HAL open science

Détermination de limite de conservation du saumon Atlantique (Salmo salar) : analyse de sensibilité à partir du cas d'étude de la Nivelle (Pays Basque, France) Rapport final

Benjamin Flais

To cite this version:

Benjamin Flais. Détermination de limite de conservation du saumon Atlantique (Salmo salar) : analyse de sensibilité à partir du cas d'étude de la Nivelle (Pays Basque, France) Rapport final. Sciences de l'environnement. 2022. hal-04115800

HAL Id: hal-04115800

https://hal.inrae.fr/hal-04115800v1

Submitted on 2 Jun 2023

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

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

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Détermination de limite de conservation du saumon Atlantique (*Salmo salar*) : analyse de sensibilité à partir du cas d'étude de la Nivelle (Pays Basque, France)

Rapport final

Benjamin FLAIS
Pôle pour la gestion des MIgrateurs AMphihalins dans leur Environnement, Université de Pau et des Pays de l'Adour, e2s UPPA, INRAE, OFB, Institut Agro, ECOBIOP, Saint-Pée-sur-Nivelle, France

Septembre – 2022

- **AUTEUR**

Benjamin FLAIS, Université de Pau et des Pays de l'Adour, e2s UPPA, INRAE, OFB, Institut Agro, ECOBIOP, Pôle pour la gestion des MIgrateurs AMphihalins dans leur Environnement, Saint-Pée-sur-Nivelle

- **CONTRIBUTEUR**

Etienne PREVOST, Directeur de recherche, Université de Pau et des Pays de l'Adour, e2s UPPA, INRAE, OFB, Institut Agro, ECOBIOP, Pôle pour la gestion des MIgrateurs AMphihalins dans leur Environnement, Saint-Pée-sur-Nivelle, etienne.prevost@inrae.fr
Laurent BEAULATON, Chef du pôle OFB-INRAE-Institut Agro-UPPA pour la gestion des MIgrateurs AMphihalins dans leur Environnement, Rennes, laurent.beaulaton@ofb.gouv.fr

Droits d'usage : accès libre
Niveau géographique : national
Couverture géographique : Pyrénées Atlantiques
Niveau de lecture : experts

- **TITRE, AUTEURS**

Détermination de limite de conservation du saumon Atlantique (*Salmo salar*) : analyse de sensibilité à partir du cas d'étude de la Nivelle (Pays Basque, France)

Benjamin FLAIS, Université de Pau et des Pays de l'Adour, e2s UPPA, INRAE, OFB, Institut Agro, ECOBIOP, Pôle pour la gestion des MIgrateurs AMphihalins dans leur Environnement, Saint-Pée-sur-Nivelle

- **RESUME**

La limite de conservation (LC) est un point de référence biologique important pour le saumon atlantique (*Salmo salar*). Dans cette étude la LC correspond à l'abondance de géniteurs permettant de contrôler le risque de faible recrutement. La relation stock-recrutement (SR), au cœur de cette définition, peut être décomposée en plusieurs composantes : courbe SR, erreur de processus, connaissance a priori du modèle SR, erreur de mesure. L'objectif principal de ce travail est d'analyser la sensibilité des LCs à ces différentes composantes. Plusieurs types de courbes SR et hypothèses sur l'erreur de processus ainsi que différents niveaux de connaissance a priori du modèle SR sont analysés. Pour ce faire une démarche séquentielle est suivie afin de limiter le nombre de combinaisons à analyser. L'impact des erreurs de mesure sur la LC est analysé à la suite de cette analyse séquentielle sur une sélection préalable de combinaisons. Ce travail est mené au moyen d'une étude de cas : la population de saumon atlantique de la Nivelle (Pays Basque, France). Dans notre cas la LC peut radicalement changer selon la représentation des différentes composantes de la relation SR. La prise en compte des erreurs de mesure ne semble pas avoir un impact important sur les LCs. Enfin nous discutons de l'importance d'analyser de manière précise la stochasticité du recrutement pour déterminer des LCs fondées sur une maitrise du risque de faible recrutement.

- **MOTS CLES (THEMATIQUES ET GEOGRAPHIQUES)**

Saumon atlantique, Nivelle, Limite de conservation, Stock, Recrutement, Relation SR

- ## SOMMAIRE

TEXTE INTEGRAL DU MEMOIRE DE FIN D'ETUDE DE BENJAMIN FLAIS .....

Le présent travail a été réalisé dans le cadre du pôle MIAME qui l'a soutenu financièrement au titre de son volet « petites études ». Il a été mené via un stage de Master 2 dont le mémoire est donné ci-après.

L'Institut Agro Rennes-Angers
☐ **Site d'Angers** ☒ **Site de Rennes**

Année universitaire : 2021.- 2022 Spécialité : SML- Biologie Parcours (et option éventuelle) : Sciences Halieutiques et aquacoles (Ressources et écosystèmes aquatiques)	**Mémoire de fin d'études** ☐ D'ingénieur de l'Institut Agro Rennes-Angers (Institut national d'enseignementsupérieur pour l'agriculture, l'alimentation et l'environnement) ☒ De master de l'Institut Agro Rennes-Angers (Institut national d'enseignementsupérieur 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)

Détermination de limite de conservation du saumon Atlantique (*Salmo salar)* : analyse de sensibilité à partir du cas d'étude de la Nivelle (Pays Basque, France).

Benjamin FLAIS

Soutenu à Rennes le 16/09/2022

Devant le jury composé de :
Président : Etienne RIVOT

Maître de stage : Etienne PREVOST

Enseignant référent : Etienne RIVOT

Autres membres du jury :

Didier GASCUEL, Directeur de Pôle halieutique mer et littoral, Institut Agro.
Marie NEVOUX, Chargé de recherche, INRAE UMR DECOD

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

Ce document est soumis aux conditions d'utilisation «Paternité-Pas d'Utilisation Commerciale-Pas de Modification 4.0 France» disponible en ligne http://creativecommons.org/licenses/by-nc-nd/4.0/deed.fr

Fiche de confidentialité et de diffusion du mémoire

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] :
(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] **Nom Prénom** Flais Benjamin

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** : 24/08/2022

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** :

(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

Je tiens tout d'abord à remercier Etienne Prevost, mon maître de stage, pour sa disponibilité, ses conseils et son aide précieuse tout au long du stage incluant la rédaction du rapport. De plus je remercie tous les membres de l'UMR ECOBIOP pour leur accueil très chaleureux. Enfin je remercie Etienne Rivot, mon enseignant référent, pour sa disponibilité durant toute la durée du stage.

