

**La numération des algues en sol submergé :  
Loi de distribution des organismes  
et densité d'échantillonnage**

PAR

Pierre A. ROGER et Pierre A. REYNAUD

O.R.S.T.O.M. BP 1 386 - Dakar - République du Sénégal

**I. — INTRODUCTION**

Le rôle des Cyanophycées dans les rizières a été étudié par de nombreux auteurs, cependant les travaux portant sur les variations qualitatives et quantitatives de la flore algale dans ces biotopes sont peu nombreux (REYNAUD et ROGER, 1976) et les problèmes d'échantillonnage liés au mode de répartition des algues en sol submergé n'ont jusqu'ici pas été abordés.

Leur étude est simple dans sa conception ; elle consiste à réaliser un certain nombre de fois le plus grand nombre possible de prélèvements sur un biotope donné et à effectuer la numération des différents types d'algues sur chaque prélèvement. L'interprétation statistique des résultats permet alors de déterminer la loi de répartition des organismes, puis les densités d'échantillonnage à adopter pour obtenir une précision donnée.

Toutefois la réalisation pratique d'une telle étude demande la mise en œuvre d'un matériel important : la numération des algues sur un seul échantillon (cf. chap. II, § 2) nécessite l'inoculation de 54 boîtes de Pétri qui occupent une surface de 0,2 m<sup>2</sup> pendant trois semaines dans une étuve éclairée. Sur les 54 boîtes mises à incuber, 9 au moins seront reprises pour un comptage des colonies sous microscope stéréoscopique ; ce comptage nécessitant trois à cinq minutes par boîte.

Il en résulte que le nombre d'échantillons pouvant être traités simultanément est forcément limité. Pour cette étude nous avons effectué la numéra-

Reçu le 7-9-77.

REVUE D'ÉCOLOGIE ET DE BIOLOGIE DU SOL VOL. 15, N° 2

AVR. 1979 M  
O. R. S. T. O. M. 88

Collection de Référence

n° 3579 Bio. Sols

tion simultanée de quarante échantillons. Cette quantité, sans poser de problèmes techniques trop importants permet de se situer au-dessus de la valeur limite de trente, requise pour une estimation satisfaisante des paramètres statistiques (SCHWARTZ, 1963).

## II. — MATÉRIEL ET MÉTHODES

### A) Mode de prélèvement des échantillons :

Les échantillons sont prélevés par carottage au moyen de tubes de 30 cm de long et de 2 cm de diamètre intérieur.

Chaque tube est enfoncé verticalement dans l'eau de submersion et le sol, puis il est bouché à sa partie supérieure et retiré. L'échantillon est constitué par le premier centimètre de sol et l'eau de submersion ; il est récupéré en chassant la carotte du tube au moyen d'un piston.

Les numérations sont rapportées à une colonne de 1 cm<sup>2</sup> de section englobant à la fois l'horizon 0-1 cm du sol et l'eau de submersion située au dessus ; l'extrapolation des résultats au niveau de la parcelle est ainsi plus rigoureuse et plus simple.

### B) Numérations :

Les numérations sont faites par la méthode des suspensions-dilutions avec étalement sur milieux gélosés à 1 % répartis à raison de 30 ml par boîte de Pétri. On effectue trois répétitions par dilution. Trois milieux ont été utilisés afin de permettre une numération sélective des Algues classées en procaryotes fixatrices d'azote, procaryotes totales et eucaryotes. La composition de ces milieux a été précédemment décrite (REYNAUD et ROGER, 1977).

Les trois boîtes d'une même dilution sont superposées et placées dans une étuve à 30° C sous un éclairage de 500 lux produit par des tubes au néon type lumière du jour (MAZDA FLUOR T.F.R. 20 W).

Au cours de l'incubation on fait subir tous les trois jours une permutation horizontale et verticale aux boîtes de Pétri afin d'éliminer toute hétérogénéité liée à la position des boîtes dans l'étuve.

Les comptages sont effectués après vingt et un jours d'incubation, sous un microscope stéréoscopique (WILD, M5).

### C) Protocole expérimental :

L'étude a été effectuée sur une parcelle expérimentale de l'Institut Sénégalais de Recherches Agronomiques, au cours d'un cycle de culture du riz. Les caractéristiques physico-chimiques du sol et le calendrier des pratiques culturales ont été précédemment décrites (ROGER et REYNAUD, 1976). La parcelle, d'une surface totale de 0,25 ha, a été divisée en quarante sous-parcelles égales, au centre desquelles nous avons prélevé dix carottes qui ont été réunies et homogénéisées. Nous avons ainsi obtenu quarante échantillons sur lesquels la numération des principaux types d'algues a été effectuée.

Le protocole a été appliqué trois fois au cours d'un cycle cultural du riz. Lors de la première série de prélèvements, le sol, après avoir été submergé, était en voie de dessiccation ; par suite d'un planage imparfait, vingt six sous-parcelles étaient desséchées en surface et quatorze étaient encore humides.

Les résultats sont d'abord étudiés séparément en fonction de ces deux situations, sous les désignations « sol sec — 26 » et « sol humide — 14 » sur le tableau général des résultats (Tabl. 1) ; ils sont ensuite regroupés sous la désignation « sol hétérogène — 40 ». Les deux autres séries ont été prélevées sur sol submergé.

Les numérations de six espèces lors de la première série de prélèvements, de quatre lors de la deuxième et de sept lors de la troisième nous ont permis d'obtenir 29 histogrammes de répartition.

### III. — RÉSULTATS ET DISCUSSION

Au total 760 résultats ont été obtenus ; ils correspondent chacun à la moyenne des comptages effectués sur trois répétitions. Tous ces résultats ne sont pas directement nécessaires à la compréhension de l'étude et n'ont pas été rapportés, à l'exception d'une série pour laquelle l'ensemble des calculs effectués est donné à titre d'exemple (cf. § III E).

#### A) Détermination de la loi de distribution des algues :

Pour chaque série de numérations nous avons tracé l'histogramme de répartition et calculé la moyenne et la variance des données.

Les histogrammes obtenus ont tous la même forme en L et il n'est pas besoin de test pour rejeter l'hypothèse de normalité. La figure n° 1 présente quelques exemples de ces histogrammes.

Le calcul de l'intervalle de confiance d'une moyenne ne peut s'effectuer directement que sur une population distribuée normalement ou dont la distribution peut s'ajuster de façon satisfaisante à une loi normale.

