

L'Institut Agro Rennes-Angers
 Site d'Angers Site de Rennes

Année universitaire : 2022 - 2023

Spécialité : SML – Biologie

Spécialisation (et option éventuelle) :

Sciences Halieutiques et Aquacoles (REA)

Mémoire de fin d'études

- d'ingénieur de l'Institut Agro Rennes-Angers (Institut national d'enseignement supérieur pour l'agriculture, l'alimentation et l'environnement)
- de master de l'Institut Agro Rennes-Angers (Institut national d'enseignement supérieur pour l'agriculture, l'alimentation et l'environnement)
- de l'Institut Agro Montpellier (étudiant arrivé en M2)
- d'un autre établissement (étudiant arrivé en M2)

Standardisation des données de captures par unité d'effort du rouget barbet de roche (*Mullus surmuletus*) du stock de la zone Ouest de l'Atlantique Nord Est (zones CIEM 6, 8, et les divisions 7.a–c, 7.e–k, et 9.a)

Par : Clémentin DENIAU



Soutenu à **Rennes** le **14/08/2023**

Devant le jury composé de :

Président : Etienne Rivot

Maître de stage : Nathalie Caill-Milly

Enseignant référent : Etienne Rivot

Autres membres du jury (Nom, Qualité)

Didier Gascuel : Enseignant-chercheur, Institut Agro

Jean-Baptiste Lecomte : Ingénieur en modélisation, Ifremer

Les analyses et les conclusions de ce travail d'étudiant n'engagent que la responsabilité de son auteur et non celle de l'Institut Agro Rennes-Angers

Confidentialité

Non Oui si oui : 1 an 5 ans 10 ans

Pendant toute la durée de confidentialité, aucune diffusion du mémoire n'est possible ⁽¹⁾.

Date et signature du maître de stage ⁽²⁾ : 14/09/2023

(ou de l'étudiant-entrepreneur)

A la fin de la période de confidentialité, sa diffusion est soumise aux règles ci-dessous (droits d'auteur et autorisation de diffusion par l'enseignant à renseigner).

Droits d'auteur

L'auteur ⁽³⁾ Deniau Clémentin

autorise la diffusion de son travail (immédiatement ou à la fin de la période de confidentialité)

Oui Non

Si oui, il autorise

la diffusion papier du mémoire uniquement (4)

la diffusion papier du mémoire et la diffusion électronique du résumé

la diffusion papier et électronique du mémoire (joindre dans ce cas la fiche de conformité du mémoire numérique et le contrat de diffusion)

(Facultatif) accepte de placer son mémoire sous licence Creative commons CC-By-Nc-Nd (voir Guide du mémoire Chap 1.4 page 6)

Date et signature de l'auteur :

15/09/2023

Autorisation de diffusion par le responsable de spécialisation ou son représentant

L'enseignant juge le mémoire de qualité suffisante pour être diffusé (immédiatement ou à la fin de la période de confidentialité)

Oui Non

Si non, seul le titre du mémoire apparaîtra dans les bases de données.

Si oui, il autorise

la diffusion papier du mémoire uniquement (4)

la diffusion papier du mémoire et la diffusion électronique du résumé

la diffusion papier et électronique du mémoire

Date et signature de l'enseignant :

14/09/2023

(1) L'administration, les enseignants et les différents services de documentation de l'Institut Agro Rennes-Angers s'engagent à respecter cette confidentialité.

(2) Signature et cachet de l'organisme

(3) Auteur = étudiant qui réalise son mémoire de fin d'études

(4) La référence bibliographique (= Nom de l'auteur, titre du mémoire, année de soutenance, diplôme, spécialité et spécialisation/Option)) sera signalée dans les bases de données documentaires sans le résumé

Remerciements

Un grand merci à Nathalie Caill-Milly de m'avoir permis de faire ce stage. Merci à elle de m'avoir accueilli dans son bureau tout au long de ce stage mais surtout pour sa présence malgré un programme très chargé, notamment le passage d'une HDR.

Merci à mes encadrantes, Nathalie Caill-Milly, Noëlle Bru et Claire Kermorvant, pour leur accompagnement récurrent qui fût d'une grande aide.

Merci à l'équipe féminine du LERAR d'Anglet de m'avoir accueilli. Je souhaite également remercier le service RH de l'IFREMER pour sa réactivité.

Merci à Etienne Rivot, professeur encadrant, pour sa réactivité et ses bons conseils.

Encore un merci pour toutes les personnes autour de moi, de près ou de loin, durant cette période si spéciale de la fin des études.

Un dernier merci au service public pour le financement de l'impression de mon manuscrit.

“Arriver tous les jours à son travail avec une heure de retard est un signe de ponctualité.”

Jacques Sternberg

Table des Matières

1	Introduction	1
2	Matériel et méthodes.....	3
2.1	Objet d'étude	3
2.2	Prise en main des données.....	4
2.3	Démarche de modélisation	6
2.4	Construction des indices d'abondances.....	15
3	Résultats	16
3.1	Sélection d'une flottille de référence.....	16
3.2	Création des variables de ciblage	19
3.3	Construction du modèle.....	20
3.4	Évaluation des modèles finaux et construction des indices d'abondances.....	22
4	Discussion	30
4.1	Division du stock.....	30
4.2	Sélection de l'engin de référence	30
4.3	Variables de ciblage	31
4.4	Prédiction des séries de CPUE	31
4.5	Comparaison avec les résultats obtenus dans le projet ROMELIGO.....	32
4.6	Utilisation des indices d'abondances.....	32
5	Conclusion.....	34
6	Bibliographie	35
	Annexes.....	37

Abréviations

ACOST : Amélioration des COonnaissances sur des STocks

ACP : Analyse en Composantes Principales

AIC : Critère d'information d'Akaike

CIEM : Conseil International pour l'Exploration de la Mer

CPUE : Captures Par Unité d'Effort

DLS : Data Limited Stock

FPC : Flotte de Pêche Communautaire

GDG : Golfe de Gascogne

GNS : Filets maillants calés de fond

GTR : Filets trémail

NA : Non Applicable

OTB : Chaluts de fond à panneaux

OTT : Chaluts jumeaux à panneaux

ROMELIGO : Amélioration des connaissances halieutiques du ROuget-barbet, du MErlan et du LLeu jaune du GOLfe de Gascogne

SDN : Sennes danoises

SIH : Système d'Informations Halieutiques

TAC : Total Admissible de Capture

VMS : Vessel Monitoring System

Liste des figures

Figure 1 A : Carte de la délimitation actuelle des stocks de rouget-barbet de roche telle que définie par le CIEM. Vert = stock Est, Rouge = stock Ouest. B : Carte de délimitation des sous-stocks du stock Ouest nouvellement délimités et pris en compte dans ce travail : Jaune = Stock Golfe, Violet = Stock Mixte.	3
Figure 2 Évolution des débarquements officiels de rouget-barbet de roche (<i>Mullus surmuletus</i>) tous pays confondus (ICES Advice, 2020).	4
Figure 3 Organigramme du processus de filtre des données utilisées issues de Sacrois.	6
Figure 4 Évolution de l'effort de pêche (temps de pêche navire en heures) des deux zones considérées.	7
Figure 5 Processus simplifié de la sélection d'une flottille de référence.	7
Figure 6 Organigramme du processus de sélection d'un groupe de navires.	8
Figure 7 Évolution des débarquements totaux annuels du <i>Mullus surmuletus</i> des zones 6 et 8, et dans les divisions 7.a-c, 7.e-k, et 9.a, tous engins confondus.	9
Figure 8 Histogramme des débarquements mensuels de <i>Mullus surmuletus</i> pour les 4 engins principaux dans la zone Mur Ouest.	9
Figure 9 Évolution des débarquements en fonction de la taille du navire et des deux sous-stock.	10
Figure 10 Diagramme en barres des débarquements annuelles de <i>Mullus surmuletus</i> en fonction des engins pour les stocks Golfe (gauche) et Mixte (droite).	16
Figure 11 Évolution des débarquements annuels de <i>Mullus surmuletus</i> des zones 6 et 8, et dans les divisions 7.a-c, 7.e-k, et 9.a avec l'engin OTB.	17
Figure 12 Nombre totale de séquence par rectangle statistique sur toute la période avant et après les filtres sur les flottilles. Les rectangles rouges sont ceux qui ont disparus de l'analyse.	19
Figure 13 Évolution des débarquements de la flottille sélectionnée pour le golfe de Gascogne selon différents seuils de débarquements annuels des navires.	18
Figure 14 Évolution des débarquements de la flottille sélectionnée en manche Ouest/mer Celtique selon différents seuils de débarquements annuels des navires.	18
Figure 15 Évolution des débarquements de la flottille sélectionnée pour le golfe de Gascogne selon différents seuils de présence des navires.	18
Figure 16 Évolution des débarquements de la flottille sélectionnée en manche Ouest/mer Celtique selon différents seuils de présence des navires.	18
Figure 17 Step plot de la sélection de modèle du stock Mixte (haut) et Golfe (bas) de la première période (gauche) et seconde période (droite) représentant les indices de CPUE à chaque étape de la sélection progressive des variables. Chaque courbe montre un indice de CPUE standardisé et normalisé par la moyenne au fur et à mesure que chaque variable explicative est ajoutée au modèle. Les lignes en pointillés rouges représentent la CPUE nominale.	22
Figure 18 Validations du diagramme Quantile-Quantile (Q-Q) pour le modèle GAM. Le diagramme QQ montre approximativement la ligne droite ($y = x$). La pente de cette ligne droite est l'écart-type ; le diagramme QQ permet de vérifier visuellement la normalité.	23
Figure 19 Histogramme des résidus de déviations des quatre modèles	24
Figure 20 Résidus de déviations en fonction du prédicteur linéaire. Prédications dans l'échelle de la fonction de spline qui explique les valeurs négatives.	24
Figure 21 Valeur réponse en fonction de la valeur prédite.	24
Figure 22 Coefficient de la spline bivariée spatiale des quatre modèles. $S(x)$ représente l'écart à la moyenne. Une couleur jaune sera interprétée comme des valeurs fortes et bleues comme des valeurs faibles. A : Stock Golfe première période. B : Stock Golfe seconde période. C : Stock Mixte première période. D Stock Mixte deuxième période.	25
Figure 23 Coefficient de la spline mensuelle des quatre modèles. $S(x)$ représente l'écart à la moyenne. L'erreur standard est représentée par la bande grise. A : Stock Golfe première période. B : Stock Golfe seconde période. C : Stock Mixte première période. D Stock Mixte deuxième période.	25

Figure 24 Effet des axes principaux représentés par des contours plots. A : Stock Golfe première période. B : Stock Golfe seconde période. C : Stock Mixte première période. D Stock Mixte deuxième période. Les isoplèthes représentent l'effet combiné des deux axes représentés.	26
Figure 25 Effet de l'interaction mois*année spécifié par l'argument 'by' qui ajuste donc une spline mensuelle par année. Chacune de ces splines sont alors pénalisée différemment. Les 4 panneaux présentent les coefficients de splines de l'effet mensuel pour chaque année et ce pour les 4 modèles. Le stock Mixte en haut, Golfe en bas pour la première période (first) et seconde période (second). Une ligne horizontale à 0 correspond à une absence d'effet mensuel pour l'année représentée.	27
Figure 26 Indices d'abondances nominale et standardisées des deux stocks. CPUE nominale calculée comme la moyenne du rapport des débarquements sur le temps de pêche navire. CPUE standardisée issues des prédictions des modèles GAMM. Les indices ont été normalisés par la moyenne. La ligne horizontale correspond à la valeur moyenne égale à un.	28
Figure 27 Indices de CPUE standardisés (lignes en pointillés) avec intervalles de confiance à 95% (bande grise) basés sur le modèle F et débarquements totaux issus de Sacrois (lignes continues). La ligne horizontale en pointillés correspond à la moyenne des CPUE. A : Stock Golfe première période. B : Stock Golfe seconde période. C : Stock Mixte première période. D Stock Mixte deuxième période...	29
Figure 28 Moyenne des CPUE obtenues à partir de la méthode du "data-filtering"	32

Liste des tableaux

Tableau 1 Coefficient de corrélation de Spearman entre les variables des caractéristiques des navires pour les données des stocks Golfe (gauche) et Mixte (droite). * = Pvalue < 0.05	10
Tableau 2 Récapitulatif du jeu de données issu des filtres et navires sélectionnés à la suite des filtres.	17
Tableau 3 Résumé des résultats de la construction de la variable de ciblage.....	19
Tableau 4 Résumé des informations relatives à la sélection des modèles. Les paramètres utilisés pour évaluer la performance des modèles comprennent le degré de liberté de chaque variable ajoutée, le delta de la déviance expliquée, le delta AIC. Les valeurs en gras représentent les modèles les plus performants pour une mesure donnée. *Modèles pour lesquels la variable spatiale a été remplacée par le produit tensoriel.	21

Liste des annexes

Annexe 1 Tableau des variables sélectionnées issues de l'algorithme Sacrois (Demanèche et al., 2015).	1
Annexe 2 Heat map des débarquements annuels par année pour les engins OTB du golfe de Gascogne.	6
Annexe 3 Heat map des débarquements annuels par année pour les engins OTB de la zone 6-7.....	6
Annexe 4 Boxplot des débarquements par taille de navire sur toute la zone.	4
Annexe 5 Cartographie des débarquements totaux de <i>Mullus surmuletus</i> cumulés sur la période 2000-2022 par rectangle statistique.	4
Annexe 6 Cartographie des débarquements totaux de <i>Mullus surmuletus</i> cumulés sur la période 2000-2022 par rectangle statistique pour les navires mettant en place l'engin OTB (Chalut de fond à panneaux).	5
Annexe 7 Évolution de l'effort de pêche total du stock Golfe en fonction de la taille des navires	12
Annexe 8 Évolution de l'effort de pêche total du stock Mixte en fonction de la taille des navires	12
Annexe 9 Graphiques des variables des ACP du stock Golfe pour la période 2000-2008.....	7
Annexe 10 Graphiques des variables des ACP du stock Mixte pour la période 2000-2008	7
Annexe 11 Graphiques des variables des ACP du stock Golfe pour la période 2009-2022.....	8
Annexe 12 Graphiques des variables des ACP du stock Mixte pour la période 2009-2022	8
Annexe 13 Processus de modification des données et de création de sous-ensembles.	13

Annexe 14 Processus de sélection des données avant modélisation du stock Mixte (processus identique pour la zone golfe de Gascogne).	13
Annexe 15 Processus de création de la variable de ciblage du stock Mixte (processus identique pour la zone golfe de Gascogne).	5
Annexe 16 Série temporelle des débarquements de rouget-barbet de roche. Haut = Deux sous-stocks étudiés. Bas = Zone totale	3
Annexe 17 Récapitulatif de l'ajout de chaque variable à partir la structure aléatoire des 4 modèles. On regarde ici variable par variable en fonction de la structure aléatoire.....	9
Annexe 18 Ajustement aux données des 4 modèles. La courbe en pointillée rouge correspond à la CPUE nominale.....	10
Annexe 19 Valeurs des indices d'abondances finaux	11

1 Introduction

Le rouget-barbet de roche (*Mullus surmuletus*, Linnaeus, 1758) est présent le long des côtes de l'Europe, du Sud de la Norvège et du Nord de l'Écosse (Atlantique Nord, mer Baltique, mer du Nord et Manche), jusqu'au Nord de l'Afrique de l'Ouest, dans le bassin méditerranéen et dans la mer Noire (Hureau, 1986 ; Mahé et al., 2005). En Atlantique Nord-Est la population peut être séparée en trois zones géographiques distinctes (Benzinou et al., 2013). Le CIEM s'est appuyé sur ces travaux pour définir deux stocks : un à l'Est (sous-zone CIEM 4 et divisions 7.d et 3.a) et l'autre à l'Ouest (sous-zones CIEM 6, 8, et divisions 7.a–c, 7.e–k, 9. a) qui est le regroupement de deux zones géographiques.

Le stock à l'Est, classé dans la catégorie 3 du CIEM, est évalué avec un modèle structuré en âge (ICES, 2019). En revanche, la gestion du stock à l'Ouest repose sur une démarche adoptée en 2012 pour les stocks présentant des données insuffisantes (Data Limited Stock, DLS) pour une évaluation analytique (ICES, 2012). Ainsi, depuis 2012, le CIEM émet des avis sur ce stock de rouget-barbet de roche pour 3 ans (ICES Advice, 2020). Le dernier datant de 2020 mentionne que les débarquements annuels ne devraient pas dépasser 1 280 tonnes pour ce stock, soit 20 % de moins que les 3 années précédentes (ICES Advice, 2020). Il n'est cependant pas soumis à un total admissible de capture (TAC), alors qu'il s'agit d'une espèce d'intérêt commercial pour les professionnels. Son exploitation dans le stock Ouest représente des débarquements annuels compris entre 800 t et 2 400 t sur la période 2000-2022. Bien que les captures récentes soient à un niveau bas (1 100 t en 2022), cette espèce représente un revenu essentiel pour les navires qui l'exploitent, avec un prix moyen à 17,6 €/kg en 2021 (FranceAgriMer, 2021). Aujourd'hui classé dans la catégorie 5 du CIEM, l'état de ce stock tend à être évalué et des indices de suivis du stock fixés afin de mettre en place une gestion adaptée.

Classiquement, la capture par unité d'effort (CPUE) est utilisée en tant qu'indice d'abondance (mesure relative de la biomasse d'individus d'un stock). Elle correspond à l'indice d'abondance du stock le plus immédiat mais est souvent considérée comme biaisée et non strictement proportionnelle à l'abondance. C'est notamment le cas des données de pêches professionnelles en raison de l'efficacité de la pêche non constante et du comportement changeant des pêcheurs. Ces facteurs peuvent, dans une certaine mesure, être contrôlés en appliquant des méthodes de standardisation ou de normalisation par sélection des données. Dans notre cas, nous disposons des données professionnelles issues de la base de données Sacrois¹ qui nous fournit une information de débarquement et d'effort ainsi que des covariables associées (caractéristiques des navires, engins de pêche, rectangles statistiques de la capture, espèces capturées conjointement, etc...).

Concernant le rouget-barbet de roche, un premier travail a été réalisé de 2014 à 2017 lors du projet ROMELIGO. Ce dernier a permis de développer des indices d'abondance en identifiant des flottilles de références et d'améliorer les connaissances biologiques sur ce stock (Léauté et al., 2018). Les indices de suivi de stock issus de ce projet ont été développés à l'aide d'une méthode de data-filtering assimilé à « *design-based* » (Caill-Milly et al., 2019). Cette méthode de standardisation utilisant de nombreux filtres sur les données a permis d'observer une diminution de l'abondance pour la période 2000-2015 au travers de deux flottilles de référence à pavillon français (chalut de fond à panneaux et filet droit) (Léauté et al., 2018).

Le projet ACOST, qui porte également sur trois autres espèces présentes dans le golfe de Gascogne (lieu jaune, merlan et maigre), se situe dans le prolongement de ROMELIGO avec pour objectif d'améliorer les connaissances des stocks concernés. Au sein de ce projet et en

¹ L'algorithme Sacrois du système d'information halieutique (SIH) croise, rapproche, vérifie et contrôle différentes sources d'informations (données des ventes en criées, suivi satellitaire de certains navires, enquêtes sur l'activité des navires de pêche) pour produire un ensemble de données de statistiques de pêches officielles de référence.

parallèle à l'approche « *design-based* », une approche dite « *model-based* » est utilisée pour construire des indices d'abondances. Plus couramment utilisée pour les évaluations de stock, une revue non exhaustive des méthodes qu'elle regroupe est présentée par Maunder et Punt (2004). Le processus de standardisation de l'indice d'abondance nominale est généralement employé via des méthodes basées sur des modèles linéaires généralisés (GLM), des modèles additifs généralisés (GAM) (Wood, 2017) ou des modèles spatiotemporels (Grüss et al., 2019). Ces approches fondées sur des modèles contrôlent l'effet des facteurs affectant les CPUE (par exemple, le type d'engin, la localisation et les caractéristiques des navires), et la tendance annuelle extraite du modèle est interprétée comme un indice standardisé de l'abondance relative (Maunder et Punt, 2004).

L'espèce étudié ici, à l'instar du merlan, est capturée dans le cadre d'une pêcherie mixte². De récents travaux ont permis de développer des méthodes de standardisation pour des espèces présentes dans des pêcheries multi-spécifiques (Okamura et al., 2018 ; Winker et al., 2013). Nous verrons comment ce comportement peut être inclus dans le cadre de la pêcherie du rouget-barbet de roche. L'espèce étudiée peut être caractérisée d'"accessoire valorisée" (Julien, 2018). Le fait qu'elle soit valorisée est une des raisons pour lesquelles nous l'étudions mais son caractère accessoire dans la pêcherie implique un nombre conséquent de zéros dans le jeu de données. Le travail de standardisation doit donc prendre en compte un comportement de ciblage ainsi qu'un grand nombre de zéros dans le jeu de données disponible.

Ainsi, ce travail à un double but. D'une part il doit répondre à la demande du projet ACOST et plus généralement du CIEM qui est d'accroître les connaissances sur ce stock, améliorer son évaluation et *in fine* sa gestion. D'autre part, il a vocation à comparer les indices d'abondance ainsi construits avec ceux déjà obtenus pour la zone du golfe de Gascogne et issus du projet ROMELIGO. La première partie présentera les données sélectionnées et utilisées. La deuxième partie expliquera la méthode de modélisation appliquée au rouget-barbet de roche pour les zones que nous définirons. Nous considérerons trois méthodes pour rendre compte du ciblage ou de la tactique de pêche. Nous verrons quelles hypothèses seront faites sur la distribution des données. Enfin, nous présenterons les indices d'abondance ainsi construits et comparerons les résultats avec ceux issues du projet ROMELIGO.

² Plusieurs flottilles, plusieurs espèces cibles et accessoires.

2 Matériel et méthodes

2.1 Objet d'étude

2.1.1 Biologie du rouget-barbet de roche

Le rouget-barbet de roche est une espèce benthique vivant sur les fonds sableux, graveleux. Il est présent le long des côtes de l'Europe, du Sud de la Norvège et du Nord de l'Écosse (Atlantique Nord, mer Baltique, mer du Nord et Manche), jusqu'au Nord de l'Afrique de l'Ouest, dans le bassin méditerranéen et dans la mer Noire (Hureau, 1986 ; Mahé et al., 2005). Les jeunes poissons se répartissent dans les zones côtières à faible salinité, tandis que les adultes ont une distribution plus hauturière. La maturité sexuelle est atteinte à la moitié de la deuxième année pour les mâles qui arrêtent alors leur croissance, et entre la deuxième et la troisième années pour les femelles, qui continuent leur croissance après ce stade (ICES WGwide, 2020). En Manche, les premiers individus sexuellement matures mesurent environ 16 cm (Mahé et al., 2005 dans ICES WGwide, 2020) alors que dans le golfe de Gascogne, ils mesurent 16 cm pour les mâles et 18 cm pour les femelles. Par ailleurs, 50 % des individus sont matures à 22 cm (Dorel et al., 1986 dans ICES WGwide, 2020). La durée de vie des rougets présents dans les eaux européennes peut aller jusqu'à 11 ans (ICES, 2012b). Les rougets se nourrissent de petits crustacés, d'annélides, de vers et de mollusques.

2.1.2 Délimitation spatiale des stocks

Aujourd'hui, deux stocks de rouget-barbet de roche sont définis et considérés par le CIEM en Europe. Un premier, le stock Est, est constitué des zones 4 et divisions 7.d et 3.a. Le stock Ouest, que nous étudions, comprend les zones 6, 8, et les divisions 7.a–c, 7.e–k, et 9.a (Figure 1, A). Cependant, une étude menée sur la morphométrie des otolithes et des individus montre que la population de *Mullus surmuletus* présente en Europe peut être divisée en trois zones (Benzinou et al., 2013). D'une part le stock Est et d'autre part le stock Ouest coupé en deux. Ainsi, la population du rouget-barbet de roche d'Atlantique Nord-Est se décomposerait de la façon suivante :

- Mer Celtique et Manche Ouest (Stock Est i.e. zones 4 et divisions 7.d et 3.a) ;
- Mer du Nord et Manche Est (zones 6 et divisions 7.a–c, 7.e–k) ;
- Golfe de Gascogne (zone 8).

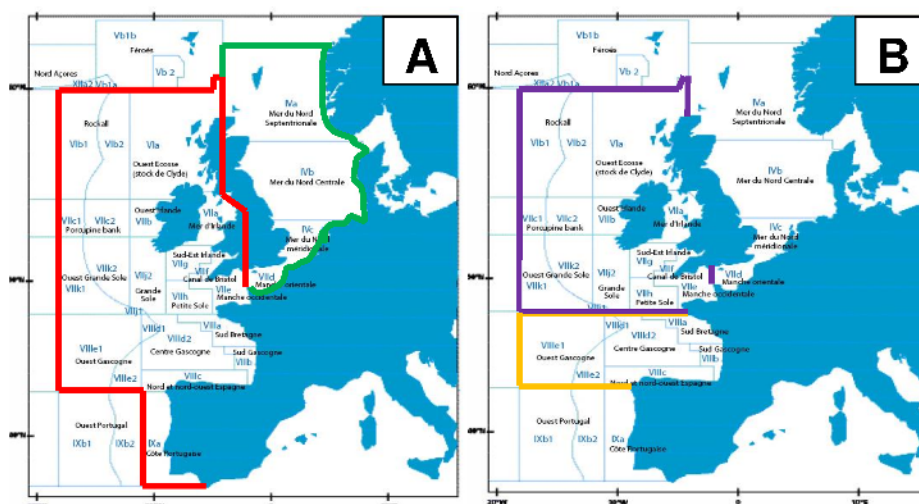


Figure 1 | A : Carte de la délimitation actuelle des stocks de rouget-barbet de roche telle que définie par le CIEM. Vert = stock Est, Rouge = stock Ouest. B : Carte de délimitation des sous-stocks du stock Ouest nouvellement délimités et pris en compte dans ce travail : Jaune = Stock Golfe, Violet = Stock Mixte.

