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

UNE MODELISATION STOCHASTIQUE DE L'IMPACT DES FACTEURS HYDROCLIMATIQUES SUR LES CAPTURES DES CIVELLES D'ANGUILLE

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

I - POSITION DES PROBLEMES

D'après les études biologiques (cf. Gascuel, 1987) , les remontées des civelles dans l'estuaire dépendent des conditions hydroclimatiques telles que la marée, les débits fluviaux, et la température. Un de nos thèmes de recherche conjoints avec les laboratoires INRA et IFREMER de Saint-Pée-sur-Nivelle est la quantification de l'influence des conditions hydroclimatiques sur la capturabilité des civelles. Cette étude s'appuie d'abord sur les données journalières (5 mois par an) recueillies depuis 1984 par le laboratoire IFREMER. Elles portent sur les captures par unité d'effort (CPUE) des pêcheries professionnelles, des relevés de débits fluviaux et de températures ainsi que les coefficients de marée.

II - LA DEMARCHE SUIVIE

Elle consiste d'abord, par une analyse exploratoire (harmonie et régression) des séries de données précédentes, à mettre en évidence l'influence des facteurs hydroclimatiques sur les CPUE, puis à améliorer la quantification de cette influence en utilisant des modèles stochastiques simples.

La marée, le débit et les CPUE partagent, bien qu'avec des amplitudes très différentes, une période quasi-commune autour de 12 jours. Une faible dépendance entre les CPUE et les débits et marées est plausible, suggérée aussi par les régressions linéaires entre les CPUE et les facteurs débit et marée. (Cf. tableau 1). Pour améliorer sa quantification, nous essayons alors de cerner les caractéristiques des CPUE avec des modèles simples. L'inspection des CPUE montre qu'elles s'expliquent par une composante g(t) liée aux flux des civelles

[a] - Pôle de recherche "Gestion des ressources aquatiques : Modélisation et Prédiction" Laboratoire de Mathématiques Appliquées, URA-CNRS 1204, Université de Pau et des Pays de l'Adour

et par la " variabilité V(t) des captures autour de g(t) ", imputable aux conditions hydroclimatiques. Nous représentons donc chaque CPUE à l'instant t, notée X(t) sous la forme X(t) = g(t) +V(t) +e(t) où e(t) inclut les autres perturbations. On suppose ici que la variabilité V(t) dépend linéairement des facteurs hydroclimatiques, que nous limitons pour l'instant à la marée et au débit. Aucune mesure sur les flux n'étant disponible, l'idée est alors "d'estimer" sa contribution g(t) dans les CPUE sous la forme g(t) = G(t,X) = G(X(t-s),...,X(t),..., X(t+s)). D'où le premier modèle:

(I) X(t) = m+ G(X(t-s),..., X(t+s)) + b M(t-d) + c D(t-d) +e(t) où M est la marée, D le débit, d un décalage, s largueur de la fenêtre d'estimation , m, b et c étant des paramètres.

Une estimation non-paramétrique de G est en cours. Nous proposons ici des choix simples suivants de G :

(Ia) G(t) = a_1 X(t-1) + a_2 X(t-2) + ...+ a_s X(t-s), où les a_j sont des paramètres;

(Ib) G est une moyenne mobile sur X(t-1),..., X(t-s);

(Ic) G est une moyenne mobile sur X(t-1+s),..., X(t-1-s);

(Id) G est une puissance k ième de X(t-1).

Une autre approche consiste à considérer la "variabilité" des CPUE traduite soit par leurs accroissements dX(t) = X(t) - X(t-1), soit par leurs taux d'accroissement relatifs TX(t) = dX(t) /(X(t-1)+r) où r est une constante positive. On peut expliquer la variabilité des CPUE par les schémas suivants :

(II) dX(t) = m + a X^k(t-1) + b M(t-d) + c D(t-d) +e(t)

(III) TX(t) = m+ a X^k(t-1) + b M(t-d) + c D(t-d) +e(t).

Le modèle (III) est une extension de celui de Richards en dynamique de populations.

III - RESULTATS ET INTERPRETATIONS

Les paramètres des modèles sont estimés par régression linéaire. Les principaux résultats des traitements sur les données de 1984 à 1992 sont reportés dans les tableaux ci-dessous ainsi que les paramètres d'intérêt : le coefficient R^2, l'écart-type estimé des résidus (S.E), les estimations des paramètres autres que m et leurs écarts-types accolés entre parenthèses. Dans ces tableaux, nous notons X pour X(t), X1 pour X(t-1), M1 pour M(t-1), etc.