Il importe donc de déterminer la transformation à appliquer aux données pour les normaliser, en étudiant la liaison moyenne-variance ( $\mu$  ;  $\sigma^2$ ). On établit (KENDALL et STUART ; 1963-68) que si une famille de distributions peut être caractérisée par une relation moyenne-variance de la forme  $\sigma^2 = f(\mu)$  la transformation rendant la variance indépendante de la moyenne est de la forme :

$$g(x) = \int^x \frac{d\mu}{\sqrt{f(\mu)}}$$

L'étude de la liaison moyenne-variance sur les dénombrements d'organismes dans les biotopes aquatiques et terrestres a abouti à une loi dite loi de TAYLOR (TAYLOR, 1961) :  $m$  et  $s^2$  étant respectivement les estimations de  $\mu$  et  $\sigma^2$ , les points ( $m$ ,  $s^2$ ) portés sur un graphique log-log se répartissent suivant un nuage rectilinéaire qui traduit une liaison de la forme  $\log s^2 = a + b \log m$  soit :

$$s^2 = a \cdot m^b.$$

On déduit la transformation à utiliser pour normaliser les données

$$g(x) = a \int^x m^{1-b/2} \cdot dm$$

— si  $b = 1$   $g(x) = \sqrt{x}$  ; ce résultat correspond à une distribution de POISSON.

— si  $1 < b < 2$   $g(x) = \frac{a}{1-b/2} \cdot x^{1-b/2}$

— si  $b = 2$   $g(x) = \log x$  ; ce résultat correspond à une distribution log-normale.

L'examen des valeurs  $m_x$  et  $s_x^2$  calculées (colonnes 3 et 4, Tab. I) montre qu'elles ne sont pas indépendantes et qu'il existe une liaison stochastique entre  $m_x$  et  $s_x^2$  (Fig. n° 2 et colonne n° 5 du Tabl. I). Le nuage de points ( $m$ ,  $s^2$ ) est parfaitement rectilinéaire et s'ajuste à une droite de pente 2 passant par l'origine. Nous trouvons là une vérification particulière ( $b = 2$ ) de la loi de TAYLOR, pour  $m$  et  $s^2$  variant dans des intervalles extrêmement importants.

Tableau synoptique des différentes étapes des calculs effectués

|                   |         | $\bar{X}$           | $S^2$                | $\frac{\log S^2}{\log \bar{X}}$ | n.m.c.c. | O. | $\bar{Y}$ | $S^2y$ | $Sy$ |
|-------------------|---------|---------------------|----------------------|---------------------------------|----------|----|-----------|--------|------|
| Navicula          | sec. 26 | 1,7.10 <sup>4</sup> | 2,7.10 <sup>8</sup>  | 2,0                             | 26       | +  | 4,06      | 0,15   | 0,39 |
|                   | hum. 14 | 8,6.10 <sup>4</sup> | 1,8.10 <sup>9</sup>  | 1,9                             | 146      | +  | 4,89      | 0,06   | 0,24 |
|                   | het. 40 | 4,1.10 <sup>4</sup> | 2,0.10 <sup>9</sup>  | 2,0                             | 68       | +  | 4,35      | 0,27   | 0,52 |
|                   | sub. 40 | 2,0.10 <sup>5</sup> | 7,2.10 <sup>10</sup> | 2,1                             | 27       | +  | 4,72      | 0,61   | 0,78 |
| Chlorel.          | sec. 26 | 8,2.10 <sup>3</sup> | 1,1.10 <sup>8</sup>  | 2,0                             | 13       | +  | 3,60      | 0,37   | 0,61 |
|                   | hum. 14 | 2,0.10 <sup>4</sup> | 1,1.10 <sup>9</sup>  | 2,1                             | 32       | +  | 4,00      | 0,29   | 0,54 |
|                   | het. 40 | 1,2.10 <sup>4</sup> | 4,7.10 <sup>8</sup>  | 2,1                             | 19       | +  | 3,74      | 0,37   | 0,61 |
|                   | sub. 40 | 2,0.10 <sup>5</sup> | 2,3.10 <sup>10</sup> | 2,0                             | 35       | +  | 4,84      | 0,54   | 0,72 |
| Oscillator.       | sec. 26 | 8,3.10 <sup>3</sup> | 3,3.10 <sup>6</sup>  | 2,2                             | 1,3      | —  |           |        |      |
|                   | hum. 14 | 6,4.10 <sup>3</sup> | 1,0.10 <sup>6</sup>  | 2,1                             | 0,9      | —  |           |        |      |
|                   | het. 40 | 7,6.10 <sup>3</sup> | 2,5.10 <sup>6</sup>  | 2,2                             | 1,2      | —  |           |        |      |
|                   | sub. 40 | 2,4.10 <sup>6</sup> | 4,1.10 <sup>12</sup> | 2,0                             | 2,5      | +  | 6,11      | 0,10   | 0,44 |
| PS. Ana-<br>baen. | sec. 26 | 5,2.10 <sup>3</sup> | 2,4.10 <sup>7</sup>  | 2,0                             | 8,2      | +  | 3,54      | 0,18   | 0,42 |
|                   | hum. 14 | 1,1.10 <sup>4</sup> | 4,0.10 <sup>7</sup>  | 1,9                             | 17       | +  | 3,96      | 0,10   | 0,31 |
|                   | het. 40 | 7,2.10 <sup>3</sup> | 3,7.10 <sup>7</sup>  | 2,0                             | 11       | +  | 3,69      | 0,19   | 0,43 |
|                   | sub. 40 | 1,2.10 <sup>6</sup> | 2,7.10 <sup>7</sup>  | 1,9                             | 117      | +  | 6,02      | 0,04   | 0,21 |
| Scyto-<br>nema    | sec. 26 | 2,2.10 <sup>3</sup> | 2,2.10 <sup>4</sup>  | 1,9                             | 0,3      | —  |           |        |      |
|                   | hum. 14 | 1,9.10 <sup>3</sup> | 5,0.10 <sup>4</sup>  | 2,0                             | 0,3      | —  |           |        |      |
|                   | het. 40 | 2,1.10 <sup>3</sup> | 1,6.10 <sup>5</sup>  | 1,8                             | 0,3      | —  |           |        |      |
|                   | sub. 40 | 4,8.10 <sup>3</sup> | 4,2.10 <sup>7</sup>  | 1,8                             | 0,5      | —  |           |        |      |
| Anabaena          | sec. 26 | 2,2.10 <sup>4</sup> | 5,2.50 <sup>8</sup>  | 2,0                             | 3,0      | +  | 3,84      | 0,49   | 0,70 |
|                   | sec. 26 | 6,3.10 <sup>3</sup> | 1,7.10 <sup>7</sup>  | 1,9                             | 9,9      | +  | 3,67      | 0,16   | 0,40 |
|                   | hum. 14 | 1,3.10 <sup>4</sup> | 4,0.10 <sup>7</sup>  | 1,9                             | 20       | +  | 4,04      | 0,09   | 0,30 |
|                   | het. 40 | 8,6.10 <sup>3</sup> | 3,4.10 <sup>7</sup>  | 1,9                             | 13       | +  | 3,80      | 0,17   | 0,41 |
| sub. 40           | sub. 40 | 4,0.10 <sup>4</sup> | 2,1.10 <sup>9</sup>  | 2,0                             | 4,3      | +  | 4,39      | 0,20   | 0,44 |
|                   | sub. 40 | 4,8.10 <sup>4</sup> | 1,8.10 <sup>9</sup>  | 2,0                             | 12       | +  | 4,50      | 0,23   | 0,48 |
| Microsyst.        | sub. 40 | 2,4.10 <sup>4</sup> | 7,4.10 <sup>8</sup>  | 2,0                             | 4,6      | +  | 4,09      | 0,33   | 0,58 |
| 1                 | 2       | 3                   | 4                    | 5                               | 6        | 7  | 8         | 9      | 10   |

