



HAL
open science

Développement et ajustement d'un modèle de dynamique des populations structuré en longueur et spatialisé appliqué au stock Nord de merlu (*Merluccius merluccius*)

Hilaire Drouineau

► To cite this version:

Hilaire Drouineau. Développement et ajustement d'un modèle de dynamique des populations structuré en longueur et spatialisé appliqué au stock Nord de merlu (*Merluccius merluccius*). Sciences et techniques des pêches. Ecole Nationale d'Agronomie de Rennes, 2008. Français. NNT : . tel-03216330

HAL Id: tel-03216330

<https://hal.science/tel-03216330>

Submitted on 3 May 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

N° Ordre : 2008-4

N° Série : H-73

THESE:

présentée devant l'Ecole Doctorale Vie-Agro-Santé

pour obtenir

Le grade de : DOCTEUR DE L'ECOLE
NATIONALE SUPERIEURE D'AGRONOMIE
DE RENNES

spécialité : Halieutique

par

Hilaire DROUINEAU



photo: IFREMER Brest

**Développement et ajustement d'un modèle de
dynamique des populations structuré en longueur et
spatialisé appliqué au stock Nord de merlu (*Merluccius
merluccius*)**

Soutenue le 08 février 2008 devant la commission d'Examen

Didier Gascuel	Agrocampus Rennes	Président
Robert Faivre	INRA, Toulouse	Rapporteur
Francis Laloë	IRD, Montpellier	Rapporteur
Jean-Claude Brêthes	ISMER, Rimouski	Rapporteur
Dominique Pelletier	IFREMER, Nantes	Directrice de thèse
Stéphanie Mahévas	IFREMER, Nantes	Membres
Michel Bertignac	IFREMER, Brest	Membres

Appuyez-vous sur des principes : ils finiront bien par céder!
(Oscar Wilde)

REMERCIEMENTS

je remercie tous ceux qui se reconnaîtront...

Comment ne pas commencer par remercier les encadrants qui m'ont accompagné au cours de ces 3 années (et quelques...) de thèse? Sachez Stéphanie, Michel et Dominique que j'ai vraiment apprécié de travailler avec vous. Je vous remercie pour la confiance et le soutien que vous m'avez accordé tout au long de cette période. J'espère que l'avenir nous permettra de nous retrouver de nouveau. J'associe également à ces pensées les membres de mon comité de thèse qui ont toujours été là quand j'ai eu besoin de leur aide et dont les conseils m'ont beaucoup apporté.

De même, un grand merci aux membres de mon jury qui ont accepté de juger mon travail, et mille excuses aux rapporteurs pour leur avoir partiellement gâché les fêtes de fin d'année.

Je suis content d'évoquer ici les fidèles de: feu « le jorky Nantes », feu « l'Opércule » (mon passage aura vraiment été destructeur...), du « Webb-Ellis » et des pauses café, où nous avons passé tant de bons moments. Parmi eux, bon nombre de stagiaires, CDD, accueil, thésards (...) que je ne citerai pas de peur d'en oublier. Une petite dédicace malgré tout à:

- mon colocataire Jérémy: ça va faire bizarre de pas voir l'euro ensemble! Je te charge des pronostics. Tu prendras bien soin de notre plante en mon absence!
- Laure: merci pour tous tes bons conseils (surtout ceux en fin de thèse qui m'ont vraiment aidé). Pour te remercier j'ai fixé ma soutenance l'après-midi...
- Youen: tu crois que le Canada, c'est aussi agréable que le Danemark? Au pire je me consolerais en pensant à Boulogne...
- Yves: en espérant que tu te souviennes de 2 ou 3 trucs quand même... Bonne chance pour mon remplacement auprès de Olivier...
- Jacques: je me permets de te rappeler que tu me dois des cours de musique...
- Sigrid: heureusement que tu t'es inscrite à l'ENSAR, mon éthique personnelle m'aurait empêché de remercier quelqu'un de l'INAPG!
- Martin: Olut!

Je voudrai également souligner les excellentes semaines à bord du Gwen Drez pendant les campagnes marquage merlu, que ce soit avec les ifremeriens de Brest ou les membres d'équipage. Une mention spéciale pour Hélène, Aurélie (merci pour les massages), Fanch (Brasil!), Stéphane (« c'est bon! »), Hervé (il vaut mieux ne rien dire...), Alex et Patoche (pour la découverte de la taverne des Korigans), Jean-René et Paco (pour les cours de ramandage). Je garderai toujours un souvenir ému du moment où notre chef de mission brésilienne décrocha son drapeau brésilien un soir de juin 2006 au mouillage devant Groix, après que la France ait brillamment éliminé les auriverde pourtant largement favoris.

Enfin je suis bien entendu très reconnaissant pour l'excellent accueil que j'ai reçu au sein du département EMH au cours de ces trois ans, qui a grandement contribué à rendre ce travail si agréable.

Pour conclure, un regret: ne pas avoir réussi à le faire progresser monsieur Berthelé d'un iota en trois (décidément « 3 » est le chiffre de l'année!) ans de ping-pong. Quel bien piètre professeur je fais! J'espère qu'il trouvera le temps de prendre un avion (au pas de sa porte...) pour venir me voir au Canada.

Ayant pris l'habitude ces derniers temps de distribuer des cartons, je voudrai juste terminer ce petit préambule par les 5 cartons jaunes de ma thèse:

1: les merlus de mon modèle qui ont obstinément refusé de grandir contrairement à ceux du LASAA

2: Jean-Luc Gripond, Rudy Roussillon et Fabien Barthez à cause de qui j'ai subi les pires moqueries de la part de Jérémy, Arnaud et Youen (des moqueries d'un rennais!...) tout au long de la saison 2006-2007

3: Nouméa parce que c'est énervant tous ces gens qui vous montrent des photos d'atolls et de poissons chatoyants alors que vous luttez contre les affres des hivers nantais

4: les juges de l'eurovision qui n'ont pas su reconnaître le talent des fatals picards

5: tous les vins blancs autres que le muscadet

*On dirait un fanatique
De la cause halieutique,
Avec sa belle canne et
Son moulinet.
Mais s'il pêche, c'est pour rire,
Et l'on peut être certain
Que jamais sa poêle à frire
Vit le plus menu fretin.*

(Georges Brassens)

Table des matières

Introduction générale.....1

Une espèce très étudiée mais dont certains processus biologiques sont mal connus.....	1
Une espèce dont le poids socio-économique est historiquement fort.....	2
Un stock dans un état inquiétant au début des années 2000.....	2
Des diagnostics remis en cause.....	3
Vers un nouveau type de modèles.....	4
Objectif du modèle et plan du mémoire.....	4

Chapitre 1 : La pêche de merlu de l'Atlantique Nord-Est.....9

1.1 Introduction.....	9
1.2 Biologie et écologie du merlu européen.....	10
1.2.1 Structure du stock.....	10
1.2.2 Alimentation.....	12
1.2.3 Croissance.....	12
1.2.4 Migration et distribution spatiale et saisonnière.....	15
1.2.5 Maturité – Fécondité – Reproduction – Recrutement.....	16
1.3 Exploitation et gestion du stock Nord de merlu.....	20
1.3.1 Exploitation du merlu: évolutions des captures et des métiers.....	20
1.3.1.1 <i>Évolution des débarquements (Figure 8)</i>	20
1.3.1.2 <i>Description par pays</i>	22
1.3.1.3 <i>Description des unités de pêche</i>	24
1.3.2 Évaluation des stocks.....	28

1.3.3 Réglementation de la pêche.....	30
1.3.3.1 Mesures techniques: tailles minimales et maillages.....	30
1.3.3.2 Les TAC (Totaux Admissibles de Captures).....	31
1.3.3.3 Mesures spatialisées.....	31
1.3.3.4 Le plan de restauration.....	31
1.4 Conclusion.....	33

Chapitre 2 : Éléments sur les modèles de dynamique des populations structurés en longueur.....37

2.1 Introduction.....	37
2.2 Objectifs et constitution des modèles structurés en longueur.....	39
2.3 Modélisation des processus.....	41
2.3.1 Maturité-fécondité.....	41
2.3.2 Mortalité naturelle – prédation.....	42
2.3.3 Croissance.....	42
2.4 Complexité des modèles.....	44
2.5 Les types de données utilisées pour ajuster les modèles structurés en longueur.....	46
2.5.1 Données commerciales.....	46
2.5.2 Données de campagne.....	48
2.5.3 Données de marquage.....	49
2.5.4 Données de contenus stomacaux.....	49
2.6 Ajustement des modèles aux données observées.....	50
2.6.1 Méthodes d'ajustement.....	50
2.6.2 Choix de l'algorithme d'optimisation.....	51
2.7 Utilisation de modèles structurés en longueur ou en âge et longueur.....	52
2.7.1 Hors Europe.....	52

2.7.2 En Europe.....	53
2.8 Conclusion.....	53

Chapitre 3 : Quantification de l'impact d'hypothèses de discrétisation sur un modèle de croissance structuré en longueur.....61

3.1 Introduction.....	61
3.2 Matériel et méthodes.....	62
3.2.1 Un modèle de croissance de population structuré en longueur.....	62
3.2.2 Simulation de jeux de trajectoires de croissance individuelles.....	68
3.2.3 Plan d'expériences (Kleijnen, 1998 ; Kleijnen, 2001).....	69
3.2.4 Analyse du plan d'expériences.....	70
3.3 Résultats.....	73
3.3.1 Ajustement des compositions en longueur (Tableau 6).....	75
3.3.2 Ajustement du taux de croissance K (Figure 15 et Tableau 7).....	75
3.3.3 Analyse de l'arbre de régressions multiples (Figure 16 et Tableau 8).....	77
3.4 Discussion.....	79

Chapitre 4 : Un modèle de dynamique des populations structuré en longueur et spatialisé appliqué au merlu de l'Atlantique Nord-Est.....87

4.1 Introduction.....	87
4.2 Matériel et méthodes.....	88
4.2.1 Modèle d'états.....	90
4.2.1.1 Dynamique de la population.....	91
4.2.1.2 Calcul de la mortalité par pêche et des captures.....	95
4.2.1.3 Calcul des indices d'abondance de campagnes scientifiques.....	100
4.2.1.4 Effectifs initiaux.....	101
4.2.2 Modèle d'observations.....	102

4.2.2.2 Paramètres inconnus et données disponibles.....	105
4.2.2.3 Fonction objectif et algorithme d'optimisation.....	106
4.2.3 Identifiabilité du modèle: calcul de la matrice hessienne et de la matrice des corrélations.....	107
4.3 Résultats.....	108
4.3.1 Ajustement aux données observées.....	109
4.3.1.1 Ajustement aux compositions en longueur des débarquements.....	110
4.3.1.2 Ajustements aux débarquements totaux en poids.....	113
4.3.1.3 Ajustements aux indices d'abondance scientifiques.....	116
4.3.2 Identifiabilité du modèle.....	117
4.3.3 Evolution de la biomasse féconde et de sa distribution spatiale, évolution du recrutement.....	119
4.3.4 Croissance estimée.....	123
4.4 Discussion.....	124
<u>Discussion générale et perspectives.....</u>	133
<u>Références.....</u>	141
<u>Annexe 1 : Présentation d'algorithmes d'optimisation.....</u>	157
Introduction.....	157
1 Gradient.....	157
2 Simplex.....	158
3 Algorithmes de Newton et de quasi-Newton.....	158
3.1 Algorithme de Newton.....	158
3.2 Algorithme de quasi-Newton.....	159

4 Algorithme génétique (Goldberg, 1989 ; Michalewicz, 1996).....	160
Conclusion.....	163

Annexe 2 : Identifiabilité des paramètres.....167

Introduction.....	167
1 Courbes de profil de vraisemblance.....	168
2 Matrice hessienne.....	168
3 Calcul formel de la matrice des dérivées.....	170
4 Simulation.....	170

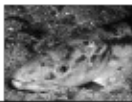
**Annexe 3 : Vers la simulation de règles de gestion à l'aide du simulateur
ISIS-Fish.....173**

Introduction.....	173
1 Adaptation du simulateur.....	173
2 Modèle de dynamique de la population.....	176

**Annexe 4 : Combiner erreur de processus et erreur d'observations dans un
modèle de dynamique des populations.....181**

Introduction.....	181
1 Les méthodes de Monte-Carlo direct.....	183
2 Importance sampling.....	184
3 Filtrage particulière.....	184
4 Algorithme d'espérance-maximisation (EM).....	186
4 Monte-Carlo kernel likelihoods (de Valpine, 2003 ; de Valpine, 2004 ; de Valpine et Hilborn, 2005).....	186
4.1 Principe général.....	186

4.2 Construction de la chaîne MCMC.....	187
4.2.1 <i>Quelques éléments sur les chaînes MCMC - algorithme de Metropolis Hastings (Gilks et al., 1995)</i>	187
4.2.2 <i>Par un échantillonneur de Gibbs (Gilks et al., 1995)</i>	189
4.2.3 <i>Par un échantillonneur hybride</i>	189
4.2.4 <i>Par un algorithme couplant algorithme génétique et algorithme de Gibbs (Yardim et al., 2005)</i>	190
Conclusion.....	190
<u>Résumé</u>	192



Introduction générale



Favet Neptunus eunti

*(Neptune favorise ceux qui voyagent – devise de la ville de
Nantes)*

Introduction générale

Une espèce très étudiée mais dont certains processus biologiques sont mal connus

Le merlu européen (*Merluccius merluccius*) est une espèce largement répandue dans l'Atlantique Nord-Est et en Méditerranée. Cette espèce a eu un rôle socio-économique fort dans les pêcheries européennes au cours du XX^e siècle, notamment en Espagne, en France et au Royaume-Uni. Ceci explique sans doute l'abondante littérature scientifique ayant trait à cette espèce et qui a permis d'accumuler une bonne connaissance sur certains des aspects de sa biologie (maturation, alimentation...). En revanche, en dépit de nombreuses études, d'autres processus de la biologie, comme la croissance et les migrations, sont encore mal connus ; de nombreuses hypothèses ont été formulées mais aucune analyse n'a réellement permis de les valider. Or ces incertitudes peuvent potentiellement avoir des conséquences fortes sur la gestion de la pêche, tant sur les diagnostics de l'état du stock que sur le choix de la réglementation de l'accès à la ressource (mesures de gestion alternatives comme les aires marines protégées). Le manque de fiabilité et de validation des techniques usuelles de détermination de l'âge a par exemple mené pour le merlu à une estimation d'un taux de croissance fortement remise en question par de récentes campagnes de marquage. Or cette sous-estimation du taux de croissance, si elle ne semble pas affecter le diagnostic sur l'état du stock, modifie fortement les perspectives quant à la réaction du stock à des mesures de gestion (Bertignac et de Pontual, 2007). De même, le comportement migratoire, bien qu'étudié au travers de l'évolution saisonnière des taux de captures, n'a jamais pu être réellement validé. Ce manque de connaissance sur les migrations et la distribution spatiale de l'espèce peut potentiellement affecter la qualité des diagnostic (Pelletier et Magal, 1996) et est surtout problématique si l'on envisage la mise en place de mesures spatialisées comme les aires marines protégées.

Une espèce dont le poids socio-économique est historiquement fort

L'exploitation du merlu par les flottilles européennes est relativement ancienne, puisque des captures de merlu sont mentionnées dès le XV^e siècle (Casey et Pereiro J, 1995). De par son poids socio-économique, cette pêcherie a joué un rôle central dans la mise en place de la réglementation française et européenne des pêches au cours du XX^e siècle. Actuellement, l'Espagne et la France représentent l'essentiel des captures, le reste se distribuant entre le Royaume-Uni, l'Irlande, le Danemark, la Norvège, la Belgique, les Pays-Bas, l'Allemagne et la Suède. Le merlu reste encore en France la huitième espèce dans les criées françaises (pour un chiffre d'affaires de 19,7 millions d'euros en 2006 - source OFIMER).

Un stock dans un état inquiétant au début des années 2000

Au début des années 2000, il est apparu que ce stock était en dehors des limites biologiques permettant une exploitation durable: le stock de géniteurs était en effet estimé en-dessous de la biomasse féconde minimale supportable par la population (biomasse de référence de l'approche de précaution classiquement notée B_{pa}) et la pression de pêche estimée au-dessus de la pression de pêche maximale supportable par la population (mortalité par pêche de l'approche de précaution F_{pa}). Un plan visant à reconstituer le stock de merlu a ainsi été promulgué par la Commission Européenne en 2004 (règlement (CE) 811/2004) ; il faisait suite à une série de mesures techniques visant à protéger les juvéniles (les règlements (CE) n°1162/2001, 2602/2001, 2256/2002, 494/2002). Ce plan définit un objectif de gestion clair en terme de biomasse féconde (atteindre B_{pa} soit 140 000t), et les règles de fixation des totaux admissibles de captures (TAC) permettant de l'atteindre.

Plusieurs arguments ont été invoqués pour expliquer le mauvais état du stock. 1) A la fin des années 1990, début des années 2000, on a observé une série de bas recrutements (en halieutique, on parle de recrutement pour désigner l'arrivée dans une pêcherie, que ce soit par une migration ou par l'acquisition d'un caractère particulier). Le recrutement du merlu étant plus lié à la survie des oeufs et des jeunes qu'au stock d'individus matures, les conditions

environnementales ont probablement été défavorables. 2) Le non-respect de la réglementation en place (Commission of the European Communities, 2002), tant des mesures techniques (maillages minimaux et taille minimale de capture) que des quotas nationaux a probablement eu, comme pour de nombreuses pêcheries, un impact fort (Jones, 2007). 3) Les diagnostics et régulations mono-spécifiques dans un contexte de pêche multi-flottes et multi-spécifiques peuvent apparaître inadaptés. 4) La qualité des diagnostics même mono-spécifiques est parfois remise en question.

Des diagnostics remis en cause

Pour le merlu, le diagnostic est réalisé par une méthode indirecte: un modèle de dynamique des populations, dans le cas du merlu Extended-Survivors Analysis (XSA) (Shepherd, 1999), est utilisé afin d'estimer les abondances et de quantifier la pression exercée par la pêche (on parle de mortalité par pêche) sur le stock. Ces modèles sont ajustés sur une série d'observations, qui sont, dans le cas du merlu et de nombreuses espèces marines, les captures historiques par classes d'âge et des séries d'indices d'abondance (mesures relatives de l'abondance) par classes d'âges.

Une des critiques souvent adressées à ces diagnostics concerne la mauvaise prise en compte des incertitudes par les modèles classiques d'évaluation de stocks. Cela concerne notamment la conversion taille-âge requise pour appliquer ces modèles, et qui apparaît peu fiable pour de nombreuses espèces, comme le merlu, dont la croissance, comme nous l'avons déjà mentionné précédemment, est mal connue. Un autre aspect est que de nombreux modèles, et notamment XSA ne prennent pas en compte les incertitudes autour de la matrice des captures aux âges (il existe cependant des modèles, dits « statistical catch-at-age », qui s'affranchissent de cette hypothèse). Enfin, la non prise en compte des hétérogénéités spatiales et saisonnières de la distribution des populations et de l'activité de pêche, est susceptible d'entraîner des biais dans l'évaluation des stocks (Pelletier et Parma, 1994), et ne permet pas l'évaluation de mesures de gestion spatialisée, comme les aires marines protégées.

Vers un nouveau type de modèles

Dans ce contexte, une nouvelle famille de modèles d'évaluation a récemment été développée. Ces modèles structurés en longueur (ou en âge et en longueur) plutôt qu'en âge, s'affranchissent de la conversion taille âge et permettent de décrire de nombreux processus biologiques ou de captures directement liés à la longueur (ICES, 2003d). De plus, la plupart d'entre eux couplent un modèle d'états, décrivant la dynamique des états du système (en général les abondances), et un modèle d'observations qui décrit le processus par lequel on obtient les observations des états. Ces deux sous modèles peuvent être stochastiques permettant ainsi de prendre en compte à la fois les incertitudes sur la dynamique des états et sur les observations. Au vu des incertitudes existantes sur la biologie du merlu et des mesures de gestion spatialisées envisagées pour réguler l'accès à la ressource, un modèle structuré en longueur et spatialisé paraît particulièrement adapté pour modéliser la dynamique de cette population.

Objectif du modèle et plan du mémoire

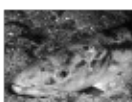
Les objectifs des modèles structurés en longueur peuvent-être assez variés. Dans notre cas, le modèle doit permettre de répondre à deux objectifs:

- améliorer la connaissance biologique sur l'espèce, en particulier sur sa croissance et sur les migrations à grande échelle
- permettre de réaliser des évaluations de stocks afin de diagnostiquer l'état de la population.

Pour cela, une synthèse bibliographique sur la biologie et l'exploitation du merlu a été réalisée afin de dégager les processus clés de la dynamique à prendre en compte dans le modèle (chapitre 1). En parallèle, une revue des modèles structurés en longueur (ou en âge et longueur) existants a été réalisée afin d'en dégager les principales caractéristiques susceptibles

de nous aider dans le développement et l'ajustement du modèle. Cette revue a notamment mis en évidence l'importance de la modélisation de la croissance dans cette famille de modèles (chapitre 2). Ce constat nous a poussé à réaliser une analyse de sensibilité d'un modèle de croissance de population aux hypothèses de discrétisation afin de proposer des hypothèses robustes réutilisables dans notre modèle (chapitre 3). Enfin le modèle a été développé et ajusté aux données du stock Nord (Figure 2) de merlu (chapitre 4).

Chapitre 1 : La pêche de merlu de l'Atlantique Nord-Est



Chapitre 1 : La pêche de merlu de l'Atlantique Nord-Est

1.1 Introduction

Le merlu (*Merluccius merluccius*) (Figure 1) est une espèce dont l'aire de distribution large s'étend de la Mauritanie jusqu'aux côtes norvégiennes (Casey et Pereiro J, 1995). Son exploitation est ancienne et a un poids socio-économique important, particulièrement en Espagne, en France et aux Royaume-Uni. De plus, cette espèce a joué un rôle de premier ordre dans la mise en place de la gestion des pêches. Ceci explique que sa biologie ait été particulièrement étudiée. Son exploitation reste encore aujourd'hui importante. Cependant des tendances récentes ont poussé à la mise en place d'un plan de restauration au début des années 2000.



Figure 1 : Photo d'un merlu capturé lors de la campagne de marquage 2005 (à gauche) et photo sous-marine d'un merlu (à droite -photo IFREMER Brest).

1.2 Biologie et écologie du merlu européen

1.2.1 Structure du stock

On trouve le merlu européen dans les eaux méditerranéennes et de l'Atlantique Nord-Est. Le détroit de Gibraltar constitue une frontière géographique limitant les échanges entre les populations atlantique et méditerranéennes, cette différenciation étant confirmée par des études génétiques (Roldan *et al.*, 1998 ; Castillo *et al.*, 2004 ; Cimmaruta *et al.*, 2005). Historiquement, pour son évaluation et sa gestion, la population de merlu de l'Atlantique Nord Est a été subdivisée en deux stocks, le Cap Breton étant supposé, comme Gibraltar, constituer une barrière naturelle limitant les échanges (ICES, 2006b). On a ainsi défini un stock Sud situé autour de la péninsule ibérique et un stock Nord s'étendant jusqu'aux côtes norvégiennes (Figure 2). Si la différenciation des populations entre l'Atlantique et la Méditerranée est démontrée, la division des stocks atlantiques est moins évidente. En effet, certaines études génétiques tendent à confirmer la séparation entre stock Sud et Nord actuellement en vigueur (Roldan *et al.*, 1998). D'autres au contraire tendent à remettre cette séparation stock Nord et stock Sud en question (Lundy *et al.*, 2000 ; Castillo *et al.*, 2004). Lundy *et al.* (2000) suggèrent que les stocks Sud et Nord pourraient effectivement être séparés à la saison de ponte, mais se mélanger le reste de l'année. Certaines études suggèrent également de distinguer la partie la plus au nord du stock Nord (Roldan *et al.*, 1998 ; Castillo *et al.*, 2004).

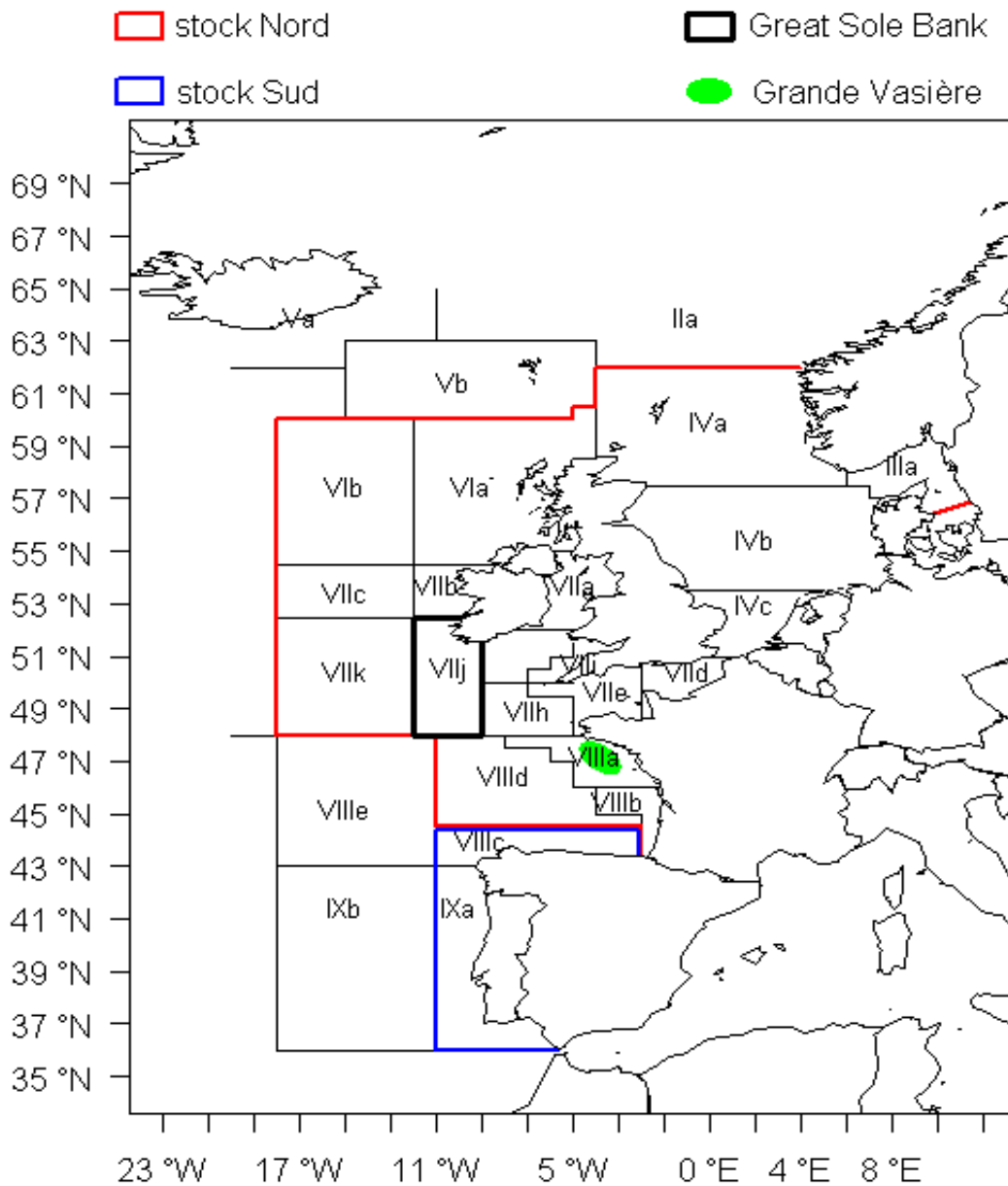


Figure 2 : Carte présentant les stocks Nord (ligne rouge) et Sud (ligne bleu) et les sous-divisions définies par le CIEM.

1.2.2 Alimentation

Le régime alimentaire du merlu est assez varié (Guichet, 1995). Les adultes se nourrissent essentiellement de poissons (ICES, 2004b), en particulier maquereau (*Scomber scombrus*), anchois (*Engraulis encrasicolus*), merlan bleu (*Micromesistius poutassou*) pour les merlus du golfe de Gascogne (Guichet, 1995). Les petits individus se nourrissent essentiellement de crustacés. Le cannibalisme est souvent mentionné pour le merlu, notamment en Méditerranée (Casey et Pereiro J, 1995). Toutefois Guichet (1995) note que dans le golfe de Gascogne, il est assez limité.

1.2.3 Croissance

La croissance du merlu est particulièrement mal connue, la détermination de son âge étant rendue complexe par les difficultés rencontrées pour interpréter les otolithes (Piñeiro et Sainza, 2003). Des ateliers réunissant des experts internationaux ont débouché à la mise en place d'une clé commune pour la lecture des otolithes (Figure 3), cependant les estimations restent peu fiables pour des individus de plus de trois ans (Piñeiro *et al.*, 2004). Les principaux problèmes résident dans la distinction entre anneaux annuels et faux anneaux, dans l'interprétation du bord de l'otolithe et dans la localisation de l'annulus (Kacher et Amara, 2005).

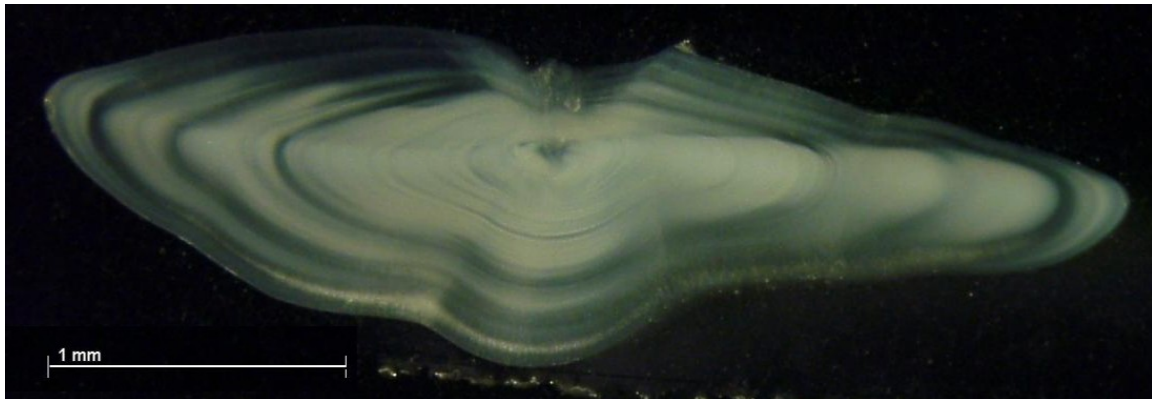


Figure 3 : Coupe transversale d'otolithe de merlu. Les stries sombres et claires correspondent à des macrostructures dont la périodicité peut être saisonnière, mensuelle. (photo IFREMER Brest).

Historiquement, le merlu a été considéré comme une espèce à croissance lente. Martin (1991) et Lucio *et al.* (2000) ont ainsi estimé à partir de lectures d'otolithes des taux de croissance K (K est assez classiquement appelé « taux de croissance » bien que cette grandeur ne soit pas à proprement parlé un « taux », nous utiliserons cette dénomination par la suite) d'une courbe de von Bertalanffy, $l(t) = L_{\infty} \cdot (1 - e^{-K \cdot (t - t_0)})$, égaux respectivement à $0,09 \text{ an}^{-1}$ et $0,124 \text{ an}^{-1}$ (Figure 4 et Tableau 1). Toutefois, au vu des difficultés liées à cette méthode, ces estimations étaient considérées incertaines. De plus, la composition en âge obtenue à partir de cette clé (Figure 5) n'était pas du tout cohérente avec la composition en longueur obtenue lors des campagnes FR-EVHOE (Bertignac et de Pontual, 2007).

	K (an ⁻¹)	L _∞ (cm)	T ₀ (an)
ICES (1993)	0.073	127.5	-1.130
Lucio <i>et al.</i> (2000)	0.124	110.0	-0.452
de Pontual <i>et al.</i> (2006)	0.362	89.9	
de Pontual <i>et al.</i> (2006) L _∞ fixé	0.250	110	

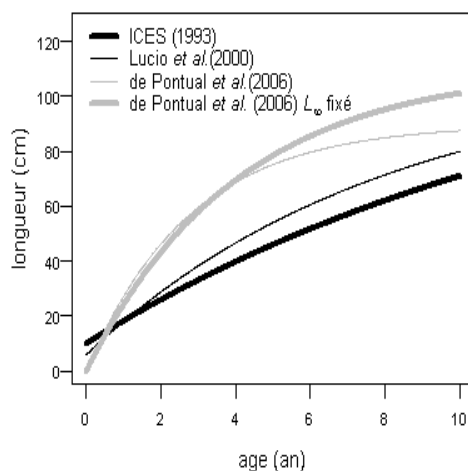


Tableau 1 : Paramètres de Von Bertalanffy estimés lors de différentes études. Toutes ces estimations concernent le golfe de Gascogne et sont sexes combinés.

Figure 4 : Courbes de croissance provenant de différentes estimations de paramètres de Von Bertalanffy estimés lors de différentes études. Toutes ces estimations concernent le golfe de Gascogne et sont sexes combinés.

Piñeiro et Sainza (2003) ont remis en cause cette hypothèse pour le stock Sud s'appuyant sur de nouvelles lectures d'otolithes. Ils notent d'ailleurs qu'une croissance rapide du merlu s'accorderait mieux aux résultats d'une étude bioénergétique du merlu du stock Sud (Riis-Vesterrgaard *et al.*, 2000). Des études récentes remettent également en cause les taux de croissance préalablement estimés pour le stock Nord. La campagne pilote de marquage effectuée par l'IFREMER dans le Golfe de Gascogne, puis les campagnes de marquages qui ont suivi en 2004, 2005, 2006 et 2007 semblent démontrer une croissance deux fois plus rapide que précédemment supposée (de Pontual *et al.*, 2003 ; de Pontual *et al.*, 2006). L'étude des stries journalières d'otolithes d'individus d'âge 0 fournit des taux de croissance comparables à ceux obtenus par la campagne de marquage et plus proches de la croissance d'autres espèces de merlu (*Merluccius hubbsi* et *Merluccius capensis*) (Kacher et Amara, 2005).

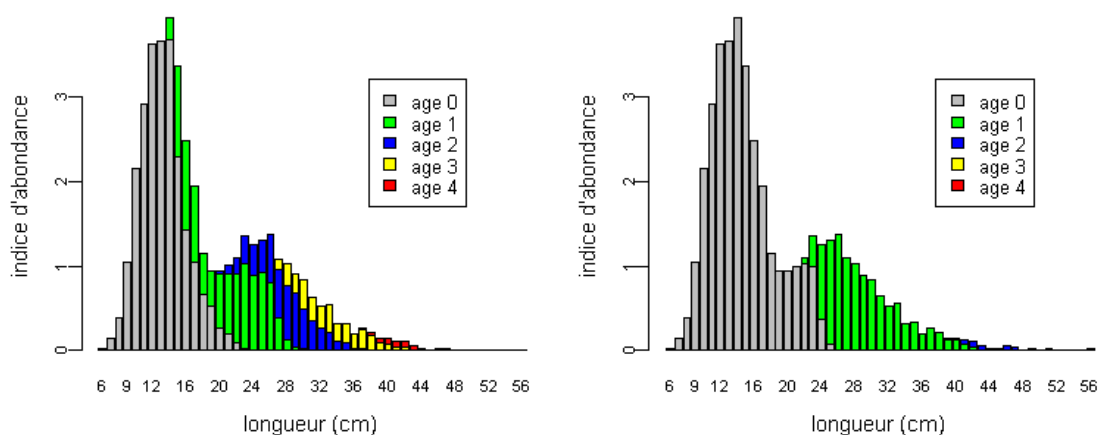


Figure 5 : Composition en longueur par groupe d'âge pour la campagne FR-EVHOE 2003 à partir de la clé taille-âge traditionnelle (à gauche) et de la clé taille-âge simulée (à droite) (Bertignac et de Pontual, 2007).

Une clé taille-âge a été simulée à partir de paramètres de croissance proches de ceux estimés à partir des données de marquage. La composition en âge obtenue à partir de cette clé (Figure 5) est beaucoup plus cohérente avec la composition en longueur observée lors de la campagne FR-EVHOE (Bertignac et de Pontual, 2007). Depuis 2003, le groupe de travail du Conseil International pour l'Exploration de la Mer (CIEM) fournit deux diagnostics, l'un basé sur les clés taille-âge, l'autre basé sur cette clé simulée (ICES, 2003c ; ICES, 2004b ; ICES, 2005b ; ICES, 2006b). Les tendances restent les mêmes, le diagnostic est peu changé. Cependant, une croissance plus rapide permet une reconstitution plus rapide du stock en cas de diminution de la pression de pêche (Bertignac et de Pontual, 2007).

1.2.4 Migration et distribution spatiale et saisonnière

Les migrations du merlu sont mal connues (Casey et Pereiro J, 1995). Elles ont été essentiellement étudiées par le suivi des taux de captures. Aucune campagne de marquage de merlu n'a eu lieu à grande échelle sur l'ensemble de l'aire de distribution. Les données récoltées lors de la campagne de marquage mentionnée plus haut sont encore limitées en nombre et ne concernent que le golfe de Gascogne (de Pontual *et al.*, 2003 ; de Pontual *et al.*, 2006).

Un premier travail réalisé par Hickling (1927) mentionnait qu'au mois de mars, les plus fortes captures de merlu avaient lieu dans des zones profondes à l'ouest de l'aire de la mer Celtique, alors qu'en juin, les captures étaient plus dispersées sur le plateau, puis de nouveau dans des eaux profondes entre septembre et décembre. Ces observations ont été ensuite corroborées par Belloc (1935) qui, en analysant les livres de bord de chalutiers rochelais, a observé que les captures les plus importantes sur le « Great Sole Bank » (Figure 2) avaient lieu de janvier à mars et entre octobre et décembre. Enfin, Poulard (2001) a étudié la distribution spatiale et saisonnière de différentes classes de taille à travers la distribution des CPUE des chalutiers français en mer Celtique et dans le golfe de Gascogne. L'étude met en évidence des différences importantes de distribution entre mer Celtique et golfe de Gascogne. Une prédominance de jeunes individus est observée dans le golfe de Gascogne, en particulier sur la Grande Vasière. Une telle abondance n'est à l'inverse pas observée en mer Celtique, probablement du fait de la différence de maillage en vigueur entre ces deux zones. Les individus adultes sont surtout observés en mer Celtique d'avril à novembre à l'est (est de 8°W), les plus vieux individus semblant migrer vers l'ouest à la fin de l'année et se concentrer au nord-ouest le long des marges du plateau continental de février à avril/mai. Des concentrations d'adultes sont également observées le long de la marge du plateau continental au nord du golfe de Gascogne en automne et en hiver, période où on les trouve en plus faible quantité en mer Celtique.

1.2.5 Maturité – Fécondité – Reproduction – Recrutement

La reproduction du merlu est étalée dans le temps. Elle s'étend de janvier à mai dans le golfe de Gascogne (Martin, 1991) avec un pic en mars (Alvarez *et al.*, 2001), d'avril à juin en mer Celtique (Fives *et al.*, 2001), d'avril à juillet dans la zone d'ouest Irlande (O'Brien, 1986) (Figure 6). On observe donc un déplacement du sud vers le nord de la zone de reproduction (Alvarez *et al.*, 2004), qui peut s'expliquer soit par une migration du nord vers le sud des reproducteurs, soit par un déplacement des conditions favorables au cours du temps (Alvarez *et al.*, 2001). Plusieurs facteurs environnementaux peuvent potentiellement affecter la reproduction. La température joue particulièrement sur la survie et le développement des

œufs, la température optimale se situant entre 10,5 et 12° (Valencia *et al.*, 1989). La salinité et les cycles de blooms phytoplanctoniques sont également susceptibles d'affecter la reproduction (Valencia *et al.*, 1989). La reproduction a lieu essentiellement le long de la marge du plateau continental (Alvarez *et al.*, 2001), une extension plus tardive vers le plateau continental est parfois mentionnée (Fives *et al.*, 2001).

Les larves sont ensuite transportées par les courants, vers les zones de nourricerie qui sont situées sur le plateau continental du golfe de Gascogne et de la mer Celtique (Figure 6). Les larves entraînées par le courant vers des zones défavorables (en particulier en dehors du plateau continental) meurent probablement (Alvarez *et al.*, 2001). Il est à noter que les courants induits par le vent jouent sur des couches trop superficielles pour influencer fortement sur le transport des œufs et des larves qui sont dans des couches plus profondes (0 à 50 mètres pour les œufs, 50-100 mètres pour les larves) (Alvarez *et al.*, 2004).

Le recrutement (à l'âge 0) a lieu sur deux nourriceries principales, l'une située sur la Grande Vasière (Figure 2) et l'autre au sud de l'Irlande (Casey et Pereiro J, 1995 ; Kacher et Amara, 2005). La distance importante entre les zones de ponte, située aux marges continentales, et les nourriceries situées plus vers les côtes entraînent une forte mortalité des œufs et larves au cours du transport. Les conditions environnementales ont donc un impact fort sur le recrutement (Alvarez *et al.*, 2004).

Jusqu'ici, aucune relation stock-recrutement n'a été établie pour le merlu, le seul essai à notre connaissance ayant consisté en un ajustement d'une relation linéaire par segments donnant des résultats peu concluant (ICES, 2003b). Le groupe de travail sur la préparation des mesures de reconstitution du stock Nord de merlu juge que l'étude de telles relations n'est pas une priorité (Commission of the European Communities, 2001).

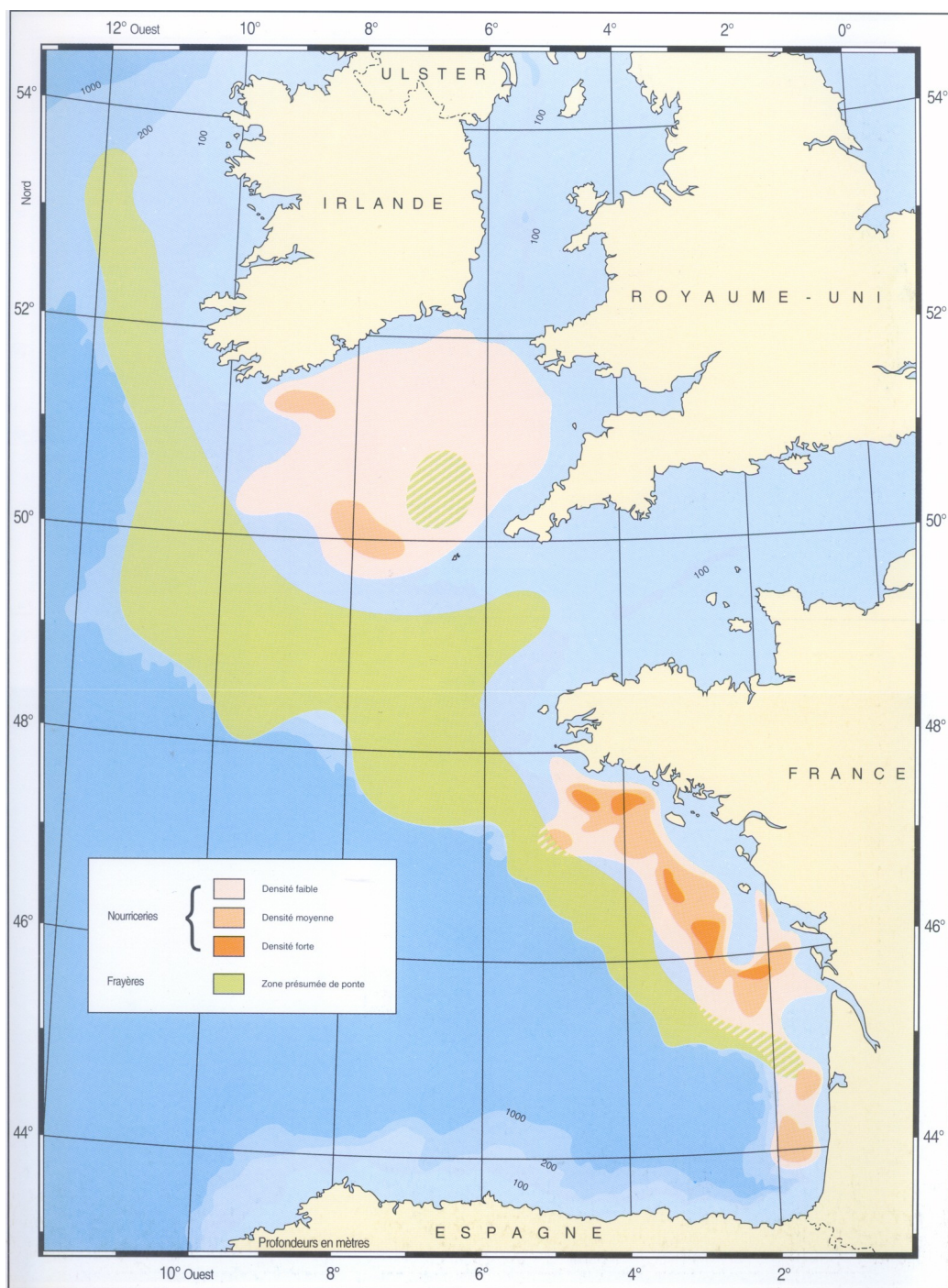


Figure 6 : Carte des principales nourriceries et zones de pontes du stock Nord de merlu (source (Abbes, 1991)).

