

SUR UNE MÉTHODE D'ANALYSE FAUNISTIQUE RAPIDE DU ZOOPLANCTON

S. FRONTIER

Océanographe Biologiste de l'O.R.S.T.O.M., Centre O.R.S.T.O.M. de Nosy-Bé (Madagascar).

Abstract: The problem of the estimation of the number of planktonic organisms in a series of collections is considered. A rough estimate of the abundance of the taxa is given by means of a scale based upon a geometric progression with a basis of 4.3. The \log^2 transformation, intermediate between the logarithmic and the square root transformations, is proposed for the treatment of data. The abundance values, which are logarithms, are then plotted on an x^2 scale on any diagrams. The mathematical justification and examples of the use of the technique are given.

Aux fins d'obtenir, dans des délais réduits, une physionomie du zooplancton d'une région encore totalement inconnue à cet égard (Nosy-bé, Madagascar), j'ai été amené à rechercher une méthode rapide d'analyse des échantillons de plancton. En raison de la variabilité considérable des récoltes (tenant à la surdispersion des organismes ainsi qu'au caractère fruste de l'échantillonnage au filet), en raison également de l'espacement des pêches dans le temps et dans l'espace pour cette première étude de plancton (récoltes espacées en moyenne de 10 milles, et répétées généralement une fois par mois), il est apparu illusoire de prétendre exploiter la précision qu'auraient fourni, au prix d'un allongement considérable du temps de travail, les comptages exacts d'organismes. En dépit des fluctuations aléatoires, des variations significatives apparaissaient au niveau de l'ordre de grandeur des abondances numériques. J'ai donc tenté de préciser cette notion d'ordre de grandeur afin de pouvoir l'utiliser dans une description semi quantitative des phénomènes. Une *cotation d'abondance* a été établie sur la base d'une progression géométrique paraissant adéquate: les classes suivantes ont été adoptées (Tab. I à gauche).

TABLEAU I
Définition des classes d'abondance.

Cotes	Effectifs	Cotes	Effectifs
0	0		
1	1 à 3	1,5	3 ou 4
2	4 à 17	2,5	environ 18
3	18 à 80	3,5	environ 80
4	80 à 350	4,5	environ 350
5	350 à 1 500	5,5	environ 1 500
6	1 500 à 6 500	6,5	environ 6 500
7	6 500 à 30 000	7,5	environ 30 000
8	30 000 à 120 000	8,5	environ 120 000
9	120 000 à 500 000	9,5	environ 500 000
10	500 000 à 2 000 000		

O. R. S. T. O. M.

Collection de Référence

n° 13053

4 MARS 1969

La suite des nombres 1; 4; 18; 80; 350; 1500; 6500; 30 000; 120 000; 500 000; 2 000 000 est voisine d'une progression géométrique de raison 4,3¹. Le choix de cette base mérite quelques commentaires. A l'origine, la raison 4,3 fut choisie intuitivement de la façon suivante. Après avoir noté, dans une centaine de cas, des unités taxonomiques présentes dans des récoltes planctoniques comme étant 'très rares', 'rares', 'assez rares', etc. (soit, au total, 7 qualifications), il a été procédé à des comptages précis des mêmes organismes. On a constaté alors une proportionnalité régulière entre les différentes expressions de l'abondance, rangées en ordre linéaire, et les effectifs précis portés en échelle logarithmique. (Cette coïncidence n'est pas extrêmement étonnante, les lois de la perception établissant que, d'une façon très générale, ce sont les variations géométriques et non les variations arithmétiques des grandeurs perçues qui sont spontanément appréciées.) Le coefficient de proportionnalité constaté se trouvait très proche du logarithme de 4,3, nombre qui fut alors adopté pour construire une échelle précise d'abondance devant donner des résultats plus rigoureux que ceux obtenus antérieurement mais cependant comparables à ces derniers.

Il est encore trop tôt pour dire si 4,3 représente un coefficient purement personnel, ou si au contraire il caractérise la perception des nombres chez l'homme du moins dans certaines conditions. Cependant un calcul d'information aboutit à découvrir une coïncidence méritant d'être signalée.

