

Une procédure systématique pour le sous-échantillonnage des prélèvements de plancton

Plancton
Sous-échantillonnage
Effort
Précision
Plancton
Sub-sampling
Effort
Precision

Constantin KOUTSIKOPOULOS ^a, Pierre PETITGAS ^b

^a IFREMER, BP 1049, 44037 Nantes Cedex.

^b Centre de Géostatistiques, 35, rue Saint-Honoré, 77305 Fontainebleau.

Reçu le 22/2/89, révisé le 23/8/89, accepté le 4/9/89.

RÉSUMÉ

Sous-échantillonner les prélèvements de plancton est souvent une nécessité même si la précision de l'estimation s'en trouve diminuée. Une optimisation méthodologique est recherchée. La variance est fonction de la procédure de construction des sous-échantillons (plan) et du nombre d'individus comptés par sous-échantillon. L'effort attribué au sous-échantillonnage est déterminé à la suite d'une stratégie de répartition de l'effort total entre échantillonnage et sous-échantillonnage. L'effort de sous-échantillonnage est défini par le temps nécessaire pour la procédure de fractionnement, le volume examiné et le nombre d'organismes comptés. Pour un effort donné, une procédure systématique est plus précise qu'une procédure aléatoire puisqu'elle parcourt mieux l'échantillon et elle est ainsi moins influencée par la répartition spatiale des organismes à l'intérieur de la cuve de fractionnement. Il est suggéré de calculer la variance du sous-échantillonnage en routine car les modèles théoriques ne sont pas toujours applicables, surtout pour les organismes abondants.

Oceanologica Acta, 1990, 13, 3, 403-409.

ABSTRACT

A systematic procedure for sub-sampling plankton samples

It is often necessary to sub-sample plankton samples even if the precision of the estimation is thereby decreased. A methodological optimization is looked for. The variance depends on the procedure of construction of the aliquot (strategy) and the number of individuals counted (effort). The sub-sampling effort is allocated in a general context of repartition of the total effort between sampling and sub-sampling. This effort depends on the time consumed by the sub-sampling procedure, the volume of the aliquot and the number of individuals counted. For the same effort, a systematic procedure for the construction of sub-samples is more precise than a random one because it travels through the entire sample and is thus less dependent on the spatial distribution of the organisms in the sub-sampling tank. The calculation of the sub-sampling variance in routine is suggested because deviations from the theoretical models are frequent, especially for the high levels of abundance.

Oceanologica Acta, 1990, 13, 3, 403-409.

INTRODUCTION

Dans les études quantitatives en planctonologie, la variance de l'estimateur d'abondance est habituellement partitionnée en trois composantes (Platt *et al.*, 1970) selon le modèle de l'analyse de variance suivant : $\sigma_{\text{tot}}^2 = \sigma_{\text{st}}^2 + \sigma_r^2 + \sigma_{\text{se}}^2$ où σ_{st}^2 (variance entre stations) représente la variabilité de l'abondance due à sa « régionalisation » (structures spatiales) à l'échelle spatio-temporelle de la campagne, σ_r^2 (variance entre répliquats) représente la variabilité due aux structures fines et aux conditions d'échantillonnage, σ_{se}^2 (variance entre sous-échantillons) représente la variabilité introduite

par le sous-échantillonnage. En effet, il est difficile de conclure sur la différence entre répliquats sans connaître la contribution de la variabilité de sous-échantillonnage à la variance des répliquats. Plus généralement, les fluctuations entre sous-échantillons et répliquats peuvent masquer les différences dues aux décalages spatiaux et temporels entre les prélèvements qui sont d'intérêt biologique.

σ_{tot}^2 est une fonction de l'effort global consenti (nombre de stations, volumes filtrés, fractions dépouillées, nombres comptés) et des plans adoptés pour l'échantillonnage (répartition des stations) et le sous-échantillonnage (construction d'un sous-échantillon). Pour l'effort

consenti, Venrick (1971) montre que l'optimisation de σ_{tot}^2 relève d'une stratégie globale de répartition de cet effort entre échantillonnage et sous-échantillonnage sans les dissocier. Les plans adoptés pour l'échantillonnage et le sous-échantillonnage peuvent être différents. La variance de sous-échantillonnage s'ajoute toujours aux autres sources de variabilité. Il est donc intéressant de l'optimiser.