Ces sous-stocks n'ayant pas été définis par le CIEM, nous les appellerons respectivement, et en accord avec les travaux de Benzinou et al., 2013, stock Est, stock Mixte et stock Golfe. Nous considérons ainsi pour cette étude le stock Golfe (Zone 8 CIEM) et le stock Mixte (Zones 6, 7.a-c, 7.e-k CIEM) (Figure 1, B). Un tel découpage permet également de comparer les résultats de la zone golfe de Gascogne avec la méthode mise en place par Caill-Milly et al., 2019. La zone 9 a été mise de côté en raison de son importance marginale pour la zone vis-à-vis des apports dans les débarquements de rouget-barbet de roche.

2.1.3 Évolution des débarquements et description de la pêche.

La pêche du rouget-barbet de roche en Europe existe depuis de nombreuses années. On sait notamment que dans les années 30, cette espèce contribuait de façon non négligeable aux apports de chalutiers (Dardignac, 1988). Les premières valeurs de débarquements officielles ne sont cependant disponibles qu'à partir de 1975 (ICES WGwide, 2020). Les données du SIH (Ifremer) commencent, elles, en 2000. Après une longue période d'apports en hausse, l'espèce voit diminuer ses débarquements depuis 2005 (Figure 2). La majorité des captures est réalisée par la flotte française, parmi laquelle les chalutiers de fond à panneaux sont les principaux contributeurs (Caill-Milly et al., 2017). Viennent ensuite les filets maillants et la senne danoise. Ces trois engins représentaient 98% des apports sur la période 2000-2015 (Léauté et al., 2018).

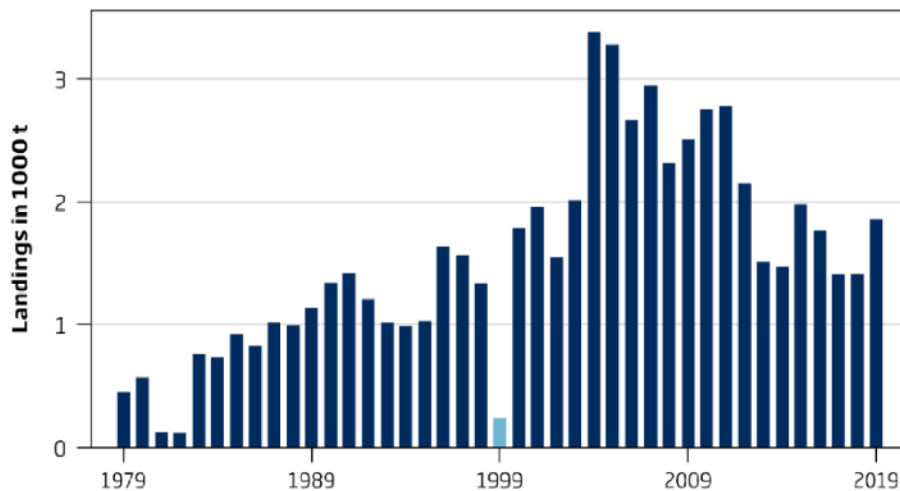


Figure 2 | Évolution des débarquements officiels de rouget-barbet de roche (*Mullus surmuletus*) tous pays confondus des sous-zones CIEM 6, 8, et divisions 7.a-c, 7.e-k, 9. a (ICES Advice, 2020).

2.2 Prise en main des données

2.2.1 Base de données Sacrois et extraction

L'algorithme Sacrois, outil du SIH, produit des séries de données de la pêche professionnelle pour les navires inscrits au fichier Flotte de Pêche Communautaire (FPC) : productions et efforts de pêche validés, consolidés et qualifiés. Cela consiste en un travail de rapprochement, vérification, contrôles de cohérence de différents flux unitaires de données (Demanèche et al., 2015). Les données d'effort et de débarquements pour la période 2000-2022 des zones 6 et 8, et des divisions 7.a-c, 7.e-k, et 9.a ont été extraites. Le jeu de données comprend l'activité de tous les navires français ayant débarqué du rouget-barbet de roche au moins une fois une année sur la période 2000-2022. Chaque ligne de données contient, entre autres :

- Les débarquements (kg) par espèce et la valeur (€) associée ;
- L'identifiant et la durée (heures) de la séquence et la marée associée ;

- L'identification du navire et ses caractéristiques (taille, puissance, tonnage de référence, quartier maritime, année de construction), l'engin de pêche utilisé, le maillage et enfin le rectangle statistique associé au débarquement ;
- La capture par unité d'effort calculée comme le rapport des débarquements sur le temps de pêche navire, que l'on note CPUE et dont l'unité est le kg/h.

La description détaillée des variables extraites de la base de données est fournie en Annexe 1. Le jeu de données, d'une taille de 2Go, compte ainsi 96 637 660 lignes et 30 variables. La mise à jour du 06/06/2023 a permis d'intégrer des corrections sur la variable du temps de pêche pour les années 2021-2022. En effet, les données de 2021 à 2023 présentaient des erreurs dans le champ « temps de pêche » pour les petits navires côtiers de moins de 10 mètres soumis à la déclaration de fiches de pêche. Ces erreurs étaient liées à une mauvaise intégration en base des données de fiches de pêche de ces navires à la suite de leur passage au format ERSv3³ et à leur saisie via l'outil VisioCaptures⁴.

2.2.2 Données anormales et filtres

Le jeu de données présentait de nombreuses valeurs manquantes identifiées avec NA. C'était notamment le cas des quantités débarquées (kg) : 14% des lignes. Depuis 2018 et la réception des données Sacrois au format ERSv3 (e-logbook), une création de « NA » s'opère pour les variables "QUANT_POIDS_VIF_SACROIS" et "MONTANT_EUROS_SACROIS" et explique cette quantité de valeurs non attribuables. En effet, ces données intègrent désormais une information déclarative sur la ZEE de l'activité du navire. Cependant, Sacrois ne dispose pas encore d'un module spécifique de croisement de cette information avec celle des données VMS, ce qui crée des NA. Avant 2018, ces NA viennent du fait que les données brutes et corrigées sont conservées afin de pouvoir retracer l'origine de ces dernières. Cependant, les informations de débarquement et de vente pour les lignes correspondantes aux données brutes qui ont été corrigées par l'algorithme Sacrois ont été supprimées et remplacées par des 'NA'. Nous avons donc supprimé ces lignes.

Le même constat a été fait pour les variables du temps de mer et du temps de pêche navire (TPS_MER/ TP_NAVIRE_SACROIS) qui correspondent à la durée de la marée et de la séquence en heures. Ces marées ou séquences, dites "orphelines", sont particulièrement nombreuses les premières années : on observe des valeurs non attribuables sur la période 2000-2008. Le temps de pêche étant notre métrique d'effort de pêche, nous ne pouvions conserver les valeurs de débarquement associées dans notre démarche. Lors de la construction du modèle, ces lignes ont été supprimées.

La variable d'identification du rectangle statistique comprenait également des NA et les séquences correspondantes ont été supprimées lors de la modélisation. D'autres variables montraient le même problème mais dans de moindres proportions. C'était le cas de la longueur des navires ou encore de l'engin de pêche. Nous avons conservé ces lignes afin de décrire la pêcherie mais elles ont été supprimées ultérieurement dans la partie modélisation.

Les données pour lesquelles les séquences de pêche avaient un poids de rouget pêché inférieur au poids d'un individu (40g) ont été jugées comme données anormales et ramenées à un poids de rouget pêché nul. Nous avons également supprimé les données ayant un temps de pêche strictement inférieur à une heure. Ces deux filtres avaient pour objectif de nettoyer les données aberrantes faibles.

³ Electronic Recording and reporting System. Dernières versions des formats de données électroniques utilisés par le système de collecte des données de pêches.

⁴ Application pour la déclaration des activités de pêche pour les navires de moins de 12 mètres. Opérationnelle depuis fin 2022, elle est déployée progressivement auprès des professionnels durant l'année 2023.

Pour certaines séquences, deux informations de débarquements et de ventes étaient disponibles pour une même espèce capturée par un type d'engin, notamment pour le rouget-barbet de roche. Ces deux informations sont le résultat d'origines différentes, les données déclaratives et de vente. Nous avons additionné ces informations pour n'avoir plus qu'une ligne par espèce, par séquence et par engin.

Nous avons aussi inspecté par la suite les valeurs aberrantes de débarquements et de CPUE de rouget. En accord avec les acteurs professionnels, nous avons supprimé les valeurs les plus extrêmes qui ne reflètent pas la réalité. Les CPUE supérieures à 500 kg/h ont ainsi été supprimées. Ces dernières présentaient un temps de mer de 155h en moyenne et ne contenaient qu'une seule séquence dont le temps de pêche était de 2h à 3h. Les débarquements associés étaient très importants. Ces constatations nous ont incité à supprimer ces données.

Les étapes de traitement des données sont résumées dans la figure ci-dessous (Figure 3)

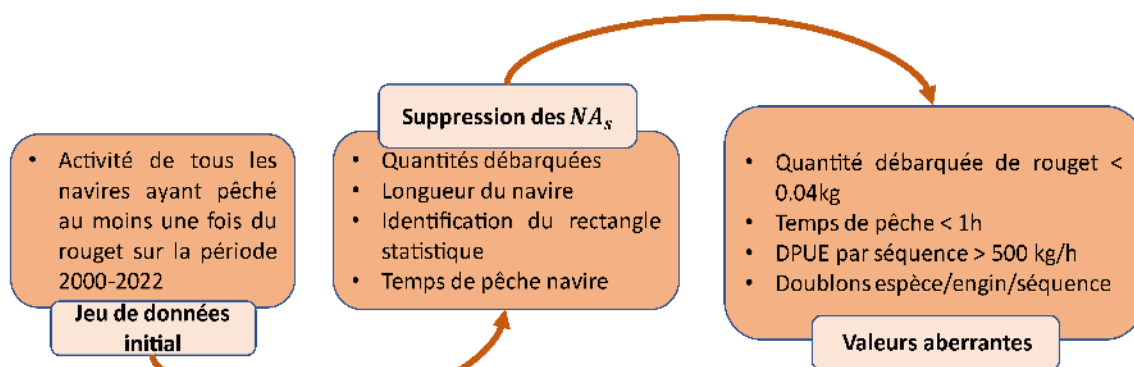


Figure 3 | Organigramme du processus de filtre des données utilisées issues de Sacrois.

2.3 Démarche de modélisation

Des modèles d'abondances pour le merlan et le lieu jaune ont déjà été construit dans le cadre du projet ACOST. Nous adaptons ici ces modèles pour le cas du rouget-barbet de roche.

2.3.1 Hypothèses de modélisation

Idéalement, les indices d'abondance devraient être dérivés d'études scientifiques indépendantes de la pêche professionnelle, telles que celles réalisées par des navires de recherche en mer. Cependant, dans de nombreux cas, en raison des coûts ou d'autres problèmes logistiques, il n'est tout simplement pas possible de réaliser des études indépendantes des pêcheries. Il faut donc utiliser des données de CPUE ou de taux de capture (Candy, 2004) dépendantes des pêcheries pour créer un indice d'abondance en vue de l'évaluation des stocks (Ducharme-Barth, 2021). Dériver un indice d'abondance à partir des données commerciales est donc une procédure commune en halieutique. Elle permet d'ajuster d'autres modèles de production de biomasse ou d'évaluation de stock. Cependant, cette technique se fonde sur des hypothèses fortes issues d'une simple relation qui relie les captures à l'abondance et à l'effort :

$$C_t = qE_tN_t, \quad (1)$$

où C_t est la capture au moment t , E_t est l'effort déployé au moment t , N_t est l'abondance au moment t et q est la portion du stock capturée par une unité d'effort (coefficient de capturabilité). Cette équation peut être réarrangée pour former la relation suivante entre CPUE et abondance :

$$\frac{C_t}{E_t} = qN_t, \quad (2)$$

ce qui rend la CPUE proportionnelle à l'abondance,

$$cpue_t \propto N_t, \quad (3)$$

à condition que q soit constant dans le temps. Malheureusement, cela se vérifie rarement sur l'ensemble de l'historique d'exploitation et peut varier pour de nombreuses raisons. On peut citer le navire, les caractéristiques de l'engin, la localisation et les conditions environnementales (Maunder et al., 2020). L'une des analyses les plus fréquemment effectuées en halieutique, la normalisation de l'effort de pêche (Maunder and Punt, 2004 ; Maunder et al., 2006), vise à normaliser les données relatives à l'effort afin de garantir que q puisse être considérée comme constante et ainsi interpréter la série de CPUE comme indice d'abondance.

2.3.2 Sélection de l'effort de pêche et découpage temporel.

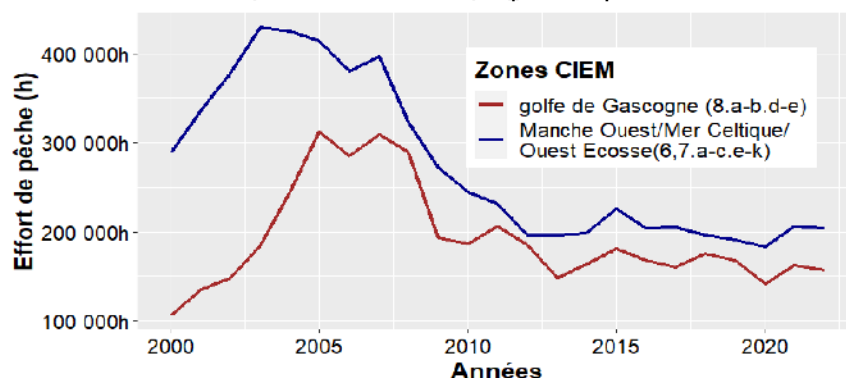


Figure 4 | Évolution de l'effort de pêche (temps de pêche navire en heures) des deux zones considérées pour les engins OTB.

L'effort de pêche est, comme vu plus haut, une métrique essentielle de la standardisation des données de débarquements pour y extraire une tendance annuelle de variation de biomasse. La base de données Sacrois nous donne accès à des valeurs d'efforts différentes que sont par exemple le temps de mer, le temps de pêche, les jours de mer et de pêche. Nous avons fait le choix d'utiliser le "temps de pêche navire" (Figure 4). Il est plus précis que le "temps de mer" et "jours de mer et pêche" car il constitue une information plus fine sur l'activité de pêche. Le calcul de cette variable dans Sacrois est présenté en Annexe 1. On distingue une rupture entre 2008 et 2009 pour les deux séries d'efforts considérées (Figure 4). Le système d'acquisition des données de pêche a en effet changé à cette date, passant de « statpêche » à « SA-CAPT » (FranceAgriMer). Cette modification ne permet pas d'utiliser la série temporelle entière. Nous avons donc décidé de travailler sur deux périodes : [2000 : 2008], [2009 : 2022].

2.3.3 Sélection d'une flottille de référence

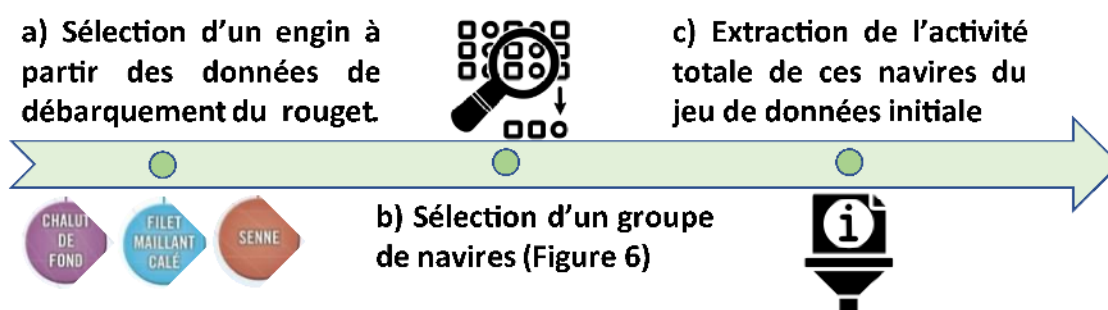


Figure 5 | Processus simplifié de la sélection d'une flottille de référence.

Une fois les données nettoyées (§ 2.2.2), nous disposons d'un jeu de données de 37 millions de lignes pour le stock Mixte et de 36 Millions de lignes pour le stock Golfe. Tel quel, il nous est difficile de standardiser les informations relatives à l'effort de pêche. Nous avons donc sélectionné une plus petite partie du jeu de données afin de rendre possible et interprétable la CPUE (Maunder et al., 2006) selon le processus ci-dessus (Figure 5). Dans cette optique, nous avons regardé en quelles quantités et proportions chaque engin capture cette espèce afin de sélectionner une flottille qui représente la plus large proportion du stock réel. Nous avons porté également notre attention sur l'étendue spatiale des débarquements ainsi que sur la stabilité dans le temps de ces derniers. Les navires mettant en place l'engin identifié lors d'une séquence de pêche sont ensuite filtrés pour affiner la représentativité de la flottille vis-à-vis de l'espèce étudiée. En effet, certains sont peu porteurs d'informations étant donné leur faible occurrence dans les débarquements et le faible tonnage de ces derniers pour l'espèce considérée/d'intérêt. Une sélection a ainsi été faite sur le nombre d'années durant lesquelles les navires ont débarqué du rouget ainsi que sur la moyenne annuelle débarquée par navire sur la période 2000-2022 (Figure 6). Ces choix ont été validés lors de groupes de travail avec les professionnels. La liste des navires issue de ces filtres a ensuite été utilisée afin d'extraire l'activité totale de ces derniers sur toute la période. Les vrais zéros du jeu de données sont les séquences des navires sélectionnés sans débarquements (ou de moins de 40g) de rouget.

2.3.4 Variabilité de la capturabilité

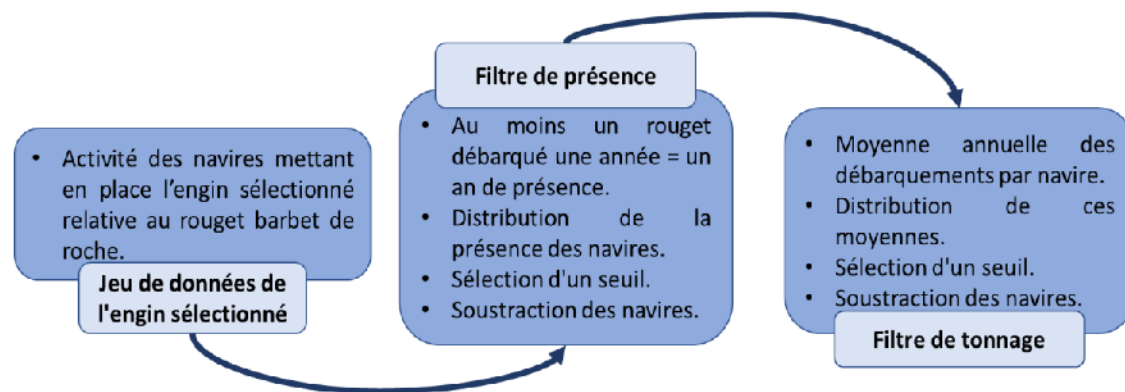


Figure 6 | Organigramme du processus de sélection d'un groupe de navires.

Comme vu plus haut, nous voulions contrôler les influences sur les débarquements autres que celle de l'évolution de l'abondance du stock. Inclure tous les effets possibles n'est pas réalisable et auquel cas ce n'est pas viable au regard du principe de parcimonie. Ainsi, nous avons souhaité sélectionner les variables expliquant le plus de variabilité. On retrouve classiquement des variables temporelles, spatiales, environnementales, techniques (caractéristiques des navires). Les variables disponibles dans la base de données Sacrois et que nous estimions avoir un effet important dans la variabilité de la capturabilité étaient les suivantes :

- Structure temporelle des données (annuelle et mensuelle) ;
- Structure spatiale (rectangle statistique) ;
- Caractéristiques techniques des navires (longueur, puissance, tonnage de référence) ;
- Comportement de pêche (ciblage).

2.3.4.1 Variabilité temporelle

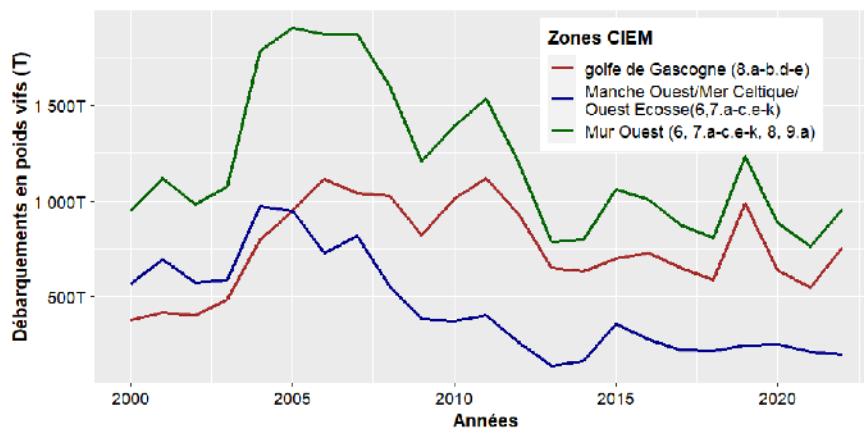


Figure 7 | Évolution des débarquements totaux annuels du *Mullus surmuletus* des zones 6 et 8, et dans les divisions 7.a-c, 7.e-k, et 9.a, tous engins confondus.

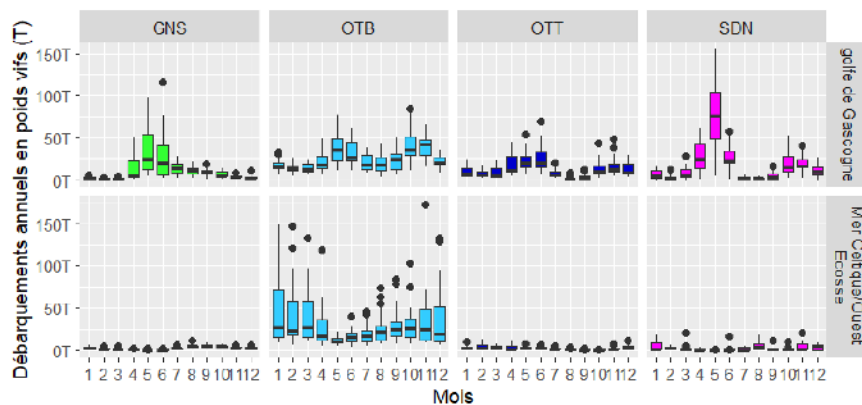


Figure 8 | Histogramme des débarquements mensuels de *Mullus surmuletus* pour les 4 engins principaux dans la zone Mur Ouest.

Plusieurs variables temporelles sont disponibles dans le jeu de données : l'année, le mois et les variables relatives aux temps de mer et de pêche. La tendance annuelle est celle que l'on suppose représenter l'évolution de la biomasse de rouget-barbet de roche. Sa représentation sur la Figure 7, qui est l'évolution de ses débarquements totaux, est bruitée par une capturabilité fluctuante. Globalement, les débarquements croissent jusqu'en 2007 puis décroissent jusqu'à aujourd'hui. Cette diminution commence plus tôt dans le stock Mixte à la différence du Golfe où la diminution des débarquements est plus graduelle (Figure 7). La variable mois, qui représente la saisonnalité, rend compte d'une certaine variabilité des débarquements (Figure 8, Annexe 8, Annexe 7). Les travaux de Kermorvant et al. (2021) sur des données de chalut de fond à panneaux ont déjà mis en évidence une saisonnalité bisannuelle in shore et annuelle offshore. Ici, selon l'engin de pêche utilisé, la saisonnalité est également différente, autour de mai et novembre pour les OTT (Chaluts jumeaux à panneaux) et SDN (Sennes danoises), moins visible pour les OTB (Chaluts de fond à panneaux) et majoritairement en avril-mai pour les GNS (Figure 8). Les travaux de Kermorvant et al. (2021) ont mis en évidence de fortes fluctuations dans le temps de la saisonnalité et confortent ce que l'on observe (Figure 7, Figure 8, Annexe 2). On voit donc la spécificité des années et des mois combinés.

2.3.4.2 Variabilité spatiale

La variable spatiale la plus fine dont nous disposons est le rectangle statistique, soit une surface de 0.5° de latitude par 1° de longitude. Cette information est renseignée automatiquement

pour les navires de plus de 12 mètres disposant d'un système VMS et par déclaration électronique ou manuelle pour les navires de plus petites tailles. Ce découpage nous permet d'observer que le rouget-barbet de roche est capturé en grande majorité au plus proche de la côte et notamment en Manche Ouest et dans la partie centrale du golfe de Gascogne (Annexe 3). Par conséquent, nous pouvons considérer une répartition hétérogène des débarquements de rouget-barbet de roche. Nous inférons les coordonnées géographiques des centroïdes de chaque rectangle à partir de leur identifiant afin de pouvoir manier cette variable de façon quantitative.