Le planctologiste se trouve dans la situation concrète suivante: considérant deux effectifs d'organismes dont l'un est K fois plus grand que l'autre, il s'agit d'estimer K 'à vue'. On cherchera, en première approximation, un nombre entier, dont on sait seulement qu'il n'est pas trop grand. On 'essaie' alors, mentalement, les valeurs successives $k = 1, 2, 3, \dots$; dans ces conditions l'indication ' $K = k$ ' contient exactement k unités d'information ou bits. (En effet pour déterminer K , il faut répondre par oui ou non à k questions successives, les deux réponses possibles étant chaque fois isoprobables *a priori*). Dès lors, ne considérer un effectif comme significativement supérieur à un autre que s'il est au moins k fois plus grand, revient à ne prendre en considération que des quantités d'information au moins égales à k bits. Dans le cas présent l'expérience intuitive des récoltes de plancton semble conduire à adopter la valeur 4 ou la valeur 5, avec comme moyenne 4,3. En d'autres termes on n'accorde, intuitivement, de signification qu'à des quantités d'information au moins égales à 4,3 bits en moyenne – les quantités inférieures sont considérées comme ne pouvant se distinguer du 'bruit', c'est-à-dire de la variabilité aléatoire.

Or le seuil de signification 5% habituellement adopté par le traitement statistique des données biologiques correspond exactement à cette valeur. La quantité d'information apportée par la réalisation d'un événement n'ayant que 5 chances sur 100 de se produire par hasard est en effet égale à $\log_2 \frac{1}{20}$ No. 4,3 bits: exclure les cas

¹ Des valeurs plus justes seraient: 1; 4,3; 18,5; 79,5; 342; 1470; 6321; 27 180; 116 800; 502 600; 2 061 000. Ces nombres ont été arrondis pour être plus faciles à retenir. Comme il s'agit de caractériser un ordre de grandeur, les résultats ne sont pas altérés par cette simplification.

n'ayant au plus que 5 % de chances de se produire par hasard, c'est accorder de signification qu'à des quantités d'information au moins égales à 4,3 bits. C'est effectuer le même 'effort mental' (mesuré en unités d'information) que lorsqu'on considère un effectif significativement plus grand qu'un autre s'il lui est 4,3 fois supérieur. Le seuil de 5 % est adopté largement non seulement en raison de sa simplicité, mais parce qu'à l'usage il se montre particulièrement apte à faire émerger la signification