Ainsi de nombreux auteurs ont proposé des appareils à sous-échantillonner le plancton. Curieusement la plupart des appareils sont conçus sans vraiment tenir compte de l'optimisation de la stratégie de construction du sous-échantillon. Wiborg a fait un inventaire des engins et des méthodes existant en 1962. Van Guelpen *et al.* (1982) ont comparé les méthodes les plus répandues en se souciant de leur précision et du temps nécessaire à leur mise en œuvre. Depuis, d'autres appareils ont été proposés (George *et al.*, 1984). Mis à part les tests qui accompagnent les descriptions techniques des appareils, leur efficacité a été examinée sur les plans théorique et expérimental (Horwood et Driver, 1976; Dahiya, 1980; Sell et Evans, 1982; Griffiths *et al.*, 1984). Les conclusions convergent vers la génération d'erreurs imprévisibles et pas toujours négligeables. En effet, malgré une « homogénéisation » de l'échantillon dans l'appareil, subsistent des structures agrégées dont nous montrerons qu'elles ne suivent pas une loi théorique classique de répartition (binomiale-Poisson). La connaissance de σ_{se}^2 étant le plus souvent nécessaire, il faut donc estimer σ_{se}^2 à chaque fois en routine en examinant au moins deux sous-échantillons. De nombreux engins, basés sur le principe dichotomique, ne permettent pas ce calcul car les deux sous-échantillons construits ne sont pas indépendants; de plus la fraction finale est forcée à une puissance de 2 ($1/2^k$ de l'échantillon) ce qui limite la souplesse d'utilisation. Mais fondamentalement la plupart des appareils construisent des échantillons selon un plan aléatoire simple; ainsi, ne différant pas sensiblement entre eux, ils ne minimisent pas l'influence des agrégats sur σ_{se}^2 .

La théorie d'échantillonnage (Cochran, 1977) établit qu'en l'absence de structures spatiales à caractère périodique un plan d'échantillonnage systématique fournit des estimations plus précises qu'une procédure aléatoire. Aucun déterminisme physique ne permet de penser qu'un échantillon de plancton dans la cuve de fractionnement est caractérisé par de telles structures. Nous avons donc entrepris de construire une cuve, le STWAG (Sub-sampling Tank With a Grid), véritable système expérimental qui nous a permis, sur la base de dénombrements, de comparer les procédures de construction de sous-échantillons. En outre, cet appareil se révèle simple et pratique.

MATÉRIEL ET MÉTHODES

L'appareil a été construit pour satisfaire les conditions suivantes :

- obtenir directement 2^k fractions de l'échantillon original, contrairement aux processus fréquemment utilisés qui nécessitent, pour obtenir le même nombre de fractions, 2^{k-1} répétitions de la procédure dichotomique,

- permettre l'utilisation d'une procédure systématique pour la construction de sous-échantillons;
- être compatible avec les échantillons récoltés par un échantillonneur type Bongo (61 cm de diamètre et vide de maille de 500 μm) durant les surveillances ichtyoplanctoniques. Le volume de plancton varie entre 200 et 400 ml, conservé habituellement dans des bocaux de 2 l.

L'appareil est composé de deux parties (fig. 1) :

- une cuve en PVC (5 mm d'épaisseur); dimension $L \times l \times h = 485 \times 245 \times 30$ mm;
- une grille en plastique; dimension des cases 30 \times 30 mm, épaisseur des séparations 0,5 mm; 128 cases (8 \times 16).

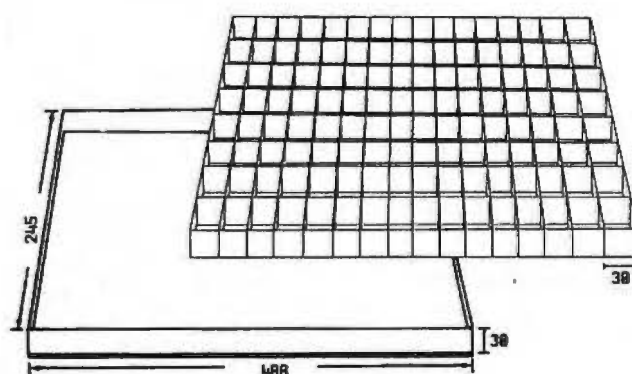


Figure 1
Plans de l'appareil.
Experimental apparatus.

L'échantillon est versé dans la cuve et il est mélangé manuellement. Le fractionnement est réalisé en poussant la grille à l'intérieur de la cuve. Nous avons construit les sous-échantillons en récupérant le contenu des cases à l'aide d'une pipette. Durant le prélèvement, les organismes ne passent pas sous la grille pour contaminer les cases contiguës.

Les données pour l'étude proviennent d'un échantillon commun récolté dans le golfe de Gascogne (47°N, 3°10'W, 20 mars 1985). Nous avons procédé sur la totalité de l'échantillon au dénombrement, case par case, des larves de sole *Solea solea* L. (N = 49), des œufs de sole (N = 302) et de l'ensemble des larves (N = 600) et des œufs des poissons restants (N = 1288).

Cette opération a été répétée trois fois. Avant chaque opération l'échantillon a été mélangé manuellement. Après le dénombrement total, nous avons compté le nombre d'organismes coincés sous la grille. Ce nombre représente environ 2,5 % du nombre total (tab. 1) et il varie selon la forme des organismes. De ce fait, la proportion du fond de la cuve occupée par les surfaces de séparation doit être minimisée. Pour chaque type d'organisme la moyenne N sur les trois dénombrements, reflète l'abondance totale de l'échantillon.

