ampleur, les résultats déduits de l'hypothèse d'une loi de dispersion absolue constituent une excellente approximation du réel (sous la condition que le  $\gamma(h)$  effectivement utilisé soit calculé sur les por-

---

(1)\* "TRAITE DE GEOSTATISTIQUE APPLIQUEE" - Editions Technip, Paris. TOME I (1962).  
Tome II (Le Krigeage), sous presse. - Tome III (L'effet de pépite, et les phénomènes transitifs), en préparation.

(2)\* "TRAITE DE GEOSTATISTIQUE APPLIQUEE", TOME III.

tions mêmes du gisement auxquelles on s'intéresse).

Lorsque cette hypothèse est vérifiée, le demi variogramme  $\gamma(h)$  prend lui-même une signification intrinsèque. On le désigne souvent sous le nom de loi de dispersion intrinsèque, ou absolue, ou, plus brièvement, de fonction intrinsèque de la variable régionalisée.

### 6/- Variances et Covariances

Considérons, en premier lieu, une variable régionalisée (que nous appellerons teneur, pour abrégé le langage), définie dans un champ  $V$  et sur un support ponctuel, et obéissant à une loi de dispersion intrinsèque  $\gamma(h)$ . Soit  $f(M)$  la valeur prise par la teneur au point  $M$  du champ  $V$ . Au lieu de la teneur ponctuelle  $f(M)$ , on a le plus souvent affaire à la teneur  $y(M)$  d'un échantillon de taille, de forme et d'orientation données, prélevé au point  $M^{(1)*}$ . Cette nouvelle variable se déduit de l'ancienne par une intégration effectuée dans le volume  $v$  centré en  $M$ .

$$(2) \quad y(M) = \frac{1}{v} \int_v f(M+h) \, dh$$

A cette variable, nous attacherons un paramètre mesurant sa dispersion dans  $V$ , et nous l'appellerons variance, comme en calcul des probabilités classiques. Désignant par  $m$  la valeur moyenne de la variable ponctuelle dans  $V$

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

nous définissons la variance de  $y(M)$  dans  $V$  comme la valeur moyenne dans  $V$  du carré de l'expression  $[y(M) - m]$ , soit

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

On notera que cette notion possède, au départ, une signification géométrique et non pas probabiliste. Cela n'empêchera pas, dans les applications, de calculer

---

(1)\* Nous entendons par là que le centre de gravité de  $v$  est implanté au point  $M$ .

cette variance, à partir des données expérimentales, par les méthodes statistiques ordinaires. Qu'elles soient prises dans leur ordre spatial, comme dans l'intégrale (3), ou regroupées au préalable sous forme d'histogramme, ce sont les mêmes expressions  $(y - m)$  qui apparaissent, avec les mêmes poids, dans les deux procédés de calcul. Mais, sur le plan conceptuel, la définition (3) possède un contenu physique que n'a pas la notion statistique. De l'expression (1) du variogramme de (2) et de la définition (3) de la variance, on déduit, en intervertissant l'ordre des intégrations :

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

Chacune de ces intégrales sextuples a un sens très clair : Elle représente la valeur moyenne dans  $V$  (ou  $v$ ) du  $\gamma(h)$ , lorsque les deux extrémités du vecteur  $h$  balayent, chacune pour son compte, le volume  $V$  (ou  $v$ )

Si l'on pose :

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

= Valeur moyenne dans  $V$  de  $\gamma(h)$

Il vient simplement

$$(5) \quad \sigma^2 = F(V) - F(v)$$

Ainsi la connaissance du variogramme des teneurs ponctuelles permet le calcul a priori de la variance d'un échantillon quelconque  $v$  dans une portion quelconque  $V$  d'un gisement. On notera bien que cette variance ne dépend pas seulement de la mesure des volumes  $v$  et  $V$ , mais aussi de leurs formes, et de leur orientation.

La signification physique de la relation (4) est très instructive. La variance d'un échantillon macroscopique  $v$ , considéré comme la juxtaposition d'un grand nombre de micro-échantillons  $dv$ , ne dépend nullement du nombre de ces micro-échantillons et de leurs variances, mais uniquement de la valeur moyenne de la fonction intrinsèque  $\gamma(h)$  dans le volume géométrique  $v$ . La statistique classique, considérant ces micro-échantillons comme indépendants, conduirait à une variance

en  $\frac{1}{v}$ . En fait, il n'existe pas de gisement dans lequel des volées de 10 tonnes aient une variance mille fois plus faible que des carottes de 10 kg. La formule (4) montre pourquoi. Les teneurs des micro-échantillons ne sont nullement indépendantes. Elles s'insèrent dans un réseau de corrélations spatiales, dont la nature est liée à la plus ou moins grande continuité de la minéralisation, et qui s'exprime justement dans la loi de dispersion intrinsèque  $\gamma(h)$ . Les micro-échantillons ont des teneurs beaucoup moins différentes, en moyenne, que ne le voudrait la statistique classique, et par suite les volées de 10 tonnes ont une variance bien supérieure au millième de la variance des carottes de 10 Kgs.

Sous la forme (5), l'expression de la variance fait apparaître une loi d'additivité. Si l'on considère dans un champ  $V$  des panneaux  $V'$  et des échantillons  $v$ , et si l'on désigne par  $\sigma^2(V', V)$ ,  $\sigma^2(v, V)$  et  $\sigma^2(v, V')$  les variances de  $V'$  dans  $V$ , de  $v$  dans  $V$  et de  $v$  dans  $V'$ , on a immédiatement :

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

Cette formule porte le nom de formule de Krige. Elle a été établie par D.G. KRIGE<sup>(1)\*</sup> dans le cas où les teneurs obéissent à une loi de distribution (statistique) lognormale. En fait, sa validité n'est pas liée à une loi de distribution (statistique) particulière, mais uniquement à l'existence d'une loi de dispersion intrinsèque.

