

# ANNALES DES MINES

DANS CE NUMÉRO :

D-152

ÉCHANTILLONNAGE SYSTÉMATIQUE  
DES GISEMENTS



MOUVEMENT DES ÉPONTES  
ÉTUDE DE LA LOI DE CONVERGENCE



DÉVELOPPEMENTS D'OUTRE-MER  
ET BESOINS CORRÉLATIFS DE NAVIRES

600 Fr.

SEPTEMBRE 1957

Revue fondée en 1795

# Théorie lognormale de l'échantillonnage systématique des gisements

G. MATHERON,

Ingénieur au Corps des Mines

**D**ans notre numéro de décembre 1955, G. MATHERON a exposé la théorie lognormale et son application à l'échantillonnage aléatoire des gisements. Mais dans la pratique, l'échantillonnage a toujours un caractère systématique, et son assimilation à un processus aléatoire est largement inexact. Par ailleurs, rien n'indique a priori qu'un échantillonnage aléatoire, préconisé pourtant par certains statisticiens, soit supérieur aux procédés systématiques pratiqués depuis toujours par les mineurs.

Dans ces conditions, G. MATHERON s'est proposé d'établir des formules simples permettant de déterminer (toujours dans le cadre de la théorie lognormale) l'efficacité d'un procédé d'échantillonnage, non plus aléatoire par hypothèse, mais quelconque.

Ces déterminations nécessitent de tenir compte de la forme du gisement et des échantillons ainsi que de la place de ceux-ci. Elles sont basées sur une généralisation de la formule de DE WIJS. On sait que cette formule donne la variance d'un échantillon lorsque celui-ci est géométriquement semblable au gisement. L'auteur, moyennant une hypothèse sur l'expression de la variance d'un échantillon circulaire, établit la formule valable pour la co-variance de deux échantillons infiniment petits et en déduit la formule de la variance d'un échantillon macroscopique de forme quelconque.

Il applique la théorie à trois cas essentiellement pratiques :

1<sup>o</sup> Cas de l'échantillonnage périphérique des panneaux. La teneur de la périphérie est une estimation sans biais de la teneur du panneau quand celui-ci est circulaire et, lorsque la teneur de la périphérie est fixée, la variance de la teneur du panneau est indépendante de sa dimension, et égale à la moitié de la « dispersion absolue ». L'auteur étudie en outre les gains de précision à attendre de prélèvements supplémentaires à l'intérieur des panneaux, notamment dans les cas (or, uranium) où la dispersion absolue est forte. Il donne un tableau des coefficients de sécurité au seuil de 95 % pour différents types de gisements;

2<sup>o</sup> Cas de sondages à mailles rigides : avec l'étude de l'influence d'une anisotropie du gisement;

3<sup>o</sup> Cas de reconnaissance d'un gisement par travaux miniers. L'auteur établit un résultat souvent méconnu : la précision de l'évaluation, à égalité de surface reconnue, est proportionnelle à la longueur totale des traçages (et non pas à la racine carrée de cette longueur). M. MATHERON donne, pour différentes valeurs de la dispersion absolue la longueur de traçages nécessaires pour une précision de 5 %.

## INTRODUCTION

Au stade actuel de son développement, la théorie de la loi lognormale (1) permet, grâce au principe de similitude et à la formule DE WIJS qui s'en déduit, de résoudre tous les problèmes posés par l'échantillonnage aléatoire des gisements. Dans la pratique, cependant, tous les modes d'échantillonnage effectivement employés, qu'il s'agisse de sondages à maille rigide ou de reconnaissance par travaux miniers, présentent au plus haut point le caractère systématique. Mis en face du problème concret posé par la pratique minière, la première réaction du statisticien est d'assimiler de façon simpliste le procédé employé à un échantillonnage aléatoire pur. Une telle assimilation entraîne vite des conclusions inacceptables et contraires à l'intuition du mineur : chaque tronçon de carotte, ou chaque volée dans un niveau étant considéré comme un échantillon aléatoire, la précision de l'évaluation du gisement paraît en effet proportionnelle à la racine carrée du nombre  $N$  de semblables échantillons. On pourrait donc s'imaginer augmenter indéfiniment la précision en découpant les carottes ou les volées en tronçons de plus en plus petits analysés séparément. En réalité, les tronçons successifs d'un même sondage ne sont pas indépendants les uns des autres et la multiplicité illusoire des données n'entraîne aucun gain de précision. La nature même du problème posé oblige à prendre en considération, outre la taille et le nombre des échantillons, leurs positions et leurs distances mutuelles. De ce point de vue, la théorie de l'échantillonnage stratifié au hasard, dans la mesure même où il se rapproche d'une maille rigide de sondage, représente un progrès considérable. Cependant des procédés aussi usuels que l'échantillonnage périphérique des panneaux ou la reconnaissance d'un gisement par traçages restent en dehors de la portée d'une telle théorie. Tout ce que le statisticien peut nous dire, c'est la précision avec laquelle est connue la teneur de la périphérie ou de la galerie elle-même : le passage de la périphérie au panneau comporte un saut dans l'inconnu dont des auteurs comme KRIGE sont parfaitement conscients, et que la statistique pure ne permet pas de chiffrer.

A la racine de ce problème se trouve l'ambiguïté de la nature de la loi de distribution des teneurs, loi physique dont les propriétés s'énoncent en langage statistique. La teneur n'est pas, en réalité, une variable aléatoire. Une portion  $v$  du gisement, une fois sa forme

et sa position fixées, a une teneur unique, physiquement déterminée et nullement aléatoire. Cependant si l'on considère le gisement comme constitué par la juxtaposition d'un nombre  $N$ , grand, mais non infini de volumes  $v_i$  et si l'on classe sous forme d'histogrammes les  $N$  teneurs ainsi obtenues, on obtient quelque chose qui ressemble beaucoup à une loi statistique continue. En effet, si  $x_1, x_2, \dots, x_N$  sont les teneurs de chacun des volumes  $v_i$  et si on tire au hasard un volume  $v$ , on a la probabilité  $\frac{1}{N}$  de tomber par exemple sur le volume n° 1, c'est-à-dire sur la teneur  $x_1$ .  $N$  étant grand, une telle loi de probabilité diffère très peu d'une loi continue, que l'expérience montre être lognormale. Cependant, outre les  $N$  volumes  $v_i$ , il existe virtuellement dans le gisement, chevauchant sur plusieurs  $v_i$ , une infinité de volumes  $v$  de même forme. On se rend compte que le tirage au sort d'un volume  $v$  fait de la teneur  $x$  une variable aléatoire obéissant à une loi réellement continue, très peu différente de la loi discontinue précédente.

Seul donc le tirage au sort permet de considérer la teneur, grandeur physique à répartition spatiale, comme une variable aléatoire. Un tel point de vue implique la possibilité de voir amener par deux tirages successifs le même volume  $v$ , ou du moins deux volumes  $v$  largement chevauchant. Ce n'est qu'à ce prix que l'on dispose d'une population infinie. Ceci étant, les teneurs  $x_1$  et  $x_2$  amenées par deux tirages au sort successifs sont deux variables aléatoires indépendantes. Cela signifie que la valeur moyenne du coefficient de corrélation de deux volumes  $v$  du gisement est nulle. Si ces deux volumes coïncident, ce coefficient est égal à  $+1$ ; il doit donc prendre des valeurs négatives lorsque les deux volumes sont très éloignés. Dans une maille rigide de sondage, où l'on exclut toute possibilité de chevauchement ou de grande proximité, on doit s'attendre à une corrélation moyenne négative, donc à une meilleure précision finale que par simple tirage au sort. Au contraire, les différentes volées d'une galerie ou les différents tronçons d'un sondage, plus rapprochés que la moyenne, ont une corrélation positive et conduisent à une moins bonne précision que le tirage au sort.

Le but ultime d'un échantillonnage n'est pas la connaissance de la loi de distribution statistique d'une variable aléatoire, mais la représentation approchée de la répartition géographique d'une grandeur physique (teneur, puissance ou accumulation). Il est commode de résumer la répartition géographique des teneurs dans une zone donnée (panneau, quartiers, etc.) sous

(1) Cf. « Application des méthodes statistiques à l'évaluation des gisements », *Annales des Mines*, décembre 1955. Les notations employées ici sont les mêmes que dans cet article, auquel le lecteur est prié de se reporter pour la définition de la loi lognormale et de ses propriétés.

forme d'histogrammes des teneurs, documents où l'on fait abstraction de la position des volumes  $v$  constitutifs de la zone choisie. Seul le mode de présentation est statistique, la signification d'un tel histogramme étant celle d'une loi physique. Ce n'est que par le biais du tirage au sort d'un volume  $v$  que cet histogramme peut prendre le sens d'une loi de probabilité. Telle est la raison de la tendance invincible qui pousse le statisticien à assimiler tout mode d'échantillonnage à un processus aléatoire. Certains statisticiens, érigeant naïvement en règle générale leur impuissance à saisir la réalité physique, soutiennent la supériorité des échan-

tillonnages aléatoires (stratifiés ou non) sur les mailles rigides — supériorité affirmée mais non démontrée. Les vieux mineurs, cependant, refusent de prendre au sérieux ces considérations et persistent à utiliser les mailles rigides. La théorie développée ci-dessous donne, en fait, raison aux vieux mineurs.

Nous nous proposons, dans ce qui suit, d'établir des formules simples permettant de déterminer l'efficacité de n'importe quel mode d'échantillonnage, systématique ou non, en tenant compte de la forme et de la position dans le gisement des échantillons prélevés.

## I. — GÉNÉRALISATION DE LA FORMULE DE DE WIJS

L'EFFICACITÉ d'un échantillonnage est liée à la variance de la teneur de l'échantillon prélevé. Cette variance est donnée, en principe, par la formule de DE WIJS :

$$(1) \quad \sigma^2 = \alpha L \frac{S}{s}$$

où  $S$  et  $s$  sont les aires (1) du gisement et de l'échantillon, tandis que  $\alpha$  est un paramètre intrinsèque, la dispersion absolue, dont la valeur est peut-être caractéristique du type de la minéralisation. La validité de cette formule est cependant soumise à une certaine condition de similitude géométrique : il faut que les aires  $S$  et  $s$  soient à peu près de même forme. Si l'échantillon  $s$  est constitué par l'ensemble d'un traçage, la formule (1) n'est certainement pas valable; elle conduirait à une valeur beaucoup trop forte. Il est donc essentiel d'écrire cette formule sous une forme plus générale, indépendante de toute condition de similitude.

Au lieu de considérer la variance d'un seul échantillon, donnons-nous, dans un gisement  $S$  d'aire et de forme connues, deux échantillons  $S_1$  et  $S_2$  dont la taille, la forme géométrique et la disposition mutuelle sont également connues. Désignons par  $x_1$  et  $x_2$  les teneurs de ces échantillons, et cherchons à calculer la covariance, ou moment rectangulaire  $\sigma_{x_1, x_2}$  de ces deux variables. Il est bien précisé que,  $\gamma_1$  et  $\gamma_2$  étant les médianes de  $x_1$  et  $x_2$ , la covariance  $\sigma_{x_1, x_2}$  se définit comme étant la valeur moyenne de l'expression  $L \frac{x_1}{\gamma_1} \cdot L \frac{x_2}{\gamma_2}$  lorsque la figure constituée par l'ensemble  $S_1 + S_2$  se déplace sans se déformer, de façon à occuper toutes les positions

possibles dans  $S$ , ou, si l'on préfère, lorsque cette même figure indéformable est implantée au hasard dans  $S$ . En particulier, si  $S_2$  coïncide avec  $S_1$ , la covariance ne sera autre que la variance  $\sigma_{x_1}^2$  de l'échantillon  $S_1$ .