L'ogive de maturité actuellement utilisée provient de l'étude de Martin (1991), confirmée depuis par des études plus récentes (Lucio *et al.*, 2000) (Figure 7). Les mâles atteignent leur maturité à des tailles inférieures à celles des femelles (respectivement 39 cm et 47 cm).

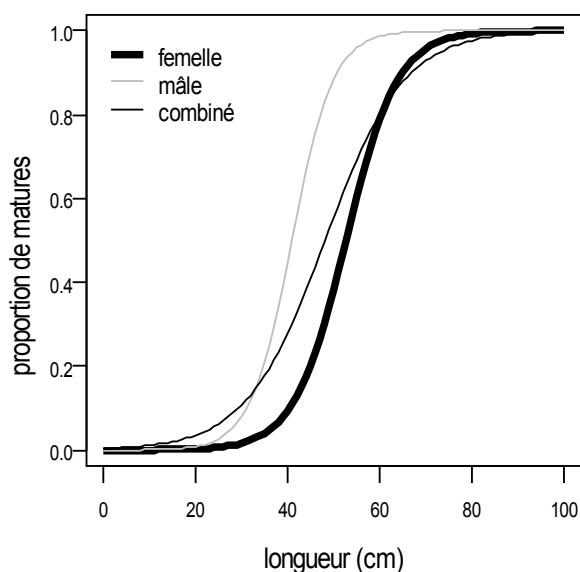


Figure 7 : Ogive de maturité en fonction de la longueur pour les femelles (en gras), les mâles (en gris) et sexes combinés (noir fin) (d'après Lucio *et al.* (2000)).

La ponte du merlu est séquentielle, Murua et Lucio (1998) estimant qu'elle a lieu en une dizaine de lots. La fécondité est par contre assez mal connue (ICES, 2004b). On ne sait notamment pas si elle est déterminée ou non (c'est à dire fixée avant la saison de ponte ou non) et si elle est constante selon les années (Murua et Lucio, 1998). La fécondité annuelle doit donc être estimée par le produit du nombre d'ovocytes émis par ponte, du pourcentage de femelles pondant par jour et de la durée de la ponte. Murua et Lucio (1998) estiment qu'une femelle mesurant 60cm doit pondre entre 1 000 000 et 3 800 000 oeufs.

1.3 Exploitation et gestion du stock Nord de merlu

1.3.1 Exploitation du merlu: évolutions des captures et des métiers

La pêche du stock Nord de Merlu se localise essentiellement dans le golfe de Gascogne et en mer Celtique (Figure 8). Toutefois des captures (au chalut et à la palangre) ont également lieu plus au nord (unité de pêche 16) (elles représentaient 11,4% des débarquements totaux en poids en 2005 (ICES, 2006b)). Le merlu est capturé principalement dans le cadre de pêcheries mixtes, comme espèce cible ou accessoire par des engins variés (chalut de fonds, chalut à perche, filet, palangres) (Casey et Pereiro J, 1995 ; ICES, 2006b). Actuellement, l'Espagne (58% des débarquements totaux en poids en 2005) et la France (29%) représentent l'essentiel des captures. Le Royaume-Uni, l'Irlande et le Danemark se partagent 11% des captures totales, le reste se distribuant entre Norvège, Belgique, Hollande, Allemagne et Suède (ICES, 2006b).

1.3.1.1 Évolution des débarquements (Figure 8)

La pêche du merlu est très ancienne puisqu'elle est signalée dès le XV^e siècle. Hickling (1935) décrit la pêche au XIX^e siècle autour des îles britanniques. Jusqu'en 1860, la pêche est dominée par des petits ligneurs côtiers. A partir du milieu du XIX^e siècle, le chalut à perche a connu un essor rapide, jusqu'à devenir l'engin le plus important, cela se traduisant par une baisse du stock (Hickling, 1935). Avant guerre, la France débarque entre 2 000 et 2 400t, quantité faible comparée aux 33 000t débarquées par les anglais à Milfordhaven (95% des débarquements anglais) (Guichet, 1996). Aucune donnée n'est disponible sur la pêche espagnole à cette époque bien qu'elle ait été probablement importante. Après guerre, le principal pays producteur est l'Espagne (63 000t en 1946, stock Nord et Sud confondus) devant l'Angleterre. Une hausse importante de l'effort de pêche au niveau européen provoque une forte augmentation de la production, qui s'élève à 95 600t (stock Nord) en 1961 (ICES, 1991b). Dans le même temps les débarquements anglais s'effondraient (8 300t en 1961), la France devenant le second pays producteur (43 000 t). En 1965, les débarquements chutent brutalement à 60 000t et restent à peu près stables jusqu'à 1990. Ils chutent de nouveau au cours des années 1990 jusqu'à atteindre 35 000t en 1998 Cette chute a été particulièrement

sensible dans le golfe de Gascogne, les débarquements passant de 28 100t à 9 200t de 1995 à 2001 (ICES, 2006b). Depuis les débarquements restent à peu près stable (41 781t en 2003).

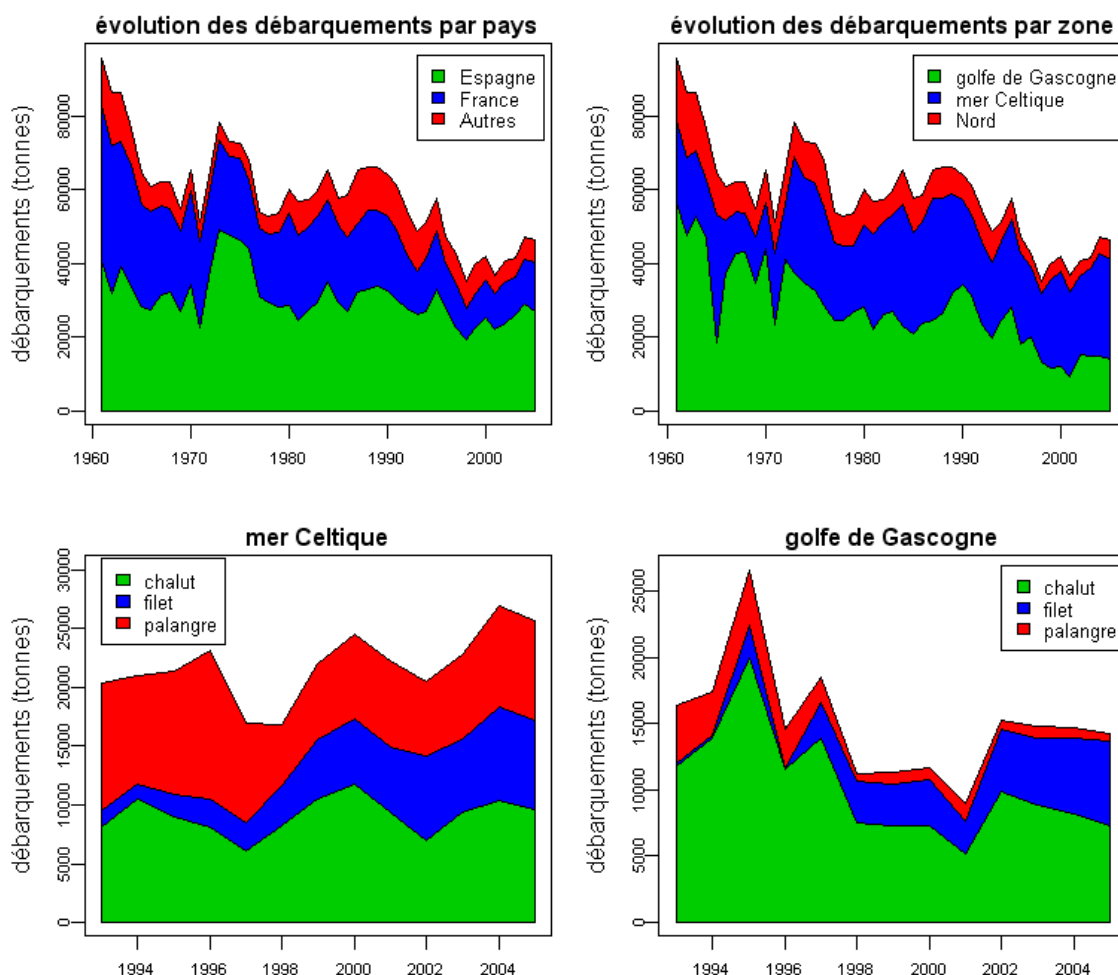


Figure 8 : Evolution des débarquements de merlu (stock Nord) depuis le début des années 1960 par pays (en haut à gauche), par zone (en haut à droite), par engins en mer Celtique (en bas à gauche) et par engin dans le golfe de Gascogne (en bas à droite) (sources : ICES (1991b) et ICES (2007)).

1.3.1.2 Description par pays

a) Espagne

Historiquement, la flottille espagnole est restée longtemps mal connue du fait d'un manque de statistiques fiables. Un règlement européen (CEE 341/78 du 20/02/1978) a entraîné une baisse importante de l'effort nominal de la flottille espagnole. Ainsi de 1973 à 1985, on a assisté à une forte diminution de l'activité des « bous » (bateaux utilisant un chalut adapté aux fonds durs), « trios », « bakas » (ou « bacas », bateaux utilisant un chalut adapté aux fonds vaseux) et « parejas » et ceci quelle que soit la zone de pêche. Une partie de ces navires se sont reconvertis à la palangre (par exemple, 41 chalutiers en 1981) alors que d'autres passaient sous pavillon britannique (par exemple 27 chalutiers + 16 palangriers en 1990) (Guichet, 1996). En 2001, des bateaux espagnols étaient recensés dans les unités de pêche 01, 03, 04, 12, 13, 14, 16, le merlu constituant la cible principale ou une des espèces cibles de pêcheries mixtes (avec notamment baudroie, cardine, langoustine). Une des principales évolutions récentes a été la création de bases avancées en Irlande, au sud de l'Angleterre et en Ecosse qui ont permis de réduire le temps de voyage pour les navires pêchant dans les sous-divisions VI et VII (Figure 2). L'Espagne a été marquée par une profonde restructuration de sa flottille chalutière au cours des années 1990, avec une réduction du nombre de chalutiers individuels, en faveur du chalutage en bœuf. Plus récemment encore, on a assisté à l'introduction du naberan, et du chalut jumeau. Si les palangriers continuent à cibler le merlu en sous-divisions VI et VII (Figure 2), en revanche le nombre de palangriers s'est réduit dans le golfe de Gascogne. Le filet s'est développé à partir de 1998 et représentait dès 1999 12% des captures espagnoles en golfe de Gascogne et 29% en mer Celtique (Commission of the European Communities, 2001).

Les débarquements proviennent aujourd'hui majoritairement des chalutiers devant les palangriers et les fileyeurs qui capturent en revanche de plus grands individus (essentiellement des matures). Dans les sous-divisions VI et VII (Figure 2), les débarquements sont dominés par les palangriers et les chalutiers individuels à panneaux ciblant le merlu (respectivement 6 385t et 2 743t en 2002), alors que dans le golfe de

Gascogne, ils sont dominés par les chalutiers en bœuf (7 014t en 2002) et les fileyeurs (2 068t en 2002), les chalutiers individuels n'y ciblant pas le merlu (Biseau, 2003).

b) France

On a observé en France une augmentation de l'effort de pêche des chalutiers industriels de 1961 à 1966. Après un pic en 1966, l'effort baisse dans le golfe de Gascogne (50% de 1961 à 1975 puis 40% de 1977 à 1985). On observe en parallèle à partir de 1975 une hausse de l'effort des chalutiers industriels en mer Celtique, et un développement du chalutage côtier par de petites unités inférieures à 10m (Guichet, 1996). De plus à partir des années 1960 s'est développée la flottille langoustinière du sud Bretagne. D'un point de vue technique, les années 1960 sont marquées par l'apparition du chalutage par l'arrière. L'utilisation du filet maillant a commencé à partir de 1967, le chalut pélagique apparaissant en 1978 (apogée en 1981) (Guichet, 1996). Les années 1990 ont été marquées par le développement du chalut jumeau, en particulier pour les flottilles langoustinières, qui s'est généralisé depuis sur cette flottille.

A la fin des années 1990, les captures françaises de merlu sont essentiellement dues aux chaluts de fond (47% des débarquements en 1990, 46% en 1998) pour qui le merlu ne constitue plus qu'une prise accessoire, alors que les débarquements des pélagiques ont chuté. En parallèle, la part des fileyeurs n'a cessé d'augmenter au cours des années 1990 (Commission of the European Communities, 2001).

A partir de 2000, le débarquement en poids des fileyeurs a dépassé celui des chalutiers de fond, que ce soit en mer Celtique ou dans le golfe de Gascogne. La flottille langoustinière du golfe de Gascogne est supposée avoir un impact important, car bien que les débarquements soient plus limités, ceux-ci sont constitués d'une part importante de juvéniles (Commission of the European Communities, 2001 ; ICES, 2004b).

La vente du merlu dans les criées a rapporté 19,7 millions d'euros en 2006, en faisant la huitième espèce en terme de valeur (la quatrième espèce si on ne considère que les poissons).

c) Royaume Uni

Le déclin des rendements de merlu après-guerre a provoqué la réduction de la flottille britannique et la reconversion vers d'autres espèces. Ainsi de 1961 à 1979, on a enregistré une baisse de 90% de l'activité des chalutiers dans les divisions IV et VIa (Guichet, 1996) (Figure 2). Actuellement, l'essentiel des débarquements du Royaume-Uni provient des unités de pêche 03 (47% en 2005) et 04 (45% en 2005) (ICES, 2006b).

1.3.1.3 Description des unités de pêche

Seize unités de pêche (« Fishery Units » en anglais) (Tableau 2) sont traditionnellement distinguées au cours de l'évaluation des stocks (ICES, 2004b). Historiquement, ces unités de pêche ont été définies en 1985 par le groupe de travail CIEM sur les sous-divisions VII et VIII (Figure 2) et ont été utilisées avec peu de modification depuis. Une unité de pêche est le croisement d'un engin, d'une zone de pêche et d'espèces cibles, elle s'apparente donc à un métier.

Tableau 2 : Description des unités de pêche définies par le CIEM (ICES, 2005b).

Nom de l'unité de pêche	Description	Sous-divisions
UP01	Palangre de moyenne à forte profondeur	VII
UP02	Palangre à faible profondeur	VII
UP03	Filet	VII
UP04	Chalut non langoustinier de moyenne à forte profondeur	VII
UP05	Chalut non langoustinier à faible profondeur	VII
UP06	Chalut à perche à faible profondeur	VII
UP08 (+UP07)	Chalut langoustinier de moyenne à forte profondeur	VII
UP09	Chalut langoustinier de faible à moyenne profondeur	VIII
UP10	Chalut de faible à moyenne profondeur	VIII
UP12	Palangre de moyenne à forte profondeur	VIII
UP13	Filet de faible à moyenne profondeur	VIII
UP14	Chalut de moyenne à forte profondeur	VIII
UP15	Divers	VII & VIII
UP16	Nord	IIIa, IV, V & VI
UP00	Français	Non connue

En mer Celtique, trois métiers représentent l'essentiel des captures (ICES, 2006b) (Figure 9):

- Les chalutages benthiques (UP04) auxquelles participent notamment des navires espagnols (73%), du Royaume-Uni (19%) et de la France (8%) ont pris de plus en plus d'importance au cours des années 1990 jusqu'à devenir le métier le plus important en capture (Commission of the European Communities, 2001). Le chalut jumeau a été introduit dès la fin des années 1980, début des années 1990 par les bateaux français (Commission of the European Communities, 2001).
- La palangre de moyenne à forte profondeur (UP01), principalement pratiquée par des espagnols ciblant le merlu a perdu de l'importance au cours des années 1990 et ne représentait plus que 16% des captures totales en 1999. De plus, une diminution des captures d'individus de grande taille a été observée (Commission of the European

Communities, 2001). Les Britanniques pratiquaient traditionnellement ce métier, mais l'ont délaissé petit à petit à partir de 1987 (ICES, 1991a).

- Le filet (UP03) était jusqu'à peu essentiellement pratiqué par le Royaume-Uni. Aux cours des années 1990, l'Espagne et la France ont également commencé à pratiquer ce métier, ce qui a expliqué une augmentation des captures, qui atteignaient 10,6% en 1999. La taille moyenne des captures a baissé au cours des années 1990. Si les Espagnols ciblent essentiellement, le merlu, les Britanniques pratiquent une stratégie plus polyvalente, le merlu restant tout de même une des prises principales (Commission of the European Communities, 2001).

Dans le golfe de Gascogne, cinq métiers se partagent l'essentiel des captures (Commission of the European Communities, 2001) (Figure 9):

- Les chalutages sur des fonds moyens à profonds (UP14) sont essentiellement pratiqués par des navires espagnols mettant en œuvre différents engins. Ils ciblent le merlu, la cardine franche et la baudroie. Traditionnellement, ce métier était pratiqué par des navires pêchant seuls (« bacas » et « bous ») ou en paires (« parejas »). Pendant les années 1990, le chalutage en paire n'a cessé d'augmenter au détriment des bacas et surtout des bous (qui ont quasiment disparu). De plus, on observe que le naberan est de plus en plus utilisé par les chalutiers en bœuf (Commission of the European Communities, 2001). Des bateaux français pratiquaient également ce métier, mais le nombre a baissé à la fin des années 1980 (ICES, 1991a).
- Le filet (UP13) était principalement pratiqué par des navires français au début des années 1990. Cependant dès 1999, les Espagnols représentaient une part équivalente des captures (10,6% au total). Le merlu est l'espèce cible de ce métier.

- Le merlu constitue une importante capture accessoire pour les langoustiniers français (UP09). La zone de pêche étant localisée sur la Grande Vasière (Figure 2), les langoustiniers capturent une part importante de merluchons hors-taille (80% étaient inférieurs à 30 cm en 1999) (Biseau, 2003). Anciennement estimé par la campagne FR-RESSGASCS, les rejets de cette unité de pêcherie sont étudiés depuis 2002 dans le cadre du programme OBSMER. Les rejets estimés en 2003, 2004 et 2005 étaient respectivement de 1 034t, 1 360t et 1 596t (ICES, 2007).
- La palangre sur des fonds moyens à forts (UP12) est majoritairement pratiquée par des espagnols ciblant le merlu. Ce métier qui était particulièrement important dans les années 1980 et jusqu'au début des années 1990, a baissé sensiblement à partir de 1994. En parallèle, la taille moyenne des captures a sensiblement augmenté.
- Le chalutage sur des fonds faibles à moyens (UP10) est surtout pratiqué par des navires français. Il a perdu beaucoup de son importance dans les années 1990. En parallèle, la taille moyenne des captures a baissé.

En 2004, les UP04, UP01, UP03, UP13 et UP14 représentaient à elles seules près de 80% des débarquements (Figure 9). Les UP09 et UP10 ne représentaient plus qu'environ 6% des débarquements chacun, mais des rejets importants sont supposés (respectivement 2 989 000 et 2 250 000 d'individus inférieurs à 32 cm pour un débarquement total dans des FU du golfe de Gascogne de 35 119 000 individus) (ICES, 2004b).

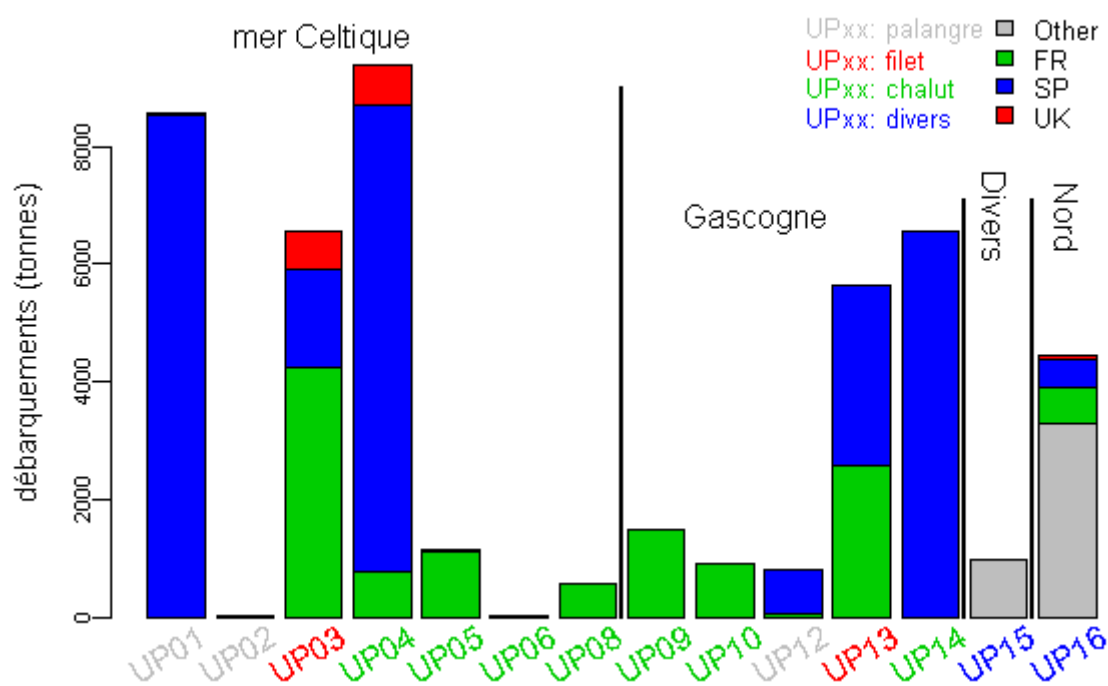


Figure 9 : Débarquements par engin, zone (Figure 2) et unités de pêche (Tableau 2) en 2004 (ICES, 2005b). Divers correspond à des débarquements mer Celtique ou golfe de Gascogne non réalloués à des unités de pêche.

1.3.2 Évaluation des stocks

L'évaluation du stock nord de merlu se fait annuellement en utilisant le modèle XSA (Shepherd, 1999). Les débarquements totaux en poids par unités de pêche et leur composition en longueur permettent de reconstituer la matrice des captures aux âges. Celle-ci fait appel à des clés tailles-âges provenant de lectures d'otolithes réalisées en France et en Espagne. Il est à souligner que l'évaluation ne prend pas en compte les rejets, au vu de l'incertitude et des manques dans la série historique. Quatre campagnes scientifiques fournissent des séries d'indices d'abondance:

- FR-EVHOE à partir de 1997 pour les âges 0 à 5
- FR-RESSGASCS de 1987 à 2002 pour les âges 0 à 5

- UK-WCGFS de 1998 à 2004 pour les âges 1 et 2
- SP-PGFS à partir de 2001 pour les âges 2 à 7

Enfin quatre séries de CPUE provenant de flottilles professionnelles (dans la sous-division VII (Figure 2) du CIEM: les chalutiers de la Corogne et les chalutiers de Vigo ; dans la sous-division VIII (Figure 2) : les chalutiers en boeuf d'Ondarroa et les chalutiers en boeuf de Pasajes) sont également utilisées dans le calibrage du modèle.

Depuis 2004, trois évaluations sont présentées: (i) une basée sur une clé taille-âge classique et utilisant un taper (le taper permet d'attribuer un poids plus important aux données récentes lors de l'ajustement), (ii) une seconde utilisant la même clé taille âge mais sans taper, les biomasses estimées en début de sortie étant largement inférieures, (iii) une basée sur une clé taille-âge simulée (Bertignac et de Pontual, 2007) liée à la remise en question du taux de croissance du merlu (de Pontual *et al.*, 2003 ; de Pontual *et al.*, 2006). Actuellement, seule la première évaluation est utilisée par l'ACFM (« Advisory Committee on Fishery Management » du CIEM).

Outre la conversion taille-âge des données, d'autres incertitudes entourent ces diagnostics. Tout d'abord, par manque de données, les rejets ne sont pas pris en compte au cours de l'évaluation, or ceux-ci sont vraisemblablement assez élevés pour certaines unités de pêche, en particulier l'unité de pêche 09 (Tableau 2). De plus, le modèle XSA suppose que les captures par année et groupe d'âge sont connues exactement, cette hypothèse étant sans doute très forte. Enfin, la non prise en compte des hétérogénéités spatiales, tant de la population que de l'activité de pêche, peut également affecter la qualité de ces diagnostics.

1.3.3 Réglementation de la pêche

De part son importance historique, le merlu a longtemps été au centre des débats, en particulier dans le golfe de Gascogne. Les premières mesures ont consisté en la mise en vigueur de mesures techniques (maillage, taille minimale de captures), auxquelles se sont ajoutés des totaux admissibles de captures (TAC). Des mesures spatialisées ont été également mises en place. Enfin depuis le début des années 2000, le stock Nord de merlu fait l'objet d'un plan de restauration.

1.3.3.1 Mesures techniques: tailles minimales et maillages

En 1983, lors de la promulgation de la politique commune des pêches, le règlement (CEE) n°171/83 met en place une série de mesures techniques, reprenant essentiellement des recommandations déjà existantes de la NEAFC (« North East Atlantic Fisheries Commission » fondée par la convention de Londres). Le maillage minimal est alors fixé à 65mm pour tout type de filets (sauf dérogations pour certaines espèces), la taille minimale s'élevant à 30cm. Après plusieurs modifications, le règlement n° 850/1998 du Conseil (lui-même modifié ultérieurement) définit des maillages minimaux variables selon les régions et les espèces cibles. Il n'y a donc plus un maillage mais une fourchette de maillages. De plus la distinction est faite entre arts dormants et arts traînants. Ainsi en mer Celtique le maillage est fixé à 80mm (120mm pour les fileyeurs) pour les chalutiers visant le merlu et d'autres espèces de poissons (parmi lesquelles baudroie, sole, plie, lieu jaune, rougets...), mais elle est de 70mm pour les pêcheurs visant la langoustine. Le maillage des chalutiers du golfe est de 70mm (100mm pour les fileyeurs) mais est de 55mm si la part de langoustine est importante. D'autre part, la taille minimale est ramenée à 27cm (sauf dans la région Skagerrak/Kattegat où elle est maintenue à 30cm).

Il est à souligner que pendant très longtemps, ces mesures n'ont été que peu (ou pas) respectées, et qu'elles constituaient encore en 2003 la principale source d'infraction à la politique commune des pêches (Commission of the European Communities, 2003).

1.3.3.2 *Les TAC (Totaux Admissibles de Captures)*

Les premiers TAC européens ont été recommandés par la NEAFC en 1974. La CEE met officiellement le système des TAC en place en 1978, définitivement formalisé lors de la promulgation de la politique commune des pêches (PCP) en 1983 (notamment le principe de stabilité relative). Là encore, le respect du TAC a été pour le moins relatif.

1.3.3.3 *Mesures spatialisées*

Historiquement, un cantonnement visant à protéger les juvéniles de merlu dans le golfe de Gascogne a été mis en place sur recommandation de la NEAFC en 1970. Toutefois, mal dimensionné (trop petit) et positionné (situé trop au large, il ne recouvrait presque pas la zone de nourricerie), il s'est montré inefficace et a été aboli en 1973. Plus récemment, la Commission Européenne a proposé deux cantonnements sur les nourriceries de la Grande Vasière (Figure 2) et de la mer Celtique en 1996 et 1997, aucune des deux n'a par contre été mise en oeuvre.

1.3.3.4 *Le plan de restauration*

Au début des années 2000, une mortalité par pêche forte, une baisse de la biomasse féconde et une série de bas recrutements (Figure 10) ont justifié la mise en place d'un plan de restauration pour le stock Nord de merlu. Une première série de règlements (les règlements (CE) n°1162/2001, 2602/2001, 2256/2002, 494/2002) a mis en place des mesures techniques visant à protéger les juvéniles de merlu. Deux box à merlu, destinés à limiter la capture de juvéniles de merlu, ont ainsi été créés dans le golfe de Gascogne et en mer Celtique (Figure 11), dans lesquels les maillages des chalutiers passent à 100mm dans le golfe de Gascogne et 120mm en mer Celtique. Ces mesures ajoutent également des contraintes sur les débarquements de grande quantité de merlu et imposent aux états membres d'organiser des observations embarquées. Le plan de restauration (règlement (CE) 811/2004) a suivi ce plan d'urgence. Il fixe comme objectif le retour à la biomasse féconde correspondant à l'approche de précaution (140 000t). Il impose pour cela les règles de fixation du TAC: le TAC ne doit pas correspondre à une mortalité par pêche supérieure à 0,25, ne doit pas entraîner une baisse de la biomasse féconde et ne doit pas varier de plus de 15% par rapport au TAC de l'année

précédente. De plus, si la biomasse féconde descend en dessous des 100 000t, les règles précédentes ne s'appliquent plus, la Commission fixant alors un TAC à la majorité qualifiée.

Les objectifs ayant été atteints, ce plan devrait être prochainement remplacé par un plan de gestion.

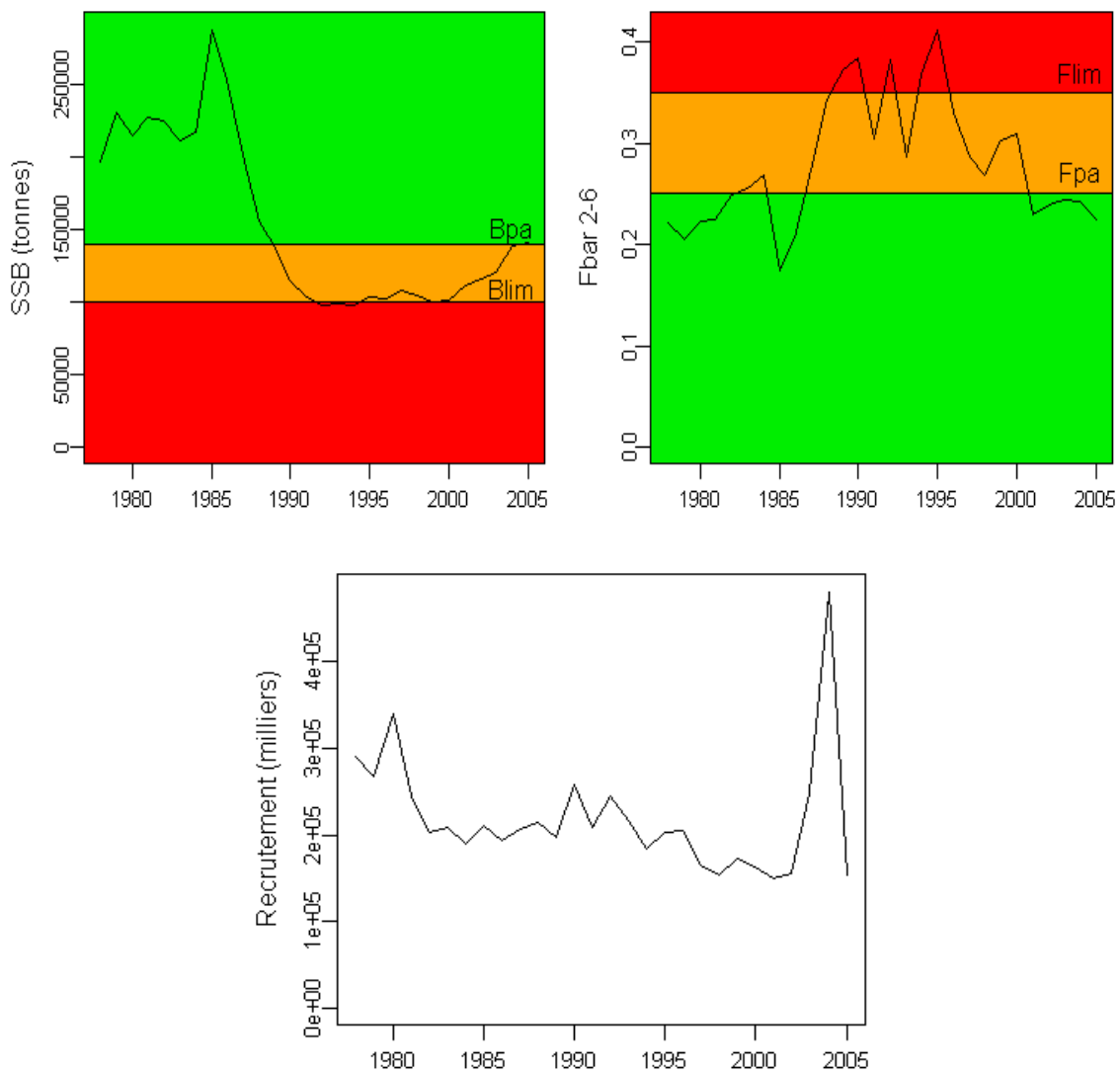


Figure 10 : Evolution des biomasses fécondes (en haut à gauche), mortalité par pêche (en haut à droite) et des recrutements estimés par XSA (ICES, 2006b) pour le stock Nord de merlu. Les zones en rouge correspondent à des biomasses fécondes (à gauche) ou à des mortalités par pêche (à droite) non durables. Les zones en orange sont des zones en dehors des limites correspondant à l'approche de précaution.

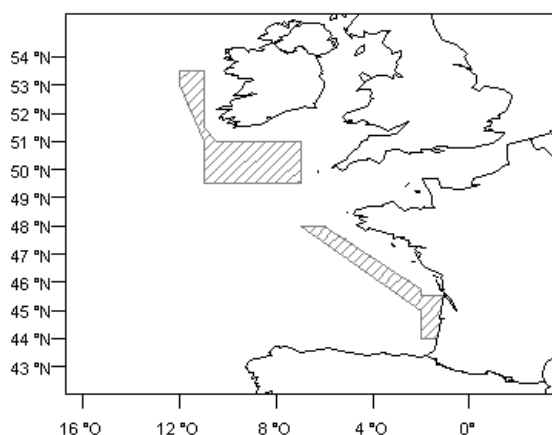


Figure 11 : Localisation des deux box à merlus (en gris) définis dans le règlement (CE) n°1162/2001.

1.4 Conclusion

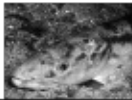
Malgré la littérature abondante existante sur la biologie du merlu, certains éléments de son cycle biologique, en particulier la croissance et les migrations, restent mal connus. La croissance est pourtant particulièrement importante pour prévoir le temps nécessaire à la restauration du stock face à l'introduction de nouvelles mesures de gestion (Bertignac et de Pontual, 2007). Une meilleure connaissance des migrations est indispensable à la mise en place de mesures de gestion spatialisées comme les aires marines protégées. A ce titre, une étude de marquages recaptures à grande échelle s'avèrerait une source d'informations précieuses pour améliorer la connaissance sur la structuration du stock, sa croissance et ses migrations, en particulier si comme lors des deux dernières campagnes des marques archives (celles-ci enregistrent toutes les 15 min les conditions de pression et de température rencontrées par l'animal marqué ; le trajet parcouru peut alors être reconstitué intégralement) peuvent être utilisées.

Si l'état du stock était alarmant au début des années 2000, les dernières tendances montrent que le celui-ci est en train de se reconstituer et que les objectifs fixés par le plan de restauration sont en passe d'être atteints. Cependant, les modèles utilisés pour réaliser ces diagnostics restent critiquables car prenant mal en compte certaines incertitudes, en particulier l'incertitude liée à la conversion taille-âge nécessaire à la reconstitution de la matrice des captures aux âges.

Ce qu'il faut retenir

- le merlu est une population ayant un poids socio-économique historiquement important
- la population située dans la partie Nord de l'Atlantique Nord-Est s'est trouvée dans une situation alarmante au début du siècle ayant justifiée la mise en place d'un plan de restauration
- certains processus biologiques restent mal connus : croissance et migration
- l'évaluation du stock est sujette à de nombreuses incertitudes pouvant remettre en cause le diagnostic

Chapitre 2 : Éléments sur les modèles de dynamique des populations structurés en longueur



Chapitre 2 : Éléments sur les modèles de dynamique des populations structurés en longueur

2.1 Introduction

Les modèles les plus classiquement utilisés en halieutique, en particulier pour l'évaluation de stocks, sont des modèles structurés en âge. Ces modèles permettent de fournir un diagnostic sur les tendances de la mortalité par pêche et de la biomasse. Cependant, comme nous l'avons mentionné au chapitre précédent, ils font actuellement l'objet de certaines critiques qui ont justifié le développement récent d'un type alternatif de modèle, les modèles structurés en longueur.

Une première critique adressée aux modèles structurés en âge est la conversion des données observées par classe de taille en données par classe d'âge, format requis pour leur utilisation (en particulier pour reconstituer la matrice des captures aux âges, et des indices d'abondance par âge, utilisés dans la plupart des modèles, voir (Megrey, 1989)). Pour de nombreuses espèces cette conversion est particulièrement incertaine et entraîne une incertitude dans les diagnostics d'évaluation des stocks. Pour le merlu, la conversion taille-âge est réalisée à partir de clés taille-âge, qui sont calculées à partir d'un échantillon de poissons mesurés et âgés par lecture d'otolithes. Or comme on l'a vu au chapitre précédent, la lecture d'otolithes fait l'objet de fortes controverses et par conséquent la clé taille-âge est elle aussi jugée peu fiable. Un autre aspect, non négligeable, est le coût nécessaire pour faire cette conversion : la construction de clé taille-âge nécessite la lecture d'un nombre élevé d'otolithes et a donc un coût important. Dans ce contexte, le groupe de travail CIEM sur les méthodes d'évaluation conclut que les modèles structurés en âge n'étaient pas nécessairement les plus adaptés pour l'évaluation lorsque l'incertitude sur la détermination de l'âge est forte (ICES, 2003a). Les modèles structurés en longueur (ou en âge et longueur) utilisant directement des données basées sur la longueur et donc permettant de s'affranchir de cette conversion,

constituent une alternative intéressante.

Un autre avantage que présentent les modèles structurés en longueur est la possibilité de pouvoir décrire, sans transformation, les nombreux processus biologiques et d'exploitation directement liés à la longueur des poissons. Cet aspect est d'autant plus important que la variabilité interannuelle ou interindividuelle de croissance est forte (ICES, 2003d).

Ainsi, parmi les processus d'exploitation, la sélectivité (probabilité qu'un poisson soit capturé par un engin de caractéristiques données) est directement liée à sa taille. Le plus souvent la sélectivité est décrite par une fonction mathématique (sigmoïde, gamma ou double gaussienne) de la taille du poisson et des caractéristiques techniques de l'engin (par exemple la maille pour un chalut ou un filet).

Parmi les processus biologiques, la prédation (dont le cannibalisme) est bien évidemment directement guidée par la taille du prédateur et de la proie. De même, la maturation d'un individu semble être autant liée à la taille de l'individu qu'à son âge (Heino *et al.*, 2002a ; Heino *et al.*, 2002b ; Grift *et al.*, 2003).

Les modèles structurés en longueur présentent donc de nombreux avantages par rapport aux modèles structurés en âge, ce qui explique leur développement récent (Schnute, 1987 ; Schnute *et al.*, 1989a ; Schnute *et al.*, 1989b ; Sullivan *et al.*, 1990 ; Banks *et al.*, 1991 ; De Leo et Gatto, 1995 ; Fournier *et al.*, 1998 ; Fu et T J Quin II, 2000 ; DeLong *et al.*, 2001 ; Froya *et al.*, 2002 ; Breen *et al.*, 2003 ; Zhang *et al.*, 2004). Cette tendance s'est également traduite par la mise en place en 2004 au sein du CIEM d'un groupe de travail sur les modèles structurés en longueur et en âge « Study Group on Age-length Structured Assessments Models » (ICES, 2003d ; ICES, 2005a ; ICES, 2006a).

2.2 Objectifs et constitution des modèles structurés en longueur

Les modèles structurés en longueur ont connu un fort développement dans les années récentes avec des objectifs, des niveaux de complexité et des données nécessaires à leur utilisation relativement variables. On peut classer les objectifs de modélisation en quatre catégories principales: (1) réaliser des évaluations de stocks (2) estimer des paramètres inconnus de la dynamique de population (3) évaluer l'impact de mesures de gestion et (4) la modélisation écosystémique avec deux sous-objectifs possibles: l'établissement de diagnostics ou le développement d'un modèle opératoire:

1. Les modèles structurés en longueur peuvent être développés afin de réaliser des évaluations de stocks (Sullivan *et al.*, 1990 ; Fournier *et al.*, 1998 ; Hampton et Fournier, 2001 ; Froysa *et al.*, 2002 ; Breen *et al.*, 2003 ; Maunder et Watters, 2003 ; Kristensen *et al.*, 2006). Le modèle de dynamique sous-jacent est en général assez simple, et généralement, seuls le recrutement, la croissance et la mortalité naturelle (mortalité autre que la mortalité par pêche) sont décrits. Selon les espèces, d'autres processus biologiques peuvent-être pris en compte, notamment les migrations (Fournier *et al.*, 1998 ; Froysa *et al.*, 2002) et cannibalisme (Froysa *et al.*, 2002). Le niveau de complexité du modèle de pêche est, quant à lui, assez variable, du simple modèle séparable (avec cette hypothèse, la mortalité par pêche est supposée être le produit d'un facteur dépendant de l'année et d'un facteur dépendant de la classe de longueur) de Kristensen *et al.* (2006), au modèle multiflottilles (pour lesquels la mortalité par pêche est décomposée en une somme de mortalités par pêches générées par les différentes flottilles du modèle) de Froysa *et al.* (2002). Contrairement aux autres modèles discutés dans ce paragraphe, la population dans le modèle de Kristensen *et al.* (2006) n'est pas découpée en classes (ni d'âge, ni de longueur), la taille et le temps étant des variables continues. Il présente en plus la particularité de ne s'ajuster que sur des données de campagnes scientifiques. Multifan-CL (Fournier *et al.*, 1998), Fleksibest (Froysa *et al.*, 2002) et A-Scala (Maunder et Watters, 2003) outre leur objectif d'évaluation de stocks, permettent l'estimation de nombreux

paramètres inconnus de la dynamique des populations dont la croissance et dans une certaine mesure les migrations (Multifan-CI propose une version spatialisée optionnelle permettant dans une certaine mesure l'estimation de migration. Ces trois modèles présentent la caractéristique de s'ajuster sur l'ensemble des données disponibles (on parle de modèles intégrés).