Outre la variance, on introduit aussi, en Géostatistique, la notion de covariance. Si  $y(M)$  et  $z(M+h)$  sont les teneurs de deux échantillons  $v$  et  $v'$  centrés en deux points  $M$  et  $M+h$ , on appelle covariance (dans  $V$ ) de  $y$  et  $z$  la fonction de  $h$  définie par :

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

Elle s'exprime au moyen du variogramme par une relation analogue à (4) :

$$(6) \quad \sigma_{yz} = F(V) - \frac{1}{VV'} \int_V dv \int_{V'} \gamma(k) dv$$

La deuxième intégrale représente la valeur moyenne de  $\gamma(k)$  lorsque les deux ex-

trémités du vecteur  $k$  balayent, respectivement, le volume  $v$  et le volume  $v'$  distants de  $h$ .

Comme cas particulier, nous examinerons le schéma de Wijsien isotrope. Il est défini par une fonction intrinsèque isotrope de la forme :

$$(7) \quad \gamma(r) = 3 \alpha \log r$$

où  $r = |h|$  représente le module de l'argument vectoriel  $h$ , c'est-à-dire la distance des deux points  $M$  et  $M + h$ . Lorsque le symbole  $\log$  représente le logarithme népérien, le paramètre  $\alpha$  prend le nom de dispersion absolue. Il caractérise, en effet, la dispersion des teneurs, indépendamment de la forme et du volume des échantillons et du gisement. Dans le cas particulier où le volume  $v$  des échantillons est géométriquement semblable au volume  $V$  du gisement, les formules (4) et (7) donnent :

$$(8) \quad \sigma^2 = \alpha \log \frac{V}{v}$$

Cette formule, ou formule de DE WIJS<sup>(1)\*</sup>, exprime un principe de similitude. Elle cesse, en général, d'être applicable dès que le gisement n'est pas géométriquement semblable aux échantillons. Cependant, à tout volume géométrique  $v$ , nous pouvons associer son équivalent linéaire  $\alpha$  défini par la relation :

$$(9) \quad \log d - \frac{3}{2} = \frac{1}{v^2} \int_v dv \int_v \log r dv'$$

Il résulte de (4) que l'échantillon  $v$  a même variance, dans n'importe quel gisement, que l'échantillon linéaire de longueur  $d$ . Si  $D$  et  $d$  désignent les équivalents linéaires du gisement et de l'échantillon, la variance se met sous la forme :

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

Les équivalents linéaires ont été calculés et tabulés pour un certain nombre de figures géométriques, et l'on dispose, en outre, de certaines formules d'approximation simples. Par exemple, pour un rectangle de côté  $a, b$ , on a :

$$d = a + b$$

---

(1)\* Le raisonnement originel de DE WIJS, qui constitue un exemple remarquable de passage de la statistique classique à la géostatistique, a servi de point de départ au développement de la théorie actuelle. On se reportera au TRAITÉ DE GEOSTATISTIQUE APPLIQUÉE, où l'on trouvera, en particulier, les références bibliographiques.

Pour un parallélogramme de cotés  $a$   $b$  et de surface  $S$  :

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

Pour un triangle, de côtés  $a$   $b$   $c$ , et de surface  $S$  :

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

Pour un trapèze de bases  $a = \frac{L + \ell}{2}$  et  $b = \frac{L - \ell}{2}$ , de médiane  $m$  et de surface  $S$ ,

$$d = \sqrt{L^2 + \ell^2 + m^2 - \frac{1}{3} \frac{\ell^2 m^2}{L^2}} + 2 S$$

Pour un parallélépipède rectangle de cotés  $a > b > c$  :

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

Pour un parallélépipède oblique d'arête  $r_1, r_2, r_3$ , de faces  $S_1, S_2, S_3$  et de volume  $V$ , on pose :

$$\left\{ \begin{array}{l} R^2 = r_1^2 + r_2^2 + r_3^2 \\ S^2 = S_1^2 + S_2^2 + S_3^2 \end{array} \right.$$

et on a l'équivalent approché suivant :

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

Cette notion d'équivalent linéaire permet une comparaison facile entre échantillonnages de natures différentes, au moins dans le cas, fréquent pour les gisements métalliques, où la loi de dispersion est de la forme (7)

#### 7/- Variances d'estimation et variance d'extension

L'un des problèmes pratiques les plus importants que la géostatistique ait à résoudre est celui de la détermination de l'ordre de grandeur de l'erreur possible commise dans l'estimation d'un gisement. Les propriétés générales des variables régionalisées nous indiquent qu'une telle erreur ne dépend pas seulement du nombre des

