G. Matheron (1961)

Note Géostatistique n° 32

- - - - - - - - - - - - - - -

Problème de la maille optimum de reconnaissance stratégique.

Le but d'une reconnaissance stratégique, ou régionale, n'est pas de trouver la totalité des indices ou anomalies existant rèellement dans la région prospectée, mais de délimiter des zones jugées dignes d'une étude plus détaillée. La maille optimum pour une telle reconnaissance ne doit pas être prise trop serrée, pour des raisons évidentes de prix de revient. Il convient de lui attribuer la dimension la plus grande possible capable de garantir la quasi certitude de ne pas manquer une zone intéressante. Une telle définition est très subjective. Tout d'abord, ce que l'on entend par quasi-certitude est matière à appréciation personnelle. On peut toutefois convenir d'un seuil de probabilité ; c'est-à-dire convenir, par exemple, que la maille à adopter est celle pour laquelle il y a une chance sur 10, ou sur cent etc... de manquer une zone intéressante. Mais surtout la notion même de zone intéressante est difficile à préciser. Combien doit-elle contenir d'indices, et de quelles dimensions pour être jugée digne d'intérêt ? Il est plus réaliste d'aborder le problème en sens inverse, c'est-à-dire, étant donnée une zone contenant des indices en nombre et en dimension supposés connus, de chercher à déterminer le rendement d'une maille de reconnaissance donnée.

Ainsi formulé, le problème revient à calculer la probabilité pour qu'un indice de caractéristiques données soit décelé. En effet, les différents indices ne sont certainement pas répartis au hasard, c'est-à-dire indépendamment les uns des autres, mais d'une façon régionalisée. Mais les distances entre anomalies voisines - à partir du moment où elles sont relativement grandes vis à vis de la maille - n'exercent pas d'influence sur le rendement de la maille. Ce ne sont pas les positions absolues dans l'espace des anomalies qui importent, mais les positions de leurs centres relativement aux noeuds les plus proches du réseau de reconnaissance : la position d'une anomalie vis à vis des 4 noeuds du réseau qui l'encadrent est indépendante de la position de toute anomalie voisine vis à vis des quatre noeuds qui la contiennent. De la même façon, les valeurs numériques des logarithmes des entiers successifs ne sont pas indépendantes, mais leurs quatoreièmes décimales le sont. D'où trois parties pour cette note :

1° - détermination de la probabilité pour qu'un anomalie de caractéristiques données soit décelée par une maille donnée.

2° - Effet d'une maille donnée sur une population d'anomalies de caractéristiques données.

3° - Problème inverse : une maille ayant décelé tant d'anomalies de telles caractéristiques, que peut-on en déduire relativement au nombre et aux caractéristiques de toutes les anomalies existant rèellement.

I. Probabilité de toucher une anomalie donnée.

Cette probabilité dépend de la forme et de la dimension de l'anomalie, de la forme et de la dimension de la maille de reconnaissance. Nous traiterons trois cas :

a/ Reconnaissance par profils continus équidistants (ex : prospection aérienne). Seule intervient la dimension a de l'anomalie perpendiculairement au profil. Un seul paramètre : le rapport a/A de la dimension de l'anomalie à l'espacement des profils.

b/ Reconnaissance par mesure ponctuelle à maille rectangulaire A x B, anomalies rectangulaires de côté a et b parallèles aux directions principales de la maille. Deux paramètres : a/A et b/B.

c/ Reconnaissance par mesures ponctuelles à maille rectangulaire A x B, anomalies elliptiques de demi axes a et b parallèles aux directions principales de la maille : 2 paramètres a/A et b/B.

A/ Profils continus équidistants.

Soient A l'équidistance des profils et a la dimension de l'anomalie perpendiculaire aux profils. La probabilité P de déceler cette anomalie a pour valeur : (voir Fig 1)

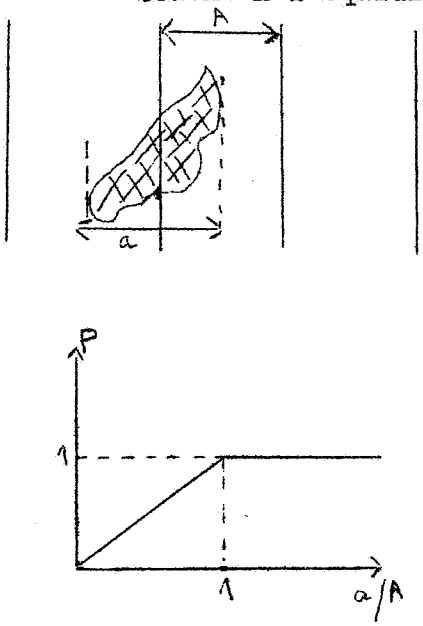


Fig. 1.