Une aire telle que  $S_1$  peut toujours être décomposée par la pensée en un grand nombre d'aires élémentaires  $dS_1$ . Soient donc  $dS_1$  et  $dS_2$  deux échantillons élémentaires de teneurs  $u_1$  et  $u_2$ , prélevés dans  $S_1$  et  $S_2$ . Désignons par  $\sigma_{u_1, u_2}$  la covariance de ces échantillons élémentaires. Il est facile de voir que la covariance  $\sigma_{x_1, x_2}$  est égale à la valeur moyenne de  $\sigma_{u_1, u_2}$  lorsque  $dS_1$  et  $dS_2$  occupent toutes les positions possibles dans  $S_1$  et  $S_2$ . En effet,  $Lx_1$ , par exemple, est égal à un facteur constant près indépendant de la position de  $S_1$ , à la valeur moyenne de  $Lu_1$  dans  $S_1$ . On écrira donc :

$$(2) \quad \sigma_{x_1, x_2} = \frac{1}{S_1 S_2} \iint_{S_1} dS_1 \iint_{S_2} \sigma_{u_1, u_2} dS_2$$

Une intégration quadruple étendue aux domaines  $S_1$  et  $S_2$  permet donc de calculer la covariance, si compliquées que soient les formes de ces domaines, pourvu que l'on connaisse la covariance élémentaire  $\sigma_{u_1, u_2}$ .

### La covariance élémentaire

Le problème fondamental est donc de connaître la covariance  $\sigma_{u_1, u_2}$  des échantillons élémentaires  $dS_1$  et  $dS_2$ . Comme il s'agit d'aires infiniment petites, on peut admettre que la forme géométrique perd toute influence et que la covariance ne dépend que de la distance  $r$  des éléments :

$$(3) \quad \sigma_{u_1, u_2} = f(r)$$

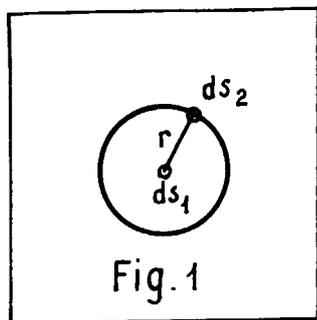
Nous ne connaissons pas la forme de la fonction  $f(r)$ . Cependant si nous introduisons le cercle de rayon  $r$  centré sur  $dS_1$ , dont la teneur  $z$  a une variance  $\sigma_z^2$

(1) Dans ce qui suit nous ne parlerons que des gisements stratifiés, c'est-à-dire à deux dimensions (couches, filons, etc.). Il va de soi que la théorie peut se généraliser au cas des gisements à trois dimensions, c'est-à-dire des amas.

dans S, le principe de similitude montre que l'on a :

$$(4) \quad \sigma_{u_1, u_2} = \sigma_s^2 + K$$

K est une constante indépendante de r, car, quel que soit r, la figure constituée par le cercle et les deux aires infiniment petites reste semblable à elle-même. En ce



qui concerne  $\sigma_s^2$ , qui dépend de r, la formule (1) ne lui est peut-être pas applicable. Admettons cependant que cette variance est de la forme plus générale :

$$\sigma_s^2 = K' - 2\alpha Lr$$

K' étant une autre constante. On est ainsi conduit à poser la formule suivante, dite de la covariance élémentaire :

$$(5) \quad \sigma_{u_1, u_2} = A - 2\alpha Lr$$

La constante A est indépendante de r, mais dépend de la taille et de la forme du gisement S. Il va de soi que les considérations précédentes ne constituent pas une démonstration de la formule (5) et n'ont d'autres buts que d'en expliquer le sens. Cette formule est en réalité un postulat qui recevra sa justification théorique, dans la mesure où elle permettra de généraliser la formule de DE WIJS, et dont le bien-fondé ne pourra être affirmé qu'après confrontation avec l'expérience.

Les formules (2) et (5) permettent de calculer la covariance de deux échantillons macroscopiques quelconques et, en particulier, la variance d'un échantillon de forme quelconque. La constante A se calcule aisément, si l'on remarque que, deux échantillons élémentaires tirés au sort dans S étant indépendants, la valeur moyenne de  $\sigma_{u_1, u_2}$  étendue à toutes les positions possibles de ds<sub>1</sub> et ds<sub>2</sub> est nécessairement nulle :

$$(6) \quad A = \frac{2\alpha}{S^2} \iint_S dS_1 \iint_S Lr dS'$$

### Formule de De Wijs généralisée

Compte tenu de cette valeur de la constante A, la variance  $\sigma_s^2$  de la teneur x d'un échantillon macroscopique s s'écrit :

$$(7) \quad \sigma_s^2 = \frac{2\alpha}{S^2} \iint_S dS \iint_S Lr dS' - \frac{2\alpha}{S^2} \iint_S ds \iint_S Lr ds'$$

Tout comme la formule (1) cette expression est de la forme F(S) — F(s). Mais la fonction F(S) ne dépend plus seulement de la valeur numérique de l'aire S; elle dépend aussi de la forme géométrique du domaine S d'intégration. La formule de DE WIJS est du reste un cas particulier de la formule (7). On trouvera en annexe I la démonstration des théorèmes suivants :

1° Si l'échantillon s est géométriquement semblable au gisement S, la formule (7) se réduit à la formule ordinaire de DE WIJS;

2° Si l'échantillon s n'est pas semblable au gisement S, la formule (7) peut toujours se mettre sous la forme :

$$(8) \quad \sigma_s^2 = \alpha L \frac{S}{s} + C$$

C restant une constante si l'on modifie la taille de l'échantillon s sans altérer sa forme. La signification de cette constante est évidente : les figures s et S n'étant pas semblables, l'échantillon s ne peut coïncider avec S lorsque son aire est égale à celle de S, et par suite sa variance n'a aucune raison de s'annuler.

En résumé, la formule ordinaire de DE WIJS reste valable si l'échantillon prélevé est géométriquement semblable au gisement. Si cette condition de similitude n'est pas remplie, la formule de DE WIJS est modifiée par l'addition d'une constante, dont la valeur numérique ne dépend que de la forme géométrique du gisement et de l'échantillon.

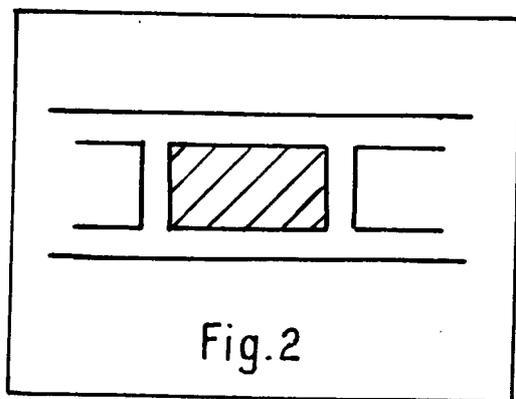
La formule de DE WIJS ainsi généralisée nous permet théoriquement d'étudier n'importe quel mode d'échantillonnage. Il va de soi, qu'en pratique, on se heurtera à certaines difficultés d'intégration. La suite de cette étude sera consacrée à un certain nombre de cas particuliers de grande importance pratique :

- échantillonnage périphérique des panneaux;
- sondages à maille rigide;
- reconnaissance d'un gisement par traçages et montages.

\*

## II. — L'ÉCHANTILLONNAGE PÉRIPHÉRIQUE DES PANNEAUX

UN panneau, de forme rectangulaire ou carrée, est généralement considéré comme reconnu lorsqu'il a été ceinturé par deux traçages et deux montages. Cependant l'extrapolation de la teneur périphérique  $x$  des traçages et des montages à la teneur  $z$  du panneau ainsi défini comporte un véritable saut dans l'inconnu, qu'aucune théorie, jusqu'ici, n'avait réussi à chiffrer. Les formules proposées au chapitre précédent permettent de calculer directement les variances des variables  $x$  et  $z$  ainsi que leur corrélation, et par suite de déterminer la loi de la variable  $z$  lorsque la variable  $x$  est supposée connue. La variance de  $z$  lorsque  $x$  est fixé donne, en particulier, une mesure exacte de la précision de l'échantillonnage périphérique.



Les résultats sont particulièrement simples si l'on assimile le panneau à un cercle et sa périphérie à une couronne circulaire infiniment mince (périphérie évanescence). Nous verrons en annexe VII que l'erreur introduite par la première assimilation est négligeable, pourvu que le panneau soit à peu près carré. Quant à la seconde, elle revient à admettre que le panneau est échantillonné par un grattage infiniment mince effectué de façon continue sur la totalité de la périphérie. Elle n'introduit jamais d'erreur appréciable. L'application de la formule de la covariance élémentaire donne lieu à une géométrie amusante, que l'on trouvera exposée en annexe II. L'analogie évidente avec la théorie des potentiels newtoniens, et en particulier avec l'électrostatique résulte du fait que la fonction  $Lr$  est l'exact correspondant, dans un espace à deux dimensions, du potentiel newtonien en  $\frac{1}{r}$  dans l'espace à trois dimensions.

Nous nous contenterons ici d'énoncer les principaux résultats relatifs à l'estimation d'un panneau circulaire par sa périphérie évanescence :

Lorsque la teneur de la périphérie est connue, la teneur  $z$  du panneau est une variable lognormale dont la valeur moyenne est égale à  $x$  et dont la variance est égale à la moitié de la dispersion absolue  $\alpha$ . Autrement dit :

1° La teneur périphérique  $x$  est un estimateur sans biais de la teneur  $z$  du panneau. La valeur probable de  $z$  coïncide avec la valeur  $x$  trouvée par échantillonnage périphérique. Ce résultat constitue la justification théorique d'un procédé séculaire.

2° La variance d'échantillonnage  $\sigma_e^2$  ou variance de  $z$  à  $x$  fixé, est indépendante de la dimension du panneau échantillonné. Elle est égale à la moitié de la dispersion absolue  $\alpha$ . Tout comme ce paramètre, elle ne dépend donc que du type métallogénique de la minéralisation étudiée.

(9)

$$\sigma_e^2 = \frac{\alpha}{2}$$

La variance d'échantillonnage fixe la précision de l'estimation du panneau. On sait, en effet, qu'il y a quatre-vingt quinze chances sur cent pour que la teneur réelle  $z$  du panneau soit comprise entre :

$$xe^{-2\sigma_e} \text{ et } xe^{+2\sigma_e}$$

Les nombres  $e^{-2\sigma_e}$  et  $e^{+2\sigma_e}$  sont les coefficients de sécurité au seuil de 95 %. Nous donnons ci-après, à titre d'illustration, la valeur de ces coefficients pour quelques valeurs de la dispersion absolue, correspondant à différents types de minéralisation.

L'échantillonnage purement périphérique d'un panneau donne des résultats acceptables dans le cas d'un gisement de substitution, excellents pour les gisements sédimentaires, mais nettement moins bons pour les gisements d'or ou d'uranium (erreur de 40 à 60 %).

En pratique le panneau n'est pas circulaire et la périphérie n'est pas évanescence. La variance d'échantillonnage diffère de la moitié de la dispersion absolue, et la teneur périphérique est un estimateur affecté d'un léger biais : la teneur probable du panneau n'est pas

tout à fait égale à celle de la périphérie. Ces points sont discutés en annexe II : les écarts restent toujours faibles et la théorie de la périphérie évanescence donne toujours une bonne approximation.

TABLEAU I  
Précision d'un échantillonnage périphérique

Type de minéralisation	Dispersion absolue	Coefficients de sécurité
Gisements d'or .....	$\frac{14}{100}$	0.61 - 1.70
	$\frac{10}{100}$	0.64 - 1.56
Uranium filonien.....	$\frac{9}{100}$	0.65 - 1.53
	$\frac{7}{100}$	0.69 - 1.45
Plomb filonien .....	$\frac{5}{100}$	0.73 - 1.37
	$\frac{3}{100}$	0.78 - 1.28
Plomb de substitution..	$\frac{1}{100}$	0.87 - 1.15
	$\frac{1}{1000}$	0.94 - 1.06
Fe ou Mn sédimentaire.	$\frac{1}{10000}$	0.98 - 1.02

La précision de l'échantillonnage périphérique est très médiocre dans le cas des gisements d'or et d'uranium. La façon la plus simple de l'améliorer consiste à tracer un sous-étage ou un montage intermédiaire. Mais on peut se demander s'il n'est pas possible, sans effectuer de travaux miniers importants, d'obtenir, par sondages intérieurs par exemple, des informations supplémentaires capables de réduire de façon appréciable la marge d'erreur. Il n'y a, en fait, que trois possibilités, suivant que l'on cherche ces données supplémentaires à l'extérieur du panneau, dans la périphérie elle-même, ou enfin dans le panneau lui-même.

