

00000775

ELEMENTS POUR LE CALCUL D'ERREUR SUR
L'ESTIMATION D'UNE BIOMASSE DE POISSONS
PAR ECHO-INTEGRATION

PAR
FRANCIS LALOE

ELEMENTS POUR LE CALCUL D'ERREUR SUR
L'ESTIMATION D'UNE BIODASSE DE POISSONS
PAR ECHO-INTÉGRATION

Francis LALOE *

L'écho-intégration sert à estimer une biomasse en utilisant l'écho renvoyé par les poissons après émission d'une onde sonore.

L'intensité reçue est transformée en une quantité exprimée en millibars, et, sur un trajet donné on estime la biomasse en multipliant le poids des réponses reçues par une valeur C supposée constante tout au long d'une campagne.

Nous n'aborderons pas ici le problème de l'incertitude sur la valeur de C . De nombreuses méthodes existent pour l'estimation de C et du coefficient de variation de cette valeur.

Nous nous consacrerons à l'étude de l'erreur sur la valeur obtenue par multiplication par C et nous distinguerons deux sources d'erreurs :

I) Nous supposons qu'il existe une distance ℓ telle que si deux points A et B sont séparés par une distance inférieure ou égale à ℓ , le type de répartition des poissons dans tout le carré $\ell \times \ell$, décrit dans la figure 1, est le même.. Nous pourrions alors admettre que pour un tel carré on dispose d'une quantité en millibars $M = m + E$, si le bateau a parcouru le segment AB, où m est la valeur "vraie" et E un écart aléatoire.

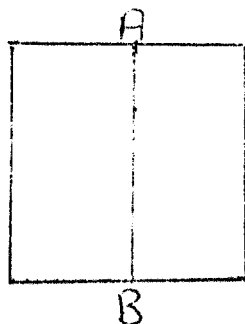


Figure 1.

Nous ferons dans toute la suite l'hypothèse que tout n des n valeurs observées dans des carrés élémentaires δ correspondants à une distance $l_0 = \delta$ au cours d'une campagne sont réparties selon une loi logarithmique. Les répartitions observées au cours des deux campagnes G.F. 76-07 et Guité II ne contredisent pas cette hypothèse. Nous ferons de plus l'hypothèse que après transformation $Y = \log(X+1)$ (on prend $X+1$ car X peut prendre la valeur 0) on a $b = \mu + \xi$ où $\xi \sim N(0, \sigma_\xi^2)$. En décomposant le carré δ en n tronçons de même longueur, nous disposons de n valeurs β_i avec $b = \sum_{i=1}^n \beta_i$. Avec l'hypothèse que le type de répartition est le même sur tout le carré, nous pouvons écrire en supposant les β_i indépendants: $\forall i \quad \beta_i = \left(\frac{\mu}{n} + \xi_i \right)$ où $\xi_i \sim N\left(0, \frac{\sigma_\xi^2}{n}\right)$

L'indépendance des β_i ne peut être satisfaite que si les recouvrements des tronçons successifs est négligeable (ceci exclut de prendre une émission comme source élémentaire si les ondes de deux émissions successives se recouvrent en partie). Nous pouvons ainsi obtenir une estimation de μ en écrivant: $\sum_{i=1}^n (\beta_i - \frac{\mu}{n})^2 \sim \frac{\sigma_\xi^2}{n} \chi_{n-1}^2$

Il est donc possible d'obtenir une estimation de σ_ξ^2 en utilisant plusieurs carrés élémentaires en tenant compte du fait que la somme de n χ_{n-1}^2 indépendants est un $\chi_{N(n-1)}^2$. À l'aide d'une table de χ^2 on peut obtenir un intervalle de confiance sur la valeur de σ_ξ^2 .

2) Au cours d'une campagne le bateau ne parcourt pas toute la zone étudiée. Si nous décomposons la zone prospectée en carrés analogues à ceux définis plus haut, nous pouvons distinguer deux types de carrés; ceux qui sont sur le trajet du bateau et ceux qui ne le sont pas. Nous estimerons la valeur de ceux qui ne sont pas touchés par la valeur observée dans le carré échantillonné le plus proche. Cette estimation implique l'existence d'une nouvelle source d'erreur.

Nous étudierons maintenant des campagnes pour lesquelles les résultats sont répartis le long de radiolignes parallèles et distantes de $k \cdot \delta$ milles où δ est la distance élémentaire définie plus haut et k

est un nombre entier, qui peut être variable (les rectangles peuvent être plus ou moins rapprochés).

Pour les campagnes GAP76-05 et Guinée II, la valeur de k sera prise égale à leur milles (chaque donnée concerne un carré de deux milles de côté).

L'ensemble des carrés élémentaires dont la valeur est estimée par une même observation constitue un ensemble que nous appellerons surface élémentaire (voir figure 2).

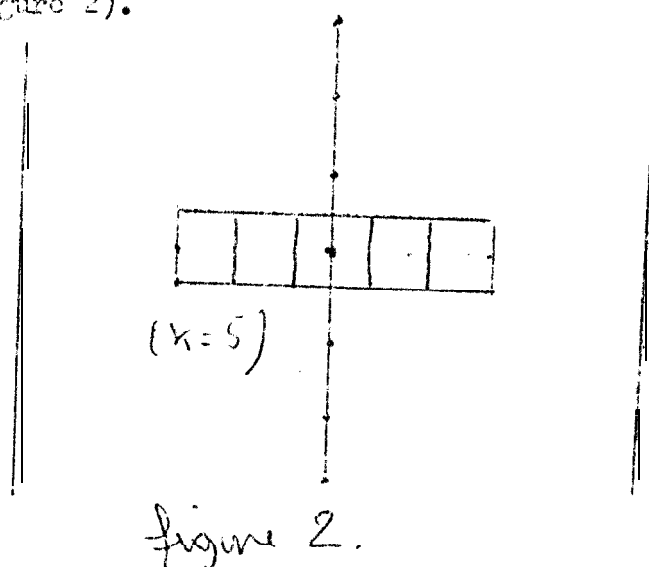


figure 2.