échantillons prélevés, mais, avant tout, de leurs formes, de leurs dimensions et de leurs implantations respectives, c'est-à-dire, globalement, de la géométrie des travaux miniers effectués. Ces indications prennent un sens précis grâce à la notion géostatistique de variance d'estimation. Supposons que, pour estimer la teneur réelle inconnue  $z$  d'un gisement ou d'un panneau  $V$ , nous disposions de la teneur  $x$  d'un certain réseau de travaux miniers  $TM$ . L'erreur d'estimation  $(z - x)$  a une valeur unique, et bien déterminée, quoiqu'inconnue, pour un panneau  $V$  donné et le réseau de  $TM$  que l'on y a implanté de manière préférentielle. Pour faire de cette erreur une variable régionalisée, la géostatistique considère le panneau ou le gisement  $V$  à estimer comme un panneau extrait d'un très grand gisement fictif  $K$  dans lequel est supposée régner la loi de dispersion intrinsèque  $\gamma(h)$  définie par le variogramme expérimental observé dans les travaux miniers  $TM$ . Nous verrons dans un instant que la forme et les dimensions attribuées à  $K$  n'interviennent pas réellement. Si l'on imagine que le panneau  $V$  à estimer se déplace dans le grand gisement  $K$  en entraînant avec lui les travaux miniers  $TM$  qui lui sont liés, l'erreur  $(z - x)$  apparaît comme une variable régionalisée de moyenne nulle et de variance :

$$(10) \quad \sigma^2 = \sigma_z^2 + \sigma_x^2 - 2 \sigma_{zx}$$

Cette variance, qui porte le nom de variance d'estimation, se calcule à partir des variances  $\sigma_z^2$ ,  $\sigma_x^2$  et de la covariance  $\sigma_{zx}$  des variables  $z$  et  $x$  dans le champ  $K$ , données elles-mêmes par des formules du type (4) et (6). Le champ  $K$  intervient dans l'expression de  $\sigma_z^2$ ,  $\sigma_x^2$  et  $\sigma_{zx}$  par la seule constante  $F(K)$  qui s'élimine de l'équation (10), de sorte que la variance d'estimation  $\sigma^2$  est indépendante du choix de  $K$ , et se calcule selon la formule :

$$(11) \quad \sigma^2 = \frac{2}{V V'} \int_V dV \int_{V'} \gamma(h) dV' - \frac{1}{V^2} \int_V dV_1 \int_V \gamma(h) dV_2 - \frac{1}{V'^2} \int_{V'} dV'_1 \int_{V'} \gamma(h) dV'_2$$

Dans (11),  $V$  désigne le volume du gisement à estimer et  $V'$  celui des travaux miniers  $TM$ . La variance d'estimation  $\sigma^2$  se calcule donc en intégrant la fonction intrinsèque  $\gamma(h)$  dans les volumes géométriques du gisement et des échantillons. De même que le variogramme donnait un contenu précis à la notion de zone d'influence d'un échantillon, on peut dire que la variance d'estimation (11) donne une mesure précise de "l'influence" des travaux miniers sur le gisement entier.

Dans les calculs pratiques, la formule (11) serait malaisée à mettre en oeuvre. Les travaux miniers constituent, en général, un réseau discontinu, dans lequel les échantillons peuvent eux-mêmes être prélevés de manière discontinue (par exemple rainurages prélevés à maille régulière dans les niveaux d'un filon reconnu par traçage à différents étages). Le volume  $V'$  intervenant en (11) est le volume discontinu constitué par le réseau des échantillons réellement prélevés et analysés. A chacun de ces échantillons distincts, on attribue traditionnellement une zone d'influence, au centre de laquelle il est implanté et dont il est censé représenter la teneur. L'erreur que l'on commet en étendant la teneur d'un tel échantillon distinct à sa zone d'influence peut être représentée par une variance du type (11), où  $V$  est le volume de la zone d'influence et  $V'$  celui de l'échantillon. Une telle variance porte le nom de variance d'extension élémentaire, et peut être tabulée, pour un  $\gamma(h)$  donné, en fonction des paramètres géométriques de l'échantillon et de sa zone d'influence. Moyennant certaines hypothèses d'approximation, on peut montrer que toute variance d'estimation du type (11) peut se calculer par composition de variances d'extension élémentaires.

En pratique, on distingue essentiellement deux cas. Pour une fonction  $\gamma(r)$  isotrope, le réseau d'échantillons élémentaires peut être isotrope (1)\* ou non. Comme exemple simple de réseau isotrope, citons le cas d'une maille carrée de sondages. Pour un réseau isotrope, les erreurs commises en étendant à chaque zone d'influence la teneur de son échantillon central peuvent être considérées comme indépendantes (c'est-à-dire de covariance géostatistique nulle). Dans ce cas, la variance d'estimation s'obtient en divisant la variance d'extension  $\sigma_E^2$  de chaque échantillon dans sa zone d'influence par le nombre  $N$  de ces zones d'influence.

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

Si, au contraire, le réseau n'est pas isotrope, on est conduit à regrouper les échantillons selon des lignes, ou des plans, de plus grande densité et à composer des variances d'extension de nature différente. Par exemple, soit un gisement filonien reconnu par des traçages eux-mêmes échantillonnés par rainurages. On considérera, en premier lieu, la variance d'extension  $\sigma_{E1}^2$  d'un rainurage dans le tronçon de traçage où il a été prélevé. Si  $N$  désigne le nombre total des rainurages, on voit que la variance d'estimation  $\frac{\sigma_{E1}^2}{N}$  représente l'erreur que l'on commet en étendant aux travaux miniers eux-mêmes la teneur déduite des rainurages. On considère ensuite la variance

(1)\*

Plus généralement, pour une fonction  $\gamma(h)$  quelconque, le réseau peut être, ou non, adapté à l'anisotropie de la fonction  $\gamma(h)$ . Pour les questions relatives aux différents types d'anisotropie, on se reportera au TRAITÉ DE GEOSTATISTIQUE ADDITIONNEL.