Ce premier résultat indique que, dans le biotope étudié, les algues se répartissent suivant une loi log-normale; il peut être rapproché de celui obtenu par FRONTIER (1973) sur le zooplancton. Cet auteur trouve que la pente de la droite de TAYLOR varie entre 1 (loi de Poisson) pour les populations très faibles et 2 (loi log-normale) pour les populations denses.

Les valeurs intermédiaires traduiraient, d'après FRONTIER « un équilibre dynamique entre les processus biologiques très généraux qui tendent à induire

EAU

I

rés

(Pour une explication détaillée concernant chaque colonne voir dans le texte)

| C    | CX'                 |    | Li                  | Ls                  | P    | Ps   | Pi   | R    | Pe   |
|------|---------------------|----|---------------------|---------------------|------|------|------|------|------|
| 1,46 | 1,7.10 <sup>4</sup> | +  | 2,4.10 <sup>4</sup> | 1,2.10 <sup>4</sup> | 0,37 | 0,44 | 0,30 | 2,06 | 0,34 |
| 1,15 | 8,8.10 <sup>4</sup> | +  | 1,2.10 <sup>5</sup> | 6,4.10 <sup>4</sup> | 0,33 | 0,39 | 0,27 | 1,90 | 0,31 |
| 1,99 | 4,5.10 <sup>4</sup> | +  | 6,5.10 <sup>4</sup> | 3,1.10 <sup>4</sup> | 0,39 | 0,46 | 0,31 | 2,11 | 0,35 |
| 4,94 | 2,6.10 <sup>5</sup> | ±  | 4,6.10 <sup>5</sup> | 1,2.10 <sup>5</sup> | 0,66 | 0,78 | 0,54 | 3,87 | 0,59 |
| 2,49 | 9,9.10 <sup>3</sup> | +  | 1,7.10 <sup>4</sup> | 5,6.10 <sup>3</sup> | 0,59 | 0,75 | 0,43 | 3,07 | 0,51 |
| 1,99 | 2,0.10 <sup>4</sup> | +  | 4,0.10 <sup>4</sup> | 9,7.10 <sup>3</sup> | 0,77 | 1,03 | 0,51 | 4,14 | 0,61 |
| 2,56 | 1,4.10 <sup>4</sup> | +  | 2,2.10 <sup>4</sup> | 9,0.10 <sup>3</sup> | 0,46 | 0,56 | 0,36 | 2,43 | 0,41 |
| 3,86 | 2,6.10 <sup>5</sup> | ±  | 4,5.10 <sup>5</sup> | 1,6.10 <sup>5</sup> | 0,55 | 0,69 | 0,41 | 2,86 | 0,48 |
| 1,56 | 2,7.10 <sup>6</sup> | +  | 3,8.10 <sup>6</sup> | 2,0.10 <sup>6</sup> | 0,33 | 0,38 | 0,28 | 1,91 | 0,31 |
| 10,6 |                     | —  |                     |                     |      |      |      |      |      |
| 1,57 | 5,4.10 <sup>3</sup> | +  | 8,1.10 <sup>3</sup> | 3,7.10 <sup>3</sup> | 0,40 | 0,48 | 0,32 | 2,17 | 0,37 |
| 1,26 | 1,1.10 <sup>4</sup> | +  | 1,7.10 <sup>4</sup> | 7,7.10 <sup>3</sup> | 0,42 | 0,51 | 0,33 | 2,25 | 0,38 |
| 2,01 | 9,7.10 <sup>3</sup> | ±  | 1,3.10 <sup>4</sup> | 7,1.10 <sup>3</sup> | 0,32 | 0,37 | 0,27 | 1,87 | 0,30 |
| 1,17 | 1,2.10 <sup>6</sup> | +  | 1,4.10 <sup>6</sup> | 1,0.10 <sup>6</sup> | 0,16 | 0,16 | 0,15 | 1,36 | 0,15 |
| 3,10 | 7,4.10 <sup>5</sup> | +  | 1,1.20 <sup>6</sup> | 4,6.10 <sup>5</sup> | 0,51 | 0,65 | 0,37 | 2,61 | 0,44 |
| 3,39 | 2,4.10 <sup>4</sup> | ±  | 4,0.10 <sup>4</sup> | 1,4.10 <sup>4</sup> | 0,54 | 0,62 | 0,40 | 2,80 | 0,47 |
| 1,49 | 6,9.10 <sup>3</sup> | +  | 1,0.10 <sup>4</sup> | 4,8.10 <sup>3</sup> | 0,37 | 0,44 | 0,30 | 2,05 | 0,34 |
| 1,25 | 1,4.10 <sup>4</sup> | +  | 2,0.10 <sup>4</sup> | 9,2.10 <sup>3</sup> | 0,41 | 0,50 | 0,32 | 2,20 | 0,37 |
| 1,53 | 9,6.10 <sup>3</sup> | +  | 1,3.10 <sup>4</sup> | 7,0.10 <sup>3</sup> | 0,31 | 0,35 | 0,27 | 1,85 | 0,30 |
| 1,66 | 4,1.10 <sup>4</sup> | +  | 5,6.10 <sup>4</sup> | 3,0.10 <sup>4</sup> | 0,32 | 0,37 | 0,27 | 1,88 | 0,30 |
| 1,78 | 5,7.10 <sup>4</sup> | +  | 8,0.10 <sup>4</sup> | 4,0.10 <sup>4</sup> | 0,35 | 0,41 | 0,29 | 1,98 | 0,33 |
| 2,32 | 2,9.10 <sup>4</sup> | +  | 4,4.10 <sup>4</sup> | 1,9.10 <sup>4</sup> | 0,43 | 0,52 | 0,34 | 2,30 | 0,39 |
| 11   | 12                  | 13 | 14                  | 15                  | 16   | 17   | 18   | 19   | 20   |

des distributions log-normales et une homogénéisation du milieu liée à la diffusion et aux mouvements turbulents ».