Pour la construction d'un sous-échantillon représentant une fraction F de l'échantillon initial, contenant K cases parmi les 128 de la grille, trois méthodes ont été utilisées. La première ignore la division de l'échantillon en cases élémentaires considérant qu'il est divisé en F

Tableau 1

Résultats des comptages montrant le pourcentage d'organismes piégés sous la grille pour chaque expérience. Les nombres d'organismes comptés et piégés sont respectivement Nc et Np. La moyenne arithmétique N des Nc reflète l'abondance totale pour chaque type d'organisme.

Results showing the percentage of organisms trapped under the grid for each experiment. Nc and Np correspond to the total number counted and trapped respectively. N is the mean of Nc and represents the level of abundance of each taxon.

Nc	Np	Np/Nc (%)	Nc	Np	Np/Nc (%)
50	0	0	309	2	1
47	2	4	306	0	0
49	2	4	291	7	2.5
N = 49			N = 302		
Nc	Np	Np/Nc (%)	Nc	Np	Np/Nc (%)
603	15	2.5	1349	26	2
611	12	2	1216	54	4.5
585	14	2.5	1299	24	2
N = 600			N = 1288		

blocs chacun constitué de K cases contiguës (F*K = 128). Chaque bloc représente un sous-échantillon. Autrement dit on considère que l'organisation de l'échantillon se limite à une grille de F éléments représentant le K/128 de l'échantillon. La procédure consiste au tirage au hasard d'un bloc parmi les F de la population. Cette méthode correspond à un échantillonnage aléatoire simple d'un bloc et peut être comparée aux résultats des processus dichotomiques ou à tout autre processus qui considère l'échantillon divisé en F unités élémentaires. La figure 2 montre cette procédure et l'équivalence pour un processus dichotomique. La fraction a représente 1/8 de l'échantillon (K = 16, F = 8).

L'autre méthode repose sur le principe du sous-échantillonnage systématique. Pour le même effort (K éléments) nous considérons la grille divisée en K blocs contigus de F cases (F*K = 128). Ensuite une case (sa position) est choisie à l'intérieur d'un bloc au hasard. Cette case et les cases correspondantes (mêmes coordonnées par rapport au bloc) sur l'ensemble de K blocs déterminent le sous-échantillon de K cases. La somme des b (fig. 2) correspond à 1/8 de l'échantillon (K = 16, F = 8). Il est évident que cette deuxième procédure permet de mieux parcourir l'échantillon et il est irréaliste, sur le plan manipulation, de tenter d'obtenir

le même résultat avec une procédure dichotomique. La troisième méthode est utilisée uniquement à titre de comparaison et elle consiste, pour la même fraction F, au tirage aléatoire de K cases parmi les 128 de la grille. La différence entre les deux approches est évidente. La première ignore l'organisation de la structure en cases élémentaires, la deuxième exploite cette organisation pour mieux répartir l'effort sur l'ensemble de l'échantillon. Pour mieux illustrer les phénomènes qui influencent la précision de l'estimation du nombre total d'organismes dans l'échantillon, la notion de la case élémentaire est conservée pour les trois méthodes. Ainsi l'opération peut être comparée à un échantillonnage par grappe. La population (l'échantillon dans la cuve) étant divisé en F unités primaires (grappes) de K éléments.

ANALYSE ET RÉSULTATS

La répartition des planctontes dans la cuve est au centre des difficultés concernant la minimisation de la variance du sous-échantillonnage (σ_{se}^2). La figure 3,

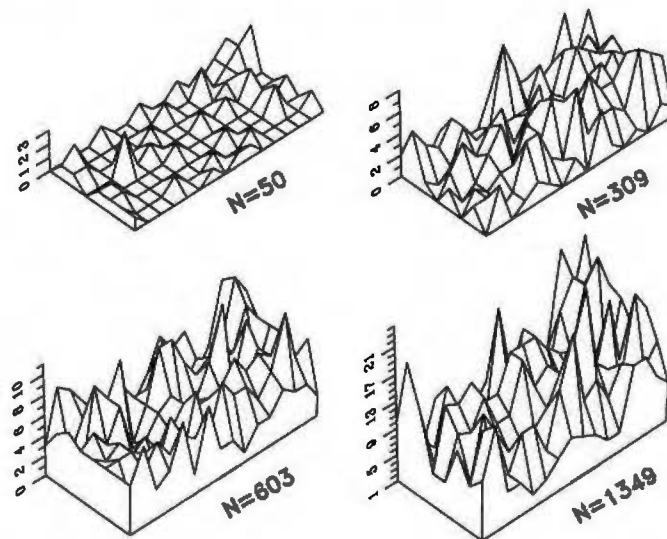


Figure 3
Illustration de la répartition spatiale des organismes dans la cuve de fractionnement durant la procédure de sous-échantillonnage.
Illustration of the spatial distribution of plankton in the sub-sampling tank.