2. Beaucoup de modèles structurés en longueur sont développés afin d'estimer certains paramètres inconnus de la dynamique de population (Zhang *et al.*, 2004), notamment ceux liés à la croissance (Banks *et al.*, 1991 ; DeLong *et al.*, 2001) et à la survie (Fu et T J Quin II, 2000 ; DeLong *et al.*, 2001). Dans ce type de modèles, le sous-modèle d'activité de pêche est en général simplifié (ou inexistant), tout comme les processus biologiques dont les paramètres ne sont pas à estimer.
3. Un autre objectif rencontré est l'évaluation de l'impact de mesures de gestion (Pet *et al.*, 1996 ; Pelletier *et al.*, 2001 ; Kvamme et Froysa, 2004 ; Mahévas et Pelletier, 2004 ; Pelletier *et al.*, submitted). C'est sans doute le type de modèles le plus complexe puisqu'il doit à la fois coupler un modèle de dynamique de population et un modèle d'activité de pêche assez détaillé. Le paramétrage de ces modèles est donc généralement complexe et il est en général plus facile de ne pas ajuster directement le modèle aux données. Une grande partie des paramètres provient en général de la littérature. Pour les autres paramètres, des analyses statistiques sur les données existantes, préalables à l'utilisation du modèle, peuvent permettre de les estimer indépendamment les uns des autres. Dans le cas contraire, on peut procéder à des simulations sous différents scénarios (hypothèses) de valeurs de paramètres.
4. Enfin, les modèles structurés en longueur étant particulièrement adaptés à la modélisation des processus de prédation, ils sont souvent utilisés dans le cadre de modèles écosystémiques. On peut ainsi citer le modèle de Hall *et al.* (2006) qui visait à évaluer la réponse d'une communauté de poissons à la pression de pêche. Le modèle GADGET, structuré en longueur ou en âge-longueur, est, quant à lui, une plate-forme

générique de modélisation écosystémique.

2.3 Modélisation des processus

Les modèles structurés en longueur permettent de modéliser de nombreux processus biologiques ou d'exploitation directement liés à la longueur, le choix des processus à modéliser étant guidé par les objectifs du modèle. Nous détaillons ci-dessous les principaux choix de modèles pour décrire des processus.

2.3.1 Maturité-fécondité

Modéliser la maturation dans des modèles a trois intérêts principaux: (i) cela permet de calculer la biomasse féconde qui est souvent utilisée comme indicateur de l'état du stock, (ii) c'est indispensable quand le modèle inclut une relation stock recrutement, (iii) ou quand certains processus biologiques (mortalité naturelle, migrations...) diffèrent entre matures et immatures (De Leo et Gatto, 1995). Plusieurs méthodes peuvent-être utilisées pour décrire la maturation:

- la méthode la plus simple est d'utiliser une ogive de maturité fonction de la longueur. Dans cette approche, on ne distingue pas à proprement parlé un stock de matures d'un stock d'immatures.
- une autre méthode consiste à modéliser la maturation comme un processus stochastique, chaque classe de longueur étant alors caractérisée par une probabilité d'acquérir la maturité. Dans cette approche, le stock doit être divisé entre matures et immatures. Dans les modèles structurés en âge-longueur, un âge minimum de maturation peut également être utilisé (Froysa *et al.*, 2002)

- Il est également possible de complexifier le modèle de maturation, le modèle GADGET propose par exemple de faire dépendre le processus de maturation du facteur de condition (Begley et Howell, 2004 ; ICES, 2005a).

A notre connaissance, le seul modèle structuré en longueur intégrant un modèle de fécondité est une application sur la morue de l'Arctique Nord-Est (Bogstadt *et al.*, 2004).

2.3.2 Mortalité naturelle – prédation

En halieutique, on définit comme mortalité naturelle, toute source de mortalité autre que la mortalité liée à la pêche. Dans la plupart des modèles, la mortalité naturelle est supposée connue et constante et n'est donc pas à proprement modéliser. Cependant, dans le cadre des modèles écosystémiques où l'on s'intéresse aux relations trophiques (Begley et Howell, 2004 ; Hall *et al.*, 2006), ou dans le cas d'espèce où le cannibalisme est fort (Froysa *et al.*, 2002), des modèles de prédation/cannibalisme peuvent-être intégrés. Dans Fleksibest (Froysa *et al.*, 2002), le cannibalisme est modélisé comme une fonction de la taille des proies, de la biomasse de prédateurs potentiels et de la biomasse de proies alternatives.

GADGET étant, comme souligné précédemment, une plate-forme de modélisation écosystémique, il permet de modéliser des relations trophiques. Pour chaque relation proie prédateur, une fonction de « préférence » dépendant de la taille du prédateur et de la taille de la proie permet de quantifier l'intensité avec laquelle le prédateur va agir sur la proie. En multipliant cette fonction par la biomasse de proie, on calcule un poids de proie disponible pour le prédateur (Begley et Howell, 2004).

2.3.3 Croissance

La croissance est un des processus majeurs de la biologie des poissons (Laslett *et al.*, 2002) et constitue un des processus clés des modèles structurés en longueur (Chen *et al.*, 2003). Un « bon » modèle doit conjointement décrire la croissance moyenne et la variabilité interindividuelle de croissance (Sainsbury, 1980 ; Laslett *et al.*, 2002 ; Chen *et al.*, 2003). Il existe classiquement trois méthodes pour modéliser la croissance et la variabilité individuelle

de croissance:

- Une première méthode, assez simple, repose sur l'hypothèse que la taille des individus au sein d'un groupe d'âge suit une distribution (en général normale) autour d'une moyenne donnée par une courbe de croissance (en général von Bertalanffy) (Fournier *et al.*, 1998 ; Maunder et Watters, 2003). Le modèle peut-être légèrement complexifié afin de modéliser une croissance densité-dépendante en liant les paramètres de la fonction de von Bertalanffy à la biomasse de la population (Fournier *et al.*, 1998). Ce type d'approche requiert soit d'avoir une connaissance de la croissance a priori, soit d'avoir au moins une partie des données structurées en âge. La taille des individus étant uniquement liée à leur âge, ce type de modèle ne permet pas de mettre en évidence des effets de sélection de la pêche (la pêche ayant tendance à enlever les individus grandissant le plus vite).
- Une seconde approche, que nous appellerons modèle d'incrément de croissance, consiste à supposer que l'incrément de taille d'un individu au cours d'un pas de temps suit une distribution stochastique (par exemple une loi gamma) de moyenne donnée par une fonction de croissance (souvent de von Bertalanffy) (Sullivan *et al.*, 1990 ; De Leo et Gatto, 1995 ; Cruywagen, 1997). Ces hypothèses permettent de calculer la probabilité de passer d'une classe à une autre au cours d'un pas de temps, l'ensemble de ces probabilités constituant la matrice de transition. On trouve de légères variations autour de cette approche, comme dans Fleksibest (Froysa *et al.*, 2002) où ce n'est pas directement l'incrément de croissance qui est une variable aléatoire mais le nombre de classes de taille franchies par un individu pendant un pas de temps. Le modèle d'incrément de croissance est probablement le plus utilisé dans les modèles matriciels structurés en longueur. Toutefois ce type de modèle ne permet pas de prendre en compte d'éventuelles différences génétiques individuelles de croissance. Une description plus détaillée de ce type de modèle de croissance sera faite dans le chapitre 3.

- La troisième méthode consiste à incorporer la variabilité inter-individuelle de croissance en faisant l'hypothèse que les paramètres de la courbe de croissance (en général von Bertalanffy) suivent une distribution statistique particulière (Sainsbury, 1980 ; Smith *et al.*, 1998 ; Smith et Botsford, 1998 ; Pilling *et al.*, 2002). Cette méthode est plus adaptée à la modélisation individu centrée, et est souvent utilisée pour l'estimation de croissance à partir de données de marquage-recapture (Laslett *et al.*, 2002 ; Eveson *et al.*, 2004).

2.4 Complexité des modèles

Les modèles structurés permettant de modéliser directement de nombreux processus liés à la longueur. En contrepartie, ce sont souvent des modèles plus complexes reposant sur de nombreux paramètres (ICES, 2003d). A titre d'exemple, le modèle Multifan-CI (Fournier *et al.*, 1998) utilisé dans le cadre de l'évaluation du stock d'albacore du Pacifique Sud (*Thunnus alalunga*) nécessite l'estimation de 876 paramètres dans sa formulation la plus simple.

Le niveau de complexité du modèle est le résultat d'un compromis entre ce que l'on souhaite étudier et ce que les données disponibles permettent d'étudier. On utilise souvent en modélisation la notion de parcimonie : un modèle doit avoir le niveau de complexité suffisant pour correctement représenter les données, mais pas plus (Box et Jenkins, 1970). De façon schématique, les modèles trop simples tendent à fournir des estimations biaisées ; la complexification des modèles tend au contraire à augmenter la variance d'estimation (Costanza et Sklar, 1985 ; Håkanson, 1995 ; Burnham et Anderson, 2002). L'idée que complexifier les modèles les rendent plus réalistes est assez communément admises. Il n'en reste pas moins qu'augmenter la complexité, accroît les difficultés d'interprétation du modèle. La comparaison avec un modèle plus simple peut dès lors permettre de plus facilement interpréter les grandes tendances et de tester l'intérêt des hypothèses d'un modèle plus complexe (Richards et Schnute, 1998).

D'un point de vue méthodologique, les performances de modèles s'ajustant par maximum de vraisemblance peuvent être étudiées suivant plusieurs critères (Burnham et Anderson, 2002). Dans ce travail, nous avons considéré les trois critères suivants:

- le rapport du maximum de vraisemblance qui permet de comparer des modèles emboîtés. L'avantage de cette méthode est qu'il existe une statistique (la différence de logvraisemblance suivant une loi du χ^2) permettant d'accepter ou de rejeter le test qu'un modèle complexe explique mieux le jeu de données qu'un modèle plus simple emboîté. Ce test a par exemple été utilisé par Kristensen *et al.* (2006) pour comparer différentes formulations de son modèle.
- le critère d'information d'Akaike (Akaike, 1973): $AIC = 2 \cdot p - 2 \cdot \log L$, avec p le nombre de paramètres à estimer et $\log L$ la logvraisemblance. La différence de logvraisemblance entre 2 modèles est donc pondérée par le nombre de paramètres. Il ne s'agit pas ici d'une statistique mais plutôt d'une mesure de distance entre le modèle considéré et le modèle « vrai », c'est-à-dire celui ayant généré les observations (Burnham et Anderson, 2002). Ce critère ne permet de comparer des modèles que si ces modèles s'ajustent sur les mêmes données.
- le critère d'information Bayésien $BIC = p \cdot \ln(n) - 2 \cdot \log L$ (Schwarz, 1978) avec p le nombre de paramètres à estimer, $\log L$ la logvraisemblance et n le nombre d'observations. La différence de logvraisemblance entre 2 modèles est donc pondérée par le nombre de paramètres et par le nombre d'observations. Le BIC est particulièrement adapté aux situations avec un nombre d'observations élevé et tend à être plus parcimonieux que l'AIC. Contrairement à l'AIC, le BIC n'est pas basé sur une mesure de distance par rapport au modèle « vrai » mais sur une estimation de la dimension (nombre de paramètres) du modèle « vrai ».

2.5 Les types de données utilisées pour ajuster les modèles structurés en longueur

Que ce soit dans le cadre des modèles développés pour estimer des paramètres inconnus, ou pour les modèles d'évaluation de stocks, des données sont nécessaires afin d'ajuster le modèle. Le type de données nécessaires dépend principalement de l'objectif du modèle et des processus qu'il inclut. Nous proposons de les classer en quatre grandes catégories.

2.5.1 Données commerciales

- Captures

Les captures commerciales sont fréquemment utilisées dans les ajustements des modèles en longueur (Schnute, 1987 ; Fournier *et al.*, 1998 ; Froysa *et al.*, 2002 ; Breen *et al.*, 2003 ; Maunder et Watters, 2003). On rencontre deux approches pour l'utilisation des captures commerciales:

1. une première approche consiste à utiliser les captures commerciales par classe de longueur (en poids ou en nombre) (Sullivan *et al.*, 1990 ; Froysa *et al.*, 2002). Selon les modèles, elle peut être encore désagrégée, comme par exemple dans Fleksibest (Froysa *et al.*, 2002) où elles sont désagrégées par pas de temps, par flottille, et par classe d'âge x longueur. Dans cette approche, on fait alors l'hypothèse que la capture observée est une variable aléatoire dont la moyenne est prédite par le modèle. Plusieurs distributions ont été envisagées (dans Fleksibest, l'utilisateur a le choix d'utiliser selon les jeux de données des distributions multinomiale, lognormale, gamma ou de Pearson) la plus fréquente étant la distribution lognormale.
2. Une seconde approche consiste à utiliser la capture totale (toutes classes confondues) en poids ou en nombre, et la composition en taille de ces captures, c'est à dire la fréquence d'individus observée dans chacune des classes de taille (Fournier *et al.*, 1998 ; Breen *et al.*, 2003 ; Maunder et Watters, 2003). Cette méthode présente l'avantage de bien faire la distinction entre deux types d'observation, la capture totale

provenant en général du flux déclaratif et la composition en longueur provenant des échantillonnages en criée. Dans cette approche, les captures totales en poids sont traitées de la même façon que les captures par classe de longueur dans le cas précédent. Pour la composition en taille des captures (Figure 12), la fréquence observée d'individus dans une classe de longueur est supposée suivre une distribution normale (sous réserve que l'échantillon soit assez grand) dont la moyenne est fournie par le modèle et la variance est une fonction de la taille de l'échantillon prélevé. Les fréquences observées dans les différentes classes de longueur sont supposées indépendantes, ce qui n'est pas le cas puisque les fréquences somment à 1. Fournier *et al.* (1990) justifie cette approximation par une meilleure robustesse : en modifiant légèrement la logvraisemblance de la loi normale, l'ajustement est moins sensible aux fortes et/ou faibles fréquences prédites par le modèle. On peut également modéliser le nombre d'individus dans chacune des classes de l'échantillon comme le résultat d'un tirage multinomial parmi la capture totale.

- Effort

Si des données d'effort sont disponibles, pour au moins une partie de la flottille commerciale, les captures par unité d'effort peuvent-être utilisées comme indices d'abondance. Elles sont alors traitées de la même manière que les indices d'abondance provenant de campagnes scientifiques (Froysa *et al.*, 2002 ; Breen *et al.*, 2003 ; Punt, 2003).

Elles peuvent également être utilisés dans le calcul de la mortalité par pêche (Fournier *et al.*, 1998), permettant de décrire la distribution spatiale et saisonnière de l'activité de pêche.

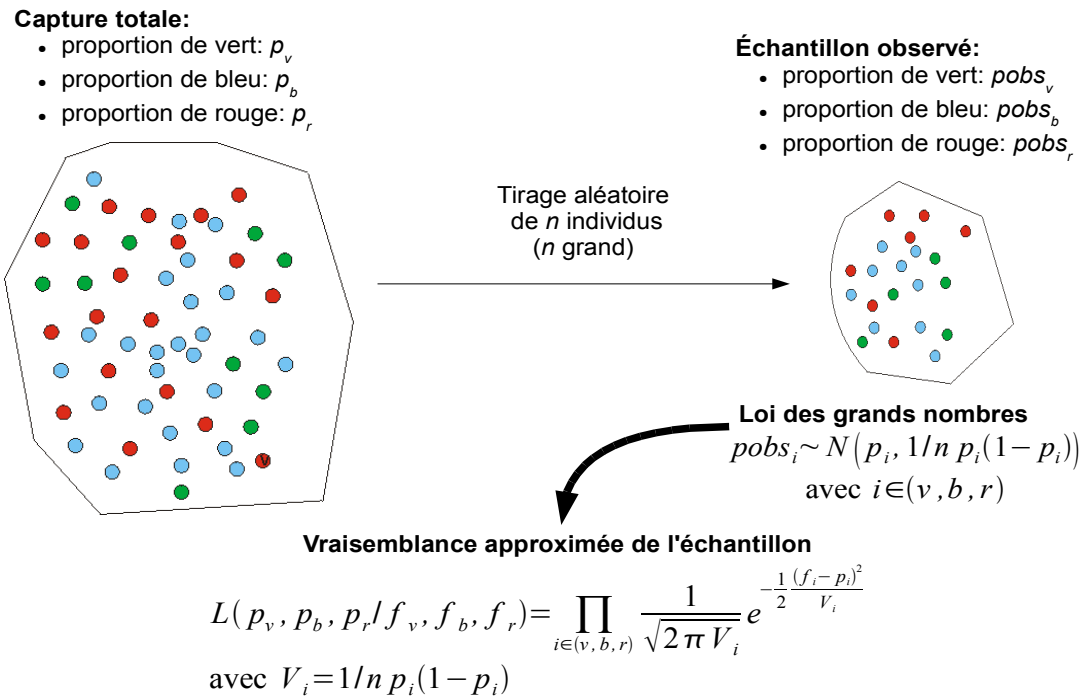


Figure 12 : Schéma expliquant les hypothèses faites sur les compositions en longueur obtenues à partir d'échantillons de captures commerciales. Les couleurs représentent ici des classes d'âge. L'indépendance n'est en théorie pas vérifiée puisque les proportions somment à 1.

2.5.2 Données de campagne

La plupart des modèles utilisent également des indices d'abondance provenant de campagnes scientifiques (Schnute, 1987 ; Fu et T J Quin II, 2000 ; Froysa *et al.*, 2002 ; Breen *et al.*, 2003 ; Kristensen *et al.*, 2006). De la même façon que pour les captures commerciales, il existe deux approches principales pour utiliser ces données:

1. supposer que les indices d'abondance scientifiques observés par classe de longueur suivent une distribution statistique (le plus souvent lognormale) autour d'une moyenne prédite par le modèle (Froysa *et al.*, 2002 ; Breen *et al.*, 2003).
2. comme pour les captures commerciales, les indices d'abondance peuvent être décomposés en un indice d'abondance global (toutes classes confondues), et une distribution probabiliste de la composition de l'indice par classe de longueur (Fu et T J Quin II, 2000 ; DeLong *et al.*, 2001). DeLong *et al.* (2001) supposent que la composition en taille des indices d'abondance suit une distribution multinomiale (voir la discussion dans le paragraphe 2.5.1). La justification de cette subdivision entre un indice global et une distribution probabiliste de la composition de l'indice par classe de longueur est moins évidente que dans le cas des captures commerciales car les deux types de données proviennent ici du même processus d'observation, l'indépendance des observations n'est donc pas assurée.

2.5.3 Données de marquage

Les données de marquage étant intéressantes pour estimer les paramètres de croissance, elles sont parfois utilisées dans les modèles structurés en longueur (Breen *et al.*, 2003 ; Punt, 2003). Elles sont en général utilisés dans un sous-modèle individu-centré, en temps continu, distinct du modèle structuré en longueur (Breen *et al.*, 2003 ; Punt, 2003), Punt (2003) séparant même complètement la phase d'ajustement de la matrice de transition à l'aide des données de marquage, de l'ajustement du reste du modèle. L'hypothèse la plus fréquente est de supposer que les incréments observés suivent une distribution stochastique de moyenne donnée par une courbe de croissance (généralement von Bertalanffy).

2.5.4 Données de contenus stomacaux

Pour les modèles intégrant des relations trophiques, des données de contenus stomacaux peuvent être utilisées. Par exemple dans Fleksibest (Froysa *et al.*, 2002) la consommation par prédateur et semestre pour chaque groupe de prédateur et chaque groupe de proie est comparée à des observations faites à partir de contenus stomacaux.

2.6 Ajustement des modèles aux données observées

Nous avons donc vu que de nombreuses données peuvent être utilisées pour ajuster le modèle. Voyons comment sont estimés les paramètres inconnus (capturabilité, taux de croissance, mortalité...) à partir de ces observations.

2.6.1 Méthodes d'ajustement

Classiquement, la plupart des modèles structurés en longueur s'ajustent par maximum de vraisemblance (Fournier *et al.*, 1998 ; DeLong *et al.*, 2001 ; Froysa *et al.*, 2002 ; Kristensen *et al.*, 2006), c'est à dire par la recherche d'un jeu de paramètres θ maximisant la vraisemblance. La vraisemblance statistique d'un modèle (ou d'un jeu de paramètres) vis-à-vis d'observations peut se définir comme la probabilité d'obtenir ces observations lorsque le jeu de paramètres θ modèle est supposé vrai (c'est la probabilité conjointe des observations sachant que le jeu de paramètres θ est vrai). Cette méthode fournit théoriquement une estimation conjointe, asymptotiquement sans biais et convergente de θ . D'autres méthodes proches sont parfois utilisées, Fu et T J Quin II (2000) par exemple utilisent des moindres carrés non linéaires, cette méthode apparaissant plus robuste aux erreurs dans les données et plus rapide. On trouve également dans la littérature des ajustements séquentiels, par exemple Punt (2003) sépare les phases d'estimation de la matrice de croissance et du modèle proprement dit. Le modèle de Breen *et al.* (2003) est à notre connaissance le seul modèle bayésien structuré en longueur.

Les méthodes basées sur le maximum de vraisemblance nécessitent l'écriture de la fonction de vraisemblance. Celle-ci découle directement des hypothèses statistiques faites sur les observations (voir paragraphe 2.5). Quelques modifications sont parfois réalisées sur la fonction de vraisemblance:

1. Ajout de pénalités: Afin d'intégrer des connaissances sur certains processus de la dynamique de la population ou de l'activité de pêche, des pénalités sont parfois ajoutées à la fonction de vraisemblance (Fournier *et al.*, 1998 ; DeLong *et al.*, 2001 ; Maunder et Watters, 2003). Ces pénalités tiennent un rôle comparable aux distributions a priori (« priors ») des modèles bayésiens. C'est également une méthode pour prendre en compte l'incertitude sur la dynamique des états du système, en forçant par exemple le recrutement à suivre une distribution lognormale ou la capturabilité à suivre une marche aléatoire. Ce type d'ajustement est parfois appelé « likelihood errors in variables ». Il faut noter que théoriquement, la fonction objectif obtenue n'est plus une fonction de vraisemblance (de Valpine, 2002 ; de Valpine, 2003 ; de Valpine, 2004 ; de Valpine et Hilborn, 2005) et que par conséquent, les tests basés sur la vraisemblance (ratio de vraisemblance, AIC, BIC) ne peuvent être utilisés.
2. Ajout de pondération: La plupart des modèles (Fournier *et al.*, 1998 ; Fu et T J Quin II, 2000 ; Froysa *et al.*, 2002 ; Punt, 2003) utilisent une pondération des jeux de données. Cette pondération est relativement arbitraire et fait appel à la connaissance des experts. Elle traduit normalement la variance (ou coefficient de variation) des observations. Cette méthode étant subjective, une analyse de sensibilité aux poids utilisés est ensuite indispensable.

2.6.2 Choix de l'algorithme d'optimisation

La fonction de vraisemblance étant généralement complexe et reposant sur un plus ou moins grand nombre de paramètres, son optimisation ne peut se faire de façon analytique et une méthode numérique est indispensable. Celle-ci doit permettre d'explorer le plus largement possible le domaine de recherche afin d'éviter des optima locaux, tout en convergeant le plus rapidement possible vers la solution optimale. L'algorithme utilisé n'est malheureusement pas toujours donné dans les articles. Multifan-CL repose sur un algorithme de type quasi-Newton adapté aux problèmes de grande dimension. Fleksibest laisse le choix entre:

- un algorithme de type quasi-Newton

- l'algorithme de Hooke et Jeeves
- un algorithme de type « simulated annealing » (Kirkpatrick *et al.*, 1983), algorithme stochastique d'optimisation basé sur un algorithme de Metropolis Hasting, et adapté au problème de grande dimension. A chaque itération, une nouvelle solution est testée, et a une probabilité d'être acceptée même si elle n'améliore pas la vraisemblance, ceci permettant une large exploration du domaine de recherche. Le domaine de recherche est restreint au fur et à mesure des itérations quand la vitesse de progression de l'algorithme diminue.

Quand le modèle est différentiable, des algorithmes de différentiation automatique peuvent-être utilisés afin d'obtenir rapidement la valeur exacte des dérivées par rapport à chacun des paramètres (Fournier *et al.*, 1998 ; Maunder et Watters, 2003 ; Kristensen *et al.*, 2006).

Une présentation de différents algorithmes d'optimisation est faite en annexe 1.

2.7 Utilisation de modèles structurés en longueur ou en âge et longueur

Bien que les modèles structurés en longueur aient connu un développement rapide ces dernières années, l'utilisation en routine de ce type de modèle n'est pas encore très fréquente.

2.7.1 Hors Europe

Des modèles structurés en longueur et/ou en âge/longueur ont été utilisés pour diagnostiquer l'état de stocks d'espèces difficiles à âger, en particulier de thonidés (Fournier *et al.*, 1998 ; Hampton et Fournier, 2001 ; Maunder et Watters, 2003) et d'invertébrés (Zheng *et al.*, 1995 ; Lai et Bradbury, 1998).

2.7.2 En Europe

L'utilisation de modèles structurés est moins fréquente en Europe. Toutefois, de plus en plus de modèles sont présentés au cours des groupes de travail CIEM, et au sein de projets européens. Le modèle *Fleksibest* (Froysa *et al.*, 2002) a été utilisé pour l'évaluation du stock de Morue de l'Arctique Nord Est en 2004 (ICES, 2004a).

Au sein du programme européen BECAUSE (<http://www1.uni-hamburg.de/BECAUSE/>), plusieurs modèles ont été développés à partir de GADGET. Dans la mer de Barents, un modèle multispécifique incluant morue, capelan, hareng et baleine de Minke a été développé. D'autres modèles multispécifiques utilisant GADGET ont été réalisés dans les eaux islandaises et sont en cours de développement pour les eaux ibériques, méditerranéennes et du golfe de Gascogne.

Le projet européen UNCOVER (<http://www.uncover.eu/>) va réutiliser les modèles GADGET développés au cours de BECAUSE comme modèle opérationnel en vue de tester l'efficacité de modèles d'évaluation de stocks et l'impact de scénarios de gestion.

2.8 Conclusion

Les modèles structurés en longueur constituent une alternative intéressante aux modèles structurés en âge particulièrement pour les espèces, comme le merlu, dont la croissance est mal connue. Ils sont potentiellement plus réalistes en intégrant différents sous-modèles de processus, mais en contrepartie, ils sont souvent relativement complexes. Le choix du modèle doit être un compromis entre les objectifs du modèle, qui peuvent-être assez variés, et les données disponibles.

Cette revue bibliographique des modèles structurés en longueur existants (Tableau 3) a permis de confirmer que les modèles structurés en longueur pouvaient être particulièrement adaptés à la dynamique du merlu. Les modèles de type Fleksibest, A-Scala et Multifan-CL développés pour réaliser des évaluations de stocks, tout en estimant des paramètres de la biologie de l'espèce (croissance et éventuellement migration) en s'ajustant sur l'ensemble des données disponibles, présentent des objectifs relativement proches des notre et pourront donc nous servir pour le développement de notre modèle. Toutefois, contrairement à ces trois modèles, une structuration de la population purement en longueur, plutôt qu'en âge et longueur, paraît plus intéressante du fait de l'absence de données fiables sur l'âge du merlu. Ce constat implique notamment que la première approche de modélisation de la croissance ne peut être appliquée (voir section 2.1), seuls les modèles d'incrémentés étant alors utilisables.

D'autre part, il est souvent fait mentionné dans la littérature ayant trait aux modèles structurés en longueur que la modélisation du processus de croissance jouait un rôle clé (voir section 2.3.3). Dès lors, une étude préliminaire visant à proposer des hypothèses robustes pour le modèle de croissance nous est apparue indispensable.

Tableau 3 : Résumé des caractéristiques principales des différents modèles structurés en longueur.

Modèle	Objectif	Processus modélisés	Données utilisées	Méthode d'ajustement
Sullivan <i>et al.</i> (1990)	Évaluation de stocks	Recrutement, croissance, sélectivité de la pêche	Captures commerciales	Maximum de vraisemblance avec filtre de Kalman
Banks <i>et al.</i> (1991)	Estimation de la croissance et de la survie (mortalité naturelle)	Croissance, recrutement, survie	Indices d'abondance scientifiques	Moindres carrés
De Leo et Gatto (1995)	Estimation de la mortalité et du taux de métamorphose	Recrutement, métamorphose, croissance, mortalité naturelle âge-dépendante, maturation, sélectivité de la pêche	Composition en longueur des captures commerciales	
Fournier <i>et al.</i> (1998)	Évaluation de stocks	Recrutement, migration, croissance, sélectivité de la pêche, dynamique de l'effort de pêche	Captures commerciales, indices d'abondance scientifiques	Maximum de vraisemblance avec ajout de pénalités
Fu et T J Quin II (2000)	Estimation de paramètres (croissance, sélectivité des engins de pêche, recrutement, mortalité naturelle, mortalité par pêche, taille au recrutement)	Croissance, variation inter-annuelle de la mortalité naturelle, recrutement, sélectivité de la pêche	Indices d'abondances scientifique, captures commerciales	Moindres carrés non linéaires

Tableau 3 : (suite).

Modèle	Objectif	Processus modélisés	Données utilisées	Méthode d'ajustement
DeLong <i>et al.</i> (2001)	Estimation de la croissance et de la mortalité naturelle	Croissance, modèle de mortalité naturelle	Indices d'abondance scientifiques, captures commerciales	Maximum de vraisemblance avec ajout de pénalités
Froysa <i>et al.</i> (2002)	Évaluation de stocks	Recrutement, croissance, maturation, (migration), cannibalisme, sélectivité de la pêche	Captures commerciales, indices d'abondance scientifiques, contenus stomacaux	Maximum de vraisemblance
Breen <i>et al.</i> (2003)	Évaluation de stocks	Croissance, recrutement, sélectivité de la pêche	Captures commerciales, indices d'abondance scientifiques, marquage-recaptures	Modèle bayésien
Maunder et Watters (2003)	Évaluation de stocks	Recrutement, migration, croissance, sélectivité de la pêche, dynamique de l'effort de pêche	Captures commerciales, indices d'abondance scientifiques	Maximum de vraisemblance avec ajout de pénalités
Punt (2003)	Évaluation de stocks	Croissance, recrutement, sélectivité de la pêche, rejets par la pêche	Marquage recapture, captures commerciales	Maximum de vraisemblance séquentiel avec pénalité

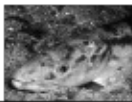
Tableau 3 : (suite).

Modèle	Objectif	Processus modélisés	Données utilisées	Méthode d'ajustement
Mahévas et Pelletier (2004)	Évaluer l'impact de mesures de gestion	Recrutement, croissance, capturabilité, mortalité naturelle, sélectivité de la pêche, dynamique de l'effort de pêche, dynamique de la gestion		Le modèle ne s'ajuste pas
Gadget	Plateforme de modélisation écosystémique	Variable selon l'application	Variable selon l'application	Maximum de vraisemblance/év entuellement extérieur au modèle
Zhang <i>et al.</i> (2004)	Estimation de la croissance et de la mortalité	Croissance (par mue), mortalité naturelle et par pêche	Indices d'abondance scientifiques	Maximum de vraisemblance
Kristensen <i>et al.</i> (2006)	Évaluation de stocks	Recrutement, croissance, sélectivité des engins, dynamique de la capturabilité	Indices d'abondance scientifiques	Maximum de vraisemblance
Hall <i>et al.</i> (2006)	Modélisation écosystémique/ évaluation d'impact	Croissance, recrutement, prédation, sélectivité de la pêche		Extérieur au modèle

Ce qu'il faut retenir

- les modèles structurés en longueur sont adaptés aux espèces à croissance mal connue comme le merlu
- ce sont des modèles flexibles mais potentiellement complexes
- ils peuvent avoir des objectifs variés : estimation de paramètres, évaluation de stocks, simulations, modélisation écosystémique...
- la modélisation du processus de croissance est un élément clé des modèles structurés en longueur

**Chapitre 3 : Quantification de l'impact
d'hypothèses de discrétisation sur un modèle de
croissance structuré en longueur**



Chapitre 3 : Quantification de l'impact d'hypothèses de discrétisation sur un modèle de croissance structuré en longueur

3.1 Introduction

Dans les modèles de dynamique des populations, le temps peut-être considéré comme continu ou discret. Les modèles à temps continu sont ceux qui représentent en théorie le mieux la réalité. Cependant ces modèles reposent sur des équations différentielles dont la résolution fait en général appel à des algorithmes numériques lourds. De plus, les données dont on dispose en halieutique sont la plupart du temps agrégées dans le temps et donc peu favorables à ce type d'approche. A l'opposé, les modèles à temps discret reposent sur des équations simples facilement programmables. Ce sont les modèles les plus fréquemment utilisés. Toutefois, les modèles à temps discrets nécessitent une hypothèse forte : au cours d'un pas de temps, les processus sont considérés homogènes. Cette hypothèse peut-être la source d'artefact de modélisation. De même, la structuration de la population en classes de longueur discrètes est également une hypothèse forte affectant la longueur des individus et leur croissance dans le modèle. Cette discrétisation requiert des hypothèses sur la taille des classes de longueur et sur la distribution de la taille des individus au sein de ces classes ainsi qu'un certain nombre d'hypothèses additionnelles sur le modèle de croissance.

Comme nous l'avons vu dans le chapitre précédent, la matrice de transition décrivant le processus de croissance, est un des éléments clés des modèles structurés en longueur. Elle doit modéliser à la fois la croissance moyenne et la variabilité interindividuelle de cette croissance (Laslett *et al.*, 2002 ; Chen *et al.*, 2003). Comme nous l'avons vu précédemment, plusieurs approches peuvent être envisagées pour construire la matrice de transition. Nous orientant vers un modèle de dynamique des populations plutôt que vers un modèle individu centré, le

Ce travail a fait l'objet d'une publication dans le journal Fisheries Research:
Drouineau, H., Mahévas, S. et Bertignac, M. in Press. Assessing the impact of discretisation assumptions in a length-structured population growth model, Fish. Res., doi:10.1016/j.fishres.2007.11.017

choix d'un modèle type « modèle d'incrément de croissance » apparaît le plus pertinent (voir section 2.3.3). Peu d'études ont quantifié l'impact de la discrétisation du temps, de la longueur et des hypothèses classiquement dans les modèles d'incrément de croissance. Nous avons choisi de développer une méthodologie permettant de quantifier la sensibilité du modèle à ces choix pour déterminer des hypothèses robustes sur le modèle de croissance. Les choix considérés pour le modèle de croissance de population structuré en longueur concernent (i) le pas de temps (ii) la taille des classes de longueur (iii) la distribution des incréments de croissance et (iv) la distribution des individus au sein des classes de longueur. Des jeux de trajectoires individuelles de croissance ont été simulés en fonction des différentes hypothèses (taux de croissance moyen des individus, un paramètre mesurant la variabilité interindividuelle de croissance, la distribution des incréments de croissance). Chaque structure (une structure correspond à une combinaison des quatre hypothèses) de modèle de croissance de population a ensuite été ajustée aux différents jeux de trajectoires individuelles de croissance simulées. Une analyse statistique de deux critères quantifiant la capacité du modèle à s'ajuster aux données individuelles a permis de mettre en évidence les hypothèses rendant le modèle de croissance plus flexible.

3.2 Matériel et méthodes

3.2.1 Un modèle de croissance de population structuré en longueur

Sous forme matricielle, un modèle structuré en longueur d'une population uniquement soumise à la croissance peut s'écrire sous la forme:

$$N(t+1) = G \cdot N(t) \quad (1)$$

avec $N(t)$: un vecteur contenant l'abondance dans chacune des classes de longueur au temps t .

On suppose ici que la croissance et donc la matrice G est constante au cours du temps (absence de phénomènes saisonniers et densité dépendants dans la croissance). La croissance moyenne est supposée suivre une courbe de von Bertalanffy, définie par trois paramètres L_∞, K, t_0 . La taille moyenne au temps t est alors donnée par la relation:

$$l(t) = L_\infty \times (1 - e^{-K(t-t_0)}) \quad (2)$$

Dans notre cas, on utilise une reformulation de cette relation qui calcule le gain de taille Δl (appelé "incrément de croissance") pendant un intervalle Δt pour un individu de taille initiale l :

$$\Delta l = (L_\infty - l) \times (1 - e^{-K \times \Delta t}) \quad (\text{Fabens, 1965}) \quad (3)$$

Pour modéliser la variabilité interindividuelle de croissance, on fait l'hypothèse que les incréments de croissance au cours d'un pas de temps sont des variables aléatoires (Sullivan *et al.*, 1990 ; Cruywagen, 1997 ; Froysa *et al.*, 2002) dont les moyennes sont calculées à partir de la relation de von Bertalanffy, et dont les variances sont des fonctions de la moyenne. Dans la littérature, on trouve deux hypothèses de relation entre variance et espérance des incréments de croissance: (i) la variance est proportionnelle à la moyenne (Sullivan *et al.*, 1990 ; Sullivan, 1992), (ii) la variance est proportionnelle au carré de la moyenne (c'est à dire que le coefficient de variation est constant) (DeLong *et al.*, 2001). Nous ne considérerons ici que la deuxième hypothèse bien que la méthode que nous présentons soit transposable à la première hypothèse.

$$\begin{cases} E(X/l) = (L_\infty - l) \cdot (1 - e^{-K \cdot \Delta t}) \\ V(X/l) = C \cdot E^2(X/l) \end{cases} \quad (4)$$

avec X/l la variable aléatoire modélisant l'incrément de croissance pour des individus de taille initiale l au cours d'un pas de temps Δt . C est appelée constante de variance.

La matrice de croissance G peut alors être calculée à partir de l'une des relations suivantes:

- si l'on suppose qu'au sein d'une classe de longueur i (dont les bornes inférieures et supérieures sont respectivement notées L_i et U_i) tous les individus font la même taille égale à la taille m_i au centre de la classe, la probabilité $g_{i,j}$ d'atteindre la classe j dont les bornes inférieures et supérieures sont respectivement notées L_j et U_j) est donc :

$$g_{ij} = \begin{cases} \int_{L_j - m_i}^{U_j - m_i} f(x) dx & \text{si } j > i \\ \int_0^{U_i - m_i} f(x) dx & \text{si } j = i \\ 0 & \text{si } j < i \end{cases} \quad (5)$$

avec f la fonction de densité des incréments de croissance.

- si l'on suppose que la taille des poissons au sein de la classe i est uniformément distribuée entre les deux bornes, la probabilité $g_{i,j}$ est alors:

$$g_{ij} = \begin{cases} \int_{L_i}^{U_i} \left(\int_{L_j - y}^{U_j - y} f(x) dx \right) dy & \text{si } i < j \\ \int_{L_i}^{U_i} \left(\int_0^{U_j - y} f(x) dx \right) dy & \text{si } i = j \\ 0 & \text{si } i > j \end{cases} \quad (6)$$

avec f la fonction de densité des incréments de croissance.

Dans les deux cas, les probabilités de transition doivent être standardisées afin de sommer à 1 par la simple relation:

$$g'_{ij} = \frac{g_{ij}}{\sum_{c=1}^C g_{ic}} \quad (7)$$

avec C le nombre de classes de longueur.

Plusieurs hypothèses ont été considérées pour décrire la distribution des incréments de croissance: la loi gamma, la loi normale et la loi lognormale (Tableau 4).

Chapitre 3 : Quantification de l'impact d'hypothèses de discrétisation sur un modèle de croissance structuré en longueur

Tableau 4 : Liste des abréviations et de leurs significations. Les paramètres en gras correspondent à des hypothèses sur les trajectoires individuelles de croissance. Les autres correspondent à des hypothèses sur le modèle de croissance de population structuré en longueur.

Abréviation	Signification	Valeurs possibles
Δt	Pas de temps du modèle de croissance de population structuré en longueur	12 : mensuel 4 : trimestriel 1 : annuel
Δl	Largeur des classes de longueur du modèle de croissance de population structuré en longueur	1 : 1cm 5 : 5cm
Lm	Loi de distribution des incréments de croissance du modèle de croissance de population structuré en longueur	0 : gamma 1 : normal 2 : lognormale
$ODep$	Distribution de la taille des individus au sein des classes de longueur du modèle de croissance de population structuré en longueur	0 : milieu 1 : uniforme
C, CV	Constante de variance du jeu de trajectoires de croissances individuelles. CV correspond au coefficient de variation associé à C	0.4 $\Leftrightarrow CV=0.63$ 1 $\Leftrightarrow CV=1$ 4 $\Leftrightarrow CV=2$
K	Taux de croissance moyen de la fonction de von Bertalanffy utilisé pour simuler le jeu de trajectoires de croissances individuelles	0.1, 0.2
Ls	Distribution statistique des incréments de croissance du jeu de trajectoires de croissances individuelles	0 : gamma 1 : normal 2 : lognormale
$Kest$	Taux de croissance estimé lors de l'ajustement du modèle de croissance de population structuré en longueur sur un jeu de trajectoires de croissances individuelles	
$Cest$	Constante de l'équation 4 estimée lors de l'ajustement du modèle de croissance de population structuré en longueur sur un jeu de trajectoires de croissances individuelles	
BIC	Critère d'information bayésien Information lors de l'ajustement du modèle de croissance de population structuré en longueur sur un jeu de trajectoires de croissances individuelles	
$Lbic$	Logarithme népérien de BIC	
LK	Logarithme népérien de la valeur absolue de la différence $Kest - K$	

La loi gamma est une distribution continue entièrement définie par deux paramètres. Sa grande flexibilité fait qu'elle est très couramment utilisée en modélisation. Les lois normale et lognormale ont également été envisagées (Tableau 5).

Tableau 5 : Fonction de densité, moyenne et variance des lois normale (N), lognormale (LN) et gamma (G).

Avec $\Gamma(x) = \int_0^{\infty} e^{-u} u^{x-1} du$ la fonction gamma.

Distrib	Density	E	V
N	$f(x) = \frac{1}{\sigma \times \sqrt{2 \times \pi}} \times e^{-\frac{1}{2} \times \left(\frac{x-\mu}{\sigma}\right)^2}$	$E(X) = \mu$	$V(X) = \sigma^2$
LN	$f(x) = \frac{1}{\sigma \times x \times \sqrt{2 \times \pi}} \times e^{-\frac{1}{2} \times \left(\frac{\log(x)-\mu}{\sigma}\right)^2}$	$E(X) = e^{\mu + \frac{\sigma^2}{2}}$	$V(X) = (e^{\sigma^2} - 1) \times e^{2\mu + \sigma^2}$
G	$f(x) = \frac{1}{\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}}$	$E(X) = \frac{\alpha}{\beta}$	$V(X) = \frac{\alpha}{\beta^2} = \frac{1}{\alpha} E^2(X)$

Plusieurs méthodes peuvent-être envisagées pour construire les classes de longueur: (i) des classes de même largeur, (ii) des classes dont le temps moyen nécessaire pour les traverser est identique (ce type de classe apparaît particulièrement adapté aux modèles de croissance utilisant une loi bêta-binomiale), (iii) des classes liées aux connaissances biologiques, (iv) des classes résultant de l'optimisation d'un critère spécifique. La méthode la plus couramment utilisée est la méthode (i) (Sullivan *et al.*, 1990 ; De Leo et Gatto, 1995 ; Froysa *et al.*, 2002). La troisième méthode est caractéristique des modèles structurés en stade. Vandermeer (1978), puis Moloney (1986), ont proposé un algorithme correspondant à la quatrième méthode. L'algorithme, détaillé ultérieurement par Caswell (2000), visait à minimiser la somme des erreurs d'échantillonnage et des erreurs de distribution. Nous n'avons par contre trouvé aucune application de la deuxième méthode. La première méthode nous est apparue la plus adaptée ; deux tailles de classe de longueur ont donc été envisagées (Tableau 4).

Enfin, trois pas de temps distincts ont été comparés: mensuel, trimestriel et annuel (Tableau 4).

Nous noterons $Mod(Lm, \Delta t, \Delta l, Odep)$ une formulation du modèle de croissance de population structuré en longueur, pour lequel les incréments de croissance sont supposés suivre une distribution Lm , le pas de temps est Δt , les classes de longueur mesurent Δl , et les individus sont distribués selon l'hypothèse $Odep$ dans les classes de longueur.

Les effectifs (10 000 individus) sont projetés à l'aide du modèle sur 10 ans (soit $T= 10 \cdot \Delta t$ le nombre de pas de temps d'une simulation), la taille initiale des individus étant distribuée selon une loi normale de moyenne 5cm et d'écart type 1.