d'extension  $\sigma_{E_2}^2$  de la teneur (supposée parfaitement connue) d'un traçage dans sa zone d'influence. La zone d'influence est ici le panneau constitué par la réunion des deux demi-étages situés à l'amont et à l'aval immédiat du traçage. Si  $n$  est le nombre des niveaux traçés, la variance d'estimation  $\frac{1}{n} \sigma_{E_2}^2$  représente l'erreur que l'on commet en étendant au gisement entier la teneur moyenne, supposée réellement connue, des travaux miniers. La variance d'estimation résultante prend la forme :

$$(13) \quad \sigma^2 = \frac{1}{N} \sigma_{E_1}^2 + \frac{1}{n} \sigma_{E_2}^2$$

Il convient, en général, d'ajouter à cette expression une variance additionnelle, représentant les erreurs d'analyse et de prélèvement. Dans une telle expression le deuxième terme est, le plus souvent, largement prépondérant. La majeure partie de l'erreur commise provient de l'extension des travaux miniers au gisement. En particulier, il ne servirait à rien de multiplier indéfiniment le nombre  $N$  de prélèvements sans effectuer de travaux miniers supplémentaires, la variance d'estimation coïncidant pratiquement très vite avec la limite  $\frac{1}{n} \sigma_{E_2}^2$ , en dessous de laquelle elle ne peut pas descendre.

Des tables et des abaques donnant les valeurs numériques des variances d'extension élémentaires ont été établies<sup>(1)\*</sup> pour un certain nombre de fonctions intrinsèques (notamment pour une fonction de WIJSIENNE du type (7)) - et permettent un calcul rapide des variances d'estimation affectées aux différents schémas de reconnaissance par sondages ou travaux miniers.

A titre d'exemple, considérons un gisement filonien soumis à un schéma de Wiji-sien isotrope de type (7) et reconnu par traçages. Supposons que les traçages aient été suffisamment bien échantillonnés pour que le premier terme de (13) soit nul. Désignons par  $h$  la relevée comprise entre deux traçages consécutifs (comptée dans le plan du filon). On montre que la variance d'extension d'un traçage de longueur  $\ell$  dans son panneau d'influence  $\ell h$  est donnée par :

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

Cette formule n'est rigoureuse que si  $h$  est petit devant  $\ell$ , mais peut pratiquement être utilisée jusqu'à  $h = \ell$ . Pour  $h > \ell$ , elle doit être remplacée par une formule

---

(1)\* TRAITE DE GEOSTATISTIQUE APPLIQUEE - Tome I pour les fonctions DE WIJSIENNES, Tome III, pour le cas d'un effet de pépite.

différente. Supposons que les longueurs  $l_1, l_2, \dots, l_n$  des traçages soient toutes supérieures à  $h$ . La variance d'estimation s'obtient en pondérant la variance d'extension de chaque traçage dans son panneau d'influence par le carré de la surface de ce panneau, soit :

$$\sigma^2 = \frac{l_1^2 \sigma_{E1}^2 + l_2^2 \sigma_{E2}^2 + \dots}{(l_1 + l_2 + \dots)^2} = \alpha \frac{\pi}{2} h \frac{l_1 + l_2 + \dots + l_n}{(l_1 + l_2 + \dots + l_n)^2}$$

En désignant par  $S = h(l_1 + l_2 + \dots + l_n)$  la surface minéralisée reconnue et par  $L = l_1 + l_2 + \dots + l_n$  la longueur tracée totale, on obtient la formule remarquable :

$$(14) \quad \sigma^2 = \alpha \frac{\pi}{2} \frac{S}{L^2}$$

Une fois calculée une variance d'estimation  $\sigma^2$ , il reste à l'interpréter, pour les usages pratiques, sous la forme d'une fourchette d'erreur conventionnelle. On y parvient en attribuant à cette variance une signification probabiliste. En se référant implicitement à un modèle gaussien, on admettra que la teneur moyenne réelle du gisement est comprise, avec une probabilité de 95 %, dans la fourchette  $m \pm 2\sigma$ ,  $m$  étant la teneur estimée. Dans d'autres cas, en particulier si  $2\sigma$  n'est pas petit devant  $m$ , on se réfère implicitement à un modèle lognormal, et on prend la fourchette  $m \exp[\pm 2\sigma]$ . Ces références implicites à des modèles probabilistes sont largement arbitraires. En fait, la notion même de distribution statistique d'une erreur d'estimation est sans doute dépourvue de signification. La seule réalité possédant un sens physique objectif est la variance. C'est pourquoi nous parlons de fourchettes conventionnelles. Leur intérêt pratique provient de ce qu'elles donnent des erreurs possibles, une image plus intuitive que les variances elles-mêmes.

### 8/- Le Krigeage

Une deuxième application pratique d'importance majeure est fournie par le procédé géostatistique du krigeage. Ce procédé consiste à estimer la teneur d'un panneau en faisant une moyenne pondérée des échantillons disponibles, les uns intérieurs, les autres extérieurs au panneau à estimer. Si les teneurs de ces échantillons sont  $x_1, x_2, \dots, x_n$ , on cherche à évaluer la teneur  $z$  inconnue du panneau par un estimateur  $z^*$  linéaire de la forme :

$$z^* = \sum a_i x_i \quad \dots / \dots$$

Les poids  $a_i$  qu'il convient d'attribuer à chaque échantillon sont déterminés par deux conditions. La première condition exprime que  $z^*$  et  $z$  doivent avoir même valeur moyenne dans tout grand champ  $V$  et s'écrit :