Dans le cas de la flore algale des sols de rizière, les populations (colonne n° 3, Tab. I) sont importantes et les facteurs d'homogénéisation du milieu sont très limités : eau stagnante sous une faible épaisseur, effet d'écran au vent et immobilisation des masses algales par les pieds de riz. Une répartition log-normale est donc en accord avec ces conditions.

On peut chercher une explication biologique à la distribution log-normale : CASSIE (1963) remarque que de petites variations des facteurs ambiants induisent très souvent une réponse exponentielle des variables biologiques.

Si les facteurs physico-chimiques du milieu ont des distributions normales, les logarithmes des variables biologiques auront eux aussi une distribution normale.

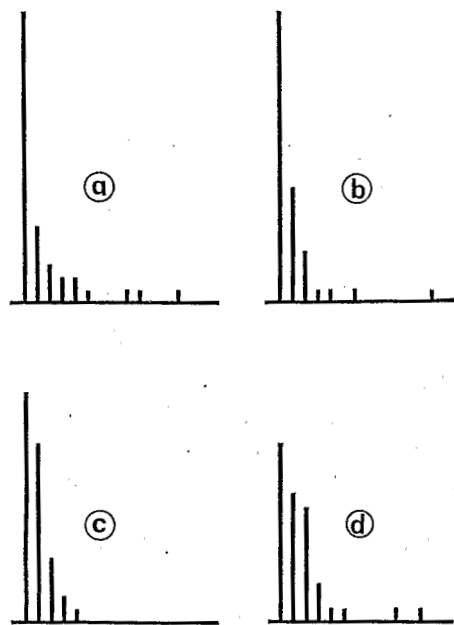


FIG. 1

FIG. 1 (a, b, c, d). — Exemples d'histogrammes de répartition de différentes espèces algales : a : *Navicula* sp. ; b : *Chlorella* sp. ; c : *Oscillatoria* sp. ; d : *Anabaena* sp.

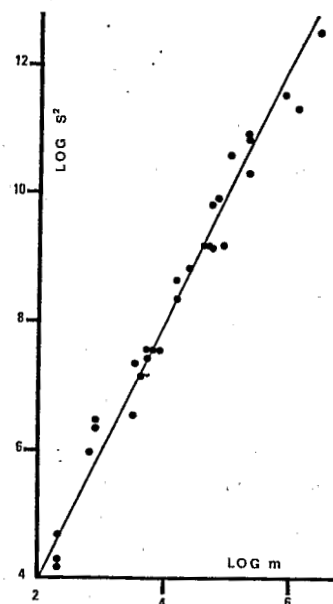


FIG. 2

FIG. 2. — Étude graphique de la liaison moyenne-variance sur les 29 histogrammes réalisés.

Nous avons vérifié cette hypothèse en étudiant les lois de répartition de quelques facteurs chimiques du biotope. Le phosphore total, le fer et l'azote organique ont été dosés sur quarante échantillons de sol prélevés dans chacune des quarante sous-parcelles. On constate (Fig. 3) que les histogrammes sont compatibles avec une distribution normale et dans les trois cas un test graphique de normalité (SCHWARTZ 1963) s'est révélé positif.

#### B) Normalisation des données et calcul de l'intervalle de confiance :

La mise en évidence d'une loi de répartition log-normale, indique que la transformation  $y = \log x$  pourra être utilisée pour normaliser les données si les répartitions observées ne comportent pas de classe zéro.

Ces répartitions ont été éliminées (colonne n° 7, Tab. I) ; elles auraient bien sûr pu être normalisées par la transformation  $y = \log (x + x_0)$  ou  $x_0$  est une constante arbitraire que l'on peut déterminer graphiquement (MERNY et DEJARDIN, 1970).

Toutefois la présence d'une classe zéro signifie principalement que les comptages ont été effectués à une dilution trop élevée. Dans ce cas on constate effectivement que le nombre moyen de colonies comptées par boîte (colonne n° 6, Tab. I) est toujours très bas; la précision des comptages est alors très faible et l'on a intérêt à reprendre la numération à une dilution inférieure si cela est possible.

La transformation  $y = \log x$  a été appliquée aux distributions ne comportant pas de classe zéro et nous en avons vérifié la validité.

Une première vérification graphique consiste à s'assurer que les points ( $x$  = centre des classes;  $y$  = fréquence cumulée relative) sont alignés sur un graphique en échelle arithmétique-gaussienne (SCHWARTZ 1963) cf. § III. E.

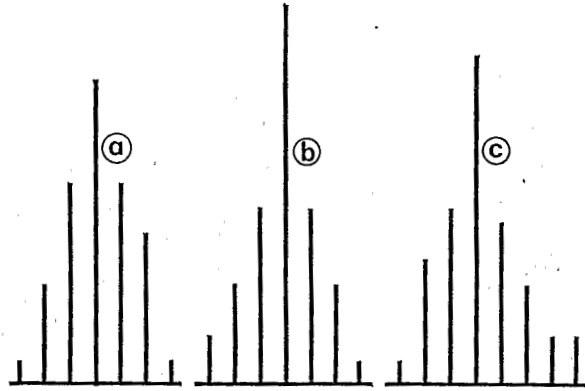


FIG. 3. — Exemples d'histogrammes de répartition de variables chimiques du milieu : a : azote organique; b : fer ; c : phosphore total.