3.2.2 Simulation de jeux de trajectoires de croissance individuelles

Pour quantifier l'impact de la discrétisation du temps et de la structuration en classes de longueur, il est nécessaire de disposer de données de croissance individuelle. Or, peu de données de ce type sont disponibles. Les données de marquage qui auraient pu éventuellement être utilisées pour cette étude sont encore insuffisamment nombreuses. Par conséquent, il a été nécessaire de simuler des jeux de trajectoires individuelles de croissance. Pour s'approcher de croissance en temps continu, un modèle à pas de temps mensuel a été utilisé. Des essais avec des pas de temps plus courts ont été réalisés, les résultats étant très proches ils n'ont pas été utilisés car les temps de simulation devenaient importants.

On note $S(K, C, Ls)$ un jeu de trajectoires individuelles simulées avec un taux de croissance K , une constante de variance C , et Ls la loi de distribution des incréments de croissance (les valeurs possibles pour chacune des hypothèses sont présentées dans le Tableau 4). Chaque jeu S comporte 10 000 trajectoires de croissance simulées sur 10 ans. Les tailles initiales des individus étaient distribuées selon une loi normale de moyenne 5cm et d'écart-type 1.

3.2.3 Plan d'expériences (Kleijnen, 1998 ; Kleijnen, 2001)

L'objectif de l'étude est de quantifier l'impact d'hypothèses, en particulier de discrétisation de la longueur et du temps, sur un modèle de croissance de population structuré en longueur. Pour cela, on cherche à évaluer la capacité de ce type de modèle à s'ajuster à des données de croissances individuelles simulées. Un plan d'expériences complet a été développé pour atteindre cet objectif. Chaque formulation du modèle $Mod(Lm, \Delta t, \Delta l, Odep)$ a été ajustée à chaque jeu de trajectoires de croissance $\mathbf{S}(\mathbf{K}, \mathbf{C}, \mathbf{Ls})$. Un ajustement consiste à trouver les paramètres K_{est} et C_{est} qui minimisent la somme des écarts au carrés:

$$LS(K_{est}, C_{est}) = \sum_{i=1}^I \sum_{t=1}^T (f_{est}(t, i) - f_{obs}(t, i))^2 \quad (8)$$

avec $f_{obs}(t, i)$ la proportion simulée d'individus dans la classe i au temps t , et $f_{est}(t, i)$ la fréquence dans la classe i au temps t estimée par le modèle de croissance de population structuré en longueur.

On a considéré ici que L_{∞} était connu et égal à 110 cm, valeur utilisée pour les estimations de taux de croissance réalisées à partir des données de marquage (de Pontual *et al.*, 2006).

La minimisation de $LS(K_{est}, C_{est})$ se fait par l'intermédiaire d'un algorithme génétique (voir annexe 1) couplé à un algorithme d'optimisation de type quasi-Newton fourni dans la librairie autodif (Otter Research Ltd. <http://otter-rsch.com>). L'algorithme génétique permet dans une première étape de fournir un point de départ, utilisé ensuite pour initier l'algorithme de quasi-Newton.

La capacité du modèle à s'ajuster aux trajectoires individuelles est quantifiée par deux critères:

- l'estimation du taux de croissance est quantifiée par la différence $\mathbf{K} - K_{est}$

- le critère d'information bayésien (*BIC*) est calculé à partir de la formule $BIC = -2 \cdot \text{Log}L + p \cdot \ln(n)$ avec $p=2$ le nombre de paramètres estimés,

$$\text{Log}L = n \cdot \ln(\sqrt{2 \times \pi}) + \frac{1}{2} \cdot \left[n \cdot \left(\ln \frac{n-1}{LS} - 1 \right) + 1 \right]$$

la logvraisemblance de l'ajustement, et n le nombre d'observations (qui dépend de Δt et Δl). Le *BIC* quantifie l'écart entre les fréquences simulées et les fréquences estimées, en les pondérant par le nombre de points d'observations (et le nombre de paramètres estimés qui est constant) de façon à ce que le critère soit comparable quelque soit le pas de temps et la taille des classes.

3.2.4 Analyse du plan d'expériences

Dans un premier temps, des graphiques type « boîtes à moustache » représentant la distribution des estimations en fonction de chacune des hypothèses du modèle de croissance de population structuré en longueur sont tracés. Elles permettent de détecter visuellement d'éventuels biais dans l'estimation du taux de croissance. Dans un second temps, une analyse de chacun des deux critères est réalisée séparément. Les critères sont au préalable transformés afin que les hypothèses du modèle linéaire soient respectées (en particulier la normalité des erreurs et la constance de la variance): on note $LBic = \log(|BIC|)$ (on n'utilise la valeur absolue car les *BIC* sont négatifs) et $LK = \log(|K - K_{est}|)$. Un « bon » modèle est donc un modèle avec une valeur *LK* faible et un *LBic* grand (du fait du changement de signe). Un modèle linéaire est ajusté sur chacun de ces deux critères, la variabilité des critères étant expliquée par les différentes hypothèses du modèle de croissance et du jeu de trajectoires, et par leurs interactions du premier ordre:

$$\begin{aligned} LBic \sim & \mu + Lm + \Delta t + \Delta l + Odep + \mathbf{K} + \mathbf{CV} + \mathbf{Ls} + Lm : \Delta t + Lm : \Delta l + Lm : Odep + \\ & Lm : \mathbf{K} + Lm : \mathbf{CV} + Lm : \mathbf{Ls} + \Delta t : \Delta l + \Delta t : Odep + \Delta t : \mathbf{K} + \Delta t : \mathbf{CV} + \Delta t : \mathbf{Ls} + \\ & \Delta l : Odep + \Delta l : \mathbf{K} + \Delta l : \mathbf{CV} + \Delta l : \mathbf{Ls} + Odep : \mathbf{K} + Odep : \mathbf{CV} + Odep : \mathbf{Ls} + \\ & \mathbf{K} : \mathbf{CV} + \mathbf{K} : \mathbf{Ls} + \mathbf{CV} : \mathbf{Ls} \end{aligned} \quad (9)$$

$$\begin{aligned} LK \sim & \mu + Lm + \Delta t + \Delta l + Odep + \mathbf{K} + \mathbf{CV} + \mathbf{Ls} + Lm : \Delta t + Lm : \Delta l + Lm : Odep + \\ & Lm : \mathbf{K} + Lm : \mathbf{CV} + Lm : \mathbf{Ls} + \Delta t : \Delta l + \Delta t : Odep + \Delta t : \mathbf{K} + \Delta t : \mathbf{CV} + \Delta t : \mathbf{Ls} + \\ & \Delta l : Odep + \Delta l : \mathbf{K} + \Delta l : \mathbf{CV} + \Delta l : \mathbf{Ls} + Odep : \mathbf{K} + Odep : \mathbf{CV} + Odep : \mathbf{Ls} + \\ & \mathbf{K} : \mathbf{CV} + \mathbf{K} : \mathbf{Ls} + \mathbf{CV} : \mathbf{Ls} \end{aligned} \quad (10)$$

On contraint les coefficients de chacun des termes des modèles linéaires à sommer à 0.

Théoriquement un plan d'expériences complet permet d'intégrer toutes les interactions (Kleijnen, 1998 ; Kleijnen, 2001), cependant dans un souci de faciliter l'interprétation, nous avons préféré nous limiter aux interactions d'ordre 1, peu de termes d'ordre supérieur étant significatifs.

Les modèles linéaires (9) et (10) permettent de quantifier l'impact des différentes hypothèses de modélisation sur chacun des deux critères individuellement. Or il est important que le modèle puisse répondre simultanément aux deux objectifs (estimation du taux de croissance et reproduction des observations). Les arbres de régressions multiples (De'Ath, 2002) permettent d'explorer l'impact de variables explicatives sur plusieurs variables quantitatives à expliquer simultanément. L'arbre est construit itérativement en séparant en deux partitions des groupes tels que la distance Euclidienne intergroupe soit maximale et que la distance euclidienne au centroïde intragroupe soit minimale. Chaque groupe est caractérisé par un critère sur une des variables explicatives. L'ordre « d'utilisation » des variables explicatives dans la construction de l'arbre reflète donc leurs impacts relatifs sur les deux variables à expliquer. Un arbre de régression a été construit selon le modèle suivant:

$$\frac{(Lbic - mean(Lbic))}{sd(Lbic)}, \frac{(LK - mean(LK))}{sd(LK)} \sim \Delta t + \Delta l + Lm + Odep + CV + K + bols Ls \quad (11)$$

La figure 13 résume schématiquement la méthode utilisée.

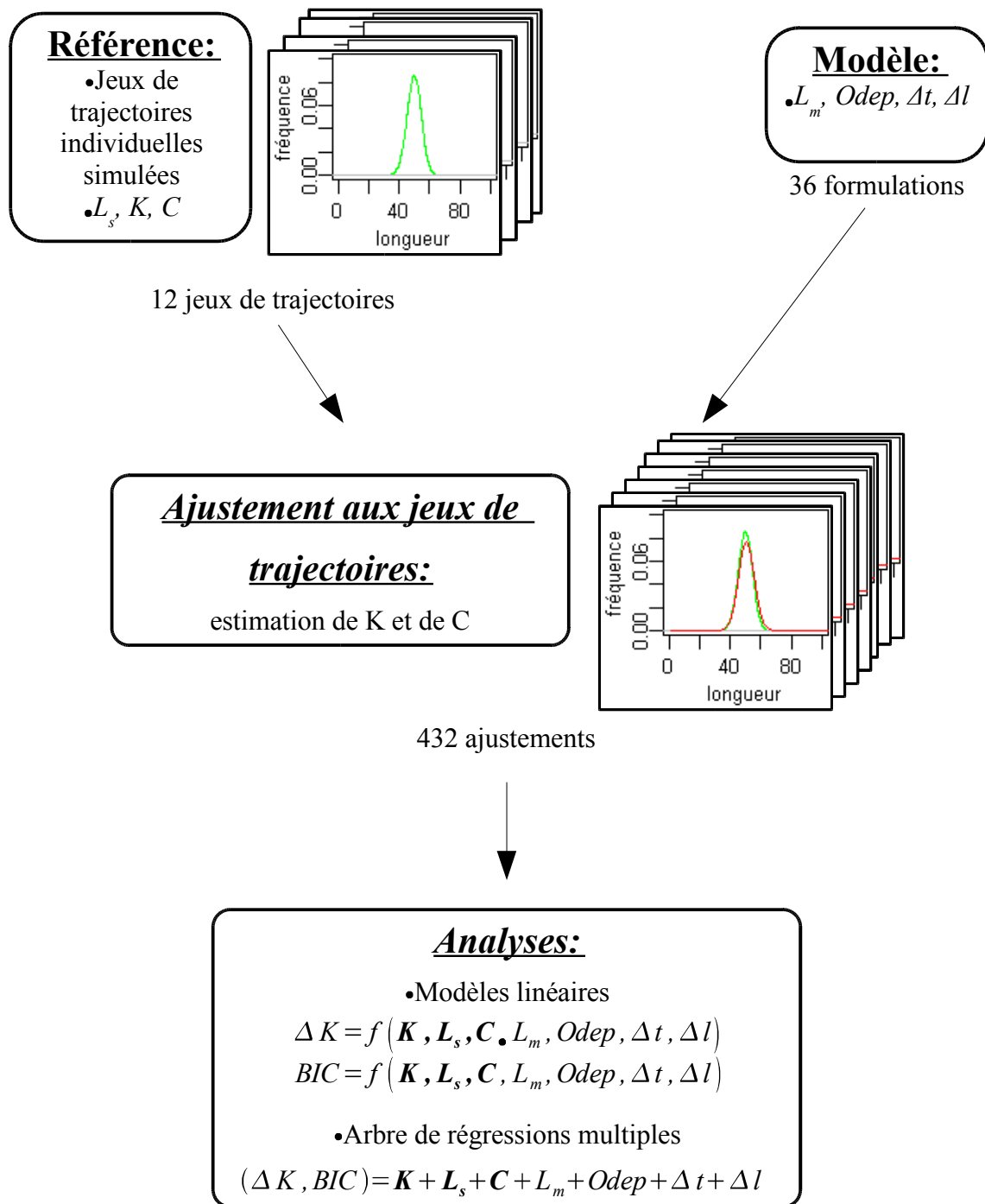


Figure 13 : Représentation schématique de la démarche utilisée.

3.3 Résultats

Les Tableaux 6 (à gauche) et 7 récapitulent les effets significatifs des deux modèles linéaires. Certaines interactions ont du être préalablement enlevées de l'analyse afin que les ajustements satisfassent les hypothèses du modèle linéaire. Les R^2 sont respectivement de 0,995 pour *LBic*, et 0,674 pour *LK*. Dans ces tableaux, les paramètres estimés quantifient l'impact de l'hypothèse correspondante. Du fait des transformations sur les variables à expliquer et du choix des contrastes:

- pour le modèle (9), les paramètres estimés positifs correspondent à des hypothèses améliorant l'ajustement aux fréquences aux longueurs.
- pour le modèle (10), les paramètres estimés négatifs correspondent à des hypothèses améliorant l'estimation du taux de croissance.

Les résidus des deux modèles linéaires sont présentés sur la figure 14. Malgré la transformation logarithmique, on observe une structure dans les résidus.

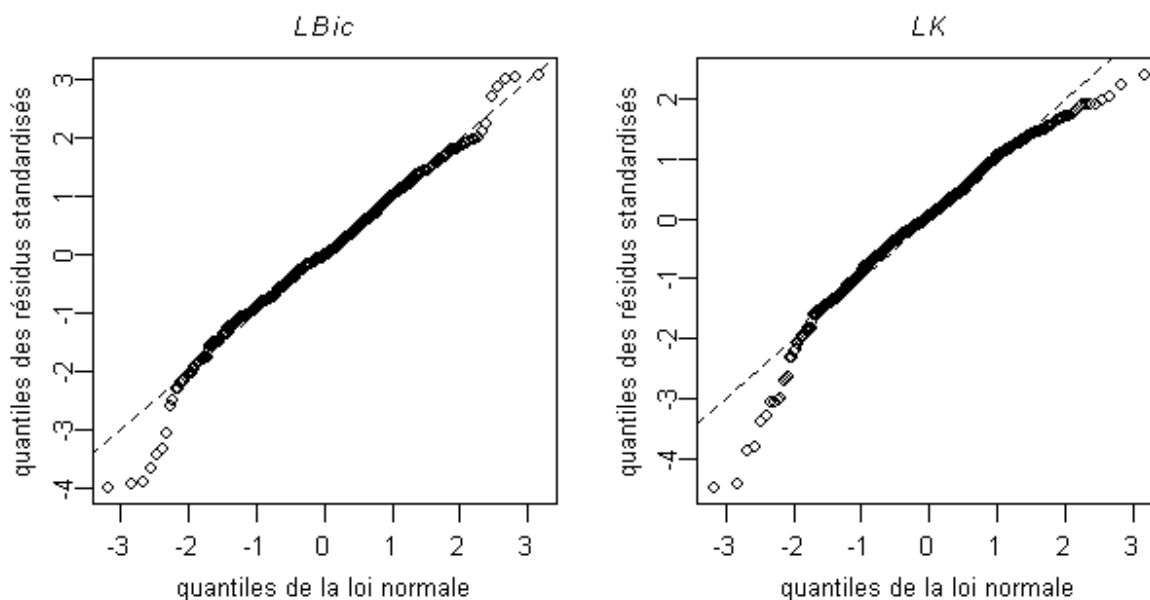


Figure 14 : Résidus des modèles linéaires 9 (à gauche) et 10 (à droite).

Chapitre 3 : Quantification de l'impact d'hypothèses de discrétisation sur un modèle de croissance structuré en longueur

Tableaux 6 (à gauche) et 7 (à droite): Effets significatifs du modèle linéaire (1) (*LBic* - à gauche) et (2) (*LK* - à droite).

*** (respectivement **) signifie $p\text{-value} < 0,01$ (respectivement $< 0,05$). Les coefficients de chacun des termes des modèles linéaires sont contraints de sommer à 0

	<i>LBic</i>			<i>LK</i>	
	Estimation	P-value		Estimation	P-value
Intercept	8.92	***	Intercept	-4.74	***
<i>Lm=0</i>	0.03	***	$\Delta t=1$	-0.49	***
<i>Lm=1</i>	-0.05	***	$\Delta l=1$	-0.38	***
$\Delta t=1$	-1.18	***	<i>Lm=0</i>	-0.35	***
$\Delta t=4$	0.06	**	<i>Lm=1</i>	0.66	***
$\Delta l=1$	1.37	***	<i>ODep=0</i>	0.21	***
<i>CV=0.63</i>	-0.21	***	<i>CV=0.63</i>	-0.65	***
K	0.89	***	K	4.93	***
<i>Ls=0</i>	0.05	**	<i>Ls=0</i>	-0.40	***
<i>Ls=1</i>	-0.06	***	<i>Ls=1</i>	0.64	***
<i>ODep=0</i>	-0.03	***	$\Delta t=1:\Delta l=1$	0.14	***
$\Delta t=1:\Delta l=1$	-0.07	***	$\Delta t=1:ODep=0$	-0.22	***
$\Delta t=4:\Delta l=1$	0.04	***	$\Delta l=1:ODep=0$	-0.13	***
$\Delta t=1:\mathbf{K}$	0.67	***	$\Delta l=1:\mathbf{Ls}=0$	-0.23	***
$\Delta t=1:Lm=0$	-0.02	**	$\Delta l=1:\mathbf{Ls}=1$	0.35	***
$\Delta t=1:Lm=1$	0.03	***	<i>Lm=0:Ls=0</i>	-0.17	**
$\Delta t=4:Lm=1$	0.02	**	<i>Lm=1:Ls=0</i>	0.42	***
$\Delta t=1:\mathbf{Ls}=1$	0.04	***	<i>Lm=0:Ls=1</i>	0.33	***
$\Delta l=1:\mathbf{K}$	-0.49	***	<i>Lm=1:Ls=1</i>	-0.68	***
$\Delta l=1:\mathbf{Ls}=0$	-0.04	***	<i>ODep=0:CV=0.63</i>	0.13	***
$\Delta l=1:\mathbf{Ls}=1$	0.06	***	<i>Lm=1:CV=0.63</i>	-0.27	***
<i>Lm=1:CV=0.63</i>	0.03	***	$\Delta t=1:\mathbf{Ls}=0$	-0.15	**
<i>CV=0.63:K</i>	0.31	**	$\Delta t=1:\mathbf{Ls}=1$	0.18	**
<i>CV=1:K</i>	0.22	**			
<i>CV=0.63:Ls=0</i>	-0.05	***			
<i>CV=0.63:Ls=1</i>	0.08	***			
<i>K:Ls=1</i>	-0.42	***			

3.3.1 Ajustement des compositions en longueur (Tableau 6)

Les effets estimés des interactions sont faibles devant ceux des effets principaux et influent peu sur les résultats. Parmi les effets principaux, presque tous sont significatifs mais la plupart ont des valeurs estimées faibles. Les hypothèses ayant le plus d'impact sont le choix du pas de temps et le choix de la taille des classes de longueur. Les meilleurs ajustements sont obtenus avec un pas de temps mensuel et des classes de 1 cm. Les résultats montrent aussi qu'une grande variabilité interindividuelle (effet de $CV=0.63$) et un taux de croissance fort (coefficient de K positif) sont plus facilement ajustés par le modèle de croissance de population structuré en longueur.

3.3.2 Ajustement du taux de croissance K (Figure 15 et Tableau 7)

Le tracé des boîtes à moustaches (Figure 15) ne met pas en évidence de biais systématique dans l'estimation du taux de croissance, mis à part pour l'hypothèse d'une distribution normale des incréments de croissance. Par contre, elles montrent que la précision de l'estimation est particulièrement faible avec un pas de temps faible, une loi normale, et des classes de 5 cm.

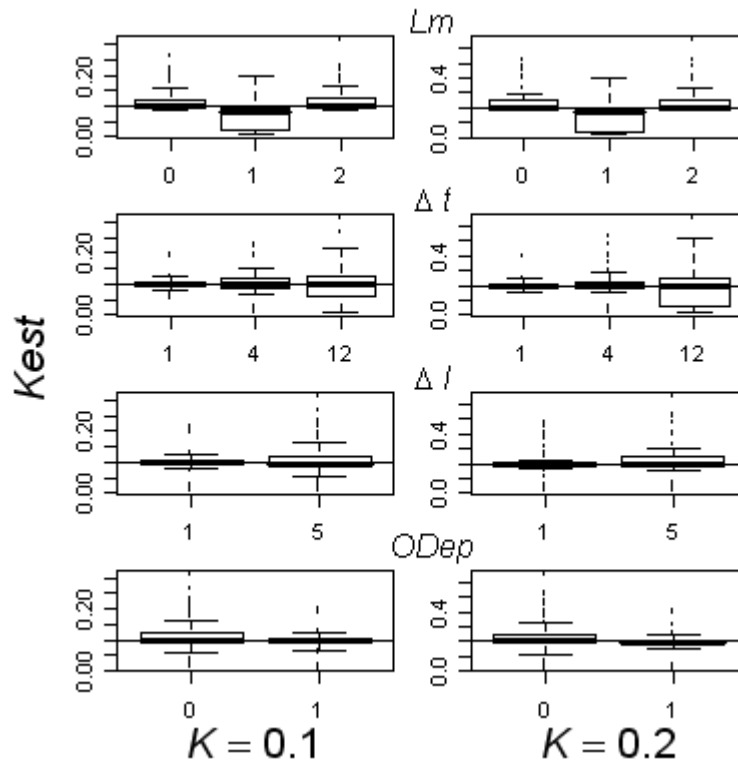


Figure 15 : Boîtes à moustaches représentant la distribution de $Kest$ en fonction des hypothèses du modèle de croissance structuré en longueur. Les lignes horizontales correspondent au taux de croissance K ayant permis de simuler les jeux de trajectoires individuelles de croissance. Sur chaque graphique, les différentes modalités de l'hypothèse considérée sont représentées en abscisses. La valeur de $Kest$ est présentée en ordonnées.

Le modèle linéaire (Tableau 7) montre que beaucoup plus d'hypothèses affectent l'estimation de K que l'ajustement aux fréquences. Au contraire de l'ajustement aux fréquences, des pas de temps courts n'améliorent pas systématiquement l'estimation du taux de croissance. En fait, avec un pas de temps mensuel, la probabilité de changer de classes devient faible, en particulier si les classes mesurent 5 cm, entraînant une surestimation du taux de croissance. Cela met en évidence que le choix du pas de temps et des classes de longueur est primordial dans un modèle structuré en longueur et que les deux choix sont liés (l'interaction étant significative).

Contrairement à l'ajustement aux fréquences, le taux de croissance est mieux estimé quand les taux de croissance sont faibles et quand la variabilité interindividuelle de croissance est faible. On note par ailleurs que supposer que les individus sont uniformément distribués au sein des classes de longueur peut améliorer l'estimation de K , en particulier pour des modèles à pas de temps courts.

On observe également un effet assez fort de K , le taux de croissance utilisé pour simuler les trajectoires individuelles de croissance. Dans le modèle linéaire, K était considéré comme une variable quantitative, contrairement aux autres variables qui étaient des facteurs. Les coefficients estimés ne sont donc pas directement comparables. Cependant, cela démontre que le taux de croissance est plus facilement estimable pour des populations à croissance lente.

Enfin, l'analyse montre que si les lois gamma et lognormale donnent des résultats très similaires, ceux-ci sont légèrement meilleurs pour la fonction gamma. Une distribution normale supposée pour le modèle structuré en longueur donne de bons résultats uniquement si une loi normale a été utilisée pour simuler les jeux de trajectoires de croissance, la loi gamma donnant aussi des résultats convenables dans ce cas. Des trois lois, la loi gamma apparaît donc comme la plus flexible pour modéliser la distribution des incréments de croissance.

3.3.3 Analyse de l'arbre de régressions multiples (Figure 16 et Tableau 8)

Le premier critère de partition est la taille des classes de longueur. L'utilisation de classes de petites tailles améliore l'ajustement aux fréquences aux longueurs. Les deux groupes sont ensuite partitionnés selon le choix du pas de temps: les forts pas de temps ne permettent pas un bon ajustement aux fréquences simulées, en particulier quand les classes de longueur sont larges. On peut noter que la partition sépare les modèles à pas de temps annuel aux modèles à pas de temps mensuels ou trimestriels. Les deux partitions suivantes concernent le choix des distributions statistiques des incréments de croissance : supposer une loi gamma ou lognormale améliorent les ajustements en particulier quand ces mêmes lois ont été utilisées

pour générer les jeux de trajectoires.

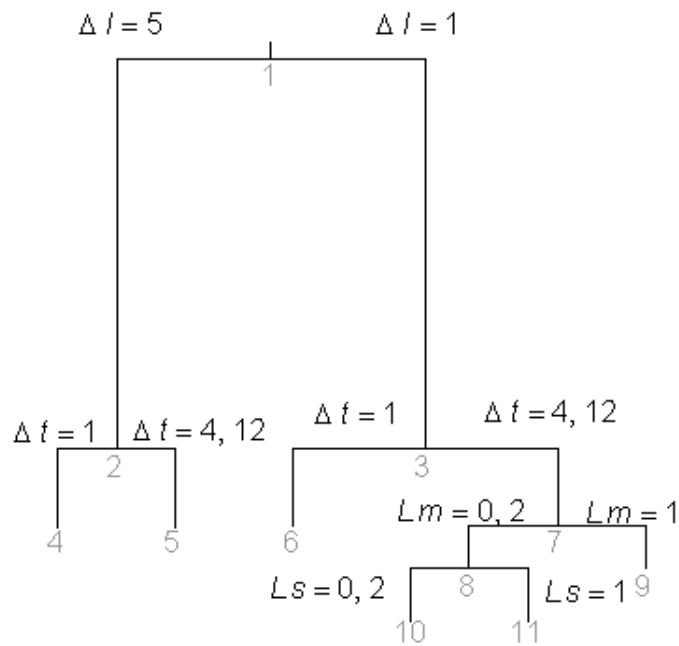


Figure 16 : Arbre de régressions multiples des ajustements du modèle de croissance de population structuré en longueur sur les jeux de trajectoires simulées de croissance individuelle. Les nombres en gris indiquent le numéro du noeud. Les caractéristiques des noeuds sont résumées dans le Tableau 8.

Tableau 8 : Caractéristiques des noeuds de l'arbre de régressions multiples (Figure 16). Les meilleurs ajustements sont ceux avec un faible LK et un fort $LBic$.

Numéro du noeud	Nombre d'expériences du noeud	$LBic$ (centré et réduit) moyen	LK (centré et réduit) moyen	Déviante
1	648	0,00	0,00	1 294
2	324	0,26	-0,83	363
3	324	-0,26	0,83	72
4	108	-0,17	-1,46	195
5	216	0,48	-0,51	445
6	108	-0,51	0,10	97
7	216	-0,14	1,19	251
8	144	-0,48	1,20	134
9	72	0,53	1,16	26
10	96	-0,96	1,23	41
11	48	0,49	1,15	69

3.4 Discussion

Cette étude visait à analyser l'impact de différentes hypothèses sur un modèle structuré en longueur de croissance de population, en particulier des choix du pas de temps et de la taille des classes de longueur, afin de détecter d'éventuels artefacts provoqués par des hypothèses inadaptées et de proposer une structure de modèle flexible. Ce type d'analyse de sensibilité aux hypothèses du modèle est essentiel pour les modèles discrets et la méthode proposée est facilement transposable à d'autres cas d'application. Dans notre cas, elle était un préliminaire au développement du modèle de dynamique des populations structuré en longueur appliqué au merlu. Nous avons décidé de nous intéresser uniquement à la modélisation de la croissance car ce processus est un des éléments clés des modèles structurés en longueur.

Le modèle repose sur un modèle d'incrément de croissance. Les incréments moyens suivent une courbe de von Bertalanffy, mais d'autres équations (Gompertz, Verhulst, Richards, Schnute....) auraient pu être utilisées. La variabilité interindividuelle de croissance est modélisée par une distribution statistique des incréments autour de cette moyenne. En théorie, une variance devrait être estimée pour chaque classe de longueur, mais nous avons choisi d'utiliser des hypothèses de la littérature liant moyenne et variance. Suivant l'hypothèse de DeLong *et al.* (2001), nous avons supposé que la variance des incréments était proportionnelle au carré de la moyenne. Nous avons également réalisé l'étude en supposant que la variance était proportionnelle à la moyenne (Sullivan *et al.*, 1990 ; Sullivan, 1992), les résultats obtenus sont parfaitement similaires. D'autre part, nous avons essayé de modéliser la variabilité interindividuelle de croissance par une loi bêta-binomiale (Ennis et Bi, 1998) modélisant le nombre de classes traversées au cours d'un pas de temps. Cette solution a été assez vite rejetée car: (i) elle nécessitait des classes de longueur peu pratiques (voir page 67), (ii) lors des essais de nombreux artefacts (en particulier la croissance moyenne n'était pas toujours respectée) et un manque de flexibilité ont été observés.

Des données individuelles de croissance sont nécessaires pour évaluer l'impact des hypothèses de discrétisation d'un modèle structuré en longueur de croissance de population. Idéalement, des données de marquage pourraient être utilisées, malheureusement, ce type de données n'était pas disponible en assez grande quantité pour le merlu. Par conséquent, nous avons choisi de simuler des données s'approchant le plus possible de données de croissance individuelles. Le modèle qui a été choisi pour simuler les jeux de trajectoires individuelles peut être discuté. Il aurait pu être intéressant notamment de simuler ces données en supposant que K et/ou L_{∞} étaient des variables stochastiques: chaque individu se voyant donc attribué un couple de paramètres en début de simulation. Les résultats sont probablement en partie liés à la méthode choisie pour simuler les données, et dépendent bien sûr également du modèle et de l'espèce étudiée. Par conséquent, les résultats que l'on présente ne peuvent en aucun cas être généralisés. La méthode est en revanche parfaitement générique et elle est adaptée à l'évaluation d'hypothèses de discrétisations sur des modèles discrets.

Dans cette étude, nous avons fait l'hypothèse que L_∞ était connu. Il peut en général être approché par une fonction de L_{max} . Cette approximation qui permet d'éviter les problèmes de corrélation entre K et L_∞ (le paramètre t_0 n'est pas à estimer si l'on utilise la formulation de Fabens (1965)), a déjà été utilisée (DeLong *et al.*, 2001 ; de Pontual *et al.*, 2006). Toutefois il aurait pu être intéressant d'utiliser une reformulation de l'équation de von Bertalanffy limitant la corrélation entre les deux paramètres à estimer (Shackell *et al.*, 1997). Les valeurs de paramètres de croissance choisies sont proches de celles estimées pour le stock nord de merlu: un jeu correspondant à peu près à l'ancienne courbe de croissance, l'autre étant proche de la courbe estimée à partir des données de marquage (de Pontual *et al.*, 2003 ; de Pontual *et al.*, 2006).

Deux critères ont été choisis pour quantifier la qualité de l'ajustement du modèle structuré en longueur de croissance de population sur les fréquences simulées. Les deux critères correspondent à deux objectifs distincts de ce type de modèle. Le premier objectif, quantifié par $Lbic$, est d'évaluer la capacité du modèle à reproduire les observations. Il s'agit donc plus d'un calibrage du modèle, les estimations n'étant pas alors primordiales. L'autre objectif, quantifié par LK , est au contraire d'estimer K en vue d'améliorer la connaissance sur la biologie de l'espèce. L'étude démontre qu'un modèle ne satisfait pas forcément simultanément ces deux objectifs et qu'un modèle s'ajustant bien aux données ne fournit pas forcément une bonne estimation des processus biologiques sous-jacents. Dans l'étude seule une analyse exploratoire a été réalisée pour quantifier l'impact des hypothèses sur les deux critères conjointement. Une autre manière d'aborder cette analyse aurait pu être d'utiliser une fonction de désirabilité (Harrington, 1965) combinant les deux critères, toutefois ce type de fonction nécessite une pondération, et qui est nécessairement subjective. Les arbres de régression multiples ont permis de réaliser l'approche exploratoire, il constitue un outil intéressant pour explorer l'impact de différentes variables explicatrices (quantitatives ou qualitatives) sur plusieurs variables quantitatives à expliquer.

L'analyse statistique par modèles linéaires permet de mettre en évidence des hypothèses rendant le modèle structuré en longueur flexible. Elle démontre surtout l'importance du choix du pas de temps et de la taille des classes de longueur dans ce type de modèle. Ces deux choix doivent être étroitement liés. Ils sont néanmoins aussi contraints par la disponibilité des données. Diminuer la taille des classes de longueur permet (sous réserve que les données disponibles le permettent) d'améliorer significativement l'ajustement aux fréquences et l'estimation du taux de croissance. Au contraire, réduire le pas de temps améliore l'ajustement aux fréquences mais tend à provoquer une surestimation du taux de croissance. La distribution gamma apparaît comme la distribution la plus flexible pour décrire la variabilité des incréments de croissance. De même, supposer que les individus sont uniformément distribués au sein des classes de longueur peut améliorer l'estimation du taux de croissance, en particulier quand le pas de temps est court. Il aurait pu être intéressant en complément de cette étude d'évaluer l'effet du nombre de trajectoires individuelles de croissance sur l'estimation du taux de croissance.

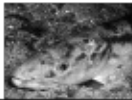
De même, nous n'avons pas considéré d'erreur d'observation, c'est à dire que l'on a supposé que la composition en longueur dans les jeux de trajectoires simulés étaient parfaitement connue. Cela n'est bien entendu le cas dans la réalité et il aurait pu être intéressant de bruite ces observations. En revanche dans ce cas, il aurait fallu simuler plusieurs jeux de trajectoires pour chaque ensemble d'hypothèses $S(K, C, Ls)$ ce qui aurait fortement augmenté les temps de simulations et d'ajustements.

Cette étude a permis de mettre en évidence des hypothèses plus robustes pour le modèle de croissance qui pourront être utilisées dans le modèle final.

Ce qu'il faut retenir

- une méthode générique d'exploration numérique des modèles pour évaluer l'impact d'hypothèses d'un modèle discret décrivant un processus continu
- les choix du pas de temps et de la largeur des classes sont des paramètres sensibles des modèles discrets
- pour le merlu de l'Atlantique Nord-Est, des classes de longueur de 1 cm et un pas de temps trimestriel sont adaptés pour la modélisation du processus de croissance
- une distribution gamma des incréments et une distribution uniforme des individus au sein des classes rendent le modèle plus flexible

**Chapitre 4 : Un modèle de dynamique des
populations structuré en longueur et spatialisé
appliqué au merlu de l'Atlantique Nord-Est**



Chapitre 4 : Un modèle de dynamique des populations structuré en longueur et spatialisé appliqué au merlu de l'Atlantique Nord-Est

4.1 Introduction

Le premier objectif du modèle est de fournir des estimations de certains des paramètres inconnus de la dynamique de la population : croissance et migration. Il doit également permettre d'établir des diagnostics sur l'état du stock et de son exploitation. Comme nous l'avons vu au cours des sections précédentes, les principales motivations pour le développement d'un modèle de dynamique des populations structuré en longueur et spatialisé pour le merlu de l'Atlantique Nord-Est sont :

- l'utilisation directe de données structurées en âge afin de limiter les incertitudes liées à la conversion taille-âge.
- une description explicite des hétérogénéités spatiales de distribution de la population et des efforts de pêche et des déplacements saisonniers de la population, dont on sait que la non prise en compte peut biaiser les estimations de la taille du stock et limite les possibilités d'évaluation de mesures de gestion spatialisées.

La revue bibliographique des modèles existants présentée au chapitre 2 a montré que les modèles de type Fleksibest et Multifan-CI avaient des objectifs proches des nôtres et nous ont servi de base pour développer un modèle structuré en longueur, à temps et espace discrets basé sur deux sous-modèles:

- Le premier sous modèle, décrit la dynamique de la population et de l'activité de pêche, on parle de modèle d'états. Il permet de calculer l'abondance dans chacune des classes et zones du modèle à chaque pas de temps t du modèle.
- Le second modèle est un modèle d'observations, qui décrit le processus générant les observations dont on dispose aux différents pas de temps t .

Pour estimer l'ensemble des paramètres du modèle, on écrit la fonction de vraisemblance des variables observées décrites par le second sous-modèle en utilisant les données observées sur le terrain. Puis, par un processus itératif, on cherche les valeurs des paramètres générant des observations qui maximisent la vraisemblance.

Dans le chapitre précédent, nous avons proposé une formulation robuste d'un modèle de croissance (voir chapitre 3). Nous présentons maintenant le modèle de dynamique de population complet, structuré en longueur et spatialisé ainsi que son ajustement aux données du stock Nord de merlu.

4.2 Matériel et méthodes

La figure 17 représente schématiquement le modèle et la démarche de modélisation utilisé.

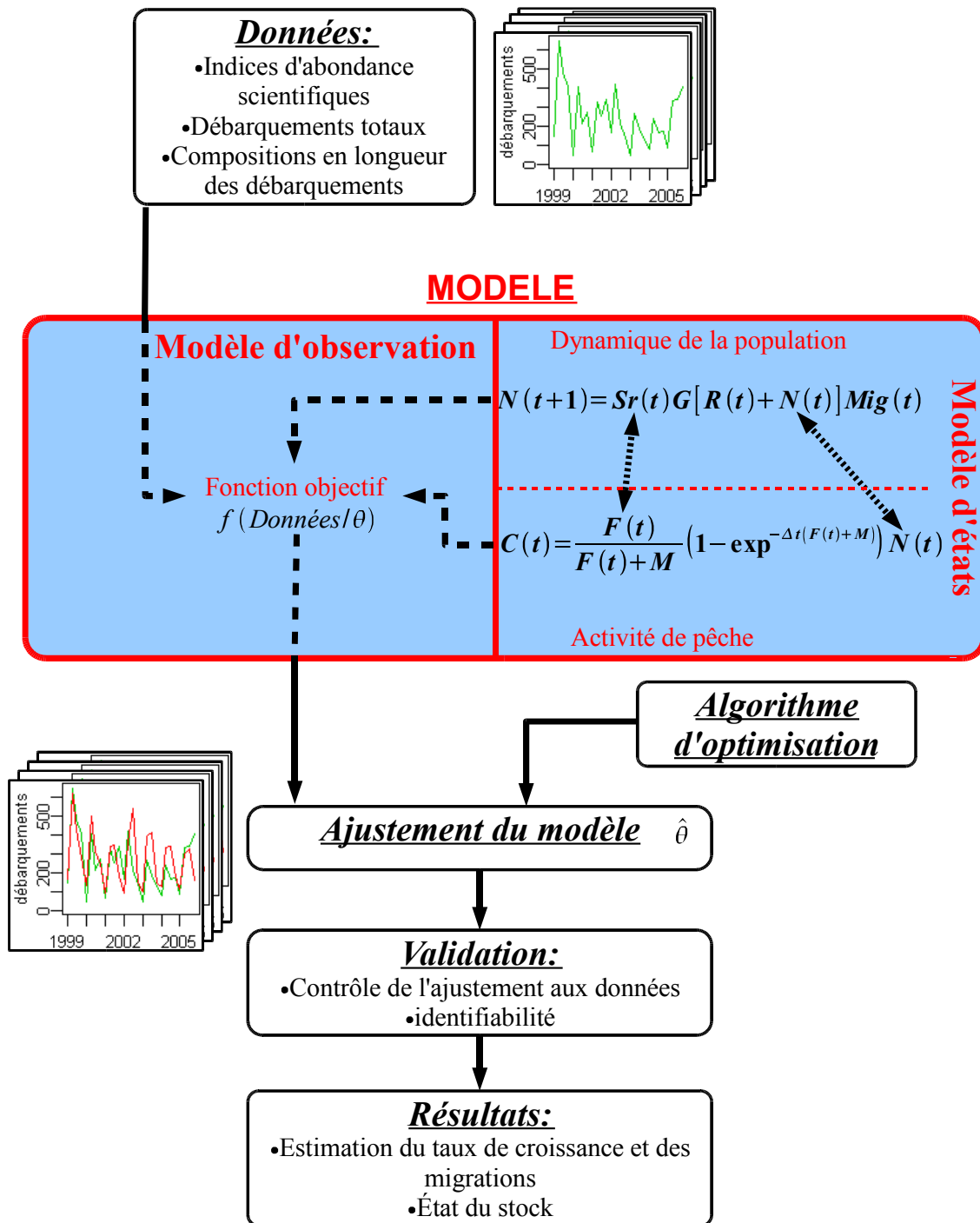


Figure 17 : Représentation schématique du modèle et de la démarche de modélisation.

4.2.1 Modèle d'états

Sous forme matricielle, un modèle structuré en longueur et spatialisé en temps et espace discrets peut s'écrire sous la forme:

$$N(t+1) = Sr(t) [Mig(t)(R(t) + N(t))G] \quad (12)$$

avec $N(t)$ la matrice contenant les abondances dans chaque zone z et chaque classe de longueur c au pas de temps t , $Sr(t)$ la matrice contenant les taux de survie au pas de temps t , $Mig(t)$ une matrice contenant les taux de migration pour chaque classe de longueur c entre les différentes zones z au pas de temps t , G la matrice de croissance (voir chapitre 3) et $R(t)$ la matrice de recrutement au pas de temps t .

Au début de chaque pas de temps la population subit successivement le recrutement, la croissance et la migration, dans un temps ε supposé négligeable. Pendant le reste du pas de temps, elle subit simultanément la mortalité naturelle et la mortalité par pêche.

Le modèle a un pas de temps trimestriel, ce choix étant à la fois basé sur la résolution des données disponibles et sur les résultats de l'analyse de sensibilité du modèle de croissance (voir chapitre 3). L'espace est découpé en 4 zones, la population et l'activité de pêche étant supposées uniformément distribuées au sein de chacune de ces zones (Figure 18). Nous utilisons finalement des classes de 2cm, essentiellement pour des raisons de temps de simulation, réparties entre 3 et 110 cm.

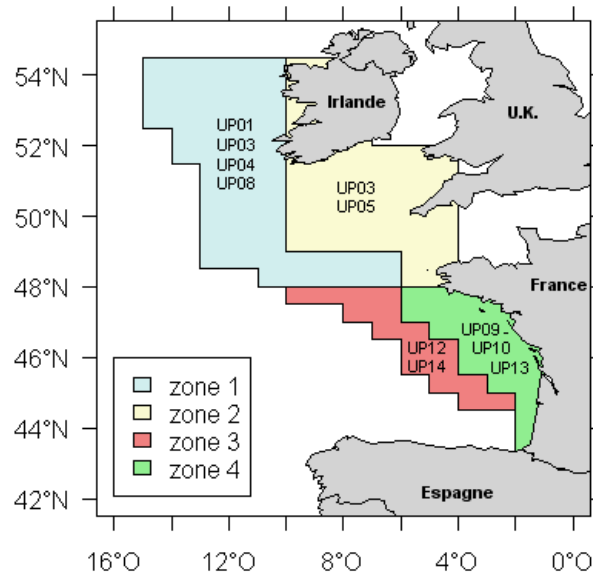


Figure 18 : Carte présentant les 4 zones du modèle et la répartition des unités de pêche.