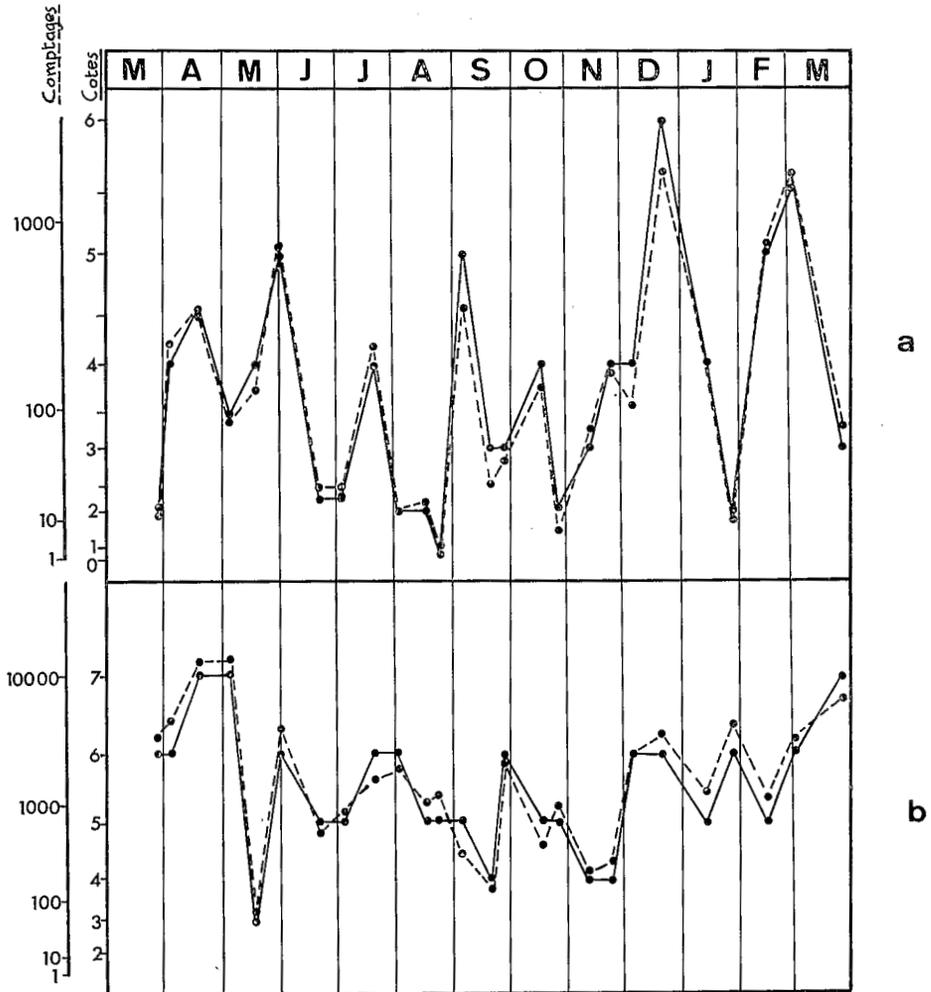


Fig. 1. Comparaison des graphiques obtenus à partir des comptages exacts (traits interrompus) et à partir des cotations d'abondance (traits pleins): a. Ptéropodes; b. Chaetognathes; Nosy-Bé, Station 10, immersion 2m, mars 1963 à mars 1964. Les comptages sont portés en échelle \log^2 et les cotes en échelle x^2 (justification ci-après). On constate entre les deux types de graphiques une coïncidence suffisante pour qu'à ce stade de la description des peuplements planctoniques, les comptages exacts soient considérés comme n'apportant pratiquement aucune information supplémentaire, et soient donc abandonnés au profit de la méthode approximative plus rapide.