Table des matières

Introduction ***1***

1 Matériel et méthode ***4***

1.1 Approche générale **4**

1.1.1 Fondement théorique de la prise en compte des différentes sources d'incertitude 4

1.1.2 Méthode numérique pour la prise en compte des différentes sources d'incertitude 6

1.1.3 Diagramme de risque et LC : 7

1.1.4 Hypothèses alternatives de modélisation de la relation SR 7

1.1.4.1 Courbes SR 7

1.1.4.2 Erreur de processus 8

1.1.4.2.1 IID selon une loi log-normale 8

1.1.4.2.2 IID selon un mélange de deux lois log-normales 8

1.1.4.2.3 IID avec une variance dépendante du stock 9

1.1.4.2.4 Autocorrélation d'ordre 1 9

1.1.4.3 Combinaisons d'hypothèses 9

1.1.5 Choix des priors 10

1.1.6 Approche séquentielle pour l'exploration des options de modélisation de la relation SR 11

1.1.6.1 Vérification à posteriori 11

1.1.6.1.1 Approche sur données répliquées 11

1.1.6.1.2 Vérification des hypothèses du modèle : analyse des résidus standardisés 12

1.1.6.1.3 Vérification a posteriori des effets 12

1.1.6.2 Critères dits « d'information » 12

1.1.6.3 Identification et mise à jour a posteriori des modèles 12

1.2 Cas d'étude **13**

1.2.1 Site d'étude 13

1.2.2 Les données d'observations ($\boldsymbol{D^S D^R}$) 13

1.2.3 Estimation de l'abondance de géniteurs (stock) et de tacons 0+ (recrutement) 14

1.2.4 Diagramme de risque et LC 15

1.2.5 Hypothèses alternatives de modélisation de la relation SR 15

1.2.6 Choix des priors 15

1.3 Abréviations **17**

1.4 Calcul Bayésien **18**

2 Résultats ***19***

2.1 Analyse de sensibilité aux priors et courbes SR **19**

2.1.1 Vérification a posteriori 19

2.1.1.1 Comparaison entre les données répliquées et observées 19

2.1.1.2 Analyse des résidus standardisés 19

2.1.2 Critère d'information : WAIC 20

2.1.3 Mise à jour et identification des modèles 20

2.1.4 Choix de combinaisons de priors et courbe SR 22

2.2 Analyse de sensibilité aux hypothèses de l'erreur de processus **22**

2.2.1 Vérification à posteriori 22

2.2.1.1 Comparaison entre les données répliquées et observées 22

2.2.1.2 Analyse des résidus standardisées 22

2.2.1.3 Vérification à posteriori des effets 23

2.2.2 Critère d'information : WAIC 24

2.2.3 Mise à jour et identification des modèles 24
2.2.4 Choix des différentes hypothèses sur l'erreur de processus 27

2.3 Prise en compte des erreurs de mesures 28
2.3.1 Mise à jour et identification des modèles 29
2.3.2 Prédictions du recrutement 29
2.3.3 Diagramme de risque et Détermination des LC's 30

2.4 Analyse rétrospective du statut de conservation 31

3 Discussion 32

3.1 Approche générale 32

3.2 Hypothèses écologiques 33

3.3 Diagramme de risque et LC 34

3.4 Statut de conservation de la Nivelle 35

Références bibliographiques .....

Liste des abréviations

dpp : distribution de probabilité a posteriori

iid : Indépendante et identiquement distribuée

BH : Beverton-Holt

DS : Deriso-Schnute

LC : limite de conservation

Lien vers les annexes et diagnostiques de convergence

https://filesender.renater.fr/?s=download&token=200d2edd-405e-4c6d-8a00-1f0ead6ebf74

Table des figures

Figure 1 : Directed acyclic graph (DAG) décrivant la structure de dépendance conditionnelle entre les données d'observations de stock (S) et recrutement (R) (D^S, D^R), les séries de S et R observées (S^{obs}, R^{obs}), les paramètres du modèle SR (θ) et la série de S (S^{pred}) utilisée pour les prédictions du recrutement (R^{pred}). ___ 4
Figure 2 : Exemple d'une relation SR. Les points bleus correspondent aux valeurs médianes des séries (S^{obs}, R^{obs}). La courbe SR est en bleu foncé. Les distributions de probabilité rouges correspondent à l'erreur de processus. Les flèches bleu clair représentent les erreurs de mesures du stock et du recrutement. _____ 5
Figure 3 : À gauche : Comparaison des prédictions de recrutement et du seuil de faible recrutement correspondant à 50 % de la capacité d'accueil de la population (ligne rouge). Les portions des distributions de recrutement inférieures au seuil de faible recrutement sont figurées en rouge. Elles correspondent au risque de faible recrutement. À droite : Diagramme de risque qui présente l'évolution du risque de faible recrutement en

fonction du stock et permet de déterminer la LC selon le niveau de risque choisi (ici 25%). Source : (Lebot, 2021). _____ 7
Figure 4 : Carte de la rivière de la Nivelle (Pays-basque, France). Source : Brun, 2011. _____ 13
Figure 5 : Valeurs médianes des estimations de recrutement (R^{obs}) en fonction du stock (S^{obs}). Le stock est exprimé en œufs·m^{-2} et le recrutement en tacons^{0+}·m^{-2}. _____ 19
Figure 6 : Résidus standardisés (ϵ) en fonction de l'année pour différentes combinaisons de priors et courbe SR. Les points rouges correspondent aux priors uniformes et les points bleus correspondent aux priors très informatifs. _____ 20
Figure 7 : Distributions marginales a priori en rouge et a posteriori en bleu du taux de survie médian maximal des recrues (α) en échelle naturelle en fonction des différentes combinaisons de prior. _____ 21
Figure 8 : Distributions marginales a priori en rouge et a posteriori en bleu du recrutement médian maximal (κ) en échelle logarithmique (log) en fonction des différentes combinaisons de priors. _____ 21
Figure 9 : Distributions de probabilité a posteriori du coefficient de d'autocorrélation des résidus iid standardisés, ϵ (eq. 14) pour les hypothèses sans autocorrélation et δ (eq. 15) pour les hypothèses avec autocorrélation. Les points et les lignes matérialisent respectivement la médiane et l'intervalle à 95% de probabilité des distributions. _____ 23
Figure 10 : WAIC en fonction des différentes combinaisons d'hypothèses et courbes SR. _____ 24
Figure 11 : Distributions marginales a priori en rouge et a posteriori en bleu du taux de survie des recrues médian maximal (α) en échelle naturelle. Les points et les lignes matérialisent respectivement la médiane et l'intervalle à 95% de probabilité des distributions. _____ 25
Figure 12 : Distributions marginales a posteriori en rouge et a priori en bleu du recrutement médian maximal (κ) en échelle log. Les points et les lignes matérialisent respectivement la médiane et l'intervalle à 95% de probabilité des distributions. _____ 26
Figure 13 : Distributions marginales a priori en rouge et a posteriori en bleu du paramètre lié au double régime de recrutement λ en échelle log. Le premier prior correspond à l'hypothèse déterministe (det), le second à l'hypothèse aléatoire (al). Les points et les lignes verticales matérialisent respectivement la médiane et l'intervalle à 95% de probabilité des distributions. _____ 27
Figure 14 : Résidus standardisés avec l'hypothèse al_Ac_0.95 et la fonction de Ricker, décomposés par étape. Régime moyen = résidus standardisés mais seulement par rapport à la courbe moyenne de recrutement. Régime al = prise en compte du double régime aléatoire mais pas de l'autocorrélation. Régime al_Ac = Résidus δ de l'hypothèse al_Ac_0.95 qui prennent en compte l'autocorrélation et le double régime de recrutement. _____ 28
Figure 15 : Distributions de probabilité a posteriori du stock (en rouge) et du recrutement (en bleu). Les lignes et les points correspondent respectivement à l'intervalle à 95% des distributions et aux médianes. _____ 29
Figure 16 : Prédictions du recrutement (R^{pred}) en fonction du stock (S^{pred}). Les lignes pleines rouges et bleues correspondent respectivement aux médianes des distributions du recrutement prédit avec EM et sans EM. __ 30
Figure 17 : Diagrammes de risque qui représentent l'évolution du risque de faible recrutement (%) en fonction d'une gamme de stock donnée. Les lignes bleues correspondent aux hypothèses sans erreur de mesure, les lignes rouges correspondent aux hypothèses avec erreurs de mesure. _____ 31
Figure 18 : Analyse rétrospective des estimations du stock de saumon sur la série temporelle (1985-2021) en fonction des LC's obtenues selon l'option de modélisation et de prior retenu. _____ 31