4.2.1.1 Dynamique de la population

a) Recrutement (voir section 1.2.5)

On suppose que chaque année le recrutement a lieu au deuxième trimestre sur le plateau du golfe de Gascogne (zone 4) et au troisième trimestre sur le plateau de la mer Celtique (zone 2). Les recrutements dans ces deux zones sont supposés être complètement indépendants. Les recrues entrent dans la population dans la plus petite classe de longueur.

$$\begin{aligned}
 &\text{trimestre 1: } N_1(z, 1) = N(t, z, 1) \\
 &\text{trimestre 2: } \begin{cases} N_1(z, c) = N(t, z, c) & \text{si } c \neq 1 \vee z \neq 4 \\ N_1(4, 1) = N(t, 4, 1) + R_{BB}(y) \end{cases} \\
 &\text{trimestre 3: } \begin{cases} N_1(z, c) = N(t, z, c) & \text{si } c \neq 1 \vee z \neq 2 \\ N_1(2, 1) = N(t, 2, 1) + R_{CS}(y) \end{cases} \\
 &\text{trimestre 4: } N_1(z, 1) = N(t, z, 1)
 \end{aligned} \quad (13)$$

avec $R_{BB}(y)$ le recrutement en zone 4 à l'année y et $R_{CS}(y)$ le recrutement en zone 2 à l'année y . Les recrutements annuels $R_{BB}(y)$ et $R_{CS}(y)$ doivent être estimés. c et z sont ici des indices représentant respectivement les classes de longueur et les zones. La classe $c=1$ est la plus petite classe du modèle, la zone $z=2$ correspond à la zone de recrutement de la mer Celtique et la zone $z=4$ représente la zone de recrutement du golfe de Gascogne.

b) Croissance (voir section 1.2.3 et chapitre 3)

La matrice de croissance contient les probabilités de transition d'une classe vers une autre au cours d'un pas de temps. Nous reprenons le modèle d'incrément de croissance présenté dans le chapitre 3. Guidés par les résultats de l'analyse de sensibilité, nous supposons que les incréments de croissance suivent une distribution gamma. En revanche, nous supposons que les individus d'une classe ont tous la longueur moyenne de la classe. Cette dernière hypothèse permet d'éviter le calcul de la double intégrale (voir équation 6) particulièrement « coûteux » en temps de simulation, et l'hypothèse de distribution des individus n'avaient qu'une influence secondaire selon l'analyse de sensibilité.

Étant données ces hypothèses, les effectifs après croissance sont donnés par la relation:

$$N_2(z, c) = \sum_{i=1}^c g_{i,c} \cdot N_1(z, i) \quad (14)$$

avec $g_{i,c}$ la probabilité de transition entre la classe de longueur i et la classe de longueur c .

Comme dans le chapitre précédent, on fixe $L_\infty=110$ cm et $t_0=0$ (de Pontual *et al.*, 2006), seuls K et la constante de variance doivent être estimés.

c) Migrations (voir section 1.2.4)

Comme nous l'avons vu dans la section 1.2.4, les migrations du merlu ont principalement été étudiées par des distributions de CPUE (Poulard, 2001) (la couverture spatiale de la campagne de marquage est encore trop limitée pour apporter beaucoup d'informations sur les migrations). Les migrations du merlu étant essentiellement liées à l'activité reproductrice, il faut dans le modèle distinguer les migrations des matures et des immatures. Les nombres d'individus matures et immatures sont calculés en utilisant une relation longueur/maturité (Piñeiro et Sainza, 2003) qui fournit la proportion d'individus matures dans chaque classe de longueur c (notée $p_{\text{mat}}(c)$). La structuration matures/immatures n'est pas conservée par la suite. On modélise donc les migrations par 4 matrices (une par trimestre) de migration pour les matures et 4 (une par trimestre) pour les immatures. Chaque matrice contient les proportions $mimm_{z_1 \rightarrow z}^s$ (respectivement $mmat_{z_1 \rightarrow z}^s$) d'individus immatures (respectivement matures) migrant d'une zone z_1 à la zone z au cours du trimestre s . Le nombre d'individus dans la zone z et la classe c après les migrations de la saison s est donné par l'équation suivante (Z_0 représente ici le nombre total de zones):

$$N_3(z, c) = \sum_{z_1=1}^{Z_0} \left[(1 - p_{\text{mat}}(c)) \cdot mimm_{z_1 \rightarrow z}^s \cdot N_2(z_1, c) \right] + \sum_{z_1=1}^{Z_0} \left[p_{\text{mat}}(c) \cdot mmat_{z_1 \rightarrow z}^s \cdot N_2(z_1, c) \right] \quad (15)$$

Des contraintes ont été ajoutées sur certains des paramètres $mimm_{z_1 \rightarrow z}^s$ et $mmat_{z_1 \rightarrow z}^s$ afin de limiter le nombre de paramètres à estimer (Tableau 9 et section 1.2.4).

Tableau 9 : Résumé des proportions d'individus partant d'une zone (en ligne) vers une autre (en colonne). Les cases « x » correspondent aux proportions à estimer.

		immatures					matures				
		1	2	3	4	justification	1	2	3	4	justification
trimestre 1	1	1	0	0	0	Immatures immobiles	1	1	0	0	Concentrations à l'ouest
	2	0	1	0	0	Immatures immobiles	2	x	x	0	Concentrations à l'ouest
	3	0	0	1	0	Immatures immobiles	3	0	0	1	Reproduction en zone 3
	4	0	0	0	1	Immatures immobiles	4	0	0	1	Reproduction en zone 3
		-----					-----				
		1	2	3	4	justification	1	2	3	4	justification
trimestre 2	1	1	0	0	0	Immatures immobiles	1	1	0	0	Reproduction en zone 1
	2	0	1	0	0	Immatures immobiles	2	1	0	0	Reproduction en zone 1
	3	0	0	1	0	Immatures immobiles	3	x	0	x	Matures retournent vers le plateau ou reproduction au nord
	4	0	0	x	x	Dispersion du recrutement	4	0	0	0	1
		-----					-----				
		1	2	3	4	justification	1	2	3	4	justification
trimestre 3	1	1	0	0	0	Immatures immobiles	1	x	x	0	Concentrations à l'est
	2	x	x	0	0	Dispersion du recrutement	2	0	1	0	Concentrations à l'est
	3	0	0	1	0	Immatures immobiles	3	0	0	1	Concentrations aux accores
	4	0	0	0	1	Immatures immobiles	4	0	0	x	x
		-----					-----				
		1	2	3	4	justification	1	2	3	4	justification
trimestre 4	1	1	0	0	0	Immatures immobiles	1	x	x	0	Concentrations à l'est
	2	0	1	0	0	Immatures immobiles	2	0	1	0	Concentrations à l'est
	3	0	0	1	0	Immatures immobiles	3	0	0	1	Concentrations aux accores
	4	0	0	0	1	Immatures immobiles	4	0	0	x	x

Avec ces contraintes, seulement 17 paramètres sont à estimer. Du fait que

$\sum_{z_i}^Z mmat_{z \rightarrow z_i}^s = 1$ et $\sum_{z_i}^Z mimm_{z \rightarrow z_i}^s = 1$, seulement 10 paramètres sont effectivement à estimer.

d) Survie

La mortalité naturelle et la mortalité par pêche sont supposées s'appliquer simultanément pendant le reste du pas de temps. Le modèle permet théoriquement de distinguer la mortalité naturelle des matures et celles des immatures. Le nombre de survivants au pas de temps $t+1$ est donné par la relation:

$$N(t+1, z, c) = (1 - p_{\text{mat}}(c)) \cdot N_3(z, c) \cdot e^{-\Delta t \cdot (M_{\text{imm}} + F_{\text{tot}}(t, z, c))} + p_{\text{mat}}(c) \cdot N_3(z, c) \cdot e^{-\Delta t \cdot (M_{\text{mat}} + F_{\text{tot}}(t, z, c))}$$

(16)

avec $F_{\text{tot}}(t, z, c)$ la mortalité par pêche s'appliquant au pas de temps t sur la classe c dans la zone z .

Pour simplifier, la mortalité naturelle étant fortement corrélée à la mortalité par pêche, nous supposons que M_{mat} et M_{imm} sont connus et égales à $0,6 \text{ an}^{-1}$. Cette valeur est plus élevée que celle utilisée par le CIEM. Cependant, dans l'hypothèse d'une croissance plus rapide, il est vraisemblable que la mortalité naturelle soit également plus forte. Nous ne prenons donc pas en compte de possibles phénomènes de densité-dépendance dans la mortalité naturelle, notamment liés au cannibalisme.

4.2.1.2 Calcul de la mortalité par pêche et des captures

Comme nous l'avons vu dans la section 1.3, le CIEM a défini, pour l'évaluation du stock de merlu, des unités de pêcherie (Tableau 2) qui correspondent à la réunion d'un engin de pêche, d'une zone de pêche et d'espèces cibles. Les données par « unité de pêcherie » correspondent à la résolution disponible la plus fine pour les données de captures. Nous avons donc choisi de baser le calcul de la mortalité par pêche sur ces unités de pêcherie. Les unités de pêcherie 2 et

6 ne capturant presque pas de merlus, elles n'ont pas été prises en compte dans l'étude. L'unité de pêche 16, pêchant dans la zone la plus au nord et dont les données sont limitées a également été rejetée. Enfin l'unité de pêche 15 (« divers ») a été rejetée au vu du peu de données et de l'absence de données d'effort.

Une zone de pêche a été attribuée à chaque unité de pêche. Seule l'unité de pêche 03 a du être répartie sur deux zones du modèle (Figure 18).

Chaque unité de pêche a ensuite été divisée en sous-unités de pêcheries nationales (Tableau 10).

Tableau 10 : Liste des sous-unités de pêche (FR : France, SP : Espagne, EW : Angleterre), et des compositions en longueurs des débarquements disponibles. T=trimestrielle A=annuelle.

Unité de pêche	Sous-unité	1999	2000	2001	2002	2003	2004	2005
UP01	UP01EW							
	UP01SP	T	A	T	T	T	T	T
UP03	UP03EW	T	T	T	T	T	T	T
	UP03FR							
	UP03SP	T	A	T	T	T	T	T
UP04	UP04EW						A	T
	UP04FR							
	UP04SP	T	A	T	T	T	T	T
UP05	UP05EW	T	T	T	T	T	T	T
	UP05FR	A	A	A	T	T	T	A
UP08	UP08FR							
UP09	UP09FR	A	A		T	T	T	T
UP10	UP10FR	A	A		T	T	T	T
UP12	UP12FR		A		T	T	T	T
	UP12SP	T	A	T	T	T	T	T
UP13	UP13FR	A	A		T	T	T	T
	UP13SP	T	A	T	T	T	T	T
UP14	UP14SP	T	A	T	T	T	T	T

La mortalité par pêche engendrée par une sous-unité de pêche sur les effectifs d'une classe de longueur de la population à un pas de temps t est le produit de la capturabilité de la sous-unité de pêche (que l'on peut définir comme la probabilité qu'un individu soit capturé par une unité d'effort non sélective), du patron de sélection (qui prend en compte à la fois la sélectivité de l'engin et d'éventuels aspects tactiques) de la sous-unité de pêche sur la classe de longueur de la population, de la dérive annuelle d'efficacité de pêche (liée au progrès technique de la sous-unité de pêche et à l'accumulation d'expérience) et de l'effort de pêche (mesure des moyens mis en oeuvre pour la capture d'individus) de la sous-unité de pêche au

pas de temps. La capturabilité et la dérive annuelle sont supposées être dépendantes de la sous-unité de pêcheur tandis que le facteur lié à la longueur est supposé dépendre de l'unité de pêcheur. Pour limiter le nombre de paramètres, ce dernier est modélisé par une fonction mathématique de la longueur : (i) une fonction sigmoïde représentant un taux de rétention croissant pour les chalutiers (par exemple (Froysa *et al.*, 2002)), (ii) une loi gamma pour les palangriers et les fileyeurs (Deriso *et al.*, 1985).

La mortalité générée par la sous-unité de pêcheur su de l'unité de pêcheur up , au cours du pas de temps t de l'année y est donc donnée par la relation:

$$F_{su}(t, z, c) = \begin{cases} q_{su} \cdot p_{su}^{y-y_0} \cdot s_{up}(c) E_{su}(t) \frac{Surface(z)}{Surface_{up}} & \text{si } z \text{ est une zone de pêche de } up \\ 0 & \text{sinon} \end{cases} \quad (17)$$

avec la q_{su} capturabilité de la sous-unité de pêcheur, p_{su} la dérive annuelle de la sous-unité de pêcheur, y_0 l'année initiale, $s_{up}(c)$ le facteur lié à la taille et $E_{su}(t)$ l'effort. Le ratio $Surface(z)/Surface_{fu}$ permet de réallouer l'effort proportionnellement à la taille des zones pour les unités de pêcheur pêchant sur plusieurs zones (UP03).

On peut alors calculer une mortalité par pêche totale par zone et par classe $F_{tot}(t, z, c)$:

$$F_{tot}(t, z, c) = \sum_{su} F_{su}(t, z, c) \quad (18)$$

La capture totale (toutes sous-unités de pêcheur confondues) est alors calculée à partir de la relation suivante:

$$C_{tot}(t, z, c) = p_{mat}(c) \cdot N_3(t, z, c) \cdot \frac{F_{tot}(t, z, c)}{M_{mat} + F_{tot}(t, z, c)} \cdot (1 - \exp^{-\Delta t \cdot (M_{mat} + F_{tot}(t, z, c))}) + (1 - p_{mat}(c)) \cdot N_3(t, z, c) \cdot \frac{F_{tot}(t, z, c)}{M_{imm} + F_{tot}(t, z, c)} \cdot (1 - \exp^{-\Delta t \cdot (M_{imm} + F_{tot}(t, z, c))}) \quad (19)$$

La capture est ensuite réallouée à chacune des sous-unités de pêcheur par un ratio de mortalité par pêche:

$$C_{su}(t, z, c) = C_{tot}(t, z, c) \cdot \frac{F_{su}(t, z, c)}{F_{tot}(t, z, c)} \quad (20)$$

On fait l'hypothèse que captures et débarquements sont strictement équivalents, sauf pour l'unité de pêche 9, cette unité de pêche étant celle qui génère le plus de rejets (ICES, 2006b) et la seule pour laquelle on dispose d'observations. Pour celle-ci, une ogive de rejets a été estimée à partir d'observations embarquées (Figure 19) (Système d'Informations Halieutiques: <http://www.ifremer.fr/sih>).

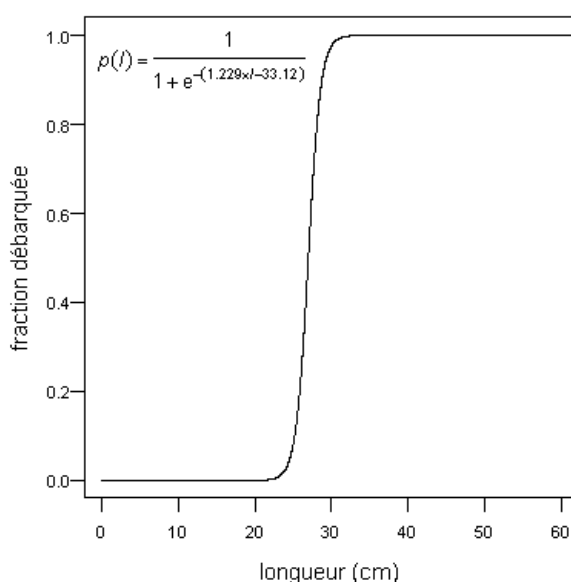


Figure 19 : Fraction de la capture totale débarquée pour la sous-unité de pêche langoustinière française du golfe de Gascogne (UP09FR).

On peut alors calculer pour une unité de pêche su : (i) $Y_{su}(t)$ son débarquement total en poids au temps t et (ii) la composition en longueur $p_{su}(période, c)$ du débarquement sur une période :

$$Y_{su}(t) = \sum_{z=1}^4 \sum_{c=1}^C w_c C_{su}(t, z, c) \quad (21)$$

$$p_{su}(période, c) = \frac{\sum_{t \in période} C_{su}(t, z, c)}{\sum_{t \in période} \sum_{c=1}^c C_{su}(t, z, c)} \quad (22)$$

Le poids des différentes classes provient d'une relation taille poids établie par Piñeiro et Sainza (2003).

4.2.1.3 Calcul des indices d'abondance de campagnes scientifiques

Les campagnes scientifiques fournissent annuellement des estimations d'indices d'abondance (mesure relative de l'abondance). On dispose pour cette étude de trois séries d'indices d'abondance obtenus par chalutage: UK-WCGFS, FR-EVHOE et FR-RESSGASCS (Tableau 11).

Tableau 11 : Présentation de la disponibilité des indices d'abondance disponibles (X) par année et zone pour les différentes campagnes scientifiques.

zone	1999				2000				2001				2002				2003				2004				2005			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
FR-EVHOE trimestre 4	X	X		X		X		X	X	X		X	X				X	X	X		X	X	X	X	X	X	X	X
FR- RESSGASCS trimestre 2				X								X																
FR- RESSGASCSS trimestre 4				X								X																
UK-WCGFS trimestre 1	X	X			X	X	X		X	X			X	X	X		X	X							X	X		
UK-WCGFS trimestre 4																								X				X

On fait l'hypothèse que les campagnes scientifiques opèrent comme une sous-unité de pêcherie sauf que (i) la mortalité engendrée par la campagne est considérée négligeable devant la mortalité par pêche et la mortalité naturelle (ii) l'abondance est considérée constante pendant la durée de la campagne. Le nombre d'individus capturés par une campagne $surv$ pendant le pas de temps t dans une zone z et une classe de longueur c est approximé par le produit d'un effort de pêche, d'une capturabilité, du nombre d'individus dans la zone, éventuellement d'un facteur de sélection lié à la longueur, et d'un ratio de surface nécessaire pour que les indices d'abondance soit comparables entre zones ($SurfaceMin$ désigne ici la surface de la plus petite zone échantillonnée par la campagne):

$$C_{surv}(t, z, c) = q_{surv} \cdot s_{surv}(c) \cdot E_{surv}(t, z) \cdot N_3(t, c, z) \frac{SurfaceMin}{Surface(z)} \quad (23)$$

Pour FR-EVHOE et UK-WCGFS, deux campagnes utilisant des chaluts à petit maillage, on suppose que la sélectivité $s_{surv}(c)$ est constante égal à 1. FR-RESSGASCS utilisant au contraire un chalut professionnel, on suppose que $s_{surv}(c)$ suit une sigmoïde. De plus, on considère que la capturabilité de FR-EVHOE est constante et égale à 10^{-6} . Enfin, nous avons décidé de ne pas prendre en compte des indices d'abondance pour lesquels l'effort d'échantillonnage (temps de pêche) pour le pas de temps et la zone considérés était trop faible.

4.2.1.4 Effectifs initiaux

Théoriquement, les effectifs initiaux par classe de longueur et zones devraient être estimés. Cela nécessiterait l'estimation de 212 paramètres supplémentaires. On préfère supposer qu'au début de la simulation, la population est en équilibre avec la mortalité par pêche et les recrutements de première année. Avant la simulation, on fait donc une présimulation de 10 ans pour laquelle

- les effectifs initiaux sont nuls
- chaque année on applique le recrutement et la mortalité par pêche de la première année
- les effectifs finaux de la présimulation donnent les effectifs de première année pour la simulation

4.2.2 Modèle d'observations

- a) Observation du débarquement total en poids par sous-unité de pêche et par pas de temps

On dispose à partir du flux déclaratif d'une estimation du débarquement total en poids par sous-unité de pêche et par pas de temps. On suppose que cette observation suit une loi normale de moyenne l'estimation du modèle et de variance constante par sous-unité de pêche. La logvraisemblance des observations des débarquements en poids pour une sous-unité de pêche est donc égale (à une constante près) :

$$\log LI_{su} = -\frac{1}{2} \cdot n_{su} \cdot \ln \left(\sum_t (Y_{su}(t) - Y_{obs_{su}}(t))^2 \right) \quad (24)$$

avec n_{su} le nombre de points d'observations pour la sous-unité de pêche su , et $Y_{obs_{su}}(t)$ $Y_{su}(t)$ les débarquements totaux en poids respectivement observés et estimés par le modèle. Cette formulation de la vraisemblance est dite vraisemblance concentrée, formulation dans laquelle la variance est remplacée par son estimateur du maximum de vraisemblance. Une explication détaillée est fournie par Deriso *et al.* (2007).

- b) Composition en taille des débarquements

Des programmes d'échantillonnage en criée fournissent une estimation des compositions en taille des débarquements par sous-unité de pêche. Selon les années et les sous-unités de pêche, on dispose soit d'une composition annuelle, soit d'une trimestrielle, soit d'aucune (Tableau 10). Quand n poissons sont échantillonnés, le nombre d'individus n_1, \dots, n_C dans chacune des classes de longueur $1, \dots, C$ suit une loi multinomiale dont les probabilités p_1, \dots, p_C sont les proportions d'individus dans chacune des classes de longueur. La vraisemblance d'un échantillon s'écrit alors (Johnson *et al.*, 1997) :

$$P(N_1 = n_1, N_2 = n_2, \dots, N_C = n_C) = n! \cdot \prod_{i=1}^C \frac{p_i^{n_i}}{n_i!} \quad (25)$$

En général, on ne dispose pas des n_1, \dots, n_C mais de fréquences observées f_1, \dots, f_C . On préfère donc écrire la relation 25 sous la forme:

$$\log(P(N_1=n_1, N_2=n_2, \dots, N_C=n_C)) = \ln(n!) + n \sum_i^C f_i \ln(p_i) - \sum_i^C \ln((n f_i)!) \quad (26)$$

qui en éliminant les termes ne dépendant pas de p_i se ramène à la simple relation:

$$\log(P(N_1=n_1, N_2=n_2, \dots, N_C=n_C)) = n \sum_i^C f_i \ln(p_i) \quad (27)$$

Si l'on considère un échantillon s des débarquements obtenu par échantillonnage en criée au cours de *période*, si p_i est la fréquence estimée par le modèle calculée par l'équation 22, et f_i la fréquence observée dans l'échantillon s . La logvraisemblance de l'échantillon est donnée par la relation (à une constante près) :

$$\log P_s = n_s \sum_i^C [f_{obs,s}(c) \ln(p_{su}(période, c))] \quad (28)$$

Pour réduire l'influence des trop gros échantillons, on fixe n_s égal au maximum à 500. Pour certaines années (ou trimestres), on ne dispose pas du nombre d'individus échantillonnés. Plutôt que de ne pas utiliser ces échantillons, nous avons décidé de leur attribuer comme nombre celui de l'année (ou du trimestre correspondant) avec le plus petit échantillon.

La fonction objectif proposée par Fournier *et al.* (1998) a également été testée :

$$\log P_s = \sum_c -\frac{1}{2} \cdot \ln(v_s(c)) + \ln \left(e^{-\frac{1}{2} \frac{(p_{su}(période, c) - f_{obs,s}(c))^2}{v_s(c)}} + 0.01 \right) \quad (29)$$

avec $v_s(c) = \left(\frac{1 - p_{su}(période, c)}{p_{su}(période, c)} + \frac{0.1}{C} \right) \cdot \frac{1}{n_s}$

Comme on l'a vu dans la section 2.5.1, les hypothèses sous-jacentes à cette formulation de la vraisemblance ne sont pas pleinement vérifiées, pourtant elle s'est avérée donner de bons résultats dans de nombreuses applications (Fournier *et al.*, 1998 ; Breen *et al.*, 2003 ; Maunder et Watters, 2003). D'après Fournier *et al.* (1990), cette formulation est plus robuste, car elle limite l'influence des classes pour lesquelles le modèle prédit des très faibles ou des très fortes fréquences.

c) Captures des campagne scientifiques

Des campagnes de chalutage (entre autres) sont menées annuellement par différents instituts de recherche européens. Celles-ci reposent sur des plans d'échantillonnage standardisés reproduits à l'identique d'année en année. Les captures obtenues sont donc comparables d'années en années et fournissent donc une mesure relative de l'abondance des espèces étudiées. On fait l'hypothèse que les captures observées d'une campagne suivent une distribution gamma de moyenne donnée par l'équation 23 et de coefficient de variation constant. La logvraisemblance peut alors s'écrire (McCullagh et Nelder, 1989):

$$\log L = n(\nu \ln(\nu) - \ln(\Gamma(\nu))) + \nu \sum_i \left(\ln\left(\frac{x_i}{\mu_i}\right) - \frac{x_i}{\mu_i} \right) \text{ avec } E(X_i) = \mu_i \text{ et } V(X_i) = \frac{\mu_i^2}{\nu} \quad (30)$$

Le coefficient de variation étant supposé constant, ν est une constante dont l'estimateur $\hat{\nu}$ du maximum de vraisemblance est la solution de l'équation:

$$2n(\ln(\hat{\nu}) - \Psi(\hat{\nu})) = D$$

avec D la déviance $D = -2 \sum_i \left[\ln\left(\frac{x_i}{\mu_i}\right) - \frac{x_i - \mu_i}{\mu_i} \right] \quad (31)$

Cependant, il n'existe aucune solution analytique à l'équation 31. On utilise donc l'approximation suivante (McCullagh et Nelder, 1989):

$$\hat{\nu}^{-1} \simeq \frac{1}{n} \sum_i \left(\frac{x_i - \mu_i}{\mu_i} \right)^2 \quad (32)$$

La contribution de la campagne $camp$ est donc:

$$\log L3_{camp} = n_{camp} (\nu \ln(\nu) - \ln(\Gamma(\nu))) + \nu \left(\sum_t \sum_z \sum_c \log L3_{camp}(t, z, c) \right)$$

$$\log L3_{camp}(t, z, c) = \begin{cases} 0 & \text{si } E_{camp}(t, z, c) = 0 \\ \ln \left(\frac{Cobs_{camp}(t, z, c)}{C_{camp}(t, z, c)} \right) - \frac{Cobs_{camp}(t, z, c)}{C_{camp}(t, z, c)} & \text{sinon} \end{cases}$$

(33)

avec n_{surv} le nombre de points d'observation. Les trois campagnes ne capturant presque pas d'individus supérieurs à 40 cm, on ne prend pas en compte les classes de longueur supérieures à 40 cm dans le calcul de la vraisemblance.

4.2.2.2 Paramètres inconnus et données disponibles

La formulation finale du modèle comporte 86 paramètres inconnus à estimer (Tableau 12). Bien que ce nombre soit plus faible que dans les modèles Multifan-CL et Fleksibest, une étude de l'identifiabilité des paramètres est indispensable.

Tableau 12 : Tableau récapitulant les paramètres inconnus du modèle.

Famille de paramètres	Paramètres	Notation	Nombre
Population	Recrutements annuels	$R_{BB}(y)$ $R_{CS}(y)$	2 x Y=14
	Croissance	K, C	2
	Migration	m_{mat}, m_{imm}	10
Activité de pêche	Capturabilité par sous-unité de pêcherie	q_{su}	18
	Dérive annuelle de l'efficacité de pêche par sous-unité de pêcherie	p_{su}	18
	Patron de sélection par unité de pêcherie	s_{up}	20
Campagnes scientifiques	Facteur d'échelle (UK-WCGFS, FR-RESSGASCS)	q_{surv}	2
	Patron de sélection (FR-RESSGASCS)	s_{surv}	2
TOTAL			86

On dispose des débarquements totaux en poids et des efforts de pêche par trimestre pour toutes les sous-unités de pêcherie. Les compositions en longueur des débarquements disponibles sont présentées dans le tableau 10. Les indices d'abondance disponibles sont résumés dans le tableau 11.

4.2.2.3 Fonction objectif et algorithme d'optimisation

La fonction objectif est la logvraisemblance des observations. Sous l'hypothèse que l'ensemble de ces observations sont indépendantes, la fonction d'objectif est égale à la somme de la logvraisemblance des débarquements totaux en poids (équation 24), de la logvraisemblance des compositions en longueur (équation 28 ou 29), et la logvraisemblance des indices d'abondance (équation 33)

$$\log L(\theta) = \sum_{su} \log L_{su} + \sum_s \log P_s + \sum_{camp} \log L_{camp} \quad (34)$$

Le programme a été codé en C++. L'optimisation de la fonction objectif se fait à l'aide d'un algorithme de quasi-Newton inclut dans la librairie autodif (<http://www.otter-rsch.com/>). Normalement l'ajustement aurait dû se faire en une maximisation directe de la fonction d'objectif 34, mais dans ce cas là seules les données de composition en longueur des débarquements étaient ajustées. Nous avons donc décidé dans une première phase de limiter le poids des structures en taille (observations les plus abondantes masquant le signal existant dans les données scientifiques) afin d'explorer une zone de l'espace de recherche dont les optima ne soient pas uniquement guidés par les débarquements en longueur.

- i) dans la première étape, on diminue le poids des compositions en longueur (divisé par 100). Le poids choisi, assez arbitrairement, était destiné à ce que la valeur des contributions des débarquements totaux en poids et celle des campagnes scientifiques soit de même ordre que celle des compositions en longueur. La fonction de vraisemblance est donc au cours de cette phase :

$$\log L(\theta) = \sum_{su} \log L_{su} + \frac{1}{100} \sum_s \log P_s + \sum_{camp} \log L_{camp} \quad (35)$$

- ii) dans la seconde étape on minimise l'équation 34 en utilisant comme point de départ pour l'algorithme, les estimations obtenus au cours de la première phase.

4.2.3 Identifiabilité du modèle: calcul de la matrice hessienne et de la matrice des corrélations

Un jeu de paramètres est dit non identifiable quand plusieurs jeux de valeurs conduisent au maximum de vraisemblance (Reboulet *et al.*, 1999). S'assurer de l'identifiabilité d'un modèle est indispensable pour pouvoir utiliser les critères d'information d'Akaike ou le critère d'information bayésien qui reposent sur le nombre de paramètres non redondants (identifiables) estimés (Gimenez *et al.*, 2004). De plus, la redondance entre paramètres peut entraîner des erreurs dans l'interprétation biologique des résultats puisque plusieurs solutions sont équivalentes (Reboulet *et al.*, 1999).

Pour contrôler l'identifiabilité du modèle on applique la méthode présentée par Viallefont *et al.* (1998), Gimenez *et al.* (2004) et Reboulet *et al.* (1999). Cette méthode repose sur le calcul de la matrice hessienne (que l'on approxime ici par la méthode des différences finies), et par la détermination des valeurs singulières de cette matrice. Le nombre de paramètres réellement identifiables est égal aux nombres de valeurs propres non nulles (Viallefont *et al.*, 1998). Une explication plus détaillée de cette méthode, ainsi que de méthodes alternatives permettant de contrôler l'identifiabilité des modèles, est faite en annexe 2.

A partir de la matrice hessienne, on calcule également la matrice de variance covariance des paramètres estimés (qui est égale à l'inverse de la matrice hessienne). On peut alors calculer pour chaque paramètre θ_i :

$$\hat{\theta}_i - t_{0.9725}^{86-2} \sigma_i \leq \theta_i \leq \hat{\theta}_i + t_{0.9725}^{86-2} \sigma_i \quad \text{avec } t_{0.9725}^{86-2} \simeq 1.987 \quad (36)$$

On calcule également la matrice des corrélations à partir de la matrice de variance covariance. Cette analyse complémentaire, à l'analyse de la matrice hessienne, permet de détecter les fortes corrélations entre paramètres estimés.

4.3 Résultats

Les premiers ajustements réalisés avec la formulation 28 de la logvraisemblance des compositions en longueur des débarquements ont donné des résultats peu satisfaisants : le modèle s'ajustait parfaitement aux compositions en longueur des débarquements mais les autres jeux de données n'étaient pas du tout ajustés, confirmant le manque de robustesse de l'hypothèse d'une distribution multinomiale (Fournier *et al.*, 1990). Nous avons donc utilisé par la suite la formulation 29. Tous les résultats présentés ci-après seront donc basés sur cette formulation.

Tableau 13 : Tableau récapitulant le poids des différents jeux de données dans la logvraisemblance totale.

Débarquements totaux en poids		Compositions en longueur des débarquements		Indices d'abondance	
sous-unités	logvrais.	sous-unités	logvrais	campagne	logvrais
UP01SP	-202.24	UP01SP	-2179.57	FR	-72.68
UP01EW	-113.00	UP01EW	0.00	EVHOE	
UP03SP	-175.51	UP03SP	-6006.21		
UP03FR	-205.77	UP03FR	0.00		
UP03EW	-160.65	UP03EW	-5838.60		
UP04SP	-229.09	UP04SP	-2555.05		
UP04FR	-175.56	UP04FR	0.00	FR	-26.22
UP04EW	-186.76	UP04EW	-673.58	RESSGASCS	
UP05EW	-64.97	UP05EW	-2536.96		
UP05FR	-177.32	UP05FR	-1967.09		
UP08FR	-150.74	UP08FR	0.00		
UP09FR	-179.65	UP09FR	-3444.00		
UP10FR	-184.54	UP10FR	-5631.03		
UP12SP	-188.38	UP12SP	-3701.94	UK	-211.19
UP12FR	-119.68	UP12FR	-2309.54	WCGFS	
UP13SP	-219.73	UP13SP	-6506.82		
UP13FR	-231.40	UP13FR	-4465.51		
UP14SP	-227.36	UP14SP	-1619.32		
TOTAL	-3192.36	TOTAL	-49435.22	TOTAL	-310.09

4.3.1 Ajustement aux données observées

Pour juger de la qualité de l'ajustement, on examine dans un premier temps l'ajustement du modèle aux différents types de données observées (compositions en longueur des débarquements, débarquements totaux en poids et captures scientifiques), puis on examine la matrice hessienne et la matrice des corrélations afin de détecter des problèmes d'identifiabilité.

4.3.1.1 Ajustement aux compositions en longueur des débarquements

Les ajustements aux distributions en longueur des débarquements sont très bons (Figure 20 et Figure 21), et le modèle reproduit bien les pics observés dans les débarquements. Ces données sont celles qui ont le plus de poids dans l'ajustement. On ne présente pas toutes les années et pour toutes les sous-unités de pêche mais tous les ajustements sont de la même qualité. Le modèle s'ajuste un peu moins bien sur les individus les plus grands ce qui peut s'expliquer par la largeur des classes.

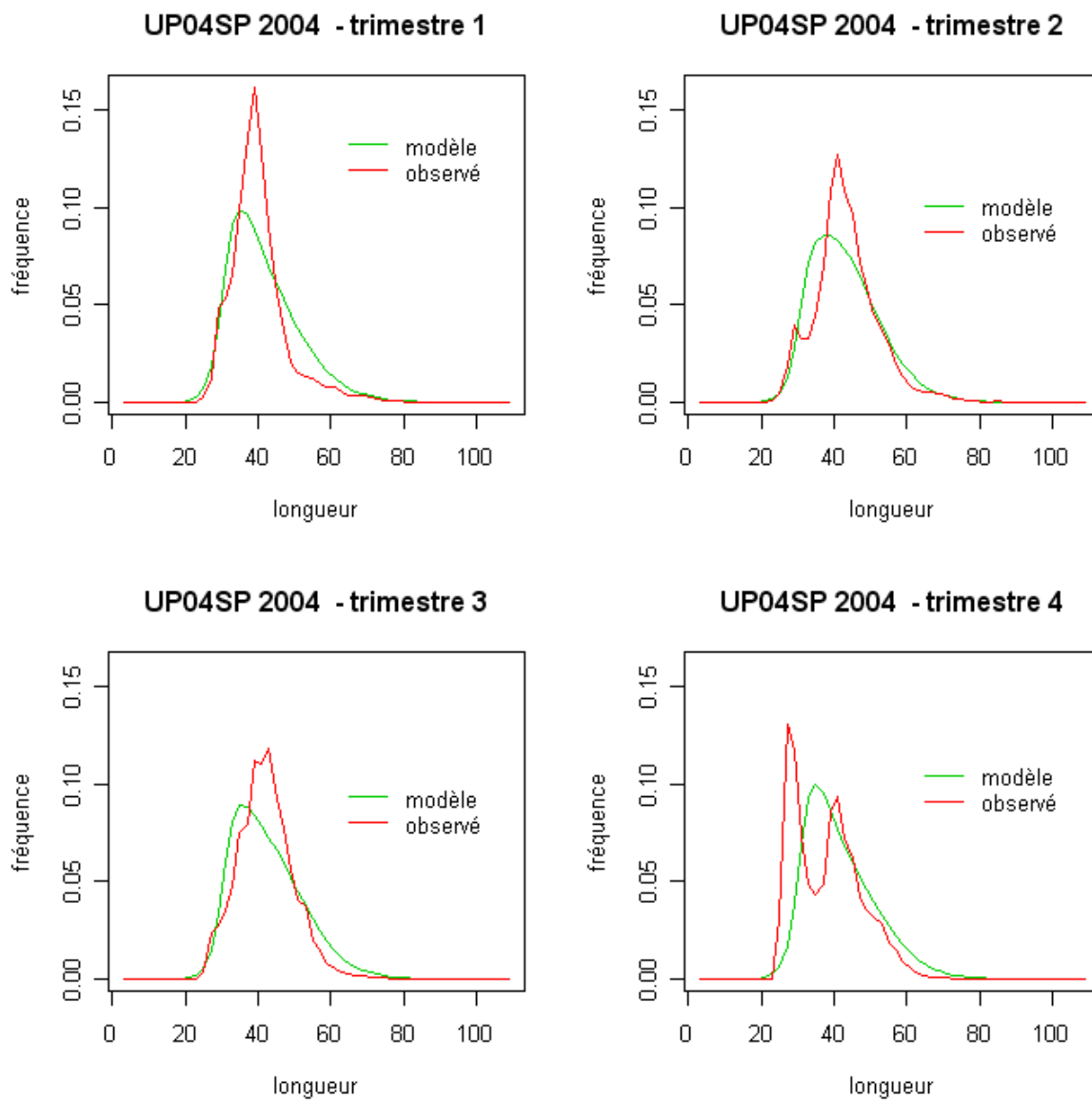


Figure 20 : Ajustement aux compositions en longueur trimestrielles des débarquements de la sous-unité de pêche chalutière espagnole non langoustinière opérant sur des fonds de moyenne à forte profondeur en mer Celtique en 2004.

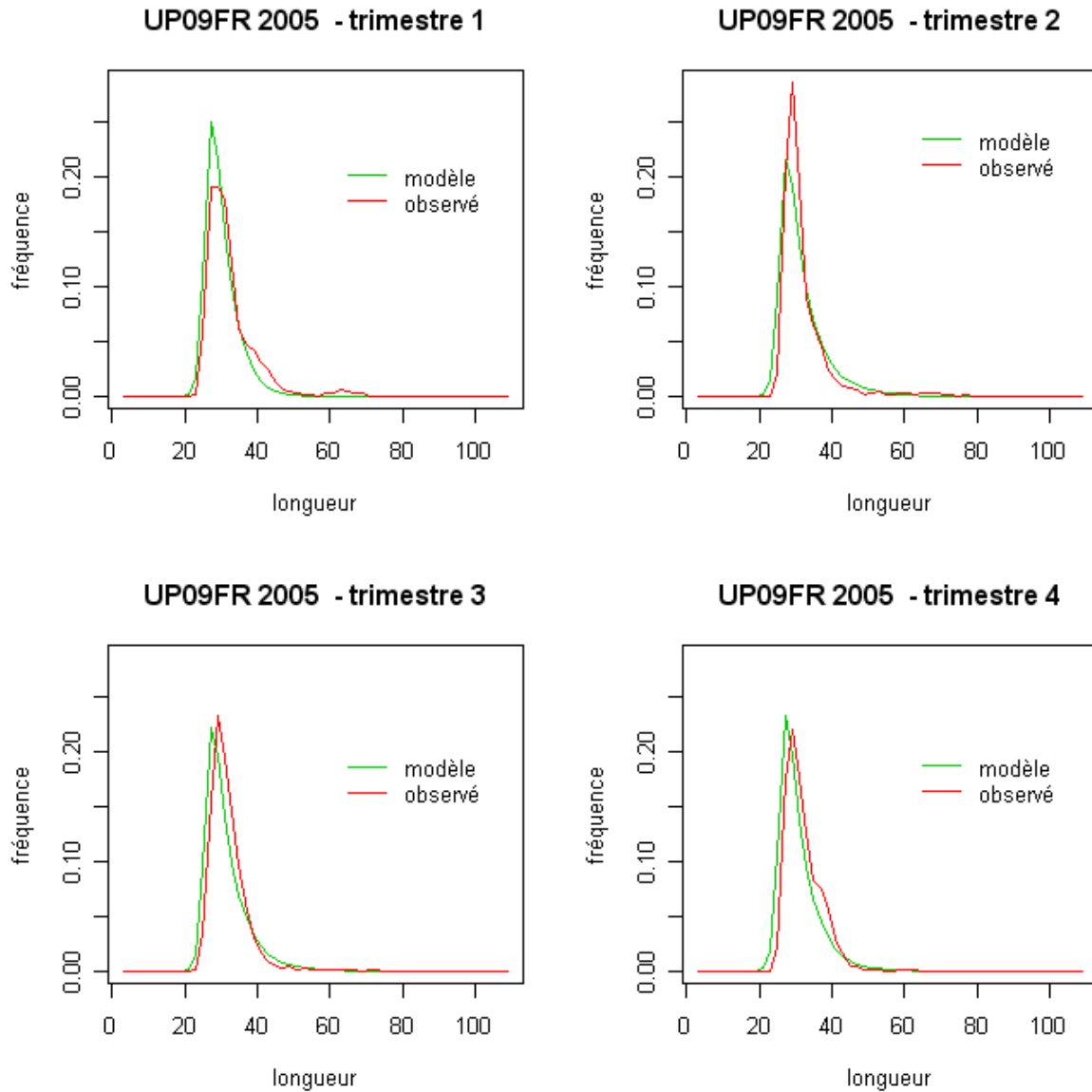


Figure 21 : Ajustement aux compositions en longueur trimestrielles des débarquements de la sous-unité de pêche chalutière langoustinière française opérant sur la « Grande Vasière » en 2005.

4.3.1.2 *Ajustements aux débarquements totaux en poids*

Si les ajustements aux compositions en taille des débarquements en poids sont bons, les débarquements totaux sont en revanche moins bien ajustés. Ces données ayant un poids plus faible dans l'ajustement (Tableau 13), il a fallu « artificiellement » leur donner un poids supérieur dans une première phase de l'ajustement en réalisant un ajustement en deux phases.

On observe quatre différents types d'ajustement (Tableau 14) selon les sous-unités de pêche (Tableau 10):

- les « bons ajustements » (ex : Figure 22 droite) pour lesquels les ordres de grandeur et les pics saisonniers de débarquements sont bien reproduits.
- pour certaines sous-unités de pêche, les ordres de grandeur sont bien reproduits, en revanche on observe un décalage saisonnier dans les pics de débarquements simulés et observés (ex : Figure 22 gauche). Cela s'explique sans doute par l'asynchronisme des pics saisonniers de captures par unité d'effort entre sous-unités de pêche pêchant dans la même zone, voire même de la même unité de pêche (Figure 23) ce qui relativise les estimations des paramètres de migrations.
- pour d'autres sous-unités de pêche, les pics saisonniers sont bien positionnés, en revanche l'ordre de grandeur n'est pas bon. Ce sont pour la plupart des flottilles avec des faibles débarquements observés.
- enfin pour trois sous-unités, l'ajustement n'est pas bon du tout. Ce sont des sous-unités pêchant dans les zones de nurricerie de la mer Celtique (UP05FR) et du golfe de Gascogne (UP13FR et UP13SP) et y capturant de grands merlus alors que les autres sous-unités pêchant dans la même zone n'en capturent presque pas. Le modèle n'arrive alors pas à reproduire simultanément leurs compositions en longueurs et leurs débarquements totaux en poids.

Tableau 14 : Tableau récapitulant les types d'ajustement pour les différentes sous-unités de pêche (Tableau 10).

