JOURNAL DE LA SOCIÉTÉ FRANÇAISE DE STATISTIQUE

BENOÎT TRUONG-VAN

PATRICK PROUZET

Méthodes de captures successives : synthèse, quelques contributions, une application

Journal de la société française de statistique, tome 142, nº 3 (2001), p. 53-87

<http://www.numdam.org/item?id=JSFS_2001__142_3_53_0>

MÉTHODES DE CAPTURES SUCCESSIVES : SYNTHÈSE, QUELQUES CONTRIBUTIONS, UNE APPLICATION

Benoît TRUONG-VAN * et Patrick PROUZET **

RÉSUMÉ

Nous présentons une revue synthétique des méthodes d'estimation de la taille de la population d'une espèce animale à partir de captures successives. Nous apportons aussi de nouvelles contributions. En particulier, nous proposons une analyse statistique des estimateurs des modèles de régression de Leslie et Davis (1939), de DeLury (1947, 1951) et de leurs extensions. Pour la méthode de captures successives fondée sur la vraisemblance, nous donnons des conditions d'existence et d'unicité des estimateurs de maximum de vraisemblance pour des modèles dont les probabilités de capture non-constantes dépendent de deux paramètres. Ce résultat généralise celui de Bedrick (1994) établi pour les probabilités de capture constantes. Ensuite, nous considérons le problème de l'estimation de la taille de la population de saumons remontant le fleuve Adour à partir des captures de pêcheries professionnelles situées seulement sur l'Adour, des efforts de pêche et des évaluations de la séparation de la population de saumons entre l'Adour et la Nive. Pour traiter cette application réelle, nous introduisons et mettons en œuvre une version adaptée de la méthode de captures successives fondée sur le maximum de vraisemblance associée à un modèle dont les probabilités de capture dépendent de deux paramètres.

Mots-clés : Captures séquentielles, captures-efforts, moindres carrés, régression, vraisemblances.

Classification AMS : 62F10, 62F25.

ABSTRACT

A synthesis of successive removal methods for estimating the size of a population is given as well as some new contributions. For regression methods by Leslie and Davis (1939) and by DeLury (1947, 1951), we propose a statistical analysis of the estimators. For successive removal procedures based on the likelihood, we give conditions for the existence and uniqueness of the maximum likelihood estimators for models whose capture probabilities depend on two parameters. This extends the result obtained by Bedrick (1994) for the case of constant capture probabilities. Then, we consider the problem of estimating the size of the population of salmons going upstream the Adour river in the South West of France, from captures and capture efforts by fisheries, situated only in the Adour river, from numbers of

* Laboratoire de Statistique et Probabilités UMR C5583, Université Paul Sabatier F-31062 Toulouse Cedex 04, E-mail : truong@gmm.insa-tlse.fr

** IFREMER Station d'Hydrobiologie, F-64310 Saint-Pée sur Nivelle.

spawning area and from evaluated proportions of the separation of the salmon population between the Adour and Nive rivers. To deal with this genuine application, we introduce and perform some suitable version of the successive removal method based on the likelihood function associated with a model whose capture probabilities depend on two parameters.

Keywords : Sequential captures, Catch-effort, Least-squares, Regression, Likelihoods.

AMS classification : 62F10, 62F25.

1. Introduction

Les méthodes que l'on appelle indifféremment de retraits successifs ou de captures séquentielles ou captures-efforts consistent à estimer la taille d'une population animale à partir d'une suite de captures et d'efforts utilisés dans ces captures. Elles sont appelées par Seber (1973, 1992) méthodes de retrait, dans le cas particulier où les probabilités de capture et les efforts sont constants. Les hypothèses générales pour leur application sont d'une part que le système est fermé (autrement dit, la population est constante) pendant la période des retraits et d'autre part que, lors d'un retrait, chaque individu a la même probabilité d'être capturé, celle-ci pouvant varier d'un retrait à l'autre. D'après Gerdeaux (1987) citant Ricker (1975), ces méthodes ne sont applicables que dans des milieux restreints où la pêche est facile ou dans des pêcheries, quand le prélèvement réduit de façon appréciable la taille de la population présente. Leur origine semble remonter à Helland (1914), qui donnait une estimation de l'effectif d'une population d'ours norvégiens à partir des statistiques de chasse. Ces méthodes fournissent des estimations statistiquement moins précises que celles obtenues avec les techniques de capture-marquage-recapture, mais ont le grand avantage du moindre coût.

Une des notions essentielles à prendre en compte dans la problématique ci-dessus est la capturabilité, c'est-à-dire la probabilité p_k qu'un individu de la population soit capturé au $k$ième retrait, sous la condition d'effort de retrait constant. Les hypothèses sur cette notion peuvent être classées en trois catégories : modèle paramétrique, fonction de variables exogènes et modèle stochastique. Un modèle paramétrique consiste simplement à exprimer les p_k par une fonction π d'un paramètre θ qui peut être multidimensionnel, avec les deux cas extrêmes : les p_k sont constants ou libres; pour le premier cas, voir par exemple Leslie & Davis (1939), DeLury (1947), Moran (1951), Zippin (1956), Crittenden & Thomas (1989), Bedrick (1994), Hirst (1994) et, dans le cas de p_k libres, Braaten (1969). Schnute (1983) propose des fonctions π simples, quand le paramètre θ est à deux dimensions. Certains auteurs comme Yamakawa et coll. (1994) et Huggins & Yip (1997) proposent d'exprimer les p_k en termes d'autres variables du phénomène étudié. Beaucoup d'auteurs parmi lesquels Otis et coll. (1978), Carle & Strub (1978), Wang & Loneragan (1996), Norris & Pollock (1996), Pledger (2000), considèrent que la capturabilité est une variable inobservable (ou cachée) qu'ils modélisent alors par une v.a. à valeurs dans $[0, 1]$.

On peut regrouper la plupart des méthodes de retraits successifs en deux classes : moindres carrés et maximum de vraisemblance.

Les premiers développements ont conduit à des méthodes de régression linéaire, dont le principe consiste à établir des relations entre les captures et les paramètres d'intérêt tels que la taille N de la population et les probabilités de capture, puis à en déduire les estimations de ces paramètres par une régression linéaire. Leslie & Davis (1939) furent les premiers à proposer une estimation de N et de la probabilité p de capture supposée constante pour tous les retraits, à partir d'une régression linéaire entre les captures et leurs cumuls. Ricker (1975) et Crittenden (1983) ont étudié certaines propriétés des estimateurs fondés sur la régression linéaire. Crittenden & Thomas (1989) étendent la régression linéaire de Leslie & Davis (1939) au cas où les efforts de pêche ne sont pas constants, et proposent d'estimer les paramètres par les moindres carrés pondérés.

La méthode de Leslie et Davis pose un problème majeur : elle fournit parfois une estimation de N inférieure au total des captures ; Schnute (1983) l'explique par le fait que, les erreurs étant supposées centrées, le modèle permet des valeurs de captures négatives ou supérieures à N. Pour remédier à cette lacune, DeLury (1947, 1951) a proposé d'estimer (p, N) à partir d'une régression linéaire entre le logarithme des captures et les efforts de pêche ; il a ensuite étendu cette méthode au cas de deux probabilités de capture et de deux efforts de pêche. Les estimateurs proposés par DeLury (1947, 1951) restent biaisés, d'après Paloheimo (1961) et Chapman (1961), qui proposent alors un modèle à effort corrigé. Beverton & Holt (1956, 1957) ont établi un modèle général de relation entre les captures et les efforts de pêche. Utilisant ce modèle de Beverton & Holt (1956, 1957), Braaten (1969) propose les estimateurs dits de DeLury généralisés et généralisés pondérés.

Moran (1951) fut le premier à proposer le modèle multinomial pour les captures et à considérer l'estimation de (p, N) par la méthode du maximum de vraisemblance (MV). Zippin (1956, 1958) donne plusieurs formules pour calculer l'estimation MV $\widehat{N}$ de N, une technique graphique pour évaluer les estimations MV de (N, p) et une expression de la variance asymptotique de $\widehat{N}$. Higgins (1985) présente un algorithme pour obtenir les estimations MV de (N, p) avec un programme en BASIC. Une expression de la variance asymptotique de $\widehat{N}$ est obtenue par Seber & Le Cren (1967) et par Robson & Regier (1968) dans le cas particulier de deux captures, et par Junge & Libosvarsky (1965) puis par Seber & Whale (1970) pour trois captures successives. Schnute (1983) étend les travaux de Zippin (1956, 1958) en introduisant un modèle de probabilités de capture dépendant de deux paramètres et en donnant les équations de vraisemblance.

Hirst (1994) construit des intervalles de confiance pour N, en employant le ratio de log-vraisemblances de Kalbfleisch & Sprott (1970) mais sous la forme proposée par Harding et coll. (1984). A partir d'une étude en simulation pour différentes valeurs de N entre 20 à 300, pour 10 000 repliques, et pour deux et trois captures, cet auteur conclut que le ratio de log-vraisemblances donne des intervalles de confiance pour N bien meilleurs en termes de probabilité de couverture que ceux déduits de la normalité asymptotique de l'estimateur MV de N.

Carle & Strub (1978), Otis et coll. (1978), Norris & Pollock (1996), Pledger (2000) considérant que la capturabilité est une variable aléatoire, proposent d'estimer N et les paramètres ou la densité de probabilité de la capturabilité par le maximum de la vraisemblance marginale. En fait, la fonction-critère utilisée dans Otis et coll. (1978) n'est pas à proprement parler une vraisemblance marginale. Notons que l'approche fondée sur des modèles de variables cachées et sur des vraisemblances marginales est bien connue dans d'autres domaines, tels que l'estimation de processus ou de champs de Markov cachés (voir Chalmond 2000). D'ailleurs, nous avons adopté la terminologie de vraisemblance marginale, utilisée dans ce domaine plutôt que celle de vraisemblance pondérée de Carle & Strub (1978).

Mentionnons aussi une autre classe de méthodes de retraits successifs : certains auteurs tels que Schnute (1985, 1994) et Reed & Simons (1996), ont utilisé le filtre de Kalman pour estimer les paramètres des modèles de captures successives.

Gerdeaux (1987) donne des résultats comparatifs intéressants entre les méthodes fondées sur les régressions, le maximum de vraisemblance et le maximum de vraisemblance marginale de Carle & Strub (1978) pour des données de captures successives de truites.

Nous présentons dans ce travail, une synthèse critique des méthodes de captures successives à partir des revues antérieures sur la question (voir Ricker 1975, Otis et coll. 1978, Soms 1985, Seber 1986, Gerdeaux 1987, Cox 1983, Crittenden 1983, Crittenden & Thomas 1989, Pollock 1991, Prouzet et coll. 1996, Seber 1992, Schwarz & Seber 1999), et une analyse des modèles d'erreurs et des estimateurs de ces méthodes. En particulier, pour les méthodes de retraits successifs fondées sur les moindres carrés, nous montrons que les erreurs du modèle de régression linéaire de Leslie & Davis (1939) forment une suite non-stationnaire de différences de martingale et que les erreurs du modèle de DeLury (1947) sont biaisées. Nous donnons aussi des résultats nouveaux sur l'existence de solutions du maximum de certaines fonctions de vraisemblance ainsi qu'une application concrète. En effet, bien que la méthode du maximum de vraisemblance ait été couramment utilisée de longue date, soit pour estimer la taille N d'une population et les probabilités de capture, soit pour construire des intervalles de confiance pour N (voir Harding et coll. 1984 ou Cormack 1992), les premiers résultats théoriques concernant l'existence et l'unicité de ces estimateurs ne furent établis qu'en 1994 par Bedrick, et seulement pour le cas de probabilité de capture constante. Nous donnons d'abord une description des résultats de Bedrick (1994) ; puis, en vue de l'application que nous considérons, nous les étendons au cas de probabilités de capture dépendant de deux paramètres. La démarche de Bedrick (1994) pour établir l'existence et l'unicité des estimateurs de maximum de vraisemblance pour le cas de probabilité de capture constante ne peut pas s'étendre à la situation des probabilités de capture non-constantes ; nous avons adopté de ce fait une approche totalement différente.

Nous considérons ensuite le problème de l'estimation de la taille de la population de saumons remontant le fleuve Adour à partir des captures de pêcheries professionnelles situées seulement sur l'estuaire Adour, des efforts de pêche développés dans ces pêcheries, des décomptes des frayères et des évaluations de la séparation de la population des saumons entre l'Adour et son affluent, la Nive. Les méthodes de retraits successifs décrites ci-dessus ne s'appliquent pas directement à ce problème réel. Pour le traiter, nous proposons alors des adaptations de la méthode de captures successives fondée sur la vraisemblance associée à un modèle dont les probabilités de capture dépendent de deux paramètres.