$$\sum a_i = 1$$

La deuxième condition exprime que les  $a_i$  ont des valeurs telles que la variance d'estimation de  $z$  par  $z^*$ , ou variance de krigeage, prenne la valeur la plus petite possible. Elle s'explicité sous forme d'un système d'équations linéaires relativement aux  $a_i$ , dont les coefficients s'expriment à l'aide des variances et des covariances des échantillons et du panneau. Pour chaque fonction intrinsèque, il est ainsi possible de tabuler les coefficients et la variance de krigeage, pour différentes configurations et en fonction des paramètres géométriques du problème. Dans le cas d'un schéma DE WIJSIEN isotrope, de nombreuses configurations de sondages et de travaux miniers ont été tabulées. Nous donnons, à titre indicatif, un exemple sur la figure 3. La configuration étudiée, utile pour l'estimation d'un gisement reconnu par sondages, ou l'exploitation sélective d'une carrière, concerne, pour un réseau de sondages à maille carrée, le krigeage du panneau d'influence d'un sondage  $A$  à l'aide de la teneur de ce sondage central  $A$  et des 8 sondages les plus proches regroupés en deux "auréoles"  $B_1 B_2 B_3 B_4$  et  $C_1 C_2 C_3 C_4$ . Si  $u$  désigne la teneur de  $A$ , et  $v$  et  $w$  les teneurs moyennes des sondages  $B$  et des sondages  $C$ , l'estimateur à utiliser est de la forme :

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

Sur la figure 3, on porte en abscisse le rapport  $\frac{h}{a}$  de la puissance  $h$  de la formation à la maille  $a$  des sondages, et on lit sur les courbes les valeurs numériques de  $\lambda$  et  $\mu$ , ainsi que celles de l'expression  $\frac{1}{3\alpha} \sigma_K^2$ . Il suffit de multiplier cette dernière par le triple de la dispersion absolue  $3\alpha$  pour obtenir la variance de krigeage.

En théorie, il y a intérêt à kriger chaque panneau par la totalité des échantillons, intérieurs ou extérieurs au panneau, prélevés dans le gisement. Outre la très grande complication des calculs auxquels on est très vite conduit, on s'aperçoit, sur les exemples numériques, qu'il est le plus souvent inutile de tenir compte des échantillons lointains. Il suffit, en général, d'une ou deux auréoles d'échantillons exté-

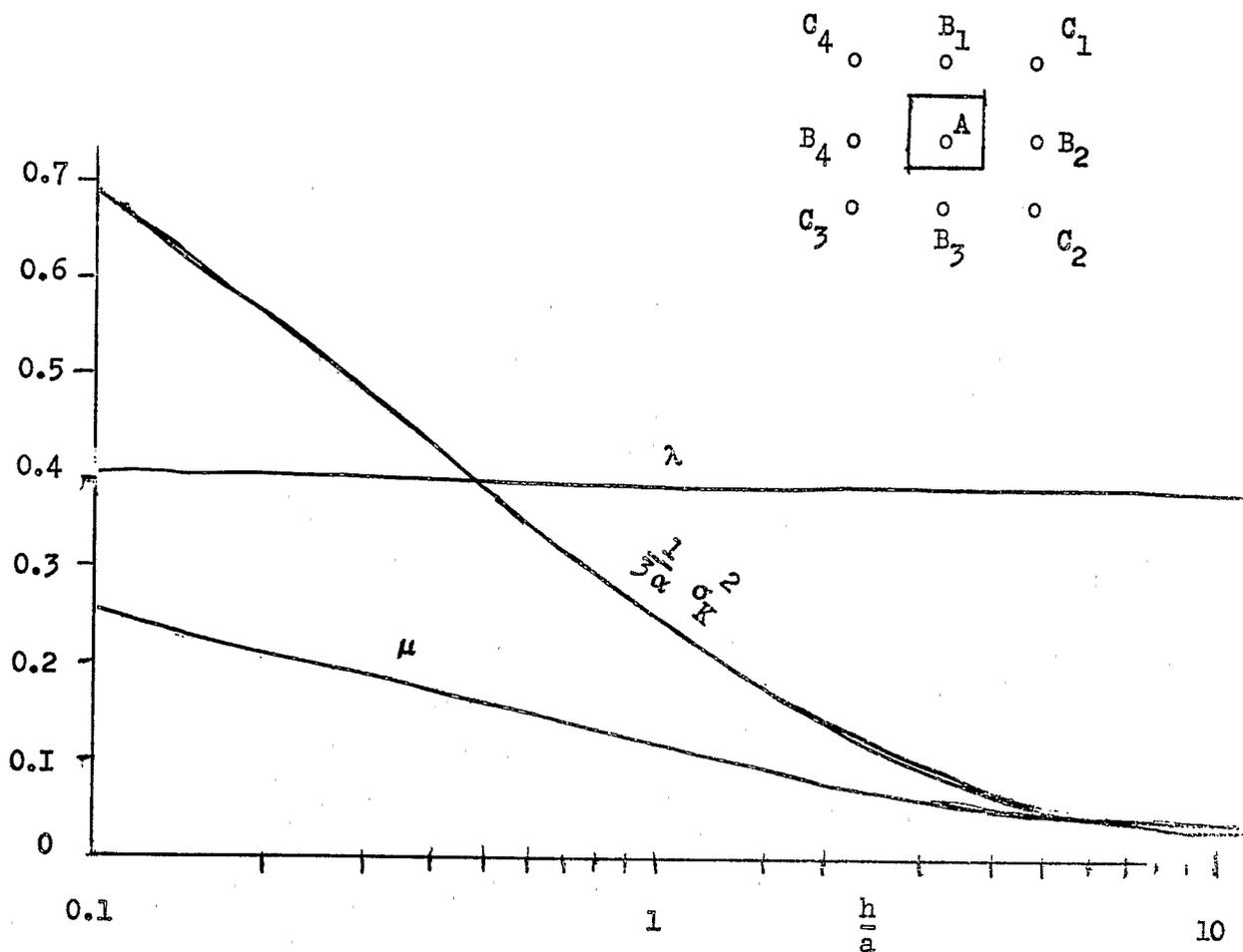


Figure 3

rieures proches pour arrêter la quasi totalité de l'influence du reste des échantillons extérieurs. Il en est ainsi, en particulier, dans la configuration étudiée sur la figure 3, où les deux auréoles des sondages B et C constituent un écran à peu près total vis-à-vis de tous les autres sondages extérieurs. On peut même voir que, pour les grandes valeurs du rapport  $\frac{h}{a}$ , le poids  $\mu$  de la deuxième auréole devient faible, de sorte que l'auréole constituée par les quatre sondages B constitue un écran à elle seule. Cet effet d'écran est un phénomène général, et joue un grand rôle dans la théorie du krigeage.

