

ÉTUDE DE LA RÉPARTITION SPATIALE DES FILONS AU CONTACT SUD DU MASSIF DE LOS PEDROCHES (ESPAGNE)

G. PANOU

Laboratoire de Géologie Appliquée de l'Université Libre de Bruxelles.

1. Introduction

1.1. Le problème étudié.

Dans le cadre d'une recherche plus large, entreprise par le laboratoire de Géologie Appliquée de l'U.L.B. dans le massif granitique de Los Pedroches [5] nous avons été amenés à étudier le problème de la répartition spatiale des filons subverticaux, situés dans la partie centrale du contact sud de ce massif.

L'expérience avait montré qu'il existait des contrôles géophysiques, tectoniques, stratigraphiques, etc... de la minéralisation. A ces critères nous avons voulu ajouter un contrôle statistique. Le problème posé était donc le suivant : étant donné une suite de n filons, dans quelle région peut-on espérer trouver, avec une probabilité donnée a priori, le $(n + 1)^{\text{e}}$ filon.

Nous avons choisi comme variable la distance EW séparant deux filons, ou leurs prolongements les plus probables, mesurée à 500 m du contact granitique. Les deux raisons qui ont déterminé notre choix sont les suivantes :

- a) Les filons ont des directions à composantes NS dominantes. Il est donc logique de mesurer la distance dans la direction EW.
- b) Les travaux de recherche et d'exploitation ont permis de cartographier dans la zone située à 500 m environ du contact granitique, la position de la plupart des filons. On peut donc, dans cette région, estimer la distance séparant deux filons sans devoir imaginer des prolongements probables et éviter ainsi des erreurs systématiques.

La région étudiée s'étend sur plus de 30 km. Elle doit nécessairement être subdivisée en deux zones :

— La région, que nous appelons principale, est actuellement la mieux connue. C'est dans cette zone que se sont concentrées les exploitations et les recherches. Elle est représentée sur de nombreux documents cartographiques dessinés à grande échelle. Nous avons pu mesurer la quasi totalité des distances sur des cartes du 1/2500 et avons dans de nombreux cas pu comparer les mesures effectuées sur diverses cartes établies à des échelles différentes (de 1/500 à 1/25000).

Malheureusement, la partie centrale de cette région est occupée par une langue de granite qui pénètre assez profondément dans le schiste encaissant ; cette zone centrale est encore très mal connue.

Nous avons décidé de ne pas tenir compte des filons situés dans cette zone. La région principale sera donc elle-même divisée en une zone est et une zone ouest.

Il faut remarquer que ni la limite occidentale de la région principale, ni les limites de la zone centrale ne sont nettes. A l'ouest, par exemple, de la zone principale, nous connaissons encore une trentaine de filons séparés par des distances qui, à quelques exceptions près, sont très largement supérieurs à celles observées dans la région principale ; il en est de même dans la zone centrale.

— La région dite marginale est située à l'est de la principale et se raccorde à celle-ci sans solution apparente de continuité. Cette zone diffère principalement de la précédente par la

nature de la minéralisation. Au lieu du seul Bi, les filons de la région marginale sont « polymétalliques » (Bi, As, Sb, W, etc...). D'autre part, nous ne disposons que d'un seul document cartographique au 1/25000 de la zone marginale et dès lors les distances mesurées pourraient être entâchées d'erreurs importantes.

Pour toutes ces raisons, nous avons cru bon d'étudier séparément les quatre populations suivantes :

Population A : représentée par les 52 distances entre filons situés dans la zone est de la région principale.

Population B : qui regroupe toutes les distances mesurées dans la région principale, à l'exclusion toutefois de la zone centrale.

N.B. Le nombre de distances (24) mesurées dans la zone ouest n'a pas été jugé suffisant pour représenter une population distincte.

Population C : qui est constituée des 80 observations dans la zone marginale.

Population D : qui couvre toute la région étudiée.

1.2. Hypothèses et conventions.

Pour mener à bien notre étude, nous avons introduit les hypothèses et convention suivantes :

1) Il n'y a pas de sens dans la direction des mesures; les distances ne peuvent donc pas être négatives. Cette convention a été introduite pour tenir compte du fait que certains filons se recoupent ou coïncident sur une certaine longueur.

2) Deux filons qui coïncident en surface restent néanmoins distincts. Une conséquence importante de cette convention est que la distance nulle peut exister. Nous discuterons les conséquences de cette convention au paragraphe 2.2.

3) Les filons de direction sensiblement EW ne font pas partie de la population considérée; nous excluons donc les distances infinies.

Remarque : trois filons EW ont été reconus; deux d'entre eux sont situés dans la zone marginale et un seul dans la région principale.

4) Un filon ne peut être encadré au maximum que par deux filons voisins. Cette hypothèse, purement formelle, est nécessaire étant donné la deuxième convention.

5) Il est indispensable de définir dès le début la nature de la variable étudiée. Faut-il considérer la distance entre filons consécutifs comme une variable aléatoire ou comme une variable régionalisée, telle qu'elle est définie par G. MATHERON [1962, p. 19-22; 1969, p. 5-9]? Ce choix est fondamental; il impose la suite du raisonnement et conditionne l'interprétation des résultats obtenus. L'estimation des gisements miniers, par exemple, ou la méthode d'échantillonnage des minerais (P. GY, 1967) sont fortement influencées par ce choix.

Nous avons admis que notre variable est une variable aléatoire. Dans les paragraphes 2 à 5 nous allons en plus admettre l'indépendance des mesures et nous étudierons la loi de répartition des distances dans les 4 populations. Dans la sixième partie, et compte tenu des conclusions de l'étude précédente, nous examinerons les corrélations éventuelles, en admettant que l'on a affaire à un processus stationnaire.

6) Nous admettons, enfin, que les distances mesurées dans chacune des 4 populations constituent des échantillons suffisamment représentatifs de ces populations.

2. Résultats des mesures

2.1. Courbes de distribution.

Nous avons tracé à la figure 1 la courbe de répartition de la population D et indiqué les points expérimentaux des courbes de distribution des 3 autres populations.

La forme asymétrique des distributions suggère, en se limitant à des fonctions simples, une loi de répartition lognormale ou gamma incomplète.