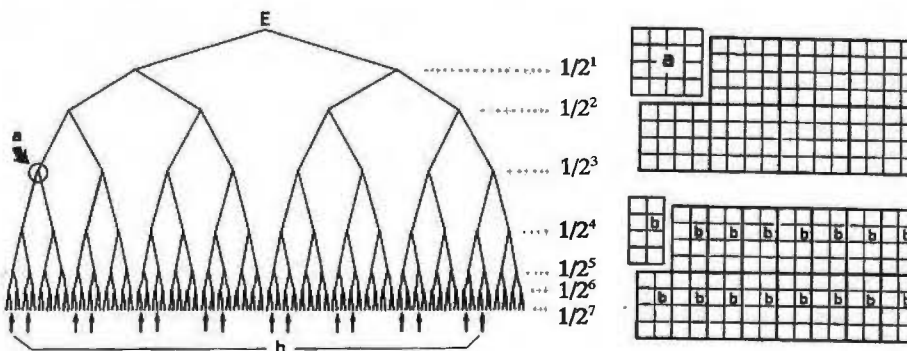


Figure 2
Procédure à suivre pour réaliser un sous-échantillonnage systématique. La fraction a (1/8^e de l'échantillon) est égale à la somme de b. L'équivalence avec une procédure dichotomique est montrée schématiquement.

The systematic sub-sampling procedure used in the study. The aliquot a (1/8 of the original sample) is equal to the sum of b. The equivalence with a dichotomic procedure is also presented.

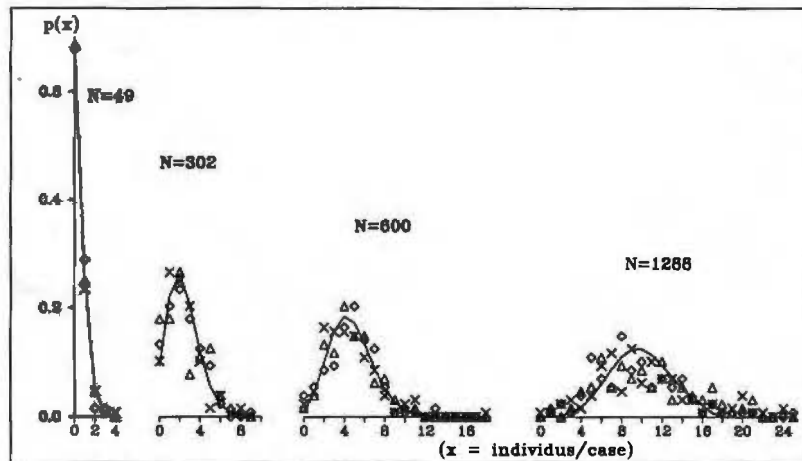


Figure 4

Description des distributions de fréquences du nombre d'organismes par case élémentaire obtenues pour les trois expériences (\square) et les lois binomiales correspondantes (---).

Frequency distributions of the number of organisms per elementary compartment (\square) for the three experiments, and the corresponding binomial distributions (---).

obtenue après le dénombrement de 128 cases, montre l'existence d'agrégats. En absence d'agrégats les trois méthodes de construction des sous-échantillons doivent fournir des résultats sensiblement identiques.

Si les organismes sont indépendants entre eux et se répartissent dans la cuve de façon aléatoire, la distribution des fréquences de l'abondance par case doit suivre une loi binomiale $B(N, 1/128)$. Les distributions spatiales, après « homogénéisation » s'écartent souvent du modèle binomial comme le montre la figure 4. Cette figure permet la comparaison entre les distributions observées au cours des trois expériences et le modèle Binomial correspondant. En moyenne, sur les trois répétitions, l'adéquation est bonne même si elle se détériore pour les grandes abondances ($N > 600$). Or en pratique, c'est à une distribution particulière que nous sommes confrontés. Celle-ci peut s'écarter considérablement du modèle moyen. Les différences observées entre les quatre types d'organismes et pour un type donné d'une expérience à l'autre suggèrent que la distribution des organismes dans la cuve est non seulement fonction de leur abondance mais aussi de leur nature et de leur proportion dans l'ensemble des organismes. Alors plutôt que de rechercher une loi expérimentale à valeur générale (Frontier 1972), ou une loi théorique (Horwood et Driver, 1976), il faut vérifier si la procédure de construction de sous-échantillons permet de s'affranchir facilement de toute distribution spatiale particulière dans la cuve. En effet, les fluctuations d'une expérience à l'autre peuvent être importantes aussi, nous proposons de ne pas faire référence à un modèle mais de calculer à chaque fois une variance de sous-échantillonnage. C'est ce que nous appelons « variance calculée en routine ».