2.3.4.3 Caractéristiques des navires

Tableau 1 | Coefficient de corrélation de Spearman entre les variables des caractéristiques des navires pour les données des stocks Golfe (gauche) et Mixte (droite). * = Pvalue < 0.05

	Loa		Power_Main	
Power_Main	0.90*	0.80*		
Ton_Ref	0.95*	0.91*	0.89*	0.78*

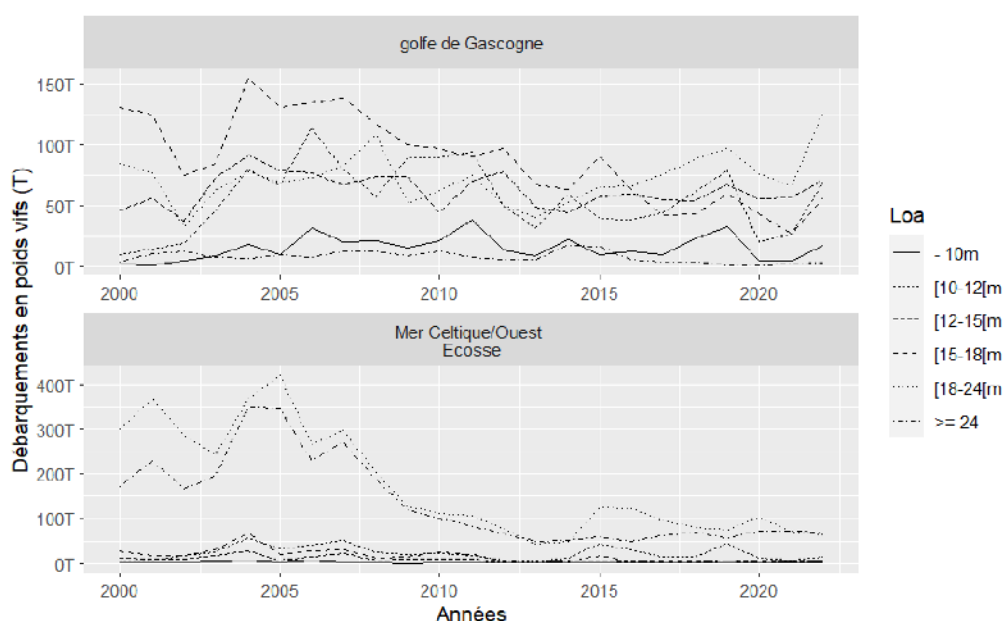


Figure 9 | Évolution annuelle des débarquements en fonction de la taille du navire (Loa) et des deux sous-stock. 'golfe de Gascogne' correspond au stock Golfe et 'Mer Celtique/Ouest Ecosse' correspond au stock Mixte.

Les caractéristiques des navires sont communément admises comme impactant significativement la relation entre taux de capture ou CPUE et abondance. Trois variables nous semblent envisageables : la longueur des navires, leur puissance et leur tonnage de référence. La longueur est renseignée par une variable quantitative dont la mesure est en mètres. Cependant, la base de données Sacrois comprend un trop grand nombre de modalités pour mener nos analyses (Annexe 4). Nous avons donc discrétisé cette variable sur la base de la synthèse des flottilles des chalutiers exclusifs des façades Mer du Nord – Manche et Atlantique (Ifremer, 2021a, 2021b) et ce choix a été validé avec les professionnels lors d'un groupe de travail. Les classes de tailles sont celles présentées sur la Figure 9. On distingue sur cette même figure une hétérogénéité des débarquements en fonction de la taille des navires. Dans le stock Mixte (Mer Celtique/Ouest-Écosse), il apparaît une prédominance de deux tailles de navires ; [18-24[& ≥ 24m ; alors que dans le golfe de Gascogne la distribution semble être plus homogène mais tout de même discriminante. La puissance du navire et le tonnage de référence sont considérés comme variables quantitatives continues du fait qu'aucune typologie n'est disponible. Cependant ces trois variables sont très corrélées entre elles (Tableau 1).

2.3.4.4 Pêcherie mixte : création d'une variable de tactiques de pêches

L'utilisation de la CPUE ou taux de capture comme un proxy de l'abondance assume l'hypothèse d'une capturabilité constante pour l'espèce étudiée et repose donc sur la capacité à normaliser les effets qui influent sur la capturabilité autres que l'abondance (Maunder and Punt, 2004). Dans les pêches plurispécifiques, la capturabilité peut varier selon l'espèce ciblée, de sorte que la tendance annuelle obtenue à partir de la CPUE normalisée est susceptible d'être biaisée. Toutefois, les informations sur les espèces ciblées ne sont généralement pas disponibles (Okamura et al., 2018). Nous avons donc décidé de construire un indicateur de ciblage en se basant sur la tactique de pêche⁵. Dans cet objectif nous avons utilisé l'information relative aux espèces capturées conjointement au rouget afin d'y inférer cet indice. Le jeu de données utilisé est relatif à l'activité totale des navires identifiés lors des étapes précédentes (§ 2.3.3). Rappelons ici que les analyses ont été réalisées sur deux périodes différentes pour des raisons liées à des changements de système d'acquisition des données (§ 2.3.2). Trois possibilités ont été identifiées pour construire l'information de ciblage :

- Méthode des composantes principales (Winker et al., 2013)

Les données de débarquements ont été agrégées par quinzaine, navire, rectangle statistique ainsi que par année. Le choix d'un découpage tous les 15 jours, décidé en groupe de travail, a été considéré comme rendant mieux compte des changements de tactiques de pêches par rapport à un découpage mensuel. Le découpage annuel a été, lui, fait sous l'hypothèse d'une évolution des tactiques de pêche à une échelle plus grande. Ainsi, les débarquements des séquences dans ces agrégations sont égaux à la valeur moyenne par regroupement (somme regroupement / nombre de séquences). Les dix espèces majoritaires en tonnage ont été conservées. Dans le cas où le rouget-barbet de roche ne faisait pas partie de ces 10 premières, il y a été ajouté et l'analyse a été réalisée sur un groupe de 11 espèces. Les quantités débarquées ont été transformées en proportions puis par la racine carrée afin de diminuer l'effet de l'abondance. L'ACP a été appliquée sur une matrice de proportions de captures avec en lignes les regroupements de séquences de pêche par quinzaine et en colonnes les espèces. Nous avons conservé les axes avec une valeur propre > 1 selon la règle de Kaiser-Guttman. Chaque séquence s'est vue ainsi attribuée son score d'ACP (coordonnées des individus). Le processus des données est disponible en Annexe 5. Le ciblage a alors été représenté par un certain nombre de composantes principales retenues qui sont des variables quantitatives.

- Méthode de la classification non hiérarchique sur composantes principales (Winker et al., 2013)

Une classification non hiérarchique a aussi été réalisée sur les composantes principales de tous les axes de l'ACP présentée au-dessus. La méthode CLARA, qui est une extension du partitionnement autour des médoïdes (PAM), a été utilisée sur toutes les composantes (Winker et al., 2013). Le nombre optimal de clusters a été défini en comparant 3 méthodes de maximisation : wss, silhouette, gap_stat. Le ciblage est alors représenté par une appartenance à un cluster qui fournit une seule variable qualitative.

- Méthodes des résidus directs du mélange (Okamura et al., 2018)

Cette dernière permet de recréer l'information du comportement de pêche à partir des résidus de modèles linéaires des espèces ciblées dans la pêcherie. Cette méthode fonctionne très bien avec un petit nombre d'espèces mais montre des faiblesses lorsque ce nombre augmente. Dans notre cas il nous a été difficile de sélectionner un nombre réduit d'espèces (3-4

⁵ Séquence de décisions prises par le capitaine qui prend en compte la technique, la zone de pêche et la saison dans le but de capturer une ou plusieurs espèces.

espèces maximum) en s'assurant que ce sont bien ces espèces qui soient ciblées et avons donc mis de côté cette approche.

3 méthodes identifiées et testées, celle des composantes principales est conservée.

2.3.5 Outils de modélisation

2.3.5.1 Modèles additifs mixtes généralisés

Les GAM, modèles additifs généralisés, ont la possibilité d'offrir de la flexibilité dans la représentation du lien entre les prédicteurs et la variable réponse en permettant de spécifier des relations non linéaires. Par définition, le GAM est un modèle linéaire généralisé avec un prédicteur linéaire impliquant une somme de fonctions de splines de lissage en tant que covariables (Wood, 2017). On peut écrire le modèle de la façon suivante :

$$g(\mu_i) = A_i\theta + f_1(x_{1i}) + f_2(x_{2i}) + f_3(x_{3i}, x_{4i}) + \dots \quad (4)$$

Où g est une fonction de lien⁶, $\mu_i \equiv E(Y_i)$ et $Y_i \sim EF(\mu_i, \emptyset)$. Y étant la variable réponse et EF une distribution de la famille des lois exponentielles de moyenne μ_i et de variance \emptyset . A_i est une ligne de la matrice du modèle pour toute composante strictement paramétrique, θ est le vecteur de paramètres correspondant.

Les fonctions $f_p(x_{pi})$ sont des fonctions de splines de lissage (équation 5). Essentiellement, une spline est une fonction polynomiale définie par morceaux. Plus précisément, elles sont une combinaison linéaire de fonctions de base b (*basis functions*) avec des poids β estimés en fonction des données. Dans l'équation ci-dessous, ' k ' représente le nombre de bases utilisées pour former la spline (Pedersen et al., 2019). Lorsque k augmente, le nombre de fonction de base augmente ce qui permet de produire des splines plus complexes, plus ondulées.

$$f(x_i) = \sum_{j=1}^k \beta_j b_j(x_i) \quad (5)$$

En pratique, il n'est pas optimal de contrôler la complexité de la courbe en variant k . Il est préférable de choisir un k plus grand que nécessaire, pour assurer que le nombre de fonctions de base ne soit pas le "facteur limitant" pour l'ajustement du modèle. Les paramètres des splines sont ainsi estimés en maximisant une version modifiée de la log-vraisemblance en ajoutant une pénalité proportionnelle au niveau de courbure de la spline. L'importance de cette pénalité est proportionnelle à un paramètre de lissage (smoothing parameter) λ_i estimés avec la méthode du maximum de vraisemblance restreint (REML) (Pedersen et al., 2019).

Pour autant les GAM permettent également de modéliser des relations linéaires i.e. composantes paramétriques. Des interactions sont également implémentables. Il est possible de modéliser des interactions entre variables non linéaires et qualitative ainsi qu'entre variables non linéaires. Pour la première, une spline est estimée pour chaque niveau du facteur. L'interaction entre variables non linéaires est différente selon si les variables sont sur la même échelle ou non. L'interaction de deux variables mesurées sur la même échelle (ex : longitude et latitude) se modélise en utilisant des splines bivariés (f_3) dans l'équation 4) caractérisé par un paramètre de lissage unique. Si nous voulons modéliser une interaction spatio-temporelle, les produits tensoriels permettent de définir une spline en plusieurs dimensions. Ces derniers sont caractérisés par un paramètre de lissage différent.

Les modèles additifs mixtes généralisés (GAMM), permettent en plus de prendre en compte des effets aléatoires. Ce type d'effet modélise la variation de la variable réponse conditionnellement aux prédicteurs et par conséquent contribue à sa variance.

⁶ La fonction de lien g permet de relier linéairement la moyenne attendue avec les variables explicatives. C'est une transformation de la moyenne attendue et non pas des valeurs observées.

Les modèles GAMM ont été implémentés sous le logiciel R (R Core Team, 2022) avec le package mgcv (Wood, 2023). La fonction bam() a été utilisée car conçue pour des jeux de données de plusieurs dizaines de milliers de lignes (Wood et al., 2015). Cette construction améliore les performances lors de l'ajustement de grands ensembles de données et accessoirement diminue le temps de calcul (Pedersen et al., 2019).

2.3.5.2 Tests statistiques pour la validation des modèles

Plusieurs critères ont été utilisés pour comparer la performance des modèles : le critère d'information d'Akaike (AIC), la déviance expliquée par le modèle et par chaque variable, le R^2 ajusté. La P-value des variables a été regardée. Une fois les variables sélectionnées nous avons opéré la validation des modèles par l'analyse des résidus et leur capacité prédictive. L'analyse des résidus nous a permis de valider ou invalider les hypothèses de modélisation. Ces résidus ont été regardés par le spectre de 4 graphiques en spécifiant quel type de résidus nous voulions observer (déviance, Pearson, bruts) :

1. Quantile-quantile des résidus. Cette technique graphique permet de comparer la distribution des résidus de déviances issus des modèles avec la distribution d'un échantillon de référence. Si les hypothèses de distribution du modèle sont respectées, ces tracés devraient être proches d'une ligne droite ;
2. Histogramme des résidus de déviances pour vérifier la normalité de la distribution des résidus ;
3. Résidus bruts en fonction du prédicteur linéaire pour vérifier la relation moyenne-variance ;
4. Valeur observée en fonction des valeurs estimées pour juger la qualité d'ajustement du modèle. Idéalement l'observation d'une droite de pente = 1 signifie que le modèle estime exactement ce qui est observé.

2.3.6 Choix et justifications des étapes de la construction du modèle

2.3.6.1 Variable réponse

La méthode de standardisation des données professionnelles (ou dépendantes) utilise classiquement la CPUE comme variable réponse. En appliquant cette méthode, nous aurions dans notre cas :

$$Y_i = \frac{C_i}{E_i}, \quad (7)$$

où Y_i est la variable réponse calculée comme le ratio de la quantité débarquée d'une espèce sur l'effort de la séquence et du navire de pêche correspondant. L'indice 'i' correspond à l'année. Cependant, ici, et comme dans les travaux de Candy (2004) notamment, la distribution du taux de débarquement (CPUE) est difficile à modéliser puisqu'il s'agit du rapport de deux variables aléatoires. Une autre approche consiste à conditionner les valeurs observées de l'effort et à modéliser les débarquements en biomasse de chaque séquence comme variable réponse (Maunder et Punt, 2004 dans Coelho et al., 2021). Pour ce faire, la variable d'effort a été spécifiée en offset⁷ du modèle. De plus, utiliser les données brutes est considéré comme bonne pratique (Thorson, 2019). L'offset doit être inclus sur la même dimension que le prédicteur linéaire et donc transformé par la fonction de lien utilisé si cette dernière est différente de la fonction canonique.

⁷ Le terme offset est un prédicteur "structurel". Son coefficient n'est pas estimé par le modèle mais est supposé avoir la valeur 1. Ainsi, les valeurs de la variable en offset sont simplement ajoutées au prédicteur linéaire de la cible.

2.3.6.2 Hypothèse sur la distribution des données et fonction de lien

Notre jeu de données comprend un assez grand pourcentage de zéros. La distribution tweedie appartient à la famille des distributions exponentielles et à l'avantage de traiter les valeurs nulles de manière unifiée. Cette dernière, de plus en plus répandue pour standardiser des séries de CPUE où la proportion de valeurs nulles est non négligeable, a donc été retenu pour sa capacité à s'adapter à la variation de valeur nulle dans la matrice de données (Shono, 2008). La distribution tweedie est définie par une variance déterminée en (6) :

$$Var(Y) = \Phi \cdot \mu^p \quad (6)$$

Où Y est la variable réponse ; Φ est le paramètre de dispersion ; μ la moyenne de Y et p le paramètre puissance. Le paramètre de puissance p , qui en prenant des valeurs entre 1 et 2, permet d'obtenir une distribution continue pour les valeurs positives en ajoutant une valeur discrète en 0. Y compris les distributions normales ($p = 0$), de Poisson ($p = 1$) et Gamma ($p = 2$). Un second paramètre règle la dispersion du modèle (Φ). Cette distribution a semblé être la plus appropriée pour modéliser des séries de débarquements directement. L'utilisation de la méthode du maximum de vraisemblance restreinte nous a permis d'estimer le paramètre de dispersion.

La fonction de lien canonique de la distribution tweedie que l'on note θ est égale à :

$$\frac{\mu^{1-p}}{1-p} \text{ pour } p \neq 1 \text{ et } \mu^{1-p} \text{ pour } p = 1 \text{ (Ohlsson and Johansson, 2006)}$$

Cependant il est communément utilisé dans le cas des GAM et GAMM, une fonction de lien log. L'échelle log permet de reformuler le problème de régression en remplaçant les CPUE par le débarquement total conditionnellement à l'effort : $Y_i | \{E_i\}$, et en traitant le log de l'effort en offset dans le prédicteur linéaire (DeLong et al., 2020).

2.3.6.3 Variables explicatives

La première étape de modélisation consiste à utiliser les variables utilisées dans les modèles de merlan et lieu jaune, c'est-à-dire l'année, le mois, le rectangle statistique mais en prenant en compte le centroïde en longitude et latitude, la longueur des navires et les variables associées aux tactiques de pêches et l'identifiant du navire. On remarque que les variables utilisées sont sensiblement identiques au GLM mise en place afin de standardiser les CPUE de rouget-barbet de roche du stock Est (Camara, 2009). Ces dernières peuvent être spécifiées comme ayant un effet fixe ou aléatoire vis-à-vis des débarquements.

L'année étant la variable reflétant l'abondance au cours du temps, information que nous voulons extraire, est spécifiée en tant que facteur en ayant un effet fixe et inclus dans tous les modèles (Maunder and Punt, 2004). Les caractéristiques des navires sont spécifiées en tant que facteur (qualitatif pour la longueur, quantitatif pour la puissance et le tonnage). Toutes les autres variables sont spécifiées avec une fonction de spline. L'identification du rectangle statistique pourrait être spécifié tel quel comme facteur qualitatif mais souffrirait dans le modèle d'un grand nombre d'estimations de paramètres (i.e. autant que de rectangles). Ainsi nous avons fait le choix d'utiliser les coordonnées géographiques du centroïde du rectangle, longitude et latitude en interaction dans une spline bivariée.

On trouve dans la littérature des effets aléatoires sur les navires et en interaction avec l'année ou la saison (Augustin et al., 2013 ; Candy, 2004 ; Thorson and Ward, 2014). Ici nous avons uniquement considéré l'effet aléatoire simple sur les navires. La variabilité que modélise cet effet dépend des attributs du navire (autres que ceux spécifiés en tant qu'effet fixe), de son équipage et de l'étendue totale des zones de pêche qu'il cible.

Le modèle contenant toutes les variables énoncées ci-dessus est le suivant :

$$Y_i \sim TW(E = E(Y_i), \emptyset) \quad (8)$$

$g(E(Y_i)) = \text{Offset}(\log(\text{TP_NAVIRE_SACROIS}_i)) + \text{year}_i + s(\text{NAVS_COD}_i, \text{bs} = 're') + s(\text{month}_i, \text{bs} = 'cc^', k = 12) + s(\text{carre.long}_i, \text{carre.lat}_i, k = 40) + \text{Loa}_i + s(\text{PC1}_i) + s(\text{PC2}_i) + s(\text{PC3}_i) + s(\text{PC4}_i) + s(\text{month}_i, \text{bs} = 'cc', k = 12, \text{by} = \text{year}_i)$

Y_i correspond aux débarquements. TW correspond à la distribution tweedie. $g()$ est la fonction de lien log. L'effort de pêche est spécifié en offset et transformé par le log. Le mois est spécifié dans une fonction de spline utilisant une régression cubique cyclique de coefficient $k = 12$ (12 nœuds i.e. 1 pour chaque mois). La variable spatiale est une spline bivariée permettant de modéliser l'interaction non linéaire de deux variables (longitude et latitude) s'exprimant dans la même unité (coordonnées géographiques). Le paramètre k est fixé à 40 pour cette fonction. La tactique de pêche est modélisée à travers des splines à plaque mince. L'identifiant du navire (NAVS_COD) est renseigné comme effet aléatoire avec une pente fixe. Lorsque k n'est pas renseigné, le modèle l'estime par la méthode de maximisation de la vraisemblance restreinte (REML). L'argument 'by' permet d'estimer une spline mensuelle pour chaque année.

Des effets spatio-temporels ont aussi été considérés à la place d'un effet spatial. Pour ce faire, des variables spatiales annuelles et mensuelles ont été spécifiées, tel que ci-dessous, dans un produit tensoriel et intégrées à la place de la spline spatiale dans le modèle.

$$te(\text{carre.long}, \text{carre.lat}, \text{year}, k = c(40,12), d = c(2,1)) \\ te(\text{carre.long}, \text{carre.lat}, \text{month}, k = c(40,12), d = c(2,1), \text{bs} = c('tp', 'cc'))$$

2.3.7 Étapes de sélection des modèles

Zuur and al. (2009) ont décrit une méthodologie à titre d'orientation générale afin de sélectionner les GAMM. Les structures aléatoires et fixes doivent être sélectionnées séparément. Cependant, la structure aléatoire n'étant composé, ici, que d'une seule variable, seule la structure optimale fixe sera recherchée. Ne pouvant pas tester toutes les variables et interactions possibles nous nous limiterons aux variables et interactions présentées dans la partie ci-dessus.

1. Premièrement, l'effet de chaque variable fixe conjointement à la structure aléatoire est regardé. Pour ce faire, un modèle a été ajusté pour chaque variable fixe. La P-value, le pourcentage de déviance expliquée et le degré de liberté nous permettent de réaliser une première sélection des variables mais avant tout d'avoir une idée de l'influence de chacune d'entre elles.
2. Deuxièmement, une sélection par étapes a été réalisée en ajoutant les variables une par une (forward selection). A chaque étape, nous avons conservée ou non la variable préalablement testée. Cette sélection se fait en prenant compte l'AIC, la significativité de la variable ajoutée ainsi que la déviance expliquée par le modèle.
3. Enfin, pour mieux comprendre et visualiser les effets des variables sur le processus de normalisation des CPUE, des 'step plot' ont été construits (Bentley et al., 2012). Pour ce faire nous avons ajouté séquentiellement les variables identifiées dans le modèle candidat en générant des prédictions de CPUE selon le paragraphe **Erreur ! Source du renvoi introuvable.**

2.4 Construction des indices d'abondances

Une fois le meilleur modèle construit pour chaque zone et période, nous souhaitons extraire une tendance annuelle qui se traduit par une série annuelle de CPUE. En raison de la forme du lien logit, les effets additifs sur l'échelle de la réponse ne le sont plus sur l'échelle de la probabilité prédite p . Dans ce cas, pour visualiser les effets de p , il est préférable de créer une grille de valeurs standardisées et de prédire sur cette dernière.

La CPUE annuelle a ainsi été standardisée en fixant toutes les covariables autres que l'année à un vecteur de valeurs normalisées X_0 et en prédisant sur ce jeu de données fictif à partir du modèle sélectionné. Les choix de variables sont la médiane du temps de pêche navire et des variables de tactiques de pêches (PC1, PC2, PC3, PC4), le mois et le rectangle (latitude et longitude) qui représentent les débarquements les plus élevés, le navire qui pêche le plus et ses caractéristiques associées (Candy, 2004). Nous avons également testé les prédictions en excluant le terme 'NAVS_COD' qui est la variable aléatoire.

Il est à noter que la distribution des valeurs prédites s'approche davantage d'une distribution normale sur l'échelle du logit, plutôt que sur l'échelle de p. C'est pourquoi nous avons calculé les intervalles de confiance dans l'échelle de la fonction de lien puis transformée par l'exponentielle la probabilité prédite afin de la visualiser dans l'échelle de la variable réponse.

Les séries temporelles de CPUE ont été normalisées par leur moyenne respective afin de regrouper les deux périodes.

3 Résultats

3.1 Sélection d'une flottille de référence

3.1.1 Sélection d'un engin de référence

Au sein des deux stocks définis pour ce travail, l'engin OTB occupe une place importante dans les débarquements (Figure 10). Il représente respectivement, 40% et 81% des débarquements sur toute la période. Avant 2008, les débarquements de rouget-barbet de roche à l'OTB étaient supérieurs dans le stock Mixte puis les tendances se sont inversées sans pour autant montrer une grande différence entre les deux zones (Figure 11). La tendance des débarquements montrent que celle pour les OTB est similaire à la tendance globale (Figure 7 ; Figure 11). Les navires mettant en place l'engin OTB sont largement répartis spatialement ce qui en fait un engin intéressant pour le travail de standardisation (Annexe 6). Les navires mettant en place l'engin OTB montrent également une saisonnalité les données de débarquements (Annexe 7, Annexe 8).

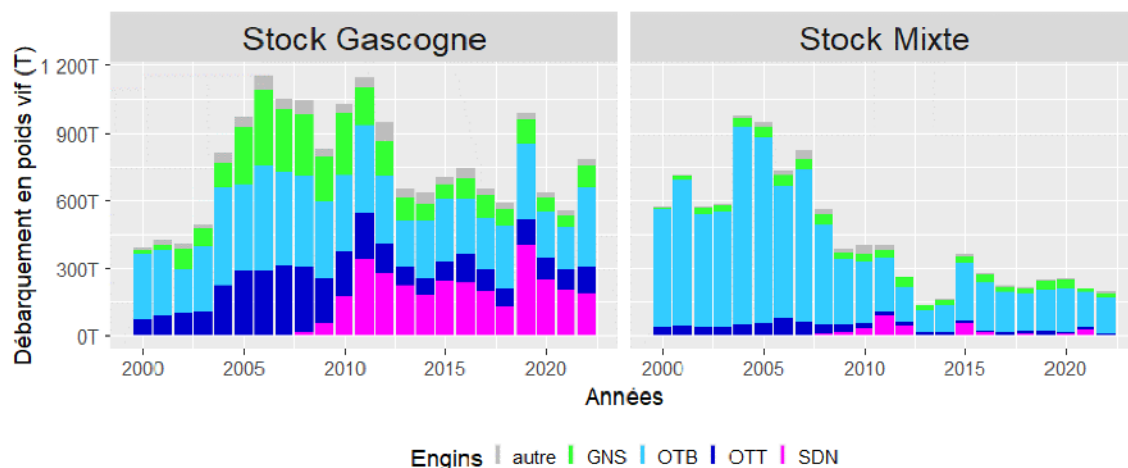


Figure 10 | Diagramme en barres des débarquements annuels de *Mullus surmuletus* en fonction des engins pour les stocks Golfe (gauche) et Mixte (droite).