2.2. Le problème des zéros.

L'existence d'une fréquence non nulle pour la valeur $x = 0$ pose une question délicate. En

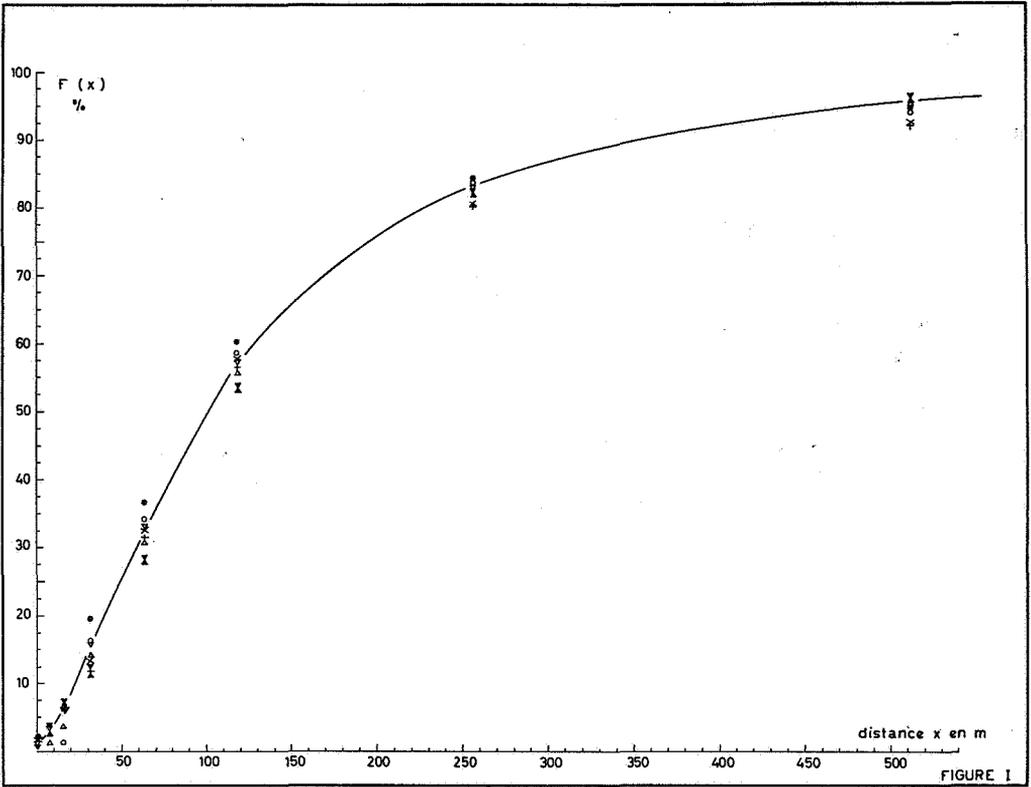


Fig. 1. Courbe de distribution dans la population D.

On tient compte des distances nulles

×
●
▼
▽

Population A
» B
» C
» D

On ne tient pas compte des distances nulles

+
○
▲
△

effet, lorsqu'on envisage des lois de répartition qui, telles les distributions lognormale ou gamma, sont définies pour des valeurs positives de la variable, la probabilité de trouver une valeur ≤ 0 est nulle et il y a donc en ce point une différence fondamentale entre le phénomène observé et le modèle mathématique choisi pour le représenter. Devant une telle situation, relativement fréquente en prospection par exemple, deux attitudes sont possibles:

a) Négliger les valeurs $x = 0$, ce qui, en dernière analyse, revient à admettre qu'elles ne font pas partie de la population étudiée. Ainsi, négliger les teneurs nulles lorsqu'on étudie la

distribution d'une minéralisation revient d'une façon ou d'une autre, à admettre, qu'elles sont étrangères au phénomène « minéralisation » étudié.

b) Rechercher un schéma ou introduire un artifice de calcul, qui permet de tenir compte, même partiellement, de ces observations; c'est l'étude de « l'effet zéro » envisagée, d'une manière d'ailleurs assez superficielle, par G. MATHERON dans son « Traité de géostatistique appliquée » (1962, p. 290-292).

Dans la suite de nos calculs, nous avons, dans la mesure du possible, tenu compte des valeurs nulles. Ainsi, en étudiant la spécification lognormale, nous avons appliqué, une

première fois le test du χ^2 en négligeant ces valeurs et ensuite, par un artifice de calcul, auquel nous n'attribuons d'ailleurs pas une grande valeur, nous avons essayé de prendre en considération ces mesures.

2.3. Calcul des moments des distributions observées.

Nous avons adopté les notations suivantes:

x = distance en m

y = $\log x$ (*)

N = effectif de l'échantillon en tenant compte des valeurs $x = 0$

n = effectif de l'échantillon en négligeant ces valeurs.

μ'_α = moment d'ordre α

μ_α = moment centré d'ordre α

$\bar{x} = \mu'_1(x) = \frac{\sum x}{N}$ et

$\bar{y} = \mu'_1(y) = \frac{\sum y}{n}$ moyennes

$\sigma^2 = \mu^2 =$ variance $\sigma =$ écart-type

Nous avons calculé directement, à partir des données numériques, les valeurs de μ'_1, σ^2 et

* Sauf indication contraire, les logarithmes utilisés sont les logarithmes naturels.

μ'_3 et avons déduit les valeurs de μ'_2 et μ_3 des formules [2, pg. 12]:

$$\mu'_2 = \mu_2 + \mu_1'^2 \quad \text{et}$$

$$\mu_3 = \mu'_3 - 3\mu_1'\mu_2' + 2\mu_1'^3$$

Le tableau A résume les résultats obtenus.