Sous-unités de pêche	Bien ajustées	Bon ordre de grandeur mais décalage saisonnier des pics	Mauvais ordre de grandeur mais pics bien positionnés	Mauvais ajustements
UP01SP	X			
UP01EW	X			
UP03SP	X			
UP03FR	X			
UP03EW	X			
UP04SP		X		
UP04FR	X			
UP04EW		X		
UP05EW	X			
UP05FR				X
UP08FR	X			
UP09FR	X			
UP10FR	X			
UP12SP			X	
UP12FR			X	
UP13SP		X		X
UP13FR		X		X
UP14SP	X			

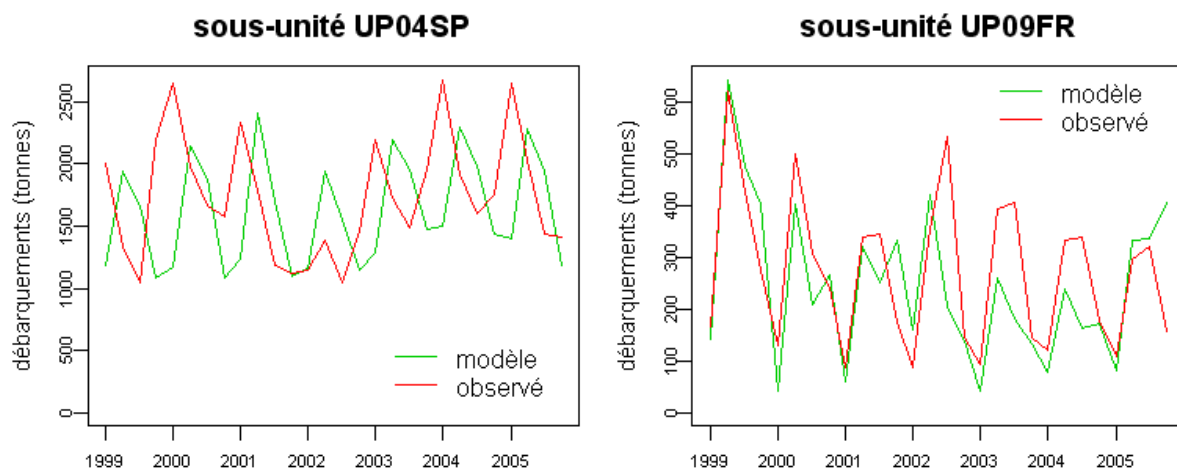


Figure 22 : Ajustement aux débarquements en poids de la sous-unité de pêche chalutière espagnole (UP04SP) non langoustinière opérant sur des fonds de moyenne à forte profondeur en mer Celtique (à gauche) et de la sous-unité de pêche langoustinière française (UP09FR) de la « Grande Vasière » (à droite).

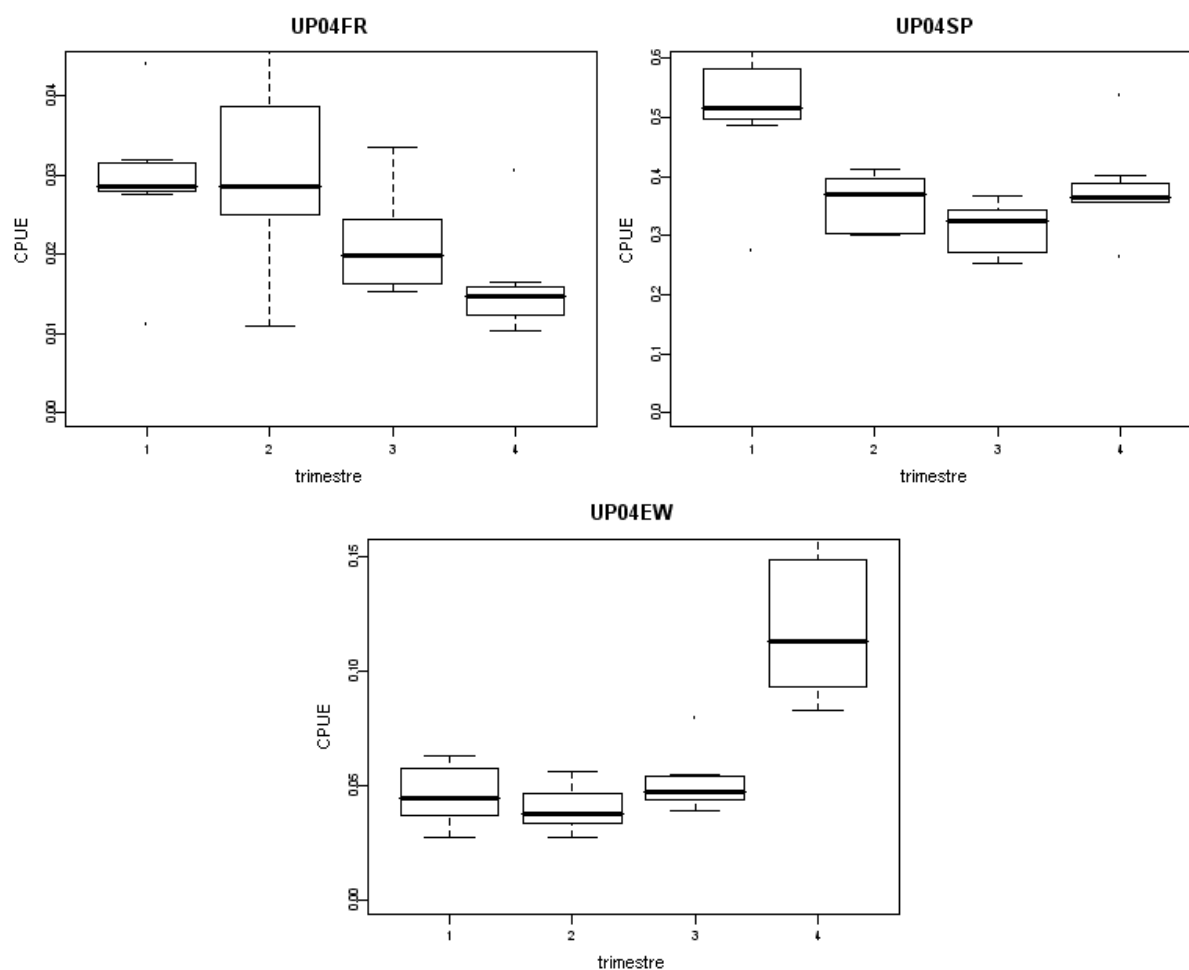


Figure 23 : Distributions des captures par unité d'effort trimestrielles observées de 1999 à 2006 pour les sous-unités de pêche française (en haut à gauche), espagnole (en haut à droite) et anglaise (en bas) de l'unité de pêche 04 (Tableau 2).

4.3.1.3 Ajustements aux indices d'abondance scientifiques

Tout comme les débarquements en poids, les indices d'abondance ont un poids assez faible dans l'ajustement. La qualité des ajustements est variable (Figure 24), les indices d'abondance de la campagne UK-WCGFS étant particulièrement mal reproduits (Tableau 13). Cependant on observe que les longueurs où l'on trouve des pics d'abondance correspondent assez bien aux observations.

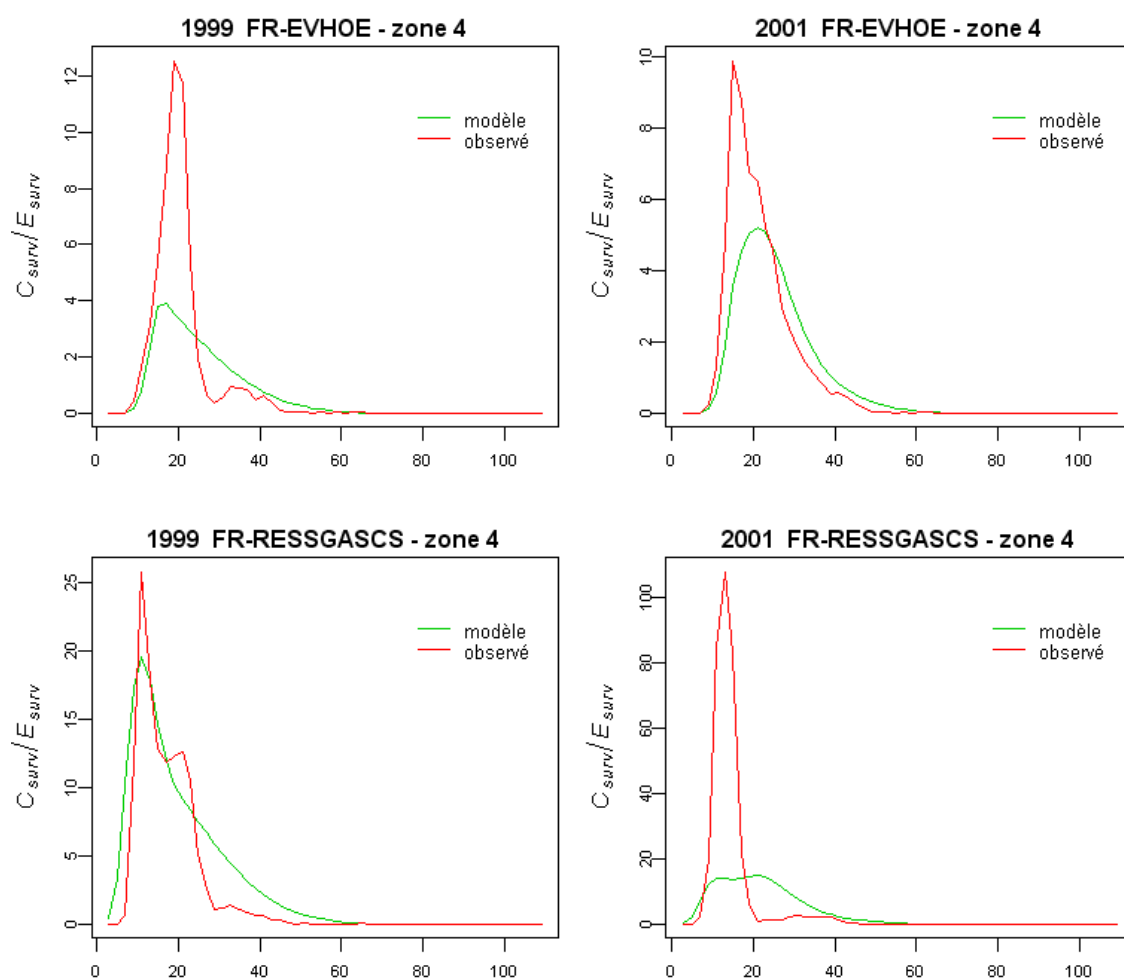


Figure 24 : Ajustements des indices d'abondance estimés par le modèle aux observations pour les campagnes FR-EVHOE (première ligne) et FR-RESSGASCS (deuxième ligne) en 1999 (première colonne) et 2001 (deuxième colonne).

4.3.2 Identifiabilité du modèle

En utilisant strictement la méthode détaillée par Reboulet basée sur les valeurs singulières de la matrice hessienne, on trouve qu'à priori seulement 4 paramètres sont réellement identifiables. Cependant il faut noter que le seuil de nullité (dans notre cas environ 10^8) choisi est basé sur la méthode proposé par Reboulet *et al.* (1999), et celui-ci est, à notre connaissance, arbitraire. L'analyse du Tableau 15 montre pourtant que le nombre de paramètres identifiables est très fortement dépendant du seuil choisi.

Tableau 15 : Nombre de valeurs singulières de la matrice hessienne par intervalles. Les cases grisées correspondent aux valeurs propres non nulles au seuil choisi, et donc à des paramètres théoriquement identifiables.

Intervalles	Nombres de valeurs singulières
[0;1]	2
[1;10]	8
[10;10 ²]	22
[10 ² ;10 ³]	24
[10 ³ ;10 ⁴]	15
[10 ⁴ ;10 ⁵]	4
[10 ⁵ ;10 ⁶]	3
[10 ⁶ ;10 ⁷]	1
[10 ⁷ ;10 ⁸]	3
[10 ⁸ ;10 ⁹]	0
[10 ⁹ ;10 ¹⁰]	0
[10 ¹⁰ ;10 ¹¹]	1
[10 ¹¹ ;10 ¹²]	0
[10 ¹² ;10 ¹³]	1
[10 ¹³ ;10 ¹⁴]	0
[10 ¹⁴ ;10 ¹⁵]	1
[10 ¹⁵ ;10 ¹⁶]	1

La matrice des corrélations montre assez peu de fortes corrélations entre les paramètres estimés. Les corrélations fortes (supérieures à 0,5) concernent:

- les recrutements interannuels pour une même zone de recrutement
- les recrutements du golfe de Gascogne et le paramètre de migration des immatures entre la zone 4 et la zone 3

- les recrutements de la mer Celtique et les capturabilités des sous-unités de pêche pêchant dans la zone de nurserie de la mer Celtique
- les paramètres de capturabilité et de dérive annuelle pour certaines sous-unités de pêche

Il existe donc des problèmes d'identifiabilité, toutefois ceux-ci semblent plus limités que ce que semblerait montrer l'analyse de la matrice hessienne. De plus, ils ne concernent que dans une faible mesure les paramètres principalement visés à savoir croissance et migration.

4.3.3 Evolution de la biomasse féconde et de sa distribution spatiale, évolution du recrutement

Le modèle permet de suivre la biomasse féconde au cours du temps. Il faut noter que la biomasse estimée est une biomasse relative puisque (i) la capturabilité de FR-EVHOE est fixée (ii) les captures sont ajustées. On a choisi de standardiser la biomasse féconde (SSB) par rapport à la SSB du premier pas de temps.

L'évolution de la SSB estimée par le modèle (Figure 25) est assez proche des résultats obtenus au cours de l'évaluation des stocks du CIEM (ICES, 2006b). Le modèle montre une augmentation récente de la biomasse, étayant l'hypothèse d'une amélioration de l'état du stock de puis le début des années 2000.

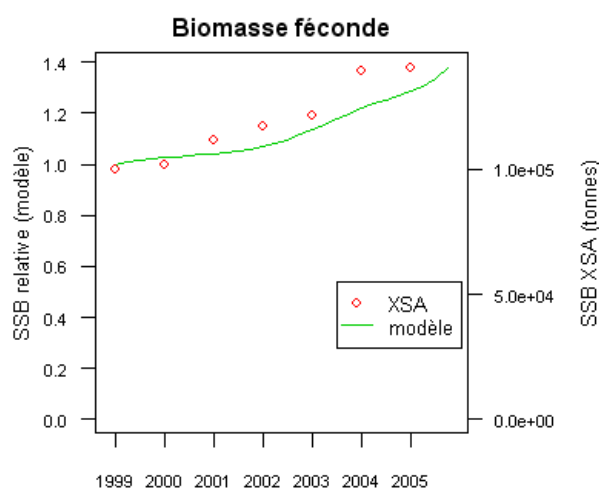


Figure 25 : Evolution de la biomasse féconde estimée par le modèle (en unité relative) et comparée au diagnostic réalisé par le CIEM (ICES, 2006b) à l'aide du modèle XSA.

Le modèle permet également d'obtenir des « cartes » de distribution de cette biomasse féconde pour chacun des pas de temps (Figure 26). On observe que cette distribution est particulièrement hétérogène. Elle est particulièrement concentrée sur le plateau de la mer Celtique. Le modèle ne reproduit par contre pas la concentration d'individus matures sur les accores du plateau continental au début de l'année, qui est pourtant mentionnée dans la littérature (Poulard et Léauté, 2002).

En ce qui concerne les migrations, le modèle prédit, comme on le supposait une forte migration du golfe de Gascogne vers la mer Celtique au cours du trimestre 2 ($0.56 \leq mmat_{3 \rightarrow 1}^2 \leq 0.58$) .

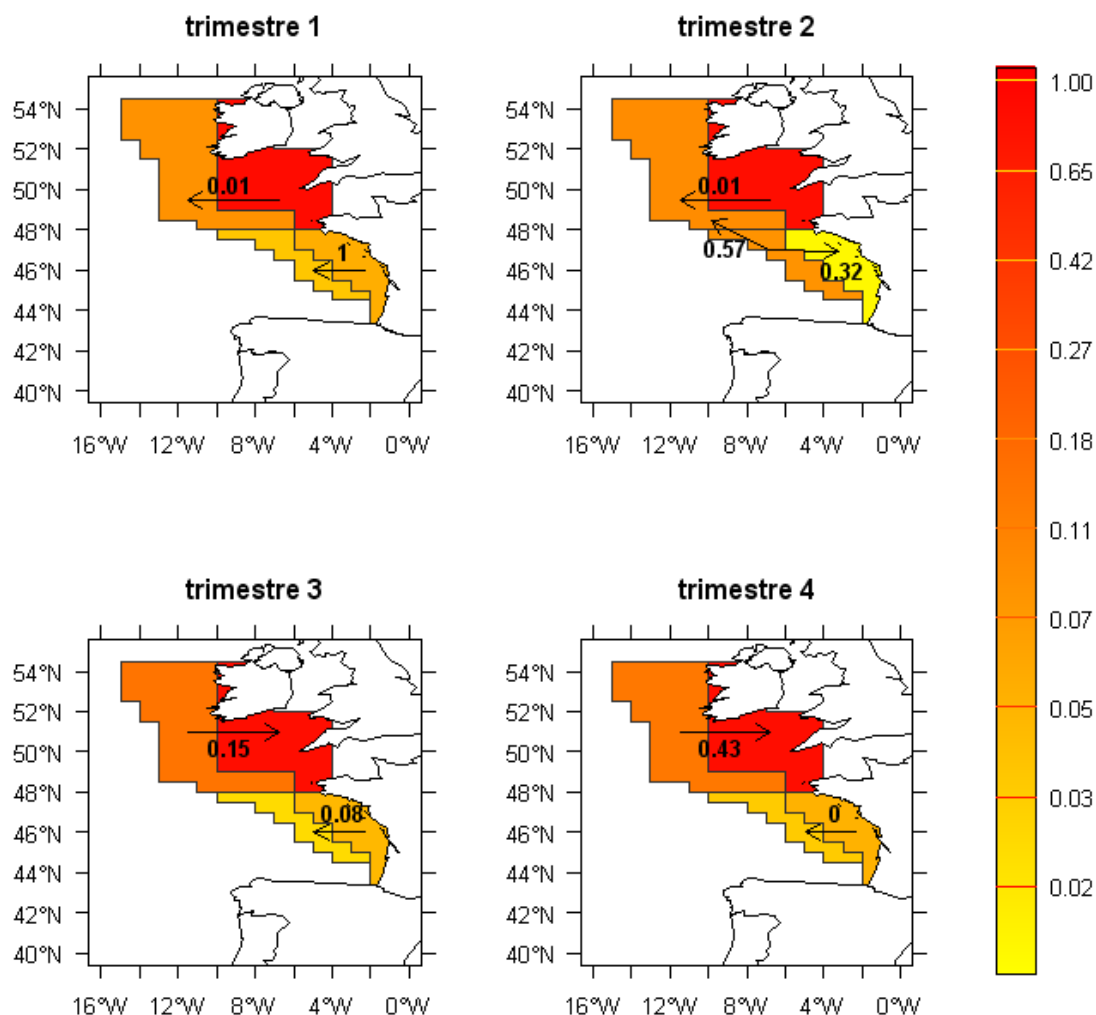


Figure 26 : Distribution de la biomasse féconde (en unité relative) par zone estimée par le modèle au début des trimestres de l'année 2001 et proportions estimées de matures migrant d'une vers une autre pour chacun des trimestres (flèches).

Si l'on s'intéresse au recrutement, le modèle diverge fortement des estimations obtenues par XSA (ICES, 2006b) que ce soit avec l'ancienne clé taille-âge ou avec la clé taille-âge simulée de Bertignac et de Pontual (2007), surtout pour l'année 2005, ce qui est également visible en regardant l'ajustement aux indices d'abondance dans les zones de recrutement pour cette

même année (Figure 27 et Figure 28). Cela peut s'expliquer par des anomalies dans les données d'effort 2005, notamment pour la sous-unité de pêche française UP05FR pêchant sur la zone de recrutement de la mer Celtique qui a un effort cinq fois supérieur en 2005 qu'en 2004 (plus de 8 fois au dernier trimestre). Il faut également souligner que les estimations du recrutement par XSA sont basées sur peu de points (uniquement sur l'indice d'abondance FR-EVHOE) et est donc considérée peu fiable lors de l'évaluation.

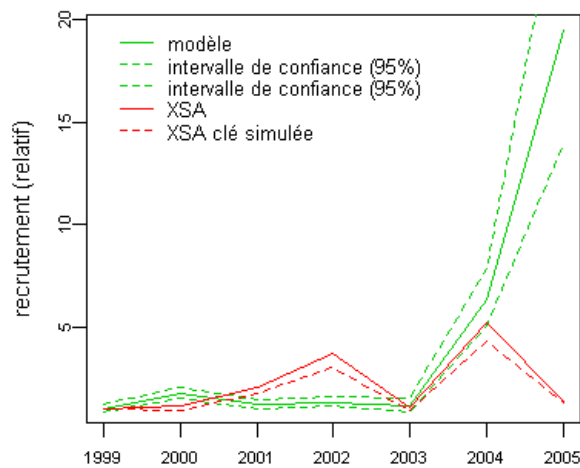


Figure 27 : Recrutement estimé (en vert) par le modèle (golfe de Gascogne et mer Celtique confondus) et son intervalle de confiance (pointillés verts), et recrutements estimés par XSA (ICES, 2006b) avec la clé taille âge classique (en rouge trait plein) et simulée (pointillés rouges). Le recrutement est standardisé par rapport au recrutement de 1999.

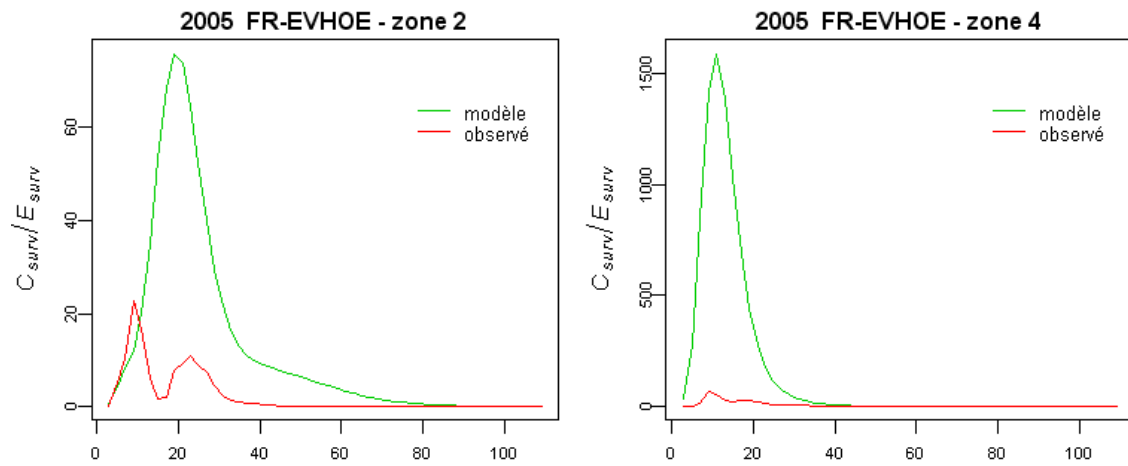


Figure 28 : Ajustement des indices d'abondance de FR-EVHOE en 2005 dans la zone de recrutement de la mer Celtique (à gauche) et dans la zone de recrutement du golfe de Gascogne (à droite).

4.3.4 Croissance estimée

Le résultat le plus étonnant du modèle est le taux de croissance estimé. Celui-ci est lent ($K=0,124 an^{-1} \pm 0,001$) comparé aux résultats obtenus à partir des données de marquage (de Pontual *et al.*, 2003 ; de Pontual *et al.*, 2006) et s'approche plus des anciennes courbes de croissance (Martin, 1991 ; Lucio *et al.*, 2000). La courbe de croissance résultante (Figure 29) est proche de celle estimée par Lucio *et al.* (2000).

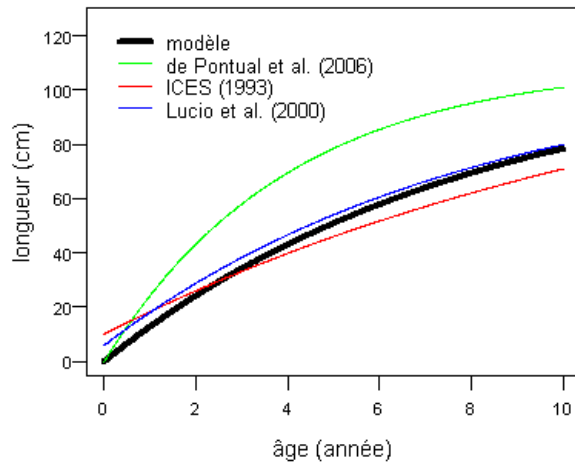


Figure 29. Courbe de croissance estimée par le modèle comparée aux modèles de différentes études.

4.4 Discussion

L'objectif de cette étude était le développement d'un modèle de dynamique de population destiné à (i) estimer certains paramètres inconnus de la dynamique de la population, notamment croissance et migration et (ii) et de permettre de dresser des diagnostics sur l'état du stock. Au cours du développement du modèle, une attention particulière a été portée sur l'écriture de la fonction de vraisemblance. En s'appuyant sur les modèles existants (voir chapitre 2), nous avons essayé de formuler les hypothèses les plus cohérentes avec les processus d'observations. Cela a notamment impliqué une séparation des données de débarquements totaux des compositions en longueur. Il a par contre été parfois nécessaire de réaliser des approximations (utilisation de la formulation 29 plutôt que 28) afin de rendre le modèle plus robuste. Le choix de respecter le plus possible les hypothèses requises pour l'application du maximum de vraisemblance, nous a également conduit à ne pas utiliser de pénalités et à ne pas pondérer les jeux de données (au moins dans la dernière phase de l'ajustement).

L'ajustement du modèle a permis d'obtenir des estimations du taux de croissance et des paramètres de migration. Ces résultats doivent toutefois être pris avec précaution. Les paramètres de migration apparaissent assez cohérents avec les connaissances à priori, mis à part l'absence de concentration saisonnière d'individus matures sur les accores du plateau de la mer Celtique. Pourtant, le fait que l'on observe des décalages saisonniers entre les pics de débarquement simulés et observés pour certaines sous-unités de pêche (Figure 22 gauche) modère ce résultat. En effet, ces décalages peuvent traduire une mauvaise estimation des migrations liée soit à des anomalies dans les données d'effort, soit à des interactions capturabilité \times saison que le modèle ne peut restituer (Figure 23). Pour les estimations des paramètres de croissance, les résultats sont plus surprenants et contrastent avec les estimations réalisées à partir des données de marquage. Pourtant les compositions en longueur des débarquements et les profils (ou au moins les pics) des indices d'abondance sont assez bien reproduits par le modèle. Cela est d'autant plus surprenant que le taux de mortalité choisi présupposait une croissance rapide. Si l'on fixe la mortalité naturelle à $0,2 \text{ an}^{-1}$, le taux de croissance estimé est alors encore plus faible (environ $0,08 \text{ an}^{-1}$). Le modèle a permis également d'obtenir un diagnostic en terme d'évolution de la biomasse féconde. Celui-ci est assez cohérent avec les diagnostics précédemment établis à partir du modèle XSA.

Plusieurs facteurs peuvent expliquer ces résultats.

- Choix de fixer L_{∞}

Comme dans le chapitre 3, nous avons choisi de considérer que L_{∞} était connu et de n'estimer que le taux de croissance afin d'éviter les problèmes liés à la corrélation négative entre ces deux paramètres. Cela permet également de directement comparer nos estimations de taux de croissance à celles obtenues dans d'autres études (notamment de marquage), qui avaient fait la même hypothèse. Toutefois la valeur de 110 cm est probablement forte au vu des tailles maximales dans les données disponibles et un choix de L_{∞} plus faible aurait vraisemblablement conduit à des estimations de taux de croissance plus fortes. À terme, on peut envisager l'utilisation de méthodes permettant l'estimation conjointe de K et L_{∞} , par

exemple en utilisant une transformation de ces deux paramètres limitant leur corrélation (Shackell *et al.*, 1997).

- Fiabilité des données

Tout d'abord la volonté d'estimer des paramètres de migration a rendu obligatoire l'utilisation des données d'effort : les paramètres de migration deviennent identifiables au travers des évolutions saisonnières des taux de captures par sous-unité de pêche. Or ces données d'effort sont supposées dans le modèle bien qu'elles puissent être particulièrement incertaines:

- Les données d'effort espagnoles et anglaises nous ont été fournies par des collègues européens qui ont émis une sérieuse réserve sur leur qualité.
- Les données françaises ont été reconstituées à partir des données livres de bord, qui ne sont que très partiellement remplis, et sont donc fortement sujettes à caution.

D'autres données, et en particulier les données de débarquements totaux en poids par trimestre et par sous-unité de pêche (la méthode utilisée pour les reconstituer est basée sur l'hypothèse que les livres de bord sont remplis de façon homogène et que toutes les sous-unités de pêche vendent équitablement en criée ce qui est discutable) sont assez peu fiables. Toutefois, contrairement au modèle XSA, ces données ne sont pas considérées comme exactes. De plus, il est intéressant de noter que les données qui contribuent le plus aux ajustements et sont les mieux ajustées sont les données de composition en longueur des débarquements qui sont à priori les données les plus fiables et celles qui apportent de l'information sur la croissance.

Pour intégrer cette hétérogénéité dans la qualité des données, on aurait pu faire le choix de fixer des pondérations dans la fonction objectif en fonction des niveaux de confiance que l'on avait sur les différents jeux de données. Cette pondération aurait pu faire l'objet de discussions avec les experts. Le risque alors aurait concerné l'exploitation statistique des résultats: signification de la vraisemblance, étude de l'identifiabilité...

- Sensibilité aux effectifs initiaux

D'autre part, nous n'avons pas pu reconstituer des données d'effort avant 1999, ce qui a limité la période d'ajustement à 1999-2005. Pour une si courte période, les effectifs initiaux sont particulièrement sensibles sur l'estimation des paramètres du modèle. Or ces derniers ont été estimés en faisant une hypothèse de système à l'équilibre et ne sont donc probablement pas corrects. A ce titre, l'utilisation d'un modèle structuré en âge et en longueur pourrait être intéressant: dans ce cas, il suffirait d'estimer les effectifs initiaux par groupe d'âge (donc seulement 8 paramètres supplémentaires si on considère huit groupes d'âge) et faire une hypothèse sur la distribution en longueur de ces groupes d'âge.

- Complexité

Outre la qualité des données et l'hypothèse sur les effectifs initiaux, le choix de spatialiser le modèle a impliqué une augmentation du niveau de complexité du modèle, à la fois du sous-modèle d'activité de pêche (décomposition de l'activité de pêche au niveau de la sous-unité de pêche), et du sous-modèle de dynamique des populations (modélisation du processus de migration, spatialisation du recrutement). En dehors des inconvénients pratiques que cela a impliqué (temps d'ajustement), cela a pu avoir des conséquences sur l'identifiabilité et la qualité des estimations. Même si les paramètres d'intérêt (croissance et migration) ne semblent pas concernés par ces problèmes, un grand nombre d'estimations de paramètres du modèles sont sujettes à caution au regard de l'analyse des valeurs propres de la matrice hessienne et de l'analyse de la matrice des corrélations. D'autre part, complexifier le modèle (rajouter des paramètres) entraîne une augmentation de la dimension du domaine de définitions des paramètres et complique la recherche de maximum de la fonction d'objectif. Une des grosses difficultés rencontrée lors de l'optimisation d'une telle fonction dans un domaine a aussi grande dimension est l'existence d'optima locaux. Parcourir l'ensemble du domaine est très coûteux en temps et selon la stratégie mise en oeuvre pour l'explorer et le point initial choisi pour démarrer l'algorithme d'optimisation, le risque de converger vers un optimum local peut être élevé. C'est d'ailleurs ce que nous avons expérimenté en réalisant d'un

côté un ajustement direct de la fonction d'objectif et d'un autre côté un ajustement en 2 étapes qui ne convergeaient pas vers les mêmes solutions. Des essais avec des algorithmes génétiques ou à stratégie évolutive, réputés mieux explorer le domaine de recherche et donc être moins sensibles à l'existence de minima locaux n'ont pas donné de résultats satisfaisants. Il est intéressant de noter que les estimations de taux de croissance obtenues lorsque l'on réalise des ajustements en une seule phase (sans aucune pondération) sont plus proches des estimations obtenues à partir des données de marquage ($K=0,178 \text{ an}^{-1}$). En revanche, les débarquements en poids ne s'ajustent plus du tout pour certaines sous-unités de pêche.

Du fait du temps nécessaire pour ajuster le modèle, nous avons également été contraint de simplifier le modèle de croissance (classes de 2cm et individus mesurant tous la taille moyenne de la classe). Toutefois, en optimisant le code du programme, il est probable que l'on puisse appliquer toutes les hypothèses recommandées par l'analyse de sensibilité (chapitre 3). Cependant, les résultats devraient être peu modifiés puisque la distribution des individus dans la classe de longueur n'avait qu'un impact limité, et que avec des classes de 2cm, l'analyse de la matrice de transition montre que les probabilités de ne pas changer de classe au cours d'un pas de temps restent assez faibles (elle n'est que de 50% pour des individus de 60cm avec le taux de croissance estimé).

Un dernier aspect est que certaines sous-unités de pêche ont un impact fort sur l'ajustement. C'est en particulier le cas pour les sous-unités de pêche de l'unité de pêche 13 qui débarquent de grands merlus en grande quantité (12% du débarquement total en poids en 2004) sur le plateau continental du golfe de Gascogne, qui est dans le modèle une zone de nourricerie, alors que les autres sous-unités pêchant dans la même zone capturent essentiellement des petits merlus. Il en va de même, dans une moindre mesure, pour les sous-unités de pêche (en particulier la sous-unité de pêche anglaise) de l'unité de pêche 5 qui pêche sur le plateau de la mer Celtique.

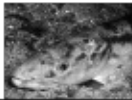
La spatialisation du modèle a permis de mettre en évidence l'hétérogénéité de la distribution spatiale de la biomasse féconde. Cela remet en question la validité des diagnostics réalisés à l'aide de modèle non spatialisé, basé sur une hypothèse d'homogénéité qui n'est pas du tout vérifiée. Les cartes de distribution spatiale de la biomasse féconde et l'estimation des paramètres de migration sont particulièrement importants pour l'évaluation de mesures de gestion spatialisées et pourront être intégrés au travail de Drouineau *et al.* (2006) visant à évaluer l'impact de scénarios de gestion sur la pêche mixte merlu langoustine du golfe de Gascogne (voir annexe 3).

Les données de marquage, encore trop peu nombreuses à ce jour pour être intégrées dans le modèle, pourraient constituer à l'avenir une source de données intéressante pour l'estimation du taux de croissance et des migrations. Cependant cela impliquerait l'organisation d'une campagne à plus grande échelle. Un modèle décrivant la dynamique de la population marquée (Bertignac *et al.*, 1999), couplé au modèle général, permettrait alors de prendre en compte ces données. Une autre alternative consisterait à utiliser les données de marquage extérieurement au modèle (ou dans une phase préliminaire de l'ajustement du modèle comme (DeLong *et al.*, 2001)) pour estimer la matrice de croissance. Cette seconde solution permettrait de diminuer le nombre de paramètres à estimer et donc faciliterait l'ajustement du modèle.

Un développement futur envisageable est l'intégration d'effets aléatoires, ou l'orientation du modèle vers une approche bayésienne, afin de mieux prendre en compte l'incertitude sur certains processus de la dynamique de la population (en particulier le recrutement) et de l'activité de pêche (dérive de la capturabilité). Cela a été envisagé au cours de ce travail, malheureusement cela n'est pas encore possible d'un point de vue méthodologique, la dimension de notre modèle étant trop élevée (voir annexe 4). Toutefois, les développements ultérieurs du modèle ne seront réellement intéressants que sous réserve d'une amélioration de la qualité des données. A ce titre, le suivi de l'activité de pêche par calendriers d'activité effectué par l'IFREMER et peut-être à l'avenir l'utilisation de données VMS pourront constituer des sources de données importantes.

Ce qu'il faut retenir

- un modèle d'estimation de paramètres de dynamique de population structuré en longueur et spatialisé
- pas de problème d'identifiabilité pour les paramètres de croissance et de migration
- un ajustement difficile du modèle aux observations pouvant s'expliquer par la complexité du modèle et la mauvaise qualité des données d'effort de pêche
- un taux de croissance pour le merlu égal à $0,124 \text{ an}^{-1}$ plus faible que celui estimé à partir des données scientifiques de marquage
- des migrations cohérentes avec la connaissance expert : départ de matures du golfe de Gascogne vers la mer Celtique au printemps
- une évolution de biomasse estimée par le modèle proche des estimations XSA



Discussion générale et perspectives



All models are wrong, some are useful (George Box)

Discussion générale et perspectives

En dépit de l'importante littérature traitant de la biologie du merlu, il demeure aujourd'hui certaines incertitudes sur cette espèce, notamment sur sa croissance et sa distribution spatio-temporelle. Ces incertitudes peuvent potentiellement affecter la qualité des diagnostics établis par les modèles de dynamique des populations usuels et limitent les possibilités d'évaluation de mesures de gestion spatialisées. Comme l'a montré la revue bibliographique du chapitre 2, les modèles structurés en longueur constituent une alternative intéressante aux modèles usuels quand, comme c'est le cas pour le merlu, la conversion taille-âge est incertaine. C'est ce qui nous a conduit à développer un modèle structuré en longueur et spatialisé pour cette espèce avec comme objectif d'améliorer la connaissance sur les migrations et sur la croissance, et d'établir des diagnostics sur l'état du stock et de son exploitation.

Une première étape, préalable au développement du modèle complet, a consisté en une analyse de l'impact de certaines hypothèses de discrétisation sur le modèle de croissance (voir chapitre 3). Ce type d'analyse est indispensable pour formuler des hypothèses de modélisation robustes et détecter d'éventuels artefacts de modélisation, et dans notre cas nous a guidés dans le choix des classes de longueur, du pas de temps et de la distribution des incréments de croissance du modèle final. La méthodologie que nous avons développée repose notamment sur l'utilisation des plans d'expériences (dans notre cas un plan complet, mais selon les situations d'autres types de plans peuvent être plus adaptés). Bien que l'utilisation de plans d'expériences pour l'analyse numérique de modèles, et plus particulièrement leur analyse de sensibilité, soit déjà répandue dans certains domaines, comme l'industrie (Ivanova *et al.*, 1999) ou l'agronomie (Henderson-Sellers et Henderson-Sellers, 1996) elle l'est assez peu en halieutique (à notre connaissance seuls Drouineau *et al.* (2006), Ginot *et al.* (2006) et Cariboni *et al.* (2007) les ont utilisés). Ils permettent de quantifier simultanément les effets des différents facteurs et leurs interactions, contrairement aux classiques analyses d'élasticité qui se limitent à l'analyse des effets des paramètres indépendamment les uns des autres. Les résultats des plans d'expériences peuvent être analysés grâce à des modèles statistiques usuels, comme la régression linéaire ou l'analyse de variance, ou plus complexes comme les

régressions PLS. L'arbre de régressions multiples est un outil complémentaire intéressant afin d'explorer simultanément les effets des variables explicatrices sur différentes variables à expliquer. Bien que ces résultats ne soient pas génériques, la méthode que nous avons développée peut aisément être transposée à toute analyse de modèle discret décrivant des processus continus.

Pour le développement du modèle, une attention particulière a été portée à l'écriture de la fonction de vraisemblance et aux algorithmes d'optimisation. A partir d'une revue bibliographique sur les ajustements de modèles structurés en longueur (section 2.6.1), nous avons cherché à construire la fonction de vraisemblance la plus robuste et la mieux adaptée à nos données et à leurs méthodes d'acquisition. Nous avons également choisi de ne pas pondérer les données et d'éviter l'utilisation de pénalités visant à prendre en compte les incertitudes sur les processus de la dynamique, car ces méthodes rendent discutables la signification de la vraisemblance et donc des propriétés statistiques qui en découlent (intervalles de confiance, AIC, BIC, identifiabilité).

Au vu de la complexité du domaine de définition des paramètres à explorer pour maximiser la fonction de vraisemblance, il nous a fallu faire particulièrement attention au choix de l'algorithme d'optimisation ; un compromis devait être trouvé entre le temps de simulation et l'exploration du domaine de recherche (afin d'éviter les optima locaux). Nous avons opté pour un algorithme de type quasi-Newton couplé à un algorithme de différentiation automatique pour deux raisons principales: (i) son efficacité en termes de vitesse de calculs a été démontrée dans différentes études (Fournier *et al.*, 1998 ; Maunder et Watters, 2003), (ii) il permet d'obtenir facilement la matrice hessienne de la fonction de logvraisemblance et donc la matrice de corrélation des paramètres, et permet ainsi d'étudier l'identifiabilité des paramètres. En revanche, ce type d'algorithme est problématique en cas d'existence d'optima locaux. A ce titre, les algorithmes génétiques ou de stratégie évolutive (Schwefel, 1995), réputés assurer une très bonne exploration du domaine de recherche, semblent des options intéressantes. Les essais d'ajustements que nous avons obtenus avec des algorithmes génétiques n'ont cependant pas donné les résultats escomptés, probablement du

fait d'un mauvais choix de paramétrage (taille de la population génétique, probabilité de mutation, nombre de générations). Il n'en reste pas moins que cette famille d'algorithmes pourraient constituer une alternative intéressante pour les ajustements de modèle.

Au vu de la complexité de notre modèle et donc du nombre important de paramètres à estimer, ce qui est le cas de la plupart des modèles structurés en longueur dans la littérature, il nous est apparu indispensable de s'assurer de l'identifiabilité du modèle. Si la comparaison des performances des modèles est souvent réalisée, celle-ci repose en général sur des critères type AIC et BIC, qui nécessitent théoriquement la connaissance du nombre de paramètres réellement identifiables dans le modèle (Viallefont *et al.*, 1998 ; Reboulet *et al.*, 1999 ; Gimenez *et al.*, 2004), or celui-ci n'est que rarement connu. Pour évaluer l'identifiabilité des paramètres, différentes méthodes peuvent-être utilisées (voir annexe 2), chacune présentant des avantages et des inconvénients. Nous avons opté pour une méthode basée sur l'analyse spectrale de la matrice hessienne et l'analyse de la matrice de corrélation des paramètres. Le nombre de paramètres redondants estimés par cette méthode est très faible. Toutefois le seuil de nullité, basé sur un critère arbitraire, est apparu dans notre cas très élevé au regard des résultats de l'analyse de la matrice de corrélation. Cette dernière a montré que peu de paramètres du modèle étaient corrélés entre eux.

Le modèle final a permis d'obtenir une estimation des taux de croissance et de migration. Si les migrations estimées paraissent relativement cohérentes avec les connaissances a priori, le taux de croissance estimé par le modèle est en revanche très différent des estimations obtenues avec les données de marquage (de Pontual *et al.*, 2003 ; de Pontual *et al.*, 2006) mais en accord avec des taux disponibles dans la littérature (Martin, 1991 ; Lucio *et al.*, 2000). Le modèle permet également d'obtenir une estimation de l'évolution de la biomasse féconde, cohérente avec celle obtenue par XSA, et une estimation du recrutement, qui est par contre assez divergente des estimation XSA. Les objectifs du modèle semblent donc satisfaits mais les résultats doivent être pris avec précaution : la série historique de données à laquelle s'ajuste le modèle est courte, il demeure des problèmes d'identifiabilité (qui concernent cependant peu les paramètres d'intérêt) et certaines flottilles sont mal ajustées.

A l'avenir, l'extension de la campagne de marquage à une plus grande échelle spatiale et saisonnière pourrait permettre de valider les différentes hypothèses de croissance et de migration existantes. Les données pourraient être directement intégrées dans le modèle, ou être utilisées en préalable pour estimer les paramètres de croissance et de migration séparément du reste du modèle (qui de fait deviendrait plus un modèle d'évaluation de stocks). Le travail actuellement mené au LASAA (IFREMER Brest) visant à mieux comprendre la formation des otolithes de merlu devrait également à terme améliorer les techniques de lecture d'âge et donc la connaissance sur la croissance de cette espèce. Cela pourrait éventuellement permettre de s'orienter vers des modèles structurés en âge et en longueur comme Multifan-CL ou Fleksibest, un des intérêts étant de pouvoir faire des hypothèses plus souples sur les effectifs initiaux.