La relation effort, abondance et variance

Avant de pouvoir comparer les méthodes il faut identifier les facteurs qui influencent la précision de l'estimation. Le premier facteur est l'effort, autrement dit la fraction examinée. La précision augmente avec la taille

du sous-échantillon. La figure 5 montre la relation entre la variance de l'estimateur du nombre total de l'ensemble des œufs des poissons ($N = 1288$) et la fraction examinée. L'organisation en cases élémentaires permet d'illustrer facilement cette liaison variance-fraction. Tout se passe comme si dans une population de 128 cases (de moyenne m par case et de variance s^2) on tirait des échantillons de taille K cases. La variance des moyennes de ces échantillons est s^2/K . Il est facile de vérifier dans la figure 5 que cette variance évolue comme une fonction de $1/K$.

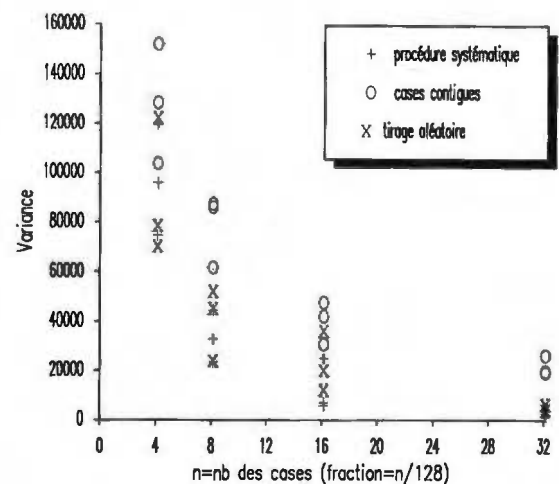


Figure 5

Relation effort-variance pour l'ensemble des œufs de poissons ($N = 1288$) et pour les trois procédures de construction du sous-échantillon.

Effort-variance relation for the total number of fish eggs ($N = 1288$) and for the different procedures of the sub-sample construction.

La relation entre la quantité totale dans la cuve d'un type d'organisme et la variance des sous-échantillons est donnée sur la figure 6, pour un effort de sous-échantillonnage fixé au $1/4$ de l'échantillon initial. Dans ce cas les quatre types d'organismes morphologiquement différents sont considérés. La variance augmente avec l'abondance totale et pour une abondance donnée,

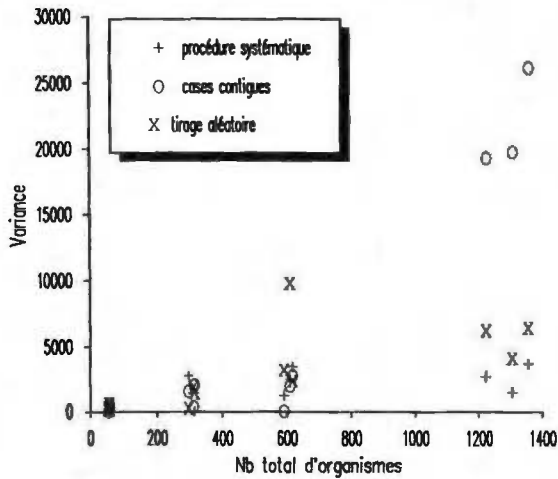


Figure 6

Relation abondance-variance pour une fraction donnée (1/4) et les différents types d'organismes qui représentent les niveaux d'abondance (larves de sole $N=49$, œufs de sole $N=302$, autres larves $N=600$ et œufs des poissons $N=1288$).

Abundance-variance relation for a given fraction (1/4) and different types of organisms representing the abundance levels (sole larvae $N=49$, sole eggs $N=302$, fish larvae $N=600$ and fish eggs $N=1288$).

la variance dépend de la procédure de construction du sous-échantillon. La distribution spatiale des organismes dans la cuve est fonction de leur abondance, de leur forme et de leur proportion dans l'ensemble des organismes. Pour éviter les interactions entre abondance, morphologie et taille de sous-échantillon, la comparaison des différentes procédures sera effectuée pour un type d'organisme et une fraction de taille fixe.

La variance des différentes méthodes

Comme on a pu le constater la répartition des organismes dans la cuve de fractionnement est loin d'être uniforme. Elle est le plus souvent agrégée, les agrégats étant des régions caractérisées par des densités relativement fortes entourées par des secteurs de faible densité. En d'autres termes, la présence d'un individu à un endroit modifie la probabilité de présence d'un autre individu à proximité du premier.

Dans la cuve, si l'agrégat a des dimensions supérieures à la case élémentaire, les effectifs par case forment une population très hétérogène. L'agrégat possède des caractéristiques (en termes de nombre d'individus par case) différentes comparé à la moyenne obtenue sur l'ensemble de la grille. Ainsi les éléments d'un sous-échantillon qui contient l'agrégat (les cases d'une grappe) pris deux à deux et comparés à la moyenne globale ont un coefficient de corrélation commun appelé coefficient de corrélation intra-grappe. Dans l'échantillonnage par grappe, la variance de l'estimateur peut être exprimée en tenant compte de la corrélation intra-grappe (Cochran, 1977). Une valeur positive indique que les éléments de la grappe ont le même comportement par rapport à la moyenne globale, c'est-à-dire qu'ils sont en majorité plus grands ou plus petits que celle-ci. Ceci signifie que la variance inter-grappe est plus grande que la variance intra-grappe. Une valeur faible de corrélation intra-grappe indique que les élé-

ments de la grappe ne montrent pas des écarts importants comparés à la moyenne globale. Par conséquent, les différentes grappes se ressemblent, la variance inter-grappe et celle de l'estimateur sont faibles. Sur la figure 7, la relation entre corrélation intra-grappe et variance des sous-échantillons est présentée pour les trois procédures de sous-échantillonnage.