2. Problème – Notations – Hypothèses

Le problème que nous considérons ici, consiste à estimer le nombre total des individus, c'est-à-dire la taille de la population d'une espèce animale, à partir d'une suite de retraits successifs et de données sur les efforts de ces retraits ou plus généralement sur les conditions de ces captures.

Pour la commodité du lecteur, nous énumérons d'abord ci-dessous les notations ainsi que les hypothèses qui seront utilisées dans toute la suite.

2.1. Notations

Nous désignons par N l'effectif total d'une population et par K le nombre total des retraits successifs. Nous utiliserons indifféremment le terme de retrait ou le terme plus spécifique de pêche, vu l'application que nous considérons dans ce travail.

Nous notons :

y_k le nombre des individus capturés à la $k$ième pêche (retrait), Y_k la variable aléatoire (v.a.) associée, $\mu_k = E(Y_k)$ pour $k = 1, \ldots, K$, où $E(\cdot)$ est l'espérance mathématique.

$s_0 = 0$, $s_i = \sum_{j=1}^{i} y_j = s_{i-1} + y_i$, $i = 1, \ldots, K$; $s = s_K$ (le nombre total de captures pour les K retraits) ; les S_i, $i = 0, \ldots, K$ sont les variables aléatoires associées.

$N_0 = N$, $N_i = N - s_{i-1}$, $i = 1, \ldots, K$,

p_k la probabilité pour un individu d'être capturé au $k$ième retrait, et $\mathbf{p} = (p_1, \ldots, p_K)'$, où V' désigne le transposé d'un vecteur V.

X la variable aléatoire (v.a.) associée à la capturabilité des individus de la population ; plus précisément, elle représente les événements qu'un individu pris au hasard de la population soit capturé ou non à l'un des K retraits ; X est donc à valeurs dans $\{1, \ldots, K+1\}$ avec la probabilité d'être capturé au $k$ième retrait $P(X = k) = p_k$, pour $k = 1, \ldots, K$ et la probabilité de ne pas être capturé $P(X = K+1) = p_{K+1} = 1 - \sum_{k=1}^{K} p_k$.

$Z_j = \sum_{i=S_{j-1}+1}^{N} I(X_i = j)$, $j = 1, \ldots, K+1$, avec une numérotation adéquate des individus, où pour $i = 1, \ldots, N$, X_i est la v.a. distribuée comme X

et correspondante à l'individu i, et où $I(A)$ est l'indicatrice d'un événement A.

f_k l'effort (évalué) de la $k$ième pêche, F_k la v.a. associée, $\phi_k = E(F_k)$ pour $k = 1, \ldots, K$, où, d'après Poinsard & Le Guen (1975), l'effort de pêche est défini comme la mesure de l'ensemble des moyens de capture mis en œuvre par les pêcheurs pendant un intervalle de temps déterminé (les pêcheries ont en général des efforts de pêche différents) ; les F_k sont donc des variables explicatives du problème considéré. L'effort d'une pêcherie peut être appréhendé soit à partir de quantités 'nominatives" (nombre et type de bateaux, nombre de pêcheurs, ...) soit à partir de quantités effectives (temps de pêche, taille des filets, nombre de sorties par jour de pêche, nombre de coups de filets, ...). Pour l'application que nous considérons, les efforts de pêche sont évalués à partir de quantités effectives.

y_k/f_k le nombre de captures par unité d'effort (CPUE) de la $k$ième pêche.

$\mathcal{F}_k$ la σ-algèbre engendrée par les v.a. Z_j, $j = 1, \ldots, k$; en pratique, $\mathcal{F}_k$ représente l'information contenue dans k premiers Z_j.

$E_\theta(\cdot|\mathcal{F}_k)$ (respectivement $E_\theta(\cdot|\mathcal{F}_k, F_{k+1})$) l'espérance conditionnelle relative à $\mathcal{F}_k$, (respectivement relative à $\mathcal{F}_k$ et à F_{k+1}) par rapport à une loi de probabilité P_θ dépendant des paramètres θ, autrement dit la moyenne théorique sachant l'information disponible jusqu'au $k$ième retrait.

La convention suivante est adoptée dans toute la suite : pour toute suite (a_i), $\sum_{j=i_1}^{i_2} a_j = 0$ si $i_2 < i_1$.

2.2. Hypothèses

On considère d'abord pour toutes les méthodes de retraits successifs le cadre général suivant :

(H0-1) La population est constante sur la période des retraits successifs étudiés (ni natalité, ni mortalité, ni émigration, ni immigration) ; autrement dit, le paramètre N est supposé fixe pendant tous les retraits.

Une première notion-clé dans le phénomène étudié est la capturabilité, c'est-à-dire la probabilité qu'un individu de la population soit capturé sous la condition d'effort de retrait constant. Pour cette notion, l'hypothèse générale suivante est souvent considérée :

(H0-2) Tous les individus ont la même probabilité d'être capturés dans un retrait donné (en particulier pas de différence selon la taille). Cette probabilité de capture peut cependant différer d'un retrait à l'autre. Plus précisément, la probabilité pour qu'un individu de la population soit capturé au $k$ième retrait est $P(X = k) = p_k$, pour $k = 1, \ldots, K$, et la probabilité pour qu'il ne soit capturé à aucun de ces retraits est $P(X = K + 1) = 1 - \sum_{k=1}^{K} p_k$.

Pour certaines méthodes, on suppose de plus que :

(H0-3) Les individus de la population ont un comportement tel que les v.a. associées X_i sont indépendantes en probabilité.

Seules les hypothèses (H0-1) et (H0-2) sont utilisées dans les méthodes des moindres carrés. L'hypothèse supplémentaire (H0-3) est requise pour les méthodes de maximum de vraisemblance. La conjugaison des trois hypothèses (H0-1), (H0-2) et (H0-3) signifie que les v.a. $X_i, i = 1, \ldots, N$ forment un échantillon de taille N, issu de la v.a. X. Ceci implique que les v.a. Z_k, $k = 1, \ldots, K+1$ suivent une loi multinomiale de paramètres $(N, p_1, \ldots, p_{K+1})$.

Donnons maintenant les hypothèses plus spécifiques sur les probabilités de captures. Elles peuvent être classées en trois catégories : modèle paramétrique, fonction de variables exogènes et modèle stochastique.

Parmi les modèles paramétriques pour les p_k, nous retenons l'une des hypothèses suivantes, qui sont adoptées dans les applications réelles et dans la quasi-totalité de la littérature (voir Seber 1973). Puisque dans la plupart des cas d'application réelle, le nombre des observations (à savoir des captures) est de six au maximum, il est statistiquement judicieux de restreindre les probabilités de capture à dépendre seulement d'un ou de deux paramètres.

(H1-1) Les probabilités de capture sont constantes pour les retraits

$$p_k = p, \quad k = 1, \ldots, K \ ;$$

(H1-2) Les probabilités de capture sont telles que

$$p_i = \pi(i; p, \alpha)$$

où π est une fonction donnée de i (généralement décroissante) dépendant de deux paramètres $p = p_1$ et α.

Schnute (1983) propose la fonction π suivante (appelée modèle 3)

(H1-2-1)

$$p_i = p_1 + (p - p_1)\left(1 - \alpha^{i-1}\right) \tag{1}$$

où $0 \leqslant \alpha \leqslant 1$, avec les cas extrêmes :

(i) $\alpha = 1$: modèle de Zippin (H1-1), les probabilités de captures sont identiques pour tous les retraits.

(ii) $\alpha = 0$: modèle 2 de Schnute.

On peut aussi considérer (H1-2) sous l'une des formes

(H1-2-2) $p_i = p\alpha^{(i-1)}$, $0 < \alpha < 1$, (décroissance exponentielle) ;

(H1-2-3) $p_i = (i-1)\alpha + p$ (décroissance linéaire) ;

(H1-2-4) $p^{(i-1)\alpha+1}$, $0 < \alpha \leqslant 1$ (décroissance polynomiale).

Certains auteurs (voir Yamakawa et coll. 1994 ou Huggins & Yip 1997) proposent d'exprimer les p_k en termes de d'efforts de retrait ou/et des variables auxiliaires T_k telles que les durées d'échantillonnage, la taille, l'âge, le sexe ou l'espèce des individus, autrement dit

(H1-3)

$$p_k = \widetilde{\pi}(\phi_k, T_k), \quad k = 1, \ldots, K$$

où $\widetilde{\pi}$ est une fonction donnée. D'après Otis et coll. (1978, p. 44), le cas particulier de (H1-3) le plus couramment considéré est simplement $p_k = 1 - e^{-\lambda\phi_k}$, $k = 1, \ldots, K$, où λ est le coefficient de Poisson de la capturabilité.

Nous discuterons des hypothèses (H1-1), (H1-2) et (H1-3) dans les paragraphes 3.3 et 4.1.

Beaucoup d'auteurs, tels Otis et coll. (1978), Carle & Strub (1978), Wang & Loneragan (1996), Norris & Pollock (1996), Pledger (2000), considèrent que la capturabilité est une variable inobservable (ou cachée) et retiennent alors au lieu de (H0-2), l'hypothèse suivante :

(H0-4) La probabilité de capture est une v.a. à valeurs dans $[0, 1]$.

Nous considérons aussi l'hypothèse naturelle ci-dessous de l'indépendance en probabilité entre les v.a. des efforts de pêche et celles associées à la capturabilité :

(H2) Les erreurs de mesure ou d'évaluation des efforts de pêche $F_k - \phi_k$, $k = 1, \ldots, K$ sont des v.a. indépendantes en probabilité des v.a. $Z_i - E(Z_i)$, $i = 1, \ldots, K + 1$ qui représentent les erreurs de modèle de capturabilité des individus de la population.

Dans les méthodes fondées sur la vraisemblance, les aléas du phénomène sont modélisés en supposant une loi multinomiale pour les v.a. Y_k et l'une des deux hypothèses (H0-2) et (H0-4). Les premiers travaux sur les méthodes de retrait basées sur les moindres carrés retiennent les hypothèses suivantes sur les erreurs de modèle :

(H3) Les erreurs de modèle de régression sont indépendantes et de même loi P_ε ;

(H3-1) P_ε est normale $N(0, \sigma^2)$;

(H3-2) P_ε est lognormale de paramètres $(0, \sigma^2)$.

La plupart des auteurs soulignent l'inadéquation de l'hypothèse (H3), mais il n'existe quasiment aucune étude proposant un modèle de corrélation pour ces erreurs de modèle de régression. Nous discuterons de ces hypothèses dans le paragraphe 3.3.

3. Méthodes des moindres carrés

3.1. Un modèle général de régression

Le principe fondamental de toutes les méthodes de captures séquentielles fondées sur les moindres carrés est très simple. Il stipule l'égalité entre le nombre de captures ramené à l'unité d'effort (CPUE) et la proportion d'individus capturés. Jusqu'à maintenant, cette égalité est exprimée en moyenne, c'est-à-dire :

$$\frac{\mu_k}{\phi_k} = p_k \left(N - \sum_{j=1}^{k-1} \mu_j \right), \quad k = 1, \ldots, K. \tag{2}$$

Nous proposons de l'exprimer en termes de v.a. comme suit :

$$Y_k = F_k Z_k, \quad k = 1, \ldots, K; \tag{3}$$

autrement dit, la v.a. nombre de captures du $k$ième retrait est le produit de la v.a. effort de ce retrait et de celle représentant le nombre d'individus de la population susceptibles d'être pris à ce retrait, par le seul fait de leur capturabilité. La relation (3) décompose clairement les aléas ou erreurs de modèle sur les Y_k en deux contributions de nature différente : erreurs de modèle ou de mesure des efforts de pêche, et variabilité aléatoire due à la capturabilité. De plus et surtout, on peut déduire aisément et rigoureusement de (3) un modèle général de régression. En effet, sous les hypothèses (H0-1), (H0-2) et (H2), la meilleure prévision de Y_k (au sens de la variance) à partir du passé et pour un effort F_k de la $k$ième pêche, considéré comme une variable aléatoire exogène, s'écrit

$$E_{(N,\mathbf{p})}(Y_k | \mathcal{F}_{k-1}, F_k) = p_k F_k (N - S_{k-1}). \tag{4}$$

En désignant par ε_k, l'erreur de prédiction de Y_k à partir de $\mathcal{F}_{k-1}$ et de F_k, c'est-à-dire en posant

$$\varepsilon_k = Y_k - E_{(N,p)}(Y_k | \mathcal{F}_{k-1}, F_k), \tag{5}$$

la relation (4) prend la forme suivante d'un modèle général de régression

$$Y_k = F_k p_k (N - S_{k-1}) + \varepsilon_k \tag{6}$$

dans lequel les F_k sont les variables exogènes et les p_k des paramètres.

On déduit de (6) le modèle général suivant pour les moyennes théoriques μ_k de capture.

PROPOSITION 1. — *Sous les hypothèses (H0-1), (H0-2) et (H2), on a pour tout* $k = 1, \ldots, K$

$$\mu_k = N \rho_k \prod_{i=1}^{k-1} (1 - \rho_i), \tag{7}$$

où $\rho_k = \phi_k p_k$, $k = 1, \ldots, K$.