Si cette vérification graphique donne un résultat positif on pourra poursuivre les calculs et l'on déterminera pour la nouvelle variable  $y = \log x$  les valeurs suivantes :

- la moyenne  $\bar{y}$  (colonne n° 8, Tab. I);
  - la variance  $s_y^2$  (colonne n° 9, Tab. I);
  - l'écart type  $s_y$  (colonne n° 10, Tab. I)
- et les deux limites de confiance inférieure ( $li$ ) et supérieure ( $ls$ ) de cette moyenne :

$$li = \bar{y} - \frac{ts_y}{\sqrt{n}} \quad ls = \bar{y} + \frac{ts_y}{\sqrt{n}}$$

( $n$  = nombre d'échantillons;  $t$  = variable de STUDENT à  $n-1$  degrés de liberté, ddl, au seuil de précision choisi).

La retransformation simple de  $\bar{y}$ ,  $li$  et  $ls$  conduit à

$$\begin{aligned} \bar{x}' &= 10^{\bar{y}} \\ Li' &= 10^{li} \\ Ls' &= 10^{ls} \end{aligned}$$

qui sont des estimations biaisées de la moyenne  $\bar{x}$  (connue) et des deux limites de confiance de cette moyenne :  $Li$  et  $Ls$

$$\left( \frac{\sum_1^n \log x_1}{n} \neq \log \frac{\sum_1^n x_1}{n} \right)$$

Le problème de la correction de ce biais a été étudié par NEYMAN et SCOTT (1960) : des valeurs non biaisées de  $\bar{x}$ ,  $L_i$  et  $L_s$  sont obtenues en multipliant les retransformées simples par un coefficient de correction C dont l'expression dans le cas de la transformation logarithmique est :

$$C = \sum_{i=0}^{\infty} \left[ \frac{(\ln 10)^2}{2} \cdot \frac{n-1}{n} \cdot S_y^2 \right]^i \cdot \frac{1}{i!} \prod_{k=1}^i \left( 1 + 2 \frac{k-1}{n-1} \right)$$

Le calcul de C peut se faire de la façon suivante : (MERNY et DEJARDIN, 1970).

1) Calculer :

$$\frac{2,65095 (n-1) S_y^2}{n} = A$$

n est la taille de l'échantillon.

$S_y^2$  est la somme des carrés des écarts de y (numérateur de  $S_y^2$ ,  $2,65095 = \frac{1}{2} (\ln 10)^2$ ).

2) Calculer les termes successifs

$$t_{i+1} = \frac{A}{(i+1)(2i+n-1)} t_i$$

Pour  $i = 0, 1, 2, \dots$  avec  $t_0 = 1$  jusqu'à ce que le dernier terme calculé soit suffisamment petit.

3) Le coefficient C est égal à la somme des t. Un exemple de ce calcul est donné au paragraphe III E.

C peut-être calculé dans chaque cas (colonne n° 11, Tab. I) toutefois si n reste constant, une détermination graphique au moyen d'une abaque  $C = f(s_y)$  est d'un emploi plus facile.

La validité de la transformation, vérifiée une première fois par le test graphique est confirmée si la retransformée corrigée  $C\bar{x}' = C \cdot 10^y$  est peu différente de  $\bar{x}$ .

On constate, dans le cas des répartitions observées, que les valeurs  $C\bar{x}'$  calculées (colonne n° 12, Tab. I) diffèrent peu des valeurs originelles  $\bar{x}$ .

Nous avons considéré la transformation comme pleinement justifiée (+) si :

$$\frac{|\bar{x} - C\bar{x}'|}{\bar{x}} < 0,2$$

et comme acceptable ( $\pm$ ) si :

$$0,2 \leq \frac{|\bar{x} - C\bar{x}'|}{\bar{x}} < 0,33$$

cf. colonne n° 13, Tab. I.

C) Évaluation de la précision :

La précision P peut se définir comme une fraction de la moyenne qui ajoutée ou soustraite à celle-ci établit les limites de son intervalle de confiance ; autrement dit, la moyenne est estimée à 100 P % près.

En écrivant que la différence entre Li et Ls vaut 2 P  $\bar{x}$  et en estimant respectivement ces trois valeurs par CL'i, CL's et C $\bar{X}$  nous avons pu calculer P (colonnes 14, 15, 16, Tab. I)

En exprimant Li et Ls en fonction de sy et n on établit l'équation (1) :

$$P = \frac{1}{2} \left[ \text{antilog} \frac{tSy}{\sqrt{n}} - \frac{1}{\text{antilog} \frac{tSy}{\sqrt{n}}} \right]$$

qui permet de calculer la précision attendue en fonction du nombre d'échantillons et de l'écart type de la variable transformée.

La figure n° 4 donne une représentation graphique de ce calcul pour des écarts types allant de 0,1 à 1,5 et des nombres d'échantillons égaux respectivement à 16 - 36 - 64 - 100 - 144.

A partir de l'équation (1) on peut exprimer n en fonction de P et l'on obtient :

$$n_p = \frac{t^2 s_y^2}{[\log (P + \sqrt{P^2 + 1})]^2} \quad (2)$$

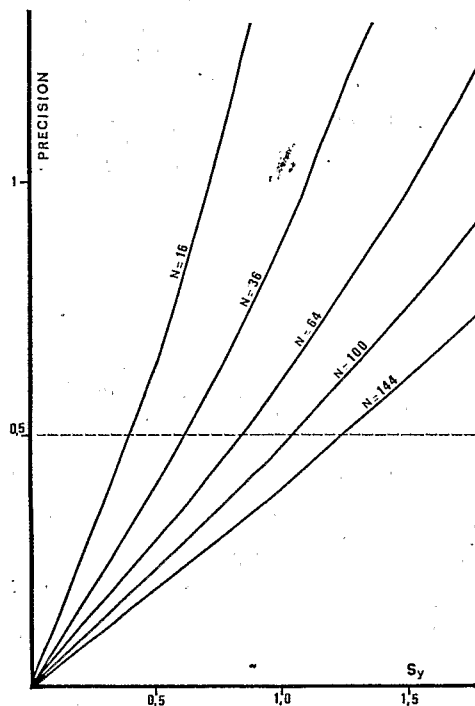


FIG. 4

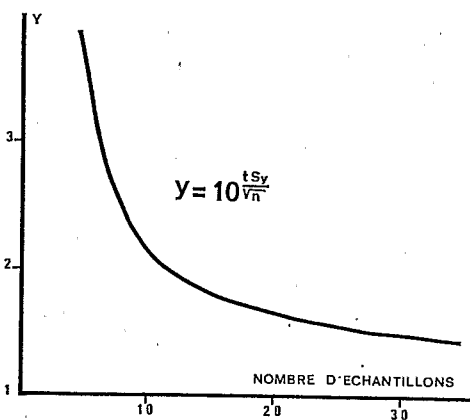


FIG. 5

FIG. 4. — Abaqués de détermination de la précision P en fonction du nombre d'échantillons n et de l'écart type des données transformées sy.