Table des tableaux

Tableau 1 : Gamme de stock (S^{pred}) utilisé pour les prédictions correspondant aux quantiles à {0 ; 0,25 ; 0,50 ; 0,75 ; 1} des médianes a posteriori des stock observés (S^{obs}). Le pas choisi entre chaque intervalle est croissant. _____ 15
Tableau 2 Espérances et écart-types des différents priors pour α et κ selon la probabilité cumulée associée. __ 16
Tableau 3 : Loi de probabilité choisi a priori (prior) pour les différents paramètres (θ). _____ 16
Tableau 4 : Abréviations des hypothèses sur l'erreur de processus. _____ 17
Tableau 5 : Abréviations correspondant aux priors sur α et κ. _____ 17

Introduction

Le saumon Atlantique *Salmo salar* (Linnaeus, 1758) est une espèce de poisson appartenant à la famille des Salmonidés. C'est une espèce anadrome qui alterne une phase de son cycle de vie en mer et en rivière. Elle est répartie dans les zones tempérées et froides de l'océan Atlantique. Les adultes ont la particularité d'avoir une forte philopatrie, c'est-à-dire qu'ils remontent majoritairement dans leur cours d'eau natal pour se reproduire. Les juvéniles (tacons) restent en rivière quelques années (1 à 2 ans en France), avant de dévaler lors de la smoltification afin de réaliser l'essentiel de leur croissance en mer (1 à 3 ans) à l'ouest du Groenland, dans la mer du labrador ou dans la mer de Norvège (Olmos et al., 2019 ; Thorstad et al., 2010). Depuis plusieurs décennies la conservation du saumon atlantique est fortement menacée dans plusieurs zones de son aire de répartition. Entre 1985 et 2016, le nombre de saumons avant toute exploitation a diminué de plus de moitié (NASCO, 2019). En 2019, parmi les 2539 rivières dans la zone Atlantique Nord qui ont été l'objet d'une évaluation de l'état des populations de saumon , 43 % sont considérées en danger ou menacées d'extinction (NASCO, 2019). En France on observe une grande diminution du nombre de cours d'eau colonisés par cette espèce depuis plus deux siècles (Thibault, 1994).

Aujourd'hui, à l'échelle internationale l'Organisation pour la Conservation du Saumon de l'Atlantique Nord (OCSAN) régit les grands principes de conservation à appliquer pour les différents stocks de saumons. L'OCSAN recommande que pour chaque rivière soient définis des points de référence biologique (NASCO, 2009). Les points de référence correspondent à des repères pour évaluer l'état de conservation des populations. Pour le saumon un point de référence important est la limite de conservation (LC). Elle est définie comme la quantité minimale de reproducteurs permettant d'assurer la conservation de la population (Lebot, 2021). L'OCSAN recommande de définir la LC comme étant le "nombre de reproducteurs permettant de maximiser les captures en moyenne à long terme". Cette définition comporte une ambiguïté fondamentale, en mélangeant la conservation et l'exploitation. Au Canada, une nouvelle définition de LC est appliquée (Chaput et al., 2015) depuis 2015. Plus récemment en Bretagne, (Lebot, 2021), a repris cette nouvelle définition de LC. Elle considère que la limite de conservation est le « nombre de reproducteurs permettant de contrôler le risque de faible recrutement ». Les notions de risque et de faible recrutement sont à définir selon le cas étudié.

La détermination de cette nouvelle LC repose sur la relation stock-recrutement (SR) qui décrit le lien entre l'abondance des géniteurs (stock, S) et l'abondance des juvéniles (recrutement, R). C'est à l'aide de cette relation que l'on va pouvoir d'une part définir le seuil de faible recrutement et d'autre part prédire les distributions des valeurs possibles de recrutement pour une gamme de stock donnée. La comparaison de ces prédictions et du seuil de faible recrutement permet d'obtenir la probabilité que le recrutement prédit tombe sous le seuil de faible recrutement en fonction du stock. Cette probabilité est le risque de faible recrutement. Son évolution en fonction du stock est exprimée à l'aide d'un diagramme de risque qui permet de déterminer la LC en fonction du niveau de risque choisi et pour un seuil de faible recrutement donné. Le calcul de ces nouvelles limites de conservation est donc fortement lié à la manière dont on prédit la distribution des recrutements possibles en fonction du stock. Ces prédictions sont réalisées au moyen d'un modèle SR et dans un cadre statistique bayésien (Lebot 2021). Ainsi, la prise en compte des différentes sources d'incertitude qui affectent ces prédictions est primordiale.

La relation SR est source de nombreux débats car le recrutement est très variable pour un niveau de stock donné. En effet, la production et la survie des œufs et des larves sont très

sensibles aux facteurs biotiques (prédation, compétition) et abiotiques (débit, température) (Maunder and Thorson, 2019). Cette relation ne doit donc pas se limiter à une formule mathématique qui décrirait précisément le recrutement en fonction du stock. Elle doit être considérée comme un ensemble de distribution de probabilité du recrutement en fonction du stock (Walters and Hilborn, 1992).

Un modèle SR est l'ensemble de deux composantes. La distribution de probabilité de recrutement pour un stock donné correspond à l'erreur de processus. Classiquement, selon des éléments empiriques et théoriques, on suppose qu'elle suit une loi lognormale (Hennemuth et al., 1980; Peterman, 1981; Shelton, 1992; Walters and Hilborn, 1992). Cependant il est important de vérifier cette hypothèse, car cela peut varier selon la population étudiée (Shelton, 1992). D'autres hypothèses peuvent être faites pour représenter l'erreur de processus, par exemple utiliser une loi gamma plutôt qu'une lognormale, ajouter une distribution de mélange ou de l'autocorrélation (Adkison, 2021; Lebot, 2021; Parent and Rivot, 2012). La courbe SR elle, correspond à la fonction mathématique qui relie les moyennes ou médianes des différentes distributions de recrutement en fonction du stock. Plusieurs fonctions mathématiques sont utilisées et impliquent des hypothèses écologiques différentes au niveau des effets densités-dépendants (Adkison, 2021). Celles de Ricker et Beverton-Holt (Beverton and Holt, 1957; Ricker, 1954) sont les plus employées.

Le modèle SR qui sert à produire les prédictions de recrutement en fonction du stock n'est pas connu exactement, il est seulement estimé à partir des données observées de stock et de recrutement. L'incertitude d'estimation de la relation SR est intégrée dans les prédictions de recrutement. Dans un cadre bayésien, cette erreur aléatoire est influencée par la façon dont on spécifie de façon probabiliste notre connaissance a priori du modèle SR (Gelman et al., 2017; van de Schoot et al., 2021).

Enfin, le stock et le recrutement ne sont généralement pas directement observables. Ils sont estimés à partir de données d'observation avec plus ou moins de précision. L'incertitude liée à ces estimations correspond à l'erreur de mesure. Ignorer l'erreur de mesure peut amener à biaiser et à sous-évaluer l'incertitude de l'estimation de la relation SR, surtout pour les populations où l'on dispose de peu de données (Walters and Hilborn, 1992; Walters and Ludwig, 1981).

Les prédictions de recrutement en fonction du stock qui permettent de définir les nouvelles LC fondées sur la maitrise du risque de faible recrutement dépendent des quatre composantes présentées ci-dessus :

- La courbe SR,
- L'erreur de processus,
- L'erreur d'estimation de la relation SR, elle-même influencée par la spécification de la connaissance a priori du modèle SR,
- L'erreur de mesure du stock et du recrutement.

Leur influence se répercute in fine sur le calcul de la LC de manière plus ou moins sensible. L'objectif général de ce travail est d'analyser la sensibilité des LC's à ces différentes composantes. Ce travail est mené au moyen d'un cas d'étude, la population de saumon atlantique de la Nivelle, fleuve côtier franco-espagnol. Cette population est suivie depuis plus de 35 ans par l'UMR ECOBIOP, ce qui permet d'avoir une quantité de données importante pour analyser la relation SR. Par le biais de cette étude de cas, on vise d'une part à (i) proposer une approche générale pour analyser la sensibilité des LC's aux différentes composantes citées au-dessus, et d'autre part à (ii) établir une nouvelle LC pour la Nivelle et évaluer le statut de conservation de la population au regard de cette dernière.

