MODÈLES MIXTES ET DONNÉES HALIEUTIQUES

Francis LALOË *, Daniel GAERTNER ** et Frédéric MÉNARD **

RÉSUMÉ

Le recours à des modèles linéaires à effets fixes est très répandu pour le traitement des données de rendements de pêche. En tenant compte d'éventuelles variations spatio-temporelles, il s'agit de comparer l'efficacité d'unités de pêche en vue d'obtenir un effort de pêche exprimé par rapport à une unité de référence. Si l'effet associé à une unité de pêche doit être fixe, la question du statut (fixe ou aléatoire) des effets spatio-temporels est moins évidente. À l'aide d'un exemple, nous discutons de ce statut dans le cadre de la prédiction des rendements de pêche, en considérant des caractéristiques de leur variabilité spatio-temporelle et des informations dont disposent les unités de pêche. Selon le type de modèle utilisé les différences entre ces prédictions et entre les variabilités qui leurs sont associées apparaissent non négligeables. Cela peut avoir des conséquences importantes en termes de conclusions émises dans le contexte de la gestion des pêches.

ABSTRACT

General linear models with fixed effects are commonly used for fishing yield data analysis. Accounting for potential sources of spatio-temporal variability, the aim is to compare the efficiency (fishing powers) of different fishing units in order to standardize their fishing effort with respect to a given fishing unit. By contrast with the fishing unit factor which must be considered as a fixed effect, the status of the spatial and temporal factors (fixed or random) remains problematic. In the context of this standardization procedure, we discuss the usefulness of mixed effects models in fishery studies. We take into account some peculiar features of the fishery system such as the spatio-temporal variability of the resource as well as the amount of information available to the fishermen. Depending on the model, the differences between the predicted values of catch per unit of effort and between their variances can be substantial. As a result, neglecting random effects in conventional fishery analysis can lead to erroneous conclusions in the context of fishery regulations.

1. Introduction

Les modèles linéaires généraux sont systématiquement utilisés pour l'analyse des données de pêche commerciales en vue de caractériser diverses sources de variations, spatiales, temporelles et liées à l'efficacité des unités de pêche

* UMR C3ED (IRD-UVSQ), LER, IRD, BP 64501, 34394 MONTPELLIER Cedex 5; Courriel : Laloe@mpl.ird.fr

** UR THETIS, CRHMT, BP 171, 34203 Sète Cedex

en général identifiées par l'embarcation utilisée. Ces analyses permettent de répondre à des questions de natures différentes.

- Il s'agit tout d'abord de caractériser les « puissances de pêche » individuelles des unités de pêche (Robson 1966). En considérant une unité étalon dont la puissance est par définition égale à 1, la puissance de pêche d'une unité quelconque est égale au rapport des espérances des rendements par action de pêche de cette unité et de l'unité étalon. L'estimation des puissances de pêche est nécessaire pour exprimer un « effort de pêche standardisé » d'un ensemble d'unités de pêche sur une période de temps donnée, selon une quantité f d'actions de l'unité étalon. On a : $f = \sum_i p_i f_i$ où p_i est la puissance de l'unité i et f_i le nombre de ses actions pendant cette période. Cette standardisation est réalisée en vue d'estimer l'impact engendré sur la ressource par un ensemble d'actions de pêche (une action peut être par exemple un jour de mer) effectivement réalisées par un ensemble d'unités de pêche.
- Cette standardisation peut également être utilisée pour prévoir l'impact futur d'une combinaison quelconque d'actions de pêche entreprises par cet ensemble d'unités de pêche de puissances « connues ».
- Des variations spatio-temporelles peuvent exister, pouvant être caractérisées par des effets zones de pêche, périodes de pêche et par les interactions entre ces deux sources. L'examen de ces sources de variation est d'un intérêt évident pour l'étude du comportement de la ressource.