La proposition se démontre en déduisant l'égalité (2), de la relation-clé (4) et de l'hypothèse (H2) en établissant par récurrence la relation suivante

$$\sum_{j=1}^{k-1} \mu_j = N \left\{ 1 - \prod_{i=1}^{k-1} (1 - \rho_i) \right\}.$$

Remarque 1

(i) La relation (7) montre que les efforts de pêche ne sont appréciés que de façon relative les uns par rapport aux autres; on peut donc choisir l'un d'entre eux pour unité. Ces efforts jouent un rôle multiplicatif par rapport aux probabilités de capture.

(ii) Si on suppose l'hypothèse (H0-3) en plus de (H0-1) et (H0-2), et si $\phi_k = 1$, $k = 1, \ldots, K$, alors la relation (7) est bien connue puisque les v.a. Y_k, $k = 1, \ldots, K+1$ sont alors de loi multinomiale de paramètres $(N, p_1, \ldots, p_{K+1})$.

3.2. Une revue des méthodes dites de régression

Les méthodes actuelles de captures séquentielles fondées sur les moindres carrés, reposent sur une relation de la forme

$$\frac{\mu_k}{\phi_k} = \psi\left(\sum_{j=1}^{k-1} \mu_j, \sum_{j=1}^{k-1} \phi_j, \phi_k\right)$$

liant les CPUE théoriques μ_k/ϕ_k aux captures moyennes cumulées $\sum_{j=1}^{k-1} \mu_j$ ou aux efforts de pêche ϕ_j, $j = 1, \ldots, k$, où ψ est une fonction linéaire ou exponentielle. Elles se rangent donc en deux classes : régressions linéaires simples (Leslie & Davis 1939) ou pondérées (Crittenden & Thomas 1989) et régressions log-linéaires (DeLury 1947, 1951, Paloheimo 1961, Chapman 1961). Toutes ces méthodes peuvent être dérivées du modèle général (7) des moyennes théoriques des captures. C'est l'objet des premiers paragraphes ci-dessous.

3.2.1. Régression linéaire de Leslie et Davis

Considérons la situation où les p_k sont tous égaux à une constante $p \in]0, 1[$ et que $\phi_k = 1, k = 1, \ldots, K$. D'après (7) et (6), il est immédiat que :

$$\mu_k = -p \sum_{j=1}^{k-1} \mu_j + pN$$

et

$$Y_k = -pS_{k-1} + pN + \varepsilon_k. \tag{8}$$

L'équation (8) est communément appelée modèle de régression de Leslie & Davis (1939). Schnute (1983) l'avait établie à partir $E_{(N,p)}(Y_k|S_{k-1}) = p(N - S_{k-1})$. Si $\widehat{p}$ et $\widehat{\beta}$ désignent les estimations des moindres carrés ordinaires respectivement (resp.) de p et $\beta = pN$, l'estimation de N est

$$\widehat{N} = \widehat{\beta}/\widehat{p}. \tag{9}$$

Cet estimateur de N défini à partir d'un ratio est en général biaisé. Ce biais peut être corrigé ou plutôt réduit en utilisant des techniques usuelles (jackknife, méthode Delta,...).

Mentionnons que, pour estimer p et N à partir de (8), Hayne (1949) a proposé une simple inspection visuelle et que les méthodes des moindres carrés (MC) ne sont utilisées qu'à partir de DeLury (1947) et Zippin (1956, 1958).

Sous les hypothèses supplémentaires (H3) et (H3-1) que les erreurs de modèle ε_k sont des v.a. indépendantes et de loi $N(0,\sigma^2)$, DeLury (1951), repris par Ricker (1975), montre que les extrêmités d'un intervalle de confiance asymptotique au niveau de confiance $1-\alpha$ pour $\widehat{N}$ sont solutions de l'équation du second degré :

$$N^2(q^2 - t^2 R_{yx} C_{22}) - 2N(q^2 N_0 - t^2 R_{yx}^2 C_{12}) + (q^2 N_0^2 - t^2 R_{yx}^2 C_{11}) = 0,$$

où

$$C_{11} = \frac{K(\sum\limits_{l=1} S_{l-1}^2)}{K(\sum\limits_{l=1} S_{l-1}^2 - (\sum\limits_{l=1} S_{l-1})^2)},$$

$$C_{12} = \frac{K(\sum\limits_{l=1} S_{l-1})}{K(\sum\limits_{l=1} S_{l-1}^2 - (\sum\limits_{l=1} S_{l-1})^2)},$$

$$C_{22} = \frac{K}{(\sum\limits_{l=1} S_{l-1}^2 - (\sum\limits_{l=1} S_{l-1})^2)},$$

$$R_{yx}^2 = \frac{(\sum\limits_{l=1} Y_l^2 - \frac{(\sum\limits_{l=1} Y_l S_{l-1})^2}{\sum\limits_{l=1} S_{l-1}^2})}{K-2},$$

et t est le quantile d'ordre $\alpha/2$ d'une v.a. de Student à $K-2$ degré de liberté.

Nous discutons de la validité des hypothèses ci-dessus dans le paragraphe 3.3.1.

3.2.2. Régression linéaire pondérée de Crittenden et Thomas

Crittenden & Thomas (1989) étendent le modèle de régression de Leslie & Davis (1939) au cas où les efforts de pêche ne sont pas constants, mais toujours avec des probabilités de capture constantes. Leur modèle, qui est un cas particulier de (6), s'écrit sous la forme

$$\frac{Y_k}{F_k} = -p\widetilde{S}_{k-1} + pN + \widetilde{\varepsilon}_k, \tag{10}$$

où $\widetilde{S}_{k-1}$ est le cumul corrigé des captures juqu'au $(k-1)$ième retrait et $\widetilde{\varepsilon}_k$ le terme d'erreur.

Les paramètres p et $\beta = pN$ sont estimés par régression linéaire pondérée, où la matrice de poids est l'inverse d'une estimation d'une matrice non-diagonale Γ des variances-covariances. Les composantes de Γ correspondant aux variances des CPUE Y_k/F_k sont estimées à partir des régressions suivantes

$$\frac{Y_{k,j}}{F_{k,j}} = \alpha_k F_{k,j} + e_{k,j},$$

où α_k est un paramètre, $Y_{k,j}$ (resp. $F_{k,j}$) est la capture (resp. la v.a. effort) de la $j$ième unité dans le k ième retrait, une unité étant par exemple un coup de filet. La plupart des autres composantes de la matrice Γ sont estimées par des techniques de bootstrap.

Pour obtenir l'estimation finale de N, Crittenden & Thomas (1989) appliquent une correction de biais basée sur la méthode Delta (voir Seber 1973), sur l'estimation $\widehat{\beta}/\widehat{p}$ de N, déduite de la régression.

3.2.3. Modèles de DeLury

D'après Schnute (1983), un des problèmes majeurs de la méthode de Leslie & Davis (1939) est qu'elle peut fournir une estimation de N inférieure au total des captures à cause de l'hypothèse que les erreurs sont supposées centrées. Pour remédier à cette lacune, DeLury (1947) a proposé un estimateur de (p, N) à partir d'un modèle linéaire entre le logarithme des CPUE théoriques et les efforts de pêche. Ce modèle ainsi que ses extensions s'appuient sur l'hypothèse suivante que la fraction de stock prélevée par chaque unité d'effort est petite, plus précisément que

$$1 - \rho_i \approx e^{-\rho_i}, \quad i = 1, \ldots, K. \tag{11}$$

Le modèle approché de (7) s'écrit alors

$$\mu_k = p_k \phi_k N \exp\left(-\sum_{j=1}^{k-1} \rho_j\right). \tag{12}$$

Dans le cas particulier où $p_k = p$ constant, on obtient après transformation logarithmique de (12), le modèle suivant de DeLury (1947) :

$$\ln\left(\frac{\mu_k}{\phi_k}\right) = -p\sum_{j=1}^{k-1}\phi_j + \ln(pN). \tag{13}$$

Si $\widehat{p}$ et $\widehat{\alpha}$ désignent les estimateurs MC ordinaires respectivement de p et de $\alpha = \ln(pN)$, associés au modèle (13), alors l'estimateur dit de DeLury $\widehat{N}$ de N est :

$$\widehat{N} = \frac{\exp\widehat{\alpha}}{\widehat{p}}. \tag{14}$$

Comme nous l'avons remarqué pour la méthode de Leslie & Davis (1939), cet estimateur de DeLury est en général biaisé.

De (12), on peut aussi déduire le modèle de DeLury (1951) dit à deux capturabilités et à deux efforts, c'est-à-dire pour $p_k = p, \phi_k = \phi$, $1 \leqslant k \leqslant K_1$ et $p_k = p', \phi_k = \phi'$, $K_1 + 1 \leqslant k \leqslant K$.

3.2.4. Modèle de DeLury à efforts corrigés

Paloheimo (1961) et Chapman (1961) argumentent que l'estimateur déduit du modèle de DeLury (1947) est biaisé et proposent alors la relation suivante, appelée modèle de DeLury à efforts corrigés :

$$\ln\left(\frac{\mu_k}{\phi_k}\right) = -p\left(\sum_{j=1}^{k-1}\phi_j + \frac{\phi_k}{2}\right) + \ln(pN). \tag{15}$$

Pour l'établir, ces auteurs utilisent la même approximation (11) mais écrivent le modèle approché de (7) plutôt sous la forme

$$\frac{\mu_k}{\phi_k} = pN\frac{1-\exp(-p\phi_k)}{p\phi_k}\exp\left(-p\sum_{j=1}^{k-1}\phi_j\right);$$

puis, pour des ϕ_k petits, en approchant une nouvelle fois, $(p\phi_k)^{-1}\{1 - \exp(-p\phi_k)\}$ par $\exp(-p\phi_k/2)$, ils obtiennent

$$\frac{\mu_k}{\phi_k} = pN\exp\left(-p\sum_{j=1}^{k-1}\phi_j\right)\exp\left(-\frac{p\phi_k}{2}\right),$$

donc (15), en passant au logarithme.

3.2.5. Modèles de DeLury généralisés

En supposant que

$$\sum_{i=1}^{k}\rho_i \approx 1-\exp\left(\sum_{i=1}^{k}\rho_i\right) \quad \text{et} \quad \prod_{i=1}^{k-1}(1-\rho_i) \approx \exp\left(-\sum_{i=1}^{k-1}\rho_i\right),$$

on peut écrire le modèle approché de (7) sous la forme

$$\mu_k = \frac{\rho_k}{\rho_k + \sum_{i=1}^{k-1}\rho_i}\left\{1-\exp\left(\sum_{i=1}^{k}\rho_i\right)\right\}\left\{N\exp\left(-\sum_{i=1}^{k-1}\rho_i\right)\right\}. \tag{16}$$

L'équation ci-dessus est dite de Beverton & Holt (1956, 1957), lesquels en donnent l'interprétation suivante : chaque retrait μ_k est le produit de la fraction

$$\frac{\rho_k}{\rho_k + \sum_{i=1}^{k-1}\rho_i}$$

des captures prises lors du $k$ième retrait, du taux du retrait

$$1-\exp\left(\sum_{i=1}^{k}\rho_i\right)$$

et du stock

$$N \exp\left(-\sum_{i=1}^{k-1} \rho_i\right)$$

restant encore disponible avant le $k$ième retrait.

S'appuyant sur l'équation (16) ci-dessus de Beverton & Holt, Braaten (1969) propose les estimateurs suivants de la taille N_k de la population après k retraits.

(i) Estimateur de DeLury généralisé :

$$\widehat{N}_k^{(g)} = \frac{F_k \exp \widehat{\alpha}}{1 - \exp(-F_k \widehat{p})}. \tag{17}$$

On retrouve l'estimateur de DeLury à partir de la relation ci-dessus pour $k = 0$ et de l'approximation $F_k^{-1}\{1 - \exp(-F_k \widehat{p})\} \approx \widehat{p}$.

(ii) Estimateur de DeLury généralisé pondéré (par les efforts de pêche) :

$$\widehat{N}_k^{(p)} = \sum_{i=1}^{k} \widehat{N}_i^{(g)} \frac{\sum_{j=1}^{i} F_j}{\sum_{i=1}^{k}\sum_{j=1}^{i} F_j}.$$

Mentionnons au passage une procédure à deux étapes, proposée par Cormack (1993), pour estimer les paramètres de modèles log-linéaires pour des marquages-recaptures. Le modèle de Cormack (1993) analogue à (13) est $\ln(\mu_k) = -p(k-1) + \ln p + \ln N$; c'est le cas particulier de (13) où $\phi_j = 1$, $j = 1, \ldots, K$ et où les Y_k, $k = 1, \ldots, K$ sont des v.a. indépendantes et de loi de Poisson. Cormack (1993) propose d'estimer par maximum de vraisemblance les paramètres définissant les μ_k, puis d'en déduire l'estimation de (p, N) par la méthode des moments.