Afin de répondre à la première partie de cette problématique, plusieurs hypothèses alternatives peuvent être formulées concernant chacune des composantes. Leur combinatoire peut résulter en un nombre très important d'hypothèses croisées, rendant leur évaluation comparative lourde à mettre en œuvre. Une démarche séquentielle a été retenue avec tout d'abord une exploration des différentes formes de courbe SR et spécification probabiliste des connaissances à priori du modèle SR sous l'hypothèse standard d'une distribution indépendante et identiquement distribuée (iid) lognormale du recrutement (i.e. erreur de processus). Ensuite après avoir réalisé une première sélection de configurations de forme de courbe et de spécifications a priori du modèle SR, une exploration des différentes hypothèses de représentation de l'erreur de processus est testée sur ces dernières. Enfin, après une sélection de quelques options à cette seconde étape, l'influence des erreurs de mesure sur les LC's est examinée.

1 Matériel et méthode

Notations

D^S : Données observées pour l'estimation du stock

D^R : Données observées pour l'estimation du recrutement

S^{obs} : Série de stock pour lequel on dispose de données d'observation D^S

R^{obs} : Série de recrutement pour lequel on dispose de données d'observation D^R

θ : Paramètres du modèle stock-recrutement

S^{pred}: Valeur de stock pour lequel on prédit le recrutement R^{pred}

R^{pred} : Recrutement prédit pour la valeur de stock S^{pred}

S_t : Stock de l'année *t*

R_t : Recrutement de l'année *t*

$\mu R(S_t)$: Recrutement médian dépendant du stock de l'année *t*

1.1 Approche générale

1.1.1 Fondement théorique de la prise en compte des différentes sources d'incertitude

La détermination des LCs fondées sur une maitrise du risque de faible recrutement requiert de connaître la distribution des recrutements possibles selon le niveau du stock reproducteur. Dans un raisonnement bayésien il s'agit d'obtenir la distribution probabilité $p(R^{pred}|S^{pred}, D^S, D^R)$ du recrutement (R^{pred}) conditionnellement à une valeur de stock (S^{pred}) et à des données d'observation du stock (D^S) et du recrutement (D^R). Les données D^S et D^R procèdent elles-mêmes des stocks et recrutements S^{obs} R^{obs} qui ont été observés. Le processus de recrutement qui produit les recrutements R^{obs} et R^{pred} est représenté par un modèle SR lui-même contrôlé par un jeu de paramètres θ. La figure 1 représente la structure de conditionnement entre les différentes composantes décrites ci-dessus.

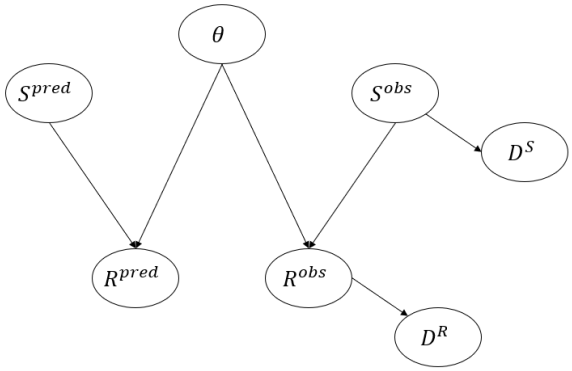

Figure 1 : Directed acyclic graph (DAG) décrivant la structure de dépendance conditionnelle entre les données d'observations de stock (S) et recrutement (R) (D^S, D^R), les séries de S et R observées

(S^{obs}, R^{obs}), les paramètres du modèle SR (θ) et la série de S (S^{pred}) utilisée pour les prédictions du recrutement (R^{pred}).

Compte tenu de la structure de conditionnement décrite dans le DAG (fig. 1), la distribution d'intérêt $p(R^{pred}|S^{pred}, D^S, D^R)$ peut être exprimée comme suit :

$$p(R^{pred}|S^{pred}, D^S, D^R) = \int_{\theta} \int_{S_{obs}} \int_{R_{obs}} p(R^{pred}|S^{pred}, \theta) p(\theta|S^{obs}, R^{obs}) p(S^{obs}|D^S) p(R^{obs}|D^R)\, d\theta\, dS^{obs} dR^{obs}$$

Avec :

$$p(\theta|S^{obs}, R^{obs}) = \frac{p(R^{obs}|S^{obs}, \theta) p(\theta)}{\int_{\theta} p(R^{obs}|S^{obs}, \theta) p(\theta)\, d\theta}$$

Cette formulation fait clairement apparaître que $p(R^{pred}|S^{pred}, D^S, D^R)$ dépend :

- De la formulation du modèle SR $p(R|S, \theta)$ (décliné ici en $p(R^{pred}|S^{pred}, \theta)$ ou $p(R^{obs}|S^{obs}, \theta)$ selon que l'on soit en situation de prédiction ou que l'on dispose de données rétrospectives d'observation), qui a lui-même deux composantes(fig. 2) :
  - La courbe SR qui décrit l'évolution du recrutement moyen ou médian en fonction du stock ;
  - L'erreur de processus qui représente la variabilité aléatoire du recrutement pour un niveau de stock et donc de recrutement moyen/médian donné.
- De la spécification probabiliste $p(\theta)$ de nos connaissances a priori concernant les paramètres θ du modèle SR.
- Des erreurs de mesures (fig. 2) $p(S^{obs}|D^S)$ et $p(R^{obs}|D^R)$ afférentes à l'estimation des séries de stock (S^{obs}) et de recrutement (R^{obs}) à partir des données observées D^S et D^R.

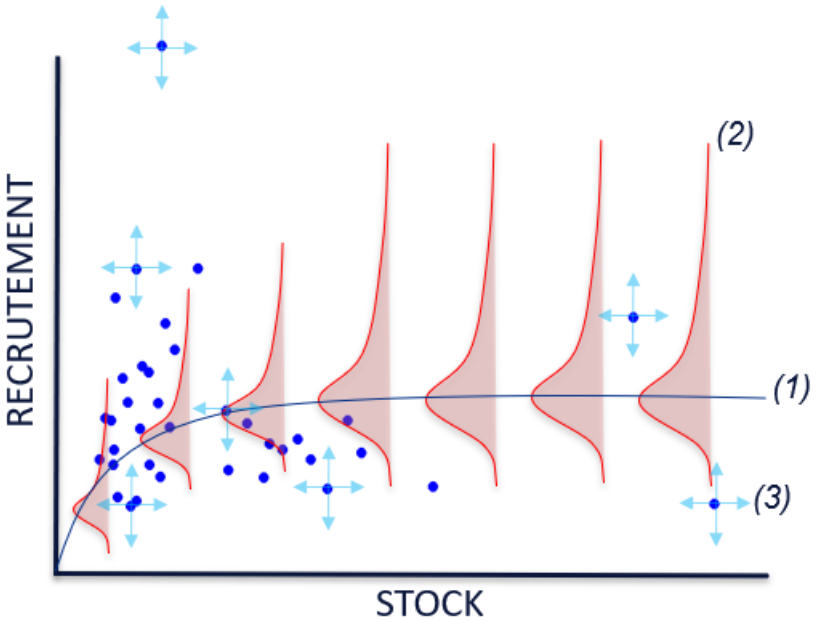

Figure 2 : Exemple d'une relation SR. Les points bleus correspondent aux valeurs médianes des séries (S^{obs}, R^{obs}). La courbe SR est en bleu foncé. Les distributions de probabilité rouges correspondent à l'erreur de processus. Les flèches bleu clair représentent les erreurs de mesures du stock et du recrutement.