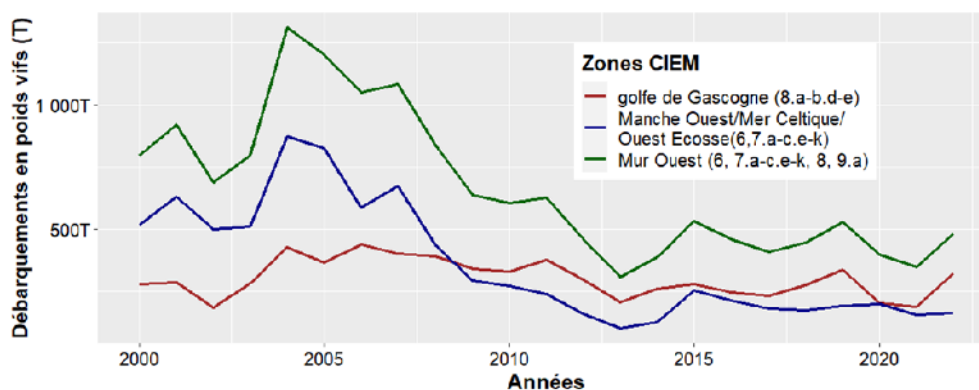


Figure 11 | Évolution des débarquements annuels de *Mullus surmuletus* des zones 6 et 8, et dans les divisions 7.a-c, 7.e-k, et 9.a avec l'engin OTB.

3.1.2 Flottille du golfe de Gascogne (8.a-b. d-e)

Les navires OTB ayant pêché du rouget dans le golfe de Gascogne au moins une fois sur la période 2000-2022 sont au nombre de 951. Une première sélection est opérée pour ne retenir que les navires ayant été actifs 5 ans ou plus sur la période. Une première sélection visuelle est opérée pour ne retenir que les navires ayant été actifs plus de 4 ans sur la période. Les tendances d'évolution des débarquements sont conservées tout en diminuant le nombre de données (Figure 12). 425 navires sont ainsi sélectionnés. Deuxièmement, une sélection des navires est appliquée en fonction de leur moyenne annuelle de débarquements : les navires les plus représentatifs en matière de tonnage débarqué sont conservés. Parmi les navires actifs 5 ans ou plus, les seuils de captures suivants ont été testés : > 250 kg, > 500 kg, > 1800 kg. Ces derniers sont définis au regard de la distribution des moyennes annuelles de débarquements. Le seuil de 500 kg est finalement conservé du fait qu'il conserve la tendance et diminue le nombre de navires présent dans la flottille (Figure 14).

3.1.3 Flottille de la Manche-Ouest, Mer Celtique et Ouest-Ecosse (6, 7a-c, e-k)

Les sélections opérées amènent à retenir les navires étant actifs plus de 6 ans (courbe verte, Figure 13) dans la pêcherie et débarquant en moyenne plus de 600 kg par an de rouget (courbe bleue, Figure 15). Cette sélection correspond à 125 navires. Le pattern d'évolution des débarquements est conservé. De plus, un nombre suffisant de navires est présent dans le jeu de données.

Tableau 2 | Récapitulatif du jeu de données issu des filtres et navires sélectionnés à la suite des filtres. Les flèches représentent le avant/après les filtres. Il est donc présenté le nombre de navires avant et après avoir appliqué les filtres, de même pour le nombre de séquences.

	Golfe de Gascogne	Manche Ouest/Mer Celtique
Filtre présence	> 4 ans	> 6 ans
Filtre tonnage	> 500 kg/an	> 600 kg/an
Nombre de navires	951 → 266	888 → 125
Nombre de séquences i.e. lignes	445 542 → 340 279	645 454 → 464 098
Proportion de zéros	33,15 %	26,76 %

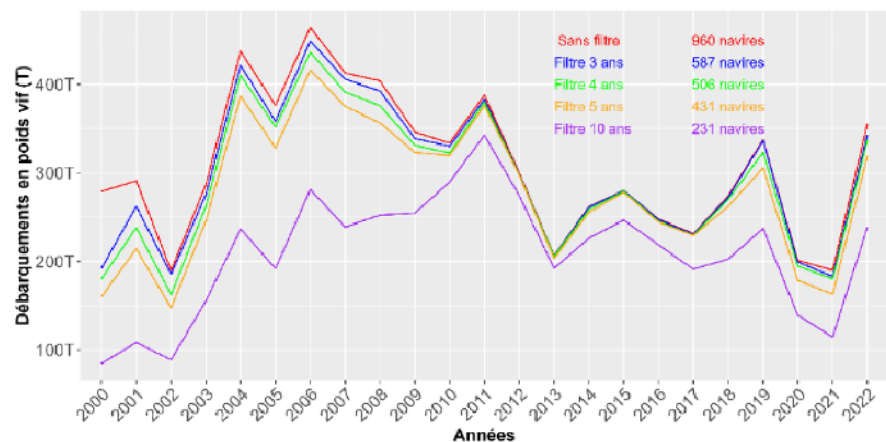


Figure 12 | Évolution des débarquements de la flottille sélectionnée pour le stock Golfe selon différents seuils de présence des navires.

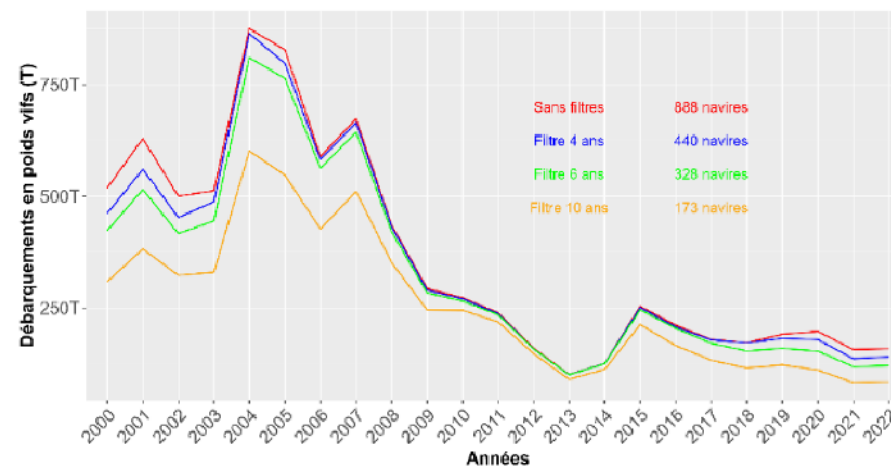


Figure 13 | Évolution des débarquements de la flottille sélectionnée pour le stock Mixte selon différents seuils de présence des navires.

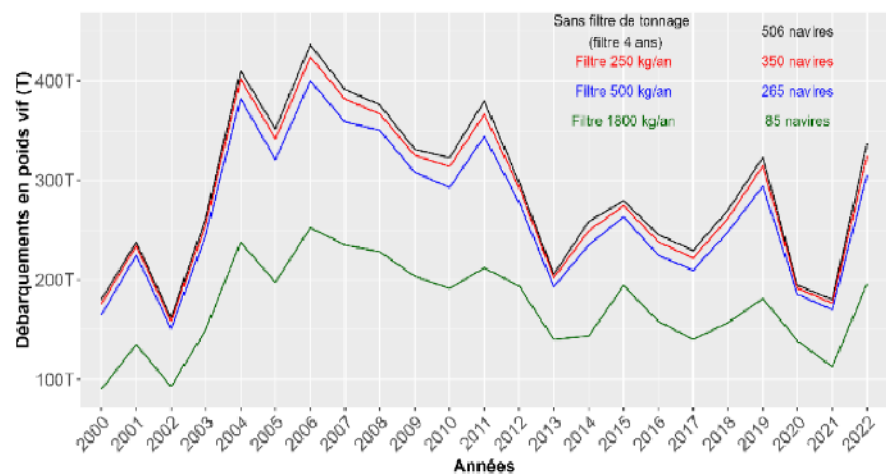


Figure 14 | Évolution des débarquements de la flottille sélectionnée pour le stock Golfe selon différents seuils de débarquements annuels des navires.

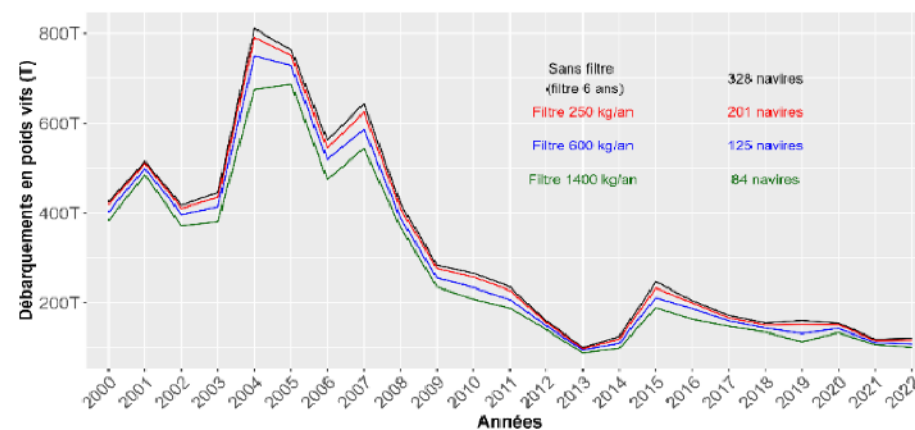


Figure 15 | Évolution des débarquements de la flottille sélectionnée pour le stock Mixte selon différents seuils de débarquements annuels des navires.

3.1.4 Influence des filtres sur la répartition spatiale des navires

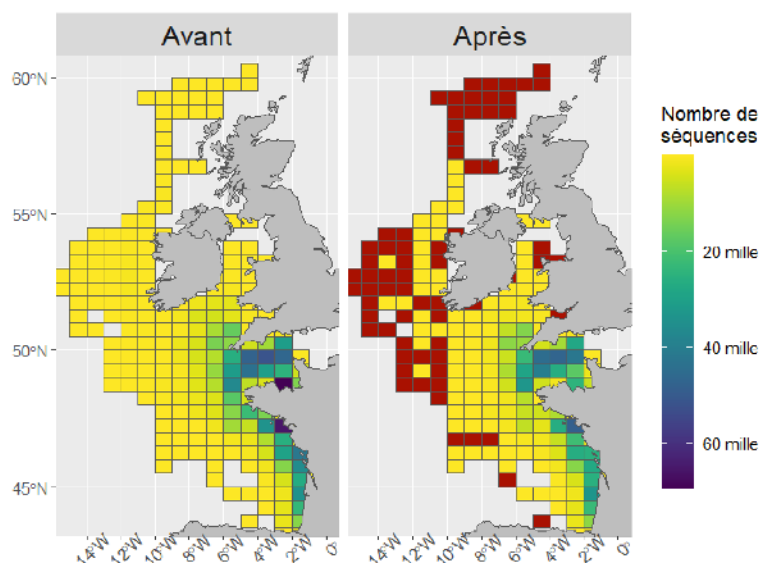


Figure 16 | Nombre totale de séquence par rectangle statistique sur toute la période avant et après les filtres sur les flottilles. Les rectangles rouges sont ceux qui ont disparu de l'analyse.

Les filtres effectués plus haut ont pour effet de retirer des navires et des séquences du jeu de données (Tableau 2). Ici nous regardons l'effet de ces filtres sur la spatialisation des captures relatives aux débarquements (Figure 16). La répartition géographique des navires débarquant du rouget-barbet de roche est regardée à travers le nombre de séquences effectuées dans chaque rectangle. Une analyse du χ^2 sur la distribution du nombre de séquences par rectangle montre que la distribution avant est significativement différente de celle d'après (P -value < 0.05). La tendance spatiale n'est donc pas statistiquement conservée à la suite de ces filtres. Il faut cependant mettre en exergue une conditions d'utilisation de cette méthodes qui est d'avoir des effectifs supérieurs à 5. Cette dernière n'étant pas remplis dans notre cas d'étude, cette différence est à moduler. Graphiquement on observe tout de même une similarité avec une forte présence sur les côtes. Certaines zones ne sont plus considérées dans l'analyse et sont représentées en rouge sur la Figure 16 ci-dessus. Ces rectangles sont d'une importance marginale dans les captures. Les filtres on ainsi pour autre effet un réajustement de la répartition spatiale des navires vis-à-vis des débarquements de rouget-barbet de roche.

3.2 Création des variables de ciblage

Tableau 3 | Résumé des résultats de la construction de la variable de ciblage. Le nombre de groupes correspond aux regroupements par quinzaines de jours (§ 2.3.4.4).

	Stock Golfe		Stock Mixte	
	1 st	2 nd	1 st	2 nd
Nombre de groupes	39 786	63 678	61 260	90 520
Nombre d'axes conservés	4	3	4	4
% de variance expliquée cumulé	63,5 %	50 %	63,5 %	63,8 %

Rappelons tout d'abord que la méthode des composantes principales a été sélectionnée pour la création de la variable de ciblage. Les 10 premières espèces en tonnage sont conservées

car elles expliquent le plus de différences dans la composition des captures. Le rouget fait partie de ces 10 espèces dans le stock Golfe à la différence de la zone correspondant au stock Mixte pour laquelle nous avons artificiellement ajouté l'espèce. Respectivement quatre et trois axes sont retenus pour les deux périodes du stock Golfe. Ces derniers expliquent respectivement 64% et 50% de la variance cumulée. Quatre axes sont conservés pour les deux périodes du stock Mixte et expliquent chacun 64% des différences dans la composition des captures. Les graphiques des variables des ACP sont présentés en Annexe 9 à Annexe 12.

3.3 Construction du modèle

Les effets de chaque variable d'intérêt ajoutée à la structure aléatoire indépendamment des autres variables sont présentées en Annexe 13. Le Tableau 4 résume la construction par 'forward selection' des quatre modèles. Nous introduisons pas à pas dans le modèle les variables mois, celles relatives au spatial, des caractéristiques navires, de variables reflétant la tactique de pêche et l'interaction année*mois. Les modèles ' G_i ' sont constitués des variables du modèle 'F' pour lequel la variable spatiale est remplacée par une variable spatio-temporelle (année * spatial ou mois * spatial).

La tactique de pêche explique entre 20 et 31% de la déviance totale de chaque modèle et est donc la variable la plus explicative. On note cependant que les degrés de libertés effectifs des variables de tactiques de pêches sont proches de k , i.e. fonction de base, fixé à 10. L'ajustement des splines n'est donc pas optimal. L'effet saison i.e. variable mois, est significatif pour les quatre modèles et explique entre 2 et 10% de la déviance totale. La variable spatiale dans sa forme de spline bivariée explique beaucoup de la variabilité dans les modèles du stock Mixte à la différence de ceux du stock Golfe pour lequel cette variable présente des déviations inférieures à 1%. La caractéristique des navires montre une faible importance dans les modèles quel que soit sa représentation (longueur, puissance ou tonnage). L'interaction entre le mois et l'année est significative pour tous les modèles et explique entre 1 et 11% de la déviance. L'interaction entre l'année et les rectangles est significative et explique entre 5 et 13% de la déviance. Il en est de même pour l'interaction entre le spatial et le mois. Ces résultats nous poussent à considérer toutes ces variables pour construire les modèles lors de la forward sélection. Les variables mensuelles, spatiale bivariée, de tactiques de pêches et l'interaction mois*année sont significatives dans tous les modèles. Ces variables améliorent plus ou moins les modèles au regard de l'AIC ou du pourcentage de déviance expliquée. Le mois semble être plus déterminant pour la première période du stock Mixte avec un pourcentage de déviance d'environ 10% alors que pour la même période mais dans l'autre stock le mois explique moins de déviance

La variable spatiale semble être plus importante pour le stock Mixte. L'ajout des caractéristiques navires n'apporte que peu d'informations dans le modèle. La taille, la puissance ou le tonnage de référence n'influent donc que très peu sur la variabilité des captures de rouget-barbet de roche. On remarque que pour les 4 modèles, la variable qui rend compte le mieux de la caractéristique des navires n'est pas la même. Cependant la puissance du navire étant significative dans tous les modèles, nous décidons de conserver celle-ci. Le modèle G1 apporte peu de déviance supplémentaire expliquée. L'ajout d'une composante spatio-temporelle mensuelle accroît la qualité statistique du modèle.

Tableau 4 | Résumé des informations relatives à la sélection des modèles. Les paramètres utilisés pour évaluer la performance des modèles comprennent le degré de liberté de chaque variable ajoutée, le delta de la déviance expliquée, le delta AIC. Les valeurs en gras représentent les modèles les plus performants pour une mesure donnée. *Modèles pour lesquels la variable spatiale a été remplacée par le produit tensoriel.

Model	Variable ajouté		Δ df-edf	% déviance	AIC	Δ df-edf	% déviance	Δ AIC
			Stock Golfe			Stock Mixte		
A	offset(log(TP_NAVIRE_SACROIS)) + year	1 ère periode	239,16	14,66%	1 163 156	119,70	27,90%	1 724 776
B	A + s(month, bs='cc', k = 12)		10,24	19,45%	1 151 108	10,13	37,76%	1 682 800
C	B + s(carre.lon, carre.lat, k = 40)		33,45	19,87%	1 150 076	37,96	45,23%	1 646 902
D1	C + Power_Main		1,09	19,90%	1 150 017	0,60	45,23%	1 646 902
D2	C + as.factor(Loa)		0,88	19,89%	1 150 039	0,99	45,25%	1 646 776
D3	C + Ton_Ref		0,87	19,88%	1 150 058	1,35	45,28%	1 646 671
E	D1 + s(RS1) + s(RS2) + s(RS3) + s(RS4)		32,50	42,96%	1 082 773	34,80	61,20%	1 557 062
F	E + s(month, bs='cc', k = 12, by = year)		77,26	45,07%	1 076 029	75,51	62,91%	1 546 966
G1	F* + te(carre.lon, carre.lat, year_bis, k = c(40, 8))	52,19	46,40%	1 071 802	15,86	64,40%	1 537 155	
G2	F* + te(carre.lon, carre.lat, month, k = c(40, 8), d = c(2, 1), bs = c("tp", "cc"))	262,38	49,50%	1 059 977	214,61	64,96%	1 532 291	
% déviance totale (modele F)				45,07%			62,91%	
			-----			-----		
A	offset(log(TP_NAVIRE_SACROIS)) + year	2nd periode	222,74	12,19%	1 931 065	123,25	17,73%	1 915 567
B	A + s(month, bs='cc', k = 12)		9,85	14,24%	1 922 503	10,01	20,61%	1 900 632
C	B + s(carre.lon, carre.lat, k = 30)		26,95	15,26%	1 918 252	35,38	28,11%	1 859 576
D1	C + Power_Main		0,54	15,26%	1 918 245	0,96	28,11%	1 859 577
D2	C + as.factor(Loa)		0,81	15,26%	1 918 254	1,12	28,17%	1 859 194
D3	C + Ton_Ref		0,19	15,26%	1 918 253	34,72	28,10%	1 859 575
E	D1 + s(RS1) + s(RS2) + s(RS3)		26,23	35,94%	1 819 381	34,72	46,85%	1 742 739
F	E + s(month, bs='cc', k = 12, by = year)		127,09	37,70%	1 810 427	122,92	48,22%	1 733 477
G1	F* + te(carre.lon, carre.lat, year_bis, k = c(30, 14)))	54,58	38,30%	1 807 545	205,41	50,45%	1 717 802	
G2	F* + te(carre.lon, carre.lat, month, k = c(30, 12), d = c(2, 1), bs = c("tp", "cc"))	173,81	40,20%	1 797 065	275,24	51,93%	1 705 710	
% déviance totale (modele F)				37,70%			48,22%	

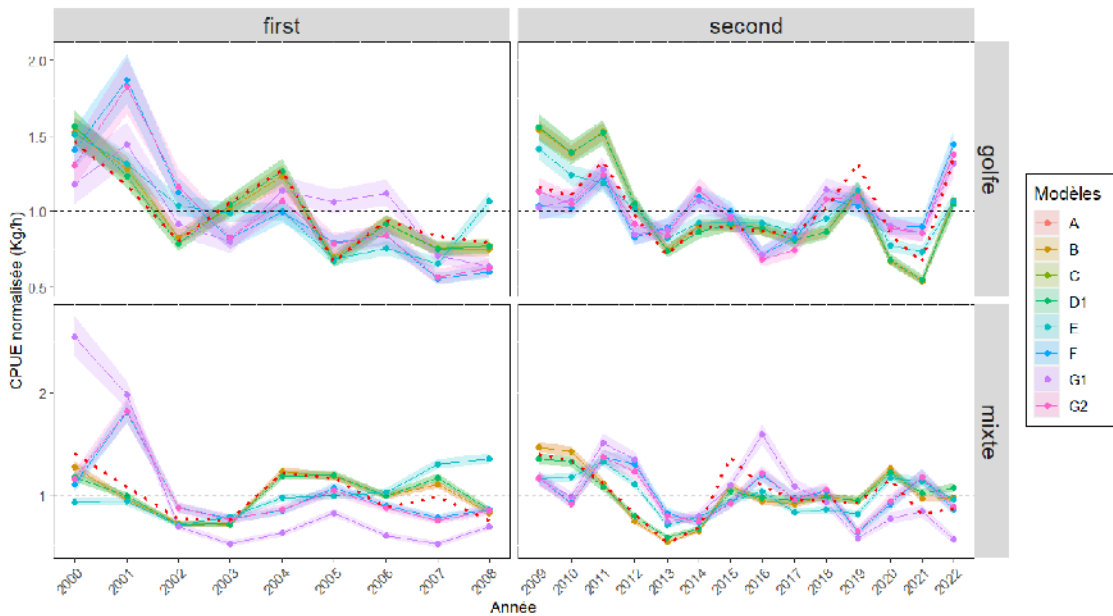


Figure 17 | Step plot de la sélection de modèle du stock Mixte (haut) et Golfe (bas) de la première période (gauche) et seconde période (droite) représentant les indices de CPUE à chaque étape de la sélection progressive des variables. Chaque courbe montre un indice de CPUE standardisé et normalisé par la moyenne au fur et à mesure que chaque variable explicative est ajoutée au modèle. Les lignes en pointillés rouges représentent la CPUE nominale.

Les « step plot » permettent d’appréhender l’effet de l’ajout de chaque variable. À la différence de la sélection des variables du modèle pour laquelle nous nous basons sur un critère statistique quantitatif, ici, on évalue les modèles sur la qualité de leur ajustement à travers une analyse visuelle. Les courbes sont les prédictions des différents modèles à partir du jeu de données fictif. Les modèles A, B, C et D1 donnent tous des prédictions similaires (Figure 17). L’ajout des tactiques de pêches ainsi que l’interaction année*mois modifie sensiblement les prédictions et donc les CPUE considérées (modèle E et F, Figure 17). Il en est de même pour le produit tensoriel spatial et annuel qui montre des prédictions différentes. Les CPUE prédites du modèle F et G2 sont sensiblement identiques. Ainsi le produit tensoriel considérant le spatial et le mois n’apporte pas d’informations dans les prédictions. Ainsi le produit tensoriel considérant le spatial et le mois n’apporte pas d’informations dans les prédictions. Le modèle G1 est écarté du fait que la variable spatio-temporelle modifie sensiblement les prédictions tout en expliquant une très faible part de la déviance expliquée. En relation avec les résultats du Tableau 4 nous sélectionnons le modèle F et excluons également le modèle G2 par simplification en raison des prédictions similaires.

3.4 Évaluation des modèles finaux et construction des indices d’abondances

Nous présentons ici l’ajustement du modèle F aux données des 2 périodes des deux stocks.

3.4.1 Qualité d’ajustement des modèles

Les valeurs extrêmes s’ajustent mal comme on peut le voir sur les quantile-quantile plots (Figure 18). La distribution des données est décalée vers la droite ; l’hypothèse de normalité n’est pas validée (Figure 19). On admet tout de même que la distribution est correcte pour le travail effectué. Le graphique des résidus bruts en fonction du prédicteur linéaire présente des valeurs négatives du fait d’une distribution continue et non bornée dans l’échelle log. On retrouve bien ici la relation moyenne variance caractéristique d’une distribution tweedie avec une hyperdispersion en forme de cône (Figure 20).

On retrouve sur cette même figure une courbe de limite inférieure qui correspond aux zéros du jeu de données. L'estimation du paramètre de dispersion de la distribution tweedie par la méthode REML rend donc correctement compte de la variance des données. Idéalement, tous les points devraient être proches d'une ligne diagonale régressée sur le graphique des valeurs réponses en fonction des valeurs prédites (Figure 21). Le R^2 est une bonne métrique de cette relation. Ici nous avons respectivement un R^2 de 0.26, 0.27, 0.47 et 0.37 pour les modèles des premières et deuxièmes périodes des stocks Golfe et Mixte. On distingue ici quelques valeurs réponses extrêmes qui impactent la qualité de l'ajustement. Cependant, globalement, l'ajustement des modèles peut être jugé comme satisfaisant pour les valeurs moyennes.

L'ajustement des modèles aux données est bon pour le stock Golfe et moins performant pour le stock Mixte (Annexe 14). Le modèle F est considéré satisfaisant pour extraire une série temporelle de CPUE. Sur ces graphiques on détecte tout de même certaines valeurs aberrantes.