Le bon sens élimine immédiatement la première possibilité. Il résulte du reste du théorème de l'écran péri-

phérique, que l'on trouvera démontré en annexe II, qu'une fois connue la teneur périphérique, la teneur du panneau est indépendante de toutes les données extérieures. La deuxième possibilité, en fait, n'existe pas non plus. La teneur  $x$  de la périphérie, teneur moyenne réelle des deux traçages et des deux montages, est la moyenne arithmétique pondérée des teneurs des différentes volées constitutives. On peut se demander s'il n'est pas possible, à partir de ces  $n$  teneurs, de former un estimateur, lognormal par exemple, plus efficace que  $x$ . On peut aussi s'imaginer augmenter indéfiniment la précision en diminuant la taille des volées analysées séparément. En fait il n'en est rien, et le théorème dit du « statisticien déçu » (annexe III) montre qu'il n'est pas possible, à partir des volées, de constituer un estimateur plus efficace que  $x$ . Si l'échantillonnage périphérique est du type continu (prélèvement sur tout venant volées par volées, par exemple, en opposition avec les types discontinus, tels que les prélèvements par rainurages équidistants) il est absolument inutile de calculer la teneur périphérique  $x$  par les procédés lognormaux : en effet, si l'on néglige les erreurs d'analyse et de quartage, toujours faibles, la moyenne arithmétique pondérée, calculée par les procédés usuels, représente la véritable teneur périphérique  $x$  et, une fois  $x$  connu, la connaissance des teneurs individuelles des volées n'apporte aucune information supplémentaire. Si, par contre, l'échantillonnage de la périphérie est discontinu il y aura tout intérêt à estimer la valeur réelle de la teneur  $x$  par les procédés lognormaux : dans ce dernier cas deux erreurs s'ajoutent : erreur sur l'estimation de  $x$  à partir des échantillons discontinus et erreur sur l'estimation du panneau à partir de la valeur réelle de  $x$ .

Reste la troisième possibilité. Il s'agit cette fois de prélever des échantillons supplémentaires à l'intérieur du panneau (par exemple, sondages obliques exécutés à partir de l'extrémité de petites recoupes). Cette possibilité existe, mais son efficacité est assez faible si l'on ne veut pas multiplier exagérément le nombre de ces sondages. Dans le cas où l'on n'exécute qu'un seul sondage, la meilleure implantation possible est au centre du panneau. Si  $u$  est la teneur de ce sondage, la valeur probable de la teneur du panneau est une moyenne géométrique pondérée de  $u$  et de  $x$  :

$$(10) \quad z' = Cx^{1-a} u^a$$

On trouvera en annexe IV la valeur des coefficients  $C$  et  $a$ .  $a$  est d'ailleurs très faible, de sorte que les travaux miniers ont une influence tout à fait prépondérante... Le gain de précision représenté par le sondage est également très faible, comme on pourra s'en rendre compte sur l'exemple numérique traité en annexe.

## III. — SONDAGES À MAILLE RIGIDE

La pratique très ancienne des mailles rigides de sondages, comme du reste tout procédé systématique d'échantillonnage, pose, pour le statisticien qui l'étudie, des problèmes assez complexes. Lorsqu'un gisement a été reconnu par  $n$  sondages, ayant fourni des teneurs  $x_1, x_2, \dots, x_n$ , il est aisé, par l'emploi d'une simple moyenne arithmétique pondérée ou d'un estimateur lognormal plus élaboré, d'évaluer la teneur moyenne,  $m$ , du gisement. Les choses se compliquent dès que l'on cherche à déterminer la précision de cette évaluation, c'est-à-dire la variance d'échantillonnage  $\sigma_e^2$ . La première idée qui vient à l'esprit du statisticien, habitué à la théorie des erreurs est d'écrire que la variance d'échantillonnage est  $n$  fois plus petite que la variance  $\sigma^2$  de la variable  $x_i$ ,

$$(11) \quad \sigma_e^2 = \frac{\sigma^2}{n}$$

Une telle équation serait parfaitement exacte si les  $n$  teneurs  $x_i$  pouvaient être considérées comme des données indépendantes les unes des autres. Il existe un schéma très simple réalisant cette indépendance, c'est celui où les emplacements des  $n$  sondages sont tirés au sort. Une telle implantation aléatoire peut amener deux ou plusieurs sondages très proches les uns des autres, voire même chevauchant. Toute possibilité de chevauchement ou de grande proximité étant exclue dans une maille rigide, hexagonale ou carrée, et la formule de la covariance nous ayant montré que la corrélation est une fonction décroissante de la distance, nous sommes certains *a priori*, que les  $x_i$  ne sont pas des variables indépendantes, mais sont, au contraire, liées par des corrélations négatives. La variance réelle d'échantillonnage est donc plus faible que l'expression (II). Pour reprendre une notion introduite par GROSJEAN (1), on peut dire que le nombre utile des sondages (nombre par lequel il faut diviser la variance  $\sigma^2$  pour obtenir la variance d'échantillonnage) est, dans une maille rigide, supérieur à leur nombre réel. La meilleure implantation possible de nos  $n$  sondages est celle qui maxime ce nombre utile, c'est-à-dire celle qui réalise la plus forte corrélation négative possible entre les variables  $x_i$ . La formule de la covariance nous montre que cette implantation doit rendre maximum l'expression :

$$\sum_{i \neq j} L r_{ij}$$

où  $r_{ij}$  représente la distance des sondages d'ordre  $i$  et  $j$ .

(1) P.-V. GROSJEAN, « L'évaluation mathématique des gisements détritiques », Bruxelles, 1953.

Il est parfaitement intuitif que ce maximum ne peut être réalisé que par une maille régulière, hexagonale ou carrée, c'est-à-dire une maille dans laquelle le « polygone d'influence » de chaque sondage est aussi voisin que possible d'une figure circulaire, avec un très léger avantage pour le réseau hexagonal, dans la mesure même où l'hexagone se rapproche davantage du cercle que le carré.

Le lecteur trouvera en annexe V le détail des calculs. En désignant par  $\sigma^2$  la variance dans le gisement S de la teneur  $x$  d'un sondage, par  $\sigma_i^2$  la variance de la teneur  $z$  d'un polygone d'influence, et par  $\alpha$  la dispersion absolue, on trouve :

$$(12) \quad \sigma_e^2 = \frac{\sigma^2 - \sigma_i^2 - \alpha/2}{n}$$

Cette formule est calculée en assimilant le polygone d'influence à un cercle de même aire. C'est pourquoi elle ne distingue pas les mailles hexagonales et carrées.

Dans les formules rigoureuses, le terme  $\frac{\alpha}{2}$  serait remplacé par une expression de quelques pour cent plus faibles, avec un léger avantage pour la maille hexagonale. Ces corrections sont négligeables, puisque  $\frac{\alpha}{2}$  est très petit lui-même et la formule (12) représente très bien la réalité.

## Comparaison avec l'échantillonnage aléatoire stratifié

Dans le schéma dit « stratifié au hasard » le gisement est divisé en  $n$  polygones d'influence, comme dans une maille systématique, mais chaque sondage est implanté au hasard dans son polygone, au lieu d'être placé au centre. Cette petite part laissée au hasard se manifeste, dans la formule (12), par la dispersion du terme  $\frac{\alpha}{2}$  qui représentait justement l'élément préférentiel de la corrélation du sondage placé au centre avec son polygone. La modification est faible, puisque, aussi bien le schéma stratifié au hasard diffère très peu d'une maille rigide, mais se traduit par une perte de précision : le prix de revient des deux méthodes étant exactement le même, cette légère différence doit inciter à adopter dans tous les cas la maille systématique.

Pour résumer cette discussion, le tableau suivant présente, par ordre de précision décroissante, les différents modes d'implantation et leur variance d'échantillonnage :

Maille rigide hexagonale ou carrée	$\sigma_k^2 = \frac{\sigma^2 - \sigma_1^2 - \alpha/2}{n}$
Implantation aléatoire stratifiée	$\sigma_k^2 = \frac{\sigma^2 - \sigma_1^2}{n}$
Implantation aléatoire pure	$\sigma_k^2 = \frac{\sigma^2}{n}$

**Le facteur de Run**

Les théories précédentes ne sont évidemment valables que dans la mesure où la formule (5) est vérifiée. Or, cette formule de la covariance élémentaire implique une hypothèse, l'isotropie de l'espace du gisement. Le terme  $Lr$  est, en effet, indépendant de la direction du vecteur joignant les aires élémentaires. Cependant, l'expérience montre, en prospection alluvionnaire par exemple, qu'il existe souvent au contraire une anisotropie très poussée, qui se manifeste par l'existence de *runs*.

L'hypothèse la plus simple que nous puissions faire, pour essayer de généraliser les résultats précédents, est celle d'une anisotropie elliptique.

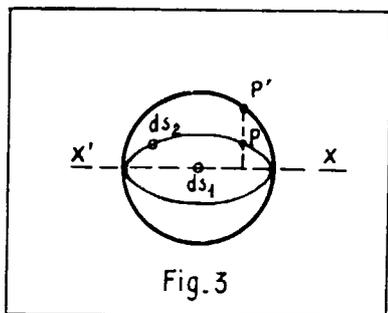


Fig. 3

Étant donné un élément  $dS_1$ , admettons que le lieu des éléments  $dS_2$  tels que la covariance  $\sigma_{u,u}$  soit constante est, non plus un cercle, mais une ellipse centrée sur  $dS_1$  et dont le grand axe est parallèle à la direction  $x'x$  du run. Si  $2a$  et  $2b$  sont les longueurs du grand axe et du petit axe de cette ellipse, il existe une transformation géométrique très simple, appelée affinité d'axe  $x'x$  et de module  $\frac{a}{b}$  capable de transformer cette ellipse en un cercle de rayon  $a$ . Dans cette affinité, l'axe  $x'x$  étant l'axe des abscisses, l'homologue  $P'$  d'un point  $P$  s'obtient en multipliant par  $\frac{a}{b}$  l'ordonnée de ce point, son abscisse restant invariante. On peut alors admettre que la formule de la covariance reste valable, à condition de prendre pour  $r$ , non pas la distance des deux éléments, mais celle de leurs homologues dans l'espace isotrope que cette affinité nous restitue. Les théories précédentes se généralisent ainsi, d'elle-mêmes.

J. P 739208.

Le module de cette affinité  $\frac{a}{b}$  porte le nom de facteur de run. Son importance pratique est évidente. En effet, le meilleur schéma possible d'échantillonnage est celui qui a pour homologue dans l'espace affine, une maille rigide hexagonale ou carrée, c'est-à-dire le schéma, classique en prospection alluvionnaire, de lignes de puits, le rapport de la distance des lignes à la distance des puits sur une même ligne étant égal justement au facteur de run  $\frac{a}{b}$  (1). Ce rapport est fixé empiriquement, en prospection alluvionnaire, à une valeur de l'ordre de 5 ou 10. Cependant, dans son ouvrage déjà cité, GROSJEAN signale que le nombre « utile » des puits est parfois inférieur à leur nombre réel, et que cet état de chose doit s'améliorer, suivant les cas, soit en resserrant l'intervalle des puits sur une seule ligne, soit, au contraire, en creusant de nouvelles lignes intermédiaires : il est clair que dans les cas signalés par GROSJEAN le rapport de maille adopté diffère notablement du facteur de run, par défaut ou par excès respectivement. La connaissance *a priori* de ce facteur, qui doit être à peu près constant pour des gisements de même type, permettrait de fixer à coup sûr la meilleure maille de prospection.

**L'estimation des panneaux**

KRICE (2) a magistralement montré que l'assimilation simpliste de la teneur d'un panneau à la teneur de l'échantillon qui le représente entraînait une erreur systématique : surestimation des panneaux riches et sous-estimation des panneaux pauvres. Des formules simples (3) permettent de corriger cette erreur. Dans le cas des sondages à maille rigide, le panneau est un polygone d'influence. Du fait de la position privilégiée du sondage central, les formules sont légèrement modifiées, tout en restant du type :

$$(13) \quad z' = Cm^{1-a} x^a$$

Dans cette formule,  $z'$  est la valeur probable de la teneur  $z$  du panneau,  $x$  la teneur du sondage et  $m$  la teneur moyenne du gisement supposée connue (4). Les valeurs des coefficients  $C$  et  $a$  sont données en annexe VI. La précision de l'estimation d'un panneau est toujours très médiocre. Il est possible, cependant, de l'améliorer quelque peu en tenant compte, non seulement du son-

(1) Les puits étant, ou non, disposés en quinconce, suivant que la maille est hexagonale ou carrée dans l'espace affine.