FIG. 5. — Graphe de la fonction  $y = \text{antilog} \frac{tSy}{\sqrt{n}}$  pour  $s_y = 0,47 = \bar{s}_y$ . Cette fonction indique la dissymétrie de l'intervalle de confiance en fonction du nombre d'échantillons.

Cette équation permet de calculer le nombre d'échantillons nécessaires pour obtenir une précision donnée en fonction de l'écart type des données transformées. On peut obtenir une solution graphique de ce calcul par extrapolation (Fig. 4).

Toutefois il convient de remarquer que la précision définie par  $P = \frac{Ls - Li}{2\bar{x}}$  laisse supposer l'existence d'un intervalle de confiance symétrique par rapport à  $\bar{x}$ . Or dans le cas d'une distribution log-normale on constate une dissymétrie de l'intervalle de confiance comme le montre le calcul du rapport entre l'écart supérieur et l'écart inférieur.

$$\frac{Ls - \bar{x}}{\bar{x} - Li} = 10 \frac{tSy}{\sqrt{n}} \quad (3)$$

La figure 5 indique les variations de  $y = 10 \frac{tSy}{\sqrt{n}}$  pour  $s = s_y = 0,47$  et montre que la dissymétrie devient particulièrement marquée pour  $n < 10$ .

Les équations (1), (2) et (3) permettent de calculer :

$$P = \frac{x'_s - x'_i}{2\bar{x}'} = \text{précision moyenne ;}$$

$$Ps = \frac{x'_s - \bar{x}'}{\bar{x}'} = \text{précision calculée sur l'écart supérieur ;}$$

$$Pi = \frac{\bar{x}' - x'_i}{\bar{x}'} = \text{précision calculée sur l'écart inférieur.}$$

Ces valeurs calculées pour les répartitions observées sont reportées aux colonnes n°s 16 - 17 - 18 du Tab. I.

Un autre mode de raisonnement consiste à faire intervenir la dissymétrie de l'intervalle de confiance dans le calcul de la précision.

Une numération ne présente que rarement de l'intérêt en tant que résultat isolé ; le problème le plus fréquemment rencontré est de savoir si deux numérations sont significativement différentes l'une de l'autre.

Considérons deux résultats  $\bar{x}_1$  et  $\bar{x}_2$  tels que  $\bar{x}_1 < \bar{x}_2$  et supposons  $n$  identique ; on peut admettre que ces deux résultats seront significativement différents si la limite inférieure de  $\bar{x}_2$  est supérieure à la limite supérieure de  $\bar{x}_1$ , soit :  $\bar{x}_1 (1 + Ps_1) < \bar{x}_2 (1 - Pi_2)$  (4)

Si  $\bar{x}_1$  et  $\bar{x}_2$  concernent le même organisme et si  $Sy_1 = Sy_2$  on démontre que (4) est réalisé si  $\bar{x}_2/\bar{x}_1 > \text{antilog} \frac{2ts_y}{\sqrt{n}}$  (5)

Posons  $R =$  valeur limite de  $\frac{\bar{x}_2}{\bar{x}_1}$  pour laquelle les valeurs  $\bar{x}_2$  et  $\bar{x}_1$  sont significativement différentes.

Cette valeur peut être calculée soit en fonction de  $P$  soit en fonction de  $Ps$  et  $Pi$  :

$$R_{(Pi, Ps)} = \frac{1 + Ps}{1 - Pi} \quad R_{(P)} = \frac{1 + P}{1 - P}$$

Sachant que  $P = \frac{Pi + Ps}{2}$  et que  $Ps > Pi$

On démontre :

$$\frac{R_{(P)}}{R_{(P_s, P_i)}} = \frac{2 + P_s + P_i - P_s P_i - P_i^2}{2 + P_s + P_i - P_s P_i - P_s^2} < 1$$

soit  $R_{(P)} > R_{(P_s, P_i)}$

La valeur P est une estimation de la précision qui ne tient pas compte de la dissymétrie de l'intervalle de confiance et conduit à le surestimer lorsque l'on compare deux mesures.

Pour cette raison nous avons défini la précision équivalente  $P_e$  telle que

$$R_{(P_s, P_i)} = \frac{1 + P_s}{1 - P_i} = R_{(P_e)} = \frac{1 + P_e}{1 - P_e} \quad (6)$$

La notion de précision équivalente présente l'avantage de tenir compte de la dissymétrie de l'intervalle de confiance et d'être utilisable directement dans les équations donnant la précision en fonction du nombre de prélèvements et de l'écart-type des données transformées ou inversement.

On démontre que :

$$P_e = \frac{\text{antilog} \frac{2ts_y}{\sqrt{n}} - 1}{\text{antilog} \frac{2ts_y}{\sqrt{n}} + 1} \quad (7)$$

En exprimant (7) par rapport à n on obtient :

$$n = \frac{4 t^2 s_y^2}{\left( \log \frac{1 + P_e}{1 - P_e} \right)^2} \quad (8)$$

qui permet de calculer le nombre d'échantillons nécessaires pour obtenir une précision équivalente donnée.

Le calcul de  $P_e$  pour les répartitions observées (Tab. I, colonne n° 20) confirme  $P_e < P$ .

#### D) Généralisation des résultats :

Les équations précédentes font intervenir, en plus du nombre d'échantillons et de la précision, l'écart type de la variable transformée :  $s_y$ .

Nous avons cherché à généraliser les équations précédentes en assignant à  $s_y$  une valeur constante calculée à partir des données expérimentales (colonne n° 10, Tab. 1). Les 20 valeurs de  $s_y$  calculées se répartissent approximativement suivant une loi normale caractérisée par :  $\bar{s}_y = 0,477 =$  moyenne  
 $s^2_{s_y} = 0,0245 =$  variance

la valeur  $\bar{s}_y$  ne permet que la détermination d'une « précision moyenne » en fonction du nombre d'échantillons.

Par contre l'utilisation de la table  $\int_0^t \Phi(t) dt$  permet de déterminer au risque de 5 %, une valeur  $s_y$  limite qui ne sera dépassée que par 100 - k % des valeurs de  $s_y$ . En fixant k arbitrairement on pourra donc déterminer le nombre de prélèvements à effectuer pour que 100 - k % des résultats aient une précision au plus égale à P ou  $P_e$ .