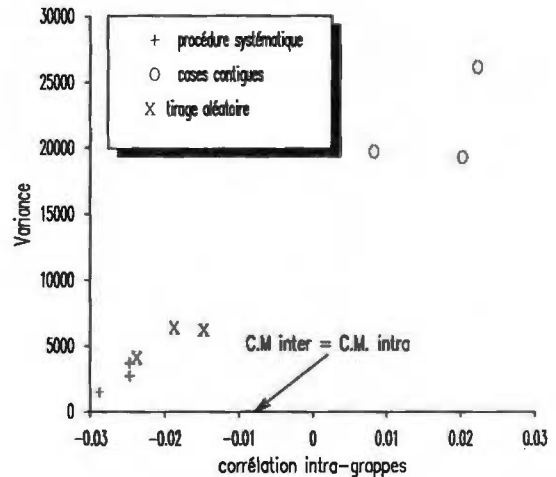


Figure 7

Relation entre la corrélation intra-grappe et la variance pour un niveau d'abondance ($N=1288$) et une fraction (1/4) donnée. La valeur du coefficient pour laquelle la variance inter-grappes est égale à la variance intra-grappes est indiquée (C.M. = carré moyen).

The inter and intra-cluster variance can be expressed as a function of the intra-cluster correlation:

$$C.M. \text{ inter} = FS^2 (1 + (K - 1)p) / (F - 1)$$

$$C.M. \text{ intra} = S^2 (1 - p)$$

Le coefficient de corrélation inter-grappes est donné par :

Relation between intra-cluster correlation and variance for a fixed level of abundance ($N=1288$) and a given fraction (1/4). The value of the coefficient giving an inter-cluster variance equal to the intra-cluster one is also indicated (C.M. = mean square).

The inter and intra-cluster variance can be expressed as a function of the intra-cluster correlation:

$$C.M. \text{ inter} = FS^2 (1 + (K - 1)p) / (F - 1)$$

$$C.M. \text{ intra} = S^2 (1 - p)$$

The intra-cluster correlation coefficient is:

$$p = \frac{\sum_{i=1}^F \sum_{j=1}^K \sum_{u=1}^K (x_{ij} - \bar{X})(x_{iu} - \bar{X})}{(K - 1)(FK - 1)S^2}$$

S^2 = variance des éléments dans la population abstraction faite des grappes.

S^2 = variance of the elements for the entire grid.

$$S^2 = \frac{\sum_{i=1}^F \sum_{j=1}^K (x_{ij} - \bar{X})^2}{(FK - 1)}$$

x_{ij} = nombre d'organismes dans la case j de la grappe i .

x_{ij} = number of organisms in the element j of the cluster i .

\bar{X} = moyenne du nombre d'individus par case sur l'ensemble de la grille.

\bar{X} = mean number of individuals per element for the entire grid.

La première méthode de construction des sous-échantillons (bloc de cases contiguës) analogue à la procédure dichotomique donne le plus fort coefficient de corrélation. L'influence des agrégats est nette. Les éléments de la grappe ont tendance à se comporter de la même façon, comparés à la moyenne générale, du fait de leur proximité dans l'espace. En d'autres termes, les

éléments se ressemblent, les grappes sont au contraire très dissemblables. Le coefficient de corrélation intra-grappe est négatif pour la procédure systématique, car celle-ci tend à maximiser la distance entre les cases élémentaires de la grappe et de ce fait augmenter la variance intra-grappe. Les grappes sont composées d'éléments hétérogènes. Ainsi, la procédure systématique minimise la variance inter-grappe. Chaque grappe incorpore le maximum d'hétérogénéité dû aux structures spatiales. C'est ce que l'on traduit par l'expression « meilleur parcours de l'échantillon ». Les grappes se « ressemblent ». Dans l'échantillonnage par grappes, seule la variance inter-grappe est considérée pour le calcul de la variance de l'estimateur d'abondance. Il est donc évident que la procédure systématique est plus précise. Le tirage aléatoire des cases donne des résultats intermédiaires.

La variance du sous-échantillonnage dépend du plan adopté. Pour un nombre donné d'individus comptés, la procédure systématique donne une meilleure précision par rapport à un plan aléatoire simple. En conséquence, comme le prévoit la théorie d'échantillonnage, pour une précision donnée, la procédure systématique demande un effort moindre.

Les résultats obtenus, pour des fractions allant du 1/128 au 1/4, par la procédure systématique et la procédure dichotomique, sont étudiés globalement à l'aide des relations moyenne-variance ($m-s^2$), et peuvent être directement comparés (fig. 8).

