

G. Mathéron 1959

GM/EC

- NOTE STATISTIQUE N° 21 -

Comparaison des Correlations théoriques avec les données expérimentales

Le but de cette étude est de comparer numériquement avec l'expérience les formules théoriques déduites de l'hypothèse d'une covariance élémentaire de forme logarithmique, hypothèse sur laquelle repose toute la théorie de l'échantillonnage des gisements.

Les données numériques de départ sont empruntées à de WIJS ( Statistics of ore distribution - Géologie en Mijnbouw - Nov 1951). Dans le but de tester sa loi log-Binomiale (équivalente pratiquement à une loi log-normale), de WIJS a choisi la Mine de Pulacayo (Bolivie). Il s'agit d'un filon de quartz et blende, reconnu au Niveau 446 par un traçage de 240 m de long. Ce traçage a été échantillonné par 118 saignées distantes de 2m. La caisse filonienne a une puissance moyenne de 0.50m, mais les épontes contiennent de la blende disséminée et des filonnets satellites. Les saignées ont été effectuées sous la puissance standard de 1.30m, qui répond aux conditions de l'exploitation ultérieure. Les teneurs, données en  $\frac{Zn}{m^2}$ , correspondent donc, en réalité, à des accumulations: elles ne diffèrent que par un facteur constant du poids de zinc au mètre carré d'épontes. Elles sont d'autre part données dans l'ordre de leur succession naturelle dans le traçage, d'où possibilité de calculer à partir d'elles les corrélations en fonction de la distance.

Le tableau I reproduit les données originales de de WIJS.

| Echantillon N° | Zn % | Ech N° | Zn % | Ech N° | ZN % |
|----------------|------|--------|------|--------|------|
| I              | 17,7 | 41     | 15,3 | 81     | 13,8 |
| 2              | 17,8 | 42     | 24,0 | 82     | 6,5  |
| 3              | 9,5  | 43     | 12,3 | 83     | 6,5  |
| 4              | 5,2  | 44     | 7,8  | 84     | 10,6 |
| 5              | 4,1  | 45     | 9,9  | 85     | 10,6 |
| 6              | 19,2 | 46     | 20,7 | 86     | 23,0 |
| 7              | 12,4 | 47     | 25,0 | 87     | 21,8 |
| 8              | 15,8 | 48     | 19,1 | 88     | 32,8 |
| 9              | 20,8 | 49     | 13,1 | 89     | 30,2 |
| 10             | 24,1 | 50     | 27,4 | 90     | 30,8 |
| 11             | 14,7 | 51     | 15,2 | 91     | 33,7 |
| 12             | 21,6 | 52     | 12,2 | 92     | 26,5 |
| 13             | 12,8 | 53     | 10,1 | 93     | 39,3 |
| 14             | 11,9 | 54     | 12,3 | 94     | 24,5 |
| 15             | 35,4 | 55     | 16,7 | 95     | 24,9 |
| 16             | 12,3 | 56     | 18,6 | 96     | 23,2 |
| 17             | 14,9 | 57     | 6,0  | 97     | 16,0 |
| 18             | 19,6 | 58     | 10,6 | 98     | 20,9 |
| 19             | 10,6 | 59     | 11,3 | 99     | 10,3 |
| 20             | 15,1 | 60     | 4,7  | 100    | 22,6 |
| 21             | 15,6 | 61     | 10,9 | 101    | 16,2 |
| 22             | 9,3  | 62     | 6,0  | 102    | 22,9 |
| 23             | 8,1  | 63     | 7,2  | 103    | 36,9 |
| 24             | 13,5 | 64     | 5,6  | 104    | 23,5 |
| 25             | 30,2 | 65     | 8,9  | 105    | 18,5 |
| 26             | 29,1 | 66     | 5,8  | 106    | 16,4 |
| 27             | 7,4  | 67     | 8,9  | 107    | 17,9 |
| 28             | 12,3 | 68     | 6,7  | 108    | 18,5 |
| 29             | 13,6 | 69     | 7,2  | 109    | 13,6 |
| 30             | 9,5  | 70     | 9,7  | 110    | 7,9  |
| 31             | 13,1 | 71     | 10,8 | 111    | 31,9 |
| 32             | 27,4 | 72     | 17,9 | 112    | 14,1 |
| 33             | 8,8  | 73     | 10,9 | 113    | 7,1  |
| 34             | 11,4 | 74     | 13,7 | 114    | 3,9  |
| 35             | 6,4  | 75     | 22,3 | 115    | 3,7  |
| 36             | 11,0 | 76     | 10,2 | 116    | 22,5 |
| 37             | 11,4 | 77     | 5,1  | 117    | 27,6 |
| 38             | 14,1 | 78     | 13,9 | 118    | 17,3 |
| 39             | 20,9 | 79     | 19,0 |        |      |
| 40             | 10,6 | 80     | 10,6 |        |      |