$$(1)\quad \begin{cases} P = \dfrac{a}{A} & \text{pour } a < A \\ P = 1 & \text{pour } a \geqslant A \end{cases}$$

Il en résulte immédiatement (comme le suggère le bon sens) que le meilleur rendement de reconnaissance est obtenu pour des profils perpendiculaires à la plus grande dimension de l'anomalie.

B/ Anomalies rectangulaires a x b, maille rectangulaire A x B.

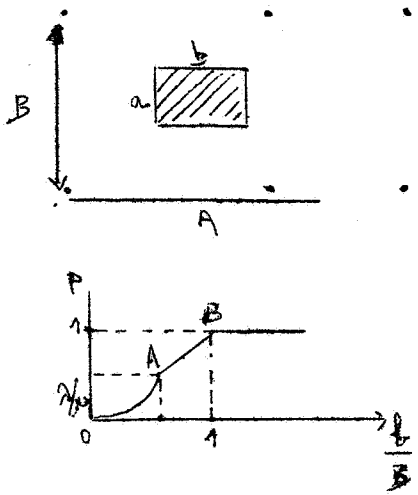


Fig 2.

Ici encore, on obtient facilement la probabilité P pour que l'anomalie soit décelée : soit :

$$(2)\quad \begin{cases} P = \dfrac{ab}{AB} & \text{pour } \begin{cases} a < A \\ b < B \end{cases} \\ P = \dfrac{b}{B} \text{ pour } \begin{cases} a \geqslant A \\ b < B \end{cases} \text{ ou } P = \dfrac{a}{A} \text{ pour } \begin{cases} a < A \\ b \geqslant B \end{cases} \\ P = 1 & \text{pour } \begin{cases} a \geqslant A \\ b \geqslant B \end{cases} \end{cases}$$

Pour les petites anomalies, c'est-à-dire celles dont les dimensions principales sont inférieures aux dimensions correspondantes de la maille, le rendement de la reconnaissance est indépendant du rapport de maille A/B.

Il ne dépend que de la surface a b de l'anomalie et du nombre $\frac{1}{AB}$ de mesures par unité de surface. Pour les petites anomalies, la formule :

$$(3)\qquad P = \frac{\text{Surface de l'anomalie}}{\text{Surface du rectangle AxB}}$$

est générale. Si l'une des dimensions de l'anomalie, par exemple a, atteint ou dépasse la dimension A correspondante de la maille, la probabilité P prend la forme b/B. Ce résultat est lié à la forme rectangulaire de l'anomalie. Pour $a > A$, la maille de reconnaissance donne le même résultat que des profils continus parallèles à A et distants de B.

Soit λ l'allongement de l'anomalie ($b = \lambda a$) et u le rapport de maille ($B = \mu A$). Si $\lambda < \mu$, par exemple, on aura $b < B$ pour $a = A$. La probabilité P a pour valeur :

$$(3)\quad \begin{cases} P = \dfrac{\lambda}{\mu}\,\dfrac{a^2}{A^2} = \dfrac{\mu}{\lambda}\,\dfrac{b^2}{B^2} & \text{pour } \dfrac{a}{A} < 1 \text{ ou } \dfrac{b}{B} < \dfrac{\lambda}{\mu} \\ P = \dfrac{\lambda a}{\mu a} = \dfrac{b}{B} & \text{pour } \dfrac{\lambda}{\mu} < \dfrac{b}{B} < 1 \\ P = 1 & \text{pour } \dfrac{b}{B} > 1 \end{cases}$$

La courbe représentative comporte l'arc de parabole O A (fig 2), la droite A B incliné à 45° et un palier horizontal. Sa dérivée présente deux discontinuités (pour $\frac{b}{B} = \frac{\lambda}{\mu}$ et $\frac{b}{B} = 1$).