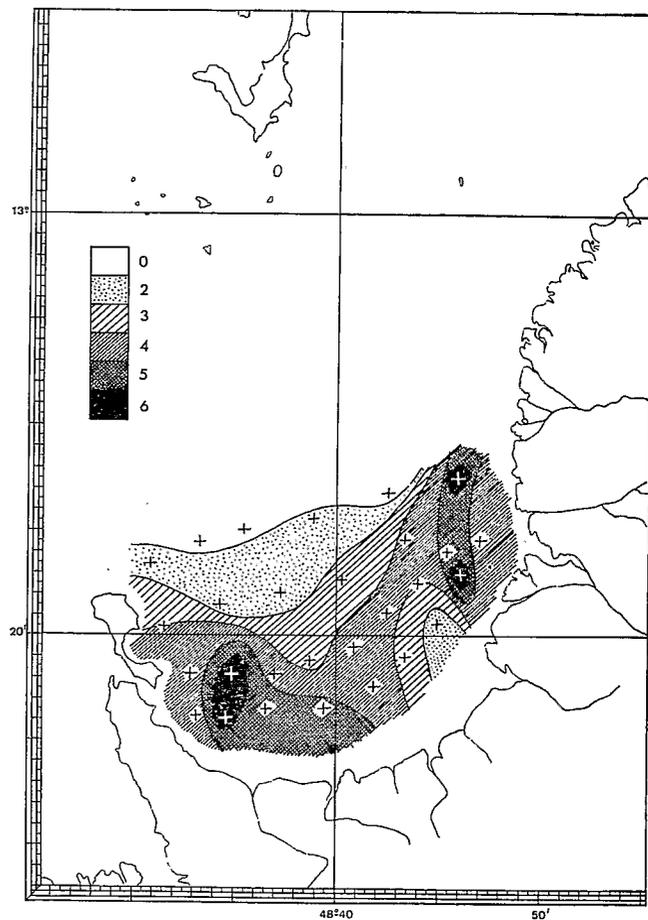
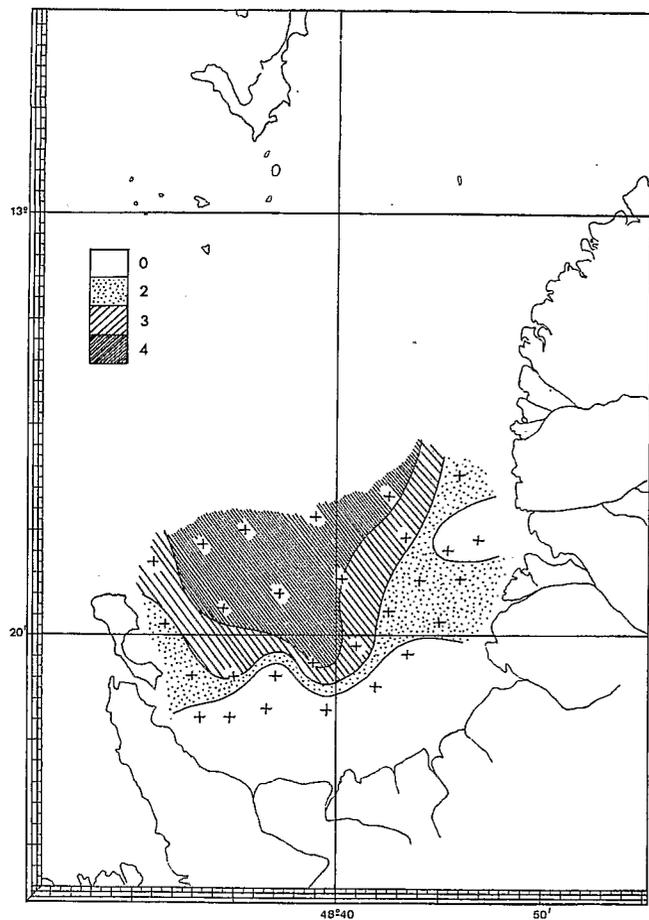


Fig. 2. Carte de répartition de deux catégories d'organismes dans une baie proche de Nosy-Bé, obtenues au moyen des cotations d'abondance: à gauche, Calycopteres; à droite, *Penilia avirostris*.

dans le domaine des faits biologiques; la coïncidence des deux valeurs n'est donc pas sans intérêt.

Ceci dit, il reste à justifier la méthode, en montrant que les résultats auxquels elle aboutit concurrencent largement ceux donnés par la méthode des comptages exacts. Quelques résultats ont déjà été publiés (Frontier, 1966); les graphiques représentant des variations mensuelles, horaires, spatiales d'abondance ont souvent des allures très interprétables, et lorsque tel n'est pas le cas, il est manifeste que des comptages précis n'auraient abouti à aucune signification supplémentaire. Deux graphiques, relatifs aux Ptéropodes et aux Chaetognathes en une station du plateau continental malgache, sont repris ici (Fig. 1). Ils permettent de comparer les résultats obtenus, sur une même série de récoltes, à partir des comptages et des cotations d'abondance. La coïncidence est manifeste. Il apparaît qu'à ce niveau d'observation, la précision supplémentaire apportée par les comptages exacts appartient à la variabilité aléatoire.

La Fig. 2 donne deux cartes de répartition obtenues à la suite d'un quadrillage de 28 stations dans une baie du plateau continental malgache. La première illustre la répartition d'un groupe zoologique qui dans son ensemble évite l'abond immédiat de la côte (Siphonophores Calycophores), la seconde celle d'une espèce (*Penilia avirostris*) très littorale au contraire.