2.4. Estimation des moments des distributions vraies.

On peut estimer, compte tenu de notre hypothèse d'indépendance des valeurs de x , les moments des distributions vraies à l'aide des formules [R.M. BECKER, 1964, p. 28; H. CRAMER, 1966, p. 352; C. MARK, 1967, p. 36; R. MURARD, 1960, p. 35]:

$$\tilde{x} = \bar{x}$$

$$\mu_2^* = \frac{N}{N-1} \mu_2 \quad \text{et} \quad \mu_3^* = \frac{N^2 \cdot \mu_3}{(N-1)(N-2)}$$

3. Spécification lognormale.

3.1. Tests utilisés et résultats.

Il existe de nombreux tests permettant de vérifier la correspondance entre une loi de référence normale ou lognormale et la distribution statistique observée. Nous avons choisi

TABLEAU A. Population

	A	B	C	D
\bar{x}	163	143	171	157
$\mu_2(x)$	25720	21265	32870	27420
$\sigma(x)$	160	146	181	166
$\mu_2'(x)$	52345	41595	62205	52165
$\mu_3'(x) \cdot 10^{-6}$	23,95	17,83	37,51	27,92
$\mu_3(x) \cdot 10^{-6}$	7,015	5,835	10,57	7,195
\bar{y}	4,69	4,56	4,67	4,62
$\mu_2(y)$	0,876	0,941	1,140	1,047
$\sigma(y)$	0,936	0,970	1,068	1,023

TABLEAU B. *Population.*

	A	B	C	
\bar{x}	163	143	171	157
$\mu^*_2(x)$	26225	21545	33284	27597
$\sigma^*(x)$	162	147	182	166
$\mu^*_3(x).10^{-6}$	7,44	6,07	10,98	7,34
\tilde{y}	4,69	4,56	4,67	4,62
$\mu^*_2(y)$	0,894	0,951	1,154	1,053
$\sigma^*(y)$	0,945	0,977	1,074	1,025

deux procédés graphiques et un test numérique. Sur les figures 2 et 3 nous avons représenté l'anamorphose de Galton et le test de la droite de Henry; les distributions observées semblent coïncider assez sensiblement avec une loi lognormale.

Le test du χ^2 a été appliqué de deux manières différentes:

Dans une première étape, nous avons négligé les mesures $x = 0$. Dans un deuxième stade, nous avons essayé de tenir compte de ces mesures. Pour cela, nous avons admis que la distribution des y conserve sa moyenne et son écart-type. Cette convention est une simple hypothèse de travail; elle revient à admettre que, pour le calcul de y et $\sigma(y)$, on choisit comme valeur approchée de $x = 0$ $x = G_x =$ moyenne géométrique (cf. tableau E).

Les résultats obtenus sont indiqués au tableau C.

La spécification lognormale peut donc être retenue.

3.2. Limites de confiance de l'estimation des moments des distributions vraies

La moyenne μ'_1 obéit [C. MARC, 1967, p. 65; R. MURARD, 1960, p. 60-62] à une loi normale de moyenne μ'_1 et d'écart-type $\frac{\sigma^*}{\sqrt{N}}$. L'intervalle de confiance est donc $\mu'_1 \pm n \frac{\sigma^*}{\sqrt{N}}$ où n représente la variable normale réduite.

Remarques: 1) La formule de SICHEL [G. MATHERON, 1962, p. 272-273; R. MURARD, 1960, p. 47] permet d'estimer d'une façon plus précise l'intervalle de confiance de la moyenne x d'une variable lognormale. Nous avons appliqué cette formule aux 4 populations étudiées; les résultats diffèrent très peu de ceux indiqués au tableau D ci-dessous.

2) La population y étant normale, on peut également définir l'intervalle de confiance de \tilde{y} par $\tilde{y} \pm t \frac{\sigma(y)}{\sqrt{n-1}}$ avec $t_v =$ variable de

Student-Fisher. Cette formule fournit dans les 4 cas des estimations pratiquement identiques à celles figurant au tableau D.

L'écart-type $\sigma^*(y)$ varie selon une loi normale de moyenne $\sigma^*(y)$ et d'écart-type

$$\sigma(\sigma(y)) = \frac{\sigma(y)}{\sqrt{2n}}$$

On peut donc dresser le tableau suivant, indiquant les limites de confiance des différents moments.

3.4. Autres paramètres caractéristiques.

Les données précédentes permettent de déterminer d'autres paramètres caractéristiques, tant des distributions observées que des distributions lognormales théoriques correspondantes.

Nous avons ainsi calculé les 3 quartiles X_1 ,

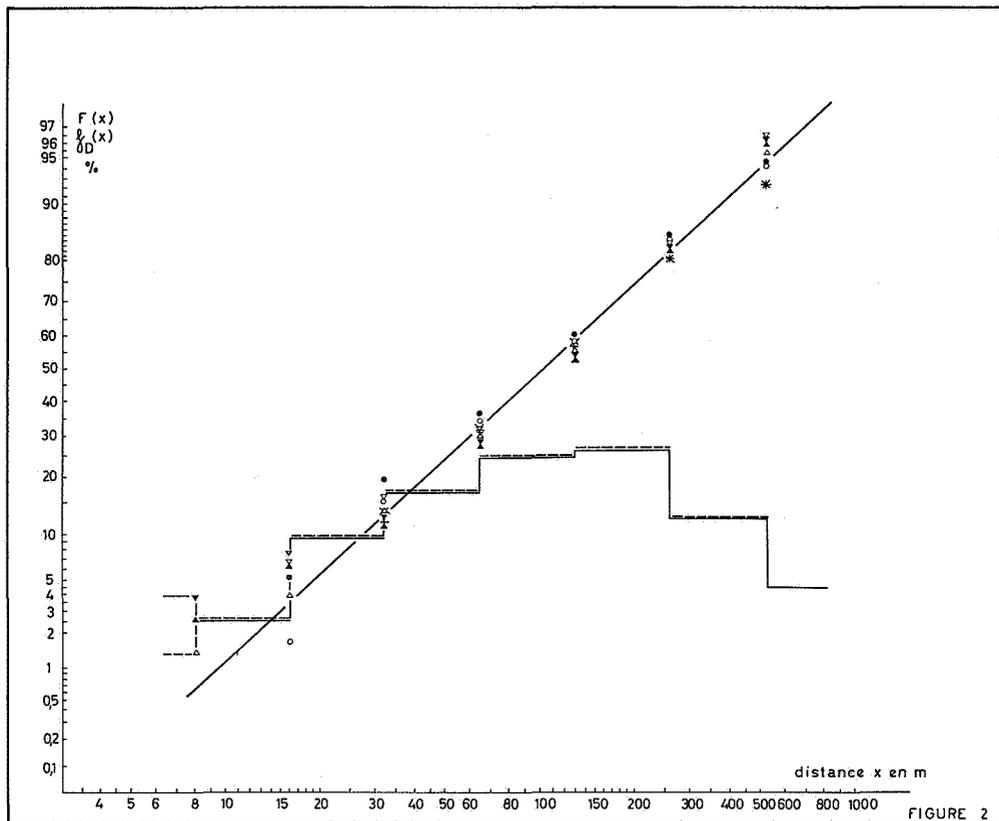


Fig. 2. Anamorphose de Galton. Droite de distribution dans la population D.