3.3. Discussion des méthodes de régression

3.3.1. Une analyse des modèles de régression linéaire

Notons tout d'abord que les erreurs ε_k de modèle telles que nous les avons définies dans (5) sont différentes des écarts $Y_k - E(Y_k)$. Cette différence pourrait expliquer l'appelation de « régression linéaire conditionnelle » que Crittenden & Thomas (1989) ont donnée à leur méthode.

La proposition suivante donne quelques résultats sur les erreurs ε_k du modèle de Leslie & Davis (1939).

PROPOSITION 2. — *Sous les hypothèses (H0-1) et (H0-2), on a pour* $k = 1, \ldots, K$

(i) $E(\varepsilon_k) = 0$;

(ii) $Var(\varepsilon_k) = Var(Y_k) = E(Y_k^2) - \mu_k^2$, *où* μ_k *est défini par (7)*
(iii) et pour $h = 1, \ldots, K$ *tel que* $k + h \leqslant K$, $Cov(\varepsilon_k, \varepsilon_{k+h}) = 0$.

Preuve de la proposition : La relation (i) est triviale, et pour l'assertion (ii), on a $Var(Y_k|\mathcal{F}_{k-1}) = E((Y_k - E(Y_k|\mathcal{F}_{k-1}))^2|\mathcal{F}_{k-1}) = E(\varepsilon_k^2|\mathcal{F}_{k-1}) = Var(\varepsilon_k|\mathcal{F}_{k-1})$ et $Var(Y_k) = E(Var(Y_k|\mathcal{F}_{k-1}))$, donc, $Var(Y_k) = Var(\varepsilon_k)$. L'assertion (iii) est immédiate puisque pour $l = 1, \ldots, K$, les ε_l étant en fait les innovations des Y_l relatives à $\mathcal{F}_{l-1}$ i.e. $\varepsilon_l = Y_l - E_{(N,p)}(Y_l|\mathcal{F}_{l-1})$, on a

$$E(\varepsilon_l|\mathcal{F}_{l-1})) = 0. \tag{18}$$

Par conséquent, pour $h = 1, \ldots, K$ tel que $k + h \leqslant K$,

$$E(\varepsilon_k \varepsilon_{k+h}) = E(\varepsilon_k E(\varepsilon_{k+h}|\mathcal{F}_{k+h-1})) = 0.$$

Remarque 2

On a en fait une propriété plus forte que l'énoncé (iii) de la proposition : les $\varepsilon_l, l = 1, \ldots, K$ forment une suite non-stationnaire de différences de martingale relatives à $\mathcal{F}_l$, dont les variances sont celles des Y_k.

Commentaires

(i) L'hypothèse retenue par DeLury (1951) et Ricker (1975), que les erreurs de modèle ε_k sont des v.a. indépendantes et de loi $N(0, \sigma^2)$, pour établir un intervalle de confiance asymptotique pour N (voir paragraphe 3.2.1) n'est pas judicieuse, car les ε_k ont la propriété énoncée dans la remarque 2 : ils ne sont pas nécessairement indépendants mais vérifient la condition (18) et leurs variances ne sont pas constantes. Schnute (1983), Crittenden (1983), Crittenden & Thomas (1989) ont déjà remarqué qu'il n'est pas réaliste de supposer que les erreurs de modèle sont indépendantes et identiquement distribuées, en argumentant que les CPUE sont hétéroscédastiques et que les variables cumulées ne sont pas indépendantes.

(ii) L'estimateur de N obtenu par la méthode de régression de Leslie & Davis (1939) est en général biaisé, car il est défini par le ratio (9).

(iii) Dans le cas où les efforts de pêche sont constants, la proposition ci-dessus indique clairement que si on peut disposer d'un modèle pour les $E(Y_k^2)$, une régression pondérée de poids $1/Var(Y_k)$ est bien plus judicieuse qu'une régression simple. On pourrait penser à prendre pour $Var(Y_k)$, les variances d'une loi multinomiale ou hypernomiale associée mais dans ce cas compte tenu du point (ii) ci-dessus, ne serait-il pas préférable d'estimer N par la méthode de maximum de vraisemblance ?

(iv) Pour les situations plus générales où les efforts de pêche ne sont pas constants, la méthode de Crittenden & Thomas (1989) constitue une approche adéquate (même si sa mise en œuvre est lourde) quand les données disponibles sont suffisamment représentatives pour que les estimations obtenues par les techniques de boostrap soient convergentes. Une analyse statistique de leur méthode demande des développements qui dépassent le cadre de ce travail.

3.3.2. Une analyse des estimateurs de DeLury

Reprenons le modèle (13) de DeLury (1947) et analysons les erreurs du modèle de régression associé.

Posons $\varepsilon'_k = Y_k/F_k - \mu_k/\phi_k$ et $\varepsilon''_k = F_k - \phi_k$, $k = 1, \ldots, K$. Le modèle de régression déduit de (13) est donc

$$\frac{Y_k}{F_k} = Np \exp\left(-p\sum_{j=1}^{k-1} F_j\right) \exp\left(p\sum_{j=1}^{k-1} \varepsilon''_j\right) + \varepsilon'_k$$

que l'on peut encore écrire sous la forme

$$\frac{Y_k}{F_k} = Npe^{-p\sum_{j=1}^{k-1} F_j} e_k, \quad k = 1, \ldots, K$$

ou en passant au logarithme, pourvu que presque sûrement (p.s.) les $\xi_k \geqslant 0$,

$$\ln\frac{Y_k}{F_k} = -p\sum_{j=1}^{k-1} F_j + \ln(Np) + \ln(e_k), \quad k = 1, \ldots, K, \tag{19}$$

où

$$e_k = \xi_k \exp\left(p\sum_{j=1}^{k-1} \varepsilon''_j\right), \tag{20}$$

et

$$\xi_k = 1 + \frac{\varepsilon'_k}{Np\exp\left(-p\sum_{j=1}^{k-1} \phi_j\right)}.$$

D'après (20), les v.a. $\ln(e_k)$, $k = 1, \ldots, K$ sont clairement corrélées et l'hypothèse (H3-2) que ces v.a. sont lognormales de paramètres $(0, \sigma^2)$ est inappropriée, car elle revient à supposer que les v.a. e_k, $k = 1, \ldots, K$ sont normales $N(0, \sigma^2)$, alors qu'elles ne sont pas centrées.

Il s'ensuit de (3) que les aléas $\varepsilon'_k = Z_k - E(Z_k)$ ne concernent que la capturabilité des individus et sont donc indépendants en probabilité des ε''_j. Il est plus judicieux de supposer que les ε''_j sont indépendants. Alors la structure de corrélation des $\ln(e_k)$ est la suivante : chaque $\ln(e_k)$ est la somme d'une martingale $p\sum_{j=1}^{k-1} \varepsilon''_j$ et d'une v.a. non-centrée $\ln(\xi_k)$ indépendante de la martingale ; ici, la martingale est simplement une somme de $k-1$ v.a. indépendantes. Ce résultat montre en particulier que les erreurs du modèle de DeLury sont biaisés car les $E\ln(e_k) = E\ln(\xi_k)$ ne sont pas nuls. Pour corriger ces biais, une solution simple consisterait à ajouter à chaque $\ln(Y_k/F_k)$ un terme de recentrage adéquat. Ce type de biais se retrouve aussi dans les modèles log-linéaires d'Evans et coll. (1994) pour des recaptures multiples. Pour réduire ces biais, Evans & Bonett (1994) proposent d'ajouter, pour $k = 1, \ldots, M$, $(1/2)^{k-1}$ à la $k$ième composante du vecteur de dimension M des observations du modèle d'Evans et coll. (1994). Pour le modèle de DeLury (19), le choix des valeurs judicieuses de réduction de biais est bien plus délicat car ces biais dépendent des efforts et des lois de probabilité des ε'_k. En effet, prenons $\xi_k = \max(W_k, 0)$, où W_k est une v.a. normale $N(1, \sigma_k^2)$. Alors $E\ln\xi_k$ a pour valeur 0.3, 0.1 ou 0.05 selon que σ_k prend la valeur 1, 0.5 ou 0.1.

Le modèle log-linéaire associé au modèle de DeLury corrigé (voir Paloheimo 1961 et Chapman 1961) a des biais similaires à ceux de (19). Ceci peut expliquer en partie le constat fait par Moran (1971), Seber (1977), Hodges & Moore (1972), Davies & Hutton (1975) et autres, que les estimateurs obtenus par la régression log-linéaire associée au modèle de DeLury corrigé, restent biaisés.

4. Modèle multinomial et vraisemblances

Nous avons vu que les estimateurs obtenus par les méthodes de retraits successifs fondées sur les moindres carrés, à l'exception de celle de Crittenden & Thomas (1989), n'ont pas de bonnes propriétés statistiques. Certains auteurs, comme Gould & Pollock (1997a), Gould & Pollock (1997b) et Gould et coll. (1997), à partir de leurs études en simulation, affirment même que les estimateurs de maximum de vraisemblance (MV) associée aux lois multinomiales sont moins biaisés, plus précis que ceux fournis par les moindres carrés ordinaires, et sont aussi plus robustes aux erreurs de mesure.

4.1. Maximum de vraisemblance

Dans ce paragraphe, nous supposons d'abord que $\phi_k = 1$, $k = 1, \ldots, K$ et nous considérons les hypothèses (H0-1)-(H0-3). Sous ces conditions, le vecteur aléatoire $\mathbf{Y} = (Y_1, \ldots, Y_{K+1})'$ est multinomial de paramètres $(N, p_1, \ldots, p_{K+1})$, ou de façon équivalente et sous une forme séquentielle, la loi de Y_0 (resp. la loi de Y_k, conditionnelle à $\mathcal{F}_{k-1}$, pour $k \geqslant 1$) est binomiale de paramètres $(N - s_{k-1}, p_k)$, où avec $\mathcal{F}_0$ réduit à l'ensemble contenant l'événement impossible et l'événement certain.

Pour une observation $\mathbf{y} = (y_1 \ldots, y_K)'$ donnée, il est alors commode de mettre la vraisemblance $l(N, \mathbf{p})$ sous la forme générale suivante

$$P_{(N,\mathbf{p})}(\mathbf{Y} = \mathbf{y}) = l_1(N; s) l_2(N, \mathbf{p}; s) l_3, \tag{21}$$

où $\mathbf{s} = (s_1, \ldots, s_K)', \mathbf{p} = (p_1 \ldots, p_K)'$ et

$$l_1(N; s) = \frac{N!}{(N-s)!}; \quad l_2(N, \mathbf{p}; \mathbf{s}) = \prod_{i=1}^{K} p_i^{y_i} (1 - p_i)^{N - s_i}; \quad l_3 = \prod_{i=1}^{K} (y_i!)^{-1}. \tag{22}$$

Le logarithme de la vraisemblance $L(N;\mathbf{p})$ s'écrit

$$L(N;\mathbf{p}) = L_0(N,\mathbf{p};\mathbf{s}) + L_3, \tag{23}$$

où

$$L_0(N;\mathbf{p}) = L_1(N;s) + L_2(N,\mathbf{p};\mathbf{s}), \tag{24}$$

et

$$L_3 = -\sum_{i=1}^{K} \ln(y_i!),$$

$$L_1(N;s) = \sum_{j=0}^{s-1} \ln(N-j),$$

$$L_2(N,\mathbf{p};\mathbf{s}) = \sum_{i=1}^{K} y_i \ln p_i + \sum_{i=1}^{K} (N-s_i)\ln(1-p_i).$$

Remarque 3

Considérons la situation où les efforts de pêche ne sont pas constants. Nous avons vu dans la remarque 1, que les efforts de pêche ϕ_k ne jouent qu'un rôle multiplicatif par rapport aux probabilités de capture p_k. On est ramené à la modélisation des $\phi_k p_k$, $k = 1,\ldots,K$. Dans le cas où la capturabilité est constante, soit $p_k = p$ constant, une modélisation des $\phi_k p_k$, $k = 1,\ldots,K$ revient à choisir un modèle approprié des p_k, $k = 1,\ldots,K$, en analysant les données sur les efforts de pêche. C'est cette approche que nous avons retenue dans l'application que nous présentons dans le dernier paragraphe.