En prenant k = 5 % on trouve  $s_y$  limite = 0,73.

La figure n° 6 indique la précision équivalente en fonction du nombre d'échantillons pour  $s_y = s_y \text{ lim } 95 \%$ .

TABLEAU II  
Valeurs des numérations utilisées pour l'exemple de calcul (cf. § III E).

|     |    |    |    |    |    |    |     |
|-----|----|----|----|----|----|----|-----|
| 3,3 | 7  | 5  | 15 | 43 | 23 | 53 | 27  |
| 2,3 | 8  | 1  | 13 | 40 | 63 | 20 | 53  |
| 1,7 | 3  | 2  | 1  | 83 | 5  | 13 | 193 |
| 1,2 | 2  | 17 | 23 | 4  | 47 | 1  | 6   |
| 6   | 23 | 7  | 27 | 43 | 50 | 2  | 227 |

(Toutes ces valeurs sont à multiplier par  $10^4$ )

È) Exemple de calculs complets pour une répartition observée :

A partir des quarante valeurs observées de  $x$  (Tab. II) on calcule :

$$T1_x = \sum x = 159,2 \cdot 10^4$$

$$T2_x = \sum x^2 = 1453,85 \cdot 10^8$$

$$\bar{x} = \frac{\sum x}{n} = 3,98 \cdot 10^4$$

$$s_x^2 = \frac{1}{n-1} \left( T2 - \frac{T1^2}{n} \right) = 21,03 \cdot 10^8$$

$$s_x = \sqrt{s_x^2} = 4,59 \cdot 10^4$$

$$\frac{\log s_x^2}{\log \bar{x}} = 2,026$$

L'histogramme de répartition (Fig. 7) permet de rejeter l'hypothèse de normalité des résultats.

Les données  $x$  sont remplacées par leur transformée  $y = \log x$  et l'on détermine :

$$T1_y = \sum y = 175,648$$

$$T2_y = \sum y^2 = 30852,219$$

$$\bar{y} = \frac{\sum y}{n} = 4,3912$$

$$s_y^2 = 0,198$$

$$s_y = 0,44$$

$$\frac{t_{sy}}{\sqrt{n}} = 0,1362 \text{ (t à 95 \% et 39 ddl)}$$

$$\bar{y} = 4,3912 \pm 0,1362 \begin{cases} y_s = 4,5274 \\ y_i = 4,2550 \end{cases}$$

On vérifie que l'histogramme des données transformées est compatible avec une distribution normale et l'on effectue un test graphique de normalité — (Fig. 7) qui se révèle satisfaisant.

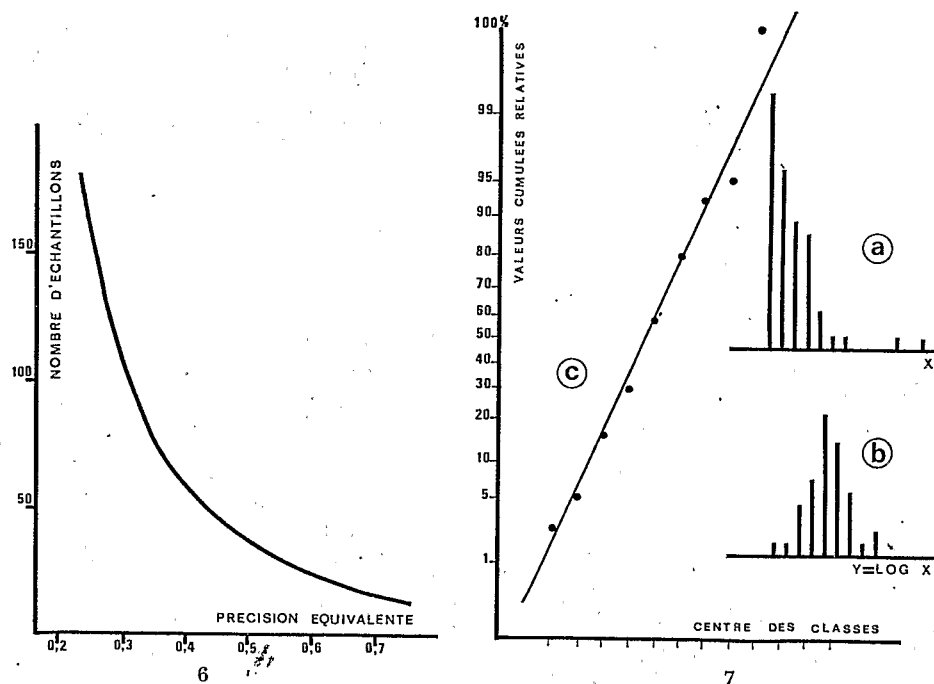


FIG. 6. — Valeurs de la précision équivalente  $P_e$  en fonction du nombre d'échantillons  $n$  pour  $s_v = 0,73 = s_v$  limite.

FIG. 7. — Histogrammes de répartition des valeurs de l'exemple (§ III, E) avant (a) et après (b) transformation logarithmiques; test graphique de normalité (c) des données transformées.

Calcul du coefficient correcteur C :

$$A = \frac{2,65095 \times 39 \times 0,198 \times 39}{40} = 19,9589$$

|         |  |           |
|---------|--|-----------|
| $t = 0$ |  | 1         |
| $t = 1$ | $\frac{19,9589 \times 1}{1 \times 39}$       | = 0,51176 |
| $t = 2$ | $\frac{19,9589 \times 0,51176}{2 \times 41}$ | = 0,12456 |
| $t = 3$ | $\frac{19,9589 \times 0,12456}{3 \times 43}$ | = 0,01927 |
| $t = 4$ | $\frac{19,9589 \times 0,01927}{4 \times 45}$ | = 0,00214 |
| $t = 5$ | $\frac{19,9589 \times 0,00214}{5 \times 47}$ | = 0,00018 |

$t = 5$  est suffisamment petit  $C = 1,6579$ .

A partir des valeurs  $y$ ,  $y_s$ ,  $y_i$  on calcule les retransformées  $10^y$ ,  $10^{y_s}$ ,  $10^{y_i}$  qui sont multipliées par  $C$  pour obtenir :

$$\bar{x}' = C \cdot 10^{\bar{y}} = 4,08 \cdot 10^4$$

$$L's = C \cdot 10^{y_s} = 5,59 \cdot 10^4$$

$$L'i = C \cdot 10^{y_i} = 2,99 \cdot 10^4$$

On constate que  $\bar{x}'$  est peu différent de  $\bar{x}$  ( $3,98 \cdot 10^4$ ), et que la transformation est justifiée.