Lorsque l'on est sur la droite AB, c'est-à-dire si l'on a $\frac{a}{A} > 1$ et $\frac{b}{B} < 1$, la maille n'est pas optimum. A nombre égal de mesure par unité de surface (soit à prix de revient constant), c'est-à-dire avec AB = C ^{te}, on a intérêt à diminuer B et à augmenter correlativement A jusqu'a ce que l'on ait :

- soit $\frac{a}{A} \leqslant 1$, et à ce moment $P = \frac{ab}{AB}$ ne dépend plus du rapport de maille,
- soit $\frac{b}{B} \geqslant 1$, et à ce moment P = 1 ne dépend plus de la maille.

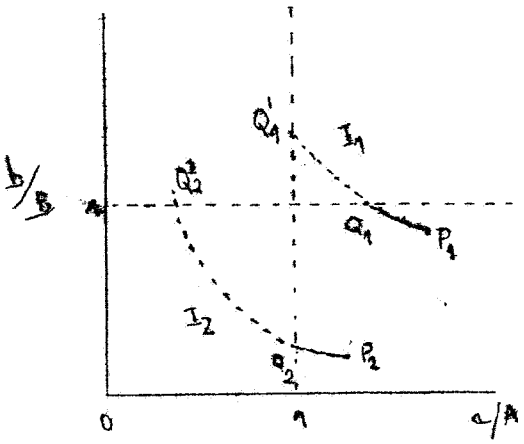


Figure 3.

Ces circonstances sont représentées sur la figure 3, où l'on a porté les rapports $\frac{a}{A}$ et $\frac{b}{B}$ en coordonnées.

Dans les quadrants I_1 et I_2, le rapport de maille est <u>indifférent.</u> Dans I_1 (grandes anomalies), on a P = 1.

Dans I_2 (petites anomalies), on a $P = \frac{ab}{AB}$ = Surface de l'anomalie X nombre de mesures par unité de surface. En dehors de I_1 et I_2, le rapport de maille est mauvais.

Si l'on part d'un point P_1, par exemple, tel que a A et b B (anomalies dont une des dimensions dépasse la maille et non l'autre), on a intérêt, à nombre de mesure 1/AB constant, à se déplacer sur l'hyperbole équilatère P, $Q_1Q'_1$ jusqu'en un point quelconque de l'arc Q_1 Q'_1 dans le quadrant I_1 (ou Q_2 Q'_2 dans le quadrant I_2 si l'on part du point P_2).

En résumé, il apparaît que pour avoir un rapport de maille A/B optimum vis à vis d'une anomalie rectangulaire donnée a/b, il suffit que A et B soient simultanément supérieurs ou inférieurs à a et b respectivement.

Il convient d'avoir :

- soit $A > a$ et $B > b$ et alors $P = \frac{ab}{AB} = \frac{\text{Surface de l'anomalie}}{\text{Surface de la maille}}$

- soit $A < a$ et $B < b$ et alors $P = 1$

Les cas intermédiaires ($A > a$ et $B < b$ par exemple) correspondent à de mauvais rapports de maille.

En pratique, celà signifie que le rapport de maille B/A doit être pris égal à l'allongement b/a des anomalies dont les dimensions sont de l'ordre de grandeur de la maille. Il n'y a pas à tenir compte des petites ou des grandes anomalies.

Ces conditions seront toujours vérifiées si l'on prend une maille telle que :

$$(4) \qquad \frac{a}{A} = \frac{b}{B}$$

mais cette condition (4) sera trop stricte pour les petites ou les grandes anomalies.

C/ Anomalie elliptique, maille rectangulaire.

Soient a et b les demi axes de l'ellipse, A et B la maille. Seuls interviennent les rapports $\frac{a}{A}$ et $\frac{b}{B}$. Il est donc loisible de transformer la figure par une affinité, de manière à remplacer l'ellipse par un cercle de rayon R = a, et la maille rectangulaire A,B par la maille rectangulaire A, $\frac{a}{b}$ B. Nous nous limitons donc au problème d'une anomalie circulaire de rayon R et d'une maille rectangulaire A X B.

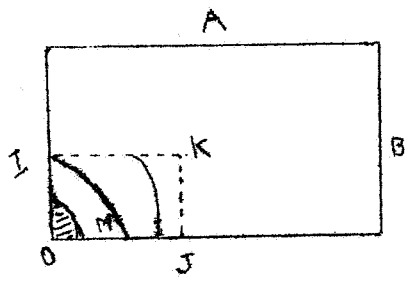


Figure 4.