Du point de vue pratique, l'intérêt du krigeage est double. Tout d'abord, et cela résulte de la définition même de ce procédé, il permet d'obtenir la meilleure estimation possible d'un panneau donnée, c'est-à-dire l'estimation de variance minimale. En améliorant la prévision des productions mensuelles, par exemple, des différents quartiers d'une mine, il peut rendre de grands services, en particulier, dans le cas où l'exploitant est tenu de livrer des produits de caractéristiques aussi constantes que possible.

Pour appréciable qu'il soit, cependant, le gain de précision qu'apporte le krigeage ne justifierait pas toujours la mise en oeuvre des calculs qu'il comporte. Dans la plupart des cas, l'intérêt majeur du procédé ne vient pas de ce qu'il réduit les variances d'estimation mais de ce qu'il permet d'éliminer une cause d'erreur systématique. Il est rare, en effet, qu'un gisement soit exploitable dans sa totalité. On n'en retient, comme rentables, que certains panneaux sélectionnés au vu des teneurs des échantillons qu'on y a prélevés. D.G. KRIGE<sup>(1)\*</sup> a montré que les résultats des seuls échantillons intérieurs conduisaient, nécessairement, à surestimer les panneaux riches et à sous-estimer les panneaux pauvres. La notion géostatistique de krigeage permet d'interpréter facilement ce phénomène et d'en corriger les effets. Du fait même que l'on sélectionne un panneau riche, l'auréole des échantillons extérieurs est, en général, à teneur plus faible que les échantillons intérieurs; et cependant son influence sur le panneau à estimer n'est pas négligeable, puisque le krigeage lui attribue un poids non nul. En ne tenant pas compte de cette auréole extérieure, on introduit donc nécessairement une cause d'erreur systématique par excès, que le krigeage permet d'éliminer.

(1)\* Le raisonnement originel de DG KRIGE constitue un deuxième exemple de passage implicite de la statistique classique à la géostatistique. Il repose essentiellement sur le fait que la variance d'un panneau est toujours plus faible que celle de l'échantillonnage intérieur. Voir références au TRAITÉ DE GÉOSTATISTIQUE APPLIQUÉE.

L'effet de pépite

Les règles générales dégagées dans les paragraphes précédents peuvent présenter des exceptions apparentes en présence d'un effet de pépite prononcé. L'effet de pépite a été défini, sur la figure 1, par un variogramme présentant une discontinuité à l'origine, et correspondant à une variable régionalisée qui ne possède pas la continuité en moyenne. Il peut être d'origine purement granulométrique, comme dans les gisements d'or et de diamant, ou, plus généralement, manifester l'existence de microstructures discontinues. Dans un stockwerk, la présence de veinules ou de microfractures à remplissage riche peut donner naissance à un effet de cette nature. Dans les gisements d'or, les teneurs de deux échantillons très voisins, ou même contigus, peuvent différer notablement si l'un d'entre eux contient " par hasard " une grosse pépite. L'effet est d'autant plus marqué que les prélèvements sont de plus petite taille, et peut atteindre une ampleur énorme pour des prélèvements de quelques litres. Il suffit d'une translation de quelques millimètres de son support géométrique pour qu'un échantillon contienne, ou non, une grosse pépite capable de modifier sa teneur dans un rapport de 1 à 10 ou de 1 à 100. Le fait qu'une pépite marginale soit, ou non, incluse dans un prélèvement apparaît comme un phénomène tout à fait aléatoire. En réalité, cependant, le comportement de la teneur ne peut être considéré comme aléatoire que localement. S'il n'en était pas ainsi, les panneaux de quelques milliers de tonnes, sur lesquels les pépites marginales n'ont plus d'action décelable, présenteraient des teneurs quasi constantes (leur variance étant alors un million de fois plus faible que celles d'échantillons de quelques kilos). On sait bien qu'en réalité - même pour les gisements d'or - il y a des panneaux riches et des panneaux pauvres. Mais, localement, cet effet aléatoire peut être si prononcé qu'il masque entièrement la régionalisation sous-jacente. On est frappé par la fréquence dans la littérature consacrée à ces gisements de certains termes comme aberrant, teneur monstre, teneur mammoth, etc ... qui font tous allusion à une hypothétique anomalie dans le comportement de la minéralisation. Certes, c'est avec raison que des statisticiens classiques ont fait observer qu'il n'y avait aucune anomalie réelle, et que ces teneurs monstrueuses, existant réellement dans le gisement, apparaissaient de temps à autre à l'échantillonnage avec une fréquence déterminée par les lois du hasard. Historiquement, c'est sans doute la fascination exercée par cet effet de pépite qui a longtemps empêché de faire une distinction claire entre les notions de variables régionalisées et aléatoires. Du point de vue de la géostatistique, il apparaît qu'en fait la terminologie naive n'avait pas tort de sug-

... / ...

gérer l'existence de quelque anomalie : seulement, le fait aberrant ne réside pas dans la présence de quelques teneurs " anormalement " élevées, mais bien plutôt dans le comportement localement aléatoire de toutes les teneurs, fortes ou faibles, et dans la détérioration du réseau des corrélations spatiales. Pour reprendre la terminologie naive, ce ne sont pas les teneurs mammoth elles-mêmes qui sont aberrantes, mais le fait qu'elles soient dépourvues de zone d'influence. Et, de leur côté, les statisticiens classiques avaient raison de souligner que l'apparition de ces teneurs aberrantes est régie par les lois du hasard. Mais ils n'avaient pas vu que le phénomène ne peut être considéré comme aléatoire que localement.

