Deuxième Forum Halieumétrique, Nantes, 1995
*Session II : **Analyse de l'information***

ESTIMATION DES PUISSANCES DE PECHE

François Gauthiez[a]

I - INTRODUCTION

Le modèle multiplicatif de Robson (Robson, 1966) d'estimation des puissances de pêche (et des indices d'abondance) est certainement l'un des modèles les plus utilisés en halieutique. Il présente cependant à nos yeux des inconvénients majeurs, c'est pourquoi nous proposons ici de le reconsidérer dans le cadre de l'analyse des données de chalutages scientifiques.

Ce modèle est d'une utilisation assez simple, puisqu'il se réduit à une analyse de variance (en général déséquilibrée) après passage au logarithme des captures. Son expression générale est la suivante:

$$C \sim \mathcal{LN}(\alpha + \beta + \cdots, \sigma^2),$$

c'est-à-dire que la capture C est supposée suivre une loi log-normale; son logarithme voit son espérance s'exprimer comme la somme d'effets (au sens de l'analyse de variance) et a une variance constante. Deux sortes de problèmes sont liés à ce modèle:

- Il existe d'une part des problèmes pratiques:

  ↪ Que faut-il faire des captures nulles?

  ↪ L'analyse de variance sur les logarithmes correspond à l'estimation du maximum de vraisemblance. Que deviennent les propriétés de cet estimateur lorsque la loi n'est plus log-normale?

  ↪ L'hypothèse de variance constante sur l'échelle log correspond à un coefficient de variation constant pour C. Que faire si cette hypothèse n'est pas réaliste?

  ↪ Si l'on considère la classe des modèles linéaires généralisés, la loi "naturelle" pour les modèles multiplicatifs (ou log-linéaires) est la loi de Poisson. Pour cette loi, la variance de C est supposée être égale (ou, du moins, proportionnelle) à sa moyenne. Que choisir?

[a]IFREMER, DRV/RH/MAERHA, BP 1105, 44311 Nantes Cedex 03, France

- D'autre part, l'utilisation d'un tel modèle pose le problème général de l'utilisation de modèles ne reposant pas sur une analyse des processus sous-jacents. Par exemple, l'hypothèse de constance du coefficient de variation de C n'est pas une hypothèse "biologique": bien que l'on ait la possibilité de contrôler *ex post* sa validité, sa justification première correspond probablement au fait qu'elle permet, après un passage au logarithme, de se trouver dans les condition s standard de l'analyse de variance, c'est-à-dire l'homoscédasticité. Plus généralement, il est difficile dans ce type de démarche d'incorporer au modèle des connaissances ou des hypothèses particulières sur le comportement de la ressource ou sur le processus de capture.

On se propose donc d'appliquer au modèle multiplicatif une démarche générale consistant à fonder un modèle statistique sur une analyse des processus sous-jacents, en particulier les processus intervenant à une échelle inférieure à l'échelle d'observation.

II - MODELISATION DU PROCESSUS DE CAPTURE

On peut distinguer deux composantes essentielles dans la variabilité des captures:

- Le chalut étant un engin qui se déploie dans l'espace, il s'ensuit que l'arrangement spatial des poissons dans la zone prospectée constitue une première composante. Plus précisément, le nombre d'individus présents dans l'aire balayée par le chalut et sa variabilité sont directement liés aux déplacements et au comportement agrégatif du poisson.

- D'autre part, les individus présents dans l'aire balayée ne sont pas tous capturés. Le comportement du poisson vis-à-vis du chalut constitue la deuxième composante.