D'un point de vue pratique l'expérience montre qu'il est aisé de voir si l'effectif d'une catégorie d'organisme dans un échantillon se trouve, par exemple, compris entre 80 et 350 ou entre 350 et 1500. Diverses méthodes d'estimation peuvent être utilisées. Le doute ne s'établit que si l'effectif est voisin de 350. Plutôt alors que reprendre l'estimation d'abondance d'une façon plus précise on adopte une cote conventionnelle intermédiaire: 4,5. Le tableau de cotation se complète donc par une liste de cotes fractionnaires (Tab. I à droite). Sur les graphiques représentant des variations d'abondance, l'adoption de ces cotes intermédiaires se traduira par des courbes moins heurtées. Sur les cartes de répartition, les courbes d'iso-abondance passeront par les points affectés d'une cote fractionnaire.

Décrivons brièvement, à titre indicatif, la méthode employée actuellement pour trouver une cote d'abondance. L'échantillon de plancton est mis en suspension¹ et homogénéisé le mieux possible par agitation. Une fraction volumétrique connue de la suspension est alors prélevée au moyen d'une poire de caoutchouc manipulée toujours de la même façon (erreur de $\pm 2\%$ sur la fraction - celle-ci est généralement comprise entre $\frac{1}{10}$ et $\frac{1}{40}$) et déposée dans une cuvette de Dollfuss². Suivant l'abondance des organismes, ceux-ci sont dénombrés sur la totalité de la cuvette ou sur une partie aliquote, par exemple $\frac{1}{2}$ (une colonne de carrés sur deux), $\frac{1}{10}$ (une double diagonale), etc. Connaissant la fraction d'échantillon représentée par le contenu de la cuvette, on en déduit immédiatement une estimation de l'effectif total dont on

¹ Il est commode que le volume total de la suspension soit égal environ à 5 fois celui du plancton sédimenté.

² Cuvette en verre, de 5 cm sur 10, dont le fond est divisé par un quadrillage en relief en 200 carrés de 5 mm sur 5. Fournisseur: LEUNE, Paris.

déduit la valeur de la cote d'abondance. Cette cote est considérée comme une *mesure*. On s'impose des lors d'effectuer trois fois la même mesure c'est-à-dire d'examiner trois cuvettes successives. Après un comptage, le contenu de la cuvette est replacé dans la récolte afin de ne pas modifier les probabilités de présence des organismes pour le comptage suivant.

La variabilité des comptages successifs dans une même récolte dépasse en général celle qui pourrait être prévue à partir de la loi binomiale. Toutefois, l'expérience montre que les comptages obtenus sur trois cuvettes varient très rarement de plus du simple au double. La cotation d'abondance étant basée sur une progression géométrique de raison 4,3, les trois estimations successives obtenues pour l'effectif total de l'échantillon ou bien se groupent à l'intérieur d'une classe, et l'on adopte la cote définissant cette classe; ou bien chevauchent une des limites, et l'on adopte une cote fractionnaire. Lorsque les résultats des comptages se montrent trop dispersés, on procède au dénombrement d'au moins deux cuvettes supplémentaires, ce qui permet en général d'éliminer une valeur aberrante (pouvant être due à un défaut d'homogénéisation, une erreur de comptage, etc.).

La plupart du temps la cote d'abondance est, ainsi, déterminée sans ambiguïté. L'incertitude apparaît cependant dans le cas des faibles effectifs. Supposons par exemple qu'une récolte contienne cinq organismes d'un certain type (cote 2), et supposons que ces cinq individus soient répartis au hasard dans la suspension. On calcule que l'on a, sensiblement, les probabilités suivantes de conclure après trois dénombrements d'un dixième de l'échantillon, aux différentes valeurs de la cote d'abondance:

$$\begin{array}{ll} P(0) = 0,21 & P(2) = 0,40 \\ P(1) = 0,00 & P(2,5) = 0,03 \\ P(1,5) = 0,34 & P(3) = 0,004 \end{array}$$

(plus une très petite probabilité d'obtenir trois valeurs trop hétérogènes). On voit que, même en considérant qu'adopter la cote 1,5 ('environ 4') ne constitue pas une erreur, il reste une probabilité non négligeable d'aboutir à une cote fautive: 0.