(2) D. G. KRICE, A statistical approach to some Mine valuation and allied problems on the Witwatersrand. Thèse présentée à l'Université du Witwatersrand, Johannesburg, 1951.

(3) Cf. *Annales des Mines*, décembre 1955.

(4) Cette teneur  $m$  est connue avec une précision bien plus grande que celle que l'on peut espérer obtenir sur la teneur d'un panneau individuel, de sorte que cette hypothèse peut être considérée comme pratiquement valable.

dage central, mais également des sondages immédiatement voisins, qui sont au nombre de 6 dans une maille hexagonale. La formule (13) est alors remplacée par une formulé du type :

$$(14) \quad z' = C'm^{1-a-b} \cdot x^a (x_1 x_2 \dots x_6)^{\frac{b}{6}}$$

#### IV. — RECONNAISSANCE D'UN GISEMENT PAR TRAVAUX MINIERS

LES «échantillons» que fournissent les travaux miniers ont des caractéristiques géométriques bien particulières. L'ensemble d'un traçage, par exemple, forme un rectangle dont un des côtés, celui de l'allongement, peut être cent fois plus grand que l'autre. Il ne saurait être question d'assimiler pareille figure à un cercle, ni par conséquent d'appliquer la formule de DE WIJS ordinaire. Cette formule, en effet, conduirait à attribuer la même variance à un panneau carré de 30 mètres sur 30 et à un traçage de 450 mètres de long sur 2 mètres de haut. Il est évident pour le mineur que le traçage donne une bien meilleure image du gisement dans son ensemble que le petit panneau carré. Cette conséquence paradoxale est évitée dans la présente théorie. L'intégration de la formule de la covariance dans un contour rectangulaire (1) fixe à la variance d'un traçage de longueur  $l$  une valeur égale à celle du panneau carré de côté  $\frac{l}{2}$  (soit 225 m au lieu de 30 dans l'exemple précédent). Le traçage est du reste assimilable à un échantillon à une seule dimension, c'est-à-dire à une ligne et non à une surface. Sa valeur, en tant qu'échantillon est uniquement fonction de sa longueur.

D'une façon générale, et en première approximation, la variance d'échantillonnage est proportionnelle à la surface échantillonnée, et inversement proportionnelle au carré de longueur tracée totale.

$$(15) \quad \sigma_k^2 = \alpha k \frac{S}{L^2}$$

$\alpha$  est la dispersion absolue, et  $k$  un facteur de forme qui dépend du problème posé.  $S$  est la surface considérée comme échantillonnée, et  $L$  la longueur totale des travaux miniers effectués dans  $S$ .

Exemples :

1° Reconnaissance d'un gisement entre deux cotes données à partir d'un seul traçage :

$$(16) \quad \sigma_k^2 = \alpha \frac{\pi S}{3 L^2}$$

Il est pratiquement inutile de tenir compte des sondages suivants, comme il résulte du reste du théorème de l'écran périphérique : les six sondages les plus proches, par suite d'un effet de cage de FARADAY, arrêtent la plus grande part de l'influence du reste du gisement. On trouvera en annexe un exemple numérique.

2° Reconnaissance d'un gisement entre deux cotes à partir d'un traçage et de plusieurs montages. Dans le cas où la longueur totale des montages est égale à celle du traçage, le coefficient  $k$  a une valeur à peu près double :

$$(17) \quad \sigma_k^2 = 2.20 \alpha \frac{S}{L^2}$$

mais  $L$  étant deux fois plus grand, la variance est deux fois plus petite : cette propriété résulte simplement du fait que traçage et montage peuvent être considérés comme données indépendantes ;

3° Reconnaissance d'un gisement à partir de plusieurs traçages. On a ici :

$$(18) \quad \sigma_k^2 = \alpha \frac{\pi S}{2 L^2}$$

4° Gisement reconnu à partir de plusieurs traçages et montages formant une grille à peu près carrée : le coefficient  $k$  est ici encore, doublé par rapport au cas précédent :

$$(19) \quad \sigma_k^2 = \alpha \pi \frac{S}{L^2}$$

Ces diverses formules sont toutes du type (15) avec des valeurs différentes du coefficient  $k$ , ce coefficient, sans dimension, varie en gros du simple au double suivant qu'il n'y a que des traçages, ou simultanément des traçages et des montages en réseau carré. On voit ainsi qu'à longueur égale, il est préférable de n'avoir que des traçages.

En deuxième lieu, et pour un même dispositif géométrique, la variance d'échantillonnage est proportionnelle à la dispersion absolue : un gisement d'uranium, de dispersion égale à  $\frac{9}{100}$ , reconnu par les mêmes travaux miniers qu'un gisement de substitution de même aire et de dispersion égale à  $\frac{1}{100}$ , est évalué avec une précision trois fois moins bonne que ce dernier.

(1) Voir Annexe VII.

En troisième lieu, la formule (15) montre que la variance est proportionnelle au rapport sans dimension  $\frac{S}{L^2}$ . La précision de l'évaluation est ainsi en raison inverse de la racine de la surface reconnue, et proportionnelle à la longueur tracée totale. C'est ainsi, par exemple, que le procédé qui consiste à tracer systématiquement des sous-étages multiplie la précision par deux, et non pas par  $\sqrt{2}$  comme l'aurait suggéré la statistique ordinaire. De plus, ce facteur homogène conserve la même valeur si l'on multiplie L par  $\lambda$  et S par  $\lambda^2$ . Il en résulte qu'à dispersions absolues égales, un gros gisement nécessite beaucoup moins de travaux qu'un petit : un gisement cent fois plus grand qu'un autre, par exemple, n'exige que dix fois plus de travaux

On déduit immédiatement des formules (15) et (20) que la longueur tracée assurant la précision de  $p$  % est égale à :

$$(21) \quad L = \frac{200}{p} \sqrt{\alpha k S}$$

Dans le cas où il n'y a que des traçages ( $k = \frac{\pi}{2}$ ) et où l'on désire une précision de 5 %, on obtient par exemple :

$$(22) \quad L = 50 \sqrt{\alpha S}$$

Le tableau suivant donne la valeur de L en mètres, en fonction de l'aire S et de la dispersion  $\alpha$ .

	$\frac{1}{10\ 000}$	$\frac{1}{1\ 000}$	$\frac{1}{100}$	$\frac{3}{100}$	$\frac{5}{100}$	$\frac{7}{100}$	$\frac{9}{100}$	$\frac{11}{100}$
5 000 m <sup>2</sup> . . . . .	35	112	350	610	790	935	1 060	1 170
10 000 m <sup>2</sup> . . . . .	50	159	500	865	1 120	1 320	1 500	1 660
50 000 m <sup>2</sup> . . . . .	115	350	1 150	1 940	2 470	2 960	3 350	3 700
100 000 m <sup>2</sup> . . . . .	159	500	1 590	2 740	3 540	4 170	4 750	5 250
500 000 m <sup>2</sup> . . . . .	350	1 120	3 500	6 100	7 900	9 350	10 600	11 700
1 000 000 m <sup>2</sup> . . . . .	500	1 390	5 000	8 650	11 200	13 200	15 000	16 600

de reconnaissance : le coût de ces travaux, rapporté à la tonne de minerai, est donc dix fois plus faible. Le gros gisement, en somme, gagne sur tous les tableaux.

D'un point de vue pratique, la question qui se pose est surtout celle-ci : étant donné un gisement d'aire et de dispersion absolue connues, quelle est la longueur totale qu'il est nécessaire de tracer pour avoir une précision donnée? On sait qu'il n'y a que cinq chances sur cent d'avoir une erreur relative supérieure à  $p$  % si l'on a l'égalité :

$$(20) \quad 2\sigma_r = \frac{p}{100}$$

Un gisement de 100 000 mètres carrés nécessite 500 mètres de traçages, si la dispersion est  $\frac{1}{1\ 000}$  (manganèse, par exemple) et 5 000 si la dispersion est  $\frac{10}{100}$  (uranium) : si le gisement est un rectangle de 500 mètres sur 200 mètres, un niveau suffit dans le cas du manganèse. Dans le cas de l'uranium, il en faudrait 10. En fait, pour les gisements très dispersés, on se contente souvent d'une précision plus faible. Si l'on admet une marge d'erreur de 10 %, au lieu de 5 %, les longueurs données par le tableau doivent être divisées par deux, et le gisement d'uranium sera reconnu à partir de cinq niveaux, soit un niveau tous les 40 mètres.

\*

CONCLUSION

La théorie développée ci-dessus est loin d'être définitive. Son principal mérite est de situer tous les problèmes et de conduire à des formules simples, directement utilisables par les praticiens. Il est certain que la réalité est plus complexe. L'hypothèse de base, suivant laquelle la variance est mathématiquement déterminée par la donnée de la taille et de la forme de l'échantillon et du panneau échantillonné — c'est-à-dire le principe de similitude géométrique — ne constitue sans doute qu'une première approximation. Il est probable que non seulement la valeur expérimentale — sujette aux fluc-

tuations statistiques normales, inséparables de tout processus d'estimation — mais aussi la valeur réelle elle-même de la variance peut différer quelque peu de la valeur attendue. La théorie précédente peut donc être notablement améliorée, mais il ne faut pas perdre de vue que ces perfectionnements conduiront à une complexité mathématique croissante.

En l'absence de critères expérimentaux décisifs (1), les formules proposées ici ont l'avantage de la simplicité et peuvent être acceptées par les utilisateurs à titre de première approximation.

ANNEXE I

Généralisation de la formule de De Wijs

Considérons la fonction :

$$(1,1) \quad G(S) = \frac{1}{S^2} \iint_S dS \iint_S LrdS$$

et deux aires S et S<sub>1</sub> géométriquement semblables, le rapport de similitude étant λ

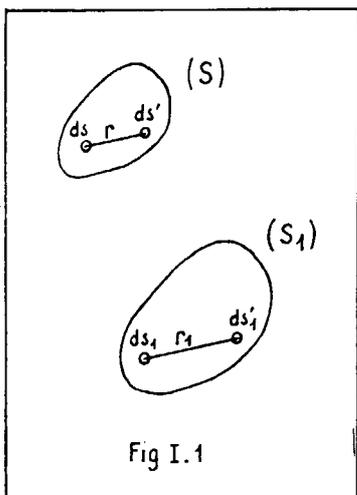


Fig I. 1

A tout couple dS dS' distant de r dans S, correspond dans S<sub>1</sub>, par similitude le couple dS<sub>1</sub> dS<sub>1</sub>' distant de r<sub>1</sub>, tel que l'on ait :

$$dS_1 = \lambda^2 dS, \quad dS_1' = \lambda^2 dS', \quad r_1 = \lambda r$$

(1) Le lecteur trouvera en Annexe VII une comparaison des formules du rectangle avec les données de l'expérience. L'accord n'est pas excellent, mais il est acceptable.

On a donc :

$$G(S_1) = \frac{1}{S_1^2} \iint_S \lambda^2 dS \iint_S \lambda^2 L(\lambda r) dS'$$

Comme S<sub>1</sub> = λ<sup>2</sup>S, on a finalement :

$$(1,2) \quad G(S_1) = G(S) + L\lambda$$

Ceci étant, envisageons les cas particuliers suivants :

1° L'échantillon est géométriquement semblable au gisement S dont il est prélevé. La formule (1, 2) montre que la variance de s dans S est :

$$\sigma_s^2 = -\alpha L \lambda^2$$

ce que l'on peut écrire aussi bien sous la forme de de Wijs :

$$(1,3) \quad \sigma_s^2 = \alpha L \frac{S}{s}$$

2° L'échantillon s n'est plus géométriquement semblable au gisement S. On cherche comment se modifie sa variance lorsqu'on fait varier la taille de s sans altérer sa forme. Soient donc deux échantillons :

s<sub>1</sub> et s<sub>2</sub>, semblables prélevés dans un gisement S de forme quelconque. On a :

$$\sigma_{s_1}^2 = 2\alpha G(S) - 2\alpha G(s_1)$$

$$\sigma_{s_2}^2 = 2\alpha G(S) - 2\alpha G(s_2)$$

D'après la formule (1, 2) on peut écrire :

$$\sigma_{s_2}^2 = 2\alpha G(S) - 2\alpha G(s_1) - 2\alpha L \lambda$$

c'est-à-dire :

$$\sigma_{s_2}^2 = \sigma_{s_1}^2 - \alpha L \frac{s_2}{s_1}$$

On peut donc écrire :

$$(1,4) \quad \sigma_s^2 = \alpha L \frac{S}{s} + C$$

C étant une constante pour un gisement S pourvu que s reste géométriquement semblable à lui-même.