Pour ce qui est de la première composante, nous nous plaçons à une échelle locale, où prédominent les comportements agrégatifs (voir Gauthiez *et al.*, 1995). Le nombre N d'individus présents dans l'aire balayée par le chalut peut alors être exprimé comme la somme du nombre X d'individus par agrégats calculée sur l'ensemble des agrégats présents dans l'aire balayée (P agrégats):

$$N = \sum_{i=1}^{P} X_i,$$

où P et les X_i sont des variables aléatoires mutuellement indépendantes. On suppose que les agrégats sont répartis au hasard, de sorte que P suit une loi de Poisson. Dans des conditions de chalutage standardisées, l'espérance de N (notée ν) représente l'abondance locale. En réponse à des variations de cette abondance locale, l'arrangement spatial des individus peut varier d'une multitude de façons. Le nombre moyen d'agrégats peut rester constant, ou bien la taille moyenne des agrégats peut rester constante, ou encore on peut avoir une situation intermédiaire. De plus, la variabilité de la taille des agrégats peut, elle aussi, varier de différentes manières lorsque la taille moyenne change. L'une ou l'autre de ces réponses possibles à une variation d'abondance locale constitue ce que nous appelons la densité-dépendance des schémas d'agrégations. Dans le cas d'une campagne de chalutages il n'est pas possible d'observer directement la forme de cette densité-dépendance, mais chaque forme se traduit par une formulation particulière de la variance de N. Un certain nombre de cas sont envisagés dans le tableau 1. Pour ce qui est de la deuxième composante, l'hypothèse la plus simple consiste à supposer que tous les individus ont la même probablité π d'être capturés. Dans ce cas, conditionnellement à un nombre N d'individus présents dans l'aire balayée, la capture C suit une loi binomiale de paramètres N et π:

$$C \sim \mathcal{B}(N, \pi).$$

De nombreuses raisons peuvent nous amener à remettre en cause un modèle aussi simple, nous n'en citerons que les principales. D'abord, la probabilité de capture π peut varier d'un trait à l'autre: des facteurs influençant cette probabilité, comme la géométrie du chalut, la longueur du trait, ou les conditions de milieu peuvent en effet subir des variations inter-traits. Ensuite, au cours d'un même trait, on peut imaginer que la taille d'un agrégat influe sur sa capacité d'échappement. Enfin, un problème de saturation peut intervenir. Certains de ces problèmes sont des sources de variabilité supplémentaires, tandis que d'autres introduisent un biais dans la relation entre capture et abondance locale (la capture moyenne peut ne plus être proportionnelle à l'abondance locale ν). Le tableau 2 présente le cas π constant. Les autres cas ne sont pas présentés ici faute de place.

Tableau 1: Relations moyenne-variance vérifiées par le nombre N d'individus se trouvant dans l'aire balayée par le chalut.

variation de la loi de P	variation de la loi de X	expression de $\mathrm{Var}(X)$	expression de $\mathrm{Var}(N)$
$\mathrm{E}(P) \propto \nu$	constante	.	$a_0\,\mathrm{E}(N)$
constante	$\mathrm{E}(X) \propto \nu$	$c\,\mathrm{E}^2(X)$	$a_0\,\mathrm{E}^2(N)$
		$\mathrm{E}(X) + c\,\mathrm{E}^2(X)$	$\mathrm{E}(N) + a_0\,\mathrm{E}^2(N)$
		$c\,\mathrm{E}^b(X)$	$a_0\,\mathrm{E}^b(N) + a_1\,\mathrm{E}^2(N)$
		$\mathrm{E}(X) + c\,\mathrm{E}^b(X)$	$\mathrm{E}(N) + a_0\,\mathrm{E}^b(N) + a_1\,\mathrm{E}^2(N)$
$\mathrm{E}(P) \propto \nu^q$	$\mathrm{E}(X) \propto \nu^{1-q}$	$c\,\mathrm{E}^2(X)$	$a_0\,\mathrm{E}^b(N)$ avec $b = 2 - q$
		$\mathrm{E}(X) + c\,\mathrm{E}^2(X)$	$\mathrm{E}(N) + a_0\,\mathrm{E}^b(N)$ avec $b = 2 - q$
		$c\,\mathrm{E}^\beta(X)$	$a_0\,\mathrm{E}^{b_0}(N) + a_1\,\mathrm{E}^{b_1}(N)$ avec $b_0 = q + \beta(1-q)$ et $b_1 = 2 - q$
		$\mathrm{E}(X) + c\,\mathrm{E}^\beta(X)$	$\mathrm{E}(N) + a_0\,\mathrm{E}^{b_0}(N) + a_1\,\mathrm{E}^{b_1}(N)$ avec $b_0 = q + \beta(1-q)$ et $b_1 = 2 - q$