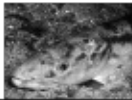
Comme déjà mentionné, nous avons choisi de négliger la stochasticité de certains processus de la dynamique (notamment recrutement et dérive de la capturabilité). Les modèles intégrés à espace d'états permettent pourtant théoriquement de combiner erreurs de processus et erreurs d'observations, mais l'estimation par maximum de vraisemblance est alors techniquement difficile (voir annexe 4) (de Valpine et Hastings, 2002 ; de Valpine, 2003 ; de Valpine et Hilborn, 2005). Aucune méthode existante ne permet de réaliser des ajustements dans un temps raisonnable pour des modèles de grande complexité tel que le nôtre. Le développement de ce type de méthodes devrait ouvrir à l'avenir des perspectives intéressantes pour le développement de modèles à effets aléatoires ou de modèles bayésiens complexes.

Bien que notre modèle prenne en compte l'incertitude entourant certaines données, il n'en reste pas moins que la qualité des données, en particulier des données d'effort, reste un problème majeur et pourrait potentiellement expliquer certains des problèmes que nous avons rencontrés lors de son ajustement. Ce type d'analyse pointe les problèmes de cohérence entre observations et met en exergue les problèmes de fiabilité de la donnée. La mise en place des calendriers d'activité et peut-être l'utilisation future des données VMS ouvrent à ce titre des perspectives intéressantes (Système d'Informations Halieutiques: <http://www.ifremer.fr/sih>).

De même, la mise en place des observations embarquées, notamment sur la pêcherie langoustinière du golfe de Gascogne, principale unité de pêcherie génératrice de rejets. De futurs développements méthodologiques et une complexification de notre modèle (plus généralement des modèles de dynamiques des populations visant à réaliser des évaluations de stocks ou à l'estimation de paramètres) ne seront réellement intéressants que sous réserve d'une amélioration de la qualité des données disponibles, en particulier des données liées à l'activité de pêche.

L'estimation des paramètres du modèle que l'on présente ici ne peut-être considérée comme validée, le modèle peut déjà de par sa flexibilité, servir de base au développement de modèles de dynamiques des populations structurés en longueur pour d'autres espèces dont la conversion taille âge des données est incertaine, comme la langoustine ou la baudroie.

Enfin, améliorer la connaissance sur la dynamique de population spatialisée et structurée en longueur ouvre la perspective d'une meilleure évaluation de l'impact des mesures de gestion régulant la sélectivité des engins et celles régulant spatialement l'accès aux ressources marines. Des mesures de gestion comme les aires marine protégées sont ainsi de plus en plus souvent préconisées pour une gestion durable des écosystèmes (Botsford *et al.*, 1997 ; Holland, 2000 ; Holland, 2003). L'efficacité de telles mesures dépend fortement de leur localisation, de leur saisonnalité (Sumaila *et al.*, 2000 ; Jamieson et Levings, 2001) et des interactions avec les autres mesures de gestion (Holland, 2003). L'évaluation des impacts de telles mesures, notamment à l'aide de simulateurs devient donc essentiel pour la gestion des pêcheries (Gulland, 1983 ; Hilborn et Walters, 1992). Les résultats d'ajustement de notre modèle peuvent être par exemple intégrés avec l'ensemble de la connaissance sur la pêcherie dans le modèle de simulation ISIS-Fish (Pelletier *et al.*, 2001 ; Mahévas et Pelletier, 2004 ; Pelletier *et al.*, submitted) afin d'évaluer l'impact de mesures de gestion spatialisées sur la pêcherie de merlu. A ce titre, une mise à jour du paramétrage du cas d'étude de Drouineau *et al.* (2006) présentée en annexe 3 permettrait d'évaluer l'impact d'une aire marine protégée sur la nourricerie de merlu du golfe de Gascogne ou de la mer Celtique.



Références



Références

- Abbes, R. 1991. Atlas des ressources et des pêches françaises dans les mers européennes. Editions Ouest-France, Rennes, Rennes. 86pp.
- Akaike, H. 1973. Information theory as an extension of the maximum likelihood principle. *In* 2nd International Symposium on Information Theory. Petrov, B. N., et Csaksi, F. (Eds.). Akademiai Kiado, Budapest, Hungary. 267-281.
- Alvarez, P., Fives, J., Motos, L., et Santos, M. 2004. Distribution and abundance of European hake *Merluccius merluccius* (L.), eggs and larvae in the North East Atlantic waters in 1995 and 1998 in relation to hydrographic conditions. *Journal of Plankton Research*, 26: 811-826.
- Alvarez, P., Motos, L., Uriarte, A., et Engaña, J. 2001. Spatial and temporal distribution of European hake, *Merluccius merluccius* (L.), eggs and larvae in relation to hydrographical conditions in the Bay of Biscay. *Fisheries Research*, 50: 111-128.
- Banks, H. T., Botsford, L. W., Kappel, F., et Wang, C. 1991. Estimation of growth and survival in size-structured cohort data: an application to larval striped bass (*Morone saxatilis*). *Journal of Mathematical Biology*, 30: 125-150.
- Begley, J., et Howell, D., 2004. An overview of Gadget, the Globally applicable Area-Disaggregated General Ecosystem Toolbox. ICESCM2004/FF:13pp.
- Belloc, G. 1935. Etude monographique du merlu (*Merluccius merluccius* L.) 3ème partie. *Revue Trav. Off. (Scienc. Tech.) Pêch. Mar.*, 8: 145-202.
- Bertignac, M., et de Pontual, H. 2007. Consequences of bias in age-estimation on the assessment of the Northern stock of European hake (*Merluccius merluccius*) and on management advice. *ICES Journal of Marine Science*, 64: 981-988.
- Bertignac, M., Hampton, J., et Coan, A. L. J. 1999. Estimates of exploitation rates for north Pacific albacore, *Thunnus alalunga*, from tagging data. *Fishery Bulletin*, 97: 421-433.

- Biseau, A. 2003. Les nouvelles de l'IFREMER: Le Merlu (*Merluccius merluccius*) stock Nord (divisions IIIa, IV, V, VI, VII et VIIIabde du CIEM). *Le Marin*, 46: 4pp.
- Bogstadt, B., Howell, D., et Asnes, M. N., 2004. A closed life-cycle model for the Northeast Arctic cod. ICES C.M. 2004/K:26pp.
- Botsford, L., Castilla, J., et Peterson, C. 1997. The management of fisheries and marine ecosystems. *Science*, 277: 509-515.
- Box, G., et Jenkins, G. M. 1970. *Time-Series Analysis: Forecasting and Control*. Holden-Day, London. 575pp.
- Breen, P. A., Kim, S. W., et Andrew, N. L. 2003. A length-based Bayesian stock assessment model for the New Zealand abalone *Haliotis iris*. *Marine Freshwater Research*, 54: 619-634.
- Burnham, K. P., et Anderson, D. R. 2002. *Model Selection and Multimodel Inference: A Practical Information Theoretic Approach*. Springer, New York. 496pp.
- Cariboni, J., Gatelli, D., Liska, R., et Saltelli, A. 2007. The role of sensitivity analysis in ecological modelling. *Ecological Modelling*, 203: 167-182.
- Casey, J., et Pereiro J 1995. European hake (*M. merluccius*) in the North-east Atlantic. *In* Hake fisheries, ecology and markets. Alheit, J., et Pitcher, T. J. (Eds.). Chapman & Hall, London. 125-147.
- Castillo, A. G. F., Martinez, J. L., et Garcia-Vasquez, E. 2004. Fine spatial structure of Atlantic hake (*Merluccius merluccius*) stocks revealed by variation at microsatellite loci. *Marine Biotechnology*, 6: 299-306.
- Caswell, H. 2000. *Matrix Population Models: Construction, Analysis, and Interpretation*. Sinauer Associates, Sunderland, Massachusetts. 710pp.
- Chen, Y., Hunter, M., Vadas, R., et Beal, B. 2003. Developing a growth-transition matrix for the stock assessment of the green sea urchin (*Strongylocentrotus droebachiensis*) off Maine. *Fishery Bulletin*, 101: 737-744.
- Cimmaruta, R., Bondanelli, P., et Nascetti, G. 2005. Genetic structure and environmental

-
- heterogeneity in the European hake (*Merluccius merluccius*). *Molecular Ecology*, 14: 2577-2591.
- Commission of the European Communities. 2001. Report of the Working Group on the preparation of the future recovery measures for Northern Hake. SEC(2001) 1193: 28pp.
- Commission of the European Communities. 2002. Comportements ayant gravement enfreint les règles de la politique commune des pêches en 2001. COM(2002)0687 final: 29pp.
- Commission of the European Communities. 2003. Respect des règles de la politique commune de la pêche: "plan et tableau de conformité". COM(2003)344 final: 67pp.
- Costanza, R., et Sklar, F. H. 1985. Articulation, accuracy and effectiveness of mathematical models: a review of freshwater wetland applications. *Ecological Modelling*, 27: 45-68.
- Cruywagen, G. 1997. The use of generalised linear modelling to determine inter-annual and inter-area variation of growth rates : the caperock lobster as example. *Fisheries Research*, 29: 119-131.
- De Leo, G. A., et Gatto, M. 1995. A size and age-structured model of the European eel (*Anguilla anguilla* L.). *Canadian Journal of Fisheries and Aquatic Sciences*, 52: 1351-1367.
- de Pontual, H., Bertignac, M., Battaglia, A., Bavouzet, G., Moguelet, P., et Groison, A. L. 2003. A pilot tagging experiment on European hake (*Merluccius merluccius*): methodology and preliminary results. *ICES Journal of Marine Science*, 60: 1318-1327.
- de Pontual, H., Groison, A. L., Pineiro, C., et Bertignac, M. 2006. Evidence of underestimation of European hake growth in the Bay of Biscay, and its relationship with bias in the agreed method of age estimation. *ICES Journal of Marine Science*, 63: 1674-1681.
- de Valpine, P. 2002. Review of methods for fitting time-series models with process and observation error and likelihood calculations for nonlinear, non-gaussian state space models. *Bulletin of Marine Science*, 70: 455-471.
- de Valpine, P. 2003. Better inferences from population-dynamics experiments using Monte
-

- Carlo state space likelihood. *Ecology*, 84: 3064-3077.
- de Valpine, P. 2004. Monte-Carlo state space likelihoods by weighted posterior kernel density estimation. *Journal of the American Statistical Association*, 99: 523-536.
- de Valpine, P., et Hastings, A. 2002. Fitting population models incorporating process noise and observation error. *Ecological Monographs*, 72: 57-76.
- de Valpine, P., et Hilborn, R. 2005. State-space likelihoods for nonlinear fisheries time-series. *Canadian Journal of Fisheries and Aquatic Sciences*, 62: 1937-1952.
- De'Ath, G. 2002. Multivariate Regression Trees: a new technique for modelling species-environment relationships. *Ecology*, 83: 1105-1117.
- DeLong, A. K., Collie, J. S., Meise, C. J., et Powell, J. C. 2001. Estimating growth and mortality of juvenile winter flounder, *Pseudopleuronectes americanus*, with a length-based model. *Canadian Journal of Fisheries and Aquatic Sciences*, 58: 2233-2246.
- Deriso, R. B., Maunder, M. N., et Skalski, J. R. 2007. Variance estimation in integrated assessment models and its importance for hypothesis testing. *Canadian Journal of Fisheries and Aquatic Sciences*, 64: 187-197.
- Deriso, R. B., T J Quinn II, et Neal, P. R. 1985. Catch-at-age analysis with auxiliary information. *Canadian Journal of Fisheries and Aquatic Sciences*, 42: 814-824.
- Drouineau, H., Mahévas, S., Pelletier, D., et Beliaeff, B. 2006. Assessing the impact of different management options using ISIS-Fish: the French *Merluccius merluccius* - *Nephrops norvegicus* mixed fishery of the Bay of Biscay. *Aquatic Living Resources*, 19: 15-29.
- Ennis, D. M., et Bi, J. 1998. The beta-binomial model: Accounting for intertrial variation in replicated difference and preference tests. *Journal of Sensory Analysis*, 13: 348-412.
- Eveson, J. P., Laslett, G. M., et Polacheck, T. 2004. An integrated model for growth incorporating tag-recapture, length frequency, and direct aging data. *Canadian Journal of Fisheries and Aquatic Sciences*, 61: 292-306.
- Fabens, A. J. 1965. Properties and fitting of the von Bertalanffy growth curve. *Growth*, 29:

265-289.

- Fives, J. M., Acevedo, S., Lloves, M., Whitaker, A., Robinson, M., et King, P. A. 2001. The distribution and abundance of larval mackerel, *Scomber scombrus* L., horse mackerel, *Trachurus trachurus* (L.), hake, *Merluccius merluccius* (L.), and blue whiting, *Micromesistius poutassou* (Risso, 1826) in the Celtic Sea and west of Ireland during the years 1986, 1989 and 1992. *Fisheries Research*, 50: 17-26.
- Fournier, D. A., Hampton, J., et Sibert, J. R. 1998. MULTIFAN-CL: a length-based, age-structured model for fisheries stock-assessment, with application to South Pacific albacore, *Thunnus alalunga*. *Canadian Journal of Fisheries and Aquatic Sciences*, 55: 2105-2116.
- Fournier, D. A., Sibert, J. R., Majkowski J, et Hampton H 1990. MULTIFAN a likelihood-based method for estimating growth parameters and age composition from multiple length frequency data sets illustrated using data for southern bluefin tuna (*Thunnus maccoyii*). *Canadian Journal of Fisheries and Aquatic Sciences*, 47: 301-317.
- Froysa, K. G., Bogstad, B., et Skagen, D. W. 2002. Fleksibest- an age-length structured fish stock assessment model. *Fisheries Research*, 55: 87-101.
- Fu, C., et T J Quin II 2000. Estimability of natural mortality and other population parameters in a length-based model: *Pandalus borealis* in Kachemak Bay, Alaska. *Canadian Journal of Fisheries and Aquatic Sciences*, 57: 2420-2432.
- Gilks, W. R., Richardson, S., et Spiegelhater, D. J. 1995. Markov chain Monte Carlo in practice. Chapman & Hall, London. 512pp.
- Gimenez, O., Viallefont, A., Catchpole, E. A., Choquet, R., et Morgan, B. J. T. 2004. Methods for investigating parameter redundancy. *Animal Biodiversity and Conservation*, 27: 561-572.
- Ginot, V., Gaba, S., Beaudouin, R., Aries, F., et Monod, H. 2006. Combined use of local and ANOVA-based global sensitivity analyses for the investigation of a stochastic dynamic model: Application to the case study of an individual-based model of a fish population. *Ecological Modelling*, 193: 479-491.
- Goldberg, D. E. 1989. Genetic Algorithms in Search, Optimization, and Machine Learning.
-

- Addison-Wesley Professional, Reading, Massachusetts. 432pp.
- Grift, R., Rijnsdorp, A., Barot, S., Heino, M., et Dieckmann, U. 2003. Fisheries-induced trends in reaction norms for maturation in North Sea plaice. *Marine Ecology-Progress Series*, 257: 247-257.
- Guichet, R. 1995. The diet of European hake (*Merluccius merluccius*) in the northern part of the Bay of Biscay. *ICES Journal of Marine Science*, 52: 21-31.
- Guichet, R.. 1996. Le merlu européen (*Meluccius merluccius* L.). RIDRV. 96 - 04: 55pp.
- Gulland, J. A. 1983. Fish stock assessment: A manual of basic methods. John Wiley and Sons, New York, New York. 223pp.
- Håkanson, L. 1995. Optimal size of predictive model. *Ecological Modelling*, 78: 195-204.
- Hall, S. J., Collie, J. S., Duplisea, D. E., Jennings, S., Bravington, M., et Link, J. 2006. A length-based multispecies model for evaluating community responses to fishing. *Canadian Journal of Fisheries and Aquatic Sciences*, 63: 1344-1359.
- Hampton, J., et Fournier, D. A. 2001. A spatially disaggregated, length-based, age-structured population model of yellowfin tuna (*Thunnus albacares*) in the western and central Pacific Ocean. *Marine Freshwater Research*, 52: 937-963.
- Harrington, E. C. J. 1965. The desirability function. *Industrial Quality Control*, 21: 494-498.
- Heino, M., Dieckmann, U., et Godo, O. 2002a. Measuring probabilistic reaction norms for age and size at maturation. *Evolution*, 56: 669-678.
- Heino, M., Dieckmann, U., et Godo, O. 2002b. Estimating reaction norms for age and size at maturation with reconstructed immature size distributions: a new technique illustrated by application to Northeast Arctic cod. *Ices Journal of Marine Science*, 59: 562-575.
- Henderson-Sellers, B., et Henderson-Sellers, A. 1996. Sensitivity evaluation of environmental model using fractional factorial experimentation. *Ecological Modelling*, 86: 291-295.
- Hickling, C. F. 1927. The natural history of the hake. Part I and II. *Fish. Invest. Ser. II*, 10: 112 pp.
- Hickling, C. F. 1935. The hake and the Hake Fishery. Edward Arnold & Co., London. 142pp.

-
- Hilborn, R., et Walters, C. J. 1992. Quantitative Fisheries Stock Assessment. Chapman and Hall, New York, New York. 570pp.
- Holland, D. S. 2000. A bioeconomic model of marine sanctuaries on Georges Bank. *Canadian Journal of Fisheries and Aquatic Sciences*, 57: 1307-1319.
- Holland, D. S. 2003. Integrating spatial management measures into traditional fishery management systems: the case of the Georges Bank multispecies groundfish fishery. *ICES Journal of Marine Science*, 60: 915-929.
- ICES. 1991a. Report of the Working Group on Fishery Units in sub-areas VII and VIII. ICES CM 1991/Assess: 24: 215pp.
- ICES. 1991b. Report of the working group on the assessment of the stocks oh hake. CM1991/Assess:20: 35pp.
- ICES. 1993. Report of the Working Group on the Assessment of Southern Shelf Stocks of Hake, Monk and Megrin. CM 2005/ACFM: 02: 92pp.
- ICES. 2003a. Report of the Working Group on Methods on Fish Stock Assessments. ICES CM 2003/D:03. : 171pp.
- ICES. 2003b. Report of the Study Group on Precautionary Reference Points For Advice on Fishery Management. ICES CM 2003/ACFM:15: 79pp.
- ICES. 2003c. Report of the Working Group on the Assessment of Southern Shelf Stocks of Hake, Monk and Megrin. ICES CM 2003/ACMF:01: 451pp.
- ICES. 2003d. Report of the Study Group on Age-Length Structured Assessment Models. ICES CM 2003/D:07: 64pp.
- ICES. 2004a. Report of the Arctic Fisheries Working Group. CM 2004/ACFM:28: 491pp.
- ICES. 2004b. Report of the Working Group on the Assessment of Southern Shelf Stocks of Hake, Monk and Megrin. ICES CM 2004/ACMF:02: 482pp.
- ICES. 2005a. Report of the Study Group on Age-length Structured Assessment Models. ICES CM 2005/D:01: 63pp.
- ICES. 2005b. Report of the Working Group on the Assessment of Hake, Monk and Megrin
-

- (WGHMM). ICES CM 2006/ACFM:01: pp.
- ICES. 2006a. Report of the Study Group on Age-length Structured Assessments Model. ICES CM 2006/RMC:10: 73pp.
- ICES. 2006b. Report of the Working Group on the Assessment of Hake, Monk and Megrin (WGHMM). ICES CM 2006/ACFM:29: 792pp.
- ICES. 2007. Report of the Working Group on the Assessment of Hake, Monk and Megrin (WGHMM). CM2007/Assess:21: 560pp.
- Ivanova, T., Malone, L., et Mollaghasemi, M. 1999. Comparison of a two-stage group-screening design to a standard 2k-p design for a whole-line semiconductor manufacturing simulation model. *In* Proceedings of the 1999 Winter Simulation Conference. Farrington, P. A., Nembhard, H. B. , Sturrock, D. T. , et Evans, G. W. (Eds.). IEEE Standards Office, Phoenix, Arizona, United States. 640 - 646.
- Jamieson, G. S., et Levings, C. O. 2001. Marine protected areas in Canada - implications for both conservation and fisheries management. *Canadian Journal of Fisheries and Aquatic Sciences*, 58: 138-156.
- Johnson, N. L., Balakrishnan, N., et Kotz, S. 1997. *Discrete Multivariate Distributions*. Wiley-Interscience, New York, United States of America. 328pp.
- Jones, P. 2007. Point-of-View: Arguments for conventional fisheries management and against no-take marine protected areas: only half of the story?. *Reviews in Fish Biology and Fisheries*, 17: 31-43.
- Kacher, M., et Amara, R. 2005. Distribution and growth of 0-group European hake in the Bay of Biscay and Celtic Sea: a spatial and inter-annual analyses. *Fisheries Research*, 71: 373-378.
- Kirkpatrick, S., Gelatt, C., et Vecchi, M. 1983. Optimization by Simulated Annealing. *Science*, 220: 671-680.
- Kleijnen, J. P. C. 1998. Experimental design for sensitivity analysis, optimization, and validation of simulations model. *In* Handbook of simulation :principles, methodology,

-
- advances, applications, and practice. Banks, J. (Ed.). Wiley-Interscience, New York. 173-224.
- Kleijnen, J. P. C.. 2001. Experimental designs for sensitivity analysis of simulation models. Discussion Paper 15 from Tilburg University, Center for Economic Research: 7pp.
- Kristensen, K., Lewy, P., et Beyer, J. 2006. How to validate a length-based model of single-species fish stock dynamics. *Canadian Journal of Fisheries and Aquatic Sciences*, 63: 2531-2542.
- Kvamme, C., et Froysa, K. G. 2004. Assessing the effects on the stocks of selectivity changes in a fishery. *Fisheries Research*, 69: 283-292.
- Lai, H., et Bradbury, A. 1998. A modified catch-at-size analysis model for a red sea-urchin (*Strongylocentrotus fransiscanus*) population. In Proceedings of the north Pacific symposium on invertebrate stock assessment and mangement. 85-96
- Laloë, F., Pech, N., Sabatier, R., et Samba, A. 1998. Model identification for flexible multifleet-multispecies fisheries: a simulation study. *Fisheries Research*, 37: 193-202.
- Laslett, G. M., Eveson, J. P., et Polacheck, T. 2002. A flexible maximum likelihood approach for fitting growth curves to tag-recapture data. *Canadian Journal of Fisheries and Aquatic Sciences*, 59: 976-986.
- Lucio, P., Murua, H., et Santurtun, M. 2000. Growth and reproduction of hake (*Merluccius merluccius*) in the Bay of Biscay during the period 1996-1997. *Ozeanografika*, 3: 325-354.
- Ludwig, D., et Walters, C. J. 1981. Measurement errors and uncertainty in parameter estimates for stock and recruitment. *Canadian Journal of Fisheries and Aquatic Sciences*, 38: 711-720.
- Lundy, C. J., Rico, C., et Hewitt, G. M. 2000. Temporal and spatial genetic variation in spawning grounds of European hake (*Merluccius merluccius*) in the Bay of Biscay. *Molecular Ecology*, 9: 2067-2079.
- Mahévas, S., et Pelletier, D. 2004. ISIS-Fish, a generic and spatially explicit simulation tool for evaluating the impact of management measures on fisheries dynamics. *Ecological*
-

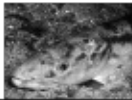
- Modelling, 171: 65-84.
- Martin, I.. 1991. A preliminary analysis of some biological aspects of hake (*Merluccius merluccius*) in the Bay of Biscay. C.M. 1991/G:54: 14pp.
- Maunder, M. N., et Deriso, R. B. 2003. Estimation of recruitment in catch-at-age models. Canadian Journal of Fisheries and Aquatic Sciences, 60: 1204-1216.
- Maunder, M. N., et Watters, G. M. 2003. A-SCALA: un analisis estadístico de captura a talla estructurado por edad para la evaluación de las poblaciones de atunes en el oceano pacifico oriental. Bulletin Inter-American Tropical Tuna Commission, 22: 531-581.
- McCullagh, P., et Nelder, J. A. 1989. Generalized Linear Models. Chapman & Hall, London. 387pp.
- Megrey, B. A. 1989. Review and comparison of Age-Structured Stock Assessment models from theoretical and applied points of view. American Fisheries Society Symposium, 6: 8-48.
- Michalewicz, Z. 1996. Genetic Algorithms + Data Structures = Evolution Programs. Springer, London. 387pp.
- Moloney, K. A. 1986. A generalized algorithm for determining category size. Oecologia, 69: 176-180.
- Murua, H., et Lucio, P. 1998. Reproductive modality and batch fecundity of the European Hake (*Merluccius merluccius* L.) in the Bay of Biscay. CalCOFI Reports, 39: 196-203.
- O'Brien, B. 1986. The distribution and community structure of larval and postlarval fish and chaetognaths off the West Coast of Ireland. Thesis Doctoral. University College Galway.
- Pelletier, D., et Magal, P. 1996. Dynamics of a migratory population under different fishing effort allocation schemes in time and space. Canadian Journal of Fisheries and Aquatic Sciences, 53: 1186-1199.
- Pelletier, D., et Parma, A. M. 1994. Spatial distribution of Pacific halibut (*Hippoglossus stenolepis*): an application of geostatistics to longline survey data. Canadian Journal of Fisheries and Aquatic Sciences, 51: 1506-1518.

-
- Pelletier, D., Mahévas, S., Drouineau, H., Vermard, Y., Thébaud, O., Guyader, O., et Poussin, B. submitted. Evaluation of the bioeconomic sustainability of complex fisheries under a wide range of policy options using ISIS-Fish. *Ecological Modelling*, : .
- Pelletier, D., Mahévas, S., Poussin, B., Bayon, J., André, P., et Royer, J. 2001. A conceptual model for evaluating the impact of spatial management measures on the dynamics of a mixed fishery. *In Spatial Processes and Management of Marine Populations*. Kruse, G. H., Bez, N. , Booth, A. , Dorn, M. W. , Hills, S. , Lipcius, R. N. , Pelletier, D. , Roy, C. , Smith, S. J. , et Witherell, D. (Eds.). University of Alaska Sea Grant, AK-SG-01-02, Anchorage, Alaska, Anchorage, Alaska. 53-66.
- Pet, J. S., Machiels, M. A. M., et Densen, W. L. T. V. 1996. A size-structured simulation model for evaluating management strategies in gillnet fisheries exploiting spatially differentiated populations. *Ecological Modelling*, 88: 195-214.
- Pilling, G. M., Kirkwood, G. P., et Walker, S. G. 2002. An improved method for estimating individual growth variability in fish, and the correlation between von Bertalanffy growth parameters. *Canadian Journal of Fisheries and Aquatic Sciences*, 59: 424-432.
- Piñeiro, C., et Sainza, M. 2003. Age estimation, growth and maturity of the European hake (*Merluccius merluccius* (Linnaeus, 1758)) from Iberian Atlantic waters. *ICES Journal of Marine Science*, 60: 1086-1102.
- Piñeiro, C., Sainza, M., Mogado, C., Pontual, H. D., Hoey, S., Beattie, S., Lucio, P. et al.. 2004. Report of 3rd Workshop on European Hake Age Reading. <http://www.ices.dk/reports/acfm/pgccdb/hke.agewk2004.pdf>: 22pp.
- Poulard, J. C. 2001. Distribution of hake (*Merluccius merluccius*, Linnaeus, 1758) in the Bay of Biscay and the Celtic sea from analysis of French commercial data. *Fisheries Research*, 50: 173-187.
- Poulard, J. C., et Léauté, J. P. 2002. Interaction between marine populations and fishing activities: temporal patterns of landings of La Rochelle trawlers in the Bay of Biscay. *Aquatic Living Resources*, 15: 197-210.
- Punt, A. E. 2003. The performance of a size-structured stock assessment method in face of
-

- spatial heterogeneity in growth. *Fisheries Research*, 65: 391-409.
- Reboulet, A. M., Viallefont, A., Pradel, R., et Lebreton, J. D. 1999. Selection of survival and recruitment models with *Surge 5.0*. *Bird Study*, 46: S148-156.
- Richards, L. J., et Schnute, J. T. 1998. Model complexity and catch-age analysis. *Canadian Journal of Fisheries and Aquatic Sciences*, 55: 949-957.
- Richards, L. J., Schnute, J. T., et Olsen, N. 1997. Visualizing catch-age analysis: a case study. *Canadian Journal of Fisheries and Aquatic Sciences*, 54: 1646-1658.
- Riis-Vestergaard, J., Velaco, F., Hill, L., et Olaso, I. 2000. Food consumption of European hake (*Merluccius merluccius*) estimated by application of a bioenergetics model: Is the growth of hake underestimated?. *ICES C.M. 2000/Q:12*: 12pp.
- Roldan, M. I., Marin, J. L. G., Utter, F. M., et Pla, C. 1998. Population genetic structure of European hake, *Merluccius merluccius*. *Heredity*, 81: 327-334.
- Sainsbury, K. J. 1980. Effect of Individual Variability on the von Bertalanffy Growth Equation. *Canadian Journal of Fisheries and Aquatic Sciences*, 37: 241-247.
- Schnute, J. 1987. A general fishery model for a size-structured fish population. *Canadian Journal of Fisheries and Aquatic Sciences*, 44: 924-940.
- Schnute, J. T., Richards, L. J., et Cass, A. J. 1989a. Fish survival and recruitment: investigations based on a size-structured model. *Canadian Journal of Fisheries and Aquatic Sciences*, 46: 743-769.
- Schnute, J. T., Richards, L. J., et Cass, A. J. 1989b. Fish growth: investigations based on a size-structured model. *Canadian Journal of Fisheries and Aquatic Sciences*, 46: 730-742.
- Schwarz, G. 1978. Estimating the Dimension of a Model. *The Annals of Statistics*, 6: 461-464.
- Schwefel, H. P. 1995. *Evolution and Optimum Seeking*. Wiley & Sons, New-York. 444pp.
- Shackell, N. L., Stobo, W. T., Franck, K. T., et Brickman, D. 1997. Growth of cod (*Gadus morhua*) estimated from mark-recapture programs of the Scotian Shelf and adjacent areas. *ICES Journal of Marine Science*, 54: 383-398.

-
- Shepherd, J. G. 1999. Extended survivors analysis: an improved method for the analysis of catch-at-age data and abundance indices. *ICES Journal of Marine Science*, 56: 584-591.
- Smith, B. D., Botsford, L. W., et Wing, S. R. 1998. Estimation of growth and mortality parameters from size frequency distributions lacking age patterns: the red sea urchin (*Strongylocentrotus franciscanus*) as an example. *Canadian Journal of Fisheries and Aquatic Sciences*, 55: 1236-1247.
- Smith, B. D., et Botsford, L. W. 1998. Interpretation of growth, mortality, and recruitment patterns in size-at-age, growth increment, and size frequency data. *Proceedings of the North Pacific Symposium on Invertebrate Stock Assessment and Management in Can. J. Fish. Aquat. Sci.*, 125: 125-139.
- Sullivan, P. J. 1992. A Kalman Filter Approach to Catch-at-Length Analysis. *Biometrics*, 48: 237-257.
- Sullivan, P. J., Lai, H., et Galluci, V. F. 1990. A catch -at- length analysis that incorporates a stochastic model of growth. *Canadian Journal of Fisheries and Aquatic Sciences*, 47: 184-198.
- Sumaila, U. R., Guénette, S., Alder, J., et Chuenpadgee, R. 2000. Addressing ecosystem effects of fishing using marine protected areas. *ICES Journal of Marine Science*, 57: 752-760.
- Valencia, V., Motos, L., et Urrutia, J.. 1989. Estudio de la variacion de la hidrografía y el plancton en la zona neritica frente a San Sebastian. : 80pp.
- Vandermeer, J. 1978. Choosing category size in a stage projection matrix. *Oecologia*, 32: 79-84.
- Viallefont, A., Lebreton, J. D., Reboulet, A. M., et Gory, G. 1998. Parameter identifiability and model selection in capture-recapture models: A numerical approach. *Biometrical journal*, 40: 313-325.
- Yardim, C., Gerstoft, P., et Hodgkis, W. S. 2005. A fast hybrid genetic algorithm-Gibbs sampler approach to estimate geoacoustic parameter uncertainties. In *Underwater acoustic measurements*.
-

- Zhang, Z., Hajas, W., Phillips, A., et Boutillier, J. A. 2004. Use of length-based models to estimate biological parameters and conduct yields analyses for male Dungeness crab (*Cancer magister*). *Canadian Journal of Fisheries and Aquatic Sciences*, 61: 2126-2134.
- Zheng, J., Murphy, M. C., et Kruse, G. H. 1995. Updated Length-Based Population Model and Stock-Recruitment Relationships for Red King Crab in Bristol Bay, Alaska. *Alaska Fishery Research Bulletin*, 2: 114-124.



Annexe 1 : Présentation d'algorithmes d'optimisation

Annexe 1 : Présentation d'algorithmes d'optimisation

Introduction

Le modèle s'ajuste par la méthode du maximum de vraisemblance, c'est à dire par la recherche d'un jeu de paramètres θ maximisant la vraisemblance. Cette méthode fournit un estimateur unique, asymptotiquement sans biais et convergent. Dans notre cas l'optimisation est complexe: la fonction de vraisemblance dépend d'un nombre important de paramètres, l'espace de recherche est donc de dimension élevée. Par conséquent, il est impossible de résoudre analytiquement le problème et une méthode numérique est indispensable. Celle-ci doit permettre d'explorer le plus largement possible le domaine de recherche afin d'éviter des optima locaux, tout en convergeant le plus rapidement possible vers la solution optimale.

Trois principales familles d'algorithmes ont été envisagés pour réaliser l'optimisation.

1 Gradient

La méthode du gradient consiste à construire itérativement des vecteurs de paramètres tels que la vraisemblance du nouveau vecteur soit supérieure à celle de son prédécesseur, la suite de vecteur converge alors vers un maximum (Ciarlet, 1982). Les nouveaux vecteurs sont construits en cherchant la direction de plus "grande pente" (de plus fort gradient) autour du vecteur courant.

Cette méthode, bien que la plus couramment utilisée, présente différents inconvénients. Tout d'abord, elle nécessite le calcul de dérivées afin de choisir la direction de plus forte pente. Cette opération peut-être difficile lorsque la fonction de vraisemblance est complexe. De plus, le résultat obtenu dépend du vecteur initial, les optima locaux ne sont pas nécessairement évités.

2 *Simplex*

La méthode du simplex consiste à choisir $n+1$ vecteurs initiaux de paramètres (si n paramètres à estimer). A chaque génération, le vecteur ayant la plus faible vraisemblance est modifié grâce à des calculs simples de géométrie dans l'espace (Freedman et al., 1998). Cette méthode permet donc d'éviter le calcul de dérivées puisque seule les fonctions de vraisemblance sont évaluées. Par contre, cette méthode ne permet pas de résoudre le problème des optima locaux et le temps de calcul augmente très rapidement avec le nombre de paramètres, ce qui rend la méthode peu adaptée à des fonctions complexes.

3 *Algorithmes de Newton et de quasi-Newton*

3.1 Algorithme de Newton

La méthode de Newton consiste à approximer à chaque itération la fonction f à optimiser par une approximation quadratique autour du point courant. Si l'on note $x=(x_1, \dots, x_n)$, le vecteur de paramètres à l'itération courant, on approxime la fonction par la fonction $q(s)$:

$$q(s) = f(x_k) + \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{bmatrix}^T s + \frac{1}{2} s^T \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \dots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \dots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix} s \quad (37)$$

$$\text{avec } H(f(x)) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \dots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \dots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix} \text{ la matrice d'Hessian}$$

on choisit le vecteur de paramètre y minimisant $q(y)=0$. Si la matrice d'Hessian est positive définie, il suffit de résoudre le système d'équations linéaires suivant:

$$\begin{bmatrix} \frac{\partial f^2(x)}{\partial x_1^2} & \dots & \frac{\partial f^2(x)}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f^2(x)}{\partial x_n \partial x_1} & \dots & \frac{\partial f^2(x)}{\partial x_n^2} \end{bmatrix} s = - \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{bmatrix} \quad (38)$$

et de prendre alors $y=x+s$.

La convergence est assurée si le point de départ est suffisamment proche de l'optimum de façon à ce que la matrice hessienne soit positive définie. Afin de limiter les risques de non convergence et d'éviter le calcul de la matrice hessienne qui est coûteux en temps, des améliorations ont été proposées, notamment la méthode dite de quasi-Newton.

3.2 Algorithme de quasi-Newton

Cet algorithme est particulièrement adapté quand la matrice hessienne est compliquée et longue à calculer. Cette méthode repose sur une approximation itérative de la matrice hessienne. On note f la fonction à optimiser, B_k l'approximation de la matrice hessienne à l'itération k , x_k le vecteur de paramètres à l'itération k , $\nabla f(x_k)$, le vecteur des dérivées premières de f au point x_k et $\nabla^2 f(x_k)$ la matrice hessienne au point x_k . A chaque itération, on résout le système suivant:

$$B_k d_k = -\nabla f(x_k) \quad (39)$$

qui est analogue au système 38.

Le vecteur d_k fournit une direction à partir de laquelle on construit (par une recherche linéaire le long de la direction)

$$x_{k+1} = x_k + \alpha_k d_k \quad (40)$$

La question est: comment construire B_{k+1} à partir de B_k ?

Si l'on considère le segment $[x_k, x_{k+1}]$, le théorème fondamental de l'analyse nous donne:

$$\left\{ \int_0^1 \nabla^2 f(x_k + t(x_{k+1} - x_k)) dt \right\} (x_{k+1} - x_k) = \nabla f(x_{k+1}) - \nabla f(x_k) \quad (41)$$

La matrice entre accolade peut-être interprétée comme l'Hessien moyen sur le segment $[x_k, x_{k+1}]$. Pour construire une matrice s'approchant de l'Hessien, on construit une matrice B_{k+1} vérifiant:

$$B_{k+1}(x_{k+1} - x_k) = \nabla f(x_{k+1}) - \nabla f(x_k) \quad (42)$$

Plusieurs transformations de B_k permettent d'assurer cette condition. Par exemple dans la méthode dite BFGS, on prend:

$$B_{k+1} = B_k + \frac{B_k s_k (B_k s_k)^T}{s_k^T B_k s_k} + \frac{y_k^T y_k}{y_k^T s_k} \quad (43)$$

Cette méthode permet de garder B_k positive définie si f est strictement convexe ou bien qu'on a choisi α_k assurant la convexité.

Cet algorithme est particulièrement efficace et rapide pour des problèmes de grande dimension. Par contre, selon la fonction f , peut se poser des problèmes de minima locaux.

4 *Algorithme génétique (Goldberg, 1989 ; Michalewicz, 1996)*

Ces méthodes sont basées sur le principe de la sélection naturelle. Un ensemble (appelé "population") de vecteur de paramètres (appelé "chromosomes") est initialement généré. A chaque chromosome est attribuée une valeur adaptative (dans notre cas la fonction de vraisemblance). Les chromosomes sont composés de gènes (un gène=un paramètre), dont la valeur est codée par une chaîne binaire (Figure 30). La valeur peut éventuellement directement être « codée » par sa valeur réelle ce qui présente certains avantages: les chromosomes sont de plus petites tailles puisque chaque gène est codé par un réel et non plus par une chaîne de binaires, ce codage est plus naturel et converge plus rapidement (Michalewicz, 1996).

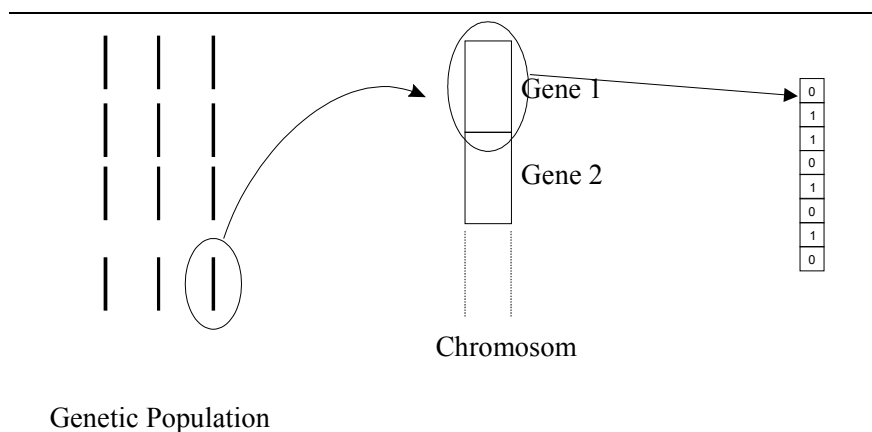


Figure 30 : Les différents niveaux d'organisation dans un algorithme génétique.

Le but est de construire une nouvelle génération dont la valeur adaptative moyenne est plus élevée, tout en maintenant une diversité afin d'explorer l'ensemble du domaine de recherche.

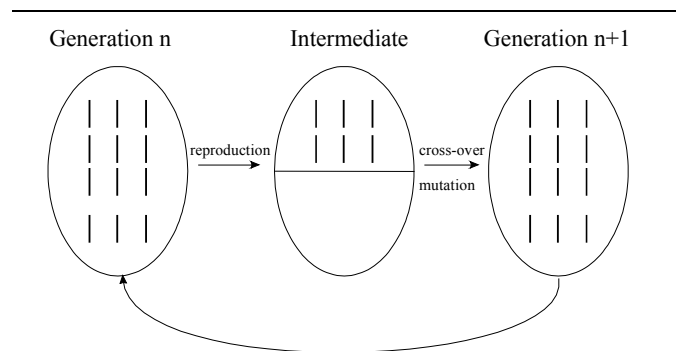


Figure 31 : Schéma général de fonctionnement d'un algorithme génétique.

Pour cela, trois opérations sont réalisées successivement (Figure 32):

- La sélection: elle consiste à sélectionner les chromosomes les plus adaptés pour construire la génération suivante. La sélection peut-être élitiste et consister à ne garder que les meilleurs chromosomes. L'inconvénient est alors que tous les "mauvais" chromosomes sont éliminés, entraînant une perte importante de diversité. L'espace de recherche risque alors de ne pas être totalement étudié. Une solution plus satisfaisante consiste à utiliser la méthode dite de la roulette biaisée dans laquelle la probabilité de conserver un chromosome est proportionnelle à sa valeur adaptative. Une troisième méthode dite du tournoi consiste à piocher deux chromosomes au hasard dans la population et à garder le meilleur (celui de plus forte fitness). Cette opération est répétée jusqu'à obtenir la population intermédiaire.
- Le crossing-over (Figure 32) permet de tester de nouvelles associations de valeurs en recombinaison différents chromosomes. De plus le point de recombinaison étant choisi au niveau du bit (élément de la chaîne binaire), il permet de créer de nouveaux chromosomes. Quand un codage réel est utilisé, le crossing over consiste à réaliser une combinaison linéaire des deux chromosomes.

- La mutation (Figure 32): à chaque bit est associé une probabilité de mutation, elle consiste à inverser le bit concerné et donc à "créer" de nouveaux allèles (version d'un gène). Pour un codage réel, la mutation consiste à piocher une nouvelle valeur dans le domaine de définition du paramètre. Il est également possible de restreindre le domaine de définition au cours des générations

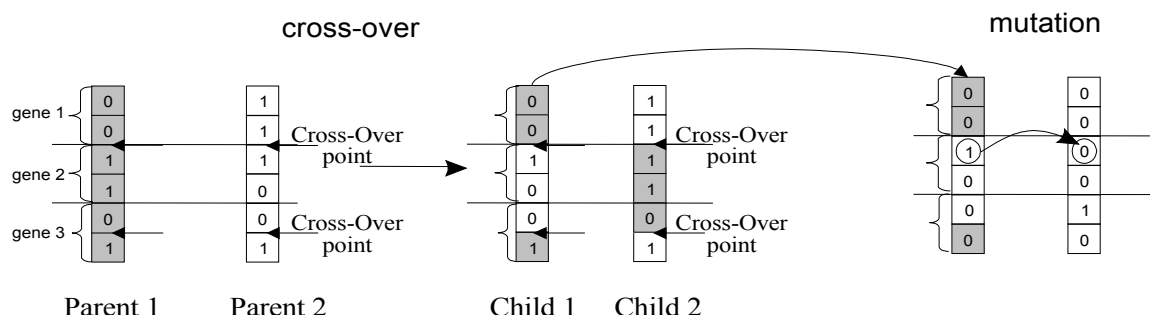


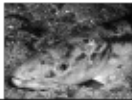
Figure 32 : Mode de fonctionnement du cross-over et de la mutation dans le cas d'un codage binaire.