1°- Ajustement des données.-

Les 118 teneurs ont été ajustées suivant la loi log-normale. Paramètres =

|                      |                     |
|----------------------|---------------------|
| Médiane              | $\mu = 13,63 \%$    |
| Ecart type           | $\sigma = 0,5286$   |
| Variance             | $\sigma^2 = 0,2794$ |
| Moyenne de Krige     | $m = 15,6 \%$       |
| Moyenne arithmétique | $m = 15,61 \%$      |

2°- Calcul du Correlogramme.-

Le coefficient de corrélation  $r_n$  des teneurs des saignées distantes de  $2n$  mètres a été calculé pour les valeurs de  $n$  de 1 à 25 par la méthode rapide qui consiste à compter le nombre  $N_1$  couples  $x_i, x_{i+2n}$ , pour lesquels les deux valeurs sont simultanément supérieures ou inférieures à la médiane et le nombre  $N_2$  de couples pour lesquels les deux valeurs encadrent la médiane. On a alors =

$$(1) \quad r_n = \sin \frac{\pi}{2} \frac{N_1 - N_2}{N_1 + N_2}$$

Le correlogramme (fig I, courbe I) obtenu par cette méthode de calcul sommaire frappe dès l'abord par son caractère non régularisé, son allure en dent de scie. Il n'est pas possible d'interpoler une courbe continue. Certaines valeurs paraissent anormalement basses ( $r_2$ ), d'autres anormalement élevées ( $r_9$ ). D'autre part, une certaine alternance de maximum et minimum, "en gros" suggère la possibilité d'une tendance périodique, de période égale à 16 mètres = maxima, en gros, observé aux alentours de  $r_8$ ,  $r_{16}$  et  $r_{24}$ . Cependant, cette pseudo périodicité pourrait n'être qu'une apparence, liée au caractère non régularisé du Correlogramme.

Cette irrégularité elle même pose un problème. Elle peut provenir de plusieurs causes =

a- elle peut-être imputée, en partie au moins, à l'emploi de la formule (I). Ce mode de calcul du coefficient de corrélation, avantageux par sa rapidité, laisse échapper une partie de l'information disponible = les valeurs réelles des teneurs sont remplacées par de simples signes = + pour toutes les teneurs supérieures à la médiane, et - pour les autres. Tout couple de teneurs encadrant la médiane est compté pour -1, tout couple situé d'un même côté de la médiane pour + 1. Autrement dit, le couple 13% - 14% apporte la même contribution négative au coefficient que le couple 4% - 35%. Cette perte d'information introduit certainement des irrégularités.

b- L'irrégularité peut provenir aussi, au moins en partie, de l'insuffisance de l'effectif. Les chenaux de confiance des fig. 1 et 2 sont très larges. Les fluctuations attachées à l'estimation de tout paramètre statistique à partir d'un effectif réduit doivent se traduire par de telles irrégularités en dent de scie. Si nous disposions de 2400 saignées distantes de 10 cm, la courbe expérimentale serait certainement plus régulière.

c- En plus de ces causes extérieures d'irrégularité, il se peut enfin que certains traits de notre courbe, comme la valeur "anormalement" basse de  $r_2$  soient dus à des irrégularités locales du correlogramme vrai, la loi "régulière" que nous postulons étant par exemple à peu près vérifiée dans l'ensemble du gisement, tandis que les lois locales peuvent en différer plus ou moins.

Il ne sera pas possible de trancher entre les causes b et c, puisque nous ne disposons pas de 2400 saignées dans le traçage, ni de données relatives à l'ensemble du gisement. Par contre il est facile de mettre en évidence l'influence de la cause a. -(seule responsable, en particulier, de la pseudo périodicité apparente.)

Nous avons pour cela calculé les valeurs des coefficients  $r_1$  à  $r_9$  par la méthode, plus longue mais plus précise, qui consiste à sommer les valeurs des produits des logarithmes des teneurs associées. Ce morceau de correlogramme est représenté par la courbe 2 de la figure I. Il est remarquablement plus régulier que le précédent.