4.1.1. Cas des probabilités de capture constantes

Ce cas a été étudié d'abord par Moran (1951) puis par Zippin (1956, 1958). Sous la condition (H1-1), la fonction L_0 s'écrit

$$L_0(N,p) = \ln l_1(N;s) + s\ln p + \ln(1-p)\left(NK - \sum_{i=1}^{K} s_i\right).$$

Sous l'hypothèse (H1), et en supposant que N prenne des valeurs réelles plutôt qu'entières, les estimateurs MV $(\widehat{N},\widehat{p})$ satisfont aux équations suivantes des extrema du logarithme de la vraisemblance, à savoir

$$\frac{\partial L(N,p)}{\partial N} = \sum_{j=1}^{s-1} \frac{1}{N-j} + K\ln(1-p) = 0, \tag{25}$$

et

$$\frac{\partial L(N,p)}{\partial p} = \frac{1}{p}\sum_{i=1}^{K} y_i - \frac{1}{1-p}\left(NK - \sum_{i=1}^{K} s_i\right) = 0. \tag{26}$$

On en déduit que

$$\widehat{p}(\widehat{N}) = \frac{s}{\widehat{N}K - \sum_{i=1}^{K-1} s_i}, \tag{27}$$

et

$$\sum_{j=1}^{s-1} \frac{1}{\widehat{N} - j} + K\ln\left(\frac{\widehat{N}K - \sum_{i=1}^{K} s_i}{\widehat{N}K - \sum_{i=1}^{K-1} s_i}\right) = 0. \tag{28}$$

Moran (1951) écrit plutôt la relation (27) sous la forme

$$\widehat{p}(\widehat{N}) = \frac{s}{d_0 + s + K(\widehat{N} - s)}$$

en posant

$$d_0 = \sum_{i=1}^{K} (i-1)y_i \tag{29}$$

En approchant $\sum_{i=1}^{s-1}(\widehat{N}-i)^{-1}$ par $\ln(\frac{\widehat{N}-s}{\widehat{N}})$, il propose d'estimer N à partir de

$$\ln\left(\frac{\widehat{N} - s}{\widehat{N}}\right) = K\ln\{1 - \widehat{p}(\widehat{N})\} \tag{30}$$

d'où

$$\widehat{N} = \frac{s}{1 - \{1 - \widehat{p}(\widehat{N})\}^K}.$$

Zippin (1956), utilisant la même approximation de $\sum_{i=1}^{s-1}(\widehat{N}-i)^{-1}$, déduit de (27)–(28), la relation

$$\frac{\widehat{N} - s}{\widehat{N}} = \left(\frac{\widehat{N}K - \sum_{i=1}^{K} s_i}{\widehat{N}K - \sum_{i=1}^{K-1} s_i}\right)^K$$

pour calculer $\widehat{N}$ par des approximation successives. Il a proposé une méthode graphique pour évaluer les estimations $\widehat{p}$ et $\widehat{N}$, Higgins (1985) a donné un algorithme et un programme en BASIC pour les calculer.

Bien que le modèle binomial avec l'une des hypothèses (H1-1) ou (H1-2) ait été couramment utilisé depuis longtemps, les premiers résultats théoriques concernant l'existence et l'unicité des estimateurs fondés sur le maximum de vraisemblance n'ont été établis qu'en 1994 par Bedrick, et seulement pour le cas (H1-1).

D'après Bedrick (1994), la maximation de $L(N,p)$ peut se réduire à un problème à un paramètre car $L(N,p)$ est maximisé pour chaque N fixé en $p(N)$ défini par l'équation (26), c'est-à-dire $p(N) = s/(NK - \sum_{i=1}^{K-1} s_i)$. Autrement dit, Bedrick (1994) a décomposé le problème de maximisation (PB1)

$$(\widehat{N}, \widehat{p}) = \operatorname{argmax}_{N \geqslant s, 0<p<1} L(N,p),$$

selon le schéma

$$\max_{N \geqslant s, 0<p<1} L(N,p) = \max_{N \geqslant s} \max_{0<p<1} L(N,p). \tag{31}$$

A notre avis, cette approche de Bedrick (1994) peut se justifier de la façon suivante : pour tout N fixé, la fonction $L_2 : p \rightarrow L(N,p)$ est convexe sur $]0,1[$ et admet admet un maximum unique en $p(N)$, sa dérivée seconde étant égale à

$$L_2''(p) = -\left\{ \frac{s}{p^2} + \frac{\sum_{k=1}^{K} s_k}{(1-p)^2} \right\}.$$

Ainsi, $L(N,p(N)) = \max_{p \in]0,1[} L(N,p)$ et le problème (PB1-1) se ramène à

$$\max_{N \geqslant s, 0<p<1} L(N,p) = \max_{N \geqslant s} L(N,p(N)), \tag{32}$$

puisque $N \geqslant s$ implique que $0 < p(N) < 1$. En effet,

$$0 < p(N) < 1 \Leftrightarrow 0 < \frac{s}{NK - \sum_{i=1}^{K-1} s_i} < 1 \Leftrightarrow N > K^{-1} \sum_{i=1}^{K} s_i.$$

Comme $s \geqslant s_i$, donc $N \geqslant s \Rightarrow N \geqslant K^{-1} \sum_{i=1}^{K} s_i$.

Nous donnons dans la proposition suivante les résultats essentiels de Bedrick (1994).

PROPOSITION 3. — *(i) Pour $N \geqslant s$, $L(N,p(N))$ est croissante pour $d_0 \geqslant s(K-1)/2 + K/2$ et décroissante pour $d_0 \leqslant c_0$ où $d_0 = \sum_{i=1}^{K}(i-1)y_i$ et $c_0 = s/\{(1+2s)^{1/K} - 1\}$.*

(ii) $L(N,p(N))$ est maximum en une valeur finie de N si et seulement si $d_0 < s(K-1)/2 + K/2$.

(iii) Le profil du logarithme de vraisemblance $L(N,p(N))$ est unimodal.

Pour obtenir ce résultat à partir du problème (PB1), Bedrick écrit la dérivée $g(N)$ de $L(N,p(N))$ sous la forme

$$\begin{aligned} g(N) = L'(N,p(N)) &= \psi(N+1) - \psi(N) - K\ln\left\{1 + \frac{s}{K(N-s)+d_0}\right\} \\ &= \sum_{j=0}^{s-1} \frac{1}{N-j} - K\ln\left\{1 + \frac{s}{K(N-s)+d_0}\right\}, \end{aligned}$$

où $\psi(x)$ est la dérivée de $\ln\Gamma(x)$, appelée fonction digamma; on aurait d'ailleurs pu déduire ce résultat directement de la relation (25). En utilisant l'approximation de $\Gamma(x)$ par la formule de Stirling, Bedrick (1194) approche

$$\ln l_1(N;s) = \ln\frac{\Gamma(N+1)}{\Gamma(N-s+1)} \approx (N+a)\ln(N+a) - (N+a-s)\ln(N+a-s) - s$$

avec $a = 1/2$ (correspondant à la meilleure approximation) et en déduit que les propriétés-clés de la fonction s'appuient sur celles de la famille de fonctions approximantes

$$g_a(N) = \ln\left(1 + \frac{s}{N+a-s}\right) - K\ln\left\{1 + \frac{s}{K(N-s)+d_0}\right\}.$$

Puis, en employant la technique de la transformée de Laplace pour étudier les zéros de $g(N)$, Bedrick (1994) prouve la proposition ci-dessus.

De ce résultat, Bedrick (1994) déduit que les estimateurs de MV $(\widehat{p}, \widehat{N}))$ sont donnés par

$$\widehat{N} = \text{argmax}_{N \geqslant s} L(N,p(N))$$

et

$$\widehat{p} = \widehat{p}(\widehat{N}) = \frac{s}{\widehat{N}K - \sum_{i=1}^{K-1} s_i}.$$

4.1.2. Cas des probabilités de capture non-constantes

Nous nous plaçons ici dans le cadre de l'hypothèse (H1-2), puisqu'elle correspond au cas d'application que nous considérons ici et qu'elle est adoptée dans la plupart des applications réelles et dans la quasi-totalité des travaux dans ce domaine (voir Seber (1973). Pour plus de commodité, écrivons l'hypothèse (H1-2) sous la forme plus générale

$$\mathbf{p} = \boldsymbol{\pi}(\theta),$$

où $\theta = (\theta_1, \theta_2)'$ est un vecteur appartenant à un sous-ensemble Θ de $\mathbb{R}^2$, par exemple $\theta_1 = p_1$ et $\theta_2 = \alpha$.

Nous donnons dans ce paragraphe quelques résultats sur l'existence et l'unicité des estimations $(\widehat{N}, \widehat{\theta})$ de maximum de vraisemblance de (N, θ) définis par le problème d'optimisation (PB-2)

$$(\widehat{N}, \widehat{\theta}) = \arg \max_{N \geqslant s, \theta \in \Theta} L_0(N, \boldsymbol{\pi}(\theta)).$$

La démarche de Bedrick (1994) décrite dans le paragraphe précédent ne peut pas permettre de résoudre ce problème. Aussi avons-nous adopté dans Truong-van (1998) et Truong-van et coll. (1998) une approche totalement différente que nous décrivons maintenant.
Nous exprimons d'abord $L_0(N, \boldsymbol{\pi}(\theta))$ sous la forme

$$L_0(N, \boldsymbol{\pi}(\theta)) = L_1(N) + Nh_1(\theta) + h_2(\theta), \tag{33}$$

où

$$h_1(\theta) = \sum_{i=1}^{K} \ln(1 - p_i) = \sum_{i=1}^{K} \ln\{1 - \pi_i(\theta)\}, \tag{34}$$

$$h_2(\theta) = \sum_{i=1}^{K} \{y_i \ln p_i - S_i \ln(1 - p_i)\} \tag{35}$$

$$= \sum_{i=1}^{K} \{y_i \ln \pi_i(\theta) - S_i \ln(1 - \pi_i(\theta))\}, \tag{36}$$

où on a simplement noté $L_1(N) = L_1(N; s)$.
Avec le changement de variables

$$\alpha_i = h_i(\theta), \quad i = 1, 2,$$

la relation (33) s'écrit plus simplement

$$L_0(N, \boldsymbol{\pi}(\theta)) = L_1(N) + \alpha_1 N + \alpha_2. \tag{37}$$

Sous la forme (37), pour tout $\boldsymbol{\theta}$ fixé tel que $\alpha_2 = h_2(\boldsymbol{\theta})$ est fini (notons que α_2 peut prendre les valeurs $\pm\infty$), la fonction $N \to L_0(N, \boldsymbol{\pi}(\boldsymbol{\theta}))$ est la somme d'une fonction $L_1(N)$ strictement concave et d'une fonction affine. Plus précisément, comme

$$L_1''(N) = -\sum_{j=0.}^{s-1} (N - j)^{-2},$$

$L_1(N)$ est une fonction strictement concave sur $[s, \infty[$, dont la dérivée est strictement positive et décroissante vers 0. Ainsi, pour tout $\boldsymbol{\theta}$ fixé tel que α_2 est fini, la fonction $N \to L_0(N, \boldsymbol{\pi}(\boldsymbol{\theta}))$ est aussi strictement concave. De plus, on a

$$\frac{\partial L_0(N, \boldsymbol{\pi}(\boldsymbol{\theta}))}{\partial N} = L_1'(N) + \alpha_1. \tag{38}$$

Comme L_1' est strictement positif et strictement décroissant pour $N \geqslant s$, et que $\alpha_1 = h_1(\boldsymbol{\theta}) \leqslant 0$, il s'ensuit que pour tout θ tel que $-L_1'(s) \leqslant \alpha_1 < 0$, il existe un unique $N_*(\alpha_1(\theta)) \in [s, \infty[$, solution de l'équation

$$L_1'(N) + \alpha_1 = 0. \tag{39}$$

Autrement dit, pour tout θ tel que $\alpha_2(\theta)$ est fini et que $-L_1'(s) \leqslant \alpha_1(\theta) < 0$, $N_*(\alpha_1(\theta)) \in [s, \infty[$ est solution du problème $\max_{N \geqslant s} L_0(N, \boldsymbol{\pi}(\boldsymbol{\theta}))$. Ceci nous conduit à résoudre le problème (PB-2) sous la forme

$$\max_{N \geqslant s, \theta \in \Theta} L_0(N, \boldsymbol{\pi}(\theta)) = \max_{\theta \in \Theta} \max_{N \geqslant s} L_0(N, \boldsymbol{\pi}(\theta)).$$

La proposition suivante donne les conditions d'existence et d'unicité des estimations de maximum de vraisemblance pour (N, θ).

PROPOSITION 4. — *(i) Pour tout θ tel que $\alpha_2(\theta)$ soit fini et que $-L_1'(s) \leqslant \alpha_1(\theta) < 0$, avec $L_1'(s) = 1 + 2^-1 + \ldots + s^{-1}$, il existe un unique $N_*(\alpha_1(\theta)) \in [s, \infty[$ tel que*

$$\frac{\partial L_0(N_*(\alpha_1(\theta)), \boldsymbol{\pi}(\boldsymbol{\theta}))}{\partial N} = 0.$$

(ii) Si $\frac{s}{K^2} \sum_{j=0}^{s-1} \dfrac{1}{(N_*(\alpha_1) - j)^2} > (1 - \exp(\alpha_1/K))^2 / \exp(\alpha_1/K)$ alors la fonction $H(\alpha_1) =: L_1(N_*(\alpha_1)) + \alpha_1\{N_*(\alpha_1) + K^{-1} \sum_{i=1}^{K} s_i\}$ admet un maximum unique.