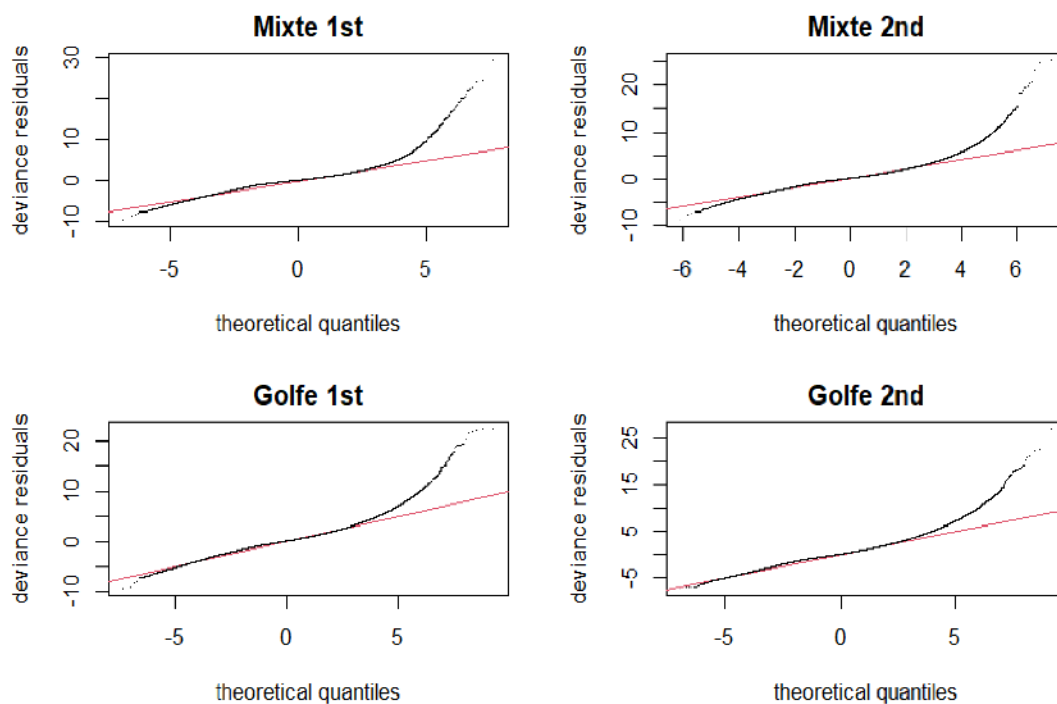


Figure 18 | Validations du diagramme Quantile-Quantile (Q-Q) pour le modèle GAM. Le diagramme QQ montre approximativement la ligne droite ($y = x$). La pente de cette ligne droite est l'écart-type ; le diagramme QQ permet de vérifier visuellement la normalité.

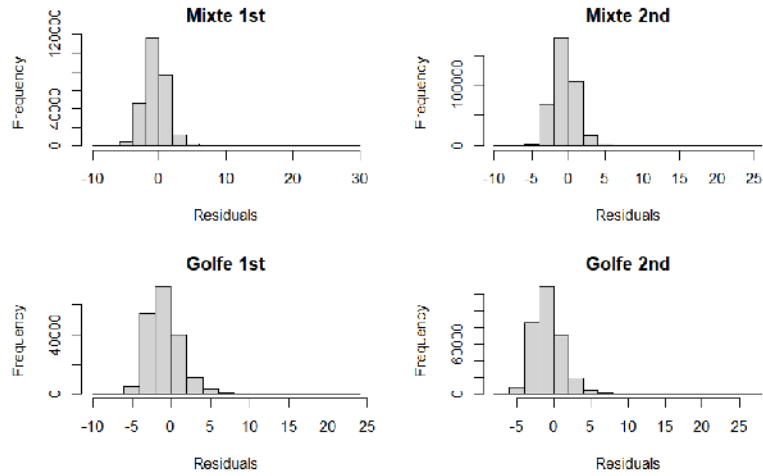


Figure 19 | Histogramme des résidus de déviations des quatre modèles

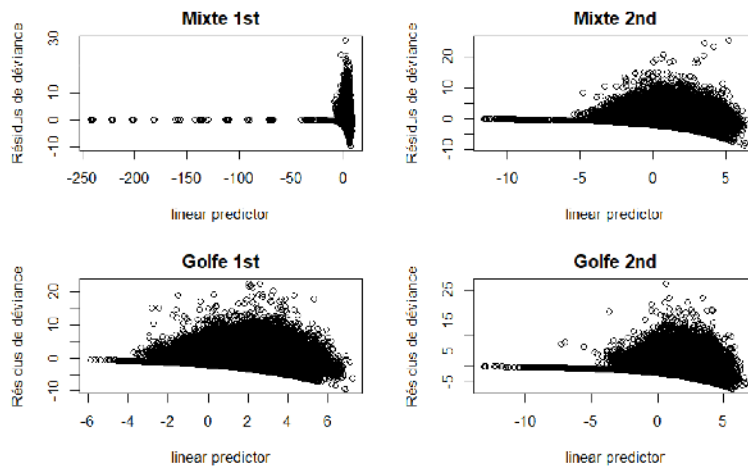


Figure 20 | Résidus de déviations en fonction du prédicteur linéaire. Prédications dans l'échelle de la fonction de spline qui explique les valeurs négatives.

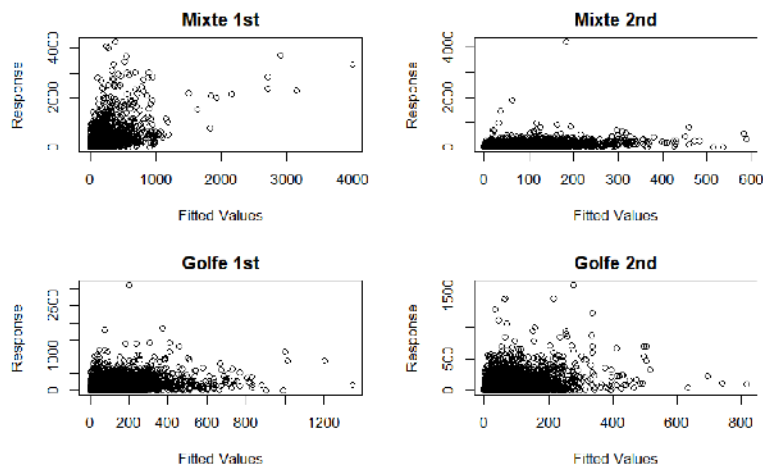


Figure 21 | Valeur réponse en fonction de la valeur estimée

L'ajustement des grandes valeurs n'est pas bon et pour cause, les modèles ont tendance à estimer de faibles valeurs de débarquements alors que ces derniers sont très conséquents (Figure 21). Il a été testé de supprimer certaines valeurs aberrantes, visibles sur la Figure 21, du jeu de données et réajuster le modèle. Ces derniers ne montrent pas ou peu de différences. Les points extrêmes contribuent donc peu à l'ajustement du modèle aux données et montre une certaine robustesse des modèles.

3.4.1.1 Tendances des CPUE

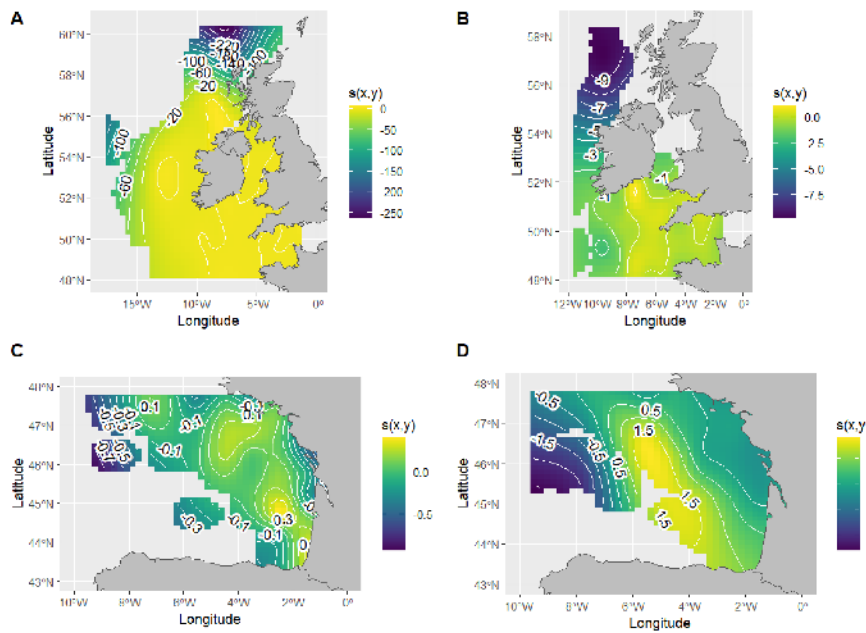


Figure 22 | Coefficient de la spline bivariable spatiale des quatre modèles. $S(x)$ représente l'écart à la moyenne. Une couleur jaune sera interprétée comme des valeurs fortes et bleues comme des valeurs faibles. A : Stock Golfe première période. B : Stock Golfe seconde période. C : Stock Mixte première période. D Stock Mixte deuxième période.

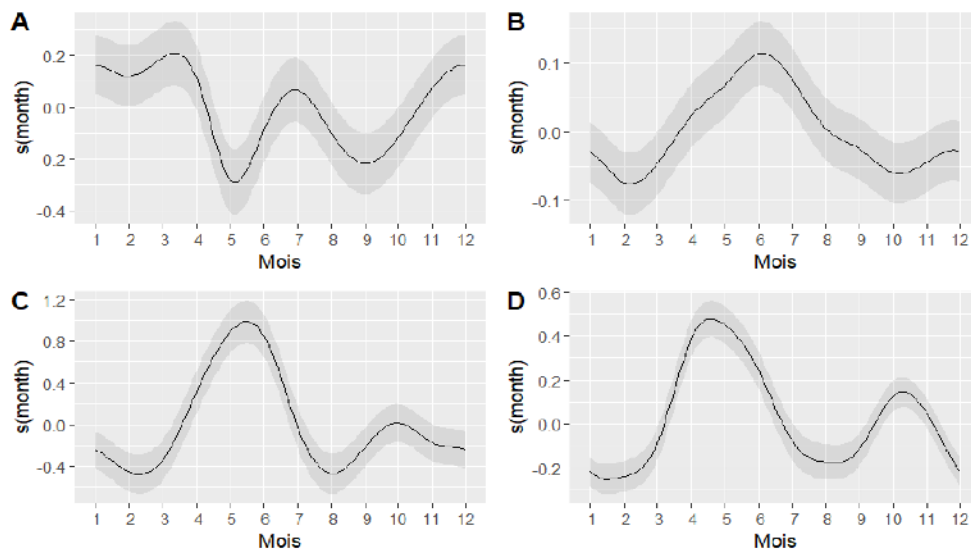


Figure 23 | Coefficient de la spline mensuelle des quatre modèles. $S(x)$ représente l'écart à la moyenne. L'erreur standard est représentée par la bande grise. A : Stock Golfe première période. B : Stock Golfe seconde période. C : Stock Mixte première période. D Stock Mixte deuxième période.

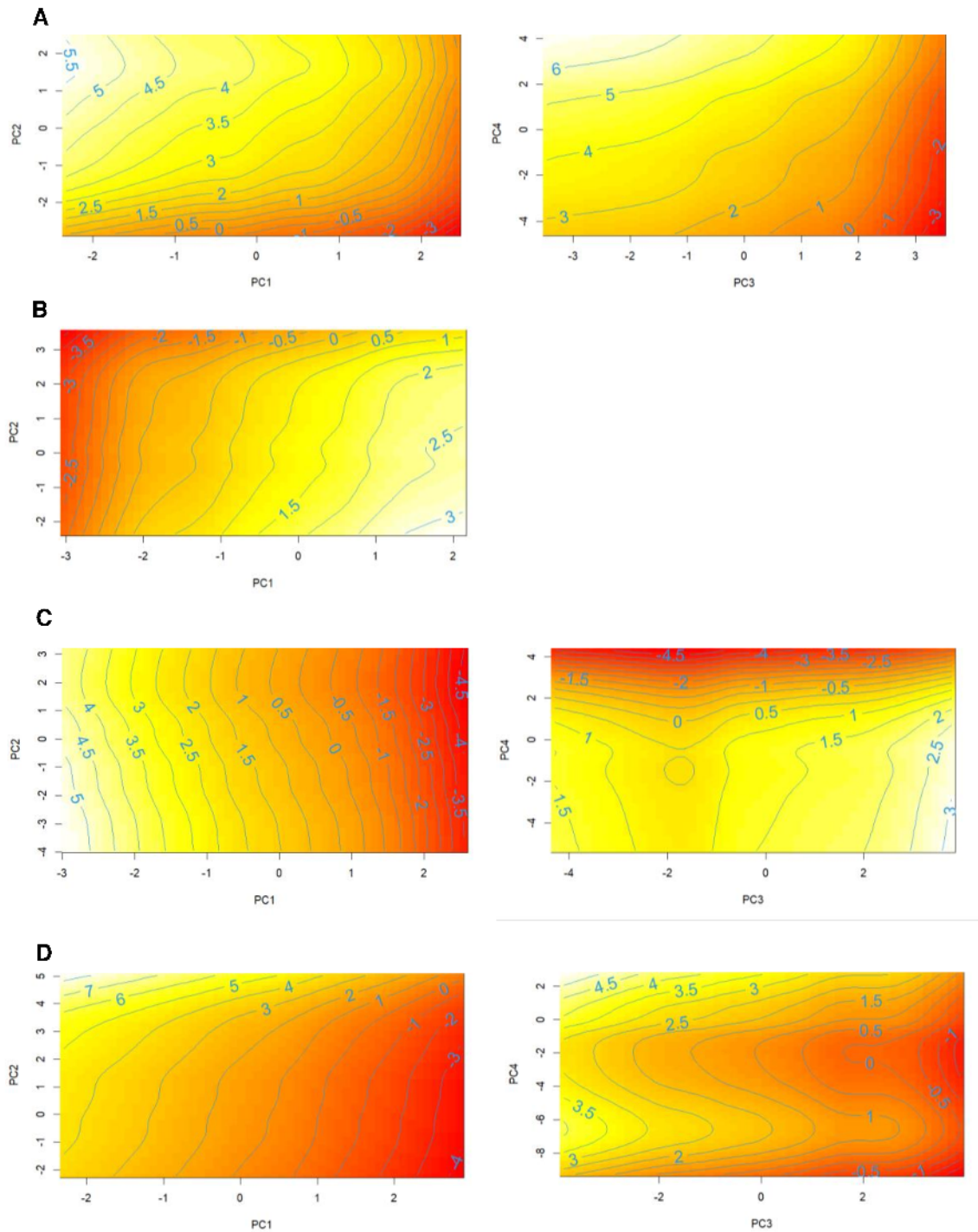


Figure 24 | Effet des axes principaux représentés par des contours plots. A : Stock Golfe première période. B : Stock Golfe seconde période. C : Stock Mixte première période. D Stock Mixte deuxième période. Les isoplèthes représentent l'effet combiné des deux axes représentés.

Ce sont les variables de ciblage, puis les variables saisonnières et spatiales qui expliquent le plus de variabilité dans les débarquements (conditionnellement à l'effort). Comme vu pour dans les données, on observe bien sur la cartographie des coefficients de la spline bivariable (Figure 22) certaines zones où les débarquements sont prédits à des niveaux plus élevés par le modèle. Ces valeurs sont le résultat de l'ajustement d'une spline modélisant la relation non linéaire entre les débarquements et les centroïdes des rectangles. On observe un effet bordure

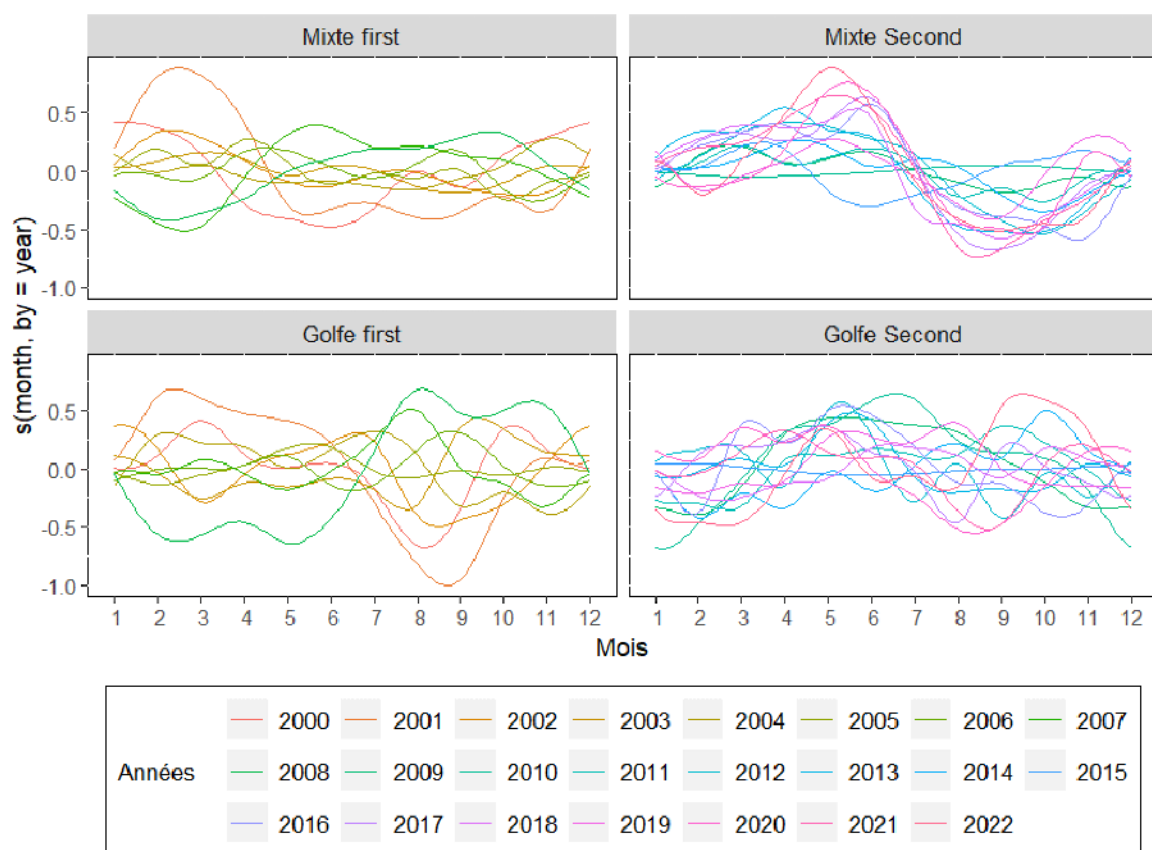


Figure 25 | Effet de l'interaction mois*année spécifié par l'argument 'by' qui ajuste donc une spline mensuelle par année. Chacune de ces splines sont alors pénalisée différemment. Les 4 panneaux présentent les coefficients de splines de l'effet mensuel pour chaque année et ce pour les 4 modèles. Le stock Mixte en haut, Golfe en bas pour la première période (first) et seconde période (second). Une ligne horizontale à 0 correspond à une absence d'effet mensuel pour l'année représentée.

au plus proche des côtes du golfe de Gascogne. Certains rectangles sont confondus avec le continent, dans l'estuaire de la Gironde par exemple. La raison est que la spline s'ajuste en fonction de la latitude et de la longitude et non pas de la limite de la côte. De ce fait, les isoplèthes de faibles valeurs sont 'tirées' vers la côte. Les surfaces aux coefficients positifs ne correspondent pas aux zones de forts débarquements que l'on observe dans les données (Figure 22, Annexe 6). Dans la première période du stock Golfe, on distingue 2 à 3 points chauds et dans la seconde, une seule zone mais plus étendue. En ce qui concerne le stock mixte, les échelles sont trop différentes entre les deux périodes et la comparaison n'est pas objective (Figure 22). On peut cependant dire que les zones septentrionales sont très peu représentées dans la pêcherie de rouget-barbet de roche par les flottilles françaises et que les débarquements sont essentiellement issus des captures déclarées en Manche et Mer Celtique.

La variable mensuelle ne montre pas le même pattern entre les deux zones. Ils sont très similaires pour les deux périodes du stock Golfe et sensiblement différents pour celles du stock Mixte (Figure 23). Dans le golfe de Gascogne, on retrouve bien une saisonnalité bisannuelle avec de plus forts débarquements au printemps et plus faibles en hiver. Dans la zone mixte, on observe un changement de saisonnalité entre les deux périodes. Les débarquements de rougets sont passés d'une saison hivernale de plusieurs mois et d'un pic en été durant la première période à une saison estivale plus étendue.

Les variables de ciblage montrent une non-linéarité dans leurs relations avec les débarquements (Figure 24). Ces résultats traduisent correctement les graphiques des variables d'ACP. On retrouve essentiellement le rouget-barbet de roche lorsque les seiches et les calamars sont

capturés. On voit par exemple sur le graphique A de droite que le rouget-barbet de roche est prédit comme plus capturé pour de fortes valeurs des axes 1 et 2 de l'ACP. En comparant au graphique des variables (Annexe 9, gauche), on remarque bien que la variable rouget-barbet de roche se trouve à l'endroit où les deux axes ont de fortes valeurs, en haut à gauche. On retrouve à ce même endroit les seiches.

La représentation de l'interaction entre le mois et l'année nous montre des changements de saisonnalité selon les zones et les périodes (Figure 25). Au cours de la première période, seules quelques années présentent un effet mensuel distinct. Au cours de la seconde période du stock Mixte, on observe une saisonnalité plus marquée ces dernières années

3.4.2 Séries d'indices d'abondance

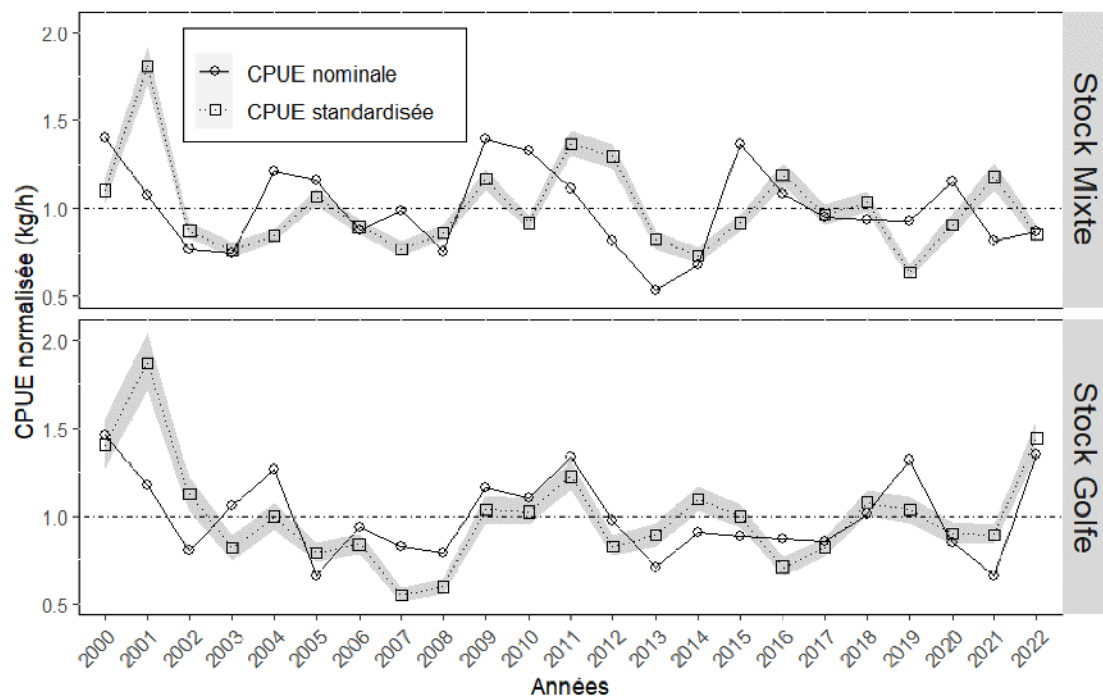


Figure 26 | Indices d'abondances nominale et standardisées des deux stocks. CPUE nominale calculée comme la moyenne du rapport des débarquements sur le temps de pêche navire. CPUE standardisée issues des prédictions des modèles GAMM. Les indices ont été normalisés par la moyenne. La ligne horizontale correspond à la valeur moyenne égale à un.

Les séries d'indices d'abondance sont présentées par la métrique de CPUE normalisées par la moyenne. En complément des figures ci-dessus, le tableau de valeurs est disponible en Annexe 15. Malgré la méthode de standardisation mise en place, les CPUE nominales sont proches des données prédites (Figure 26). Une décroissance se dessine dans le stock Golfe jusqu'en 2008 puis le niveau des CPUE apparaît stable. Dans le stock mixte il n'y a pas de tendance remarquable observable. En 2001 et pour les deux zones, les CPUE standardisées sont marquées par un pic et s'écartent fortement des valeurs nominales. La dynamique de l'abondance du stock semble être globalement assez stable au cours de temps et suivre la CPUE nominale.