1.1.2 Méthode numérique pour la prise en compte des différentes sources d'incertitude

Outre l'identification des différentes composantes d'incertitude à prendre en compte pour produire la distribution d'intérêt ultime $p\left(R^{pred}\middle|S^{pred}, D^S, D^R\right)$, l'intégrale associée permet aussi de concevoir une approche numérique par échantillonnage séquentiel pour approximer $p\left(R^{pred}\middle|S^{pred}, D^S, D^R\right)$.

Le recrutement R^{obs} et le stock S^{obs} sont estimés et l'incertitude associée (i.e. erreurs de mesure) est décrite par les distributions de probabilité a posteriori (dpp) $p\left(R^{obs}\middle|D^R\right)$ et $p\left(S^{obs}\middle|D^S\right)$. Ces erreurs de mesures sont spécifiques à chaque protocole d'observation du stock (e.g. piégeage avec marquage-recapture) et du recrutement (e.g. indices d'abondance sur un réseau de sites par pêche électrique). Pour les prendre en compte, on peut s'inspirer de la méthode proposée par (Rivot et al., 2001) qui consiste à générer différents jeux de données SR possibles par tirage aléatoire dans leurs distributions d'estimation a posteriori.

In fine, pour obtenir les distributions prédictives a posteriori de recrutement d'intérêt ultime $p\left(R^{pred}\middle|S^{pred}, D^S, D^R\right)$, intégrées sur les erreurs de mesures, l'approche séquentielle suivante peut être utilisée :

1. Tirer dans les distributions de probabilité a posteriori $p\left(R^{obs}\middle|D^R\right)$ et $p\left(S^{obs}\middle|D^S\right)$ des séries de stock et de recrutement possibles.
2. Apparier ces séries pour constituer un échantillon de jeux de données (S^{obs}, R^{obs}) possibles.
3. Analyser chacun de ces jeux de données séparément au moyen des différentes options de modèle SR étudiés pour en déduire des dpp $p(\theta|S^{obs}, R^{obs})$.
4. Agréger (i.e. moyenner) les différentes distributions a posteriori $p(\theta|S^{obs}, R^{obs})$ ainsi obtenues pour en déduire :

$$p(\theta|D^S, D^R) = \int_{S_{obs}} \int_{R_{obs}} p(\theta|S^{obs}, R^{obs}) p(S^{obs}|D^S) p(R^{obs}|D^R)\, dS^{obs} dR^{obs}$$

5. Utiliser la distribution a posteriori $p(\theta|D^S, D^R)$ ainsi obtenue pour générer les prédictions de recrutement a posteriori d'intérêt ultime $p\left(R^{pred}\middle|S^{pred}, D^S, D^R\right)$.

Dans cette séquence, les dpp $p\left(\theta\middle|S^{obs}, R^{obs}\right)$ (étape 3) et $p\left(R^{pred}\middle|S^{pred}, D^S, D^R\right)$ (étape 5) peuvent être approximées par échantillonnage. Des techniques MCMC sont employées pour les dpp $p\left(\theta\middle|S^{obs}, R^{obs}\right)$ relatives à chaque jeu de donnée SR possible issus des étapes 1 et 2. Après agrégation des échantillons ainsi obtenus (étape 4), un tirage de Monte-Carlo simple dans $p\left(R^{pred}\middle|S^{pred}, \theta\right)$ (i.e. le modèle probabiliste de la relation SR) pour chaque valeur de θ permet d'obtenir un échantillon de valeurs de R^{pred} relatives à la dpp d'intérêt ultime $p\left(R^{pred}\middle|S^{pred}, D^S, D^R\right)$.

On notera également que, dans le cas où les séries de stock et de recrutement observées (S^{obs}, R^{obs}) sont supposées connues sans erreur, les seules étapes 3 et 5 permettent d'obtenir la dpp prédictive $p(R^{pred}|S^{pred}, S^{obs}, R^{obs})$.

1.1.3 Diagramme de risque et LC :

Le modèle stock-recrutement SR ajusté avec les estimations de stock (S^{obs}) et de recrutement (R^{obs}) observées permet de prédire le recrutement (R^{pred}) en fonction du stock (S^{pred}) en prenant en compte les différentes sources d'incertitudes associées. En reprenant l'approche proposée par (Lebot, 2021), il est possible de déterminer la probabilité que les prédictions du recrutement soient inférieures à un seuil de faible recrutement (figure 3). Le diagramme de risque représente l'évolution de ces probabilités en fonction du stock. La LC correspond à la valeur du stock associée à un risque choisi préalablement que le recrutement prédit soit sous le seuil de faible recrutement (fig. 3).

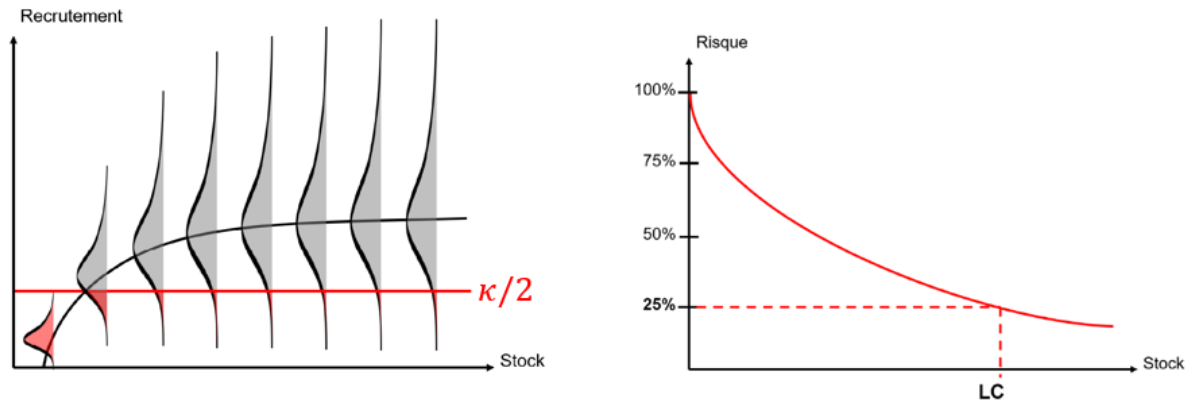

Figure 3 : À gauche : Comparaison des prédictions de recrutement et du seuil de faible recrutement correspondant à 50 % de la capacité d'accueil de la population (ligne rouge). Les portions des distributions de recrutement inférieures au seuil de faible recrutement sont figurées en rouge. Elles correspondent au risque de faible recrutement. À droite : Diagramme de risque qui présente l'évolution du risque de faible recrutement en fonction du stock et permet de déterminer la LC selon le niveau de risque choisi (ici 25%). Source : (Lebot, 2021).

Un des paramètres clé des modèles SR est le recrutement médian maximal (κ) ; i.e. la capacité d'accueil de notre population. Ce paramètre peut servir de référence pour définir la notion de faible recrutement, avec des seuils définis comme une proportion de κ. Pour les populations de saumon de Bretagne et du Canada Atlantique (hors Québec) la LC correspond à un risque de faible recrutement de 25% pour un seuil de faible recrutement égal à 50% de la capacité d'accueil κ (i.e $\kappa/2$).

1.1.4 Hypothèses alternatives de modélisation de la relation SR

Les hypothèses liées au modèle SR peuvent être séparées en deux composantes : d'une part la manière dont on représente l'erreur de processus et d'autre part les différentes fonctions choisies pour la courbe SR.