On tient compte des distances nulles

- ×
-
- ▼
- ▽

Population A
 » B
 » C
 » D

On ne tient pas compte des distances nulles

- +
-
- ▲
- △

Histogramme de fréquence. Population D.

X_2 , X_3 , le mode $Mx = X_2 e^{-\sigma^2}$ et l'écart probable moyen $Epm = \frac{X_3 - X_1}{2}$

N.B. La médiane X_2 est aussi égale à la moyenne géométrique $G_x = e^{\bar{y}}$; on peut donc estimer éventuellement l'intervalle de confiance de ce paramètre.

4. Specification gamma

Les quatre distributions considérées sont très

proches de la loi lognormale et nous n'avons dès lors pas jugé utile d'étudier en détails la spécification gamma incomplète. Nous nous sommes limités à calculer les valeurs des deux membres de l'égalité $\mu_3^*(x) \cdot \mu_1'(x) = 2\mu_2^{*2}(x)$.

La différence entre ces deux valeurs donne une idée de l'écart à la loi gamma, mais il est bien sûr impossible de conclure que cet écart est significatif ou non.

Nous avons profité de cette étude pour calculer les paramètres de cette loi de référence.

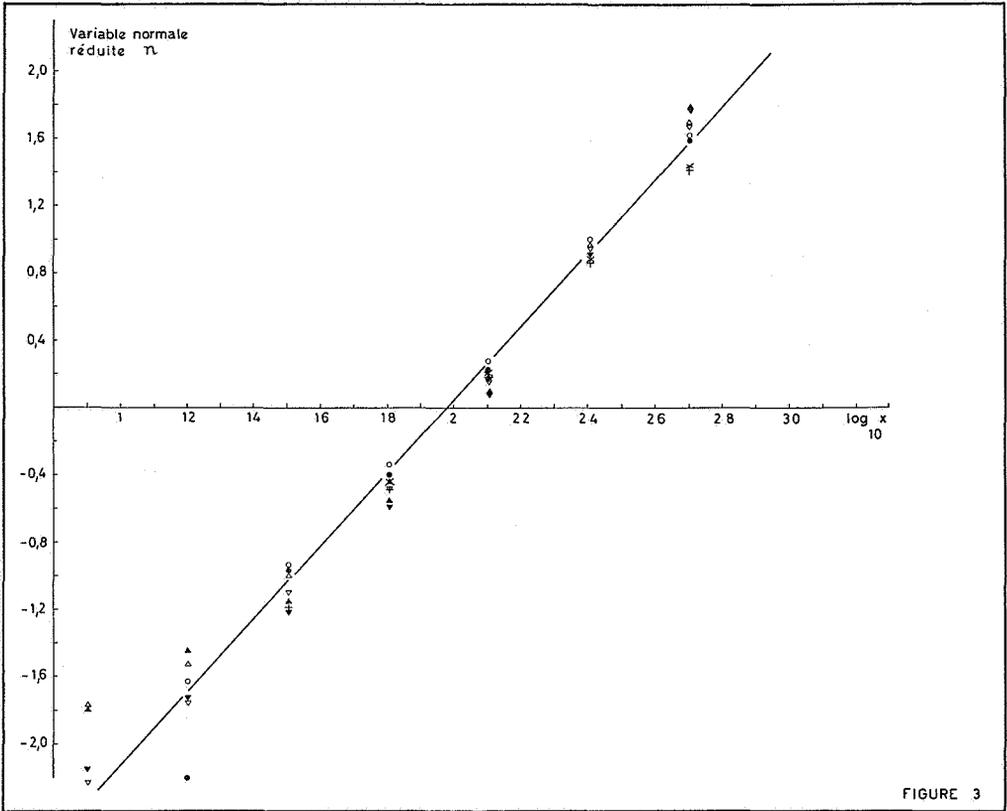


FIGURE 3

Fig. 3. Droite de Henry.

On tient compte des distances nulles

- ×
-
- ▼
- ▽

Population A

- » B
- » C
- » D

On ne tient pas compte des distances nulles

- +
-
- ▲
- △

Si on adopte la forme $f(x) = \frac{1}{\Gamma(\rho + 1)}$

$$\left(\frac{x}{b}\right)^{p-\frac{x}{b}} \frac{1}{b}$$

$$\text{on a } b = \frac{\mu_2^*(x)}{\mu_1(x)} \text{ et } p = \frac{\mu_1'^2(x)}{\mu_2^*(x)} - 1$$

Connaissant p et b , on peut calculer $\mu_3(x) = 2b^3(p+1)$ et la comparaison de cette valeur et de l'estimation μ_3^* donne également une idée de l'écart entre la loi gamma de Pearson et les distributions observées.

Il semble donc que la spécification Γ doit être rejetée au profit d'une loi lognormale.

Toutefois, si on adoptait une loi de référence gamma, celle-ci se réduirait à une exponentielle.

5. Identité statistique des populations.

Nous avons testé l'identité statistique des 4 populations en deux temps:

1) Dans la première étape, nous avons comparé chacune des trois premières populations normales y à la quatrième, qui, rappelons-le, est la somme des populations B et C [C. MARC, 1967, p. 46-50; R. MURARD, 1960, p. 74-77].

TABLEAU C. Population.

	A	B	C	D
Nombre de degrés de liberté	3	4	4	4
Valeur de κ^2 au seuil 5 % au seuil 10 %	7,815 6,261	9,488 7,779	9,488 7,779	9,488 7,779
On néglige les valeurs $x = 0$				
n κ^2	51 0,950	73 3,924	79 5,407	152 2,250
On tient compte des valeurs $x = 0$				
N κ^2	52 1,526	76 3,779	80 7,273	156 5,754

TABLEAU D. Population.