ANNEXE II

L'échantillonnage périphérique des panneaux circulaires

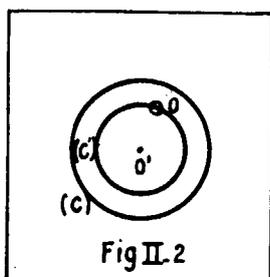
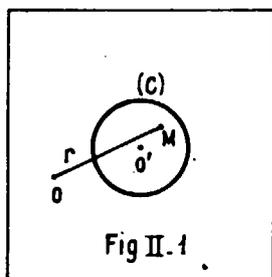
Le choix de contours circulaires est motivé par le théorème classique suivant : « Si une fonction est harmonique dans un cercle et sur son contour, la valeur moyenne de cette fonction sur le contour est égale à sa valeur au centre ». Il en résulte immédiatement que sa valeur moyenne dans le cercle est, elle aussi, égale à la valeur au centre.

**Cas extérieur.** — En particulier, soit un cercle (C) et un point O extérieur à (C). Si l'on pose  $r = OM$ , la fonction  $Lr$  est harmonique dans (C) et sur son contour. La valeur moyenne de  $Lr$  dans le cercle est donc égale à  $Lr_0$ ,  $r_0$  désignant la distance  $OO'$  du point O et du centre du cercle.

Il en résulte immédiatement que la covariance de deux cercles extérieurs l'un à l'autre est égale à la covariance élémentaire  $\sigma_{u_1 u_2}$  des aires infiniment petites  $dS_1$  et  $dS_2$  placées au centre de chacun des cercles :

$$(II,1) \quad \sigma_{r_1 r_2} = \sigma_{u_1 u_2} = A - 2\alpha Lr_0$$

La covariance de deux cercles ne dépend que de la distance de leurs centres, et reste indépendante de leurs rayons tant que ces cercles n'ont pas de point commun.



**Cas intérieur.** — Si au contraire le point O est intérieur à (C), la fonction  $Lr$  n'est plus définie lorsque M vient en O et le théorème ne s'applique plus. On considère alors le cercle (C') de même centre O' et de rayon  $O'O = r_0$ . La valeur moyenne de  $Lr$  dans le cercle (C') est  $Lr_0$ . Dans la couronne circulaire comprise entre C et C',  $Lr$  a la même valeur moyenne que si le point O était en O'. En effet, cette valeur moyenne est la même quelle que soit la position de O sur le contour de (C') et peut être considérée comme une fonction harmonique du point O — elle est donc égale à sa valeur au centre, c'est-à-dire,  $R$  désignant le rayon de (C), à :

$$\frac{1}{\pi(R^2 - r_0^2)} \int_{r_0}^R 2\pi r Lr dr = \frac{R^2 Lr - r_0^2 Lr_0}{R^2 - r_0^2} = \frac{1}{2}$$

La valeur moyenne de  $Lr$  dans (C) entier s'en déduit facilement :

$$(II,2) \quad E(Lr) = LR - \frac{R^2 - r_0^2}{2R^2}$$

Il en résulte que la covariance  $\sigma_{uz}$  de l'aire élémentaire  $dS$  et du cercle (C) de rayon  $R$  est égale, en fonction de la distance  $r$  du centre de (C) et de  $dS$ , à :

$$(II,3) \quad \begin{cases} \sigma_{uz} = A - 2\alpha Lr & \text{pour } r > R \\ \sigma_{uz} = A - 2\alpha LR + \alpha \frac{R^2 - r^2}{R^2} & \text{pour } r < R \end{cases}$$

La teneur  $z$  du cercle (C), de son côté, a une variance égale à :

$$\sigma_z^2 = \int_0^R \frac{2\pi r \sigma_{uz}^2}{\pi R^2} dr$$

La valeur de  $\sigma_{uz}$  étant celle du cas intérieur. Il vient :

$$(II,4) \quad \sigma_z^2 = A - 2\alpha LR + \frac{\alpha}{2}$$

Compte tenu de cette relation, les formules (II, 3) peuvent s'écrire :

$$(II,5) \quad \begin{cases} \sigma_{uz} = \sigma_z^2 - 2\alpha L \frac{r}{R} + \frac{\alpha}{2} & \text{pour } r > R \\ \sigma_{uz} = \sigma_z^2 - \alpha \frac{r^2}{R^2} + \frac{\alpha}{2} & \text{pour } r < R \end{cases}$$

**Cercle critique.** — Le lieu des centres des cercles dont les teneurs sont indépendantes de la teneur  $z$  de (C) est un cercle concentrique à (C), appelé cercle critique, de rayon  $R_c$  tel que l'on ait :

$$2\alpha L \frac{R_c}{R} = \sigma_z^2 - \frac{\alpha}{2}$$

En particulier, si le gisement  $S$  est lui-même un cercle de rayon  $R_0$ , on a :

$$\sigma_z^2 = \alpha L \frac{R_0^2}{R^2}$$

et le rayon du cercle critique est :

$$R_c = R_0 e^{-1/4} = 0.78 R_0$$

Les éléments intérieurs au cercle critique sont en corrélation positive avec  $z$ , et les éléments extérieurs en corrélation négative.

**Les fonctions  $k, \eta, \omega$ , et  $h$ ,**

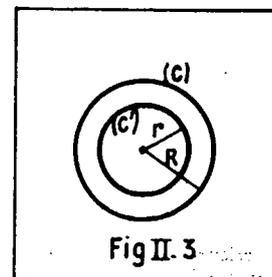
Soit un cercle C de rayon  $R$ , et un cercle C' concentrique, de rayon  $r$  plus petit. Dans ce qui suit,  $z$  désigne la teneur de C,  $y$  celle de C', et  $x$  la teneur de la couronne comprise entre C et C', c'est-à-dire la teneur périphérique.

Posons :

$$(II,6) \quad \begin{cases} \sigma_z^2 = \sigma^2 - k \\ \sigma_{xy} = \sigma^2 - h \\ \sigma_{xz} = \sigma^2 - \eta \\ \sigma_{yz} = \sigma^2 + \omega \end{cases}$$

Prenons comme paramètre le rapport de l'aire de la couronne périphérique à l'aire du cercle (C) :

$$(II,7) \quad \lambda = \frac{R^2 - r^2}{R^2}$$



On trouve, par des intégrations élémentaires analogues à celles du paragraphe précédent :

$$(II,8) \quad \left\{ \begin{aligned} k &= \alpha \frac{1-\lambda}{\lambda^2} \left[ (1-\lambda)L(1-\lambda) + \lambda \right] \\ h &= -\frac{\alpha}{2} - \alpha \frac{(1-\lambda)}{\lambda} L(1-\lambda) \\ \eta &= \frac{\alpha}{2} (1-\lambda) \\ \omega &= \frac{\alpha}{2} \lambda \end{aligned} \right.$$

Ces formules résolvent le problème de l'échantillonnage périphérique. On remarque que  $k$  est toujours positif. La variance de la périphérie est toujours inférieure à la variance du grand cercle, bien que son aire soit plus petite. La raison en est que les éléments constitutifs de la couronne sont plus éloignés, en moyenne, les uns des autres que ceux du grand cercle.

**Théorie de la périphérie évanescence.**

La périphérie évanescence correspond au cas où  $\lambda$  est infiniment petit : le panneau est échantillonné par un grattage infiniment mince effectué de façon continue sur la totalité de la périphérie. On trouve alors, en faisant  $\lambda = 0$  dans les formules (II, 8) :

$$(II,9) \quad \boxed{\begin{aligned} k &= h = \eta = \frac{\alpha}{2} \\ \omega &= 0 \end{aligned}}$$

La dernière relation résulte du fait que  $y$  et  $z$  sont égaux. Des relations (II, 9) découlent les conséquences suivantes :

1° *Théorème de l'écran périphérique.* — Lorsque la teneur  $x$  de la périphérie évanescence est fixée, la teneur  $z$  du panneau est stochastiquement indépendante de la teneur de tout échantillon prélevé à l'extérieur du panneau. Autrement dit, le panneau circulaire n'est en corrélation avec le reste du gisement que par l'intermédiaire de sa périphérie évanescence.

Il suffit de démontrer qu'il en est bien ainsi pour toute aire élémentaire  $dS$ , de teneur  $u$ , extérieure au panneau.

Pour un tel élément, extérieur à (C), le théorème sur les fonctions harmoniques montre que l'on a :

$$(II,10) \quad \sigma_{ux} = \sigma_{uz}$$

Considérons alors l'expression :

$$H = \sigma_x^2 \sigma_{uz} - \sigma_{ux} \sigma_{xz}$$

D'après (II, 10), on a :

$$H = \sigma_{uz} [\sigma_x^2 - \sigma_{xz}]$$

L'égalité de  $k$  et de  $\eta$  nous montre que  $\sigma_x^2$  et  $\sigma_{xz}$  sont égaux, donc  $H$  est nul. En divisant  $H$  par  $\sigma_x^2 \sigma_{uz}$  et en introduisant les coefficients de corrélation, il vient :

$$(II,11) \quad \rho_{uz} = \rho_{ux} \rho_{xz}$$

Cette relation démontre le théorème.

Du point de vue pratique, ce théorème signifie qu'une fois connue la teneur périphérique, le prélèvement d'échantillons nouveaux extérieurs au panneau n'apporte aucune information supplémentaire quant à la teneur du panneau.

2° *Estimation du panneau par la périphérie évanescence.* — Supposons connue la teneur  $x$  de la périphérie. Quelle est l'efficacité de  $x$  comme estimateur de la teneur du panneau ? Il suffit, pour répondre à cette question, de connaître la loi de probabilité de  $z$  à  $x$  fixé (c'est-à-dire la loi de probabilité de la teneur  $z$  d'un panneau (C) tiré au sort parmi l'infinité de panneaux ayant une teneur périphérique égale à  $x$ ). Désignons par  $\gamma_x$  et  $\gamma_z$  les médianes dans le gisement  $S$  de  $x$  et de  $z$ : A  $x$  fixé,  $z$  est une variable lognormale de médiane.

$$\gamma_{z|x} = \gamma_z \left[ \frac{x}{\gamma_x} \right]^{\rho_{xz} \frac{\sigma_z}{\sigma_x}}$$

et de variance :

$$\sigma_{z|x}^2 = \sigma_z^2 (1 - \rho_{xz}^2)$$

Compte tenu des relations (II, 9), ceci peut s'écrire :

$$(II,12) \quad \left\{ \begin{aligned} \gamma_{z|x} &= \gamma_z \frac{x}{\gamma_x} \\ \sigma_{z|x}^2 &= \frac{\alpha}{2} \end{aligned} \right.$$

Il en résulte que la valeur moyenne de  $z$  à  $x$  fixé est :

$$z' = \gamma_z \frac{x}{\gamma_x} e^{\alpha/4}$$

Soit  $m$  la teneur moyenne du gisement (S), on a :

$$m = \gamma_z e^{\sigma_z^2/2} = \gamma_x e^{\sigma_x^2/2}$$

d'où l'on déduit :

$$\frac{\gamma_z}{\gamma_x} = e^{-\alpha/4}$$

et par suite :

$$(II,13) \quad \boxed{z' = x}$$

On voit ainsi que la teneur  $x$  de la périphérie évanescence est un estimateur sans biais de la teneur  $z$  du panneau circulaire. La précision de cette estimation est mesurée par la variance de  $z$  à  $x$  fixée, variance égale à la moitié de la dispersion absolue.

Il est remarquable que cette précision soit indépendante non seulement de la taille et de la forme du gisement  $S$ , mais également des dimensions du panneau circulaire.

