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

ÉTUDE DE L'INFLUENCE DES FACTEURS HYDROCLIMATIQUES SUR LA CAPTURE DE LA CIVELLE D'ANGUILLE (*ANGUILLA ANGUILLA L.*) PAR L'ANALYSE CANONIQUE FONCTIONNELLE

Noëlle Bru, S. Dossou-Gbété, Benoît Truong-Van [a]

I- INTRODUCTION

En raison de l'importance économique que revêt la pêche à la civelle sur l'Adour et de la nécessité d'assurer le renouvellement de cette ressource, le laboratoire IFREMER implanté à la station INRA de Saint-Pée-sur-Nivelle (64) a initié un programme de recherche dont un aspect consiste à étudier l'influence des facteurs hydroclimatiques sur la capturabilité des civelles.

II- MATÉRIEL ET MÉTHODES

A- Étude biologique

L'aire de reproduction de l'Anguille se situe dans la mer des Sargasses (Gascuel D., 1987). Après plusieurs migrations transocéaniques, les larves, à l'état civelle, atteignent leur zone de grossissement: les zones côtières et les cours d'eau de l'Europe et de l'Afrique du Nord. On remarque principalement que dans les zones estuariennes saumâtres: les civelles montent dans la colonne d'eau durant le flot: la migration vers l'amont est portée, et elles s'enfouissent dans le sédiment ou se plaquent sur le fond durant le jusant. Les principaux facteurs abiotiques qui influent sur la migration, donc sur la capturabilité, sont: la lumière, les coefficients de marée, les débits fluviaux, la température de l'eau. Il paraît intéressant de mettre en évidence par des méthodes statistiques, les formes des liaisons entre ces différents facteurs.

a - Laboratoire de Mathématiques Appliquées URA-CNRS 1204 - Université de Pau et des Pays de l'Adour - 64000 PAU (FRANCE); Membres du Pôle de Recherche sur la Gestion des Ressources Aquatiques en Environnement Sensible

B- L'Analyse Canonique

1) L'Analyse Canonique Multivariée Classique

a) Problématique

On cherche à mettre en évidence, lorsqu'elles existent, les liaisons linéaires (ou affines) entre deux suites finies de variables aléatoires réelles (v.a.r.) $X=(X_1,...,X_p)$ et $Y=(Y_1,...,Y_q)$. Si l'une des suites est effectivement une fonction linéaire de l'autre, alors on cherche à spécifier cette relation. Dans le cas contraire, on peut envisager de construire une approximation linéaire de l'une en fonction de l'autre. Il est alors nécessaire d'évaluer la qualité d'une telle approximation.

L'analyse canonique multivariée classique propose une méthode d'évaluation de ces approximations en faisant intervenir une suite décroissante de coefficients de corrélation linéaires dits canoniques qui servent d'indicateurs pour évaluer la qualité globale des approximations ci-dessus en fonction de la dimension retenue.

b) Notation

Soient : Γ_{11} (resp. Γ_{22}) la matrice des covariances associée à X (resp. Y) et Γ_{12} la matrice des covariances croisées.

On notera F_X (resp. F_Y) le sous-espace vectoriel engendré par les $X_1,...,X_p$ (resp. $Y_1,...,Y_q$). Soit l'application: Φ_X: $a \mapsto {}^tX.a = \sum_{i=1}^{p} a_i X_i$ et l'application Φ_Y correspondante pour Y.

c) Définitions et propriétés

Définitions:

1- On appelle coefficient de corrélation canonique maximal entre X et Y (ou premier coefficient de corrélation canonique), le réel positif ρ égal au maximum des coefficients de corrélation entre U et V où $U \in F_X$ et $V \in F_Y$.

2- On appelle couple de variables canoniques, tout couple (U,V) de variables centrées réduites qui réalise le maximum.

3- On dira que (u,v) est un couple de vecteurs canoniques associé au couple de variables canoniques (U,V) si: $U=\Phi_X(u)$ et $V=\Phi_Y(v)$.

On se rend compte que le coefficient de corrélation canonique maximal ne suffit pas pour évaluer la qualité de l'approximation (au sens des moindres carrés) d'un élément de F_X à partir de ceux de F_Y. L'information apportée par ce premier coefficient de corrélation canonique est alors complétée à l'aide des coefficients canoniques de rang suivant, ce qui conduit à la définition suivante de l'analyse canonique:

Définition: On appelle analyse canonique entre X et Y la recherche des variables centrées réduites $U \in F_X$ et $V \in F_Y$ telles que la corrélation entre U et V soit maximale avec itérations sous contraintes de non corrélation.

Proposition: Soit ρ_k le coefficient de corrélation canonique, ρ_k est la k-ième valeur propre de la matrice $\Gamma_{12}\, \Gamma_{22}^{-1}\, {}^t\Gamma_{12}$ (resp. $\Gamma_{21}\, \Gamma_{11}^{-1}\, {}^t\Gamma_{21}$), dans l'ordre décroissant, et les variables canoniques (U_k ,V_k) sont les vecteurs propres associés à la valeur propre ρ_k^2.