Considérons un noeud O du réseau, et le rectangle O I K J de côtés A/2 et B/2 ($B < A$). La varaible aléatoire est la position du centre M de l'anomalie dans ce rectangle. Si M tombe à l'intérieur du cercle de centre O et de rayon R, l'anomalie est décelée. On a donc :

$$P = \frac{\text{Aire commune au cercle O de rayon R et au rectangle O I J K}}{\text{aire du rectangle O I J K}}$$

On trouve facilement :

$$P = \frac{\pi R^2}{A\,B} \quad \text{pour } 2R < B < A$$

$$P = \left[2 \operatorname{Arc\,Sin} \frac{B}{2R} + B \sqrt{\frac{4R^2 - B^2}{2R^2}} \right] \frac{R^2}{AB} \quad \text{pour } B < 2R < A$$

$$P = \left[2 \operatorname{Arc\,Sin} \frac{B}{2R} + B \sqrt{\frac{4R^2 - B^2}{2R^2}} - 2 \operatorname{Arc\,cos} \frac{a}{2R} + A \sqrt{\frac{4R^2 - A^2}{2R^2}} \right] \frac{R^2}{AB} \quad \text{pour } A < 2R < \sqrt{A^2 + B^2} \tag{5}$$

$$P = 1 \quad \text{pour} \quad 2R > \sqrt{A^2 + B^2}$$

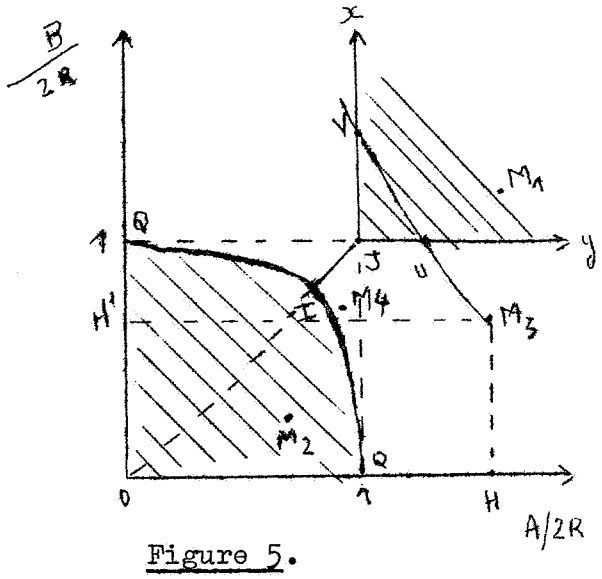


Figure 5.

La discussion se mène graphiquement comme au paragraphe précédent. Sur la figure 5, les axes de coordonnées portent $\frac{A}{2R}$ et $\frac{B}{2R}$

- Si A/2R et B/2R sont plus grands que 1, (point M_1), on a $P = \frac{\pi R^2}{A\,B}$ et le rapport de maille est indifférent : le quadrant hachuré xJy est le quadrant d'indifférence pour les petites anomalies, et la probabilité P est égale au rapport des aires de l'anomalie et de la maille.

Si $\frac{\sqrt{A^2 + B^2}}{2R}$ est inférieur à 1, (point M_2), on a P = 1 quel que soit le rapport de maille : le quart de cercle O P I Q est le secteur d'indifférence des grandes anomalies. Celles-ci sont détectées à coup sûr. (P=1).

Si B/2R < 1 < A/2R (point M_3), on a intérêt à se déplacer sur l'hyperbole équilatère M_3 U V (AH = C^{te}, c'est-à-dire à dépense constante) jusqu'à venir dans le quadrant Jy ou le secteur O P I Q, ou enfin sur la droite I J, de façon à rendre maximum l'aire interceptée par le rectangle O H M_3 H sur le cercle d'indifférence. Il en est de même si $\frac{A}{2R} < 1 < \sqrt{A^2+B^2}$,

sauf que le point d'aboutissement est nécessairement sur I J.

Les rapports de maille optimum correspondent donc à un quadrant et à un secteur d'indifférence, plus le segment I J : <u>la maille carrée n'est obligatoire que pour les anomalies dont le rayon est compris entre A et $A\sqrt{2}$.</u>

On voit par affinité que les mêmes raisonnements s'appliquent à une anomalie elliptique de demi axe a et b. Il suffit de remplacer R^2 par ab, $\frac{A}{2R}$ par $\frac{A}{2a}$ et $\frac{B}{2R}$ par $\frac{B}{2b}$ dans les formules (5) et la figure (5).