1.1. Données et modèles usuels

Un jeu de données caractéristique de ce type d'analyse est constitué par un ensemble de n observations d'actions de pêche, avec pour chaque action : la capture, le bateau, une zone de pêche et la date (ou période). Il convient d'insister sur le fait que le bateau permet généralement d'identifier sans ambiguïté l'unité de pêche ayant réalisé l'action de pêche. En parlant d'un effet bateau, on considère, au delà de ce bateau, le centre de décision, les moyens techniques et le savoir faire de l'unité de pêche correspondante. Dans toute la suite, effets « bateau » et « unité de pêche » sont donc un seul et même effet.

Les modèles s'écrivent généralement sous la forme :

$$\log(Y_{ijlk} + 1) = m + b_i + z_j + p_l + c_{jl} + e_{ijlk}$$

où Y_{ijlk} est la capture (en quintaux dans l'exemple traité plus loin) obtenue au cours d'une action, m est une constante, b_i est l'effet bateau, z_j l'effet zone, p_l l'effet pas de temps (pour distinguer des effets intra et interannuels, cet effet est parfois décomposé en effet année, effet mois ou autre période de durée inférieure à l'année, et interaction entre ces effets), c_{jl} est l'interaction spatio-temporelle.

La transformation Log présente un double intérêt (Robson 1966). Elle corrige la très forte asymétrie généralement observée dans les distributions de captures et elle stabilise les variances (avec de bonnes chances). Elle conduit

surtout à une interprétation naturelle des effets bateaux en termes de puissances de pêche, en imposant une contrainte de nullité pour l'effet du bateau étalon.

1.2. Problèmes et questions

Plusieurs problèmes se posent de façon naturelle.

On ne considère pas d'interactions entre bateau et autres sources de variation. Cela est imposé par la question sur les puissances de pêche : si une unité a une puissance p, elle doit toujours et partout pêcher p fois plus que l'unité étalon. Cependant l'existence de telles interactions ne peut être rejetée *à priori*.

- Les plans sont très déséquilibrés. Ils sont de fait choisis par les pêcheurs qui cherchent à maximiser leurs revenus et ont donc tout intérêt à pêcher à chaque moment là où les rendements sont les plus élevés. À l'extrême s'il existe une très forte hétérogénéité spatio-temporelle des rendements, toutes les unités de pêche devraient être à chaque période dans la zone où les rendements sont les plus élevés (ces zones peuvent dépendre des unités de pêche s'il existe des interactions entre effets bateaux et effets spatio-temporels). Si elles y parviennent, rien ne pourra être dit à partir de ces données sur la variabilité spatio-temporelle de la ressource. En d'autres termes, non seulement le plan d'expérience est très déséquilibré, mais les pêcheurs, en poursuivant leur propre objectif, ont intérêt à offrir un plan qui va à l'encontre des capacités d'analyse des variabilités spatio-temporelles. Cela conduit également à considérer que l'efficacité d'une unité de pêche peut être la combinaison de plusieurs aspects (Laurec 1977) : capacité à se rendre dans les zones les plus poissonneuses et capacité à pêcher les poissons présents.
- Les modèles utilisés sont à effets fixes. Si la justification de ce choix paraît évidente pour l'effet principal « bateau » au vu de la question relative aux puissances de pêche d'unités de pêche « identifiées et nommées », elle l'est beaucoup moins pour les effets spatio-temporels et pour les éventuelles interactions impliquant les bateaux.

Les possibilités offertes depuis quelques années par les logiciels et ordinateurs modernes dans le domaine des modèles mixtes peuvent être mises à profit. L'objet de cette présentation est de faire une première évaluation de l'intérêt des modèles mixtes dans ce domaine en tenant compte des problèmes décrits plus haut. Nous nous intéressons d'abord à un examen général des sources de variation en considérant des modèles entièrement aléatoires (Genstat 1993). Nous examinons ensuite les estimations (et leurs écarts-types) relatives aux rendements moyens prédits pour chaque navire selon divers modèles avec des statuts (fixes ou aléatoires) différents pour les effets spatio-temporels et les interactions entre ces effets et l'effet bateau. Ces statuts correspondent à diverses situations caractéristiques de la variabilité spatio-temporelle et de la connaissance qu'ont les unités de pêche de cette variabilité.