La valeur anormalement basse de  $r_2$  est notablement relevée, l'anomalie de  $r_9$  disparaît complètement. La courbe expérimentale donne encore l'impression d'être légèrement déprimée vers  $r_2$ ,  $r_3$  et  $r_4$ . Mais on ne peut affirmer que cette dépression corresponde à un phénomène réel: la comparaison avec l'ouverture du chenal de confiance montre qu'il n'y a là rien de significatif.

### 3°- Calcul du Correlogramme théorique.

Soit  $h$  la puissance que nous attribuons à la saignée. La variance théorique de la saignée dans le traçage de longueur 240 est :

$$(2) \sigma^2 = 3\alpha \frac{L \cdot 240}{h}$$

La covariance théorique de deux saignées distantes de  $d$  est :

$$(3) \sigma_d = 3\alpha \left[ L \cdot 240 - \frac{3}{2} - Ld \right]$$

Cette formule (3) suppose que  $d$  soit déjà grand par rapport à  $h$ . L'examen des formules du rectangle montre qu'il suffit que  $d$  soit de l'ordre du triple de  $h$  (voir formule 6, p.34, Note St. N° 20).

Lorsque  $d$  au contraire devient très petit devant  $h$ , l'approximation au premier ordre en  $\frac{d}{h}$  de la formule correspondante du rectangle (formule 2, id.) nous donne :

$$\sigma_d = \sigma^2 - 3\alpha \pi \frac{d}{h}$$

d'où l'on déduit le comportement du correlogramme au voisinage de  $d=0$

$$\rho_d = 1 - \frac{3\alpha}{\sigma^2} \pi \frac{d}{h}$$

le correlogramme présente un sommet anguleux et non pas parabolique, et l'on déduit de l'équation précédent la pente de la tangente au sommet.

Le coefficient de corrélation est indépendant de la dispersion absolue. On a en effet =

$$(4) \rho_d = \frac{\sigma_d}{\sigma^2} = \frac{L \cdot 240 - \frac{3}{2} - Ld}{L \cdot 240 - Lh}$$

a- Prenons d'abord pour  $h$  la puissance réelle du filon, soit 0.50 m. La courbe 3 de la figure I donne le correlogramme théorique correspondant. On a représenté également son chenal de confiance à 95 %. Les 9 points de la courbe expérimentale 2 (valeurs de  $r$  calculées par la méthode précise) tombent à l'intérieur du chenal, sauf  $r_2$  qui est marginal = la valeur de  $r_2$ , rejetée au seuil 95 % serait acceptée au seuil 96%. La courbe I (valeurs de  $r$  calculées par la méthode rapide) donne deux points extérieurs,  $r_2$  et  $r_9$ . Toutefois, le chenal de confiance est établi pour des coefficients calculés par la méthode précise. Il est trop sévère pour la courbe I.

Bien qu'on ne puisse conclure au rejet, on a l'impression que la courbe théorique est trop haute par rapport aux deux courbes expérimentales.

b- Pour essayer de réduire cet écart, on peut introduire un coefficient d'anisotropie (facteur de run), c'est à dire admettre que la dimension puissance ne joue pas le même rôle que les deux autres, ou plus simplement admettre que la puissance de 50 cm doit être considérée comme équivalente à une puissance  $h$ , que l'on choisira de façon à approcher le mieux possible le correlogramme expérimental.

L'examen de la formule (4) montre que pour diminuer dans l'ensemble les coefficients théoriques, il faut diminuer  $h$ . La puissance équivalente doit être prise inférieure à la puissance réelle. On se serait attendu peut-être à l'inverse. Cependant la caisse filonienne est remplie d'une blende pratiquement massive d'une éponte à l'autre, ce qui explique peut être que la variabilité soit moins grande dans le sens de la traversée que parallèlement aux épontes.

Modifier  $h$  se traduit par la multiplication de tous les  $P_d$  par un même facteur  $\beta$ . Ce coefficient a été déterminé de façon à rendre minimum l'écart quadratique entre les 9 valeurs expérimentales  $r_1 \dots r_9$  de la méthode précise et les 9 valeurs théoriques correspondantes. On trouve:

$$\beta = 0.74$$

On en déduit la valeur de  $h$

$$0.74 = \frac{L 240 / 0.5}{L 240 / h}$$

soit:

$$h = 5,9 \text{ cm}$$

Cette valeur, remarquablement petite, ne peut plus être considérée

comme représentant une puissance équivalente. Pour des valeurs aussi petites de la puissance, celle-ci cesse d'être dominante dans l'expression de la variance vis à vis des dimensions de la section de la saignée. Il est probable que nos 6 cm représentent plutôt la dimension de la section (aucune indication n'est donnée sur celle-ci par de VIJS), et que la puissance équivalente est encore plus faible

La figure 2 représente le correlogramme théorique correspondant à  $h_s = 5,9$  cm, avec son chenal de confiance et les courbes expérimentales. L'accord est cette fois nettement plus convaincant.