La maille $\frac{A}{a} = \frac{B}{b}$ homothétique de l'anomalie n'est obligatoire que si $\frac{A}{2a}$ est compris entre $\frac{1}{\sqrt{2}}$ et 1. Inversement, on voit que le rapport de maille doit être pris égal au rapport des axes des anomalies qui sont de la taille de la maille, il est indifférent aux grandes et aux petites anomalies.

Si la maille est telle que $\frac{a}{A} = \frac{b}{B}$, on se ramène, par affinité, à une anomalie circulaire reconnue par une maille carrée. Les formules (5) se simplifient :

$$(6)\qquad \begin{cases} P = \dfrac{\pi R^2}{A^2} & \text{pour} \quad 2R < A \\ P = \left[\pi - 4 \operatorname{Arc\,cos} \dfrac{A}{2R} + \dfrac{A\sqrt{4R^2 - A^2}}{R^2}\right] \dfrac{R^2}{A^2} & \text{pour } A < 2R < A\sqrt{2} \\ P = 1 & \text{pour } 2R > A\sqrt{2} \end{cases}$$

La courbe représentative est donnée sur la figure 6 (en abscisse le rapport 2R/A du diamètre de l'anomalie à la maille). Cette courbe et sa dérivée sont continues aux deux points de transition 2R/A = 1 et $2R/A = \sqrt{2}$.

Si l'on a affaire surtout à des petites anomalies, il n'y aura pas d'inconvénients à remplacer la courbe rèelle par la courbe obtenue en prolongeant la parabole $P = \frac{\pi}{4}\left(\frac{2R}{A}\right)^2$ jusqu'au point P d'ordonnée 1. (en pointillé sur la fig.6).