Cette présentation ne prétend en aucun cas couvrir l'ensemble des questions pouvant être abordées. Elle est faite à partir d'un jeu limité de données, réunissant 3444 observations réalisées par 14 navires thoniers dans l'Océan Atlantique Tropical, dans 6 zones au cours des 12 mois de l'année 1988. Chaque observation est relative à un jour de mer pour un navire. Les traitements présentés ont été réalisés avec le logiciel Genstat (Genstat 1993).

Il s'agit pour nous de mettre en évidence l'intérêt et l'importance de ces modèles dans un domaine où ils sont largement méconnus et de susciter d'éventuelles collaborations avec des statisticiens spécialistes.

2. Examen des sources de variation. Modèles à effets tous aléatoires

La première analyse consiste en un examen des diverses sources de variations en considérant aléatoires tous les effets (principaux et interactions). L'examen des données brutes montre une distribution affectée d'une masse importante de valeurs nulles. Le modèle linéaire est donc mal adapté et sera complété par deux analyses distinctes. L'une considère un modèle généralisé (loi binomiale) avec des valeurs 0 ou 1 selon que les captures sont nulles ou non. L'autre considère un modèle linéaire ajusté sur les seules données non nulles. Les résultats sont présentés dans le tableau 1.

TABLEAU 1. - Résultats des modèles à effets aléatoires.

Modèle	linéaire toutes données		généralisé binomial captures nulles (0) ou non (1)		linéaire exclusion des captures nulles	
Composante	Variance estimée	Écart Type	Variance estimée	Écart Type	Variance estimée	Écart Type
bateau	0.069	0.060	0.058	0.050	0.016	0.015
mois	0.001	0.057	0.011	0.059	0.017	0.024
zone	0.386	0.294	0.361	0.274	0.033	0.032
bateau.mois	0.001	0.074	0.000	0.067	0.059	0.036
bateau.zone	0.135	0.072	0.103	0.064	0.000	0.018
mois.zone	0.320	0.121	0.315	0.117	0.045	0.029
bat.mois.zone	0.386	0.120	0.331	0.105	0.000	0.041
résidus	5.590	0.142	1.000	Fixée	1.200	0.049

Les résultats des deux premiers ajustements indiquent des valeurs très faibles pour les variances des effets mois et mois.bateau. L'analyse des seules données non nulles conduit à des valeurs beaucoup plus réduites pour les variances

associées aux résidus et aux divers effets. La probabilité de trouver ou non du poisson apparaît donc un élément déterminant.

Par ailleurs la nullité des variances associées aux effets mois et bateau.mois peut s'interpréter comme une redistribution, chaque mois, des mêmes valeurs de rendements selon les diverses zones. L'ampleur de cette redistribution est mise en évidence par la variance de la composante mois.zone. Les variabilités associées aux interactions entre bateaux et zones et entre bateaux et mois.zone peuvent indiquer une efficacité spécifique à chaque bateau, pouvant dépendre d'un « comportement » de la ressource. La nullité de l'interaction « bateau.mois » peut alors encore indiquer que la redistribution des mêmes abondances selon les zones est également observée pour chaque unité de pêche.

3. Prédiction des impacts futurs

Il s'agit ici d'obtenir des prédicteurs des rendements par action de pêche pour chaque unité de pêche. Il s'agit d'étudier dans quelle mesure les rendements d'une unité de pêche dépendent de son comportement spatio-temporel, selon la nature de la variabilité correspondante. L'effet bateau est donc fixe. Le statut des autres effets doit être discuté à partir du comportement de la ressource et de celui des unités de pêche.