T A B L E A U II

Coefficients de Correlation

| Distance<br>(Unité = 2m) | r expérimental<br>méthode rapide | r expérimental<br>méthode précise | $\rho$ théorique<br>$h_s = 50$ cm | $\rho$ théorique<br>$h_s = 5,9$ cm |
|--------------------------|----------------------------------|-----------------------------------|-----------------------------------|------------------------------------|
| I                        | 0.468                            | 0.474                             | 0.53                              | 0.393                              |
| 2                        | 0.108                            | 0.242                             | 0.42                              | 0.311                              |
| 3                        | 0.257                            | 0.226                             | 0.354                             | 0.262                              |
| 4                        | 0.218                            | 0.157                             | 0.306                             | 0.226                              |
| 5                        | 0.261                            | 0.245                             | 0.270                             | 0.200                              |
| 6                        | 0.346                            | 0.225                             | 0.242                             | 0.179                              |
| 7                        | 0.127                            | 0.139                             | 0.216                             | 0.160                              |
| 8                        | 0.225                            | 0.154                             | 0.194                             | 0.144                              |
| 9                        | 0.406                            | 0.125                             | 0.176                             | 0.130                              |
| 10                       | 0.144                            | .                                 | 0.158                             | 0.117                              |
| 11                       | 0.016                            | .                                 | 0.142                             | 0.105                              |
| 12                       | 0.                               | .                                 | 0.125                             | 0.092                              |
| 13                       | 0.045                            | .                                 | 0.116                             | 0.086                              |
| 14                       | 0.210                            | .                                 | 0.103                             | 0.076                              |
| 15                       | 0.137                            | .                                 | 0.094                             | 0.070                              |
| 16                       | 0.152                            | .                                 | 0.081                             | 0.060                              |
| 17                       | 0.046                            | .                                 | 0.073                             | 0.054                              |
| 18                       | 0.125                            | .                                 | 0.065                             | 0.048                              |
| 19                       | 0.110                            | .                                 | 0.055                             | 0.041                              |
| 20                       | 0.097                            | .                                 | 0.047                             | 0.035                              |
| 21                       | 0.113                            | .                                 | 0.039                             | 0.029                              |
| 22                       | 0.161                            | .                                 | 0.030                             | 0.022                              |
| 23                       | 0.148                            | .                                 | 0.024                             | 0.018                              |
| 24                       | 0.                               | .                                 | 0.018                             | 0.013                              |
| 25                       | 0.151                            | .                                 | 0.010                             | 0.007                              |

## 4° - Applications -

Valeurs de la Dispersion absolue.

La variance est égale à  $3\alpha L^2 40/h$

$$\text{d'où} \quad 3\alpha = \frac{0.279}{L^2 40/h} = \frac{3.32}{100}$$

$$\alpha = \frac{1.1}{100}$$

Précision avec laquelle est connue la teneur moyenne du traçage.

Chaque saignée commande une "zone d'influence" de 2m de long. Elle en fixe la teneur avec une variance d'échantillonnage donnée par la formule (44), p. 15 de la Note 20=

$$\sigma_{E_1}^2 = 3\alpha \left( 1 - 2L^2 + \frac{L^2 200}{5.9} \right) = \frac{1}{10} \quad \left( e^{\pm 2\sigma_{E_1}} = \begin{cases} 1.88 \\ 0.53 \end{cases} \right)$$

La teneur du traçage est connue avec une variance d'échantillonnage 118 fois plus petite.

$$\sigma_{E_1}^2 = \frac{1}{1180} = \frac{8.5}{10000} \quad e^{\pm 2\sigma_{E_1}} = \begin{cases} 1.06 \\ 0.94 \end{cases}$$

C'est à dire à 6 % près.

Précision sur une tranche de 40m.