Figure - 6-

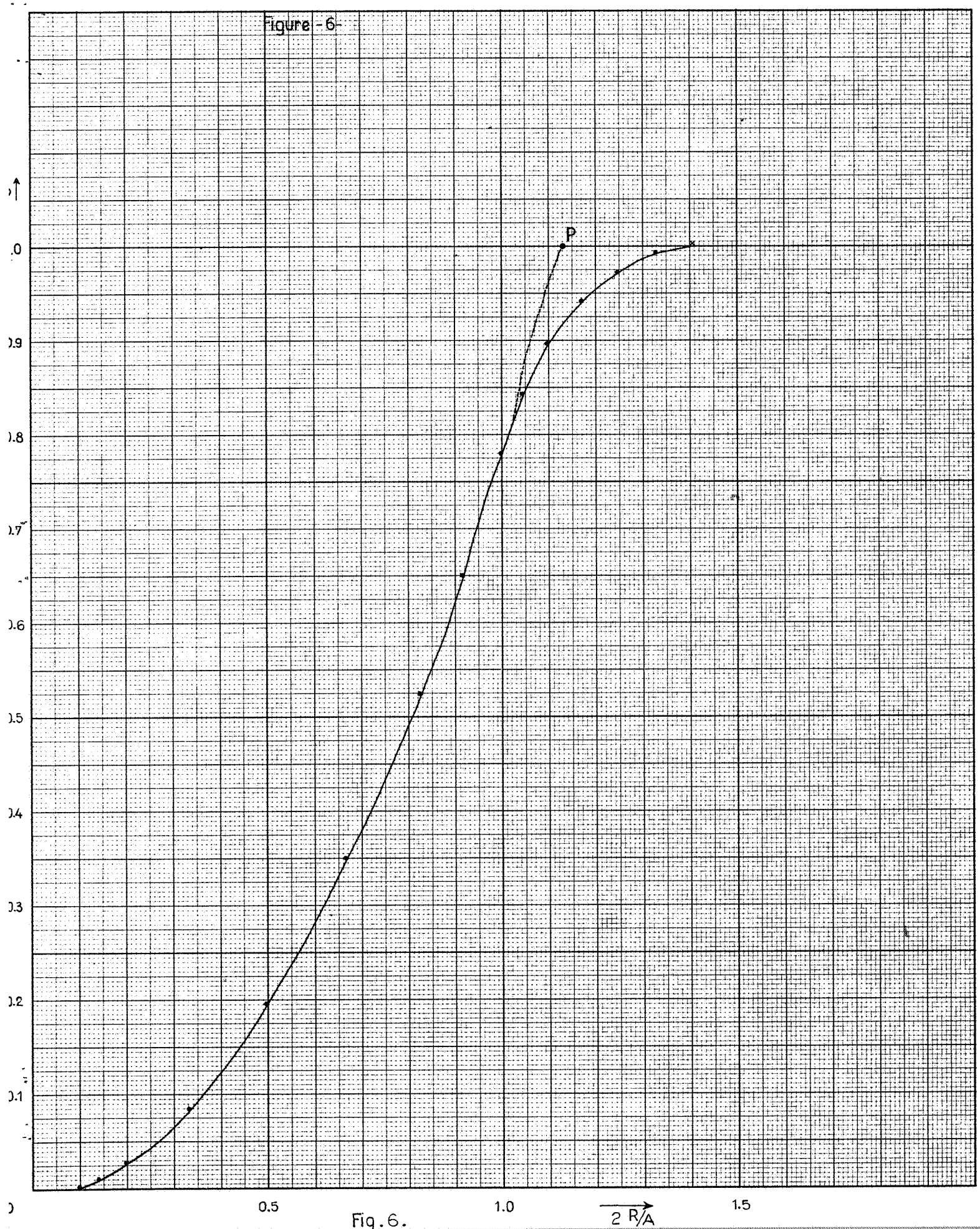


Fig. 6.

II. Effet d'une maille sur une population d'anomalies.

L'effet d'une maille donnée sur une population d'anomalie peut être prévu à partir de la fonction P donnant la probabilité pour qu'une anomalie de dimensions déterminées soit touchée par la maille. Les différentes anomalies, en effet, ont vis à vis de la maille des comportements indépendants, sous réserve seulement que la distance moyenne entre deux anomalies voisines soit relativement grande vis à vis de la maille.

Nous nous limiterons, dans ce qui suit, au cas où chaque anomalie est caractérisée par une seule dimension a. (Celà revient à admettre que les anomalies elliptiques ou rectangulaires ont même allongement b/a, de sorte que la deuxième dimension b est proportionnelle à a). Il serait facile de généraliser le problème aux cas ou les deux dimensions a et b interviendraient. Une population d'anomalie est alors définie par la donnée du nombre ni d'anomalies de dimensions a_i, le nombre total N d'anomalies étant :

$$N = \sum n_i$$

Pour une maille (A,B) donnée, la probabilité pour que l'anomalie de taille a_i soit touchée est donnée par la fonction P (a_i). Le nombre x_i d'anomalies de taille a_i décelées sur les n_i existant au total est une variable binomiale de moyenne n_i P (a_i) et de variance n_i P (a_i) $\left[1 - P\ (a_i)\right]$. Enfin, le nombre total $X = \sum x_i$ d'anomalies de tailles quelconque est une variable aléatoire dont la valeur probable et la variance sont :

$$(7) \qquad \begin{cases} E\ (X) = \sum n_i\ P_i\ (a_i) \\ D^2(X) = \sum n_i\ P_i\ (a_i) \left[1 - P\ (a_i)\right] \end{cases}$$

En première approximation, X, somme de variables binomiales indépendantes, pourra être considérée comme une variable normale ayant les paramètres ci-dessus.

Il sera souvent plus commode de définir la population d'anomalies par une loi de distribution continue. On dira, par exemple, qu'il y a au total N anomalies, dont N f (a) da ont une taille comprise entre

a et a+da. La fonction f (a) définit une véritable loi de distribution des anomalies selon leur taille. Elle pourra souvent être assimilée à une fonction log normale. Dans ces conditions, la moyenne et la variance du nombre X d'anomalies décelés ont pour expression :

$$(8)\qquad \begin{aligned} E(X) &= N \int_0^{\infty} P(a)\, f(a)\, da \\ D^2(X) &= N \int P(a) \left[1 - P(a)\right] f(a)\, da \end{aligned}$$

et la variable X elle-même est approximativement normale. Examinons deux exemples.

A/ Profils continus. (prospection aérienne). Dans ce cas, la fonction P (a) est donnée en (1), A désignant l'espacement des profils, et a la dimension de l'anomalie perpendiculaire aux profils. Admettons que cette dimension a soit distribuée lognormalement avec une médiane γ et une variance σ^2.