De cette proposition, on déduit la condition suffisante suivante : soit $\Delta = [-L_1(s), a_M] \times [b_m, b_M]$ où a_M, b_m, b_M sont des réels tels que $-L_1(s) < a_M < 0, -\infty < b_m < b_M < \infty$. Si $h(\theta) = (h_1(\theta), h_2(\theta))$ est un difféomorphisme (c'est-à-dire une application bijective de classe C^1 telle que sa fonction réciproque est aussi de classe C^1, ce qui est réalisé pour toutes les fonctions π usuelles) entre Θ et Δ, alors les estimations de maximum de vraisemblance $(\widehat{N}, \widehat{\theta})$ existent et sont uniques.

Notons que notre approche permet aussi de résoudre le problème de maximum de vraisemblance dans le cas de probabilités de capture constantes. Nous illustrons la proposition ci-dessus en considérant ce cas. En effet, sous l'hypothèse (H1-1), on a

$$L_0(N, \boldsymbol{\pi}(\theta)) = L_1(N) + Nh_1(q) + h_2(q) \tag{40}$$

où

$$h_1(q) = K\ln(q), \tag{41}$$

$$h_2(q) = \ln(1-q)\sum_{i=1}^{K} y_i - \ln(q)\sum_{i=1}^{K} s_i. \tag{42}$$

Ici, le paramètre θ est égal à $q = 1 - p$. L'équation (38) devient

$$L_1'(N) + K\ln(q) = 0.$$

D'après la proposition, pour tout q tel que

$$-L_1'(s) \leqslant K\ln(q) < 0 \Leftrightarrow \exp(-L_1'(s)/K) \leqslant q < 1,$$

cette équation admet une solution unique $N_*(h_1(q))$.
En remplaçant N par $N_*(q)$ dans $L_0(N, \boldsymbol{\pi}(\theta))$, on est alors conduit au problème

$$\max_{q\in\Theta}\{L_1(N_*(h_1(q))) + N_*(h_1(q))h_1(q) + h_2(q)\}.$$

On montre que ce problème admet une solution unique si $\exp\{-L_1'(s)/K\} \leqslant q < 1$ et si

$$\left(\frac{K}{q}\right)^2 \frac{N_*(h_1(q)) - \sum_{i=1}^{K} s_i}{K} + \frac{s}{(1-q)^2} > N_*'(h_1(q)).$$

On retrouve les résultats de Bedrick (1994) en faisant des approximations pour N grand.

4.1.3. Remarques sur les propriétés des estimateurs

1. Cas de probabilité de capture constante

   Zippin (1956) établit la variance asymptotique de $\widehat{N}$ sous la forme suivante :

   $$var_a^{(Z)}(\widehat{N}) = \frac{Nq^K(1-q^K)}{(1-q^K)^2 - K^2p^2q^{K-1}}, \tag{43}$$

   où $q = 1 - p$.

   Dans le cas particulier où on ne dispose que de deux captures, $var_a^{(Z)}(\widehat{N})$ étant égal à $\dfrac{Nq^2(1+q)}{p^3}$, Seber & Le Cren (1967) proposent alors, pour la variance asymptotique de $\widehat{N}$, l'expression ci-dessous qui, d'après ces auteurs, est plus précise que $var_a^{(Z)}(\widehat{N})$:

   $$var_a^{(SL)}(\widehat{N}) = \frac{Nq^2(1+q)}{p^3} + \frac{2q(1-p^2-q^3)}{p^5} - b^2, \tag{44}$$

   où $b = q(1+q)/p^3$.

   Junge & Libosvarsky (1965) puis Seber & Whale (1970) donnent une formule analogue à (44) pour trois captures successives.

   Sous la condition que N et S_K tendent vers l'infini ($S_K \to \infty$ p.s.), Seber & Le Cren (1967) énoncent que $(\widehat{N} - N)/\sqrt{var_a(\widehat{N})}$ converge en loi vers la normale $N(0,1)$ et en déduisent des intervalles de confiance pour N.

   Hirst (1994) construit des intervalles de confiance pour N, en employant le ratio de log-vraisemblances de Kalbfleisch & Sprott (1970) mais sous la forme suivante proposée par Harding et coll. (1984) :

$$rL(N,p;\widehat{N},\widehat{p}) = \frac{\ln(N,p)}{\ln(\widehat{N},\widehat{p})},$$

quantité dont le double a la distribution asymptotique d'une loi du khi-deux avec le nombre de degrés de liberté égal au nombre de paramètres estimés. A partir d'une étude en simulation pour différentes valeurs de N entre 20 à 300, pour 10 000 répliques, et pour deux et trois captures, cet auteur calcule les fréquences d'appartenance de N aux différents intervalles de confiance, obtenus avec sa technique et avec celles de Seber & Le Cren (1967) et de Seber & Whale (1970). Il en conclut que le ratio de log-vraisemblances donne des intervalles de confiance bien meilleurs en termes de probabilité de couverture que les autres méthodes.

2. Cas des probabilités de capture définies par l'hypothèse (H1-2)

   Pourvu que les conditions d'application de la théorie asymptotique des estimateurs MV soient remplies, on obtient les expressions des variances asymptotiques des estimateurs MV de (N,θ) à partir de la matrice hessienne de $L_0(N,\boldsymbol{\pi}(\theta))$, laquelle se déduit facilement de (33) puisque l'on a :

$$\frac{\partial^2 L_0(N,\boldsymbol{\pi}(\theta))}{\partial N^2} = -\sum_{r=1}^{s-1}\frac{1}{(N-r)^2},$$

$$\frac{\partial^2 L_0(N,\boldsymbol{\pi}(\theta))}{\partial\theta_i\partial\theta_j} = N\frac{\partial^2 h_1(\theta))}{\partial\theta_i\partial\theta_j}+\frac{\partial^2 h_2(\theta))}{\partial\theta_i\partial\theta_j}, \quad 1\leqslant i,j\leqslant 2$$

$$\frac{\partial^2 L_0(N,\boldsymbol{\pi}(\theta))}{\partial N\partial\theta_j} = \frac{\partial h_1(\theta))}{\partial\theta_j}, \quad j=1,2.$$

3. On a tendance à admettre que les estimateurs de maximum de vraisemblance $(\widehat{N},\widehat{\theta})$ ont les mêmes propriétés que celles données par la théorie classique. Cependant les résultats asymptotiques généraux de cette théorie ne sont établis que pour des paramètres autres que la taille de l'échantillon. Le problème ici est que le paramètre N, taille de la population est en même temps la taille de l'échantillon. Considérons par exemple le cas où la vraie taille N de la population est égale à 100. Quelle est alors la validité des résultats asymptotiques sur son estimateur $\widehat{N}$ quand on fait tendre N vers l'infini alors que N est fini ? L'approche de Hirst (1994) fournit des éléments de réponse à cette question.

5. Maximum de vraisemblance marginale

Nous supposons ici que $\phi_k = 1$, $k=1,\ldots,K$ et nous considérons l'hypothèse (H0-4), c'est-à-dire la situation où la capturabilité est une v.a. U (dite cachée) à valeurs dans $[0,1]$, dont la loi de probabilité est définie par la densité de probabilité g_U relative à une mesure de référence ν (mesure de Lebesgue ou de dénombrement). Soit $U_1, \ldots, U_K$ un échantillon de taille K issu de U. Nous supposons aussi que la loi de $\mathbf{Y} = (Y_1, \ldots, Y_{K+1})'$ conditionnelle à $U_1 = p_1, \ldots, U_K = p_K$ est multinomiale de paramètres $(N, p_1, \ldots, p_{K+1})$. La densité de la loi conjointe de $(\mathbf{Y}, U_1, \ldots, U_K)$ relative à une mesure de référence adéquate est donc :

$$g_{(\mathbf{Y},U_1,\ldots,U_K)}(\mathbf{y}, u_1, \ldots, u_K) = \left\{ \prod_{k=1}^{K} g_U(u_k) \right\} g_{\mathbf{Y}}(\mathbf{y}|u_1, \ldots, u_K),$$

où $g_{\mathbf{Y}}(\mathbf{y}|u_1, \ldots, u_K)$ est la densité de probabilité relative à la mesure de dénombrement (c'est-à-dire la distribution de probabilité) de $\mathbf{Y}$ conditionnelle à $U_1 = u_1, \ldots, U_K = u_K$.

Comme les seules observations que l'on peut disposer sont les K captures $y_1, \ldots, y_K$, il est naturel de considérer la vraisemblance marginale en $\mathbf{Y}$:

$$g_{\mathbf{Y}}(\mathbf{y}) = \int_0^1 \cdots \int_0^1 g_{(\mathbf{Y},U_1,\ldots,U_K)}(\mathbf{y}, u_1, \ldots, u_K) d\nu(u_1) \cdots d\nu(u_K).$$

Cette vraisemblance marginale en $\mathbf{Y}$ n'est donc autre que l'espérance mathématique de $g_{\mathbf{Y}}(\mathbf{y}|u_1, \ldots, u_K)$ relative à la loi de $(U_1, \ldots, U_K)$.

Carle & Strub (1978) ont considéré le cas particulier où la v.a. U est de loi bêta à deux paramètres α, β. D'après ces auteurs, pour une observation donnée $\mathbf{y}$ du vecteur aléatoire $\mathbf{Y}$, la vraisemblance marginale

$$L_w(N, \alpha, \beta) = g_{\mathbf{Y}}(\mathbf{y}) = P_{(N,\alpha,\beta)}(\mathbf{Y} = \mathbf{y})$$

s'écrit sous la forme

$$L_w(N, \alpha, \beta) \propto \frac{N!}{\prod y_i!(N-s)!} \frac{(\alpha+\beta-1)!}{(\alpha-1)!(\beta-1)!} \frac{(s+\alpha-1)!(KN-d_0-s+\beta-1)!}{(KN-d_0+\alpha+\beta-1)!}.$$

Ils en déduisent l'estimation de (N, α, β) par

$$\left(\widehat{N}^w, \widehat{\alpha}, \widehat{\beta}\right) = \mathrm{argmax}_{N \geqslant s, \alpha, \beta} L_w(N, \alpha, \beta).$$

D'après Carle & Strub (1978), $\widehat{N}^w$ est alors le plus petit entier n tel que

$$n - \frac{n+1}{s+1} \prod_{k=1}^{K} \frac{Kn - d_0 - s + \beta + (K-k)}{Kn - d_0 + \alpha + \beta + (K-k)} \leqslant 1,$$

et $\widehat{N}^w$ donne une estimation bien meilleure que celle fondée sur la vraisemblance. Sur les résultats d'une simulation réalisée pour différentes valeurs de N inférieures à 500, ces auteurs obtiennent avec $\widehat{N}^w$ une réduction de biais de 60% et une réduction de l'écart-type empirique de 70% par rapport à l'estimateur de maximum de vraisemblance $\widehat{N}$. Ils ajoutent que les résultats pour $\widehat{N}^w$ et pour $\widehat{N}$ sont similaires quand N et K croissent.

Le modèle appelé M_{bh} (pour « *h*eteroscedastic *b*ehavior ») d'Otis et coll. (1978) est une extension du modèle de Carle & Strub (1978). Il correspond en fait au cas où la loi de U dépend de paramètres θ.

Otis et coll. (1978) (p. 45 et Annexe H), en se référant à une thèse inédite de K. H. Pollock, affirment que

$$P_{(N,\theta)}(\mathbf{Y} = \mathbf{y}) = l_1(N; s)\widetilde{l}_2(N, \boldsymbol{\theta}; s)l_3, \tag{45}$$

où $l_1(N; s)$ et l_3 sont définis dans (22) et où

$$\widetilde{l}_2(N, \boldsymbol{\theta}; s) = \prod_{i=1}^{K} \{E_\theta(U(1-U)^{i-1})\}^{y_i} \times \left[1 - \sum_{i=1}^{K} \{E(U(1-U)^{i-1})\}^K\right]^{N-s_K}$$

où E_θ est l'espérance relative à la loi de U.

Notons que la quantité (45) ne correspond à aucune notion de vraisemblance classique. De toute évidence, elle n'est pas une vraisemblance conditionnelle ou non conditionnelle. Mais elle n'est pas non plus une vraisemblance marginale puisque

$$\widetilde{l}_2(N, \boldsymbol{\theta}; s) \neq E_\theta \left[\prod_{i=1}^{K} \{(U_i(1-U_i)^{i-1})\}^{y_i} \left\{1 - \sum_{i=1}^{K} (U_i(1-U_{i-1})^{i-1})^K\right\}^{N-s_K}\right].$$

Norris & Pollock (1996) étendent les résultats de Carle & Strub (1978) au cas où la densité de probabilité de U n'est pas paramétrée. Ils utilisent l'algorithme EM (voir Chalmond 2000) pour une estimation non-paramétrique de g_U (pour différentes valeurs fixées de N) et appliquent leur méthode aux modèles M_b et M_{bh} d'Otis et coll. (1978). Pledger (2000) généralise le travail de Norris & Pollock (1996) aux huit modèles d'Otis et coll. (1978) en utilisant les mêmes techniques.