	A	B	C	D
$\sigma(\bar{x})$	22,5	16,8	20,4	13,3
Moyenne \bar{x} au seuil 80 % au seuil 95 %	163 ± 29 163 ± 44	143 ± 22 143 ± 33	171 ± 26 171 ± 40	153 ± 17 153 ± 26
$\kappa(\bar{y})$	0,132	0,143	0,121	0,083
Moyenne \bar{y} au seuil 80 % au seuil 95 %	4,69 ± 0,17 4,69 ± 0,26	4,56 ± 0,18 4,56 ± 0,22	4,67 ± 0,15 4,67 ± 0,24	4,62 ± 0,11 4,62 ± 0,16
$\sigma(\sigma(y))$	0,0935	0,081	0,0855	0,059
Ecart-type $\sigma^*(y)$ au seuil 80 % au seuil 95 %	0,95 ± 0,12 0,95 ± 0,18	0,98 ± 0,10 0,98 ± 0,16	1,07 ± 0,11 1,07 ± 0,17	1,03 ± 0,08 1,03 ± 0,12

TABLEAU E. *Population.*

	A	B	C	D
~ Distribution observée				
X ₁	50	41	58	49
X ₂	107	95	115	104
X ₃	208	168	215	205
Epm	79	72,5	78,5	78
~ Distribution log N théorique correspondante				
X ₁	56,5	49,5	52	50,5
X ₂	109	95,5	107	101
X ₃	210	184	220	201
Epm	76,8	67,3	84	75,3
M _x	45	36,8	33,7	35,5

TABLEAU F. *Population.*

	A	B	C	D
$\mu^*_3(x) \cdot 10^{-6}$	7,44	6,07	10,98	7,34
$\mu^*_3\mu'_1 \cdot 10^{-6}$	1214	865,8	1880,2	1157,2
$2\mu^*_2 \cdot 10^{-6}$	1375	929,5	2215,1	1532,2
p	0	- 0,056	- 0,120	- 0,103
b	161	151,1	194,3	175,4
$\mu_3(x) \cdot 10^{-6}$	8,345	6,51	12,91	9,68

La différence $y_i - y_D$ ($i = A, B, C$) suit une loi normale de moyenne nulle et dont la variance peut être estimée par

$$\text{Var}_{iD} = \frac{n_i \mu_{2i}(y) + n_D \mu_{2D}(y)}{n_i + n_D - 2}$$

Donc la variable de Student-Fisher

$$t_v = \frac{|\bar{y}_i - \bar{y}_D|}{\sqrt{\frac{\text{Var}_{iD}}{n_i} + \frac{\text{Var}_{iD}}{n_D}}} \text{ correspond à}$$

$v = n_i + n_D - 2$ degrés de liberté.

Pour comparer les variances, nous avons

considéré le rapport de Snédécour $F = \frac{\mu_{2i}^*(y)}{\mu_{2D}^*(y)}$

à $v_i = n_i - 1$ et $v_D = n_D - 1$ degrés de liberté.

On peut également considérer la différence $\sigma_i(y) - \sigma_D(y)$ qui suit une loi normale de moyenne nulle et de variance $\text{Var}(\Delta\sigma) =$

$$= \frac{\mu_{2i}^*(y)}{2n_i} + \frac{\mu_{2D}^*(y)}{2n_D}$$

Les résultats de cette étude sont reproduits au tableau G.

TABLEAU G. Population.

	A	B	C
Comparaison des moyennes			
v	201	223	229
t_{∞} au seuil 5 %	1,96	1,96	1,96
t_{∞} au seuil 10 %	1,645	1,645	1,645
t_v observé	0,449	0,408	0,375
Comparaison des variances			
v_i	50	72	78
Valeur de F pour $v_D = \infty$			
au seuil 10 %	1,34	1,27	1,22
au seuil 5 %	1,47	1,36	1,29
F observé	1,17	1,10	1,10
$\sigma (\Delta\sigma)$	0,11	0,10	0,104
$\Delta\sigma$ observé	- 0,087	- 0,053	0,044

On peut donc admettre l'identité statistique des populations comparées.

2) L'identité de chacune des 3 premières populations à la quatrième n'implique nécessairement pas l'identité de ces populations entre elles. Pour tester cette hypothèse, nous avons effectué une analyse de variance (P. GY, 1967, p. 97-102; C. MARC, 1967, p. 58-60; R. MURARD, 1960, p. 78-80):

La variance totale $\text{Var}_T = 1,014$

La variance entre groupes $\text{Var}_e = 0,378$

La variance résiduelle $\text{Var}_r = 1,012$

Le rapport de Snédécour vaut $F = \frac{\text{Var}_r}{\text{Var}_e} = 2,68$; le nombre de degrés de liberté $v_r = 202$ et $v_e = 2$. Or, même pour $v_e = \infty$ $F = 19,5$ au seuil 5 % et $F = 9,5$ au seuil 10 %. On peut donc admettre l'identité des 4 populations.

6. Etude de l'autocorrélation

6.1. Test d'hypothèse

Nous avons jusqu'à présent admis que les distances successives étaient indépendantes

entre elles. Examinons la validité de cette hypothèse pour les variables normales y . A titre d'exemple, nous avons tracé à la figure 4 la variation de $y - \tilde{y}_D$ en fonction du rang de la mesure. Peut-on admettre qu'une telle série est de « bruit blanc »?

Rappelons que nous avons admis avoir affaire à un processus stationnaire; l'identité statistique des 4 populations peut être considérée comme un argument en faveur de cette hypothèse. Appelons Y_i la variable aléatoire, dont la mesure y_i est un échantillon.