La variance d'extension pour une telle tranche est donnée par la formule (43) de la Note (20).

$$\sigma_{E_2}^2 = 3\alpha \frac{\pi}{6} \frac{40}{240} = \frac{29}{10000}$$

La variance totale d'échantillonnage est =

$$\sigma_E^2 = \sigma_{E_1}^2 + \sigma_{E_2}^2 = \frac{37.5}{10000} \quad e^{\pm 2\sigma_E} = \begin{cases} 1.13 \\ 0.88 \end{cases}$$

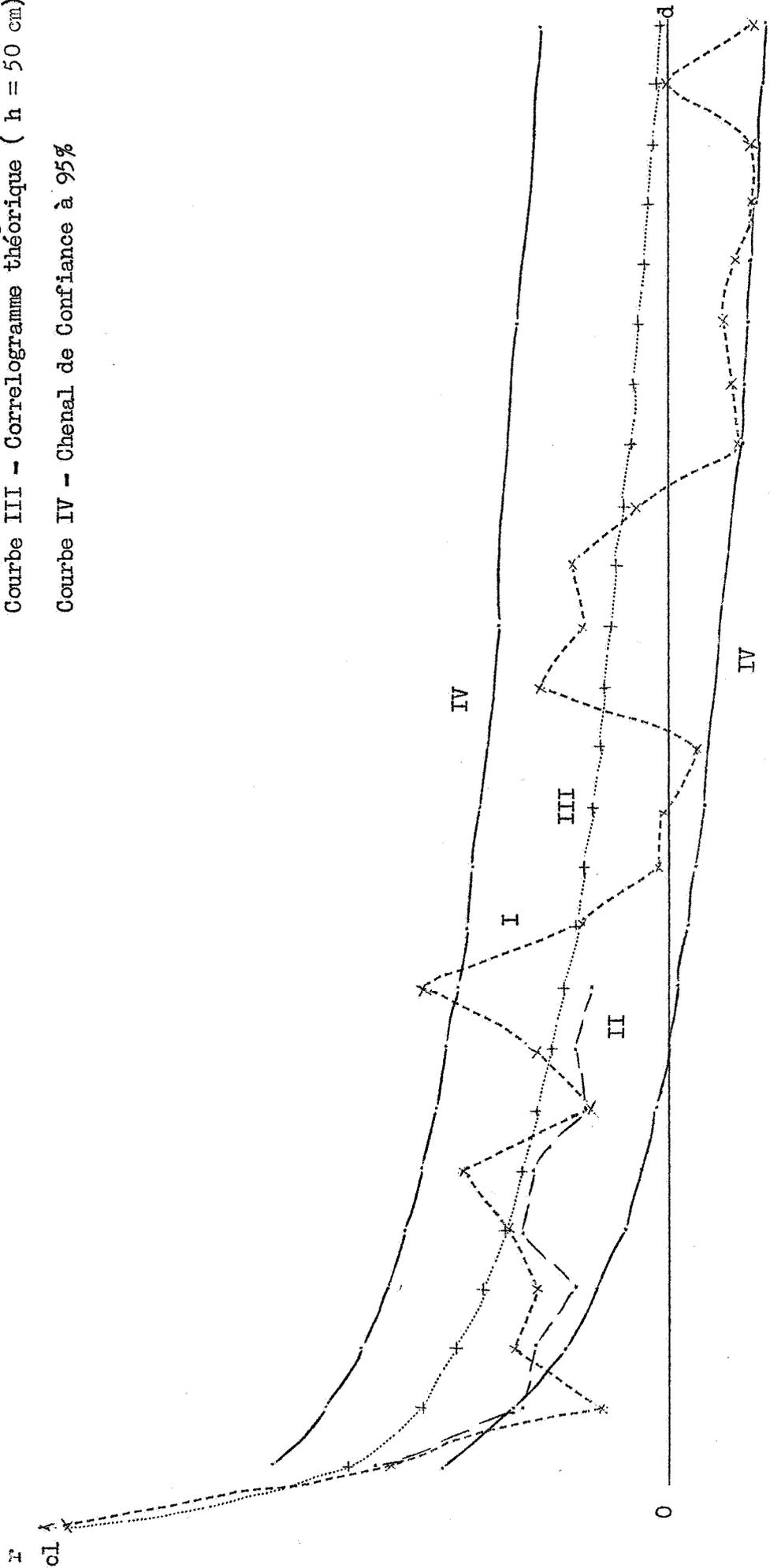
La tranche n'est connue qu'à 13% près.

- Alger - juillet 1959 -

NOTE STATISTIQUE N°21

Pulacaya - Figure I

- Courbe I - Correlogramme expérimental  
( méthode rapide)
- Courbe II - Correlogramme expérimental  
(méthode précise)
- Courbe III - Correlogramme théorique (  $h = 50 \text{ cm}$ )
- Courbe IV - Chenal de Confiance à 95%



NOTE STATISTIQUE N°21

Pulacaya - Figure 2 -

Courbe I - Correlogramme expérimental  
(valeurs précises)  
Courbe II - Correlogramme théorique  
( $h = 5,9 \text{ cm}$ )  
Courbe III - Chenal de confiance à 95 %

