

LE COVARIOGRAMME: UN OUTIL STRUCTURAL

Nicolas Bez^a, Jacques Rivoirard ^a, Jean-Charles Poulard^b

I -INTRODUCTION

L'utilisation d'outils structuraux répond à deux objectifs : décrire, en la résumant, la structure spatiale d'un phénomène régionalisé, et utiliser cette structure pour effectuer par exemple des cartographies, des estimations, ou encore des simulations. Le but de ce travail est de présenter le covariogramme jusque là utilisé dans des conditions limitées.

Nous verrons qu'il se présente comme une alternative au variogramme, dans des cas où celui-ci se révèle peu adapté. En particulier lorsque se pose la question de savoir comment décrire la distribution spatiale d'une densité de poissons quand l'abondance totale du stock provient de concentrations fortes et rares, et qu'il est difficile de délimiter un champ à l'intérieur duquel les variations sont homogènes.

II -LE VARIOGRAMME

L'analyse variographique porte sur des données de merlu (*Merluccius merluccius*) issues de la campagne de chalutage organisée en 1990 par IFREMER dans le Golfe de Gascogne. L'échantillonnage est stratifié aléatoire dans la moitié Sud du Golfe, et presque régulier dans la moitié Nord (figure 1). Les 137 densités de merlu d'âge 0 sont réparties dans la partie centrale du plateau; l'ensemble du stock de la région semble avoir été observé. L'irrégularité spatiale apparaît importante puisque de fortes densités sont entourées de densités très faibles voire nulles. La distribution des densités est très disymétrique : moins de 10% des densités engendrent 50% du total (forte sélectivité).

Le variogramme est défini dans le cadre de la géostatistique intrinsèque; c'est-à-dire lorsque le phénomène régionalisé est modélisé par une fonction aléatoire $Z(x)$, où x est un point de l'espace :

^aCentre de Géostatistique, ENSMP, 77305 Fontainebleau

^bIFREMER, BP 1105, 44311 NANTES

$$\gamma(\vec{h}) = \frac{1}{2} \times \text{Var}(Z(x) - Z(x + \vec{h}))$$

Dans la pratique, il est souvent considéré sous une autre forme, plus commode à estimer, car il correspond à une espérance dont l'équivalent dans les données est la moyenne du carré des écarts des échantillons distants de \vec{h} . :

$$\gamma(\vec{h}) = \frac{1}{2} \times E[(Z(x) - Z(x + \vec{h}))^2]$$

Des variogrammes ont été calculés perpendiculairement et parallèlement aux isobathes (figure 2). Au travers du plateau continental, le variogramme expérimental augmente rapidement jusqu'à 100 milles nautiques où il dépasse la variance, et décroît ensuite; ce qui n'en fait pas un variogramme "typique". Parallèlement au plateau, le variogramme augmente plus lentement et dépasse largement la variance expérimentale à partir de 100 milles nautiques.

La chute du variogramme aux grandes distances s'explique par la baisse des densités lorsqu'on s'éloigne du "coeur riche". Les différences entre valeurs deviennent alors de plus en plus faibles, jusqu'à devenir nulles quand les distances excèdent les dimensions du stock. Il faut donc définir le champ des valeurs non nulles à l'intérieur duquel le variogramme représente la variabilité intrinsèque du merlu, dans la mesure où la géométrie du stock n'a pas d'incidence sur elle. Cependant, la définition du champ entraîne une modification du variogramme et implique le choix d'un seuillage permettant de considérer des densités très faibles comme nulles et de faire la distinction entre zéros extérieurs et intérieurs à la régionalisation (problème du traitement des densités quasi-nulles observées au Nord).

La signification et l'intérêt du variogramme est de mesurer une variabilité à l'intérieur d'un champ en faisant une moyenne de quantités qu'on juge d'une certaine façon comparables. Cette démarche est formalisée par l'hypothèse dite intrinsèque où les incréments de la variable sont stationnaires (écarts entre densités supposés de même loi partout dans le champ). Or, dans la mesure où un champ serait fixé, il reste que les fortes densités ont généralement une localisation préférentielle à l'intérieur du champ. Il devient alors difficile de supposer que l'ensemble des paires $z(x), z(x + \vec{h})$, dont on fait la moyenne pour calculer le variogramme, sont comparables. Le recours au variogramme n'est donc judicieux que si toutes les parties du champ jouent le même rôle dans la variabilité du phénomène.

III -LE COVARIOGRAMME

A -Définition et estimation

La géométrie du champ étant souvent un élément structurant des populations halieutiques, il devient délicat de chercher à exhiber une variabilité des densités interne et indépendante au champ. De plus, en probabilisant l'information, on suppose que les densités auraient pu être observées en d'autres endroits du champ, tout en conservant les caractéristiques structurales moyennes de la variable. Ce qui n'est probablement pas très réaliste en halieutique.