Notons que l'approche utilisant la vraisemblance marginale (voir Carle & Strub 1978, Norris & Pollock 1996, Pledger 2000) est bien connue dans le domaine de l'estimation de processus ou de champs de Markov cachés (voir Chalmond 2000).

6. Application au cas des saumons de l'Adour-Nive

Nous considérons ici un problème d'estimation de la taille N de la population de saumons de l'Atlantique remontant le fleuve Adour situé dans le sud-ouest de la France, à partir des données des trois saisons de pêches de 1988 à 1990 (Prouzet et coll. 1996). Pour chaque campagne de pêche, on dispose du nombre K de pêcheries (professionnelles) ($K = 5$ pour l'année 1988 et $K = 6$ pour les deux années 1989 et 1990), des captures par pêcherie (y_i, $i = 1, \ldots, K$), du nombre de pêcheurs par pêcherie, du nombre de coups de filet, des captures des pêches (amateurs) à la ligne $y = y_n + y_g$, où y_n sont celles sur la Nive et y_g dans les Gaves après la Kième pêcherie, du nombre de frayères de grands salmonidés $\chi = \chi_n + \chi_g$, où χ_n concerne la Nive et χ_g les Gaves de l'Adour, des proportions ϖ_1 des castillons (S1) et ϖ_2 des grands saumons (S2), et des proportions ρ_1 et ρ_2 des femelles des deux catégories de saumons (S1) et (S2). Les pêcheries professionnelles successives, $\imath = 1, \ldots, K$, sont situées uniquement sur l'Adour. La séparation entre l'Adour et son affluent, la Nive, se trouve entre la deuxième et la troisième pêcherie (voir graphiques).

Les données de retraits séquentiels que nous considérons sont les captures des pêcheries professionnelles, toutes situées sur l'Adour. Il est clair que pour le problème considéré, nous ne sommes pas dans les conditions d'application directe des méthodes de retraits séquentiels. Pour satisfaire au mieux la condition de population fermée, nous prenons en compte la séparation entre l'Adour et la Nive de la façon suivante :

Nous supposons que le taux r de la séparation de la population de saumons entre l'Adour et la Nive est connu (sa valeur estimée par les experts du Conseil Supérieur de la Pêche à partir de leurs expériences de terrain et de plusieurs années de données de pêches à la ligne sur l'Adour et la Nive, étant de 3/4 de la population partant vers l'Adour) et ne considérons que les données des K pêcheries.

Nous admettons les hypothèses (H0-1)–(H0-3), autrement dit que la loi de Y_0 (resp. la loi de Y_k, conditionnelle à $\mathcal{F}_{k-1}$, pour $k \geqslant 1$) est binomiale $B(c_k N - s_{k-1}, p_k)$ de paramètres $N - s_{k-1} - c_k, p_k$ où, pour tenir compte de la séparation entre l'Adour et la Nive, $c_2 = 1$ pour $k = 1, 2$ et $c_k = 3/4$ pour $k = 3, \ldots, K$.

Comme le nombre des observations est de six au maximum, il est statistiquement judicieux de se restreindre à un ou deux paramètres. Pour prendre en compte les efforts de pêche, nous choisissons l'approche décrite dans la remarque 3. L'étude des données sur les efforts de pêche nous a conduits à envisager chacune des deux d'hypothèses (H1-1) et (H1-2-4).

Les captures $y_\imath$ des différentes pêcheries $\imath = 1, \ldots, K$ sont présentées dans le Tableau 1 ci-dessous. Ce sont des données agrégées sur 100 à 120 jours pour les grands saumons et de 30 à 60 jours pour les castillons.

TABLEAU 1. – La quatrième pêcherie ne s'est installée qu'à partir de 1989.

Pêcherie	captures	en 1988	captures	en 1989	captures	en 1990
	S1	S2	S1	S2	S1	S2
1	445	810	112	93	807	119
2	87	204	40	52	32	28
3	123	220	56	44	139	47
4			5	1	42	10
5	39	29	42	4	75	29
6	17	19	26	16	39	25
Capture totale	711	1282	281	210	1134	258

Les données sur les efforts de pêche, sur les frayères et sur les différentes proportions de la population sont résumées dans les Tableaux 2, 3 et 4 suivants. Un effort de pêche corrigé (efc) est le nombre de filets par jour susceptibles d'être mis à l'eau au même instant. Dans le Tableau 2, $\widetilde{S}1$ (resp. $\widetilde{S}2$) correspond à la période de remontée de castillons (resp. de saumons) et T est la saison totale de pêche.

TABLEAU 2. – Données sur les moyennes des efforts corrigés (efc). Les quantités entre parenthèses sont les écart-types estimés.

Pêcherie	efc	1988	efc	1989	efc	1990
	$\widetilde{S}1$	$\widetilde{S}2$	$\widetilde{S}1$	$\widetilde{S}2$	$\widetilde{S}1$	$\widetilde{S}2$
1	3.1	3.2	1.9	2.4	2.9	2.5
	(1.2)	(1.2)	(1.3)	(1.2)	(1.4)	(1.3)
2	3.0	3.6	1.4	2.9	0.7	1.1
	(1.5)	(1.9)	(1.0)	(1.7)	(0.8)	(0.8)
3	2.7	3.6	2.6	3.6	4.4	5.9
	(1.9)	(2.3)	(1.2)	(1.7)	(2.3)	(2.8)
4			0.1	0.2	1.4	1.5
			(0.2)	(0.1)	(1.0)	(1.1)
5	1.7	1.5	1.1	1.2	1.7	1.7
	(1.7)	(1.5)	(0.9)	(0.8)	(0.6)	(0.7)
6	2.6	2.9	2.7	2.9	3 3	3.1
	(1.3)	(1.3)	(1.5)	(1.6)	(1.4)	(1.5)

TABLEAU 3. – Données sur les pêches à la ligne sur l'Adour (CLA) et sur la Nive (CLN), sur le ratio de femelles pour les castillons (RfC) et pour les saumons (RfS), sur le ratio de castillons (RC) et de saumons (RS) dans la population.

Année	1988	1989	1990
CLA	210	103	166
CLN	100	32	23
RfC	0.47	0.42	0.48
RfS	0.85	0.83	0.65
RC	0.25	0.52	0.65
RS	0.75	0.48	0.35

TABLEAU 4. – Données sur les frayères pour l'Adour (FA) et pour la Nive (FN).

Période	88	89	90
FA	1100	648	246
FN	500	197	206

TABLEAU 5. – Résultats dans le cas (H1-1) de probabilités de capture constantes.

	1988 S1	1988 S2	1989 S1	1989 S2	1990 S1	1990 S2
$\widehat{p}$	0.50	0.54	0.22	0.38	0.48	0.25
$\widehat{N}$	784	1403	397	239	1237	353
Taux d'exploitation	90.7%	91.4%	70.8%	87.9%	91.7%	73.0%

TABLEAU 6. – Résultats dans le cas (H1–2-4) de probabilités de capture à décroissance polynomiale.

	1988 S1	1988 S2	1989 S1	1989 S2	1990 S1	1990 S2
$\widehat{p}$	0.39	0.43	0.18	0.38	0.48	0.20
$\widehat{\alpha}$	0.40	0.47	0.06	0.07	0.47	0.08
$\widehat{N}$	1050	1836	516	250	1447	474
Taux d'exploitation	67.7%	69.8%	54.4%	84.0%	78.4%	54.5%

Nous présentons dans les deux tableaux 5 et 6 les estimations de maximum de vraisemblance obtenues pour un coefficient de séparation Adour-Nive $r = 3/4$ et sous les deux conditions de probabilités de capture (H1-1) et (H1-2-4). Les taux d'exploitation que sont les ratios $s_K/\widehat{N}$ sont exprimés en pourcentage.

Pour les estimations dans les deux tableaux 5 et 6, les écart-types empiriques que nous avons déduits des formules asymptotiques du maximum de vraisemblance conduisent à des coefficients de variation estimés inférieurs à 10%.

Le problème majeur soulevé par les résultats obtenus dans les tableaux 5 et 6 ci-dessus est que les taux d'exploitation obtenus, surtout ceux du tableau 5, sont trop élevés par rapport aux évaluations des experts de la pêche fondées sur le comptage des frayères et des captures de pêche (amateur) à la ligne qui sont en amont dans les Gaves. Ceci est dû à l'inadéquation du modèle multinomial. En effet, les résultats numériques obtenus dans la littérature (voir par exemple Gerdeaux 1987) et les études en simulation que nous avons menées par ailleurs, montrent que pour les modèles multinomiaux, les estimations de maximum de vraisemblance du paramètre N sont toujours voisines des valeurs du total s des captures. Ainsi, pour l'application considérée ici, en supposant que les v.a. Y_k sont binomiales de paramètres $N - s_{k-1}, p_k$, on sous-entend que chaque pêcherie k effectue $N - s_{k-1}$ tirages (avec remise), ce qui est irréaliste.

Nous utilisons alors une approche proposée par Prouzet et coll. (1996) pour adapter le modèle multinomial au problème considéré. Cette approche consiste à prendre en compte la variation de la taille N de la population dans des intervalles $[N_{\min}, N_{\max}]$ dont les extrêmités sont déterminées comme suit : à partir de données des Tableaux 2, 3 et 4, pour chacune des deux catégories «castillon» et «saumon», on évalue d'abord les extrêmités de l'intervalle $[N_{\min}^{(Tf)}, N_{\max}^{(Tf)}]$ contenant la population totale $N^{(Tf)}$ (par catégorie) qui a réussi à frayer, où $N_{\max}^{(Tf)}$ est le quotient du nombre de frayères par le nombre de femelles (de la catégorie) et $N_{\min}^{(Tf)} = N\max^{(Tf)}/1.4$, car on trouve en moyenne 1.4 femelles par frayère (le minimum étant 1). Puis toujours pour chaque catégorie, on ajoute à la population $N^{(Tf)}$ qui a réussi à frayer, le nombre de captures par les pêcheries et par les pêches à la ligne, donc

$$N_{\min} = N_{\min}^{(Tf)} + \rho_C T_{CL} + s_K$$

et

$$N_{\max} = N_{\max}^{(Tf)} + \rho_C T_{CL} + s_K,$$

où ρ_C est le proportion de la catégorie dans la population et T_{CL} le nombre total (pour les 2 catégories) de captures par les pêches à la ligne. Par exemple, pour l'année 1988 et pour la catégorie des castillons, la seconde formule donne $N_{\max} = \frac{(1100 + 500)}{0.47} 0.25 + (210 + 100) \times 0.25 + 711.$

Les résultats que nous avons obtenus avec cette approche, en considérant le modèle multinomial avec l'hypothèse (H1-2-4) et avec une prise en compte des contraintes d'inégalités sur la taille N de la population, sont résumés dans le Tableau 7 ci-dessous.

TABLEAU 7. – Résultats obtenus avec le modèle multinomial sous contraintes d'inégalités sur N et avec l'hypothèse (H1-2-4).

	1988		1989		1990	
	S1	S2	S1	S2	S1	S2
$\widehat{N}$	1396	2539	1105	627	1693	407
Taux d'exploitation	51%	50%	25%	33%	66%	63%

Les résultats obtenus dans le Tableau 7 sont similaires aux évaluations des experts. Cependant, les maxima obtenus avec cette approche sont très dépendants des bornes choisies des contraintes et sont proches de la borne inférieure des contraintes. Ces deux lacunes de la procédure utilisée, en plus du fait qu'elles ne permettent pas d'obtenir une estimation des variances des estimateurs par les formules asymptotiques associées à la vraisemblance, les maxima étant au voisinage des bornes des contraintes, nous amènent à réexaminer les conditions d'application de la méthode. Parmi les sept violations des hypothèses d'application de méthodes de captures séquentielles, énumérées par Seber (1973, p. 305-308), les deux qui nous paraissent majeures dans notre méthode sont :

(i) l'impossibilité pour l'application considérée de respecter de façon stricte l'hypothèse de population fermée,

(ii) une modélisation encore inadéquate des probabilités p_k pour prendre en compte la variabilité dans la capturabilité et dans les efforts de pêche.