Ces trois processus permettent donc une amélioration de l'adaptation moyenne de la population (sélection) tout en testant de nouvelles combinaisons (cross-over) et de nouvelles valeurs (cross-over et mutation).

Cette méthode permet donc théoriquement d'explorer tout le domaine de recherche. Elle est particulièrement adaptée pour des fonctions complexes, dépendant d'un grand nombre de paramètres.

Conclusion

Les algorithmes génétiques, ou d'autres algorithmes proches comme les stratégies évolutives (Schwefel, 1995), au vu de la complexité de la fonction de vraisemblance semblaient particulièrement adaptés pour maximiser la fonction de vraisemblance. Ils se sont effectivement montrés très efficaces pour réaliser les ajustements lors de l'analyse de sensibilité aux hypothèses de discrétisation. En revanche, nous avons préféré utiliser un algorithme de type quasi-Newton pour ajuster le modèle. Couplé à un algorithme de différentiation automatique et en fournissant un point initial « vraisemblable », il s'est avéré très rapide. De plus, ce type d'algorithme permet le calcul de la matrice hessienne et donc de la matrice variance-covariance des paramètres. Se pose malgré tout le problème d'optima locaux.



Annexe 2 : Identifiabilité des paramètres



Annexe 2 : Identifiabilité des paramètres

Introduction

Du fait du grand nombre d'inconnues dans notre modèle, il est essentiel de s'assurer de son identifiabilité. Un jeu de paramètres est dit non identifiable quand plusieurs jeux de valeurs conduisent au maximum de vraisemblance (Reboulet *et al.*, 1999). Cela a plusieurs conséquences. Tout d'abord, si deux paramètres sont redondants, l'estimation que l'on obtient n'a aucun sens puisque des solutions équivalentes peuvent-être obtenues en modifiant des valeurs. De plus, si des paramètres sont redondants, l'utilisation des critères statistiques comme l'AIC et le BIC sont problématiques puisqu'ils reposent sur le nombre de paramètres réellement identifiables du modèle (Viallefont *et al.*, 1998 ; Reboulet *et al.*, 1999 ; Gimenez *et al.*, 2004). Enfin, la perception que l'on a du système peut-être complètement modifiée selon le modèle retenu en cas de redondance des paramètres (Schnute, 1987 ; Laloë *et al.*, 1998) On recense deux types d'identifiabilité: on parle d'identifiabilité intrinsèque (ou à priori) quand la redondance de deux paramètres est due à la structure même du modèle, c'est à dire que les deux paramètres ont implicitement le même rôle dans le modèle. Dans ce cas, la redondance des paramètres existe quelque soit le jeu de données disponibles et peut en théorie être détecté à priori par analyse du modèle (Schnute, 1987 ; Gimenez *et al.*, 2004). La redondance extrinsèque est elle liée à la structure des données disponibles: dans certaines conditions (mauvaise qualité des données, quantité de données insuffisantes par exemple (Reboulet *et al.*, 1999)) l'identifiabilité des paramètres n'est pas assurée bien qu'intrinsèquement les paramètres ne soient pas redondants (Schnute, 1987 ; Gimenez *et al.*, 2004). Nous ferons ici un recensement de différentes méthodes, en se basant sur la revue réalisée par Gimenez et al. (2004) permettant d'analyser l'identifiabilité des modèles, les atouts et faiblesses de chacune, et leur applicabilité dans notre modèle.

1 Courbes de profil de vraisemblance

Cette méthode ne permet pas de distinguer redondance intrinsèque et extrinsèque. Elle consiste à tracer graphiquement en fonction d'un paramètre dont on fixe la valeur, la vraisemblance à l'optimum (maximisation en faisant varier les autres paramètres). Une courbe plate autour de l'optimum obtenu est caractéristique d'un paramètre redondant. On peut éventuellement distinguer redondance intrinsèque et redondance extrinsèque en travaillant sur plusieurs jeux de données (la redondance extrinsèque provenant d'une mauvaise configuration des données, en multipliant les jeux de données, on diminue la probabilité de ne tomber que sur des mauvais jeux de données).

Cette technique simple est assez efficace sur des modèles simples, ou au moins sur des modèles dont on a une assez bonne idée des paramètres redondants. Si ce n'est pas le cas, il faut tracer le profil de vraisemblance pour chaque paramètre inconnu (avec une optimisation nécessaire pour chaque point du profil!), analyser tous les graphiques. Cela ne semble pas applicable notre cas.

2 Matrice hessienne

La matrice hessienne d'une fonction est la matrice contenant l'ensemble des dérivées partielles secondes de la fonction par rapport aux paramètres. Dans le cas d'optimisation, on s'intéresse à la matrice hessienne $H(\Theta)$ de la logvraisemblance par rapport aux paramètres Θ . Au maximum de vraisemblance, on peut approximer la distribution de Θ par: $\hat{\Theta} \sim N(\Theta, -H^{-1})$. L'analyse de la matrice hessienne permet d'analyser l'optimum: en cas d'optimum fort, la matrice hessienne est positive définie (dans ce cas là toutes les valeurs propres de H sont strictement positives), si au contraire la matrice l'optimum est « faible », la matrice hessienne est semi-positive définie (les valeurs propres de H sont positives ou nulles). La matrice hessienne permet également d'estimer le nombre de paramètres effectivement estimables dans le modèle, et de détecter quels paramètres du modèle sont effectivement

redondants. En pratique, le nombre de paramètres identifiables d'un modèle est égal au rang de la matrice hessienne, c'est à dire au nombre de valeurs propres non nulles de cette matrice (Viallefont *et al.*, 1998). Viallefont et al. (1998) recommandent de calculer les valeurs singulières de $H(\hat{\Theta})$ et fournissent un critère pour déterminer si la valeur singulière est nulle ou non. Cependant, nous n'avons pu trouver aucune justification sur ce seuil qui apparaît empirique et n'est pas nécessairement adapté à tous les cas. Géométriquement, les vecteurs propres associés à des valeurs propres nulles correspondent à des directions planes de la fonction de vraisemblance (Reboulet *et al.*, 1999). On peut également en examinant les vecteurs propres associés aux valeurs propres nulles déterminer quels sont les paramètres effectivement redondants. Les paramètres non identifiables sont les paramètres de coordonnées non nulles dans les différents vecteurs propres correspondants aux valeurs propres nulles (Reboulet *et al.*, 1999). Cette méthode est donc assez facilement applicable. Elle souffre toutefois de quelques inconvénients plus ou moins graves. Tout d'abord, cette méthode ne permet pas de distinguer redondance intrinsèque et extrinsèque. Un autre problème vient du calcul de la matrice hessienne, le calcul analytique étant en général impossible (notamment dans notre modèle), on doit approximer la matrice hessienne par différences finies, or cette méthode est sensible aux choix du pas des différences. Le dernier inconvénient est que théoriquement cette méthode n'est valable que pour des ajustements sans contraintes. On peut l'appliquer si la solution obtenue est loin des bornes de l'espace de recherche (puisque dans ce cas-là la contrainte ne joue pas ou peu). Par contre si la solution est proche des bornes du domaine de définition, la solution n'est plus forcément un optimum mathématique (puisque le calcul de la dérivée prend en compte une pénalité liée à la borne) et donc rien ne garantit que la matrice soit positive ou semi-positive définie. Les résultats de cette méthode ne seront donc valables que sous réserve que la solution soit éloignée des bornes.

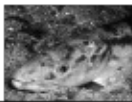
3 Calcul formel de la matrice des dérivées

Cette méthode consiste à calculer formellement la matrice D composée des dérivées de la densité de probabilité de chacune des observations par rapport à chacun des paramètres du modèle. Le rang de cette matrice D donne alors le nombre de paramètres réellement identifiables (Gimenez *et al.*, 2004). Cette technique est donc particulièrement adaptée à des modèles simples avec un nombre de paramètres limités. En pratique, cette solution ne paraît pas envisageable dans notre cas, le modèle étant beaucoup trop complexes et le nombre d'inconnues trop élevées.

4 Simulation

Avec l'analyse de la matrice hessienne, c'est la seconde méthode qui paraît applicable à notre cas d'étude. Cette méthode consiste à simuler un jeu de données pour des valeurs de paramètres connus. On réestime les paramètres à partir des données simulées et on analyse alors l'écart type des estimations ainsi que le biais des estimations. Cette méthode permet de détecter des redondances intrinsèques, puisqu'elle travaille à partir de données simulées ; mais peut aussi permettre de détecter des redondances extrinsèques en simulant des données avec des paramètres estimés à partir d'un jeu de données réelles.

Laloë *et al.* (1998) simule des données avec un modèle complexe et ajuste un modèle imbriqué plus simple sur les données simulées. La capacité du modèle simple à s'ajuster sur les données simulées démontre la non-identifiabilité du modèle complexe. Toutefois, il observe que si les deux modèles s'ajustent de la même façon aux données, les projections réalisées à partir des deux modèles peuvent différer fortement. Il souligne que le choix d'un modèle sur le seul critère statistique tendrait à choisir le modèle simple (parcimonie), mais que la connaissance de la pêcherie peut au contraire amener à sélectionner le modèle le plus complexe.



Annexe 3 : Vers la simulation de règles de gestion à l'aide du simulateur ISIS-Fish

Annexe 3 : Vers la simulation de règles de gestion à l'aide du simulateur ISIS-Fish

Introduction

Le simulateur ISIS-Fish (Pelletier *et al.*, 2001 ; Mahévas et Pelletier, 2004) est un outil générique visant à évaluer l'impact de scénarios de gestion, notamment spatialisés sur les population et leur exploitation. Le modèle repose sur trois sous-modèles, explicites dans l'espace et les saisons : un modèle décrivant la dynamique de la population, un modèle décrivant la dynamique de l'activité de pêche et un modèle décrivant la dynamique de gestion. Drouineau *et al.* (2006) ont proposé un paramétrage du simulateur afin d'évaluer l'impact d'aires marines protégées, de TAC et d'interdiction d'engin sur la pêche mixte merlu – langoustine du golfe de Gascogne. Toutefois, il existait certaines lacunes liées à un manque de connaissance de la biologie du merlu. Nous proposons ici quelques modifications à la fois dans le code du simulateur et dans le paramétrage utilisant les résultats du modèle. La meilleure connaissance des migrations devrait permettre d'améliorer les résultats précédemment obtenus.

1 Adaptation du simulateur

Le code du simulateur doit-être légèrement modifié afin qu'il s'adapte à notre modèle. En particulier, l'ordre des processus biologiques doit-être modifié pour qu'ils reprennent celui de notre modèle (cela consiste en une inversion du recrutement et de la croissance). Une autre modification doit-être faite: la croissance ne doit avoir lieu que dans le premier mois de chaque saison (c'est déjà le cas pour les migrations). Cela est tout à fait faisable grâce la flexibilité du simulateur. Le code du simulateur doit être modifié ainsi

```

protected void computeMonth(SimulationContext context, SiMatrix siMatrix, Date date,
Population pop) throws IsisFishException, TopiaException {
    // to add result
    ResultStorage resManager = context.getSimulationStorage().getResultStorage();

    PopulationMonitor popMon = context.getPopulationMonitor();
    MatrixND N = popMon.getN(pop);

    if (log.isInfoEnabled()) {
        log.info("===== begin "+ date + " - " + pop + "
===== " + System.currentTimeMillis());
        log.info("N: " + N);
    }

    // add N and biomass result now, before computation
    // N is reassigned during computation
    resManager.addResult(date, pop, N);

    if (resManager.isEnabled(ResultName.MATRIX_BIOMASS)) {
        MatrixND biomass = siMatrix.matrixBiomass(N, pop, date);
        resManager.addResult(date, pop, biomass);
    }

    Month month = date.getMonth();
    PopulationSeasonInfo info = pop.getPopulationSeasonInfo(month);

    //Recrutement
    MatrixND R = info.getReproductionMatrix(month, N);
    log.debug("R: " + R);

    // ajout de la matrice R dans le suivi de la pop comme etant
    // la reproduction pour le mois courant.
    popMon.setReproduction(date, pop, R);

    // recrutement
    MatrixND recruitment = popMon.getRecruitment(date, pop);
    log.debug("recruitment: " + recruitment);

    // mortalite de la reproduction
    popMon.applyReproductionMortality(pop);

    N = N.add(recruitment);
    log.debug("N after recru: " + N);

    // group change
    MatrixND CA = null;
    if (date.getDate() == 0 || !date.getMonth().equals(info.getFirstMonth())) {
        CA =
MatrixFactory.getInstance().matrixId(pop.sizePopulationGroup()*pop.sizePopulationZone());
    } else { //si on n'est pas au premier pas de temps et qu'on est au premiers mois de la saison,
on change de classe
        CA = info.getGroupChangeMatrix(month);
    }
    log.debug("CA: " + CA);

```

```

//calcul des migrations
//migration
MatrixND M = info.getMigrationMatrix(month, N);
log.debug("M: " + M);
//emigration
MatrixND EM = info.getEmigrationMatrix(month, N);
log.debug("EM: " + EM);
//immigration
MatrixND IM = info.getImmigrationMatrix(month, N).transpose();
log.debug("IM: " + IM);

// pour les premiers calculs on met N en une matrice 1D
MatrixND N1D = pop.N2DToN1D(N);
log.debug("N1D: " + N1D);
MatrixND tmp0 = N1D.mult(CA);
MatrixND tmp1 = M.minus(EM);
MatrixND tmp2 = tmp0.mult(tmp1);
MatrixND tmp3 = tmp2.add(IM);

log.debug("N1D after mig: " + tmp3);

// On reconvertie en une matrice Semantique
N = pop.split2D(tmp3);

log.debug("N after mig: " + N);

MatrixND abundance = N;
// compute fishing matrix only if there are one or more strategy
if (siMatrix.getStrategies(date).size() > 0) {
    // compute some Matrix and add result
    MatrixND matrixFishingMortality = siMatrix.matrixFishingMortality(date, pop);
    resManager.addResult(date, pop, matrixFishingMortality);

    abundance = siMatrix.matrixAbundance(N, pop, date, matrixFishingMortality);

    // this matrix is necessary for matrixCatchPerStrategyMet
    MatrixND catchRatePerStrategyMet = siMatrix.matrixCatchRatePerStrategyMet(pop, date,
matrixFishingMortality);
    resManager.addResult(date, pop, catchRatePerStrategyMet);

    // this matrix is necessary for PopulationMonitor.holdCatch (reused in rule)
    MatrixND catchPerStrategyMet = siMatrix.matrixCatchPerStrategyMet(N, pop, date,
catchRatePerStrategyMet);
    popMon.holdCatch(pop, catchPerStrategyMet);
    resManager.addResult(date, pop, catchPerStrategyMet);

    if
(resManager.isEnabled(ResultName.MATRIX_CATCH_WEIGHT_PER_STRATEGY_MET)) {
        MatrixND catchWeightPerStrategyMet = siMatrix.matrixCatchWeightPerStrategyMet(pop,
date, catchPerStrategyMet);
        resManager.addResult(date, pop, catchWeightPerStrategyMet);
    }
}

```

```
    }  
  }  
  log.debug("abundance: " + abundance);  
  
  // Keep new N  
  popMon.setN(pop, abundance);  
  
  log.debug("=====  
  ===== end " + date + " - " + pop + "  
  =====");  
}
```

2 Modèle de dynamique de la population

Le modèle de dynamique des populations est celui qui doit être le plus réactualisé. En particulier, la population ne doit plus être structurée en âge mais en longueur. La population comporte donc 53 classes de 2 cm (entre 4 et 110 cm). Quatre saisons doivent être définies reprenant les quatre trimestres du modèle. Pour la croissance, la matrice de transition estimée par le modèle est directement réutilisable. La reproduction a lieu en trimestre 1 et en trimestre 2, les deux zones de reproduction se situant aux accores du golfe de Gascogne et de la mer Celtique. Le recrutement doit avoir lieu trois mois après la reproduction sur les plateaux du golfe de Gascogne et de la mer Celtique. N'ayant pas utilisé d'équation stock-recrutement une solution est de définir un recrutement constant égal à la moyenne estimée par notre modèle. L'équation de reproduction devient donc:

```
for(Zone zr : zoneRepro){  
    if (zr.getName.equals(new String('accores_celtique')) && month.getNumMois==3)  
        result.setValue(zr,3.08e8);  
    else if (zr.getName.equals(new String('accores_gascogne')) &&  
month.getNumMois==0)  
        result.setValue(zr,2.67e8);  
    else  
        result.setValue(zr,0);  
}  
return 0;
```

Pour les migrations, il est nécessaire d'écrire une nouvelle équation distinguant matures et immatures:

Pour la mortalité, une équation permet de définir un recrutement constant à 0.6 an⁻¹:

```
if (group == null) return 0; //pas de mortalité sur la classe naissance
else return 0.6;
```

L'équation de poids devient:

```
return 0.00733*group.getLength^2.981/1e6; //retourne un poids en tonne
```

Enfin, l'équation de migration devient:

```
//on calcule la proportion de mature
double l50=37.9;
double sr=11.5;
double beta=2*Math.log(3)/sr;
double alpha=-l50*beta;

double Taille=group.getLength;
double Maturite=beta*Taille;
Maturite*=-1;
Maturite-=alpha;
Maturite=Math.exp(Maturite);
Maturite+=1;
Maturite=1/Maturite;

Date=context.getSimulationControl.getDate();
int NumMois=Date.getMonth.getMonthNumber;

double proportion=0;

java.lang.String arrivee=arrivalZone.getName;
java.lang.String depart=departureZone.getName;

if (NumMois==0){//migration du premier trimestre
  if (depart.equals('accores_celtique') && arrivee.equals('accores_celtique'))
    proportion=1;
  else if (depart.equals('plateau_celtique') && arrivee.equals('accores_celtique'))
    proportion=(Maturite*0.0062969);
  else if (depart.equals('plateau_celtique') && arrivee.equals('plateau_celtique'))
    proportion=(1-Maturite)+Maturite*0.993703;
  else if (depart.equals('accores_gascogne') && arrivee.equals('accores_gascogne'))
    proportion=1;
  else if (depart.equals('plateau_gascogne') && arrivee.equals('plateau_gascogne'))
    proportion=Maturite;
  else if (depart.equals('plateau_gascogne') && arrivee.equals('plateau_gascogne'))
    proportion=1-Maturite;
}
else if (NumMois==3){//migration du second trimestre
  if (depart.equals('accores_celtique') && arrivee.equals('accores_celtique'))
    proportion=1;
  else if (depart.equals('plateau_celtique') && arrivee.equals('accores_celtique'))
    proportion=(Maturite*0.0054742);
```

```
else if (depart.equals('plateau_celtique') && arrivee.equals('plateau_celtique'))
    proportion=(1-Maturite)+Maturite*0.994526;
else if (depart.equals('accores_gascogne') && arrivee.equals('accores_celtique'))
    proportion=Maturite*0.570065;
else if (depart.equals('accores_gascogne') && arrivee.equals('accores_gascogne'))
    proportion=(1-Maturite)+Maturite*0.245239;
else if (depart.equals('accores_gascogne') && arrivee.equals('plateau_gascogne'))
    proportion=Maturite*0.184696;
else if (depart.equals('plateau_gascogne') && arrivee.equals('accores_gascogne'))
    proportion=(1-Maturite)*0.131248;
else if (depart.equals('plateau_gascogne') && arrivee.equals('plateau_gascogne'))
    proportion=(1-Maturite)*0.868752+Maturite;
}
}
else if (NumMois==6){//migration du 3eme trimestre
    if (depart.equals('accores_celtique') && arrivee.equals('accores_celtique'))
        proportion=(1-Maturite)+Maturite*0.853575;
    if (depart.equals('accores_celtique') && arrivee.equals('plateau_celtique'))
        proportion=Maturite*0.146425;
    if (depart.equals('plateau_celtique') && arrivee.equals('accores_celtique'))
        proportion=(1-Maturite)*0.0711041;
    if (depart.equals('plateau_celtique') && arrivee.equals('plateau_celtique'))
        proportion=(1-Maturite)*0.928896+Maturite;
    else if (depart.equals('accores_gascogne') && arrivee.equals('accores_gascogne'))
        proportion=1;
    else if (depart.equals('plateau_gascogne') && arrivee.equals('accores_gascogne'))
        proportion=Maturite*0.0818838;
    else if (depart.equals('plateau_gascogne') && arrivee.equals('plateau_gascogne'))
        proportion=(1-Maturite)+0.918116*Maturite;
}
else if (NumMois==9){//migration du 4eme trimestre
    if (depart.equals('accores_celtique') && arrivee.equals('accores_celtique'))
        proportion=(1-Maturite)+Maturite*0.572232;
    if (depart.equals('accores_celtique') && arrivee.equals('plateau_celtique'))
        proportion=Maturite*0.427768;
    if (depart.equals('plateau_celtique') && arrivee.equals('plateau_celtique'))
        proportion=1;
    else if (depart.equals('accores_gascogne') && arrivee.equals('accores_gascogne'))
        proportion=1;
    else if (depart.equals('plateau_gascogne') && arrivee.equals('plateau_gascogne'))
        proportion=1;
}
}
return proportion;
```



**Annexe 4 : Combiner erreur de processus et
erreur d'observations dans un modèle de
dynamique des populations**

*Il n'est pas déshonorant de perdre, ce qui est déshonorant
est de refuser le combat*

(devise du FC Lorient)

Annexe 4 : Combiner erreur de processus et erreur d'observations dans un modèle de dynamique des populations

Introduction

Lors de l'ajustement d'un modèle de dynamique des populations par maximum de vraisemblance, on néglige souvent l'incertitude existant sur les processus de la dynamique des populations devant l'incertitude autour des observations, hypothèse pouvant mener à des biais dans l'estimation des paramètres et dans les tests d'hypothèse si elle n'est pas vérifiée. Cependant, considérer simultanément ces deux sources d'incertitude entraîne d'importants problèmes méthodologiques pour l'ajustement des modèles par maximum de vraisemblance (de Valpine et Hastings, 2002). Les modèles dits à « espace d'états » reposent sur deux sous-modèles : un modèle stochastique décrivant la dynamique du système (population et pêche) et un modèle stochastique d'observations (captures, indices d'abondance...) (voir la Figure 33 pour une représentation schématique de notre modèle). Si on note θ le vecteur des paramètres inconnus, X_t l'état du système au temps t et Y_t les observations du système au temps t , le modèle d'espace d'états peut se résumer en deux équations:

$$X_{t+1} = F_{\theta}(X_t, v_t) \quad (44)$$

$$Y_t = G_{\theta}(X_t, \epsilon_t) \quad (45)$$

avec F_{θ} l'équation de paramètres θ déterminant le nouvel état X_{t+1} de la population comme une fonction de l'état précédant X_t et d'une variable aléatoire v_t , et G_{θ} la fonction de paramètres θ déterminant les observations à partir de l'état X_t et d'une variable aléatoire ϵ_t .

Dans notre modèle, la connaissance sur le recrutement et la dérive de capturabilité sont mal connues. Le recrutement est ainsi souvent modélisé comme une variable aléatoire suivant une distribution lognormale tandis que la dérive de capturabilité est modélisée comme une marche aléatoire. Ils constitueraient alors des « erreurs de processus » affectant les états (effectifs

pour le recrutement, mortalité par pêche et effectifs pour la dérive des puissances de pêche).

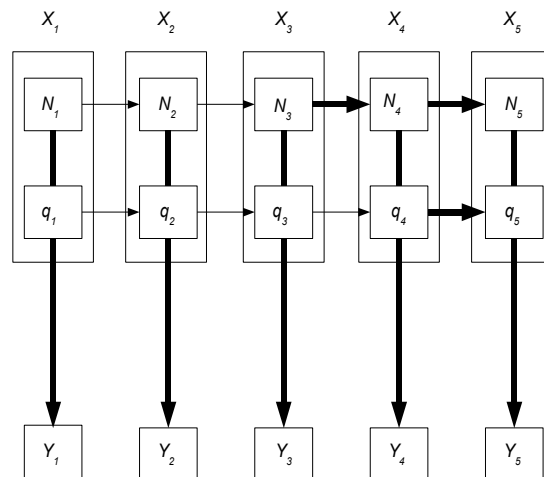


Figure 33 : Représentation schématique du modèle sur 5 trimestres. Les flèches en gras représentent une transition stochastique (recrutement et progrès technique). N_t désigne les effectifs, q_t les capturabilités et Y_t les observations (captures).

Le but de l'ajustement est de déterminer le jeu de paramètres Θ rendant le plus vraisemblable les observations ; c'est à dire si on note $L(\Theta/Y)$ la vraisemblance de Θ sachant les observations, que l'on cherche Θ tel que $L(\Theta/Y) = P(Y/\Theta)$ soit maximum (P dénotant ici la densité de probabilité d'observer Y sachant les paramètres). Une première méthode, encore très souvent utilisée en halieutique, est la méthode dite « errors-in-variables » (EV) dans laquelle on maximise le produit des vraisemblances des observations et des états (Ludwig et Walters, 1981 ; Richards *et al.*, 1997):

$$L_{EV}(\Theta/Y) = P(Y/\Theta, X) \times P(X/\Theta) \quad (46)$$

La vraisemblance EV ne considère donc qu'une seule trajectoire d'états (celle de vraisemblance maximum). Celle-ci n'est pourtant pas la vraisemblance « classique », qui elle intègre la vraisemblance sur l'ensemble des états possibles du système (de Valpine et Hastings, 2002 ; de Valpine et Hilborn, 2005):

$$L(\Theta|Y)=P(Y|\Theta)=\int P(Y, X|\Theta)dX=\int P(Y|X, \Theta)P(X|\Theta)dX \quad (47)$$

Il est démontré que la vraisemblance EV est différente de la vraisemblance. Par conséquent, les propriétés liées à l'estimation par maximum de vraisemblance (convergence, sans biais...) ne sont pas assurées. De plus, la vraisemblance EV ne permet théoriquement pas ensuite de réaliser de tests d'hypothèses (ratio de vraisemblance notamment, AIC, BIC).

Cette solution ne paraît pas satisfaisante. Toutefois l'intégrale (47) étant hautement multidimensionnelle (autant de dimensions que d'effets aléatoires dans le modèle), son intégration n'est pas triviale. Nous allons présenter ici différentes méthodes existantes pour réaliser ce calcul, leurs avantages et leurs inconvénients.

1 Les méthodes de Monte-Carlo direct

Une première méthode pour réaliser cette intégration consiste à tirer un large échantillon suivant $P(X|\Theta)$, c'est à dire dans notre cas qu'il faut tirer des recrutements annuels et des dérivées annuelles de capturabilité dont les lois sont définis Θ . On peut alors approximer l'équation 47 par:

$$L(\Theta)\approx\frac{1}{M}\sum_{i=1}^M P(Y|X_i, \Theta) \quad (48)$$

Un algorithme d'optimisation est couplé afin de trouver le vecteur Θ qui maximise l'équation 48. Si cette technique peut marcher dans des cas assez simples, elle se révèle assez inefficace quand le nombre de dimensions du problème augmente : il faut tirer des échantillons et en calculer la vraisemblance à chaque itération de l'algorithme d'optimisation,

ce qui peut s'avérer assez coûteux. De plus, de nombreux individus de l'échantillon seront peu vraisemblables et auront un poids très faibles dans le calcul de l'équation 48.

2 Importance sampling

Une amélioration possible est la technique dite « importance sampling » (voir par exemple (Maunder et Deriso, 2003), (Gilks *et al.*, 1995)). Dans cette méthode, au lieu d'échantillonner suivant $P(X/\Theta)$, on échantillonne suivant une densité dite d'importance $P_S(X/\Theta)$ et l'on réalise l'approximation suivante:

$$L(\Theta) = \int P(Y/X, \Theta) P(X/\Theta) dX = \int P(Y/X, \Theta) \frac{P(X/\Theta)}{P_S(X/\Theta)} P_S(X/\Theta) dX$$

$$\approx \frac{1}{M} \sum_{i=1}^M P(Y/X_i, \Theta) \frac{P(X_i/\Theta)}{P_S(X_i/\Theta)} \quad (49)$$

Il faut donc savoir évaluer $P(X/\Theta)$, et simuler facilement selon la loi $P_S(X/\Theta)$. Pour être efficace, il faut que $P_S(X/\Theta)$ soit assez proche de $P(X/Y, \Theta)$, or nous avons peu d'informations sur cette densité de probabilité et de plus, on reste obligé de coupler l'échantillonneur de Monte-Carlo a un algorithme d'optimisation...

3 Filtrage particulière

Le filtrage particulière est basé sur la construction d'un échantillon de « particules » suivant la distribution $P(X_t/Y_t, \Theta)$ pour chaque pas de temps t . On calcule la vraisemblance d'un jeu de paramètres séquentiellement en calculant:

$$L(\Theta) = P(Y_1/\Theta) \prod_{t=2}^T P(Y_t/Y_{1:t-1}, \Theta) \quad (50)$$

1. On génère un échantillon $X_t^{(i)} = (X_{1:t}/Y_{1:t-1})^{i=1, \dots, m}$ suivant $P(X_{1:t}/Y_{1:t-1}, \Theta)$.

2. On peut alors calculer par approximation de Monte Carlo

$$P(Y_t / Y_{1:t-1}, \Theta) \approx \frac{1}{m} \sum_{i=1}^m P(Y_t / X^{(i)}, \Theta)$$

3. Pour chaque particule $X_t^{(i)}$, on attribue un poids

$$w^{(i)} = P(Y_t / X^{(i)}, \Theta) / \sum_{j=1}^m P(Y_t / X^{(j)}, \Theta)$$

4. On tire alors un échantillon $X_{t+1}^{(l)} = (X_{t+1} / Y_{1:t}, \Theta)^{l=1, \dots, m}$ en piochant parmi les $X_t^{(i)}$ avec une probabilité $w^{(i)}$ (permet d'éliminer les particules à poids trop faibles pour se concentrer sur celles avec un poids élevé).

5. On tire alors un échantillon de $P(X_{t+1} / Y_{1:t})$ en simulant des $X_{t+1}^{(l)}$ à partir des particules $X_t^{(i)}$.

6. Retour à l'étape 1 jusqu'au pas de temps final

7. On aboutit à une mesure de $L(\Theta)$

Cet algorithme n'est pas très compliqué à mettre en oeuvre et converge bien vers $L(\Theta)$ quand m tend vers l'infini. Il existe quelques variantes de l'algorithme, mais toutes reposent sur le même principe général. Il faut quand même apparemment un assez grand nombre de particules (de plus, on fait des calculs pour des particules que l'on n'utilise plus ensuite...) pour que la convergence soit assurée, et on doit coupler le filtre particulaire à un algorithme d'optimisation ; le filtrage devant être recommencé à chaque itération de l'algorithme d'optimisation.

4 Algorithme d'espérance-maximisation (EM)

Ces algorithmes reposent sur deux étapes : la phase dite d'expectation fournit un échantillon d'états $(X)^{(1:m)}$ suivant la distribution $P(X/Y, \Theta_0)$. Plusieurs méthodes peuvent permettre d'obtenir ce type d'échantillons (par exemple un filtrage particulière, ou une chaîne MCMC). On construit alors dans la phase de maximisation, un jeu de paramètres Θ_l qui

maximise l'espérance de la logvraisemblance des observations $\frac{1}{M} \sum_{j=1}^M \log(P(Y, X^{(j)} / \Theta_l))$.

Cet algorithme est convergent, par contre il est apparemment assez sensible aux optima locaux... et assez coûteux en temps de calcul (générer un échantillon à chaque itération de l'algorithme d'optimisation).

4 Monte-Carlo kernel likelihoods (de Valpine, 2003 ; de Valpine, 2004 ; de Valpine et Hilborn, 2005)

4.1 Principe général

De Valpine, constatant que les méthodes précédentes étaient peu efficaces dans le cas d'intégration multidimensionnelle, a développé une nouvelle méthode basée sur des chaînes MCMC et des estimateurs de densité. Pour cela, il suppose que les paramètres Θ sont des variables aléatoires, auxquels il attribue des priors peu (pas) informatifs. On peut alors calculer la densité à posteriori des paramètres sachant les observations:

$$\pi(\Theta/Y) = \int \pi(\Theta, X/Y) dX = \int \frac{P(Y/X, \Theta) P(X/\Theta) P(\Theta)}{P(Y)} dX \quad (51)$$

$$\text{avec } \pi(\Theta, X/Y) = \frac{P(Y/X, \Theta) P(X/\Theta) P(\Theta)}{P(Y)} \quad (52)$$

$P(\Theta)$ constituent les priors des paramètres (constants si l'on a choisi des priors uniformes) et $P(Y)$ est une constante. La chaîne MCMC est destinée à fournir un échantillon (X, Θ) suivant $\pi(\Theta, X/Y)$, les (Θ) de cet échantillon suivent alors approximativement la distribution $\pi(\Theta/Y)$. A partir de cet échantillon, on peut par une méthode de « Kernel density estimation » obtenir une approximation de $\pi(\Theta, X/Y)$:

$$\hat{\pi}(\Theta/Y) = \frac{1}{M} \sum_{j=1}^M K_h[\Theta - \Theta^{(j)}] \quad (53)$$

avec K_h le noyau de l'estimateur. La méthode de « Kernel density estimation » revient en fait à calculer une moyenne des paramètres de l'échantillon $\Theta^{(l:m)}$ en attribuant plus de poids aux échantillons proches de Θ (le poids venant du noyau de l'estimateur de densité). L'étape suivante est de recalculer $L(\Theta)$ à partir de $\hat{\pi}(\Theta/Y)$. D'après l'équation 45, on a:

$$L(\Theta) = \int P(Y/X, \Theta) P(X/\Theta) dX = \pi(\Theta/Y) \frac{P(Y)}{P(\Theta)} \quad (54)$$

On peut donc obtenir une estimation de $L(\Theta)$ par la relation:

$$\hat{L}(\Theta) = \frac{P(Y)}{M} \sum_{j=1}^M K_h \frac{[\Theta - \Theta^{(j)}]}{P(\Theta^{(j)})} \propto \frac{1}{M} \sum_{j=1}^M K_h \frac{[\Theta - \Theta^{(j)}]}{P(\Theta^{(j)})} \quad (55)$$

Il suffit donc de chercher Θ qui maximise l'équation 44.

Techniquement, ce calcul ne pose donc aucun problème majeur **si on arrive à construire une chaîne MCMC suivant $\pi(\Theta, X/Y)$.**

4.2 Construction de la chaîne MCMC

4.2.1 Quelques éléments sur les chaînes MCMC - algorithme de Metropolis Hastings (Gilks et al., 1995)

Les chaînes de Markov sont des chaînes au cours desquelles une variable (qui peut être un vecteur, une matrice...) d'état évolue de façon itérative (notons S_t l'état à l'itération t). Le passage de S_t vers S_{t+1} se fait par un noyau de transition qui traduit les probabilités de passer d'un état vers un autre. Le but de la chaîne MCMC est de construire à partir d'une chaîne de

Markov un échantillon convergeant vers une distribution souhaitée $\pi(S)$. Metropolis et Hastings ont proposé un premier algorithme assez simple tels que la distribution engendrée par la chaîne MCMC converge à l'infini vers la distribution. Une itération de l'algorithme se décompose en deux étapes:

- A partir de l'état S_t (qui dans notre cas serait l'union de X_t et Θ_t), on propose un nouvel état S' suivant une loi de proposition $q(.|S_t)$.
- On accepte ou pas le nouvel état avec une probabilité égale à

$$\alpha(S', S_t) = \min\left(1, \frac{\pi(S')q(S_t|S')}{\pi(S_t)q(S'|S_t)}\right) \quad (\text{dans notre cas il faut remplacer } \pi(S) \text{ par } \pi(X, \Theta/$$

$Y))$. Si l'état est refusé, la chaîne reste dans l'état S_t .

On peut démontrer que par cet algorithme, on converge vers la distribution $\pi(S)$ quelque soit la distribution $q(.|.)$ choisie (voir (Gilks *et al.*, 1995)). Toutefois, l'efficacité de l'algorithme est fortement lié au choix de $q(.|.)$.

L'algorithme de Métropolis réside sur des distributions symétriques, c'est à dire que $q(S_t|S')=q(S'|S_t)$. La probabilité d'acceptation se réduit donc à $\alpha(S', S_t) = \min\left(1, \frac{\pi(S')}{\pi(S_t)}\right)$. Cela s'apparente donc à une marche aléatoire. L'efficacité de l'algorithme dépend donc essentiellement du choix du paramètre d'échelle de la marche. Au vu de notre méconnaissance de la fonction de vraisemblance, un choix pertinent de paramètre d'échelle paraît difficile. Il faut souligner que les éléments de S peuvent également être mis à jour de façon séquentielle par blocs ou par éléments individuels (on parle de « single component Metropolis-Hastings algorithm »).

4.2.2 Par un échantillonneur de Gibbs (Gilks et al., 1995)

L'échantillonneur de Gibbs est un cas particulier de single component Metropolis-Hastings algorithm. Les éléments de S_i de S sont traités séquentiellement individuellement et la distribution de proposition est $q(S'_i/S_{t,i}, S_{t,-i}) = \pi(S'_i/S_{t,-i})$ avec $S_{t,-i} = \{S_{t+1,1} \dots S_{t+1,i-1}, S_{t,i-1} \dots S_{t,h}\}$. On démontre alors que la probabilité d'acceptation est toujours 1, donc toute proposition est acceptée ce qui rend l'algorithme très efficace. Le problème par contre est qu'il faut savoir simuler selon $\pi(S'_i/S_{t,-i})$ ce qui est loin d'être notre cas...

4.2.3 Par un échantillonneur hybride

Il est possible dans certains cas (en l'occurrence dans le notre) d'adapter l'algorithme de Metropolis-Hastings avec l'échantillonneur de Gibbs. Supposons que l'on puisse écrire:

$$\pi(S'_i/S_{t,-i}) \propto p(S'_i/S_{t,-i}) \pi_{pro}(S'_i/S_{t,-i}) \quad (56)$$

que l'on puisse simuler selon la distribution $\pi_{pro}(S'_i/S_{t,-i})$ et que l'on sache évaluer à une constante près $p(S'_i/S_{t,-i})$. On peut alors procéder comme suit:

- générer un candidat potentiel suivant $\pi_{pro}(S'_i/S_{t,-i})$
- on l'accepte avec la probabilité $\alpha(S'_i/S_{t,-i})$ avec

$$\alpha(S'_i, S_{t,-i}) = \min\left(1, \frac{p(S'_i/S_{t,-i})}{p(S_{t,-i}/S'_i)}\right) \quad (57)$$

Cette solution est théoriquement applicable à notre modèle, par contre des essais démontrent que les temps de simulations sont beaucoup trop importants pour leur utilisation.

*4.2.4 Par un algorithme couplant algorithme génétique et algorithme de Gibbs
(Yardim et al., 2005)*

On a donc vu que la construction d'un échantillon par MCMC pouvait être assez long, en partie du fait de l'exploration de régions de l'espace à peu de poids dû à la méconnaissance que l'on a de la fonction de vraisemblance. L'algorithme génétique est une méthode d'optimisation qui permet d'explorer rapidement l'espace en se concentrant sur les régions à forte vraisemblance. La population ainsi créée n'ayant pas de propriétés statistiques, elle n'est pas directement utilisable autrement que pour de l'optimisation. Par contre il pourrait être intéressant d'utiliser les points de l'espace explorés par l'algorithme génétique pour construire une chaîne MCMC. Cette idée a été utilisée par Yardim et al. (2005) (Yardim *et al.*, 2005) : l'algorithme génétique est utilisé pour fournir une approximation de $P(X_i/X_{-i})$. Pour cela, l'espace est découpé en autant de volumes (appelé cellules de Voronoï) que de solutions évaluées par l'algorithme génétique. Une cellule autour d'un point x_i de l'algorithme génétique contient l'ensemble des points de l'hyper-espace tels que x_i soit plus le point de l'algorithme génétique le plus proche en distance euclidienne. Il considère que les points dans cette cellule ont la même vraisemblance que x_i . On peut alors facilement simuler selon $\pi(S'_i/S_{t-i})$ (voir <http://www.mpl.ucsd.edu/people/gerstoft/crete05/2gerstoft.pdf>) et donc construire un échantillonneur de Gibbs. Le problème est que au vu du nombre de dimensions de notre problème, la construction des cellules paraît impossible.

Conclusion

Pour le moment aucune technique recensée ne semble pouvoir s'adapter à un problème aussi multidimensionnel. En l'état, seule une alternative semble possible : considérer la L_{EV} comme c'est encore souvent fait, ou ne pas prendre en compte d'effets aléatoires dans le modèle avec le risque de problème de convergence que cela implique. Le développement de nouvelles méthodes comme le couplage avec l'algorithme génétique, semble constituer une perspective intéressante pour l'avenir. Une autre possibilité consisterait à faire une analyse de sensibilité du modèle. Les paramètres sur lesquels le modèle serait peu sensible seraient

estimés par la L_{EV} , les paramètres restant étant estimés par la vraisemblance « classique ».

Un autre aspect est qu'il n'existe aucun critère permettant de contrôler « parfaitement » la convergence des chaînes MCMC vers la distribution cible, cela étant d'autant plus problématique que la dimension du problème est importante.

Résumé

Bien que le merlu de l'Atlantique Nord-Est ait été largement étudié, certaines caractéristiques de sa biologie sont encore assez mal connues. C'est en particulier le cas de ses migrations saisonnières et de sa croissance, dont une campagne récente de marquage montre qu'elle était probablement largement sous-estimée. Ces incertitudes ont plusieurs conséquences tant sur la qualité des diagnostics établis à l'aide des modèles de dynamique des populations usuels, que sur la gestion en limitant notamment les possibilités d'évaluation de mesures de gestions spatialisées. Pour améliorer l'estimation des paramètres de ces processus, nous avons développé un modèle de dynamique de population intégré spatialisé et structuré en longueur reposant sur un modèle d'états décrivant la dynamique de la population et de l'activité de pêche, et sur un modèle décrivant les processus d'observation, permettant l'écriture d'une fonction de vraisemblance.

Une revue bibliographique des modèles structurés en longueur existants a mis en évidence le rôle clé de la modélisation du processus de croissance dans ce type de modèles. Pour s'assurer de la robustesse des hypothèses du modèle liées à la discrétisation de ce processus continu et à la variabilité individuelle de croissance, nous avons réalisé une analyse de sensibilité du modèle de croissance. Cette méthode d'exploration numérique du modèle repose sur (i) le développement de plans d'expériences et (ii) l'ajustement de modèles statistiques aux sorties du modèle pour quantifier l'impact de ces hypothèses. Cette démarche générique est transposable à toute analyse de sensibilité d'un modèle discret décrivant un processus continu. Dans notre cas, elle a permis de mettre en évidence que pour le merlu de l'Atlantique Nord-Est il était préférable de choisir un pas de temps trimestriel, des classes de 1 cm, une distribution gamma des incréments de croissance et une distribution uniforme des individus au sein des classes.

Au cours du développement du modèle, une attention particulière a été portée à l'écriture de la fonction de vraisemblance afin qu'elle soit robuste et cohérente avec les processus d'observation. Différents algorithmes ont également été testés pour réaliser l'optimisation numérique de cette fonction, le choix se portant finalement vers un algorithme de type quasi-Newton.

L'ajustement du modèle de dynamique de population aux observations disponibles a permis l'estimation des taux de migration et du taux de croissance, plus faible que celui estimé à partir des données de marquage. Il a permis également d'établir un diagnostic en terme d'évolution de la biomasse féconde cohérent avec les diagnostics préalablement établis à l'aide du modèle actuel d'évaluation. Cependant, du fait de sa complexité et très certainement de la mauvaise qualité de certaines observations, l'ajustement du modèle est difficile. Le modèle souffre de certains problèmes d'identifiabilité même si les paramètres de migration et de croissance ne semblent pas concernés.