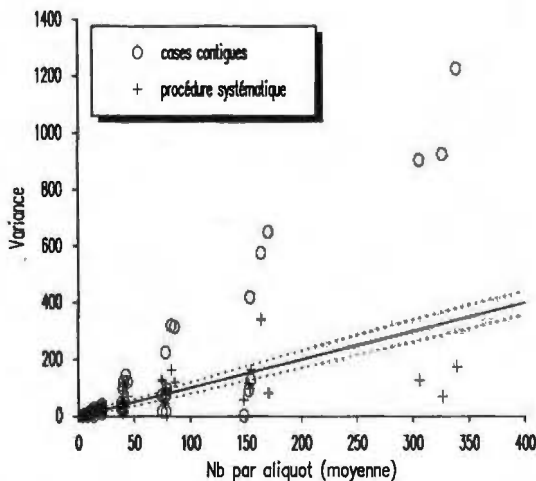


Figure 8

Relations moyenne-variance pour les deux procédures de sous-échantillonnage et comparaison avec la loi de Poisson ($m=s^2$). Les lignes pointillées donnent les intervalles de confiance autour de la loi de Poisson (Tables de Pearson et Hartley, 1966). Une comparaison directe de la précision des deux procédures peut être réalisée dans la mesure où il s'agit des expériences couplées.

Mean-variance relations for the two sub-sampling procedures and the Poisson distribution ($m=s^2$) with its 5% confidence limits (tables from Pearson et Hartley, 1966). The precision of the two procedures can be compared directly because the experiments in question were conducted in parallel.

Ces paramètres sont calculés sur les trois répétitions du sous-échantillonnage, soit trois couples $m-s^2$ par fraction et par procédure.

La relation $m-s^2$ reste proche de la loi de Poisson ($m=s^2$) pour les effectifs faibles dans les sous-échantillons et/ou les petites fractions. Ce phénomène était déjà décrit dans la littérature (Kott, 1953, George *et al.*,

1984) et commenté par Cassié, 1963. Dans ces expériences, le manque d'effectifs élevés permettait d'avancer que la loi de Poisson, qui pour des fractions faibles est une bonne approximation de la loi Binomiale, était une approximation satisfaisante pour la description du phénomène.

Pour les fortes abondances et/ou les grandes fractions l'influence du modèle de répartition dans la cuve sur la variance de l'estimateur est nette. Le contenu d'une case est fortement influencé par le contenu des cases voisines. Il est alors évident que sans référence possible à un modèle de distribution spatiale dans la cuve de fractionnement, l'estimation de σ_{sc}^2 à partir des sous-échantillons devient nécessaire. Le tirage de grappes indépendantes permet ce calcul dans la routine.

DISCUSSION

Lors de la construction de la cuve, la taille et la nature des échantillons déterminent le nombre des cases sur la grille. En effet, les dimensions d'une case doivent être inférieures à celles des agrégats potentiels. Si en théorie la réduction des dimensions des cases permet de mieux s'affranchir des structures spatiales dans la cuve, elle entraîne en retour une augmentation du nombre d'organismes qui resteront piégés sous la grille. Un compromis est donc nécessaire.

Dans l'échantillonnage par grappes, pour la variance de l'estimateur, seule la variabilité inter-grappe est prise en compte (variabilité du nombre compté par fraction) et elle est estimée par sondage (nombre des grappes examinées). Lors de la construction d'un sous-échantillon une même fraction peut être obtenue par le tirage d'une grande grappe ou de plusieurs petites. Dans le deuxième cas, l'estimation de la moyenne sera moins précise car une petite grappe « parcourt » moins bien l'échantillon. Ainsi, quand les grappes sont petites le plan se rapproche d'un tirage aléatoire de cases. Un compromis doit être trouvé entre la taille des grappes (pour une estimation précise de la moyenne) et le nombre des grappes (pour une estimation précise de la variance). Nous suggérons le dépouillement de deux grappes moyennes pour satisfaire ces deux points et calculer la variance de sous-échantillon en routine. Ce même problème se pose lors d'une approche séquentielle (dont une application est évoquée chez Möller et Bernhard, 1974) : pour un même nombre d'individus comptés, la valeur de σ_{sc}^2 dépend du parcours de l'échantillon au moment de la construction du sous-échantillon. En dernier lieu, la notion d'effort de sous-échantillonnage mérite sans doute une clarification. L'effort est défini par le temps nécessaire à la procédure de fractionnement, le volume examiné et le nombre d'individus compté. La seule utilisation du nombre compté comme indice d'effort de sous-échantillonnage est certes quelque peu imprécise.