La fonction d'autovariance est définie par $\gamma(h) = E [Y_i - \mu'_1(y)] [Y_{i+h} - \mu'_1(y)]$ la fonction d'autocorrélation par $\rho(h) = \frac{\gamma(h)}{\sigma^2}$ et la

fonction de densité spectrale $f(\omega)$ [D.R. COX and P.A.W. LEWIS, 1966, p. 70] est telle que

$$\gamma(h) = \int_{-\pi}^{+\pi} \cos h\omega \cdot f(\omega) d\omega = 2 \int_0^{\pi} \cos h\omega f(\omega) d\omega$$

On peut estimer [D.R. COX and P.A.W. LEWIS, 1966, p.89 -90; la moyenne

$$\mu'_1(Y) \text{ par } \bar{y} = \frac{\sum y_i}{n} \text{ avec une variance } \frac{2\pi}{n}$$

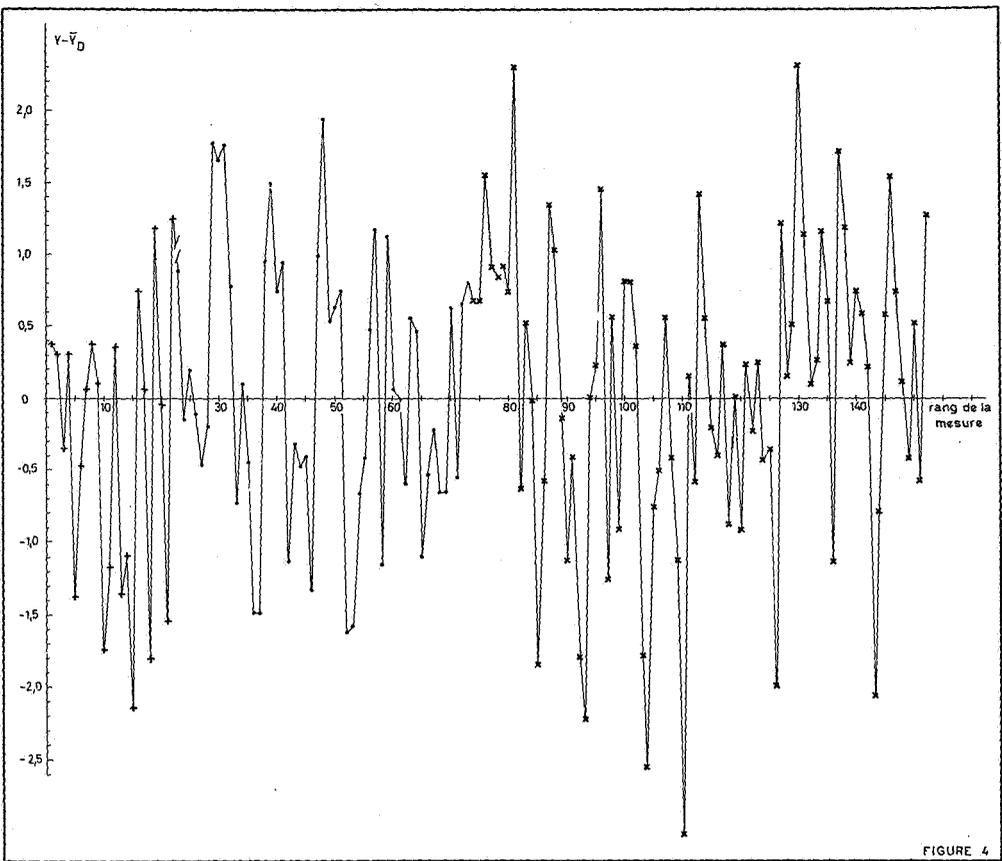


Fig. 4. Population D. Variation de $y - \bar{y}$ dans la direction WE.

+ zone ouest ● zone est × région marginale.

$f(0)$ et $\gamma(h)$ par $C(h) = \frac{1}{n-h} \sum_{i=1}^{n-h}$

$(y_i - \bar{y})(y_{i+h} - \bar{y})$. Toutefois, cette dernière estimation est biaisée puisque

$$E(C(h)) = \gamma(h) - \frac{2\pi}{n} f(0) + \frac{R}{n}$$

Si le dernier terme de cette relation est généralement petit lorsque n est grand, il n'en est pas nécessairement de même du deuxième.

Nous avons tracé à la figure 5 la variation de l'estimation biaisée $\frac{C(h)}{\sigma^2}$ de la fonction l'autocorrélation en fonction du retard h pour les quatre populations considérées. On constate qu'il existe probablement une corrélation positive pour $h = 1$. Dans la zone est de la

région principale, il existe également une corrélation négative pour $h = 4$ et 5 . Cette dernière constatation est peu importante pour nous, étant donné la position géographique de cette zone et les valeurs élevées de ces deux retards.

Nous examinerons donc si l'autocorrélation pour $h = 1$ est significativement différente de zéro. Pour cela, nous avons calculé le coefficient proposé par R.L. ANDERSON (1942)

$$R = \frac{\sum_{i=1}^n (y_i - \bar{y})(y_{i+1} - \bar{y})}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