Calcul de la précision :

$$P = \frac{5,59 \cdot 10^4 - 2,99 \cdot 10^4}{2 \times 4,08 \cdot 10^4} = 0,32$$

$$P_s = \frac{5,59 \cdot 10^4 - 4,08 \cdot 10^4}{4,08 \cdot 10^4} = 0,37$$

$$P_i = \frac{4,08 \cdot 10^4 - 2,99 \cdot 10^4}{4,08 \cdot 10^4} = 0,27$$

Calcul de la précision équivalente :

$$R_{(P_s, P_i)} = \frac{1 + 0,37}{1 - 0,27} = 1,88$$

$$R_P = \frac{1 + 0,32}{1 - 0,32} = 1,94$$

On vérifie  $R_{(P)} > R_{(P_s, P_i)}$

$$\text{Posons } \frac{1 + P_e}{1 - P_e} = R_{(P_s, P_i)} = 1,88$$

$$P_e = 0,3$$

Dans cet exemple la différence est peu marquée entre  $P$  et  $P_e$  par suite d'une valeur relativement faible de  $\frac{t_{sy}}{\sqrt{n}}$  due au fait que  $n$  est élevé et  $s_y$  bas (cf. équation (3) et Fig. 5).

#### IV. — CONCLUSION

L'étude du mode de répartition des algues en sol submergé a montré que la distribution de ces organismes suivait approximativement une loi log-normale.

La transformation  $y = \log x$  permet de normaliser les données et de calculer un intervalle de confiance pour la moyenne de  $n$  numérations. A partir de ces données nous avons pu déterminer les équations donnant la précision des mesures en fonction du nombre d'échantillons et de l'écart type de la variable transformée.

Deux formules de la précision ont été fournies ; la première, classique, ne tient compte que de la différence entre la limite inférieure de la moyenne, la

seconde tient compte de la dissymétrie entre l'écart supérieur et l'écart inférieur.

L'étude de la distribution des écarts types des données transformées a permis de montrer que cette valeur se répartit approximativement suivant une loi normale. Il est donc possible de calculer une valeur limite qui ne sera dépassée que par 5 % des valeurs de  $s_y$  et, en reportant cette valeur dans les formules établies, de calculer une valeur seuil de la précision qui n'est fonction que du nombre d'échantillons.

Les résultats obtenus peuvent être utilisés pratiquement de deux façons :

- soit pour déterminer le nombre  $n$  d'échantillons à prélever et à utiliser pour effectuer les numérations dont la moyenne aura une précision au plus égale à  $P$  ;
- soit pour déterminer le nombre  $n$  d'échantillons à prélever et à mélanger entre eux pour obtenir un échantillon moyen représentatif de la parcelle avec une précision au plus égale à  $P$ .

#### RÉSUMÉ

L'étude de la flore algale dans une rizière de la zone sahélienne a montré que la distribution de ces organismes suit approximativement une loi log-normale et que la transformation logarithmique permet de normaliser les données.

Il devient alors possible de calculer un intervalle de confiance pour la moyenne de  $n$  numération. Nous avons déterminé les équations donnant la précision des moyennes en fonction du nombre d'échantillons et de l'écart-type de la variable transformée.

Deux expressions de la précision ont été calculées ; la première, classique, ne tient compte que de la différence entre les limites inférieures et supérieures de la moyenne, la seconde tient compte de la dissymétrie entre l'écart supérieur et l'écart inférieur.

Ces résultats permettent de déterminer le nombre  $n$  d'échantillons à prélever pour effectuer des numérations dont la moyenne aura une précision donnée ; ils permettent aussi de calculer la représentativité d'un échantillon moyen obtenu en mélangeant  $n$  sous-échantillons.

#### SUMMARY

Study of the algae in a paddy field in Senegal indicated that the logarithms of numbers of these organisms in core samples (3,14 cm<sup>2</sup> in section, including the submersion water and the first centimeter of soil) were normally distributed.

Two mathematical formulas were developed showing the relation between accuracy of the mean of counts and the number of samples. The first takes into account only the upper and the lower limits of the mean ; the second one utilises these limits and, in addition, the dissymetry of the confidence interval.

## BIBLIOGRAPHIE

- CASSIE (R. M.), 1963. — Microdistribution of plankton. *Oceanogr. Mar. Biol. Ann. Rev.*, 1: 223-252.
- FRONTIER (S.), 1973. — Étude statistique de la dispersion du zooplancton. *J. exp. mar. Biol. Écol.*, 12: 229-262.
- KENDALL (M. G.) & STUART (A.), 1963-68. — The advanced theory of statistics. 3 vols. Griffin London, 1: 232, 3: 90.
- MERNY (G.) & DEJARDIN (J.), 1970. — Les nématodes phytoparasites des rizières inondées de Côte-d'Ivoire : II Essai d'estimation de l'importance des populations. *Cah. ORSTOM*, sér. Biol., 11: 45-67.
- NEYMAN (J.) & SCOTT (E.), 1960. — Correction for bias introduced by a transformation of variables. *Ann. Math. Stat.*, 31: 643-655.
- REYNAUD (P. A.) & ROGER (P. A.), 1976. — N<sub>2</sub> Fixing Algal Biomass in Senegal Rice Fields. Proc. of the Int. Symp « Environmental rôle of N<sub>2</sub>-fixing bluegreen Algae and Asymbiotic Bacteria ». *Ecol. Bull* (Stockholm), 26: 148-157.
- REYNAUD (P. A.) & ROGER (P. A.), 1977. — Milieux sélectifs pour la numération des Algues Eucaryotes, Procaryotes et fixatrices d'azote. *Rev. Écol. Biol. Sol*, 14: 421-428.
- ROGER (P. A.) & REYNAUD (P. A.), 1976. — Dynamique de la population algale au cours d'un cycle de culture dans une rizière sahélienne. *Rev. Écol. Biol. Sol.*, 13: 545-560.
- SCHWARTZ (D.), 1963. — Méthodes statistiques à l'usage des médecins et des biologistes. Ed. Méd. Flammarion, Paris, 289 p.
- TAYLOR (R. L.), 1961. — Aggregation, variance and the mean. *Nature*, London, 189: 732-735.