Nous avons donc fait appel au covariogramme. En s'en tenant au formalisme des variables régionalisées, le covariogramme est défini par :

$$g(\vec{h}) = \int z(x)z(x + \vec{h})dx$$

Il permet de résumer l'information structurale bivariable contenue dans les données. Le lien avec le variogramme est donné par la relation :

$$g(0) - g(\vec{h}) = \int_{S \cap S_{-\vec{h}}} (z(x) - z(x + \vec{h}))^2 dx$$

où S est le champ des valeurs non nulles et $S_{-\vec{h}}$ son translaté de $-\vec{h}$.

En somme, on ne s'intéresse plus à une moyenne mais à une somme d'écartés quadratiques. Le covariogramme est donc un outil structural, équivalent au variogramme, qui ne nécessite aucune hypothèse de stationnarité, et pour lequel ne se posent plus les problèmes d'occurrence (reproductible) des fortes valeurs, et de traitements de valeurs nulles (toujours source de tracas pour les variogrammes). La seule réelle contrainte pour utiliser le covariogramme est que l'échantillonnage doit déborder le champ.

Dans le cas d'une maille régulière (cadre habituel de son utilisation), le covariogramme expérimental est fourni par :

$$g^*(k.a) = a \sum_i z(x_o + i.a)z(x_o + i.a + k.a)$$

où x_o est l'origine de la grille et a la maille. La structure expérimentale est donc discrète, connue pour des distances multiples de la maille. Son ajustement par une fonction continue fournit le covariogramme recherché.

Cependant, l'estimation d'un covariogramme pose des problèmes lorsque l'échantillonnage n'est pas régulier. Deux solutions sont proposées pour mener à bien les calculs dans la pratique (Bez *et al.*, 1995):

- la création d'une grille régulière fine informée en chaque noeud par la donnée la plus proche. Le covariogramme calculé est, aux erreurs de discrétisation près :

$$g^*(\vec{h}) = \sum_i^n \sum_j^n z_i z_j |S_i \cap S_{j-\vec{h}}|$$

où S_i est la surface d'influence de l'échantillon z_i . Chaque couple de données (z_i, z_j) intervient globalement avec un poids $S_i S_j$. Cette méthode présente l'inconvénient de linéariser le covariogramme, et d'attribuer toute une gamme de distances à chaque produit de densités.

- une pondération appropriée par les surfaces d'influence de chaque échantillon :

$$g^*(\vec{h}) = K(\vec{h}) \times \frac{\sum \sum_{i,j \approx \vec{h}} z_i z_j S_i S_j}{\sum \sum_{i,j \approx \vec{h}} S_i S_j}$$

où $K(h)$ est le covariogramme géométrique du champ. Le poids total de chaque paire est toujours $S_i S_j$ mais la valeur $z_i z_j$ est dorénavant affectée à la distance correspondant à celle de leur observation. Cependant, l'échantillonnage pouvant être plus resserré autour des endroits à fortes densités, les surfaces d'influence de chaque point peuvent être liées aux densités; ce qui entrainerait alors un biais dans le calcul du covariogramme.

Le covariogramme expérimental (figure 3), obtenu par la première méthode, sont plus réguliers que les variogrammes correspondants. Leurs modélisation est donc plus commode. Leurs portées ne s'interprètent plus en termes de corrélations mais en termes géométriques (étendue des stocks).

B -Utilisation du covariogramme et approches plus fines

A maille régulière, le covariogramme s'introduit naturellement pour l'estimation d'une variance d'estimation globale. Lorsque la reconnaissance d'un phénomène déborde suffisamment le champ, il est possible, à condition de manipuler des estimateurs linéaires et invariants par translation, d'utiliser le covariogramme pour des estimations locales par krigeage (Bez *et al.*, 1995).

On définit également des outils structuraux plus fins :

- les lois bivariées $K_{z,z'}(h) = \text{Surf} \{x/z(x) > z' \text{ et } z(x+h) > z\}$ - les fréquences conditionnelles $K_{z,z'}(h)/K_{z'}(0) = f_c(z(x+h) > z | z(x) > z')$ avec laquelle la densité en $x+h$ dépasse z sachant qu'elle dépasse z' au point x , permettent de décrire la structure spatiale du point de vue des fortes concentrations,
- les régressions bivariées $m[z(x+h)|z(x)=z]$, moyenne de $z(x+h)$ quand $z(x)=z$,

dont le covariogramme est la synthèse :

$$g(h) = \int \int K_{z,z'}(h) dz dz' = \int m[z(x+h)|z(x)=z] S(z) dz$$

IV - CONCLUSION

Le covariogramme est un outil structural synthétique, global, et ne requérant pas d'hypothèse particulière. Il est généralement présenté dans le cadre d'un échantillonnage régulier qui est la situation la plus confortable pour son estimation. Dans la pratique halieutique, se pose le problème de l'inférence du covariogramme à partir d'une maille souvent peu régulière.

En proposant une intégrale plutôt qu'une espérance, l'approche transitive (externe et globale) pourrait également être appliquée aux questions relatives aux liens entre densité et variables accessoires (température, bathymétrie, ...).