Si tous les carrés d'une surface étaient échantillonnés, nous aurions k valeurs $b_1, \dots, b_k$. Soit $i=k/2$ si k est pair et $i=(k+1)/2$ sinon, alors

$$\sum_{j=1}^k b_j = \sum_{j=1}^k (\mu_j, \xi_j) \text{ donc } k(b_i - \bar{\mu}) = k(b_i - \bar{b}) + k \bar{\xi}$$

$$\text{où } \bar{\mu} = \sum_{j=1}^k \frac{\mu_j}{k}, \quad \bar{b} = \sum_{j=1}^k \frac{b_j}{k} \quad \text{et} \quad \bar{\xi} = \sum_{j=1}^k \frac{\xi_j}{k}$$

$$\text{remarquons que } \text{cov}[(b_i - \bar{b}), \bar{\xi}] = 0$$

$$\text{en effet } \text{cov}[(b_i - \bar{b}), \bar{\xi}] = \text{cov}(b_i, \bar{\xi}) - \text{cov}(\bar{b}, \bar{\xi}) = \frac{1}{k} \sigma_{\xi}^2 - \frac{1}{k^2} \cdot k \sigma_{\xi}^2 = 0$$

(on a utilisé le fait que $\text{cov}(b_i, \xi_j) = \delta_{ij} \sigma_{\xi}^2$ avec $\delta_{ij} = 1$ si $i=j$ et 0 sinon.

En reprenant l'hypothèse que les valeurs observées sont réparties selon une loi lognormale; nous pouvons écrire $b \sim N(\bar{b}, \sigma_b^2)$ où b est la valeur observée en un carré quelconque de la ~~campagne~~ zone. Les paramètres de cette loi sont estimables à partir des données de la campagne, (si k est variable il faut introduire des pondérations pour les estimations). On a: $E(b - \bar{b}) = 0$ et $\text{var}(b - \bar{b}) = \sigma_b^2$ alors $E(k(b - \mu)) = 0$ et $\text{var}(k(b - \mu)) = k^2 \sigma_b^2 + k \sigma_\varepsilon^2$. Un estimateur de σ_b^2 peut être donné en considérant les surfaces de k

carrés consécutifs le long des radiales (si les radiales sont perpendiculaires à la côte, on aura probablement une surestimation de σ_b^2). On a également $\text{var}(b_i - \bar{b}) = \text{var}(b_i) + \text{var}(\bar{b}) - \frac{2}{k} \sum_{j=1}^k \text{cov}(b_i, b_j)$ où $\text{var}(\bar{b}) = \frac{1}{k^2} \sum_{j,h=1}^k \text{cov}(b_j, b_h)$. Les $\text{cov}(b_i, b_j)$ peuvent être estimés par les covariances entre points b_{pi} et b_{pj} sur les radiales avec $q = |i - j|$.

Pour obtenir une estimation de la variance de l'estimateur associé à une radiale, on prendra l'estimation de la ~~max~~ variance de la somme des estimations des surfaces correspondant et aux carrés situés sur cette radiale, les termes de covariances ne sont pas nuls:

Pour deux surfaces 1 et 2, on a $\text{cov}(b_{11} - \bar{b}_1), (b_{21} - \bar{b}_2)$ (c'est à une combinaison linéaire de termes élémentaires du type $\text{cov}(b_{11}, b_{21})$):

$$\text{cov}[(b_{11} - \bar{b}_1), (b_{21} - \bar{b}_2)] = \text{cov}(b_{11}, b_{21}) + \frac{1}{k_1} \sum_{j=1}^{k_1} \text{cov}(b_{11}, b_{2j}) - \frac{1}{k_2} \sum_{h=1}^{k_2} (\text{cov}(b_{11}, b_{2h}) + \text{cov}(b_{21}, b_{1h}))$$
 en utilisant un estimateur de cette combinaison en remplaçant chaque terme $\text{cov}(b_{11}, b_{2j})$ par l'estimation de la covariance entre les valeurs b_{11} et b_{2j} situées sur les mêmes radiales avec $j = |k_1| + 1$ (il y a n surfaces entre les surfaces 1 et 2) (on obtient une sous estimation pour certains termes si on suppose que la liaison est positive et d'croissante avec la distance)

Soit X la matrice de terme $C_{ij} = \omega(\bar{b}_i, \bar{b}_j)$ où b_i et b_j sont les erreurs effectuées sur les mesures des surfaces i et j , (X est une matrice carrée (p, p) où p est le nombre de carrés situés sur la radiale). Tous les calculs ayant été effectués avec les données transformées $X = \log(X)$, $Y = \log(Y)$, nous estimerons les variances et covariances en posant les données brutes en écrivant: $\text{cov}(X_1, X_2) = X_1 X_2 \text{cov}(Y_1, Y_2)$. Pour une radiale l'estimation de la variance sera donc $b' H b$ où b est le vecteur des données brutes.

En supposant indépendantes les erreurs commises pour chaque radiale, la variance totale sera obtenue en sommant les estimations de variance pour chaque radiale.