La CPUE prédite lors de la première période est nettement supérieure à celle prédite lors de la seconde, avec une moyenne de 23 kg/h pour le stock Mixte et de 78 kg/h pour le stock Golfe (Figure 27). La seconde période est caractérisée par des CPUE plus homogènes entre les

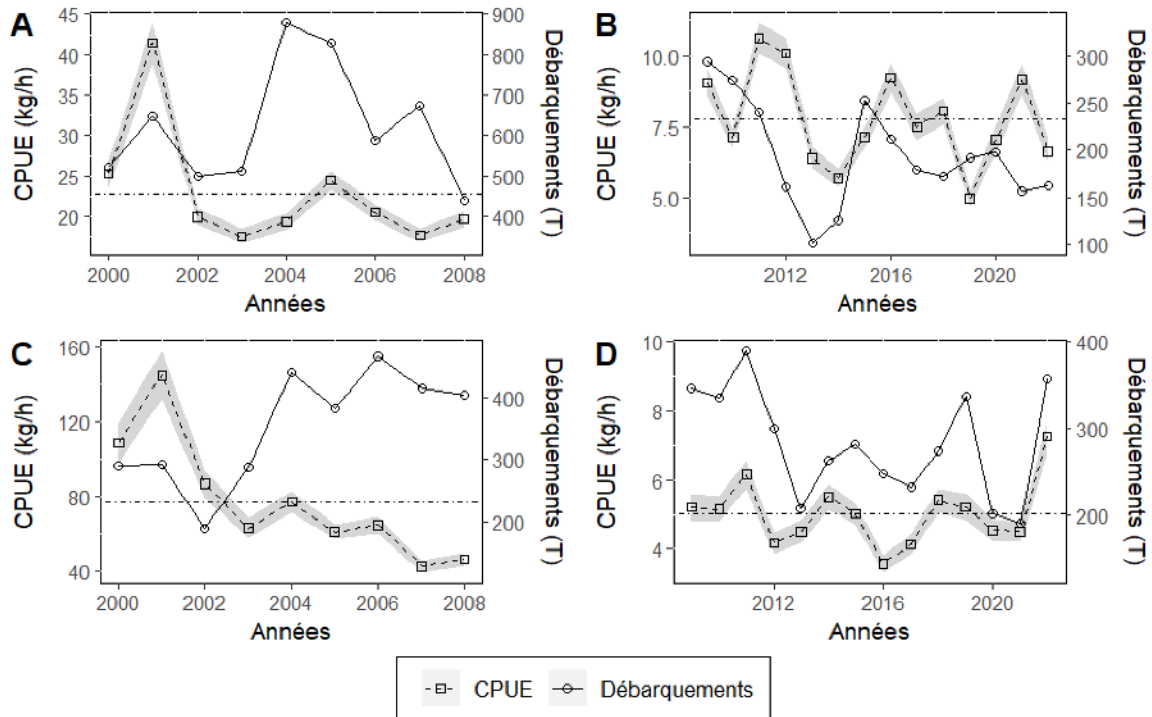


Figure 27 | Indices de CPUE standardisés (lignes en pointillés) avec intervalles de confiance à 95% (bande grise) basés sur le modèle F et débarquements totaux issus de Sacrois (lignes continues). La ligne horizontale en pointillés correspond à la moyenne des CPUE. A : Stock Golfe première période. B : Stock Golfe seconde période. C : Stock Mixte première période. D Stock Mixte deuxième période.

deux zones avec une moyenne respective de 8 kg/h et 5 kg/h. Lors de de la deuxième période et ce pour les deux zones, l'évolution des débarquements suit la tendance des CPUE. Ce schéma se retrouve pour la première période du stock Mixte mais est moins flagrant (Figure 27, A). De 2000 à 2008 dans la zone du golfe de Gascogne, l'évolution positive des débarquements est inversement corrélée à la CPUE.

4 Discussion

4.1 Division du stock

Le choix d'avoir divisé en deux le stock Ouest de rouget-barbet de roche pose question à la vue des résultats qui dessinent une tendance similaire. Rappelons que ce choix était également motivé par la volonté de pouvoir comparer les tendances entre les indices de ce présent travail et ceux issus de la méthode du « *data-filtering* » (Caill-Milly et al., 2019). Mis à part cette comparaison, la similarité des tendances dans les CPUE nous fait penser qu'il faudrait construire un indice global prenant en compte le stock considéré actuellement par le CIEM.

Rappelons que le projet ACOST comprend également un volet de génétique des populations qui vise à mieux identifier les populations et stocks de *Mullus surmuletus* dans les eaux européennes en utilisant la méthode RAD Seq (Restriction site Associated DNA). Ces travaux permettront de conforter ou non notre choix et viendront compléter les travaux existants sur l'identification des populations de rouget-barbet de roche (Benzinou et al., 2013).

4.2 Sélection de l'engin de référence

La CPUE ne mesure que la composante de la population qui est vulnérable à l'engin sélectionné (Maunder et al., 2006). Ainsi, les indices d'abondance construits à partir de la CPUE (Figure 26) ne sont le reflet que de la partie du stock exploitée par les OTB. Bien que notre sélection ait eu pour but d'identifier les composantes les plus représentatives de la dynamique du stock Ouest de l'Atlantique Nord-est de rouget-barbet de roche, certains engins et navires présentent des tendances et comportements différents de ceux des OTB (Léauté et al., 2018) et devraient ainsi être pris en compte.

D'un point de vue spatial, l'emprise géographique est différente suivant les engins. On peut notamment observer que la séparation entre le Golfe de Gascogne et la Manche Est beaucoup moins marquée, si ce n'est pas du tout, pour les GNS. C'est-à-dire que l'on observe une continuité dans les débarquements à la différence des OTB. Les GNS ont cependant une étendue plus faible sur la zone géographique du stock (Léauté, 2018).

De plus, les GNS présentent une proportion importante de rouget-barbet de roche dans leurs débarquements. En revanche, ils ne sont arrivés dans la pêcherie de *Mullus surmuletus* qu'à partir de 2009. De plus, la métrique d'effort de pêche est plus difficilement accessible pour des arts dormants. Le temps de pêche navire ne correspond pas au temps qu'a passé le filet en mer. On remarque également des tailles d'individus moyennes différentes selon les engins ainsi qu'une saisonnalité des débarquements différente.

Les GTR (Filets trémail), qui mettent en place du tramail⁸, souffrent d'une part importante de rejets d'individus de tailles commerciales à des niveaux élevés et fluctuant dans le temps (Léauté, 2018). Ceci les rend difficilement utilisables sans prendre en compte les rejets.

L'OTT, 4^{ème} engin le plus important dans la pêcherie de rouget barbet de roche du stock étudié ne représente qu'une trop faible part des débarquements pour le considérer.

L'OTB constitue alors un engin de référence cohérent et fiable, cependant, des méthodes de standardisation intégrant une dimension pluri-métiers seraient intéressantes à regarder (Gibson-Reinemer et al., 2017).

⁸ Filet mouillé verticalement, composé de trois nappes, deux extérieures à larges mailles et une intérieure à mailles beaucoup plus petites

4.3 Variables de ciblage

Il est difficile de séparer la puissance de pêche de l'abondance en analysant uniquement les données des journaux de bord (Bishop, 2008). Nous avons donc incorporé une information externe dans le modèle de standardisation qui est une variable de ciblage. Cette dernière correspond aux composantes principales d'une ACP réalisée sur une matrice de proportions de captures transformées par la racine carrée. Les relations non linéaires entre les débarquements et la valeur de ces axes est significative et explique la majorité du pourcentage de déviance dans le modèle. On retrouve la 'seiche commune' et/ou le groupe 'seiche-sépiole-nca' dans les 10 principales espèces sélectionnées. Le regroupement de ces deux groupes taxonomiques pourrait consolider les résultats donnés par les axes d'ACP en donnant plus de poids aux seiches et aux sépioles.

Les zéros ont été conservés dans le jeu de données afin de construire les indices d'abondances. Vis-à-vis du ciblage nous avons donc deux types de zéros : (1) l'espèce est ciblée mais aucune capture n'est effectuée ("vrais zéros") et (2) les captures nulles sont dues à un effort non ciblé ("faux zéros"), par exemple en pêchant au large des côtes ou l'espèce à de très faibles chances de se trouver. L'exclusion de captures nulles entraîne des CPUE positivement biaisées, notamment pour les espèces accessoires telle que le rouget-barbet de roche. Dans notre cas, les faux zéros peuvent amener à des CPUE négativement biaisées. Il serait intéressant de savoir en quelle proportion ces indices peuvent être biaisés.

4.4 Prédiction des séries de CPUE

4.4.1 Significativité des variables

L'utilisation du critère AIC a tendance à sélectionner les modèles les plus complexes malgré la condition de parcimonie. Nous nous sommes donc également intéressés à la qualité prédictive à travers des 'step-plot' afin d'identifier le modèle le plus adéquat. La variable spatio-temporelle mensuelle accroît significativement la qualité statistique du modèle mais ne modifie que peu ses prédictions. En revanche la variable spatio-temporelle annuelle présente des tendances qui s'écartent du modèle précédent en améliorant moindrement la qualité statistique. De ce fait nous privilégions la variable spatiale simple. Les prochains travaux devraient intégrer des variables spatio-temporelles additives à la variable spatiale afin de rendre compte d'un déplacement des captures ou du stock.

La variable spatiale est conservée du fait de sa significativité dans le modèle. Les coefficients les plus forts de la spline bivariée se situent assez loin des côtes (Figure 22). Or on sait que le rouget-barbet de roche est capturé en grande majorité au plus proche de ces dernières comme on peut le voir en Annexe 6 et Annexe 3. On se demande ainsi raisonnablement si la variable spatiale rend correctement compte de la spatialisation des débarquements. L'effet bordure que l'on peut voir sur la figure précédente influe également sur ce décalage. Les CPUE prédites en normalisant par le rectangle statistique le plus pêché sont affectées par cet effet. Ce rectangle ne bénéficie pas du coefficient de spline le plus élevé. Il est possible de corriger à l'avenir cet effet en tenant compte des masses continentales par l'utilisation de champs markoviens aléatoires ou d'un '*soap film smoother*' (Baudron et al., 2020 ; Rue and Held, 2005; Wood, 2017)

4.4.2 Tendances annuelles

Les prédictions sur un jeu de données fictif puis la normalisation permet d'avoir des indices d'abondances comparables. Les CPUE se montrent assez stables et très proches de la valeur nominale (Figure 26). L'évolution des débarquements est très similaire à celle de l'effort considéré (temps de pêche navire). A la vue de ces résultats, utiliser l'effort comme offset et donc comme niveau explique fortement cette constance des CPUE. On remarque donc que c'est le temps de pêche qui explique majoritairement l'évolution des débarquements. L'évolution des

débarquements suit globalement la tendance des CPUE (Figure 27, A, B, D). En revanche, lors de la première période (2000-2009) du stock Golfe on distingue des courbes antagonistes. On a alors une CPUE qui décroît pour des débarquements croissants.

Le jeu de données fictif permet de prédire annuellement mais il peut être modifié pour prédire à chaque mois et avoir une prédiction plus fine. La prédiction du modèle E sur le jeu de données qui ne contient que les années avec un seul mois sélectionné est très ressemblante à la prédiction faite sur le jeu de données mensuelles à partir du modèle F. Ainsi, ajouter l'interaction année*mois et prédire sur chaque mois revient à prédire chaque année à partir du modèle sans interactions. Il n'y a pas d'explication pour l'instant.

4.5 Comparaison avec les résultats obtenus dans le projet ROMELIGO



Figure 28 | CPUE normalisées obtenues à partir de la méthode de standardisation (ligne pointillée) et de data filtering (ligne pleine).

La méthode mise en place dans ce travail est une approche dite « *model-based* » à la différence du data-filtering que l'on dit « *design-based* ». Cette dernière méthode se base également sur les chaluts de fonds à panneaux mais uniquement pour la période 2005-2022 et pour la zone Golfe de Gascogne. Il est ainsi observé une décroissance de l'abondance de rouget-barbet de roche dans le Golfe de Gascogne avec la méthode du data filtering alors qu'une augmentation des CPUE est visible avec la seconde méthode (Figure 28). Les deux indices sont cependant très proches pour la période 2009-2022 et montrent une même dynamique. Seules les années 2005-2008 montrent une tendance antagoniste. L'augmentation que l'on peut voir avec la méthode de modélisation viens probablement du fait de la séparation de la série temporelle entre les années 2008 et 2009.

4.6 Utilisation des indices d'abondances

La mise au point d'un indice d'abondance permet de retracer l'évolution de l'abondance relative du stock ainsi que le niveau d'exploitation de la ressource. Exploitation relative à un engin donnée, ici les OTB. Dans un objectif de gestion des pêches, la CPUE permet également de calibrer d'autres modèles afin de prédire l'évolution des captures suivant différents scénarios. Différents modèles peuvent être utilisés suivant les informations dont nous disposons de la pêcherie et de la biologie de l'espèce. Disposant de peu de données biologiques, nous ne pouvons pas effectuer de structure en âge des captures. Par conséquent, l'utilisation d'un modèle d'analyse de cohorte (exemple : VPA) est exclue. Un modèle de production, lui, permet de modéliser les valeurs absolues de biomasses des ressources concernées. Au vu des données disponibles, on peut évaluer l'impact de la pression de pêche, mesurée par l'effort de

pêche (le temps de pêche), sur l'abondance du stock, mesurées ici par la CPUE. Sans chercher à prendre en compte les phénomènes biologiques qui interviennent pour cause de manque de données, il serait intéressant de tester un modèle global afin de modéliser les indicateurs qui pourraient nous permettre une gestion au RMD. Dans le cas du rouget-barbet de roche pour lequel peu d'informations sur sa biologie sont disponibles et pour lequel nous disposons désormais d'une flottille de référence, un modèle du type Spict (Stochastic Surplus Production model in Continuous Time) pourrait être utilisé (Pedersen et Berg, 2017). Il serait également intéressant de regarder si certaines limites souvent identifiées (nombre de navires limités, série de données avec peu de variation dans les niveaux, longueur de la série courte) seront des freins ou non à l'utilisation de ce modèle.

5 Conclusion

Le présent travail avait pour objectif de construire un indice d'abondance en utilisant une approche de modélisation pour le stock Ouest de l'Atlantique Nord-Est de *Mullus surmuletus*. Ce stock n'étant pas évalué, sa répartition géographique n'est pas bien définie et nous avons par précaution décidé de diviser la zone du stock en deux sur la base d'analyses otolithométriques. Cela étant fait, la CPUE nominale pour la période 2000-2022 a été standardisée en utilisant un modèle additif mixte généralisé. La série temporelle a été découpée en deux parties du fait d'un changement dans le système d'acquisition des données de débarquements courant 2008. Un indice par stock nouvellement défini a donc été construit. Les deux périodes ont été réunies en normalisant les indices d'abondances par leur moyennes respectives. La CPUE standardisée et normalisée suit globalement la tendance de la CPUE nominale. Seule la première période du golfe de Gascogne montre une évolution différente entre CPUE et débarquements. Une CPUE décroissante pour des débarquements croissants peut indiquer un accroissement dans l'efficacité de la capture du rouget barbet de roche. Ces indices maintenant construits ont vocation à être injectés dans des modèles de production (tel que SPiCT) afin de réaliser une première évaluation du ou des stocks.

Les indices d'abondances normalisés construits à partir des deux méthodes donnent des tendances similaires pour la période 2009-2022. Le data-filtering fait ainsi concurrence aux méthodes de modélisation utilisées classiquement dans la construction d'indices d'abondances.

6 Bibliographie

- Augustin, N.H., Trenkel, V.M., Wood, S.N., Lorange, P., 2013. Space-time modelling of blue ling for fisheries stock management. *Environmetrics* 24, 109–119. <https://doi.org/10.1002/env.2196>
- Baudron, A.R., Brunel, T., Blanchet, M.-A., Hidalgo, M., Chust, G., Brown, E.J., Kleisner, K.M., Millar, C., MacKenzie, B.R., Nikolioudakis, N., Fernandes, J.A., Fernandes, P.G., 2020. Changing fish distributions challenge the effective management of European fisheries. *Ecography* 43, 494–505. <https://doi.org/10.1111/ecog.04864>
- Bentley, N., Kendrick, T.H., Starr, P.J., Breen, P.A., 2012. Influence plots and metrics: tools for better understanding fisheries catch-per-unit-effort standardizations. *ICES J. Mar. Sci.* 69, 84–88. <https://doi.org/10.1093/icesjms/fsr174>
- Benzinou, A., Carhini, S., Nasreddine, K., Elleboode, R., Mahé, K., 2013. Discriminating stocks of striped red mullet (*Mullus surmuletus*) in the Northwest European seas using three automatic shape classification methods. *Fish. Res.* 143, 153–160. <https://doi.org/10.1016/j.fishres.2013.01.015>
- Bishop, J., 2008. Standardizing catch rates: is logbook information by itself enough? *ICES J. Mar. Sci.* 65: 255-266.
- Caill-Milly, N., Lissardy, M., Bru, N., Dutertre, M.-A., Saguet, C., 2019. A methodology based on data filtering to identify reference fleets to account for the abundance of fish species: Application to the Striped red mullet (*Mullus surmuletus*) in the Bay of Biscay. *Cont. Shelf Res.* 183, 51–72. <https://doi.org/10.1016/j.csr.2019.06.004>
- Camara, D., 2009. Evaluation préliminaire de stock de rouget barbet de roche (*Mullus surmuletus*) en Manche-sud mer du Nord (Mémoire de fin d'étude).
- Candy, S.G., 2004. MODELLING CATCH AND EFFORT DATA USING GENERALISED LINEAR MODELS, THE TWEEDIE DISTRIBUTION, RANDOM VESSEL EFFECTS AND RANDOM STRATUM-BY-YEAR EFFECTS.
- Dardignac, J., 1988. Les pêcheries du golfe de Gascogne. Bilan des connaissances.
- Delong, L., Lindholm, M., Wuthrich, M.V., 2020. Making Tweedie's compound Poisson model more accessible.
- Demanèche, S., Bégot, E., Gouëlle, A., Merrien, C., Leblond, E., Berthou, P., 2015. Département Sciences et Techniques Halieutiques.
- Ducharme-Barth, N.D., 2021. What is CPUE standardisation and why is it important for stock assessments? *SPC Fish. Newsl.*
- ICES, 2019. Striped red mullet (*Mullus surmuletus*) in Subarea 4 and divisions 7.d and 3.a (North Sea, eastern English Channel, Skagerrak and Kattegat) (report). *ICES Advice: Recurrent Advice.* <https://doi.org/10.17895/ices.advice.4863>
- ICES, 2012. ICES Implementation of Advice for Data-limited Stocks in 2012 in its 2012 Advice. <https://doi.org/10.17895/ICES.PUB.5322>
- ICES Advice, 2020. Striped red mullet (*Mullus surmuletus*) in subareas 6 and 8, and in divisions 7.a-c, 7.e-k, and 9.a (North Sea, Bay of Biscay, southern Celtic Seas, and Atlantic Iberian waters). <https://doi.org/10.17895/ICES.ADVISE.5772>
- ICES WG WIDE, 2020. Working Group on Widely Distributed Stocks.
- Julien, T., 2018. Identification des espèces cibles et accessoires dans la pêche mixte démersale du golfe de Gascogne. (Mémoire de fin d'étude). Ifremer, Nantes.
- Léauté, I.J., 2018. Amélioration des connaissances halieutiques du ROuget-barbet, du MERlan et du LLeu jaune du GOLfe de Gascogne.
- Maunder, M.N., Punt, A.E., 2004. Standardizing catch and effort data: a review of recent approaches. *Fish. Res.* 70, 141–159. <https://doi.org/10.1016/j.fishres.2004.08.002>
- Maunder, M.N., Sibert, J.R., Fonteneau, A., Hampton, J., Kleiber, P., Harley, S.J., 2006. Interpreting catch per unit effort data to assess the status of individual stocks and communities. *ICES J. Mar. Sci.* 63, 1373–1385. <https://doi.org/10.1016/j.icesjms.2006.05.008>
- Ohlsson, E., Johansson, B., 2006. Exact Credibility and Tweedie Models. *ASTIN Bull. J. IAA* 36, 121–133. <https://doi.org/10.2143/AST.36.1.2014146>

- Okamura, H., Morita, S.H., Funamoto, T., Ichinokawa, M., Eguchi, S., 2018. Target-based catch-per-unit-effort standardization in multispecies fisheries. *Can. J. Fish. Aquat. Sci.* 75, 452–463. <https://doi.org/10.1139/cjfas-2016-0460>
- Pedersen, E.J., Miller, D.L., Simpson, G.L., Ross, N., 2019. Hierarchical generalized additive models in ecology: an introduction with mgcv. *PeerJ* 7, e6876. <https://doi.org/10.7717/peerj.6876>
- Rue, H., Held, L., 2005. *Gaussian Markov Random Fields: Theory and Applications*. CRC Press.
- Shono, H., 2008. Application of the Tweedie distribution to zero-catch data in CPUE analysis. *Fish. Res.* 93, 154–162. <https://doi.org/10.1016/j.fishres.2008.03.006>
- Thorson, J.T., Ward, E.J., 2014. Accounting for vessel effects when standardizing catch rates from cooperative surveys. *Fish. Res.* 155, 168–176. <https://doi.org/10.1016/j.fishres.2014.02.036>
- Winker, H., Kerwath, S.E., Attwood, C.G., 2013. Comparison of two approaches to standardize catch-per-unit-effort for targeting behaviour in a multispecies hand-line fishery. *Fish. Res.* 139, 118–131. <https://doi.org/10.1016/j.fishres.2012.10.014>
- Wood, S.N., 2017. *Generalized Additive Models: An Introduction with R*, Second Edition, 2nd ed. Chapman and Hall/CRC, Boca Raton. <https://doi.org/10.1201/9781315370279>
- Wood, S.N., Goude, Y., Shaw, S., 2015. Generalized Additive Models for Large Data Sets. *J. R. Stat. Soc. Ser. C Appl. Stat.* 64, 139–155. <https://doi.org/10.1111/rssc.12068>

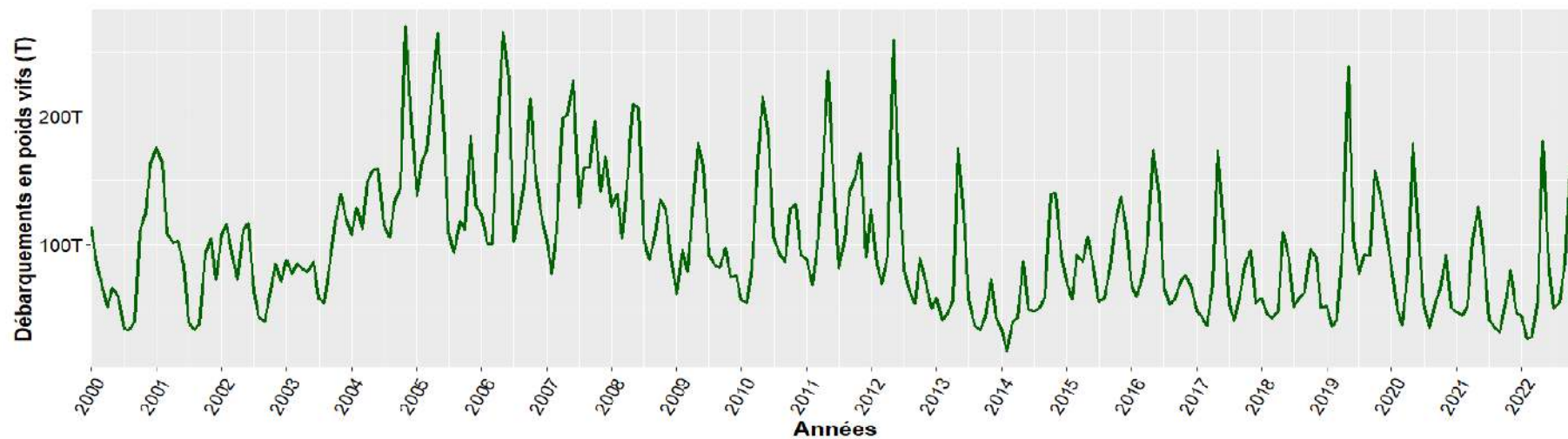
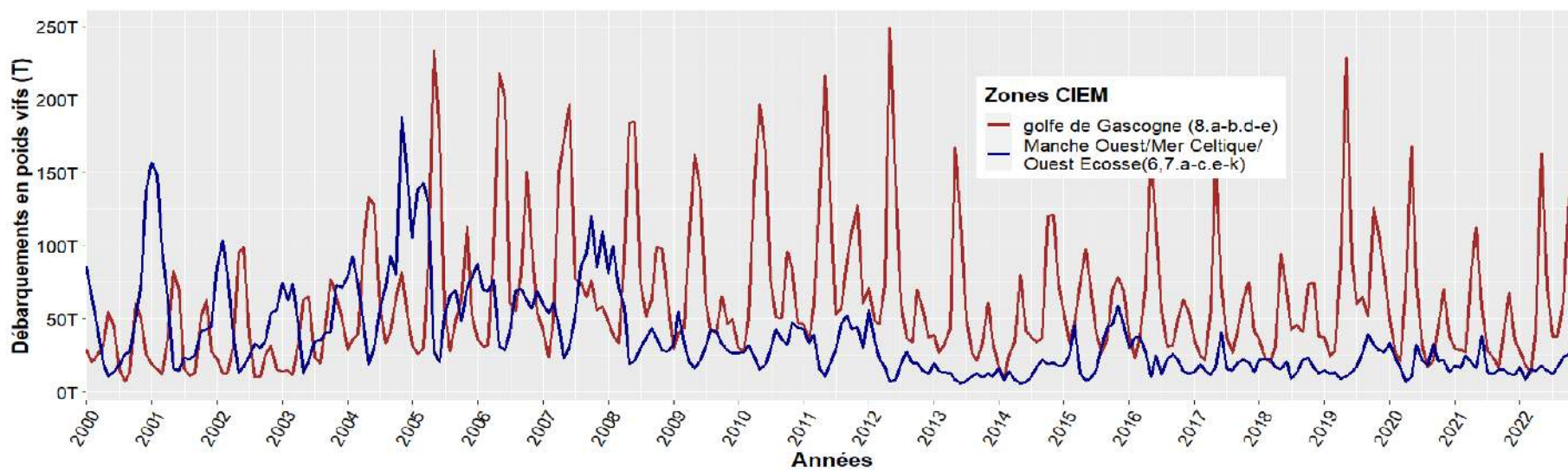
Annexes

Annexe 1 | Tableau des variables sélectionnées issues de l'algorithme Sacrois (Demanèche et al., 2015).