Dans la pratique, le panneau n'est pas circulaire et la périphérie n'est pas évanescence. Toutefois, pourvu que le panneau soit à peu près carré, les résultats précédents constituent une première approximation. En particulier, il est pratiquement inutile de tenir compte pour l'évaluation d'un panneau, des données extérieures à sa périphérie. Les données intérieures, par contre, apportent un supplément appréciable d'information. De même, la teneur  $x$  de la périphérie, en tant qu'estimateur de la teneur  $z$  du panneau n'est que très faiblement biaisée, et sa précision est donnée par une variance d'échantillonnage dont l'ordre de grandeur est toujours celui de la moitié de la dispersion absolue.

**Périphérie non évanescence.**

Le paramètre  $\lambda$  n'est plus égal à 0. En pratique, sa valeur est de l'ordre de 1/5 à 1/20, suivant la taille des panneaux. Déterminons la loi de distribution de la teneur  $y$  du panneau résiduel lorsque la teneur  $x$  de la périphérie est connue. A  $x$  fixé,  $y$  est une variable lognormale de paramètre.

$$(II, 14) \quad \begin{cases} \gamma_{y,x} = \gamma_y \left[ \frac{x}{\gamma_x} \right]^{\rho_{xy} \frac{\sigma_y}{\sigma_x}} \\ \sigma_{y,x}^2 = \sigma_y^2 (1 - \rho_{xy}^2) \end{cases}$$

La formule (I, 4) nous permet de calculer  $\sigma_y^2$  en fonction de  $\sigma_x^2$  et de  $\lambda$  :

$$(II, 15) \quad \sigma_y^2 = \sigma_x^2 + p = \sigma_x^2 - \alpha \lambda (1 - \lambda)$$

A l'aide de cette équation et des formules (II, 8), les formules (II, 14) permettent de calculer la valeur probable  $y'$  et la variance de  $y$  à  $x$  fixé. Tous calculs faits, on trouve :

$$(II, 16) \quad \begin{cases} y' = Kx \\ \sigma_{y,x}^2 = p + 2h - k - \frac{(k - 2h)^2}{\sigma_x^2 - k} \end{cases}$$

Le facteur de biais, K est égal à :

$$(II, 17) \quad K = e^{\frac{h-k}{2} - \frac{(h-k)^2}{2(\sigma_x^2 - k)}} \left[ \frac{x}{m} \right]^{\frac{k-h}{\sigma_x^2 - k}}$$

L'expression  $(hk - h)$  est positive et petite :

$$k - h = \alpha \left[ \frac{\lambda}{6} + \frac{\lambda^2}{12} + \dots \right]$$

Pour une dispersion de 10/100 et un  $\lambda$  de  $\frac{1}{10}$ , cette expression est de l'ordre de  $\frac{1}{600}$ . La correction introduite par le facteur de biais K n'excédera donc jamais quelques millièmes. Il n'y a donc pratiquement pas lieu d'en tenir compte. Il est intéressant de remarquer que ce biais très léger joue en sens inverse de la correction de

Krige, une périphérie très riche encadrant, en moyenne, un panneau plus riche encore, et réciproquement. De même la variance dépend en principe de  $\sigma_x^2$ , c'est-à-dire de la taille du panneau, mais en pratique, l'expression  $(k - 2h)^2$ , qui est de l'ordre de  $\frac{\alpha^2}{4}$  est négligeable, de sorte que la variance d'échantillonnage peut être considérée comme indépendante de la taille du panneau et égale à :

$$(II, 18) \quad \sigma_{y,x}^2 = p + 2h - k = -\frac{\alpha}{\lambda^2} \left[ \lambda(1 - \lambda) + \lambda \right]$$

Cette expression se développe en série :

$$(II, 19) \quad \sigma_{y,x}^2 = \alpha \left[ \frac{1}{2} + \frac{\lambda}{3} + \frac{\lambda^2}{4} + \dots \right]$$

C'est une fonction croissante de  $\lambda$ . Plus épaisse est la périphérie enlevée, moins bonne est la précision sur le panneau résiduel. Étant donné que les dimensions d'une galerie ou d'un montage sont à peu près toujours les mêmes, le rapport  $\lambda$  est d'autant plus grand que le panneau est plus petit. D'où cette conclusion un peu paradoxale : la précision d'un échantillonnage périphérique est d'autant meilleure que le panneau échantillonné est plus grand. Il s'agit là, toutefois, d'un effet différentiel, et la variance d'échantillonnage ne diffère que très peu de  $\frac{\alpha}{2}$ . D'ailleurs, si l'on avait considéré  $x$  comme estimateur, non du panneau résiduel, mais du panneau entier, périphérie comprise, dont la teneur est  $z$ , on aurait trouvé :

$$(II, 20) \quad \sigma_{z,x}^2 = 2\eta - k$$

Égale à  $\frac{\alpha}{2}$  pour  $\lambda = 0$ , cette expression est une fonction décroissante de  $\lambda$  et la conclusion est inversée. On voit que la théorie de la périphérie évanescence est utilisable dans la mesure même où l'on estime qu'il n'y a pas lieu de distinguer le panneau intact et le panneau résiduel.

ANNEXE III

Théorème du statisticien d'échantillonnage

Soit un échantillon S, de forme quelconque et dont la teneur  $x$  doit servir à l'estimation de la teneur  $z$  d'un panneau S', dont la forme et la disposition relativement à S sont arbitraires. Soit également un élément  $dS$  de teneur  $u$  supposé choisi au hasard à l'intérieur de S. Lorsque  $x$  est fixé, les variables  $u$  et  $z$  sont indépendantes

En effet,  $dS$  étant choisi au hasard, on a :

$$(III, 1) \quad \sigma_{uz} = \sigma_{uz} \quad \text{et} \quad \sigma_{ux} = \sigma_x^2$$

et l'expression :

$$H = \sigma_x^2 \sigma_{uz} - \sigma_{ux} \sigma_{xz}$$

est nulle, d'où résulte :

$$(III, 2) \quad \rho_{uz} = \rho_{ux} \rho_{xz}$$

égalité qui démontre le théorème.

Il va de soi que s'il était possible de choisir un certain nombre de  $dS$  occupant des positions privilégiées telles que les égalités (III, 1) ne soient pas vérifiées, il subsisterait une possibilité d'amélioration. Toutefois, le long d'une périphérie, les différentes volées occupent des positions équivalentes et les équations (III, 1) sont vérifiées.

La signification pratique de ce théorème est évidente. Il est impossible à partir des  $n$  volées successives d'une galerie, ou des  $n$  tronçons de carottes d'un sondage analysés séparément, de former un estimateur dont l'efficacité soit supérieure à celle de la teneur moyenne de la galerie ou du sondage.

ANNEXE IV

Amélioration d'un échantillonnage périphérique par sondages intérieurs

Soit un échantillon  $dS$  de teneur  $u$ , intérieur au panneau. Limitons-nous au cas de la périphérie évanescence. On a tout d'abord, en vertu des propriétés de la fonction  $L_r$ , quelle que soit la position de  $dS$  :

$$\sigma_{uz} = \sigma_{uz} = \sigma_x^2$$

Par contre la covariance  $\sigma_{uz}$  dépend de la position de  $dS$ . Ceci étant, déterminons la loi de  $z$ , teneur du panneau, lorsque  $u$  et  $x$  sont connus. On trouve facilement, pour la valeur probable de  $z$  à  $u$  et  $x$  fixés.

$$(IV, 1) \quad z' = x^{1-a} u^a e^{\dots}$$

Avec :

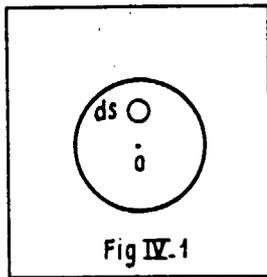
$$(IV,2) \quad \begin{cases} a = \frac{\sigma_{uz} - \sigma_z^2}{\sigma_z^2 - \sigma_x^2} \\ b = \frac{\sigma_{uz} - \sigma_x^2}{2} \left[ 1 - \frac{\sigma_{uz} - \sigma_z^2}{\sigma_z^2 - \sigma_x^2} \right] \end{cases}$$

Et pour la variance de  $z$  à  $x$  et  $u$  fixés :

$$(IV,3) \quad \sigma_{z,x}^2 = \frac{\alpha}{2} \frac{(\sigma_{uz} - \sigma_z^2)^2}{\sigma_z^2 - \sigma_x^2}$$

**Choix de l'implantation du sondage supplémentaire.** — L'implantation la plus favorable est celle qui rend maximale l'expression  $\sigma_{uz} - \sigma_z^2$ , c'est-à-dire le centre du cercle. On a alors, d'après la formule (II, 5) :

$$\sigma_{uz} - \sigma_z^2 = \alpha$$



Posons, pour simplifier :

$$\sigma_z^2 = \sigma_{uz}^2 - \sigma_x^2$$

$\sigma_x^2$  est la variance dans le panneau d'un échantillon de la taille

du sondage  $ds$ . On trouve :

$$(IV,4) \quad \begin{cases} a = \frac{\alpha}{\sigma_z^2 + \alpha/2} \\ b = \frac{\alpha}{2} \left[ 1 - \frac{\alpha}{\sigma_z^2 + \alpha/2} \right] \\ \sigma_{z,x}^2 = \frac{\alpha}{2} - \frac{\alpha^2}{\sigma_z^2 + \alpha/2} \end{cases}$$

Le gain de précision apporté par le sondage supplémentaire, effectué au centre est représenté par le terme  $\frac{\alpha^2}{\sigma_z^2 + \frac{\alpha}{2}}$  : il est donc très faible.

**Exemple numérique :** Soit un panneau de 1 600 mètres carrés dans un gisement de dispersion absolue  $\alpha = \frac{9}{100}$  (gisement d'uranium, par exemple).

L'échantillonnage périphérique conduit à une variance d'échantillonnage de  $\frac{4,5}{100}$  et à un intervalle de confiance = 0,65 — 1,53.

On effectue un sondage supplémentaire au centre (diamètre 56 mm). On a :

$$\sigma_z^2 = \alpha L \frac{1600 \cdot 4 \cdot 10^4}{\pi \cdot 5,6^2} = 1,2$$

La variance d'échantillonnage passe de  $\frac{4,5}{100}$  à  $\frac{3,8}{100}$  et l'intervalle de confiance à 0,67 — 1,48 (le gain est très faible).

La valeur probable de la teneur du panneau est :

$$z = 1,04 u^{0,073} x^{0,927}$$

donc très peu différente de  $x$ . Pour  $x = 1$  et  $u = 10$ , par exemple, on prendra pour  $z$  la valeur 1,23.

### ANNEXE V

#### Précision des sondages à maille rigide

Nous nous limitons ici au cas où la variance  $\sigma^2$  des sondages dans le gisement, ou, si l'on préfère, la dispersion absolue  $\alpha$  est connue a priori. La variance d'échantillonnage est :

$$(V,1) \quad \sigma_z^2 = \frac{\sigma^2}{n} + \frac{1}{n^2} \sum_{i \neq j} \sigma_{x_i, x_j}$$

$\sigma_{x_i, x_j}$  représente la covariance des sondages d'ordre  $i$  et  $j$ .