On déduira de (1) et (8) :

$$\begin{cases} \dfrac{E(X)}{N} = \displaystyle\int_0^A \frac{a}{A} f(a)\, da + \int_A^{\infty} f(a)\, da \\ \dfrac{D^2(X)}{N} = \displaystyle\int_0^A \frac{a}{A}\left(1 - \frac{a}{A}\right) f(a)\, da \end{cases}$$

soit, d'après le formalisme lognormal classique :

$$(9)\qquad \begin{cases} \dfrac{E(X)}{N} = \left[1 - G(z - \sigma)\right] \dfrac{\gamma}{A}\, e^{\sigma^2/2} + G(z) \\ \dfrac{D^2(X)}{N} = \left[1 - G(z - \sigma)\right] \dfrac{\gamma}{A}\, e^{\sigma^2/2} - \left[1 - G(z - 2\sigma)\right] \dfrac{\gamma^2}{A^2}\, e^{2\sigma^2} \end{cases}$$

Avec :

$$z = \frac{1}{\sigma} \log \frac{A}{\gamma}$$

La valeur probable et la variance du nombre X d'anomalies touchées sont toutes deux proportionnelles à N, nombre total des anomalies existantes. Le coefficient de proportionnalité ne dépend que de la médiane γ/A et de la variance σ^2 de la taille des anomalies rapportée à la maille A, et

peut être facilement tabulé.

B/ Anomalies Elliptiques.

Nous supposons que les anomalies sont des ellipses semblables ($\frac{b}{a} = C^{te}$), et que la maille rectangulaire est adaptée à l'allongement de ces ellipses, soit $\frac{B}{A} = \frac{b}{a}$. La fonction P (a) est donnée par (6).

Nous simplifierons l'expression de (6) en admettant que l'on a :

$$(10) \qquad \begin{cases} P(a) = \dfrac{\pi a^2}{A^2} & \text{pour } a < A < \sqrt{\pi}\,a \\ P(a) = 1 & \text{pour } a > A/\sqrt{\pi} \end{cases}$$

Celà revient à remplacer la courbe réelle de la figure 6 par la parabole qui vient mourir au point P : une telle approximation est admissible, en particulier, si la majorité des anomalies sont de dimensions petites vis à vis de la maille.

Dans ces conditions, les formules (8) nous donnent :

$$\frac{E(X)}{N} = \frac{\pi}{A^2}\,\gamma^2 e^{26^2}\left[1 - G(z - 26)\right] + G(z)$$

$$\frac{D^2(X)}{N} = \frac{\pi}{A^2}\,\gamma^2 e^{26^2}\left[1 - G(z - 26)\right] - \frac{\pi^2}{A^4}\,\gamma^4 e^{86^2}\left[1 - G(z - 46)\right]$$

Avec :

$$z = \frac{1}{6}\log\frac{A}{\gamma\sqrt{\pi}}$$

Valeur probable et variance de X sont encore proportionnels au nombre N d'anomalies, le coefficient de proportionnalité ne dépendant que de γ/A et de 6^2 - et pouvant donc être facilement tabulé.

III - Problème Inverse.

Inversement - et c'est le problème qui se posera en pratique - sachant que la reconnaissance a décelé X anomalies, quelle fourchette peut-on en déduire relativement au nombre N d'anomalies existant réellement ? Ce n'est là rien d'autre que le problème habituel de la probabilité des causes. Il se posera de façon très différente selon que les dimensions rèelles des

anomalies décelées seront, ou non, connues.

A/ Supposons connues les dimensions des anomalies décelées. Il en sera ainsi en pratique, lorsque la reconnaissance générale est immédiatement suivie d'une étude à maille plus serrée de chacune des anomalies trouvées. Dans ce cas, il est possible de reconstituer approximativement la loi de distribution f (a) des anomalies réelles. Il suffira de pondérer par $\frac{1}{P(a)}$ la fréquence des anomalies de taille a décelées par la reconnaissance. On a donc un moyen d'estimer γ et 6_1^2 et de calculer les intégrales qui figurent en (8). En posant :

$$f_1 = \int_0^{\infty} P\ (a)\ f\ (a)\ da \qquad f_2 = \int_0^{\infty} P\ (a) \left[1 - P\ (a)\right]\ f\ (a)\ da$$

les équations (8) nous donnent :