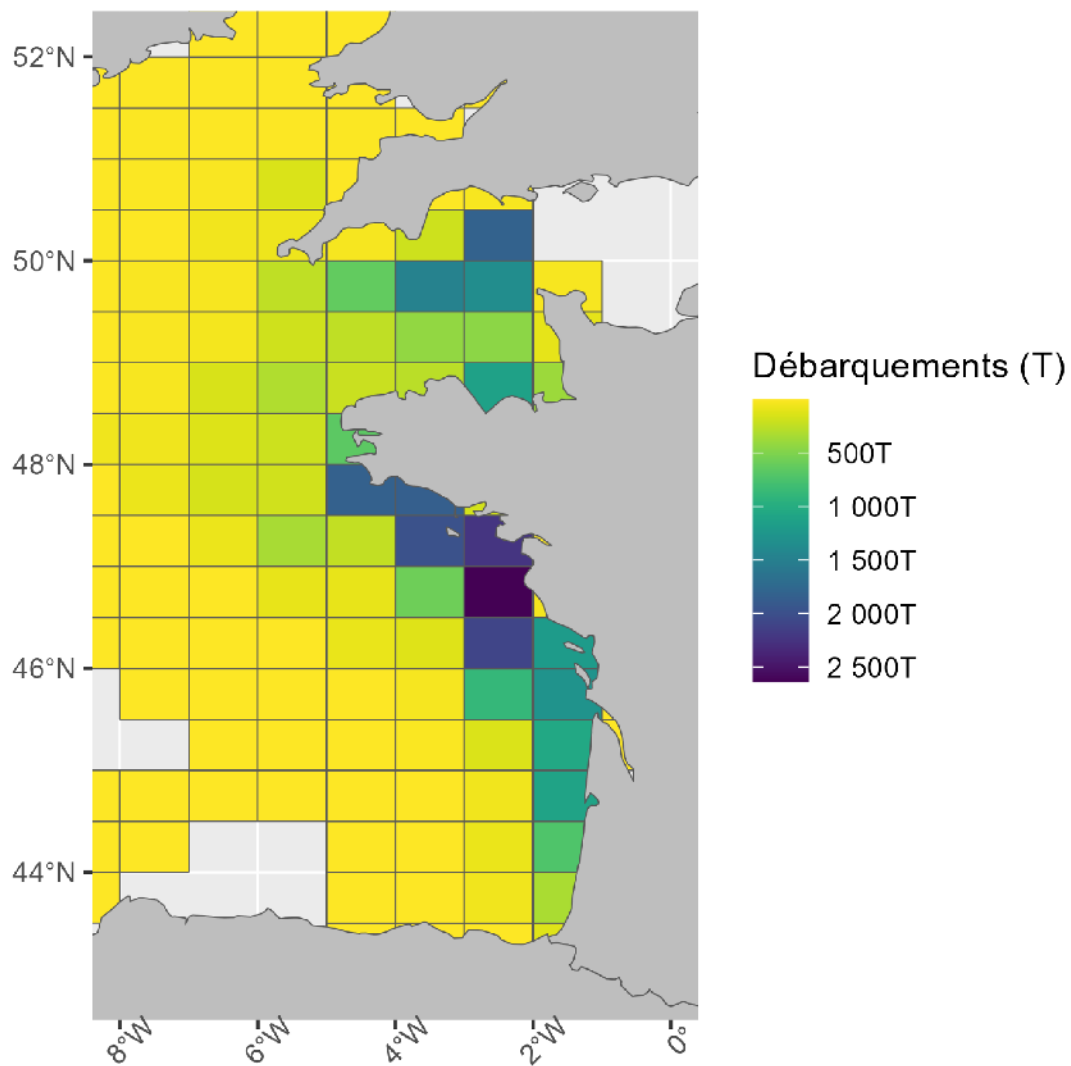
Variable	Description
Year	Années.
Month	Mois.
NAVS_COD	Immatriculation nationale du navire (identifiant unique pour le navire).
MAREE_ID	Identifiant unique de la marée SACROIS (identifiant attribué à l'ensemble des marées SACROIS quel que soit leur origine).
SEQ_ID	Identifiant unique des séquences de pêche SACROIS. Une séquence de pêche SACROIS est définie par les informations suivantes. Ainsi un changement d'engin ou de zone de pêche sur une même journée aura pour conséquence de changer de séquence de pêche.
SECT_COD_SACROIS_NIV4	Code de la zone FAO.
SECT_COD_SACROIS_NIV5	Code du carrée statistique.
METIER_DCF_5_COD	Code du métier au niveau 5 de la DCF (cf. Décision de la commission N°2010/93/EU, Appendix IV) estimé pour la séquence de pêche. Métier DCF niveau 5 lié au métier SIH estimé par l'algorithme SACROIS selon un référentiel d'agrégation des métiers du SIH.
ENGIN_COD	Code FAO de l'engin de pêche associé à la séquence de pêche SACROIS. L'engin de pêche est défini selon l'algorithme suivant : - Dans les cas où une marée JBE-SACAPT/SACAPT (origine : « SIPA », « ERS », « SIPA-OFI », « ERS-OFI », « SIPA-VMS », « ERS-VMS », « SIPA-VMS-OFI » ou « ERS-VMS-OFI ») est liée à la marée SACROIS, l'engin est déduit directement de celle-ci. Lorsque l'engin déclaré dans la marée JBE-SACAPT/SACAPT est « manquant » ou indéfini (engin MIS ou NK), alors les cas suivants de l'algorithme s'appliquent, - Engin déduit de l'historique (recherche dans les 6 mois précédents et les 6 mois suivants) des données « SACROIS » du navire (engin de pratique principal du navire pour la séquence de pêche SACROIS la « plus proche » renseignée), - Pour les navires où aucun historique SACROIS pour l'information "engin" n'est disponible, l'engin de pêche est déduit des enquêtes activité du mois concerné. Seuls les mois où un seul engin de pêche est déclaré dans l'activité sont considérés. Les enquêtes activité des 2 années précédentes sont utilisées, - Pour les navires où aucun historique SACROIS pour l'engin n'est disponible et où les enquêtes « activité » ne permettent pas de définir un engin, l'engin

	MIS (divers engins de pêche) est associé à la séquence de pêche.
ENGIN_COD_SACROIS	Engin déployé par le navire estimé par SACROIS.
TP_NAVIRE_SACROIS	Temps de pêche navire (en heures décimales) calculé par l'algorithme SACROIS/Jour. Ce temps de pêche est issu : - Soit d'une consolidation des deux temps de pêche VMS et SACAPT disponibles pour les jours où, à la fois, une séquence de pêche SIPA/SACAPT a été déclarée et une séquence de pêche VMS a été calculée. - Soit du temps de pêche navire tel que déclaré dans le document déclaratif associé (logbook ou fiche de pêche) pour les jours où seule une séquence de pêche SIPA/SACAPT a été déclarée. Ce champ n'est pas renseigné pour les séquences de pêche VMS « orphelines ». Donnée d'effort qui fait le pendant de la donnée de capture « QUANT_POIDS_VIF_REF_SACROIS ».
TPS_MER	Temps de mer en heure décimale de la marée SACROIS (MAREE_DAT_RET - MAREE_DATE_DEP). Ce champ n'est pas renseigné pour les ventes orphelines (origine : « OFI » ou « OFI-Histo »).
ESP_COD_FAO	Code FAO des différentes espèces vendues/.
QUANT_POIDS_VIF_SACROIS	Quantité totale en poids vif et en kg de la capture « estimée » issue de l'application de l'algorithme SACROIS/Jour pour l'espèce et la séquence de pêche (au sens ici des critères d'agrégation retenus soit les critères suivants : navire*marée*date_ref*lieu_ref*engin*zone de pêche*zpt*jour) considérées. Cette quantité SACROIS est répartie selon une formule qui fait intervenir à la fois le temps de pêche navire « de référence » VMS et la répartition des espèces ou groupes d'espèces observée par séquence de pêche dans le flux déclaratif SACAPT/SIPA.
MONTANT_EUROS_SACROIS	Montant total en euros de la capture « estimé » par catégorie commerciale. Ce montant correspond à la quantité « QUANT_POIDS_VIF_REF_SACROIS » auquel un prix moyen ou un montant a pu être associé selon les modalités précisées dans le champ «ORIGINE_MONTANT_EUROS »
Annee_construction	Année de construction du navire.
Loa	Longueur du navire.
Power_Main	Puissance du navire.
Ton_Ref	Tonnage de référence du navire.
Vms_Code	Variable binaire de la mise en place de la VMS sur le navire (oui ou non).

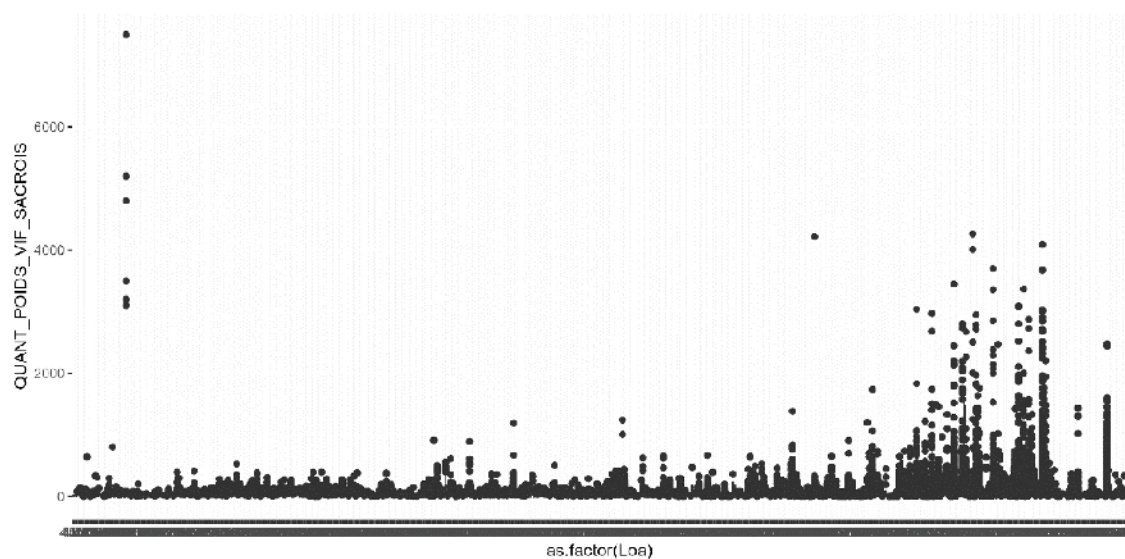
Annexe 2 | Série temporelle des débarquements de rouget-barbet de roche. Haut = Deux sous-stocks étudiés. Bas = Zone totale



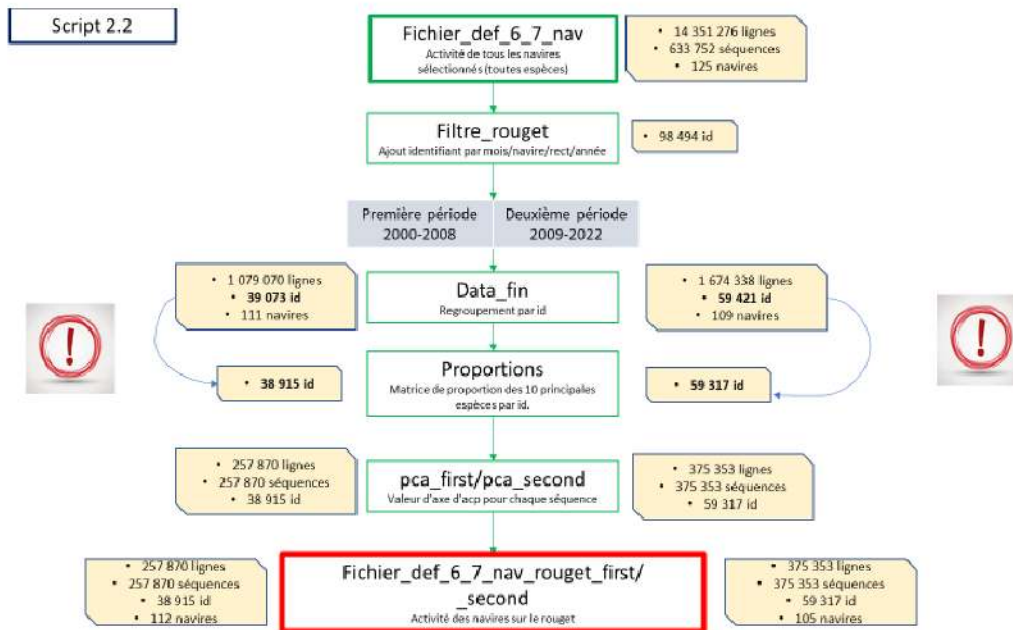
Annexe 3 | Cartographie des débarquements totaux de *Mullus surmuletus* cumulés sur la période 2000-2022 par rectangle statistique.



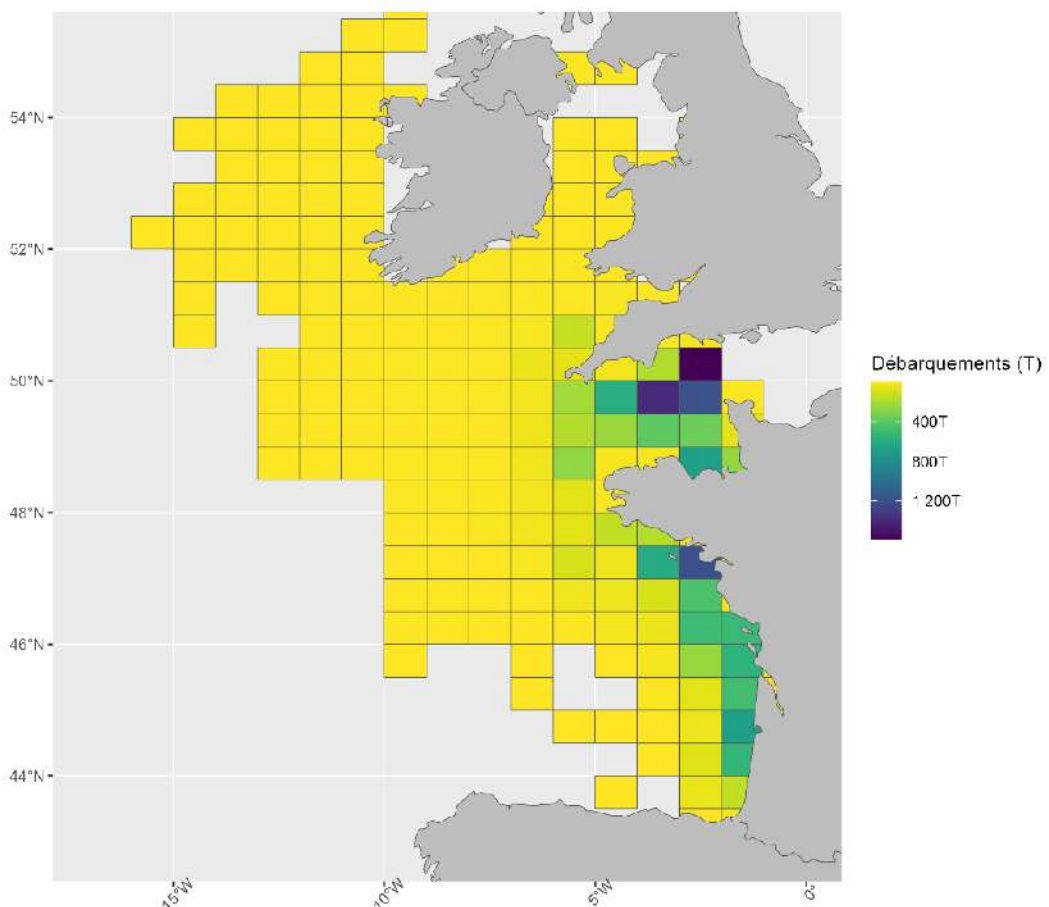
Annexe 4 | Boxplot des débarquements par taille de navire sur toute la zone.



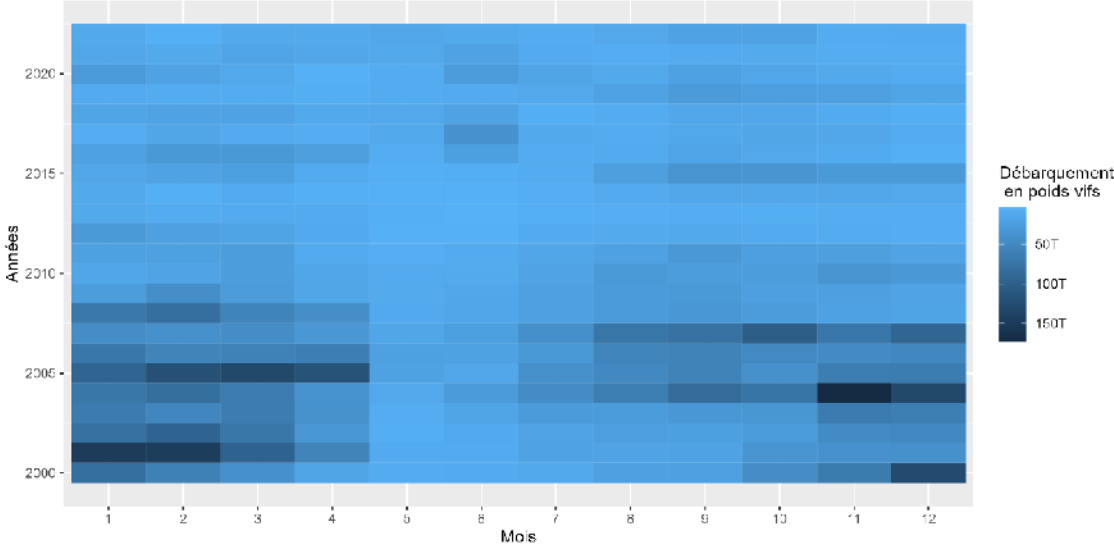
Annexe 5 | Processus de création de la variable de ciblage du stock Mixte (processus identique pour la zone golfe de Gascogne). Voir Annexe 18 et Annexe 19 pour les processus précédentes.



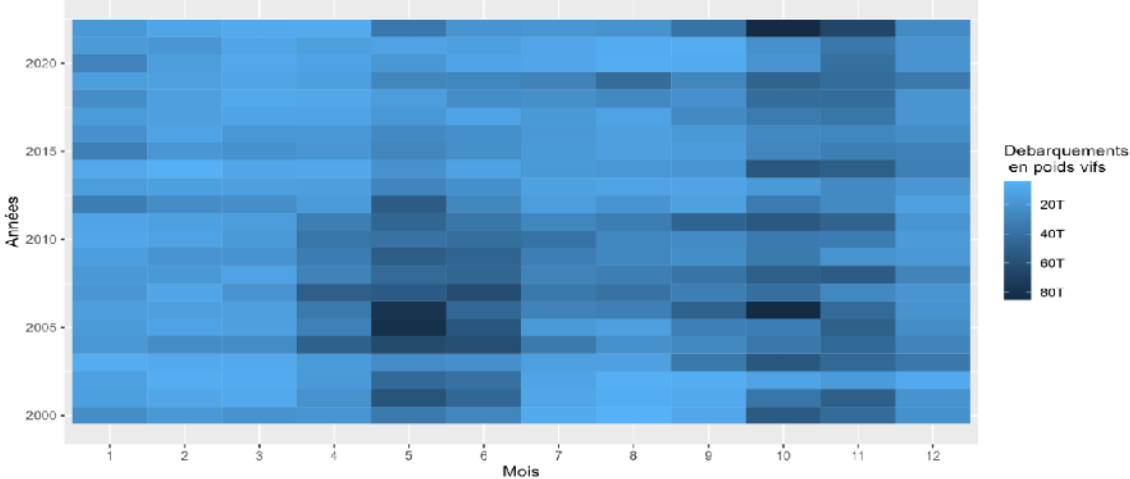
Annexe 6 | Cartographie des débarquements totaux de *Mullus surmuletus* cumulés sur la période 2000-2022 par rectangle statistique pour les navires mettant en place l'engin OTB (Chalut de fond à panneaux).



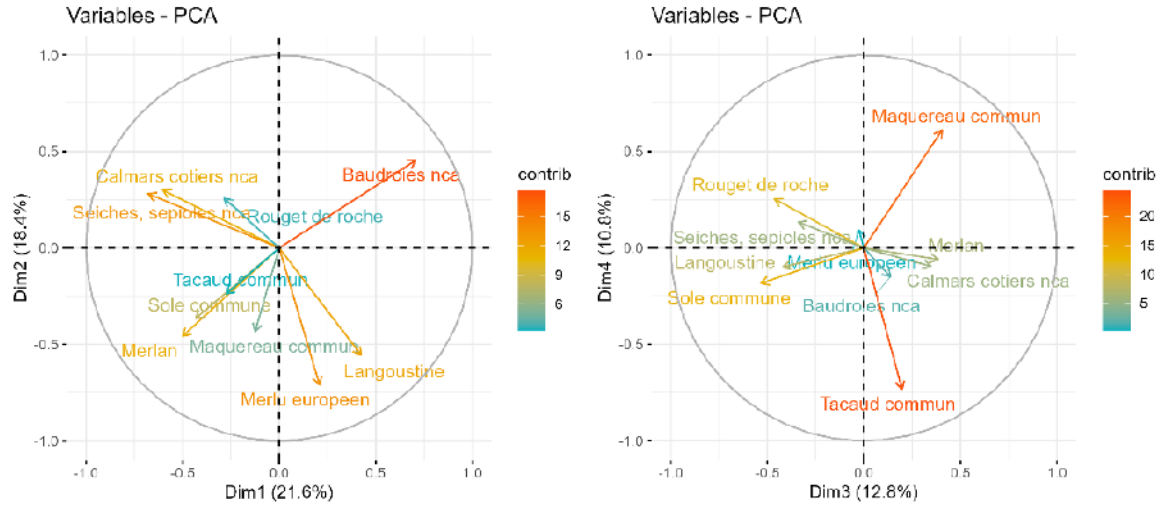
Annexe 7 | Heat map des débarquements annuels par année pour les engins OTB du stock Mixte.



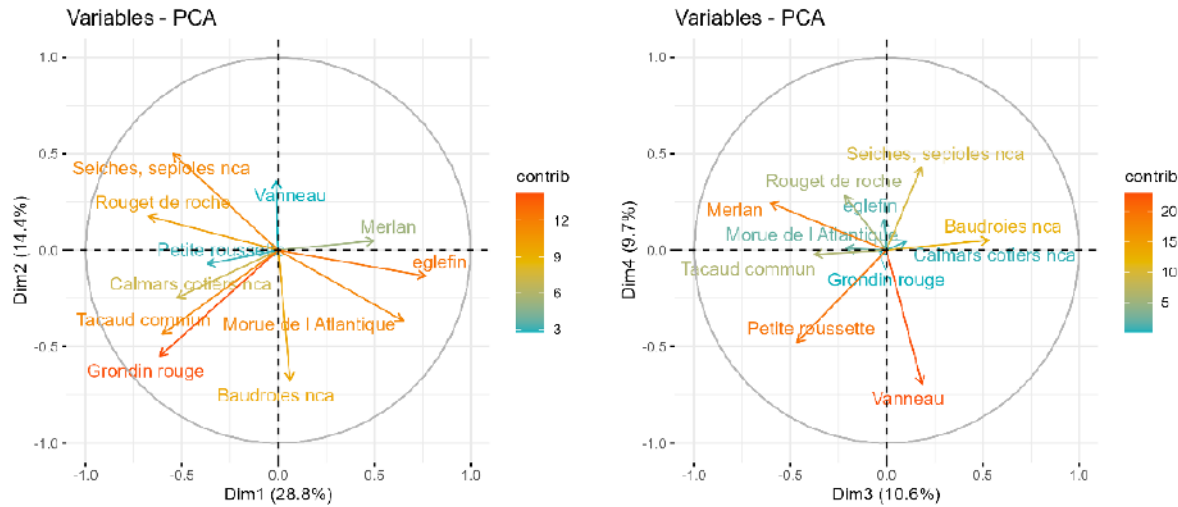
Annexe 8 | Heat map des débarquements annuels par année pour les engins OTB du stock Golfe.