Pour les effectifs moyens et grands au contraire, des comptages répétés des dizaines de fois à titre de test ont montré que les résultats ne varient que du simple au double: on ne risque alors que des erreurs d'une demi-unité de cotation. Il s'ajoute à cela, qu'en raison de la dispersion aléatoire des organismes dans leur milieu naturel les faibles effectifs représentent moins sûrement leur valeur moyenne que les grands (une variation de 1000 à 5000 est très certainement significative; une variation de 1 à 5 est souvent due au hasard). Au total, la méthode exposée ci-dessus aboutit à une description des peuplements planctoniques telle que les mêmes variations relatives contiennent moins de signification dans le domaine des petits effectifs que dans celui des grands effectifs.

L'adoption d'une échelle logarithmique (logarithmes de base 4,3) pour les effectifs d'organismes se trouve en accord d'une part, semble-t-il, avec les lois de la perception

chez l'homme; d'autre part avec les lois biologiques puisque d'une façon très générale les organismes et peuplements vivants répondent quantitativement à une variation arithmétique des caractères de l'environnement par des variations géométriques (Cassie, 1962).

Mais l'échelle logarithmique attribue une importance égale à une variation relative de même valeur, qu'elle intéresse les petits nombres ou les grands et c'est un défaut lorsque l'on recherche une signification dans les variations constatées. Pour pallier partiellement cet inconvénient et pouvoir décrire des variations spatio-temporelles au moyen de graphiques dans lesquels tout ce qui apparaît nettement puisse être considéré comme significatif, j'ai représenté les résultats de comptages en échelle \log^2 (dans le cas des cotations d'abondance, c'est le carré de la cote qui est porté en ordonnées). On dispose alors d'une échelle intermédiaire entre les échelles arithmétiques et logarithmique. Par rapport à cette dernière échelle, les effectifs de petite taille sont rapprochés les uns des autres, et les grands effectifs au contraire éloignés les uns des autres de sorte que les fluctuations dont ils sont le siège, plus significatives, sont plus distinctes (la Fig. 1 en donne un exemple d'application).

Il convient d'insister sur le fait que la transformation \log^2 n'est proposée ici que sur une base empirique, et à titre de technique de dessin apte à réaliser une 'mise au point', au sens de l'optique, sur les fluctuations se produisant dans le domaine des grands et moyens effectifs seulement. Les variations affectant les petits nombres peuvent, en réalité, avoir une signification biologique, mais les aléas de l'échantillonnage ne permettent pas de les distinguer de la variabilité due au hasard: le graphique tend alors à les estomper. Il ne s'agit pas d'un modèle mathématique de distribution. Toutefois l'on remarque que la transformation \log^2 pourrait n'être pas statistiquement absurde: elle se situe en effet entre les transformations logarithme et racine carrée, toutes deux utilisées classiquement pour normaliser des distributions surdispersées et stabiliser la variance (Barnes, 1952; Cassie, 1962).

En effet, N étant un effectif brut, comparons l'effet des transformations, $y_1 = \sqrt{N}$, $y_2 = \log N$, $y_3 = \log^2 N$. La variation d'effectif dN représentée par une même petite variation d'ordonnée dy sur le graphique, se calcule en différenciant les expressions ci-dessus. On trouve (à un facteur près);

$$dN = \sqrt{N}dy_1 = Ndy_2 = (N/\log N)dy_3.$$

Les courbes représentatives des trois fonctions \sqrt{N} , N et $N/\log N$ (Fig. 3) sont respectivement, pour $N > 1$: un arc de parabole ayant comme direction l'axe des abscisses, une droite passant par l'origine, et une courbe possédant une branche parabolique parallèle à la première, et située pour N assez grand entre les deux premières courbes. Sur un segment limité, la courbe $N/\log N$ peut être ajustée à une courbe d'équation $y = a\sqrt{N} + b$. Si les transformations racine carrée et logarithme normalisent certaines distributions rencontrées dans le plancton et stabilisent la variance, il est alors probable que la transformation \log^2 aura un effet analogue.