1) Les résultats des régressions linéaires de X sur M(t-d) et D(t-d) sont illustrés dans le tableau 1. Il est réaliste de prendre le pas de décalage d entre 0 et 1. Les valeurs de R^2 entre 0.002 et 0.04 indiquent la faible dépendance linéaire des CPUE vis à vis des facteurs, corroborées par une étude des corrélations croisées. Ainsi, les travaux antérieurs utilisant ces corrélations croisées avec décalage conduisent à des résultats peu significatifs. D'autre part, les estimations des coefficients de M(t-d) et D(t-d) sont souvent proches, voire inférieures à leur écart-type et donc faiblement significatives. On note surtout que les régressions précédentes conduisent sur certaines séries à des résultats contradictoires (e.g série 1986).

Tableau 1: Régressions de X ou de dX sur M(t-d), D(t-d); d = 0, 1.

		R^2	S.E	coef (M)	coef (D)
1)	X : M, D	.008	1.8	-3E-4 (7.5E-3)	-5.4E-5 (6.2E-4)
	dX : M1, D1	.017	1.6	8.1E-3 (7.12E-3)	-1.0E-4 (5.8E-4)
2)	X : M, D	.038	1.6	0.01 (.006)	-6.6E-4 (5.2E-4)
	dX : M, D	.013	1.9	- 7.7E-4 (74E-4)	-8.3E-4 (6E-4)
3)	X : M, D	.048	2.5	0.007 (.010)	2.6E-3 (9.8E-4)
	X : M1, D1	.041	2.5	0.010 (.010)	22E-4 (9.9E-4)

1) Série 1984 2) Série 1985 3) Série 1986

2) Des résultats sur le modèle (I) sont reportés dans le tableau 2, où sY1 (resp. sZ1) désigne une moyenne mobile sur X(t-1),..., X(t-s) (resp. sur X(t-1+s),...., X(t-1-s)). Les meilleurs choix obtenus sont s = 1 pour (Ia); s = 2, 3, 4 pour (Ib). Le modèle (Ic) ne donne pas de résultat satisfaisant. L'impact des facteurs M et D sur X devient sensible via ces modèles. Selon les critères R^2 et S.E, le modèle (Ia) fournit les meilleurs ajustements linéaires mais les coefficients estimés sont moins précis. Par contre, le modèle (Ib) donne, sur la plupart des séries, des estimations plus précises.

Tableau 2: Régressions de X sur G, M(t-d) et D(t-d)

	Facteurs	R^2	S.E	coef (M)	coef (D)	coef(G)
1)	X1	.38	1.50			.61 (.07)
	M1,D1,X1	.39	1.4	7.3E-3 (6.2E-3)	-7.3E-4 (5.1E-4)	.61 (.07)
	M1,D1,2Y1	.15	1.71	12E-3 (7.5E-3)	-10E-4 (6.1E-4)	.40 (.09)
	M1,D1,2Z1	.89	.60	3.0E-4 (26E-4)	4E-5 (21E-5)	1.1 (.03)
2)	M,D,X1	.15	1.53	9.2E-3 (6.2E-3)	-7.4E-4 (5E-4)	.33 (.08)
	M,D,2Y1	.11	1.57	9.4E-3 (6.4E-3)	-9.4E-4 (5.1E-4)	.33 (.1)
	M,D,4Y1	.06	1.61	11E-3 (6.6E-3)	-9.0E-4 (5.2E-4)	.24 (.1)
3)	M, D, X1	.59	1.65	5.8E-3 (6.7E-3)	4.3E-4 (6.7E-4)	.76 (.05)
	M1, D1, X1	.59	1.65	4.1E-3 (6.7E-3)	2.5E-4 (6.7E-4)	.76 (.05)
	M,D,4Y1	.20	2.33	19E-3 (10E-3)	5.3E-4 (9.8E-4)	.33 (.1)
	M1,D1,$X^{1/2}$1	.48	1.86	2.9E-4 (76E-4)	2.9E-4 (7E-4)	2.0 (.18)

1) Série 1984, 2) série 1985 et 3) série 1986

3) Dans le but de chercher un compromis entre un bon ajustement linéaire de modèle et des estimations plus précises des coefficients des facteurs, nous introduisons les modèles (I) et (III). Les résultats obtenus sont illustrés dans les tableaux 3 et 4. L'impact des facteurs est mieux mis en évidence avec (II). Le modèle (II) avec k = 1 "équivaut" au modèle (Ia) avec s = 1. Cependant, les simulations (dans le but de validation) de modèles semblent indiquer que les données sont mieux représentées par (II) que par (Ia). Une des raisons est que le modèle (Ia) est similaire à un processus autorégressif d'ordre 1, dont le comportement ne peut pas traduire la grande asymétrie des valeurs des CPUE. Les séries dX ont une modulation d'amplitudes que nous attribuons aux flux. Ceci conduit à introduire le facteur $X^k(t-1)$. Sur la plupart des série, le meilleur choix de k est entre ½ et ¼ et le modèle (II) réalise le meilleur compromis cherché. Les taux d'accroissement relatifs TX sont des notions utilisées en dynamique de population. Nous reportons dans le tableau 4, les résultats des traitements avec (III). Ils indiquent que le meilleur choix de k de l'ordre de 1/10, que ce modèle conduit à des ajustements meilleurs que (II) mais donne des estimations des coefficients moins précises.