En réalité, du reste, seule nous intéresse la valeur moyenne  $\overline{\sigma_{x_i, x_j}}$  de cette covariance. La sommation étant étendue à  $n(n-1)$  termes, nous pouvons donc écrire :

$$(V,2) \quad \sigma_z^2 = \frac{\sigma^2}{n} + \frac{n-1}{n} \overline{\sigma_{x_i, x_j}}$$

Étant donné un sondage  $x_i$ , la covariance  $\sigma_{x_i, x_j}$  peut être assimilée à la covariance  $\sigma_{x_i, z_j}$  de  $x_i$  et de la teneur  $z_j$  du polygone d'influence du sondage  $x_j$ . (Cette assimilation serait rigoureusement exacte si ce polygone était un cercle.) Or la valeur moyenne de  $\sigma_{x_i, z_j}$  s'obtient très facilement en écrivant que la covariance de  $x$

et du gisement entier est nulle :

$$\sigma_{x, z_i} + (n-1) \overline{\sigma_{x_i, z_j}} = 0$$

Quant à la covariance  $\sigma_{x_i, z_i}$  de  $x_i$  et de son propre polygone d'influence, elle est donnée par l'équation (II, 8) où l'on fait  $\lambda = 1$ .

$$\sigma_{x_i, z_i} = \sigma_x^2 + \frac{\alpha}{2}$$

On a donc :

$$\overline{\sigma_{x_i, x_j}} = \overline{\sigma_{x_i, z_j}} = -\frac{\sigma_x^2 + \alpha/2}{n-1}$$

et, en portant dans (V, 2) :

$$(V,3) \quad \sigma_z^2 = \frac{\sigma^2 - \sigma_x^2 - \alpha/2}{n}$$

Dans le cas de l'échantillonnage stratifié au hasard, on a simplement :

$$\sigma_{x_i, z_i} = \sigma_x^2$$

et par suite :

$$(V,4) \quad \sigma_z^2 = \frac{\sigma^2 - \sigma_x^2}{n}$$

**Exemple numérique**

Soit un gisement de 100 000 mètres carrés reconnu par 40 sondages à la maille de 50 mètres. La dispersion absolue est de 9/100 (gisement d'uranium par exemple). Si l'on suppose la forme du gisement assez ramassée pour que la formule de DE WIJS ordinaire soit valable, on trouve :

$$\sigma^2 = 1.50, \quad \sigma_x^2 = 0.33, \quad \frac{\alpha}{2} = 0.045$$

On trouve les valeurs suivantes pour la variance d'échantillonnage et l'intervalle de confiance :

	$\sigma_z^2$	Coefficients de sécurité
Maille rigide.....	0.028	0.715 - 1.40
Schéma stratifié au hasard.	0.029	0.71 - 1.41
Schéma aléatoire pur....	0.037	0.68 - 1.47

ANNEXE VI

Estimation de la teneur d'un polygone d'influence

**a. Estimation à partir de la seule teneur x du sondage central.**

Soient Z la teneur du panneau et x celle du sondage central. De l'équation :

$$\sigma_{z,x} = \sigma_z^2 + \frac{\alpha}{2}$$

on déduit les paramètres de la loi de z à x fixé.

$$(VI,1) \quad \left\{ \begin{aligned} z'_{z,x} &= m^{1-a} x^a e^{\alpha/4} \\ u &= \frac{\sigma_z^2 + \alpha/2}{\sigma^2} \\ \sigma_{z,x}^2 &= \sigma_z^2 - \frac{(\sigma_z^2 + \alpha/2)^2}{\sigma^2} \end{aligned} \right.$$

m représente ici la teneur moyenne du gisement (supposée connue). La valeur probable de z à x fixé est :

$$(VI,2) \quad z' = m^{1-a} x^a e^{\alpha/4} + \frac{\sigma_{z,x}^2}{2}$$

la variance d'échantillonnage étant  $\sigma_{z,x}^2$ . Dans l'exemple numérique précédent, on trouverait :

$$z' = 1.15 m^{0.75} x^{0.25}$$

avec une variance d'échantillonnage  $\sigma_{z,x}^2 = 0.236$  et un intervalle de confiance 0.38 — 2.62.

*Remarque.* — Si le sondage était implanté au hasard dans le panneau, on aurait  $\sigma_{z,x} = \sigma_z^2$ , et on retomberait sur la formule de KRICE, formule que l'on obtient en faisant disparaître des formules

(VI, 1) le terme en  $\frac{\alpha}{2}$ , soit numériquement, dans le cas du même exemple :

$$\left\{ \begin{aligned} z' &= 1.138 m^{0.78} x^{0.22} \\ \sigma_{z,x}^2 &= 0.258 \end{aligned} \right.$$

soit un intervalle de confiance : 0.37 — 2.73, nettement moins bon que dans le cas où le sondage est implanté au centre.

**b. Estimation à partir du sondage central et des six sondages les plus proches.**

Si la maille est hexagonale, les six sondages les plus proches, de teneur  $x_1, x_2, \dots, x_6$  sont équidistants du sondage central et interviennent symétriquement. Il suffit donc de poser :

$$(VI,3) \quad u = \sqrt[6]{x_1 x_2 \dots x_6}$$

et de déterminer la loi de Z lorsque x et u sont fixés. On voit immédiatement que la variance a priori de u est :

$$(VI,4) \quad \sigma_u^2 = \frac{1}{6} \left[ \sigma^2 + 2\sigma_{x_1, x_2} + 2\sigma_{x_1, x_3} + \sigma_{x_1, x_4} \right]$$

et que la covariance de x et de u est :

$$(VI,5) \quad \sigma_{u,x} = \sigma_{x_1, x_1} = \sigma_{x_1, x_2} = \sigma_{u,z}$$

Il est commode d'introduire la variance  $\sigma_z^2$  du cercle circonscrit à l'hexagone  $x_1, x_2, \dots, x_6$ , c'est-à-dire :

$$(VI,6) \quad \sigma_z^2 = \sigma_x^2 - aL \frac{2\pi}{\sqrt{3}}$$

On a en effet, d'après les formules (II, 5) :

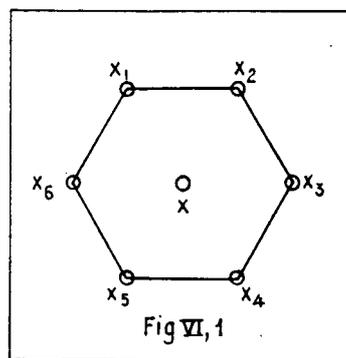
$$(VI,7) \quad \left\{ \begin{aligned} \sigma_{u,x} &= \sigma_{u,z} = \sigma_{x_1, x_1} = \sigma_x^2 - \frac{\alpha}{2} \\ \sigma_{x_1, x_2} &= \sigma_x^2 - \frac{\alpha}{2} - aL3 \\ \sigma_{x_1, x_3} &= \sigma_x^2 - \frac{\alpha}{2} - aL4 \end{aligned} \right.$$

De sorte que l'on a :

$$(VI,8) \quad \sigma_u^2 = \frac{\sigma^2}{6} + \frac{5}{6} \sigma_x^2 - \frac{5\alpha}{12} - \frac{\alpha}{3} L6$$

Ceci étant, il est facile de calculer les paramètres de la loi de z lorsque u et x sont fixés. La variance liée de z est :

$$(VI,9) \quad \sigma_{z,x}^2 = \frac{[\sigma_z^2 \sigma_x^2 - \sigma_{u,x}^2] [\sigma^2 - \sigma_x^2 - \alpha] - \alpha^2/4 \sigma_x^2}{\sigma^2 \sigma_x^2 - \sigma_{u,x}^2}$$



Et la valeur probable de  $z$  à  $x$  et  $u$  fixés :

$$(VI,10) \quad \begin{cases} z' = m^{1-a-b} x^a u^b c^c \\ a = \frac{\sigma_x^2 \sigma_{xz} - \sigma_{xz}^2}{\sigma_x^2 \sigma_u^2 - \sigma_{xu}^2} \\ b = \frac{\sigma_u^2 \sigma_{xz} - \sigma_{xz} \sigma_{xu}}{\sigma_x^2 \sigma_u^2 - \sigma_{xu}^2} \\ c = \frac{(\sigma_z^2 - \sigma_{xz}) [\sigma_x^2 \sigma_{xz} + \sigma_u^2 \sigma_{xz} - 2\sigma_{xu}^2]}{2 [\sigma_x^2 \sigma_u^2 - \sigma_{xu}^2]} \end{cases}$$

d'où :

$$a = 0.20 \quad b = 0.40 \quad c = 0.38$$

$$z' = 1.46 \quad m^{0.4} \quad x^{0.2} \quad u^{0.4}$$

avec :

$$\sigma_{z'} = 0.187$$

Dans l'exemple numérique traité plus haut :

$$\begin{cases} \sigma_{ux} = \sigma_{uz} = \sigma_{x_1 x_2} = 0.169 \\ \sigma_{x_1 x_3} = 0.070 \\ \sigma_{x_1 x_4} = 0.044 \\ \sigma_x^2 = 0.337 \\ \sigma_{xz} = 0.375 \end{cases}$$

L'intervalle de confiance est : 0.42 — 2.37 : ce n'est pas brillant mais nettement meilleur que dans le cas précédent. On peut penser améliorer encore un peu cette précision en tenant compte d'une deuxième auréole de sondages. Cependant les 6 sondages de la première auréole représentée par le terme  $u$  constituent une périphérie évanescence, un peu trouée, mais suffisante, par suite d'un effet analogue à la cage de FARADAY, pour permettre au théorème de l'écran de jouer en partie : il est en fait inutile de tenir compte d'une deuxième auréole.

ANNEXE VII

A. Les formules du rectangle

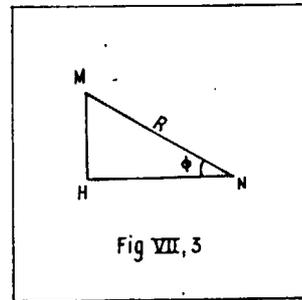
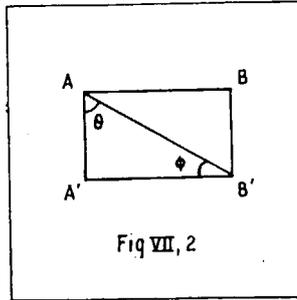
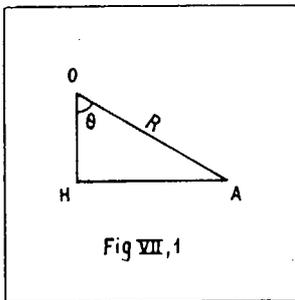
L'application aux tracés et aux montages de la formule de DE WIJS généralisée oblige à intégrer l'expression  $Lr$  dans un contour rectangulaire. Les étapes d'une telle intégration sont les suivantes :

1° Covariance d'une droite et d'un point.

Le point O est supposé situé sur la perpendiculaire en H à la droite HA. La valeur moyenne de  $Lr$ ,  $r$  étant la distance du point O à un point quelconque de HA, est :

$$(VII,1) \quad I_1(R, \theta) = LR - 1 + \frac{\theta}{\text{tg} \theta}$$

R représente la distance OA, et  $\theta$  l'angle en O.



4° Variance du rectangle.

La valeur moyenne de  $Lr$  dans le rectangle est :

$$(VII,4) \quad I_1(R, \Phi) = LR - \frac{25}{12} + \frac{1}{6} \frac{L \cos \Phi}{\text{tg}^2 \Phi} + \frac{1}{6} \text{tg}^2 \Phi L \sin \Phi + \frac{2}{3} \frac{\Phi}{\text{tg} \Phi} + \frac{2}{3} \left( \frac{\pi}{2} - \Phi \right) \text{tg} \Phi$$

Comme on devait s'y attendre, cette expression ne change pas si l'on remplace  $\Phi$  par  $\frac{\pi}{2} - \Phi$ .

2° Covariance des deux côtés, AB A'B', parallèles d'un rectangle.

$r$  étant la distance de deux points quelconques, l'un sur AB, l'autre sur A'B', la valeur moyenne de  $Lr$  est :

$$(VII,2) \quad I_2(R, \theta) = LR - \frac{3}{2} + \frac{2}{\text{tg} \theta} + \frac{L \cos \theta}{\text{tg}^2 \theta}$$

3° Covariance du rectangle et d'un de ses côtés AB.

Les deux points quelconques sont pris l'un sur AB, et l'autre dans le rectangle. La valeur moyenne de  $Lr$  est :

$$(VII,3) \quad I_3(R, \Phi) = LR - \frac{11}{6} + \left( \frac{\pi}{2} - \Phi \right) \text{tg} \Phi + \frac{\text{tg}^2 \Phi}{3} L \sin \Phi + \frac{\Phi}{3 \text{tg} \Phi}$$

Lorsque l'angle  $\Phi$  est petit,  $I_1$  a pour partie principale  $LR - \frac{3}{2}$

Le premier terme est  $\frac{\pi}{2} \text{tg} \Phi$ . Pour une galerie de 200 mètres de long sur 2 mètres de haut, ce terme est de l'ordre de 1/100 : il est donc négligeable, et la variance de la galerie se calcule, comme si la hauteur était nulle, à partir de :

$$(VII,5) \quad I_4 = Ll - \frac{3}{2}$$

$l$  étant la longueur de la galerie, il en résulte qu'une galerie infiniment mince, donc d'aire nulle, a une variance à peu près égale à celle d'un panneau carré de côté égal à  $\frac{l}{2}$ .