Si la répartition spatio-temporelle de l'abondance se reproduit à l'identique d'année en année, les effets spatio-temporels doivent également être considérés comme fixes. Dans ce cas il s'agit de savoir si les unités de pêche ont ou non une connaissance de ces effets, se traduisant également par une sélection des zones de pêche se renouvelant à l'identique d'année en année. Un prédicteur peut alors être obtenu en considérant un modèle à effets fixes accordant à chaque combinaison bateau.mois.zone un poids égal à celui présent dans le jeu de données utilisé (modèle 1 du tableau 2). Cette solution peut être adoptée en négligeant l'existence d'interactions entre les unités de pêche et les variations spatio-temporelles (modèle 2, tableau 2). Si on fait l'hypothèse que les unités de pêche ont un comportement spatial totalement aléatoire, on peut réaliser la prédiction en donnant des poids égaux aux divers effets spatio-temporels. Cette hypothèse est manifestement fausse au vu du nombre relativement important (10 sur 72 possibles) de combinaisons mois.zone pour lesquelles il n'y a pas de données. En supposant que ces désaffections correspondent à des décisions des unités de pêche on peut retenir le modèle 3 (tableau 2) avec des prédictions donnant un poids nul aux combinaisons d'effectif nul et un poids égal aux autres.

Si la répartition des mêmes effets spatio-temporels se reproduit de façon totalement aléatoire d'année en année, ces effets doivent être en première analyse considérés comme aléatoires. C'est ce que traduisent les modèles 4 et 5 tenant ou non compte d'interactions entre effets bateaux et effets spatio-temporels.

En fait les choses ne sont pas aussi évidentes. En effet les moyens actuels d'observation et de communication permettent dans une certaine mesure de connaître en temps réel la distribution spatio-temporelle de la ressource. Si

TABLEAU 2. – Définition de divers modèles et pondérations d'effets pour la prédiction des rendements moyens par navire.

- Modèle 1 Fixe : bateau*(mois*zone), pondération par les nombres de visites dans chaque combinaison bateau. zone.mois
- Modèle 2 Fixe : bateau+(mois*zone), pondération par les nombres de visites dans chaque combinaison zone.mois

Modèle 3 Fixe : bateau+(mois*zone), poids égaux pour les combinaisons « zone.mois » présentes, nuls pour les autres

- Modèle 4 Mixte : bateau fixe, bateau.(mois*zone) aléatoire

Modèle 5 Mixte : bateau fixe, mois*zone aléatoire.

Modèle	1		2		3		4		5	
Bateau	estimée	erreur	estimée	erreur	estimée	erreur	estimée	erreur	estimée	erreur
1	2.148	0.137	1.939	0.146	1.643	0.150	1.589	0.362	1.694	0.314
2	1.810	0.169	1.591	0.179	1.379	0.182	1.419	0.379	1.383	0.331
3	1.757	0.140	1.634	0.148	1.373	0.149	1.425	0.353	1.407	0.313
4	1.996	0.145	1.884	0.151	1.604	0.156	1.718	0.361	1.645	0.317
5	2.241	0.128	2.217	0.137	1.914	0.138	1.913	0.350	1.979	0.309
6	3.037	0.131	3.646	0.155	2.998	0.151	2.841	0.403	3.329	0.318
7	2.206	0.205	2.377	0.229	1.729	0.253	2.081	0.416	2.149	0.356
8	1.250	0.142	1.558	0.156	1.001	0.156	1.224	0.380	1.266	0.319
9	1.846	0.143	1.845	0.158	1.385	0.159	1.588	0.359	1.579	0.316
10	1.352	0.166	1.581	0.197	1.112	0.184	1.085	0.424	1.305	0.338
11	2.247	0.166	1.985	0.177	1.799	0.180	1.723	0.385	1.770	0.330
12	2.202	0.154	2.007	0.165	1.872	0.163	1.674	0.368	1.772	0.322
13	2.108	0.153	1.930	0.163	1.796	0.169	1.770	0.366	1.718	0.321
14	1.890	0.214	1.626	0.236	1.352	0.228	1.410	0.418	1.419	0.359

les mêmes effets spatiaux sont redistribués chaque mois et si les unités de pêche peuvent en avoir une connaissance instantanée, alors ces effets peuvent de nouveau être considérés comme fixes pour la prédiction des rendements et le recours au modèle 1 peut de nouveau être justifié.