Tableau 3: Série 1984 : Régressions de dX sur $X^k(t-1)$, M(t-d), D(t-d)

Facteurs	R²	S.E	coef (M)	coef (D)	coef(X^k)
X1	.19	1.60			-.38 (.07)
M1, D1, X1	.39	1.44	7.3E-3 (6.2E-3)	-7.3E-4 (5.1E-4)	-.38 (.07)
M1, D1, $X^{1/2}1$	.19	1.46	8.5E-3 (6.3E-3)	-7.4E-4 (5.2E-4)	-.85 (.16)

Tableau 4: Série 1984 : Régressions de TX sur $X^k(t-1)$, M(t-d), D(t-d)

Facteurs	R²	S.E	coef (M)	coef (D)	coef(X^k)
M1, D1, $X^{1/4}1$	.25	4.56	0.019 (0.019)	-0.004 (0.0016)	-4.5 (0.7)
M1, D1, $X^{1/10}1$	.27	4.50	0.024 (0.019)	-0.004 (0.0016)	-5.5 (0.8)

Tableau 5: Régressions obtenues avec les modèles (I), (II) et (III)

		R²	coef (M)	coef (D)	coef(G)
1)	X: M, D,4Y1	.11	9.4E-3 (6.4E-3)	-9.4E-4 (5E-4)	-.33 (.01)
	dX:M,D,X1	.34	9.2E-3 (6.2E-3)	-7.4E-4 (5E-4)	1.0 (.10)
	dX: M,D,$X^{1/2}1$	.30	0.010 (0.006)	-8.0E-4 (5E-4)	1.33 (.17)
	TX: M1, D1, $X^{1/10}1$	.28	0.024 (0.019)	-0.004 (0.001)	-5.8 (.84)
2)	X: M, D, 4Y1	.20	.02 (.01)	5.3E-4 (10E-4)	0.50 (.10)
	dX: M1, D1, X1	.12	0.004 (0.007)	2.5E-4 (6.7E-4)	-0.24 (.05)
	TX:M1,D1,$X^{1/10}1$	.11	0.01 (0.02)	-2.3E-4 (20E-4)	-3.7 (.85)
3)	X : M, D, 4Y1	.45	0.02 (0.007)	-13E-4 (4.8E-4)	0.71 (.07)
	dX : M, D, $X^{1/2}1$	.17	0.010 (0.006)	-10E-4 (4.1E-4)	-0.80 (.16)
	TX : M, D, $X^{1/20}1$	.23	0.003 (0.007)	-4.4E-4 (4.5E-4)	-2.4 (0.4)

(1) Série 1985, (2) série 1986, (3) série 1987

IV - CONCLUSIONS ET DISCUSSIONS

Il ressort des résultats présentés dans le tableau 5 que : (i) les meilleurs modèles réalisant le compromis cherché sont (Ib) et (II) avec k entre 1 et 1/2, (ii) le modèle (I) est adéquat pour l'estimation des flux de civelles à partir des CPUE, (iii) le modèle (II) convient davantage à la prédiction.

Dans cette étude, nous avons adopté des modèles linéaires, mais il apparaît que la dépendance entre CPUE et les facteurs est probablement non-linéaire. Nous avons commencé des traitements avec des transformations (puissances, polynômes homogènes de degré 2 sur X(t-1), M(t-d) et D(t-d)).

Un autre effet pouvant affecter les estimations est la corrélation temporelle des facteurs stochastiques. Le débit est bien représenté par un processus AR(1) (le facteur marée est ici déterministe). Pour une validation des modèles, il sera nécessaire d'introduire cet effet et d'utiliser des régressions pondérées.

BIBLIOGRAPHIE

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, Ecole Nationale Supérieure Agronomique de Rennes, Laboratoire de biologie halieutique

TRUONG (B.) & DOSSOU-GBETE (S.), 1995 - Etude de la subsidance de plateformes pétrolières. Rapport de contrat avec ELF

TRUONG (B.) & DOSSOU-GBETE (S.), 1994 - Une modélisation stochastique de données hydroclimatiques. Journées Med. Campus "Modèles et problème mathématique liés à la gestion des ressources renouvelables". Pau, 14-26 Oct. 1994