III - INFERENCE STATISTIQUE

Dans la partie précédente, les différents modèles probabilistes envisagés n'ont été présentés qu'à travers la formulation de l'espérance et de la variance de la capture. Ce parti pris corrrespond à la nécessité de formuler les problèmes d'une façon qui soit facilement transférable en termes de modèles statistiques. Or, s'il est très facile de calculer une moyenne et une variance, il est en général hasardeux en halieutique de postuler que des captures observées sont issues d'une loi donnée. C'est pourquoi nous préférons caractériser les processus par l'expression des deux premiers moments et ne pas aller plus loin dans la spécification d'une loi. Du point de vue du statisticien, cela signifie que l'on se place d'emblée dans un cadre *semi-paramétrique.* Dans le cadre d'un modèle multiplicatif, on peut par

Tableau 2: Expressions de la variance de C, en fonction des différents relations moyenne-variance vérifiées par N.

Expression de $\mathrm{Var}(N)$	capture moyenne	expression de $\mathrm{Var}(C)$
$a_0\nu$	$\pi\nu$	$\mathrm{E}(C) + a_1\pi\,\mathrm{E}(C)$ $= \pi\nu + a_1\pi^2\nu$ avec $a_1 = a_0 - 1$
$a_0\nu^2$	$\pi\nu$	$(1-\pi)\,\mathrm{E}(C) + a_0\,\mathrm{E}^2(C)$ $= \pi(1-\pi)\nu + a_0\pi^2\nu^2$
$\nu + a_0\nu^2$	$\pi\nu$	$\mathrm{E}(C) + a_0\,\mathrm{E}^2(C)$ $= \pi\nu + a_0\pi^2\nu^2$
$a_0\nu^b$	$\pi\nu$	$(1-\pi)\,\mathrm{E}(C) + a_0\pi^{2-b}\,\mathrm{E}^b(C)$ $= \pi(1-\pi)\nu + a_0\pi^2\nu^b$
$\nu + a_0\nu^b$	$\pi\nu$	$\mathrm{E}(C) + a_0\pi^{2-b}\,\mathrm{E}^b(C)$ $= \pi\nu + a_0\pi^2\nu^b$

exemple postuler le modèle suivant:

$$\begin{aligned} \mathrm{E}(C) &= \varphi\alpha\nu && (1) \\ \mathrm{Var}(C) &= \varphi\alpha\nu + a_0\varphi\alpha^2\nu^b, && (2) \end{aligned}$$

où la capture moyenne est proportionnelle à une mesure φ de l'effort de pêche, à une puissance de pêche α et à un indice d'abondance ν valable pour une strate spatio-temporelle donnée. La fonction de variance (2) englobe un certain nombre de cas du tableau 2. Sa dépendance linéaire par rapport à φ s'explique par l'hypothèse de distribution au hasard des agrégats. Le coefficient b, quant à lui, est aussi un paramètre à estimer et représente le mode de densité-dépendance des schémas d'agrégation.

Pour choisir un estimateur nous adoptons le point de vue asymptotique: les propriétés minimales requises sont la convergence et la normalité asymptotique. Le cadre semi-paramétrique impose que ces propriétés soient vérifiées **quelle que soit la loi de C vérifiant (1) et (2).** Si la loi de

C vérifie (1) et (2) mais n'est pas log-normale, l'estimateur log-normal du modèle de Robson n'est en général convergent ni pour les paramètres de moyenne (puissances de pêche, indices d'abondance) ni pour les paramètres de variance (a_0 et b). Par contre, l'utilisation d'une vraisemblance correspondant à une loi appartenant à la famille exponentielle linéaire fournit un estimateur satisfaisant dans le cadre semi-paramétrique. De plus, l'estimation des paramètres de la fonction de variance permet d'obtenir un estimateur optimal dans cette classe (Gouriéroux *et al.*, 1984). L'utilisation d'une vraisemblance gaussienne permet d'estimer de façon convergente et asymptotiquement normale les paramètres de variance (Carroll et Ruppert, 1988). En pratique on effectue une estimation alternée des deux groupes de paramètres: par exemple l'estimation de a_0 et b se fait en maximisant une vraisemblance gaussienne où sont supposés fixés les paramètres de moyenne. De façon générale, on regroupe sous le nom de méthodes de *pseudo-vraisemblance* les méthodes consistant à utiliser une vraisemblance en dehors du cadre paramétrique qui a permis de la construire.