Sans chercher à faire un exposé<sup>(1)\*</sup> systématique, indiquons rapidement comment la géostatistique permet de représenter un effet de pépite. Examinons le demi-variogramme  $\gamma(r)$  représenté comme troisième type sur la figure 1. Nous nous limiterons ici au cas où  $\gamma(r)$  est une fonction isotrope, (c'est-à-dire ne dépendant que du module  $r$  de l'argument vectoriel  $h$ ). La discontinuité  $C$ , ou saut, observée à l'origine sur le  $\gamma(r)$  d'une variable à support ponctuel porte le nom de constante de pépite. Si l'on désigne par  $H(r)$  la fonction d'Heaviside, définie par :

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

le demi-variogramme peut être séparé en deux composantes :

$$(15) \quad \gamma(r) = C H(r) + \gamma_1(r)$$

La première composante,  $C H(r)$ , représente l'effet de pépite à l'état pur. La deuxième,  $\gamma_1(r)$ , qui est continue à l'origine, représente la régionalisation sous-jacente. Toutes les variances et covariances que l'on a besoin d'introduire peuvent alors se calculer comme si la variable  $x(M)$  à support ponctuel était la somme :

$$(16) \quad x = x_0 + \xi$$

d'une variable régionalisée théorique  $x_0$  obéissant à la loi de dispersion  $\gamma_1(r)$  continue à l'origine, et d'une variable aléatoire  $\xi$  de moyenne nulle et de variance  $C$ ,

(1)\* Voir TRAITE DE GEOSTATISTIQUE APPLIQUEE, Tome III.

les  $x_0$  et les  $\mathcal{E}$  étant indépendants, et les  $\mathcal{E}$  affectés à deux points distincts, même très proches, étant également indépendants. Si l'on se borne à étudier la variation de la teneur ponctuelle  $x$  au voisinage d'un point donné, c'est-à-dire si l'on ne considère que des valeurs petites de la distance  $r$ ,  $\gamma_1(r)$  varie très peu, de sorte que  $\gamma(r)$  peut être regardé comme une constante égale à  $C$ . Les variations observables localement sont dues presque exclusivement à  $\mathcal{E}$ . C'est ce que nous exprimons en disant que la variable régionalisée se comporte localement comme une variable aléatoire. Mais à plus grande échelle, c'est-à-dire pour des valeurs plus grandes de  $r$ , la croissance de la composante continue  $\gamma_1(r)$  ~~ne peut plus être négligée~~ et la régionalisation de  $x_0$  se manifeste de façon sensible.

En réalité, la fonction d'Heaviside ne représente pas de façon tout à fait satisfaisante l'aspect aléatoire du comportement d'une variable ponctuelle. A moins de supposer la constante  $C$  infinie, le terme  $C H(r)$  perdrait toute influence sur la variance d'un échantillon de taille non nulle. Il s'élimine automatiquement dans la formule (4). Cela signifie que la valeur moyenne des variables aléatoires indépendantes  $\mathcal{E}$  contenues en nombre infini dans un support non ponctuel a nécessairement une variance nulle. En réalité, la notion d'une variable aléatoire  $\mathcal{E}$  à support ponctuel est dépourvue de signification physique. Le phénomène physique réel ne comportera jamais une véritable discontinuité à l'origine, mais une zone de transition rapide au voisinage de  $r = 0$ . On est conduit à remplacer  $H(r)$  par la fonction de transition  $T(r,a)$ , définie par :

$$\left\{ \begin{array}{l} T(r,a) = \frac{r}{a} \quad \text{pour } r \leq a \\ T(r,a) = 1 \quad \text{pour } r > a \end{array} \right.$$

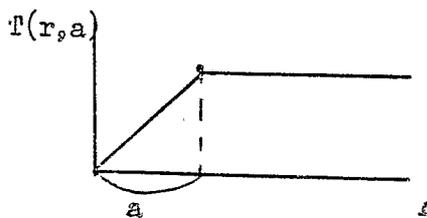


Figure 4

La constante  $a$ , ou portée, donne l'ordre de grandeur de la zone de transition, c'est-à-dire la dimension des pépites. Dans le cas de pépites homogranulaires de même volume  $u$ , on montre que l'on a :

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

La fonction intrinsèque  $\gamma(r)$  d'une teneur ponctuelle se décompose de la manière suivante :

$$\gamma(r) = C T(r,a) + \gamma_1(r)$$

$C$  est encore la constante de pépite, et  $\gamma_1(r)$  la composante continue. La teneur ponctuelle  $x$  s'interprète par une somme du type (16), dans laquelle  $\xi$  est une variable régionalisée admettant la fonction intrinsèque  $CT(r,a)$ . Cette fois les  $\xi$  ne sont indépendants que pour des distances supérieures à la portée  $a$ . Aux distances inférieures, ils sont liés par un variogramme linéaire. Par suite, l'effet de pépite va se répercuter sur les échantillons de taille  $v$  non nulle. Si  $v$  est grand devant le volume  $a^3$  des grains, la zone de transition est diluée dans le volume  $v$  d'intégration, et l'effet de pépite donne naissance à une variance additionnelle de la forme  $\frac{a^3}{v}$ . En effet, désignons par  $\sigma_p^2$  (variance de pépite) la contribution de  $CT(r,a)$  à la variance de l'échantillon  $v$ . Conformément à (4), nous devons calculer des intégrales du type :