1.1.4.1 Courbes SR

Trois fonctions ont été retenues pour décrire le recrutement médian $\mu R(S_t)$. La fonction de Beverton-Holt (BH) (1) est la plus couramment utilisée pour le saumon atlantique (Adkison, 2021; Michielsens and McAllister, 2004; Walters and Hilborn, 1992; Beverton and Holt, 1957).

$$\mu R(S_t) = \frac{S_t}{\frac{1}{\boldsymbol{\alpha}} + \frac{S_t}{\boldsymbol{\kappa}}} \qquad (1)$$

Les deux paramètres de cette fonction sont la pente à l'origine α et l'asymptote de la courbe κ. Ils correspondent respectivement au taux de survie médian maximal des recrues et à l'abondance médiane maximale du recrutement. La fonction de BH suppose que l'abondance médiane des recrues tend de manière asymptotique vers κ quand le stock augmente.

La deuxième fonction utilisée pour décrire le recrutement médian, couramment utilisée pour décrire les relations stock-recrutement de plusieurs espèces marines, est celle de Ricker (2) (Adkison, 2021; Pulkkinen and Mäntyniemi, 2013; Walters and Hilborn, 1992; Ricker, 1954).

$$\mu R(S_t) = \boldsymbol{\alpha} S_t e^{-\beta S_t} \quad (2)$$

Avec :

$$\boldsymbol{\kappa} = \frac{\alpha}{\beta} e^{-1}$$

Comme pour la fonction de BH, α est le taux de survie médian maximal des recrues et κ le recrutement médian maximal. La spécificité de la courbe de Ricker est qu'elle va atteindre son maximum κ avant que le recrutement médian ne diminue avec l'augmentation du stock.

La troisième courbe SR est la fonction de Deriso-Schnute (DS) (3) (Schnute, 1985). Elle est composée de trois paramètres, les deux paramètres clés $\alpha\,;\kappa$ et un paramètre de forme γ.

$$\mu R(S_t) = \alpha S_t (1 - \beta\gamma S_t)^{1/\gamma} \quad (3)$$

Avec :

$$\boldsymbol{\kappa} = \alpha\beta^{-1}(1+\gamma)^{-(1+\gamma)/\gamma}$$

Le paramètre de forme est compris entre -1 et 0, il offre tout une gamme de possibilités entre une fonction de Ricker ($\gamma = 0$) ou une fonction de BH ($\gamma = -1$).

Les différentes courbes SR choisies impliquent différentes hypothèses concernant le taux de mortalité instantané au cours du processus de recrutement des juvéniles (recrues). Il est proportionnel à l'abondance des géniteurs (Ricker) ou à l'abondance des recrues (BH). Un taux de mortalité qui est proportionnel à l'abondance des recrues induit que l'effet densité-dépendance sera moins marqué pour BH que pour la fonction de Ricker car il s'atténue au fil du temps au cours du processus de recrutement (Adkison, 2021; Walters and Hilborn, 1992).

1.1.4.2 Erreur de processus

1.1.4.2.1 <u>IID selon une loi log-normale</u>

L'hypothèse la plus couramment utilisée (Adkison, 2021; Shelton, 1992; Peterman, 1981) est celle d'un recrutement indépendamment et identiquement distribué selon une loi log-normale (LN) de médiane $\mu R(S_t)$ et d'écart type σ (en échelle logarithmique).

$$R_t \sim LN(\log(\mu R(S_t)), \sigma) \quad (4)$$

1.1.4.2.2 <u>IID selon un mélange de deux lois log-normales</u>

Une distribution de mélange peut être employée pour modéliser un double régime dans l'environnement ou des évènements qui auraient un impact de manière ponctuelle sur le recrutement le long d'une série temporelle (Adkison, 2021; Lebot, 2021; Maunder and

Thorson, 2019). Deux niveaux de recrutement, haut et bas, sont alors distingués. Un paramètre λ strictement positif est combiné au recrutement médian $\mu R(S_t)$ (5). Ce paramètre est activé à l'aide d'une indicatrice I_t qui est égale soit à 1 soit à 0.

$$R_t \sim LN(\log(\lambda^{I_t}[\mu R(S_t)/\big(1+p(\lambda-1)\big)]), \sigma) \ (5)$$

L'indicatrice I_t est tirée dans une loi de Bernoulli de probabilité p

$$I_t \sim Bern(p) \ (6)$$

En restreignant λ dans l'intervalle [0 ; 1], on obtient le régime bas du recrutement quand I_t = 1 et le régime haut quand I_t = 0. La probabilité p correspond alors à la fréquence moyenne d'occurrence du régime bas ($1-p$ pour le régime haut).

Une version dégradée de ce modèle peut être obtenue quand l'indicatrice I_t est connue (i.e. observée).

1.1.4.2.3 <u>IID avec une variance dépendante du stock</u>

D'après certaines études (Pulkkinen and Mäntyniemi, 2013) l'écart-type du recrutement σ, peut varier en fonction de la taille du stock.

$$\sigma(S_t) = e^a {S_t}^b \ (7)$$

Selon cette formulation, a et b sont les paramètres de la relation linéaire entre les logarithmes de l'écart-type et du stock.

1.1.4.2.4 <u>Autocorrélation d'ordre 1</u>

Une partie importante de la variabilité du recrutement peut être due aux fluctuations environnementales. Or ces dernières sont souvent auto-corrélées dans le temps (Adkison, 2021; Maunder and Thorson, 2019; Subbey et al., 2014). Le recrutement médian de l'année t est alors lié au recrutement médian de l'année $t-1$. L'autocorrélation est plus ou moins forte selon la valeur d'un coefficient ρ compris entre -1 et 1.

Pour $t=1$, le recrutement est tiré dans la loi LN de médiane $\mu R(S_t)$ et d'écart type σ (8). Quand $t>1$ le recrutement est tiré dans la loi LN dont la moyenne dépend du recrutement de l'année $t-1$ et d'écart-type σ' (9).

$$R_1 \sim LN(\log(\mu R(S_1)), \sigma) \ (8)$$

$$\textbf{Pour t } > \mathbf{1}, \ R_t \sim LN(\log(\mu R(S_t)(R_{t-1}/\mu R(S_{t-1}))^{\rho}), \sigma') \ (9)$$

$$\text{Avec } \sigma' = \sqrt{1-\rho^2} \times \sigma$$

Dans cette formulation, l'autocorrélation est supposée être seulement dépendante de l'année précédente c'est-à-dire d'ordre 1.

1.1.4.3 Combinaisons d'hypothèses

Des combinaisons des quatre hypothèses précédemment décrites peuvent être envisagées. Par exemple avoir une distribution de mélange et de l'autocorrélation d'ordre 1 ou avec un écart type dépendant du stock. Chaque combinaison est envisageable mais cela peut amener rapidement à un nombre de paramètres relativement important, d'où l'importance de vérifier ensuite la pertinence des combinatoires.

1.1.5 Choix des priors

Une étape cruciale de l'analyse bayésienne est le choix des priors, c'est-à-dire la spécification probabiliste $p(\theta)$ de la connaissance a priori des paramètres θ du modèle (van de Schoot et al., 2021; Van Dongen, 2006). Ici, deux paramètres clés des différents modèles, en particulier dans une perspective d'évaluation du risque de faible recrutement, sont le taux de survie médian maximal des recrues (α) et l'abondance médiane maximale des recrues (κ). Une attention plus particulière doit être portée à la spécification de différents niveaux d'informations a priori relatifs à ces deux paramètres. D'un point de vue théorique, la spécification de priors devrait être conduite indépendamment des données d'observation. Plusieurs techniques existent pour choisir ces priors. Des priors de référence ou peu informatifs sont souvent utilisés quand on ne dispose pas d'information préalable (Millar, 2002). Cependant un prior « trop éloigné » des données, par exemple selon une loi uniforme censée être non-informative, peut s'avérer moins adapté qu'un prior construit en tenant compte des données (Gelman et al., 2017). Quand on dispose d'informations a priori sur certains paramètres du modèle sans lien direct avec les données, l'analyse bayésienne peut permettre de prendre cette information en compte. Le saumon atlantique, et particulièrement la relation SR chez cette espèce, étant très étudiée, il est pertinent d'utiliser l'information acquise préalablement (Pulkkinen and Mäntyniemi, 2013).