Cette démarche a été appliquée aux données de campagnes IBTS en mer du Nord, où sont impliqués les navires de recherche halieutique de plusieurs pays et où se pose naturellement le problème de l'estimation des puissances de pêche. Les estimations trouvées ainsi que les intervalles de confiance calculés peuvent être différents de ceux obtenus avec un modèle de Robson.

IV - DISCUSSION

On s'est efforcé dans cette étude de montrer que le choix d'une méthode statistique doit s'appuyer en premier lieu sur un cadre fixé par une modélisation des processus sous-jacents. Ici, une analyse de l'arrangement spatial des individus et du processus de capture conduit à formuler une expression de la variance de C, en même temps qu'elle montre qu'une hypothèse paramétrique particulière ne semble pas pouvoir s'imposer. Les méthodes de pseudo-vraisemblance ont, dans ce cadre, des propriétés satisfaisantes. D'un point de vue pratique, ce type de méthode est en général facile à mettre en œuvre. En effet, la forme log-linéaire de $\mathrm{E}(C)$ et la contrainte d'utiliser des vraisemblances issues d'une loi appartenant à la famille exponentielle linéaire permettent d'utiliser les fonctionnalités des modèles linéaires généralisés. L'estimation des paramètres de variance requiert l'utilisation d'une routine de minimisation.

Un certain nombre d'extensions peuvent être envisagées. Premièrement, notre objectif étant d'analyser les processus déterminant la variabilité d'une observation, on a adopté une description particulièrement simple de l'abondance, en attribuant à ν une valeur par strate. On pourrait naturellement envisager de la relier à des variables d'environnement ou de représenter ses variations spatiales à grande échelle par des méthodes du type géostatistique. Deuxièmement, ce cadre devrait permettre l'appréhension des puissances de pêche pour les navires commerciaux. Ce transfert n'est pas sans poser certains problèmes. D'une part, l'interaction entre l'agrégation de la ressource et la recherche menée par un navire commercial peut modifier la fonction de variance et, surtout, rendre non linéaire la relation entre capture moyenne et abondance locale (Gauthiez, 1995). D'autre part, un navire commercial réalise à l'évidence des captures différentes selon qu'il cible ou ne cible pas une espèce, de sorte que sa puissance de pêche vis-à-vis d'une epèce donnée dépend de son groupe d'espèces cibles. Une approche de ce problème pourrait être d'utiliser les composantes principales issues d'une analyse factorielle sur les captures pour estimer des puissances de pêche dépendant du groupe d'espèces cibles.

BIBLIOGRAPHIE

CARROLL (R. J.), & RUPPERT (D.), 1988 - *Transformation and weighting in regression.* Chapman & Hall. New-York. 249 p.

GAUTHIEZ (F.), 1995 - Is CPUE a reliable index of abundance? Lessons from a probabilistic model coupling fish aggregation with efficient fishing tactics. *Can. J. Fish. Aquat. Sci.* (soumis).

GAUTHIEZ (F.), POULARD (J.C.), & KOUTSIKOPOULOS (C.), 1995 - Analyse de la distribution spatiale à petite échelle des poissons benthiques et démersaux en Mer Celtique. *Deuxième forum halieumétrique.*

GOURIEROUX (C.), MONTFORT (A.), & TROGNON (A.), 1984 - Pseudo-maximum likelihood methods: theory. *Econometrica* 52: 681-700.

ROBSON (D.S.), 1966 - Estimation of the relative fishing power of individual ships. *Int. Comm. Northw. Atlant. Fish. Res. Bull.* 3: 5-14.