5° Covariance d'un traçage et d'un montage.

Il résulte de la remarque précédente qu'il suffit de calculer la covariance de deux segments HM et HN perpendiculaires. La valeur moyenne de  $Lr$  s'obtient en intégrant, sur HM, la fonction  $I_1$  du point M et du segment HN. On trouve :

$$(VII,6) \quad J(R,\Phi) = LR - \frac{3}{2} + \frac{\Phi}{2 \operatorname{tg} \Phi} + \frac{1}{2} \left( \frac{\pi}{2} - \Phi \right) \operatorname{tg} \Phi$$

Comme  $I_1$ , cette expression est invariante lorsque l'on change  $\Phi$ , en  $\frac{\pi}{2} - \Phi$ .

Exemple. — Variance d'un panneau carré de côté  $a$  et de sa périphérie.

Pour un carré, on a  $\Phi = \frac{\pi}{2}$  et la formule (VII, 4) donne :

$$I_4 = La + \frac{1}{2} L2 - \frac{25}{12} - \frac{1}{6} L2 + \frac{\pi}{3} = La - 0.81$$

La variance du carré est donc :

$$\sigma_0^2 = \Lambda - 2\alpha La + 1.62\alpha$$

D'autre part un cercle de rayon  $R$  a pour variance, d'après la formule (II, 4) :

$$\sigma_0^2 = \Lambda - 2\alpha LR + \frac{\alpha}{2}$$

Prenons un cercle de même aire  $a^2$  que le carré précédent :

$$(\pi R^2 = a^2).$$

On a pour ce cercle :

$$\sigma_0^2 = \Lambda - 2\alpha La + \alpha \left( L\pi + \frac{1}{2} \right) = \Lambda - 2\alpha La + 1.645\alpha$$

On a donc :

$$(VII,7) \quad \sigma_0^2 = \sigma_0^2 - \frac{2.5}{100} \alpha$$

De toutes les surfaces de même aire, le cercle conduit à la variance la plus grande, puisque c'est pour le cercle que  $Lr$  a la plus petite valeur moyenne : il en résulte que le panneau circulaire est, de tous, le plus facile à échantillonner, puisque c'est dans un tel panneau qu'un échantillon de forme donnée prend la plus faible variance.

Du point de vue pratique, le terme  $\frac{2.5}{100} \alpha$  étant parfaitement négligeable, il est tout à fait légitime d'assimiler carré et cercle. Si, au lieu d'un carré nous avions pris un hexagone, la différence eût été

encore plus faible. Cette remarque montre que, de tous les modes possibles d'implantations de sondages, le meilleur est encore la maille rigide, hexagonale ou carrée, avec un petit avantage pour la maille hexagonale, dans la mesure même où l'hexagone se rapproche davantage du cercle que le carré (en pratique, les deux mailles sont du reste équivalentes à un ou deux centièmes de  $\alpha$  près).

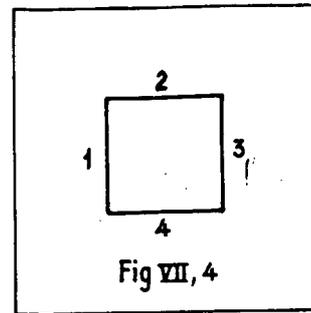


Fig VII, 4

A titre d'exercice, et pour justifier la théorie de la périphérie évanescence, calculons également la variance de la périphérie du carré, c'est-à-dire la variance d'un échantillon constitué des quatre côtés d'un carré.

En désignant par 1, 2, 3 et 4 les côtés du carré, cette variance est égale à :

$$\sigma_0^2 = \frac{1}{16} [4\sigma_1^2 + 8\sigma_{12} + 4\sigma_{13}]$$

Les formules (VII, 2), (VII, 5) et (VII, 6) montrent que l'on a :

$$\sigma_0^2 = \Lambda - 2\alpha \left[ La + \frac{1}{4} L2 - \frac{3}{2} + \frac{\pi}{4} \right] = \Lambda - 2\alpha La + 1.084\alpha$$

On a donc :

$$\sigma_0^2 = \sigma_0^2 - 0.536\alpha$$

Dans le cas du cercle on a :

$$\sigma_0^2 = \sigma_0^2 - 0.5\alpha$$

ici encore la différence est de quelques % de  $\alpha$  : on peut conclure que la théorie de la périphérie circulaire est pratiquement applicable au carré.

B. Évaluation d'un gisement entre deux cotes données

Soit un gisement reconnu à plusieurs niveaux distants de  $2a$ . On considère généralement la teneur moyenne d'un niveau  $N$  comme représentative de la teneur du panneau compris entre les cotes  $N + a$  et  $N - a$ . Soit  $x$  la teneur de la galerie et  $z$  celle du panneau, et  $l$  la longueur de la galerie. On a :

$$(VII,8) \quad \left\{ \begin{array}{l} \sigma_z^2 = \Lambda - 2\alpha Ll + 3\alpha \\ \sigma_x^2 = \Lambda - 2\alpha I_1(\Phi) \\ \sigma_{xz} = \Lambda - 2\alpha I_3(\Phi_1) \end{array} \right. \quad \left\{ \begin{array}{l} \Phi = \operatorname{Arc} \operatorname{tg} \frac{2a}{l} \\ \Phi_1 = \operatorname{Arc} \operatorname{tg} \frac{a}{l} \end{array} \right.$$

Il n'est malheureusement plus possible de reprendre le raisonnement de KRICE et d'étudier la loi de  $z$  à  $x$  fixé — des formules du

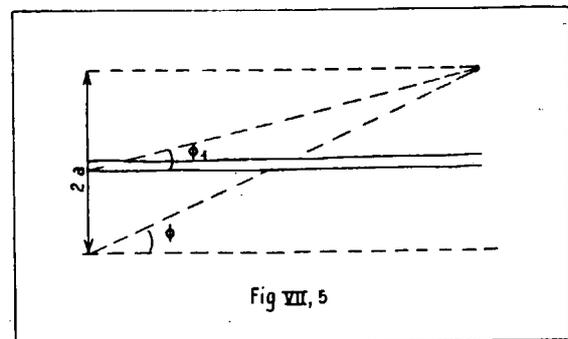


Fig VII, 5

type de KRIGE supposent, en effet, la teneur moyenne du gisement connue *a priori*. Une telle hypothèse, légitime lorsqu'il s'agit du polygone d'influence d'un sondage parmi une cinquantaine, car alors la moyenne générale est connue avec une précision bien meilleure que la teneur du polygone, cesse d'être valable lorsque le panneau occupe à lui seul le tiers ou le quart du gisement : la moyenne générale est alors, en effet, connue avec une précision du même ordre de grandeur que celle du panneau. A la limite, si le panneau était identique au gisement, le raisonnement de KRIGE conduirait à  $z' = m$  et  $\sigma_z^2 = 0$  ce qui est vrai physiquement, la teneur du panneau étant alors bien déterminée et égale à  $m$ , mais n'avance à rien puisque en fait, on ne connaît pas  $m$ . Dans ces conditions, négligeant le biais, qui, s'il existe, est nécessairement très léger, nous prendrons la teneur  $x$  de la galerie comme estimateur de la teneur  $z$  du panneau. La précision de cet estimateur sera considérée comme représentée

par la variance de la variable  $\frac{z}{x}$ , soit :

$$(VII,9) \quad \sigma_z^2 = \sigma_x^2 + \sigma_z^2 - 2\sigma_{xz}$$

expression indépendante de  $A$ , c'est-à-dire de la taille et de la forme du gisement, en accord avec notre volonté de ne pas tenir compte de données extérieures au panneau.

On trouve facilement :

$$(VII,10) \quad \sigma_z^2 = 4\alpha I_3(\Phi_1) - 2\alpha I_4(\Phi) - 2\alpha Ll + 3\alpha$$

En remplaçant  $I_3$  et  $I_4$  par leurs valeurs, on s'aperçoit que les termes en  $Ll$  disparaissent et que cette expression ne dépend que des angles  $\Phi$  et  $\Phi_1$ . En pratique, d'ailleurs, l'équidistance  $2a$  des niveaux est par exemple de 40 mètres, alors que la longueur de galeries est de plusieurs centaines de mètres.  $\text{tg } \Phi$  est de l'ordre de  $\frac{1}{10}$

et  $\text{tg } \Phi_1$  de  $\frac{1}{20}$ . Il semble donc légitime de développer en  $\Phi$  et  $\Phi_1$  les fonctions  $I_3$  et  $I_4$ , en se limitant au terme en  $\Phi$  et au terme suivant qui est en  $\Phi^2$   $L\Phi$  : on a ainsi :

$$(VII,11) \quad \begin{cases} I_3(\Phi_1) = Ll - \frac{3}{2} + \frac{\pi}{2} \Phi_1 + \frac{\Phi_1^2}{3} L\Phi_1 \\ I_4(\Phi) = Ll - \frac{3}{2} + \frac{\pi}{3} \Phi + \frac{\Phi^2}{6} L\Phi \end{cases}$$

De plus, et toujours en négligeant les termes en  $\Phi^2$ , nous pouvons écrire :

$$\Phi_1 = \frac{\Phi}{2}$$

ce qui donne pour  $I_3$  :

$$I_3\left(\frac{\Phi}{2}\right) = Ll - \frac{3}{2} + \frac{\pi}{4} \Phi + \frac{\Phi^2}{12} L\Phi$$

On trouve ainsi, pour la variance d'échantillonnage :

$$\sigma_z^2 = \frac{\pi}{3} \Phi \alpha$$

c'est-à-dire, en confondant  $\Phi$  et  $\text{tg } \Phi$  :

$$(VII,12) \quad \sigma_z^2 = \alpha \frac{\pi 2a}{3 \cdot 7}$$

Il suffit de multiplier par  $l$  numérateur et dénominateur pour faire apparaître la surface  $S = 2 \alpha l$  du panneau échantillonné :

$$(VII,13) \quad \sigma_z^2 = \alpha \frac{\pi S}{3 L^2}$$

Les autres cas, correspondant aux formules (17), (18) et (19) du texte ne sont pas développés ici, mais se traitent de façon analogue, à partir des formules du rectangle.

### C. Comparaison avec l'expérience

Les formules du rectangle permettent de tester la valeur de notre hypothèse de base, c'est-à-dire la formule de la covariance. En effet, le rapport  $\frac{x}{z}$  de la teneur  $x$  d'une volée à la teneur  $z$  d'un tronçon de galerie contenant  $n$  volées est une variable lognormale dont la variance, calculée à l'aide des formules du rectangle, apparaît comme une fonction de la position de la volée dans le tronçon : elle est d'autant plus élevée que la volée occupe une position moins centrale. Il est aisé, dans le cas d'un échantillonnage volée par volée du tout-venant, de calculer la valeur expérimentale de cette variance, et de la comparer à la valeur théorique.

Cette confrontation a été faite pour le petit gisement de plomb de Bou-Kiama, près de Bougie (Algérie). La dispersion absolue, calculée sur l'ensemble du gisement, est de  $\frac{5,2}{100}$ . On a fait le choix

d'un tronçon de galerie d'une soixantaine de mètres, où l'échantillonnage semblait avoir été particulièrement soigné, et où l'on disposait de 53 volées consécutives. Ces volées ont été, de toutes les façons possibles, regroupées, par suites de 2, 4, 8 et 16, et les variances ont été calculées pour chaque position dans chaque suite. On dispose ainsi, pour chaque position, de 104, 100, 92 et 76 données respectivement dans les groupements par 2, 4, 8 et 16. Les positions sont numérotées à partir de 1 à l'intérieur vers l'extérieur. Les résultats sont présentés sur le tableau suivant :

Nombre de volées groupées	Position	Variance théorique	Variance expérimentale
2		$0.634 \alpha = 0.033$	0.0338
4	1	$0.96 \alpha = 0.05$	0.061
	2	$2 \alpha = 0.104$	0.091
8	1	$1.87 \alpha = 0.097$	0.14
	2	$2.06 \alpha = 0.107$	0.15
	3	$2.71 \alpha = 0.141$	0.17
	4	$3.65 \alpha = 0.19$	0.20
16	(valeur moyenne)	$3.8 \alpha = 0.197$	0.195

L'accord est loin d'être parfait. Il est certain cependant que le sens et l'ordre de grandeur des variations sont bien représentés.