Deux options principales sont utilisées pour construire des priors informatifs. Il est possible de s'inspirer des travaux similaires antérieurs, via en particulier des méta-analyses qui permettent de synthétiser l'information de plusieurs études (Avci, 2017; Pulkkinen and Mäntyniemi, 2013). L'information peut aussi être récupérée auprès d'expert ayant une expérience dans le domaine d'étude (Wongnak et al., 2022; Kuhnert et al., 2010; O'Hagan et al., 2006). On peut pour cela prendre un intervalle de valeurs possible déterminé « à dire d'expert » auquel on associe un certain niveau de probabilité cumulée pour restreindre plus ou moins le prior dans cet intervalle.

Différents niveaux d'informations peuvent ainsi être testés pour le couple de paramètres ($\alpha; \kappa$), allant de non-informatif (loi uniforme) à très informatif. À partir d'un intervalle de valeurs plausibles préalablement choisi à dire d'expert, des priors plus ou moins informatifs peuvent être construits selon des lois normales, en échelle logit pour α (grandeur comprise entre 0 et 1) et en échelle log pour κ (grandeur strictement positive), en fonction de la probabilité cumulée associée a priori à cet intervalle. La moyenne de ces lois normales correspond au centre de l'intervalle choisi $\frac{x_{sup}+x_{inf}}{2}$, x_{sup} et x_{inf} étant les bornes supérieure et inférieure de l'intervalle retenu. La valeur de l'écart-type elle, dépend du niveau d'information, i.e. de probabilité cumulée p que l'on souhaite donner à l'intervalle. Si on considère k_p le quantile correspondant à la probabilité $\frac{1-p}{2}$ pour une loi normale centrée-réduite, l'écart type du prior normal est égal à :

$$\frac{x_{sup} - x_{inf}}{2k_p} \quad (10)$$

Concernant les autres paramètres des modèles, pour lesquels il plus difficile de formaliser un prior informatif à dire d'expert, des priors uniformes peu informatifs peuvent être choisis par défaut.

1.1.6 Approche séquentielle pour l'exploration des options de modélisation de la relation SR

La combinaison des différentes hypothèses qui peuvent être utilisées pour les composantes du modèle de probabilité jointe $p(R^{obs}, \theta | S^{obs}) = p(R^{obs} | S^{obs}, \theta) p(\theta)$, i.e. courbe SR, erreur de processus et connaissances a priori (cf. section 1.1.4, 1.1.5), peut rapidement conduire à examiner un très grand nombre de formulations possibles. Pour réduire le nombre de combinaisons à considérer, une approche séquentielle peut être adoptée en examinant tout d'abord différentes hypothèses combinées de courbes SR et de spécification a priori des paramètres ($p(\theta)$), avec des erreurs de processus iid selon l'hypothèse standard d'une loi lognormale. Sur une sélection d'un sous-ensemble d'hypothèses combinées issues de cette première étape, différentes formulation possibles de l'erreur de process peuvent ensuite être examinées. Ces deux étapes sont conduites en ignorant les erreurs de mesure $p(S^{obs} | D^S)$ et $p(R^{obs} | D^R)$, en considérant que les recrutements observés sont connus ($p(S^{obs} | D^S)$ = $p(R^{obs} | D^R)$ = 1) et égaux aux médianes de leur distribution a posteriori $p(S^{obs} | D^S)$ et $p(R^{obs} | D^R)$. Sur une sélection de combinaisons issues de la deuxième étape, l'effet de la prise en compte des erreurs de mesure est finalement évalué.

Pour mettre en œuvre cette approche séquentielle de sélection d'hypothèses de modélisation, différentes approches et critères sont successivement considérés. Ils sont présentés à suivre dans un ordre qui reflète leur pertinence pour comparer puis sélectionner les modèles à retenir.

1.1.6.1 Vérification à posteriori

De manière générale la vérification a posteriori vise à tester l'adéquation entre le modèle et les données d'observation qui ont servies à sa mise à jour a posteriori. Elle peut procéder à une comparaison entre les données d'observation et des données répliquées sous le modèle a posteriori ou bien d'une analyse du respect a posteriori de certaines hypothèses fondatrices du modèle.

1.1.6.1.1 Approche sur données répliquées

La distribution à posteriori des données répliquées D^{rep}, conditionnellement à la série de données observées D, se traduit par (Gelman et al., 2003) :

$$p(D^{rep}|D) = \int_{\theta} p(D^{rep|}|\theta) p(\theta|D)\, d\theta \quad (11)$$

La distribution a posteriori des valeurs répliquées résulte donc du modèle probabiliste $p(D^{rep|}|\theta)$ supposé être à l'origine des données intégrées sur les valeurs possibles a posteriori des paramètres θ. La comparaison des données répliquées vs observées est faite à l'aide de statistiques tests T calculés avec les données observées $T(D)$ et répliquées $T(D^{rep})$. Une p-value (p_c) peut être associée à cette comparaison (Gelman et al., 2003) :

$$p_c = \Pr(T(D^{rep}) \geq T(D)|D) \quad (12)$$

Cette p-value permet de résumer la différence entre les données répliquées et observées selon la statistique choisie. Deux statistiques sont retenues, celle classique du khi-deux (χ^2) et une plus spécifique à notre étude relative à $\kappa/2$.

La statistique du χ^2 mesure l'ajustement du modèle aux données. Elle est égale à la somme des résidus standardisés au carré.

$$\chi^2 = \sum_{t=1}^{t} \left(\frac{\log(R_t) - \mu R(S_t))}{\sigma} \right)^2 \quad (13)$$

Dans une perspective de quantification du risque de faible recrutement défini relativement à une valeur seuil égale à la moitié du recrutement médian maximal κ, la capacité du modèle à évaluer correctement la proportion de R_t inférieur $\kappa/2$ est un critère important. Cette proportion est comparée entre les recrutements observés R_t^{obs}et des recrutement répliqués R_t^{rep}pour les mêmes valeurs de stock observées S_t^{obs}.

1.1.6.1.2 Vérification des hypothèses du modèle : analyse des résidus standardisés

Quel que soit le modèle utilisé, les résidus standardisés sont supposés iid selon une loi normale centrée-réduite. Ils sont donc théoriquement, de moyenne nulle, réparties équitablement de part et d'autre de 0, essentiellement compris entre -1.96 et 1.96 et indépendants séquentiellement. La vérification de ces attendus peut être testée par une analyse graphique des résidus et le calcul de statistiques telles que la moyenne empirique des résidus ou leur coefficient d'autocorrélation d'ordre 1. Sauf sous l'hypothèse d'autocorrélation du processus de recrutement, les résidus iid standardisés du modèle sont :

$$\epsilon_t = \frac{\log(R_t) - \log(\mu R(S_t))}{\sigma} \quad (14)$$

Pour les modèles avec hypothèse d'autocorrélation les résidus iid sont :

$$\delta_t = \frac{\log(R_t) - (\log(\mu R(S_t)) + \rho(R_{t-1} - \log(\mu R(S_{t-1}))))}{\sigma'} \quad (15)$$

1.1.6.1.3 Vérification a posteriori des effets

Certaines grandeurs introduites dans les modèles correspondent à des effets attendus de covariables observées ou de structuration interne du modèle. L'introduction d'un effet dans un modèle est associée à une hypothèse. La vérification de cette dernière conditionnellement aux données d'observations peut être jugée par comparaison de la distribution a posteriori de l'effet avec une valeur de référence. C'est le cas par exemple du paramètre *b* dans la formulation d'un écart-type du recrutement dépendant du stock (eq. 7) ou du coefficient *ρ* dans un modèle autocorrélé (eq. 9). Dans ces deux exemples, il est attendu a posteriori une probabilité élevée d'être soit strictement positif, soit strictement négatif. À défaut, on considérera que l'hypothèse de l'effet n'est pas vérifiée et on privilégiera une formulation de modèle plus parcimonieuse sans l'effet considéré.