2) L'Analyse Canonique Fonctionnelle: Une généralisation de l' Analyse canonique Multivariée Classique.

a) Motivation

Les données disponibles pour cette étude correspondent aux observations journalières (du 1er Novembre au 31 Mars, soit environ 150 jours par saison) de plusieurs caractéristiques: capture moyenne des professionnels (c.p.u.e.), coefficients de marée, débits, pluviométrie, température de l'air, sur une période correspondant à 11 saisons de pêche. Considérant chaque saison comme un individu, chacune des séries de données correspond donc à des points d'une courbe décrivant l'évolution en fonction du temps d'une caractéristique de cet individu. La donnée associée à un individu par cette courbe n'appartient pas à un espace vectoriel de dimension finie mais à un espace vectoriel de fonctions dont la dimension est infinie. Pour des raisons techniques, on se restreint à une classe de fonctions possédant une certaine régularité (par exemple celles qui sont continues et dérivables). L'analyse canonique fonctionnelle, en tant que généralisation de l'analyse canonique multivariée classique, fournit un cadre adapté pour analyser les liaisons entre deux familles de v.a.r. appelées fonctions aléatoires.

b) Notation

soient: $(X_t)_{t\in T}$ et $(Y_t)_{t\in T}$, des processus aléatoires auxquels on fait correspondre les applications, $\Phi_x : a \mapsto \int_T a(t).X_t\, dt$ et respectivement Φ_Y.

c) Formulation de l'Analyse Canonique Fonctionnelle

On définit de manière analogue à ce qui précède le coefficient de corrélation canonique maximal et l'analyse canonique fonctionnelle (Leurgans S.E., Moyeed R.A. et Silverman B.W.,1993):

Définition: On appelle analyse canonique fonctionnelle entre les fonctions aléatoires $(X_t)_{t\in T}$ et $(Y_t)_{t\in T}$, la recherche des variables centrées réduites $\Phi_X(u)$ et $\Phi_Y(v)$ telles que la corrélation entre $\Phi_X(u)$ et $\Phi_Y(v)$ soit maximale, avec itérations sous contraintes de non corrélation.

On montre que les résultats énoncés pour une famille finie de v.a. se généralisent sous certaines conditions au cas des fonctions aléatoires et que l'on a les relations de passage suivantes entre les fonctions canoniques.

Proposition: $$\begin{pmatrix} 0 & \Gamma_{12} \\ {}^t\Gamma_{12} & 0 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix} = p\begin{pmatrix} \Gamma_{11} & 0 \\ 0 & \Gamma_{22} \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix}$$

III- DISCUSSION

L'interprétation des résultats se base sur l'analyse des fonctions canoniques u_k et v_k associées au k-ième coefficient de corrélation canonique. Cette interprétation est facilitée par les représentations graphiques simultanées des fonctions u_k et v_k, obtenues à l'aide d'estimations des opérateurs Γ_{ij}. Les estimateurs usuels de ces opérateurs nécessitent que le nombre de trajectoires observées soit plus important que le nombre de points de discrétisation de ces trajectoires. Pour cette raison, le nombre limité de trajectoires dont nous disposons (11), ne nous permet pas d'obtenir des résultats significatifs. Nous espérons résoudre cette difficulté par la modélisation des processus étudiés.

BIBLIOGRAPHIE

BESSE (P.), 1990 - *Approximation spline de l'Analyse en Composantes Principales d'une Variable Aléatoire Hilbertienne,* Laboratoire de Statistique et Probabilité, U.A. C.N.R.S. 745, Université Paul Sabatier, Toulouse.

BIRITXINAGA (E.), 1987 - *Estimation Spline de la Moyenne d'une Fonction Aléatoire,* Thèse, Université de Pau et des Pays de l'Adour.

BRU (N.), 1994 - *Étude de l'influence des facteurs hydroclimatiques sur la capture des civelles par l'Analyse Canonique Fonctionnelle,* Mémoire de DEA, Université de Pau et des Pays de l'Adour, Pau.

GASCUEL (D.), 1987 - *La civelle d'Anguille dans l'estuaire de la Sèvre Niortaise:biologie,écologie, exploitation,* Rapport Général, Les Publications du Département d'Halieutique, n°4/1, École Nationale Supérieure Agronomique de Rennes, Laboratoire de Biologie Halieutique.

LEURGANS (S.E.), MOYEED (R.A.), SILVERMAN (B.W.), 1993 - Canonical Correlation Analysis when the Data are Curves, *J.R. Statist. Soc.* B, 55, n°3, p 725-740.

TRUONG-VAN (B.), DOSSOU-GBETE (S.), 1994 - *Une modélisation stochastique de données hydroclimatiques,* Communication orale aux Journées Med.Campus "Modèles et problèmes mathématiques liés à la gestion des ressources renouvelables". Pau 24-26 Octobre 1994.