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

Si toutes les dimensions de  $v$  sont grandes vis-à-vis de  $a$ , chacun des points intérieurs à  $v$  apporte à l'intégrale sextuple la contribution :

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

Ceci n'est vrai que pour les points situés à une distance supérieure à  $a$  de la frontière de  $v$ . Mais, si  $v$  est grand, les points frontaliers n'interviennent que par des termes d'ordre supérieur. Avec cette approximation, l'intégrale sextuple est égale à  $C(1 - \frac{\pi a^3}{3v})$ . Comme l'intégrale dans  $V$  se calcule de la même manière, il vient finalement :

$$(17) \quad \sigma_p^2 = C \frac{\pi}{3} \left[ \frac{a^3}{v} - \frac{a^3}{v} \right]$$

En pratique,  $\frac{a^3}{v}$  est négligeable, et la variance de pépite est en  $\frac{a^3}{v}$ , c'est-à-dire en raison inverse du nombre de grains contenus dans l'échantillon. C'est là une circonstance très générale et très caractéristique. Chaque fois qu'il y a effet de pépite, c'est-à-dire chaque fois qu'une variable régionalisée présente un comportement localement aléatoire, les échantillons macroscopiques sont affectés d'une variance addition-

nelle, dite variance de pépite, inversement proportionnelle à leur taille.

La variance de ces échantillons apparaît comme la somme :

$$\sigma^2 = \sigma_p^2 + \sigma_\theta^2$$

de la variance de pépite et de la variance théorique  $\sigma_\theta^2$ , calculée à l'aide de la composante continue  $\gamma_1(r)$  de la fonction intrinsèque. Lorsque  $v$  augmente, la variance théorique décroît beaucoup plus lentement que la variance de pépite. En présence d'un très fort effet de pépite, il peut arriver que  $\sigma_p^2$  soit largement prédominant pour des échantillons de quelques kilo. La régionalisation sous-jacente est presque entièrement masquée à l'échelle de ces échantillons. Si l'on se contente de faire varier le volume  $v$  dans un intervalle de quelques litres à quelques dizaines de litres, les variations expérimentalement observables de la variance seront celles de la variance de pépite seule, et l'on risquera de conclure que la variance est en raison inverse du volume. Mais si l'on passe à des échantillons de quelques dizaines de mètres cube, le terme  $\sigma_p^2$  s'amortit et disparaît, et la variance théorique  $\sigma_\theta^2$  devient prédominante. La régionalisation sous-jacente fait à nouveau sentir son effet, et la variance continue à décroître lorsque  $v$  augmente, mais beaucoup plus lentement que  $\frac{1}{v}$ .

Nous avons un peu insisté sur la théorie de l'effet de pépite pour montrer, sur un exemple crucial, comment les concepts de la géostatistique permettent de retrouver des résultats locaux que l'on obtient classiquement par le raisonnement statistique ordinaire (variance de pépite inversement proportionnelle au volume), mais en les insérant dans la perspective globale d'une régionalisation sous-jacente. En ce qui concerne l'utilisation pratique de cette théorie, nous indiquerons sommairement les deux points suivants :

- En présence d'un effet de pépite, les variances d'extension et d'estimation sont toutes majorées d'un terme  $C \frac{\pi a^2}{3 v}$ , inversement proportionnel au volume total des échantillons utilisés, donc, en particulier, au nombre  $n$  de ces échantillons. A cet égard, la variance additionnelle d'estimation due à l'effet de pépite se comporte comme les variances de prélèvement et d'analyse, et peut être regroupée avec elles.

- En ce qui concerne le krigeage, l'effet de pépite a pour effet de lever partiellement tous les écrans. On est conduit, pratiquement, à utiliser des formes particulières de krigeage, dites "krigeage aléatoire", qui ne sont autres que celles qui

furent originellement proposées par D.G. KRIGE lui-même à propos du gisement d'or du Rand, gisement où, justement, l'effet de pépite doit être très prononcé.

10/- Recherche d'optimum dans la reconnaissance d'un gisement.

Sous la forme de variances d'estimation, la géostatistique est capable de donner une mesure précise de la quantité d'information obtenue sur un gisement à l'issue d'un ensemble donné de travaux de reconnaissance. Ces travaux coûtent cher, en général et l'on doit mettre en balance leur prix de revient avec la valeur économique de l'information qu'ils apportent. On voit ainsi apparaître la possibilité de déterminer le volume optimum des crédits qui doivent être consacrés à la reconnaissance d'un gisement donné, et, en particulier, de déterminer le moment où il convient d'arrêter les recherches, et de prendre une décision, positive, ou négative, quant à la mise en exploitation du gisement. Ces méthodes, qui permettent de résoudre, au moins partiellement, l'un des problèmes majeurs posés par la pratique de la recherche minière, seront publiées par ailleurs, et ne peuvent pas être exposées ici. Soulignons seulement, et ce sera notre conclusion, qu'elles apparaissent comme le prolongement naturel de la géostatistique. La possibilité de leur mise au point était liée à l'éclaircissement préalable, et à l'approfondissement scientifique des diverses notions qui ont été sommairement résumées dans cet article.

G. MATHERON.