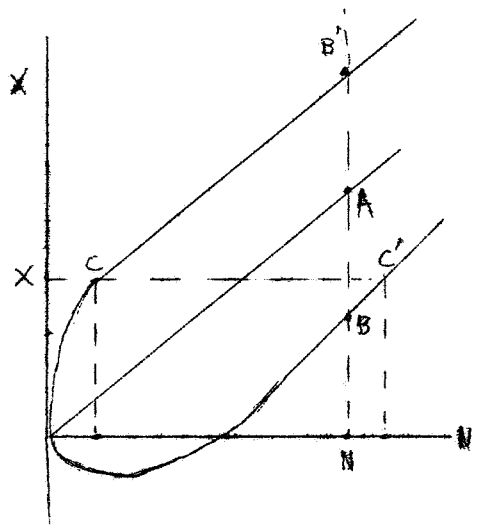
$$E\ (X) = f_1\ N \qquad D^2\ (X) = f_2\ N$$

f_1 et f_2 étant connue. On fait ensuite le raisonnement habituel : si le nombre réel d'anomalies était N, il y aurait 95 chances sur cent pour que la reconnaissance ait donné une valeur de X comprise dans la fourchette :

$$f_1 N \pm 2 \sqrt{f_2\ N}$$

c'est-à-dire pour que le point de coordonnées (NX) soit compris entre B et B'. Lorsque N varie, le lieu de B et B' est une parabole :

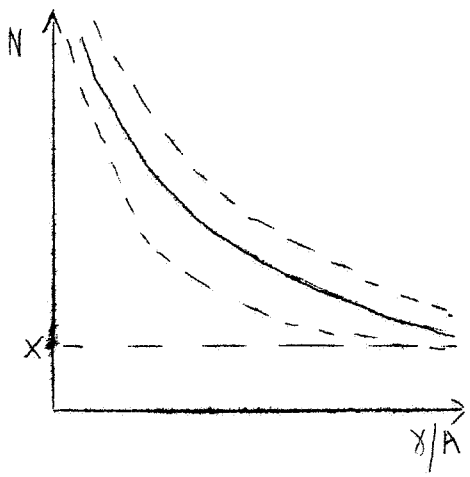
$$(X - f_1\ N)^2 = 4\ f_2\ N$$



Inversement, la reconnaissance ayant donné un nombre X d'anomalies, on admettra que le nombre total N d'anomalies existant est compris entre les deux points C et C' ou l'horizontale d'ordonnée X coupe cette parabole, c'est-à-dire dans la fourchette :

$$(12) \qquad N = \frac{X}{f_1} + \frac{2 f_2}{f_1^2} \pm \frac{2}{f_1} \sqrt{\frac{f_2 X}{f_1} + \frac{f_2^2}{f_1^2}}$$

B/ En général, cependant, on ne connaîtra pas les dimensions réelles des anomalies décelées, si la plupart d'entre elles ne sont touchées qu'en un seul point, et si l'on n'a pas fait de reconnaissance détaillée complémentaire. Dans ces conditions, il conviendra de faire des hypothèses vraisemblables sur la médiane γ et la variance σ^2 de la taille des anomalies. A chaque hypothèse (γ, σ^2) la formule (12) fait correspondre une fourchette pour N. En pratique, on pourra souvent se donner a priori la variance σ^2, par analogie avec des cas mieux étudiés. La fourchette (12) ne dépend plus alors que du rapport γ/A, et l'on peut porter sur un graphique en abscisse γ/A, en ordonnée la fourchette, comme sur la figure ci-jointe. La construction d'un tel graphique sera assez rapide si l'on dispose d'un jeu de paraboles représentant N en fonction de X, selon la formule (12), pour diverser valeurs du rapport γ/A et pour la valeur numérique choisie de σ^2.



Un tel graphique montre bien qu'un même nombre X d'anomalies décelées peut correspondre soit à un très grand nombre N d'anomalies très petites, soit à un nombre beaucoup plus réduit d'anomalies de dimensions plus grandes. On aura parfois une indication sur l'ordre de grandeur du rapport γ/A de la dimension médiane des anomalies à la maille A d'après le pourcentage des anomalies touchées en plus de deux points. C'est surtout l'expérience qui suggéra dans quelle zone du graphique il convient de se placer.

-:-:-:-:-:-:-