Aussi, nous proposons dans nos travaux en cours les axes de développement suivants fondés essentiellement sur une approche de variables cachées (c'est-à-dire non observables directement) de loi paramétrique (l'estimation non paramétrique d'une densité de probabilité requérant beaucoup plus de données que celles dont nous pouvons disposer pour l'application considérée) :

(i) Une modification du modèle multinomial pour prendre en compte des contraintes sur la taille de la population à partir des données en amont de l'Adour et de la Nive :

Soit R est une v.a. de proportion dont la loi a priori g_β pouvant dépendre d'un vecteur de paramètres β, est judicieusement choisie à partir de données de frayères et des pêches à la lignes en amont dans les Gaves. On peut choisir une loi bêta pour g_β, comme dans Carle & Strub (1978). Alors, sous l'hypothèse (H0-3) et conditionnellement à $\mathcal{F}_{k-1}$ et à $R = r$, les v.a.

$$Y_k = \sum_{i=S_{k-1}+1}^{\nu_k} X_i, k = 1, \ldots, K, \tag{46}$$

sont binomiales de paramètres ν_k, p_k, où $\nu_k = [RN - S_{k-1}] + S_{k-1}$ est une v.a. correspondant plus judicieusement au nombre de tirages effectués par la $k$ième pêcherie et où $[a]$ désigne la partie entière d'un nombre a.

Comme Carle & Strub (1978), Norris & Pollock (1996) et Pledger (2000), la vraisemblance que nous choisissons est associée à la probabilité marginale de $\mathbf{Y}$.

Nous étendons également cette approche à la prise en compte de la séparation Adour-Nive, en supposant que le taux r est une v.a. de loi bêta d'espérance 3/4.

(ii) Une meilleure modélisation des probabilités de capture :

En effet, on peut constater à partir d'un examen des Tableaux 1 et 2, que les CPUE se comportent en moyenne comme une fonction décroissante entre la première et la deuxième pêcherie, puis croissante ou discontinue entre la deuxième et la troisième, et décroîssante rapidement ensuite. Nous choisirons pour la capturabilité une v.a. dont la loi de probabilité représentera le mieux le comportement ci-dessus des CPUE.

BIBLIOGRAPHIE

BEVERTON R. J. H. & S. J. HOLT (1956), A review of methods for estimating mortality rates in fish populations with special references to sources of bias in catch sampling. *Rapport P.V. C.I.E.M.* 140, pp. 67–83.

BEVERTON R. J. H. & S. J. HOLT (1957), On the dynamics of exploited fish populations. *U.K. Min. Agr. Fish., Fish. Invest.* 19 (Ser. 2), London : Her Majesty's Stationery Office.

BEDRICK E. J. (1994), Maximum-likelihood estimation for the removal method. *The Canadian Journal of Statistics*, **22**, pp. 285–293.

BRAATEN D. O. (1969), Robustness of the DeLury population estimator. *J. Fish. Res. Bd. Can.*, **26**, pp. 339–355.

CARLE F. L. & STRUB M. R. (1978), A new method for estimating population size from removal data. *Biometrics*, **34**, pp. 621–630.

CHALMOND B. (2000), *Eléments de modélisation en analyse d'images*. Springer Verlag SMAI, Paris.

CHAPMAN D. G. (1961), Statistical problems in dynamics of exploited fish populations. *Proceedings of the Fourth Berkely Symposium, 1960 Univ. California Publ. Statist.* 4, pp. 153–168.

CORMACK R. M. (1992), Interval estimation of mark-recapture studies of closed populations. *Biometrics*, **48**, pp. 567–576.

CORMACK R. M. (1993), Variances of mark-recapture estimates. *Biometrics*, **49**, pp. 1188–1193.

COWX I. G. (1983), A review of the methods for estimating fish population size from survey removal data. *Fish. Management*, **14**, pp. 67–82.

CRITTENDEN R. N. (1983), An evaluation of the Leslie–DeLury method and of a weighted method for estimating the size of a closed population. *Fish. Res.*, **2**, pp. 149–156.

CRITTENDEN R. N. & G. L. THOMAS (1989), A conditional generalized least squares methods for estimating the size of a closed population. *Can. J. Fish. Aquat. Sci.*, **46**, pp. 818–823.

DAVIES R. B. & B. HUTTON (1975), The effect of errors in the independent variables in linear regression. *Biometrika*, **62**, pp. 383–391.

DELURY D. B. (1947), On the estimation of biological population. *Biometrics*, **3**, pp. 145–147.

DELURY D. B. (1951), On the planning of experiments for estimation of fish populations. *J. Fish. Res. Bd. Can.*, **8**, pp. 281–307.

EVANS M. A. & D. G. BONETT (1994), Bias reduction for multiple-recapture estimators for closed population size. *Biometrics*, **50**, pp. 388–395.

EVANS M. A., D. G. BONETT & L. L. MC DONALD (1994), A general theory for modeling capture-recapture data from a closed population. *Biometrics*, **50**, pp. 396–405.

GATZ A. J. & J. M. LOAR (1988), Petersen and removal population size estimate; combining methods to adjust and interpret results when assumptions are violated. *Environ. Biol. Fish.*, **21**, pp. 293–307.

GERDEAUX D. (1987), Revue des méthodes d'estimation de l'effectif d'une population par pêches successives avec retrait. Programme d'estimation d'effectif par la maéthodes de Carle et Strub. *Note Technique Bull. Fr. Pêche Piscic.*, **304**, pp. 13–21.

GOULD W. R. & K. H. POLLOCK (1997a), Catch-effort estimation of population parameters under the robust design. *Biometrics*, **53**, pp. 207–216.

GOULD W. R. & K. H. POLLOCK (1997b), Catch-effort maximum likelihood estimation of important population parameters. *Can. J. Fish. Aquat. Sci.*, **54**, pp. 890–897.

GOULD W. R., L. A. STEFANSKI & K. H. POLLOCK (1997), Effects of measurement error on catch-effort estimation. *Can. J. Fish. Aquat. Sci.*, **54**, pp. 898–906.

HAYNE D. W. (1949), Two methods for estimating populations from trapping records. *J. Mammal.*, **30**, pp. 399–411.

HARDING C. M., A. W. HEATHWOOD, R. G. HUNT & K. L. Q. READ (1984), The estimation of animal population size by the removal method. *Applied Statistics*, **33**, pp. 196–202.

HELLAND A. (1914), Rovvdyrene i Norge, Tidskrift for Skogbruk 1913-1914. In *The Optimum Catch*, publié sous la direction de Hjort et coll. Hvalradets Skr., **7**, pp. 92–127.

HIGGINS P. J. (1985), An interactive computer program for population estimating the Zippin method. *Aquat. Fish Management*, **16**, pp. 287–295.

HIRST D. (1994), An improved removal method for estimating animal abundance. *Biometrics*, **50**, pp. 501–505.

HODGES S. D. & P. G. MOORE (1972), Data incertainties and least squares regression. *Applied Statistics*, **21**, pp. 185–195.

JUNGE C. O. & L. LIBOSVARSKY (1965), Effects of size selectivity on population estimates based on successsives removal with electrical fishing gear. *Zool. Listy*, **14**, pp. 171–178.

KALBFLEISCH J. D. & D. A. SPROTT (1970), Application of likelihood methods to models involving large number of parameters. *J. Royal Statist. Soc, Series B*, **32**, pp. 175–208.

HUGGINS R. M. & P. S. F. YIP (1997), Statistical analysis of removal experiments with the use of auxilary variables. *Statistica Sinica*, **7**, pp. 705–712.

LESLIE P. H. & D. H. S. DAVIS (1939), An attempt to detrmine the absolute number of rats in a given aera. *J. Anim. Ecol.*, **8**, pp. 93–113.

MAHON R. (1980), Accuracy of catch-effort method for estimating fish density and biomass in streams. *Biol. Fish*, **5**, pp. 343–360.

MORAN P. A. P. (1951), A mathematical theory of animal trapping. *Biometrika*, **38**, pp. 307–311.

NORRIS J. L. & K. H. POLLOCK (1996), Nonparametric MLE under two closed capture-recapture models with heterogeneity. *Biometrics*, **52**, pp. 639–649.

OTIS D. L., K. P. BURNHAM, G. C. WHITE & D. R. ANDERSON (1978), *Statistical Inference From Capture Data on Closed Animal Populations*. Wildlife monographs. Wildlife Society, Washington, 62 (135 pages).

PALOHEIMO J. E. (1961) Studies on the estimation of mortalities : I. Comparisons of a method described by Beverton and Holt and a new formula. *J. Fish. Res. Bd Can.*, **18**, pp. 645–662.

PLEDGER S. (2000), Unified maximum likelihood estimates for closed capture-recapture models using mixtures. *Biometrics*, **56**, pp. 434–442.

POINSARD F. & J. C. LE GUEN (1975), Observations sur la définition d'une unité d'effort de pêche applicable à la pêcherie de thon tropical africain. *Rapport P. V. Cons. Int. Explor. Mer.*, **168**, 39–43.

POLLOCK K.H. (1991), Modelling capture, recapture, and removal statistics for estimation of demographic parameters for fish and widlife populations : past, present, and future. *Journal of the American Statistical Association*, **86**, pp. 225–238.

PROUZET P., J.-P. MARTINET & F.-X. CUENDÉ (1996), Rapport sur la pêche des marins pêcheurs dans l'estuaire de l'Adour en 19*95. Rapport IFREMER, Saint-Pée sur Nivelle.*

PROUZET P., C. KOKONENDJI & B. TRUONG-VAN (1998), Evaluation du taux d'exploitation du saumon atlantique sur le Bassin de l'Adour ; Définition d'une méthode d'estimation- Application numérique pour 3 années de captures. *Rapport MATE-IFREMER-UPPA Contrat 9700057 24 pages + Annexes.*

REED W. J. & C. M. SIMONS (1996), Analyzing catch-effort data by means of the Kalman filter. *Can. J. Fish. Aquat. Sci.*, **53**, pp. 2157–2166.

RICKER W. E. (1975), Computation and interpretation of biological statistics of fish populations. *Bull. Fish. Bd Can.* 191, (382 pages).

RICKER W. E. (1980), Calcul et interpréation des statistiques biologiques des populations de poissons. *Bull. Fish. Bd Can.* 191F, (409 pages).

ROBSON D. S. & H. A. REGIER (1968), Estimation of population number and mortality rates. In Methods for assessment of fish production in freshwater. *Ricker eddition I. B. P. Handbook, 3. Blackwell, Oxford* pp. 124-138.

SCHNUTE J. T. (1982), A manual for easy non-linear parameter estimation in fishery research with interactive microcomputer programs. *Canadian Rech. Rep. Fisheries and Aquatic Sciences*, 1140 : xvi (115 pages).

SCHNUTE J. T. (1983), A new approach to estimating populations by the removal method. *Can. J. Fish. Aquat. Sci.*, **40**, pp. 2153–2169.

SCHNUTE J. T. (1985), A general theory for analysis of catch and effort data. *Can. J. Fish. Aquat. Sci.*, **42**, pp. 414–427.

SCHNUTE J. T. (1994), A general framework for developing sequential fisheries models. *Can. J. Fish. Aquat. Sci.*, **51**, pp. 1676–1688.

SCHWARZ C. J. & G. A. F. SEBER (1999), Estimating animal abundance : Review III. *Statistical Science*, **14**, pp. 427–456.

SEBER G. A. F. & E. D. LE CREN (1967), Estimating population parameters from catches large relative to the population. *Journal of Animal Ecology*, **36**, pp. 631–643.

SEBER G. A. F. & J. F. WHALE (1970), The removal method for two and three samples. *Biometrics*, **26**, pp. 393–400.

SEBER G. A. F. (1973), *The Estimation of Animal Abundance and Related Parameters.* Hafner Press.

SEBER G. A. F. (1986), A review of estimating animal abundance. *Biometrics*, **42**, pp. 267–292.

SEBER G. A. F. (1992), A review of estimating animal abundance. II. *International Statistical Review*, **60**, pp. 129–166.

SOMS A. P. (1985), Simplified point and interval estimation for removal trapping. *Biometrics*, **41**, pp. 663–668.

TRUONG-VAN B. (1998), Estimation de la biomasse des saumons dans le bassin de l'Adour. *Rapport du contrat IFREMER-ADERA 1997-1998 et Annexe technique du 28/07/1998.*

TRUONG-VAN B., C. KOKONENDJI & P. PROUZET (1998), Méthodes de maximum de vraisemblance. Applications aux saumons de l'Adour-Nive. *Biométrie et Halieutique*, **15**, pp. 123–128.

WANG Y. G. & N. R. LONERAGAN (1996), An extravariation model for improving confidence intervals of population size estimates from removal data. *Can. J. Fish. Aquat. Sci.*, **53**, pp. 2533–2539.

YAMAKAWA T., Y. MATSUMIYA & S. KITADA (1994), Comparison of statistical models for expanded Delury method. *Fisheries Sciences*, **60**, pp. 405–409.

ZIPPIN C. (1956), An evaluation of the removal method of estimating animal populations. *Biometrics*, **12**, pp. 163–189.

ZIPPIN C. (1958), The removal method of population estimation. *Journal of Wildlife Management*, **22**, pp. 82–90.