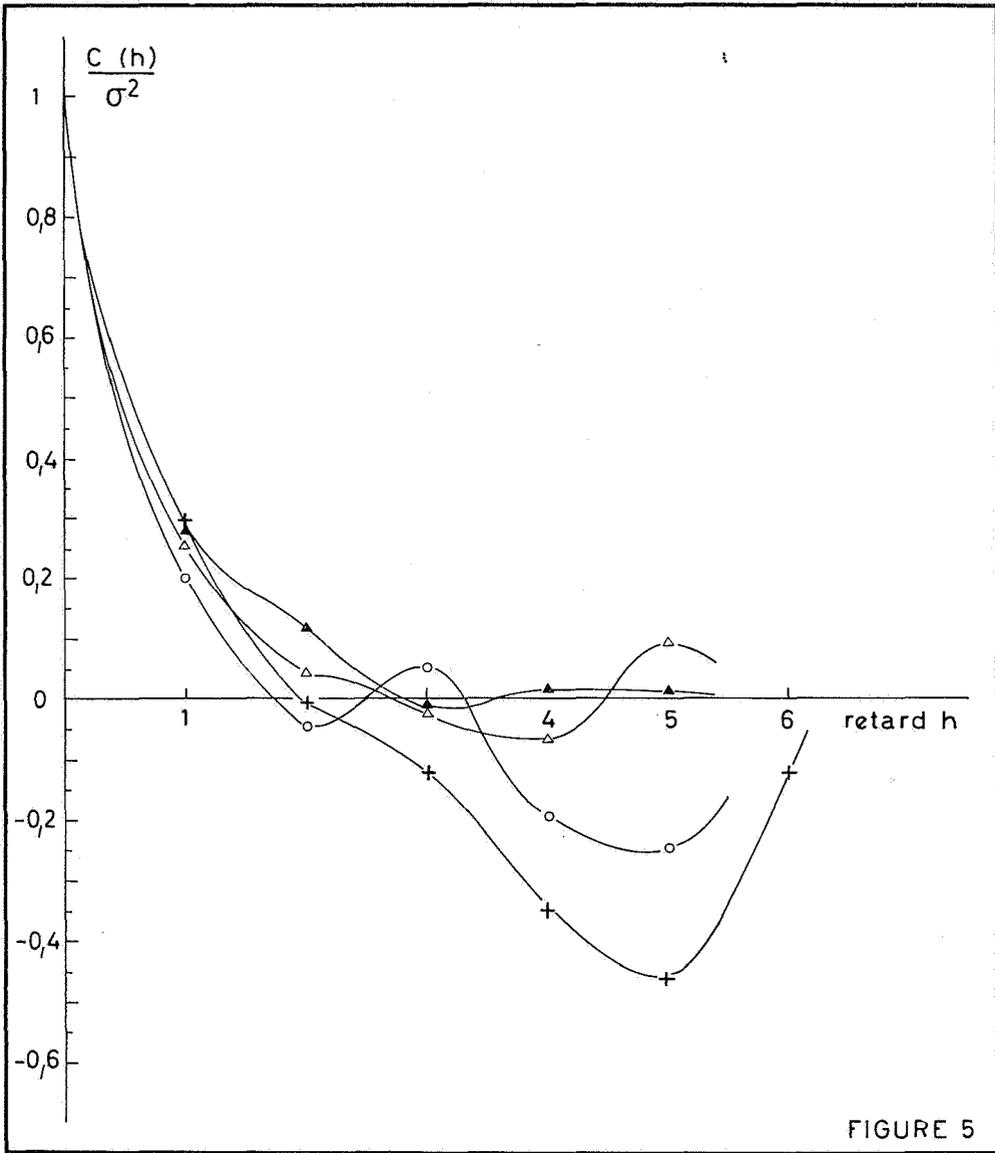


FIGURE 5

Fig. 5. Estimation (biaisée) $\frac{C(h)}{\sigma^2}$ de la fonction d'autocorrélation.

+ population A
o population B

▲ Population C
△ population D

Remarque : la théorie d'ANDERSON ne s'applique rigoureusement qu'à une série circulaire. Or, les séries considérées ne sont pas circulaires et le problème se complique encore dans le cas des populations B et D à cause de la présence de la zone centrale inconnue.

Toutefois, compte tenu de la valeur élevée de n , on peut sans erreur grave [VAN DE WIELE, communication personnelle] assimiler nos séries à des séries circulaires dans lesquelles $y_{n+1} = y_1$.

On peut donc retenir l'hypothèse d'auto-corrélation non nulle pour un retard $h = 1$.

6.2. Estimation de la distance suivante.

Nous venons de constater qu'il existe une autocorrélation non nulle pour le retard $h = 1$ et avons négligé la corrélation pour les retards 4 et 5 dans la zone est.

Dans ces conditions, nous avons choisi pour représenter en première approximation les séquences $\{Y_i\}$ le processus autorégressif du premier ordre. C'est un processus markovien caractérisé par la relation

$$y_i - \mu'_1 = \lambda (y_{i-1} - \mu'_1) + Z_i$$

dans laquelle Z_i sont des variables non corrélées de moyenne zéro et λ une constante telle que $|\lambda| < 1$.

Les estimateurs des moments et les fonctions caractéristiques définies au paragraphe précédent sont donnés par

$$f(\omega) = \frac{1 - \lambda^2}{2\pi(1 - 2\lambda \cos \omega + \lambda^2)} \sigma^2 \text{ et en}$$

$$\text{particulier } f(0) = \frac{1 + \lambda}{2\pi(1 - \lambda)} \sigma^2$$

$$\tilde{Y} = \bar{y} \quad \sigma^2 \simeq \frac{1}{n} \sum_i (y_i - \bar{y})^2 + \frac{2\pi}{n} f(0)$$

$$\gamma(h) \simeq \frac{1}{n-h} \sum_i (y_i - \bar{y})(y_{i+h} - \bar{y}) +$$

$$\frac{2\pi}{n} f(0) \quad \gamma^*(h) = \sigma^{*2} \cdot \rho^*(h) \text{ et}$$

$$\lambda^* = \rho^*(1)$$

Ces formules permettent de calculer par approximations successives, les fonctions et paramètres caractéristiques (tableau I).

Nous avons tracé à la figure 6 la fonction $\rho^*(h)$ et la fonction théorique $\rho(h) = \lambda^{|h|}$ correspondant au processus adopté. On constate que la coïncidence est très bonne pour les séries C et D et, comme on pouvait s'y attendre assez mauvaise pour $h = 4$ et 5 dans la

TABLEAU H. Population.

	A	B	C	D
R calculé au seuil 5 %	0,207	0,175	0,172	0,127
R observé	0,323	0,212	0,259	0,248

TABLEAU I. Population.

	A	B	C	D
λ^*	0,226	0,34	0,271	0,25
σ^{*2}	0,961	0,909	1,167	1,059
$\gamma^*(1)$	0,217	0,310	0,318	0,264
$f(\omega)$	4	4,5	5,5	5
	$\pi(11,5-6,8 \cos \omega)$	$\pi(10,5-4,5 \cos \omega)$	$\pi(10,5-5,5 \cos \omega)$	$\pi(10,5-5 \cos \omega)$