De fait, reprenant les données de Winsor & Clarke (1940), je trouve les effets de

stabilisation de la variance suivants pour les trois transformations en question (effets exprimés par le rapport $A = 100 \times$ intervalle de variation de la variance/valeur moyenne de la variance):

effectifs bruts	$A = 355 \%$
transformation racine carrée	$A = 266 \%$
transformation log	$A = 101 \%$
transformation \log^2	$A = 61 \%$

La transformation \log^2 a donc apporté la meilleure stabilisation de la variance.

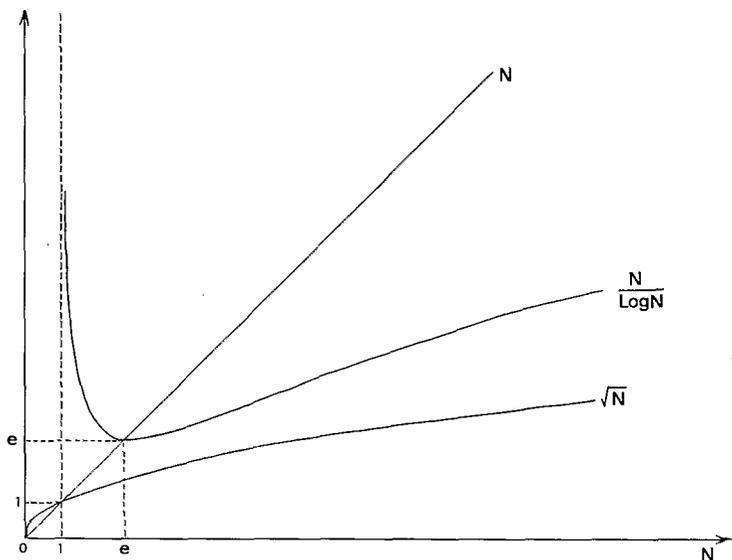


Fig. 3. Courbes représentatives des fonctions \sqrt{N} , N et $N/\log N$ illustrant respectivement l'effet des transformations \sqrt{N} , $\log N$, et $\log^2 N$.

Reprenons les données de Cassie (1962) concernant des dénombrements de *Podon polyphemoides* (p. 74, Fig. 5), on trouve que les effectifs cumulés portés sur un graphique log-probabilités s'alignent assez bien le long d'une parabole d'équation;

$$y = 0,285x^2 + 0,305x - 0,75,$$

c'est-à-dire qu'en effectuant sur les comptages la transformation

$$y = (\log N + 0,535)^2 = \log^2 3,42 N,$$

les points se trouveraient alignés le long d'une droite: les valeurs de $\log^2 3,42 N$ sont normalement distribuées.

Il est encore trop tôt pour dire si ces coïncidences ont lieu, ou non, par hasard. Mais il est significatif que, sur le graphique log-probabilités cité, un arc de parabole

semblant s'ajuster correctement aux points obtenus peut être placé entre la droite représentative de la distribution log-normale et la courbe représentative de la distribution binomiale négative. Une distribution \log^2 -normale appropriée semble donc répondre aux souhaits de l'auteur, et constitue peut-être une approximation utilisable de la distribution Poisson-log-normale qu'il propose dans le même article.

Je remercie vivement le Dr R. M. Cassie pour ses critiques et suggestions, qui ont beaucoup contribué à la mise au point de cet exposé.

BIBLIOGRAPHIE

- BARNES, H., 1952. The use of transformations in marine biological statistics. *J. Cons. perm. int. Explor. Mer*, Vol. 18, pp. 61-71.
- CASSIE, R. M., 1962. Frequency distribution models in the ecology of plankton and other organisms. *J. Anim. Ecol.*, Vol. 31, pp. 65-92.
- FRONTIER, S., 1966. Zooplancton de la région de Nosy-Bé. I) Programme des récoltes et techniques d'étude. II) Plancton de surface aux stations 5 et 10. *Cah. ORSTOM, Sér. Océanogr.*, T. 4, pp. 3-36.
- WINSOR, C. P. & G. L. CLARKE, 1940. A statistical study of variation in the catch of plankton nets. *J. mar. Res.*, Vol. 3, pp. 1-34.