Cette étude s'est déroulée dans le cadre du programme européen "Shelf Edge Fisheries and Oceanography Study" (référence : AIR 93 1105)

BIBLIOGRAPHIE

- BEZ (N.), RIVOIRARD (J.), & POULARD (J.C.), 1995 - Représentation externe de densités de poissons. *Cahiers de Géostatistique*, Comptendu des journées de Géostatistique 15-16 juin 1995, Paris, E.N.S.M.P. (A paraître)
- MATHERON (G.), 1970 - La théorie des variables régionalisées et ses applications. *Les Cahiers du Centre de Morphologie Mathématique*, Fasc. 5, Paris, E.N.S.M.P. 212 p.

MATHERON (G.), 1978 - Estimer et choisir. *Les Cahiers du Centre de Morphologie Mathématique*, Fasc. 7, Paris, E.N.S.M.P. 175 p.

PETITGAS (P.), 1991 - Contributions géostatistiques à la biologie des pêches maritimes. Thèse Dr. en Géostatistique, Paris, E.N.S.M.P. 211 p.

POULARD (J.C.), 1990 - Evaluation des ressources halieutiques de l'Ouest de l'Europe, Deuxième Phase, Contrat IFREMER-CEE, n°88/1210834/BF.

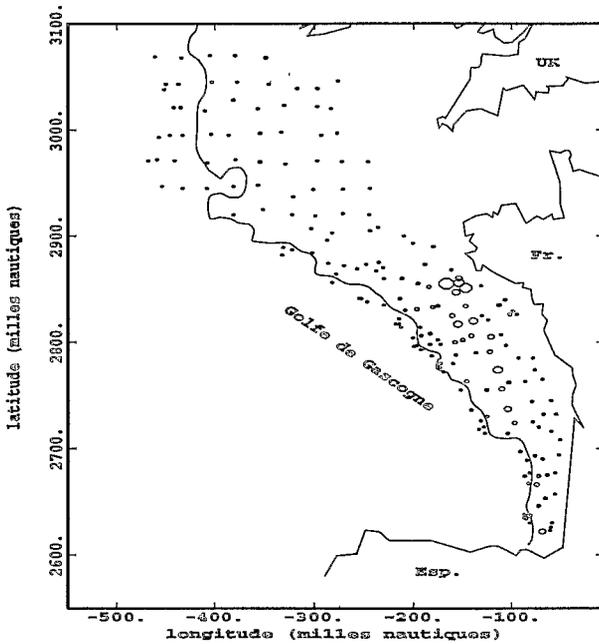


Figure 1. Densités de merlu d'âge 0.

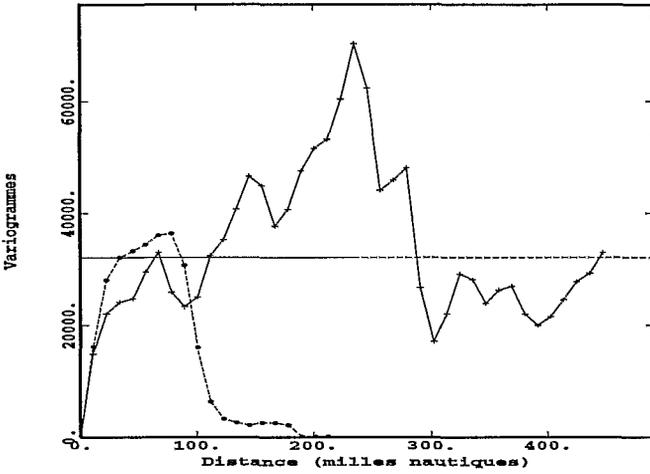


Figure 2. Variogramme expérimental selon 2 directions : perpendiculaire (•) et parallèle (+) aux isobathes.

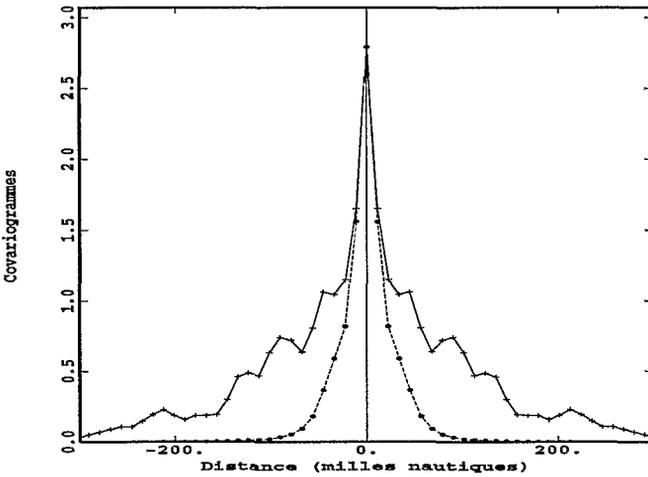


Figure 3. Covariogramme expérimental selon 2 directions : perpendiculaire (•) et parallèle (+) aux isobathes.