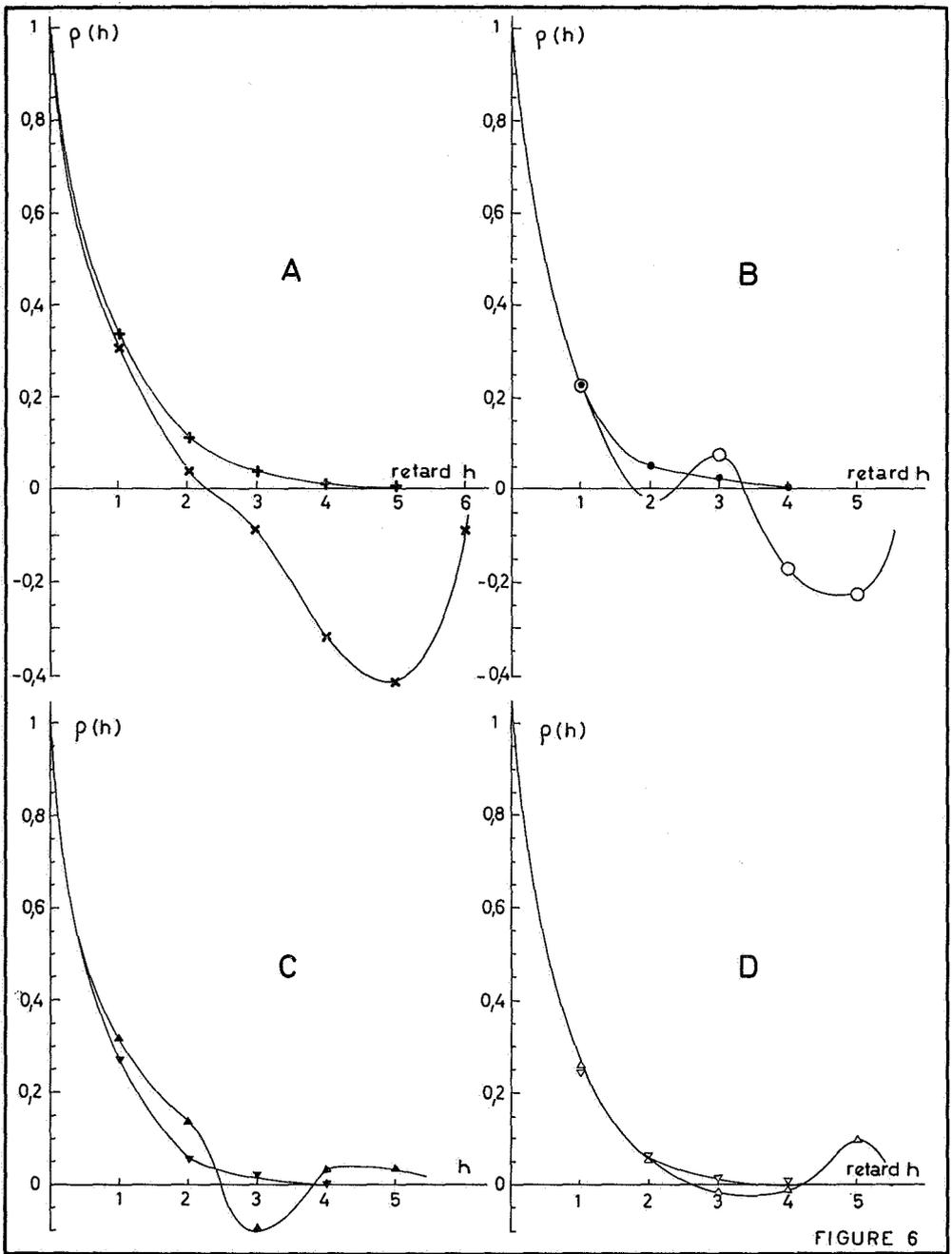


Fig. 6. Fonction d'autocorrélation $\rho(h)$.

Fonction calculée

+

●

▼

▽

Population A

» B

» C

» D

Fonction observée

×

o

▲

△

zone est (population A). L'influence de la zone est se fait encore sentir dans la série B.

On peut donc dans une large mesure adopter l'hypothèse d'un processus autocorrélatif du premier ordre et dès lors on peut estimer la $(n+k)^{\text{e}}$ distance à l'aide de $y_{n+k}^* = \bar{y} + \lambda^k (y_n - \bar{y})$.

En pratique, on choisira $y_{n+1}^* = \bar{y} + \lambda^* (y_n - \bar{y})$ avec une variance $\text{Var} (y_{n+1}^*) = (1 - \lambda^2) \sigma^{*2}$. Le tableau J indique cette estimation dans chacune des 4 régions.

7. Conclusions

Nous avons constaté que les distances entre filons sont distribuées selon une loi lognormale et que l'hypothèse d'autocorrélation pouvait être retenue. Toutefois, cette hypothèse ne modifie souvent que légèrement la prévision de la $(n+1)^{\text{e}}$ distance. Ainsi, si on considère que la n^{e} distance aboutit au dernier filon connu à l'est de la région marginale, le logarithme de la distance au filon suivant peut être calculé, en négligeant les erreurs d'estimation, par

TABLEAU J.

Population	y_{n+1}^*	$\sigma^* (y_{n+2}^*)$
A	$3,52 + 0,25 y_n$	0,94
B	$3,42 + 0,25 y_n$	0,89
C	$3,50 + 0,25 y_n$	1,05
D	$3,47 + 0,25 y_n$	1,01

	Hypothèse d'autocorrélation nulle	Hypothèse d'un processus autorégressif
Population C	$y_{n+1} = 4,67 \pm n_{\alpha} 1,07$	$y_{n+1} = 4,81 \pm n_{\alpha} 1,05$
Population D	$y_{n+1} = 4,62 \pm n_{\alpha} 1,03$	$y_{n+1} = 4,78 \pm n_{\alpha} 1,01$

avec n_{α} = variable normale réduite.

REFERENCES

- ANDERSON, R.L. (1942). Distribution of the serial correlation coefficient (*Annals of Mathem. statistics*, vol. 13, 1-13).
- BECKER, R.M. (1964). Some generalised probability distributions with special reference to the mineral industries, Part 1: sampling to n items. (*U.S.B.M. R.I. 6329*).
- CRAMER, H. (1966). *Mathematical methods of statistics* (Princeton Univ. Press).
- COX, D.R. and LEWIS, P.A.W. (1966). *The statistical analysis of series of events*. (J. WILEY).
- DEFALQUE, G., DUMONT, P. et PANOU, G. (1971). La géologie du massif de Los Pedroches (note préliminaire) (*Soc. Belge de Géol.*).
- GY, P. (1967). L'échantillonnage des minerais en vrac. (*Mémoires BRGM*, n° 56).
- MARC, C. (1967). *Essentials of statistics for scientists and technologists*. (Plenum press).

MATHERON, G. (1962). Traité de géostatistique appliquée (t1). (*Mémoires BRGM*, n° 14).

MATHERON, G. (1967). Les variables régionalisées et leur estimation (MASSON).

MURARD, R. (1960). Probabilités et statistique (*R.I.M. Doc. SIMa*).

Communication présentée le 18 mai 1971.