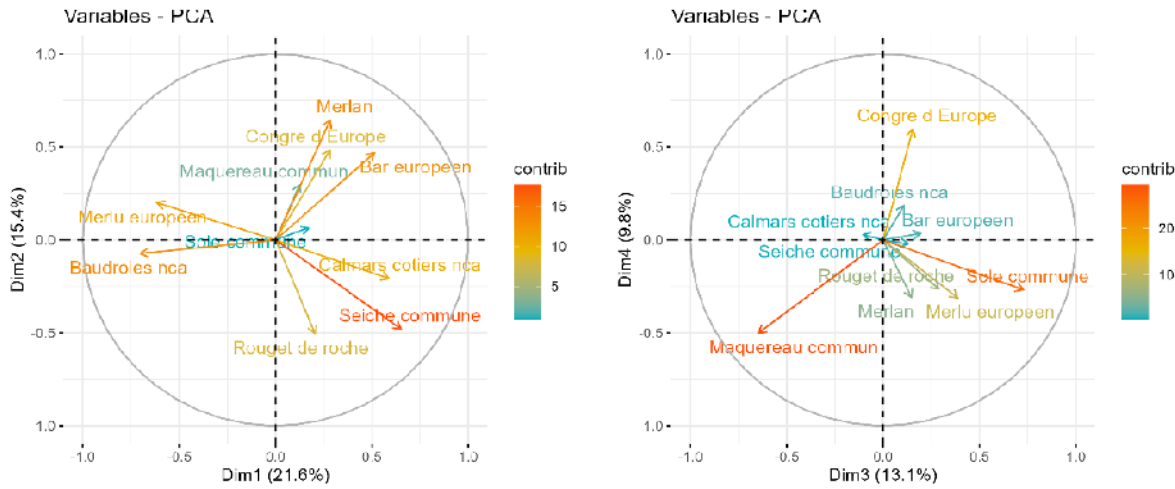
Annexe 9 | Graphiques des variables des ACP du stock Golfe pour la période 2000-2008



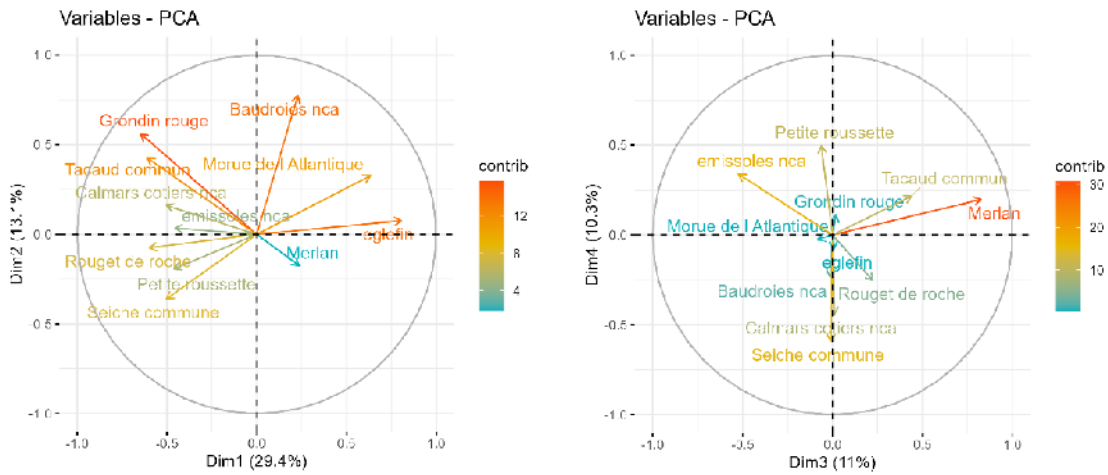
Annexe 10 | Graphiques des variables des ACP du stock Mixte pour la période 2000-2008



Annexe 11 | Graphiques des variables des ACP du stock Golfe pour la période 2009-2022



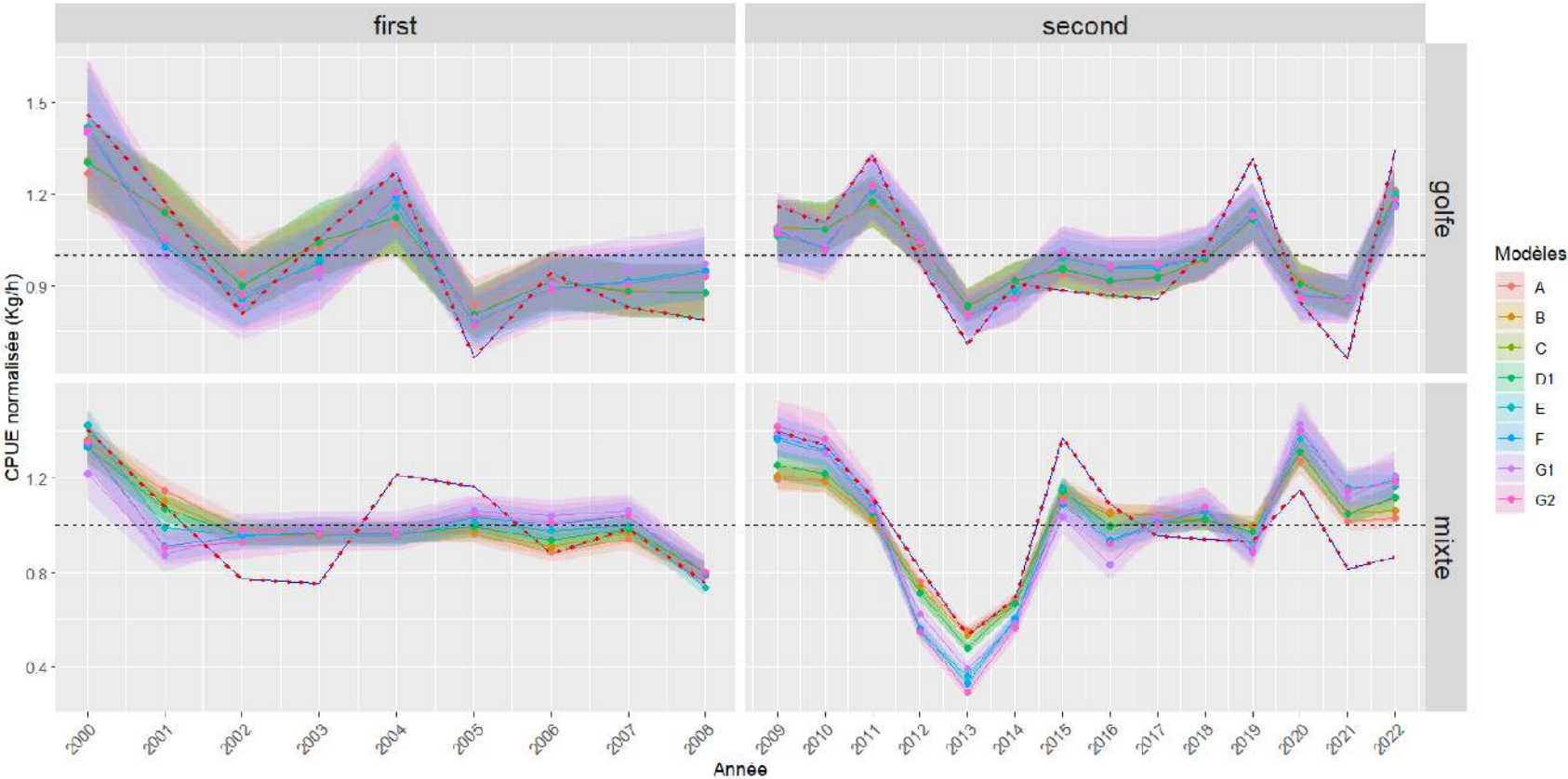
Annexe 12 | Graphiques des variables des ACP du stock Mixte pour la période 2009-2022



Annexe 13 | Récapitulatif de l'ajout de chaque variable à partir la structure aléatoire des 4 modèles. On regarde ici variable par variable en fonction de la structure aléatoire.

	Variable ajouté	df-edf	% déviance	p value	df-edf	% déviance	p value
		golfe de Gascogne			Zone 6-7		
1ère période	offset(log(TP_NAVIRE_SACROIS)) + year + s(NAVS_COD, bs = 're')	228,1 + 8	14,66%	<2e-16 ***	119,70 + 8	27,90%	< 2e-16 ***
	s(month, bs='cc', k = 12)	9,94	19,45%	<2e-16 ***	10,13	37,76%	< 2e-16 ***
	s(carre.lon, carre.lat, k = 40)	33,84	15,01%	<2e-16 ***	37,74	39,19%	< 2e-16 ***
	Power_Main	1	14,68%	3,54e-06 ***	0,80	27,90%	0,70213
	as.factor(Loa)	5	14,68%	1/5 significatif	1,00	27,91%	2/5 significatif
	Ton_Ref	1	14,67%	4,88e-05 ***	1,12	27,95%	< 2e-16 ***
	s(RS1) + s(RS2) + s(RS3) + s(RS4)	8,8 (x4)	38,74%	<2e-16 ***	35,31	59,31%	< 2e-16 ***
	s(month, bs='cc', k = 12, by = year)	9,6 (x8)	22,19%	<2e-16 ***	89,31	41,19%	< 2e-16 ***
	te(carre.lon, carre.lat, year_bis, k = c(40, 8), d = c(2,1))	37,53	18,50%	<2e-16 ***	36,02	39,20%	< 2e-16 ***
te(carre.lon, carre.lat, month, k = c(40,12), d = c(2,1), bs = c('tp', 'cc'))	38,92	30,29%	<2e-16 ***	38,15	47,48%	< 2e-16 ***	
2nd période	offset(log(TP_NAVIRE_SACROIS)) + year + s(NAVS_COD, bs = 're')	206,7 + 12	12,19%	< 2e-16 ***	107,2 + 12	17,73%	< 2e-16 ***
	s(month, bs='cc', k = 12)	9,917	14,24%	< 2e-16 ***	10,01	20,61%	< 2e-16 ***
	s(carre.lon, carre.lat, k = 30)	27,5	13,18%	< 2e-16 ***	35,46	25,96%	< 2e-16 ***
	Power_Main	1	12,19%	0,43	0,96	17,75%	< 2e-16 ***
	as.factor(Loa)	5	12,19%	2/5 significatif	1,04	17,81%	3/5 significatif
	Ton_Ref	1	12,19%	0,00658 **	0,00	17,73%	0,0994
	s(RS1) + s(RS2) + s(RS3)	8,8 (x4)	32,27%	< 2e-16 ***	35,34	44,03%	< 2e-16 ***
	s(month, bs='cc', k = 12, by = year)	9,6 (x12)	17,43%	< 2e-16 ***	135,86	23,14%	< 2e-16 ***
	te(carre.lon, carre.lat, year_bis, k = c(30, 8), d = c(2,1))	28,57	15,50%	< 2e-16 ***	227,54	29,30%	< 2e-16 ***
te(carre.lon, carre.lat, month, k = c(30,12), d = c(2,1), bs = c('tp', 'cc'))	38,93	22,61%	< 2e-16 ***	315,5387	38,11%	<2e-16 ***	

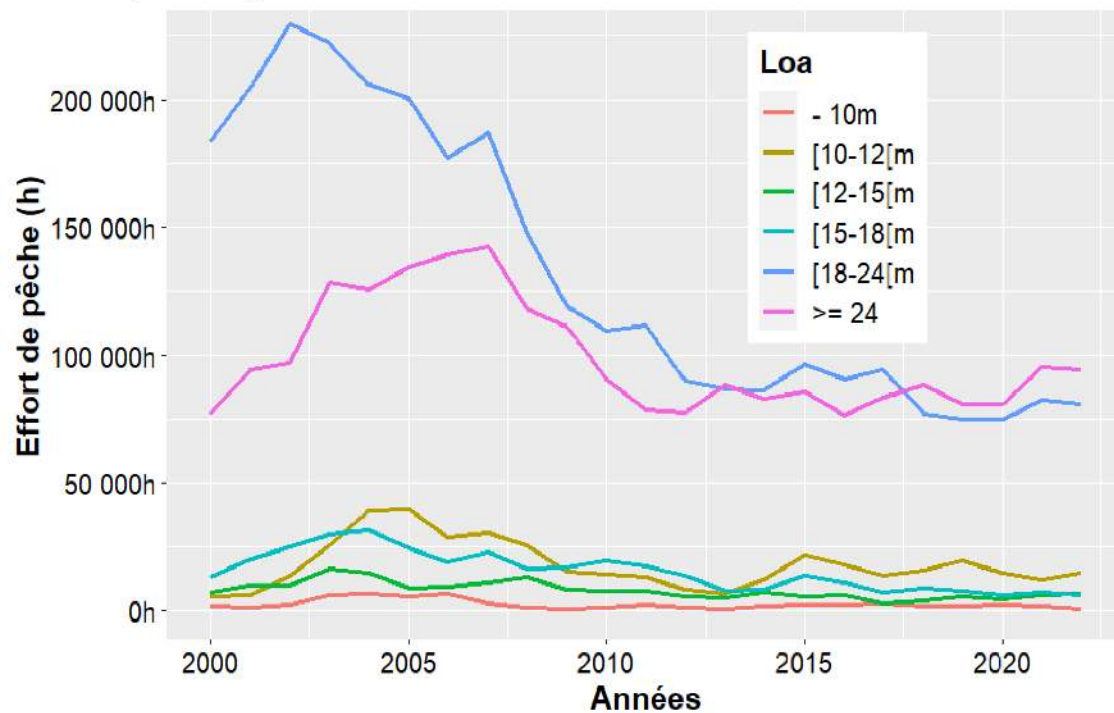
Annexe 14 | Ajustement aux données des 4 modèles. La courbe en pointillée rouge correspond à la CPUE nominale.



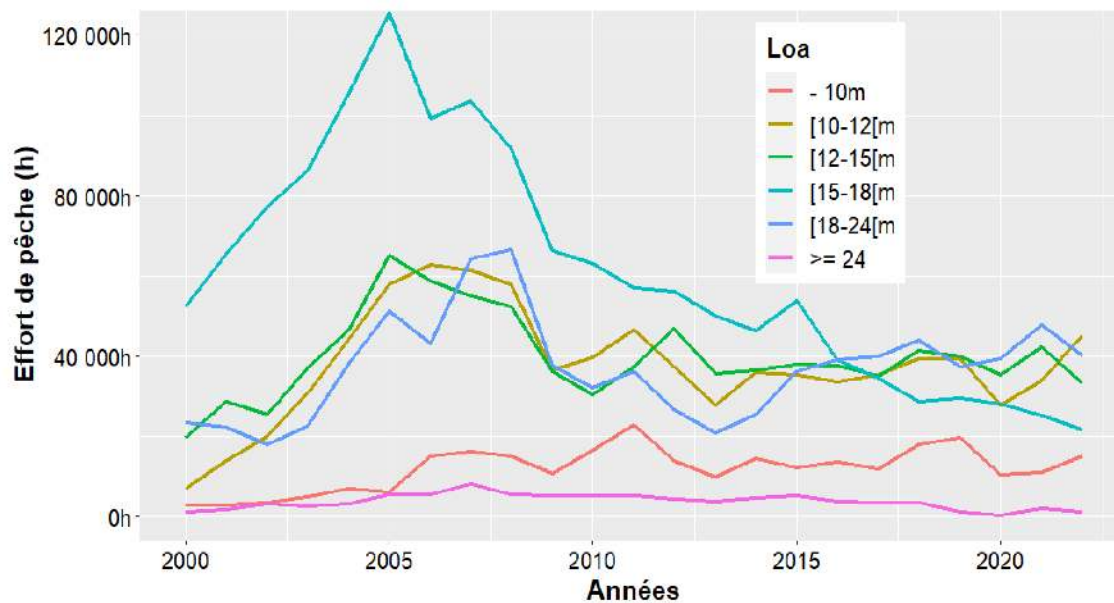
Annexe 15 | Valeurs des indices d'abondances finaux issue du processus de standardisation des CPUE (Kg/h) de la flottille de navire mettant en place l'engin OTB. Indices par sous-stock (golfe et mixte) et par période (first : 2000-2008 ; second : 2009-2022). Le tableau comprend les valeurs prédites de CPUE ('mu'), l'erreur standard ('se'), les intervalles de confiance à 95% supérieurs ('uci') et inférieurs ('lci').

year	mu	se	lci	uci	zone	periode
2000	1,41	0,05	1,27	1,55	golfe	first
2001	1,87	0,04	1,71	2,04	golfe	first
2002	1,12	0,04	1,03	1,22	golfe	first
2003	0,82	0,05	0,75	0,90	golfe	first
2004	1,00	0,04	0,93	1,07	golfe	first
2005	0,79	0,04	0,74	0,85	golfe	first
2006	0,84	0,04	0,78	0,90	golfe	first
2007	0,55	0,04	0,51	0,59	golfe	first
2008	0,60	0,04	0,56	0,65	golfe	first
2009	1,04	0,04	0,96	1,12	golfe	second
2010	1,03	0,04	0,96	1,10	golfe	second
2011	1,23	0,03	1,15	1,31	golfe	second
2012	0,83	0,04	0,77	0,89	golfe	second
2013	0,89	0,04	0,83	0,96	golfe	second
2014	1,10	0,03	1,03	1,17	golfe	second
2015	1,00	0,03	0,94	1,06	golfe	second
2016	0,71	0,04	0,66	0,76	golfe	second
2017	0,82	0,03	0,77	0,88	golfe	second
2018	1,08	0,03	1,01	1,15	golfe	second
2019	1,04	0,04	0,96	1,11	golfe	second
2020	0,90	0,03	0,85	0,97	golfe	second
2021	0,90	0,03	0,84	0,96	golfe	second
2022	1,44	0,03	1,35	1,54	golfe	second
2000	1,10	0,03	1,04	1,18	mixte	first
2001	1,81	0,03	1,70	1,92	mixte	first
2002	0,87	0,03	0,83	0,92	mixte	first
2003	0,77	0,03	0,73	0,81	mixte	first
2004	0,85	0,02	0,81	0,89	mixte	first
2005	1,07	0,02	1,02	1,12	mixte	first
2006	0,90	0,02	0,85	0,94	mixte	first
2007	0,77	0,03	0,74	0,81	mixte	first
2008	0,86	0,02	0,82	0,90	mixte	first
2009	1,17	0,03	1,11	1,23	mixte	second
2010	0,92	0,02	0,88	0,96	mixte	second
2011	1,37	0,03	1,30	1,44	mixte	second
2012	1,30	0,03	1,23	1,37	mixte	second
2013	0,82	0,03	0,78	0,87	mixte	second
2014	0,73	0,03	0,69	0,78	mixte	second
2015	0,92	0,03	0,87	0,97	mixte	second
2016	1,19	0,03	1,13	1,26	mixte	second
2017	0,97	0,03	0,91	1,03	mixte	second
2018	1,04	0,03	0,98	1,10	mixte	second
2019	0,64	0,03	0,60	0,68	mixte	second
2020	0,91	0,03	0,85	0,97	mixte	second
2021	1,18	0,03	1,11	1,25	mixte	second
2022	0,85	0,03	0,80	0,91	mixte	second

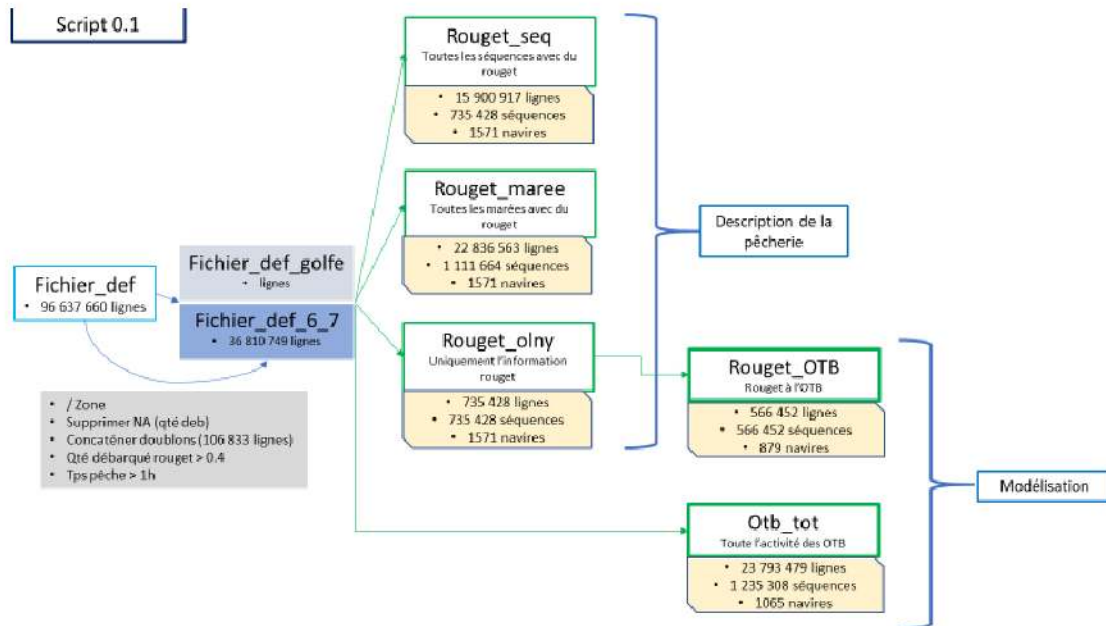
Annexe 16 | Évolution de l'effort de pêche total du stock Mixte en fonction de la taille des navires mettant en place l'engin OTB



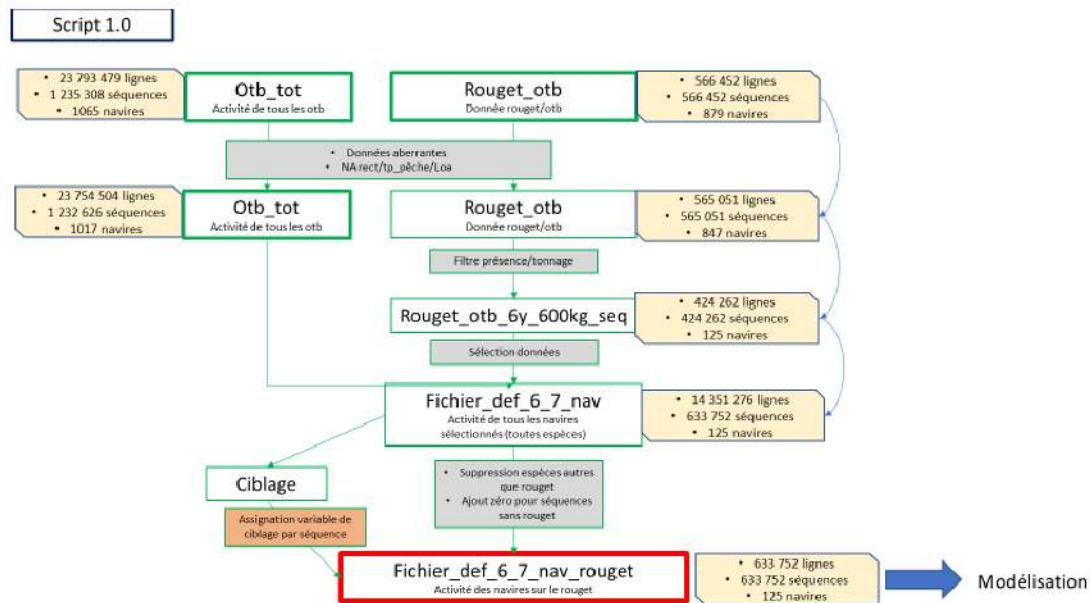
Annexe 17 | Évolution de l'effort de pêche total du stock Golfe en fonction de la taille des navires mettant en place l'engin OTB.




Annexe 18 | Processus de modification des données et de création de sous-ensembles.



Annexe 19 | Processus de sélection des données avant modélisation du stock Mixte (processus identique pour la zone golfe de Gascogne).



	<p>Diplôme : Master Science de la Mer et du Littoral</p> <p>Spécialité : Science Halieutique et Aquacole</p> <p>Spécialisation / option : REA</p> <p>Enseignant référent : RIVOT Etienne</p>
<p>Auteur(s) : DENIAU Clémentin</p>	<p>Organisme d'accueil : Ifremer</p>
<p>Date de naissance* :12/06/1999</p>	<p>Adresse : Rue de l'île d'Yeu - BP 21105 - 44311 Nantes cedex 03</p>
<p>Nb pages : 33 Annexe(s) : 13</p>	<p>Maître de stage : CAILL-MILLY Nathalie</p>
<p>Année de soutenance : 2023</p>	
<p>Titre français : Standardisation des données de captures par unité d'effort du rouget barbet de roche (<i>Mullus surmuletus</i>) du stock de la zone Ouest de l'Atlantique Nord Est (zones CIEM 6, 8, et les divisions 7.a–c, 7.e–k, et 9.a)</p> <p>Titre anglais : Standardization of catch-per-unit-effort data for striped red mullet (<i>Mullus surmuletus</i>) in the Western North-East Atlantic stock (ICES areas 6, 8, and divisions 7.a-c, 7.e-k, and 9.a)</p>	
<p>Résumé (1600 caractères maximum) : Le stock de rouget-barbet de roche (<i>Mullus surmuletus</i>) des sous-zones et divisions 6, 8, 7.a–c, 7.e–k, et 9.a est aujourd'hui très peu étudié. Ce manque de connaissance a conduit à le classer dans la catégorie 5 du CIEM dite 'data limited stock'. Le projet ACOST a pour objectif d'améliorer les connaissances sur ce stock ainsi que sur trois autres présents dans le golfe de Gascogne (merlan, lieu jaune, maigre). Cela passe notamment par la construction d'indices de suivis du stock. Le stock du rouget-barbet de roche étudié ici a été divisé en deux zones : golfe de Gascogne et Mer du Nord/Mer Celtique. Les données professionnelles issues de la base de données Sacrois sont utilisées pour construire un indice d'abondance. Une méthode de standardisation des Captures Par Unité d'Effort (CPUE) est mise en place à l'aide d'un modèle additif mixte généralisé utilisant une distribution « tweedie ». Nous sélectionnons une flottille spécifique constituée des chaluts de fond à panneaux. Les variables du mois, des rectangles statistiques, de caractéristiques de navires et de tactiques de pêches sont utilisées. Le navire est spécifié comme variable aléatoire. L'ajustement des modèles se montre faible pour les fortes valeurs de débarquements mais correct pour les faibles et moyennes valeurs. Les indices d'abondances ainsi créés ne montrent pas de fortes différences avec la tendance observée à partir des données brutes. Ces derniers ont pour vocation à être utilisés dans des modèles de productions de biomasse tel que Spict, non développé ici.</p>	
<p>Abstract (1600 caractères maximum) : The striped red mullet (<i>Mullus surmuletus</i>) stock in sub-areas and divisions 6, 8, 7.a-c, 7.e-k, and 9.a is currently very poorly studied. This lack of knowledge has led to its classification in ICES category 5, known as 'data limited stock'. The ACOST project aims to improve knowledge of this stock and three others in the Bay of Biscay (whiting, pollack and meagre). This includes the construction of stock monitoring indices. The red mullet stock studied here has been divided into two zones : Bay of Biscay and North Sea/Celtic Sea. Professional data from the Sacrois database are used to construct an abundance index. A method for standardizing Catch Per Unit Effort (CPUE) is implemented using a generalized mixed additive model with a "tweedie" distribution. We select a specific fleet of bottom otter trawls. The variables of month, statistical rectangles, vessel characteristics and fishing tactics are used. Vessel is specified as a random variable. Model fit was poor for high landings, but good for low and medium landings. The abundance indices thus created do not show strong differences with the trend observed from the raw data. They are intended for use in biomass production models such as Spict, not developed here.</p>	
<p>Mots-clés : Façade Atlantique, Système d'Informations Halieutiques (SIH), Indice d'abondance, stock à données limitées, modélisation. Key Words: Atlantic coast, Fisheries Information System (HIS), abundance index, data-limited stock, modeling.</p>	

* Élément qui permet d'enregistrer les notices auteurs dans le catalogue des bibliothèques universitaires