CONCLUSION

La variance du sous-échantillonnage est fonction de la procédure de construction (plan) et du nombre d'indivi-

des comptés par sous-échantillon. Or l'effort ne peut être optimisé dans ce seul cadre. En effet, il est établi à la suite d'une stratégie de répartition de l'effort total entre échantillonnage et sous-échantillonnage. Le modèle d'analyse de variance donné en introduction suggère que pour estimer la variabilité de la population dans le milieu, il faut minimiser la différence entre le volume de l'échantillon initial et le volume de la fraction finale. Ainsi une fois l'effort fixé, l'optimisation consiste à choisir un plan de sous-échantillonnage.

Pour un nombre donné d'individus comptés, la variance du sous-échantillonnage n'est fonction que de la procédure de construction des sous-échantillons. Pour un même effort, nous avons vérifié expérimentalement qu'une procédure systématique (par tirage des grappes), est plus précise qu'une procédure aléatoire puisqu'elle « parcourt mieux l'échantillon ». Dans cette optique, la règle à plancton de Ibanez (1975) et le « spin splitter » de Zo (1978) permettaient déjà de « parcourir l'échantillon »; mais l'optimisation réside ensuite dans l'utilisation routinière d'une méthode simple tant sur le plan statistique que pratique. Les importantes fluctuations observées, même après homogénéisation, d'une réalisation à l'autre font que la référence

à un modèle théorique ou expérimental est peu satisfaisante surtout dans les grandes abondances et/ou les grandes fractions. Or, dans la pratique c'est le nombre des organismes abondants que nous voulons estimer à partir d'un sous-échantillon. Le tirage des grappes indépendantes construites selon une procédure systématique diminue l'influence des agrégats en régularisant les structures et permet d'estimer la variance du sous-échantillonnage.

Enfin l'intérêt d'une réduction *a priori* de la variance, grâce au choix adéquat d'un plan de sous-échantillonnage, est toujours fondé car cette source de variation s'ajoute aux autres. L'attribution au sous-échantillonnage d'un effort donné dépend de l'objectif final sur la variance de l'estimateur.

Remerciements

Nous tenons à remercier A. Laurec pour avoir orienté l'étude à ses débuts, R. Chevalier pour ses remarques constructives et M^{lle} D. Pelletier pour ses commentaires sur le manuscrit.

RÉFÉRENCES

- Cassie R. M. (1963). Microdistribution of plankton. *Oceanogr. mar. Biol. a. Rev.*, **1**, 223-252.
- Cochran W. G. (1977). *Sampling techniques*. London John Wiley and Sons ed., 3rd edition.
- Dahiya R. (1980). Estimating the population size of different types of organisms in a plankton sample. *Biometrics*, **36**, 437-446.
- Frontier S. (1972). Calcul de l'erreur sur un comptage du zooplankton. *J. expl. mar. Biol. Ecol.*, **8**, 121-132.
- George D. G., M. A. Hurley et B. Winstanley (1984). A simple plankton splitter with a note on its reduced subsampling variance. *Limnol. Oceanogr.*, **29**, 2, 429-433.
- Griffiths F. B., G. H. Brown, D. D. Reid et R. R. Parker (1984). Estimation of sample zooplankton abundance from Folsom splitter subsamples. *J. Plankt. Res.*, **6**, 5, 721-731.
- Horwood J. W. et R. M. Driver (1976). A note on a theoretical subsampling distribution of macroplankton. *J. Cons. int. Explor. Mer.*, **36**, 274-276.
- Ibanez F. (1975). Contribution à l'analyse mathématique des événements en écologie planctonique : organisations méthodologiques; étude expérimentale en continu à petite échelle de l'hétérogénéité du plancton côtier. *Thèse de Doctorat d'État en Sciences naturelles*, Université Paris-VI.
- Kott P. (1953). Modified whirling apparatus for the subsampling of plankton. *Aust. J. mar. Freshw. Res.*, **4**, 387-393.
- Möller F. et W. Bernhard (1974). A sequential approach to the counting of plankton organisms. *J. expl. mar. Biol. Ecol.*, **15**, 49-68.
- Pearson E. S. et H. O. Hartley (1962). *Biometrika Tables for Statistics*. Cambridge University Press, 2^e édition.
- Platt T., L. M. Dickie et R. W. Trites (1970). Spatial heterogeneity of phytoplankton in a nearshore environment. *J. Fish. Res. Bd Can.*, **27**, 1453-1473.
- Sell D. W. et M. S. Evans (1982). A statistical analysis of subsampling and an evaluation of the Folsom plankton splitter. *Hydrobiologia*, **94**, 223-230.
- Van Guelpen L., D. F. Markle et D. J. Duggan (1982). An evaluation of accuracy, precision and speed of several zooplankton subsampling techniques. *J. Const. int. Explor. Mer.*, **40**, 226-236.
- Venrick E. L. (1971). The statistics of subsampling. *Limnol. Oceanogr.*, **16**, 811-818.
- Wiborg K. F. (1962). Estimations of number in the laboratory. *Rapp. P.-v. Réun. Cons. perm. int. Explor. Mer.*, **153**, 74-77.
- Zo Z. (1978). Zooplankton spin splitter. *Deep-Sea Res.*, **25**, 571-576.