Evidemment de très nombreuses situations intermédiaires peuvent être envisagées, par exemple en considérant des processus d'apprentissage, avec des unités qui acquièrent au cours du temps une meilleure connaissance de la répartition de la ressource. Dans ce cas, pour chaque unité, on peut imaginer

un passage d'un modèle mixte vers un modèle à effets fixes. Cette transition entre modèles peut bien sûr être différente selon les unités de pêche.

Les prédictions issues des divers modèles énoncés ci-dessus sont fournies, avec les erreurs types correspondantes dans le tableau 2. Les résultats sont conformes à l'attente selon laquelle les valeurs les plus élevées sont obtenues pour les modèles à effets fixes accordant des poids proportionnels aux nombres de visites présents dans le jeu de données. Ce modèle correspond en effet à l'hypothèse selon laquelle chaque unité de pêche exerce son activité dans les zones qui lui sont les plus favorables. De même, les variances des prédicteurs sont logiquement les plus élevées pour les modèles mixtes puisque dans ce cas la variabilité spatio-temporelle est prise en compte par les paramètres de variance des effets aléatoires. Ce résultat doit être nuancé selon les unités. La capacité à trouver les zones les plus intéressantes doit se traduire par des estimations de rendements plus élevées avec le modèle 1. C'est le cas pour l'unité 2, contrairement à l'unité 6, même si cette dernière obtient, globalement, de meilleurs résultats.

L'intérêt de ces comparaisons réside dans la mise en évidence des différences non négligeables en termes d'estimation – et de précision d'estimation – de l'impact de la pêche sur la ressource. La meilleure prédictibilité possible des prélèvements réalisés par chaque unité de pêche apparaît souhaitable dans le cadre de la gestion. Elle est également souhaitable pour pouvoir estimer le plus précisément possible les fluctuations de l'abondance de la ressource à partir des données de rendements. Mais le problème n'est pas de choisir la méthode qui donne les plus faibles variances d'estimation, mais celle correspondant au mieux à la « réalité ».

Le choix d'un modèle doit donc reposer sur de bonnes connaissances de la variabilité spatio-temporelle de l'abondance et sur de bonnes connaissances des stratégies des unités de pêche. Même si l'analyse de données telles que celles présentées ici peut contribuer à produire et illustrer de telles connaissances, elle ne peut y parvenir à elle seule. L'étude des exploitations halieutique impose d'articuler des problématiques de disciplines relevant des sciences de la nature, des sciences du milieu physique, des sciences sociales et de sciences de l'information, telle que la statistique dans le cas présenté ici.

Cette présentation conduit également à insister sur le besoin d'approfondir la recherche sur l'usage des modèles mixtes dans le domaine abordé. Il conviendrait ainsi de mieux tenir compte de la quantité importante de données nulles, peut-être en séparant l'analyse de la probabilité d'obtenir des captures non nulles de celle des distributions des données non nulles. Les déterminants de ces deux aspects peuvent en effet être de natures différentes. Il conviendrait également de bien prendre en compte le fait que le plan d'expérience est le produit des décisions des pêcheurs et que son étude constitue une problématique statistique et halieutique à part entière.

RÉFÉRENCES

GENSTAT 5 (1993), Reference manual, *Oxford Science Publications*, 796 p. *Bull.*, 3, 5-15.

LAUREC A. (1977), Analyse et estimation des puissances de pêche, *J. Cons. Int. Explor. Mer*, 37, 173-185.

ROBSON P. (1966), Estimation of the relative fishing power of individual ships, *ICNAF Res.*