1.1.6.2 Critères dits « d'information »

Les critères dits « d'information » permettent d'ordonner un ensemble de modèles selon leurs performances en termes de précision prédictive. Gelman et al. (2014) recommande l'usage du critère de Watanabe-Akaike (WAIC ; Watanabe, 2010). Il est important de subordonner l'emploi de ces critères à une étape préalable de vérification a posteriori pour éviter de comparer des modèles tous inadéquats au regard des données d'observation disponibles.

1.1.6.3 Identification et mise à jour a posteriori des modèles

L'actualisation a posteriori des paramètres (θ) peut être évaluée en comparant leur distribution a posteriori $p(\theta|S^{obs}, R^{obs})$ avec leur distribution a priori $p(\theta)$. Ceci permet d'apprécier si les données sont suffisamment informatives pour permettre une mise à jour des paramètres. À défaut, les inférences a posteriori reflètent la spécification a priori et dépendent

donc directement de la pertinence de cette dernière au regard des connaissances disponibles. Elle doit alors faire l'objet d'une attention toute particulière.

En complément, la corrélation entre les paramètres d'un modèle dans leur distribution jointe a posteriori est un indicateur de confusion entre paramètres qui témoigne d'une possible sur-paramétrisation. Il est important d'y prêter attention lorsque l'on complexifie les modèles. Dans un souci de parcimonie, on privilégiera des formulations limitant la confusion entre les paramètres.

À performances en matière de vérification a posteriori et de critères d'information égales, on préférera les modèles qui permettent une mise à jour et une identification a posteriori des paramètres.

1.2 Cas d'étude

1.2.1 Site d'étude

La Nivelle est un fleuve côtier du pays-basque, dans le sud de la région Nouvelle-Aquitaine (France). Long d'une quarantaine de kilomètres, il prend sa source en Espagne et a son embouchure en France à Saint-Jean de Luz. Elle a deux affluents principaux, le Lapitxuri et le Lurgorrieta. Plusieurs barrages ont été construits sur la Nivelle, divisant ce cours d'eau en différentes zones (fig. 4). Celui d'Uxondoa est situé en Basse-nivelle et celui d'Olha correspond à la limite entre la Basse et Haute Nivelle. Avant 1992 la partie à l'amont du barrage d'Olha était inaccessible aux géniteurs. Suite à la construction d'une passe à poissons, les zones en amont ont pu être colonisées jusqu'aux barrages Arrieta sur le Lurgorrieta et Urrutinea sur le cours principal de la Nivelle (Royer and Prévost, 2016; Brun, 2011).

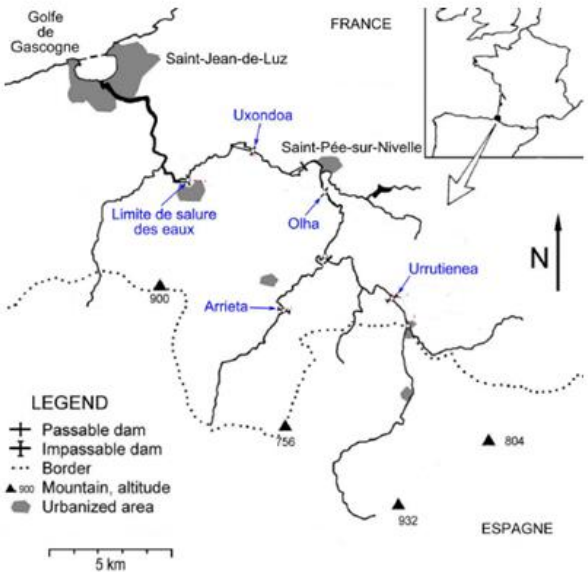

Figure 4 : Carte de la rivière de la Nivelle (Pays-basque, France). Source : Brun, 2011.

1.2.2 Les données d'observations ($D^S D^R$)

La population de saumon atlantique de la Nivelle est suivie depuis plus de 30 ans par l'UMR INRAE/UPPA ECOBIOP. Plusieurs protocoles ont été établis afin d'obtenir des données

d'observation des géniteurs (D^S) depuis 1984 et des juvéniles de l'année (tacons 0+) (D^R) depuis 1985.

Deux sources d'informations principales permettent d'obtenir les données d'observation de géniteurs de saumons D^S (Royer and Prévost, 2016) :

- Opérations de capture-marquage-recapture (CMR) réalisées principalement au niveau des deux stations de capture d'Uxondoa » (piège partiel) et d'Olha (piège total).
- Comptage de frayères sur l'ensemble de la zone colonisée par les géniteurs.

Les données d'observation D^R concernant les abondances de tacons 0+ sont acquises à l'aide de protocole de pêche électrique (Brun, 2011). L'échantillonnage est réalisé chaque automne depuis 1985. Pour cela la rivière est divisée en trois types d'habitats (i.e. faciès d'écoulement ; Malavoi and Souchon, 2002) plus ou moins propices aux juvéniles de saumon : les radiers – rapides, les plats et les profonds. Les surfaces disponibles par type d'habitat ont été mesurées. Deux protocoles de pêche électrique ont été utilisés (Brun, 2011 pour plus de détails) :

- De 1985 à 2005, les pêches électriques étaient effectuées par enlèvement successifs sur les habitats « plat » et « radiers-rapides », les profonds étant considérés comme non utilisés par les juvéniles de saumon.
- À partir de 2003, la méthode de capture par unité d'effort (CPUE) selon un protocole standardisé (Prévost and Baglinière, 1994) est appliquée seulement sur les habitats les plus propices aux juvéniles de saumon (radiers-rapides).

Les données d'observations (D^S, D^R) proviennent donc de différentes techniques d'échantillonnage au cours du temps et la surface colonisée par les saumons varie aussi au cours du temps. Des modèles hiérarchiques bayésiens ont été conçus pour permettre d'avoir une estimation de l'abondance de géniteurs et de juvéniles dans la Nivelle homogène le long de la série temporelle.

1.2.3 Estimation de l'abondance de géniteurs (stock) et de tacons 0+ (recrutement)

La surface colonisée par le saumon et les zones échantillonnés varient au cours de la série temporelle (1985-2021, année de recrutement ; cf. 1.2.1). L'abondance des tacons 0+ et géniteurs est ainsi standardisée en densité exprimé en m^2 d'équivalent radiers-rapides, rapportée à la surface colonisée par le saumon sur la Nivelle selon l'année.

Les densités de géniteurs sont exprimées en œufs potentiellement pondus par les femelles, qui est le premier facteur contrôlant la production de descendants à la génération suivante. Les valeurs en œufs sont obtenues en utilisant les proportions de femelles et les fécondités de ces dernières dépendant de l'âge de mer et des différences de taille associées. Les fluctuations annuelles de composition des adultes échantillonnées (âge/taille et proportions de femelles) sont ainsi prises en compte dans la quantification du stock reproducteur.

Les estimations de l'abondance de géniteurs et de tacons 0+ sont obtenues à l'aide de modèles hiérarchiques bayésiens décrit par (Brun, 2011). Ils permettent d'obtenir les distributions de probabilité a posteriori de l'abondance du stock $p(S^{obs}|D^S)$ (œufs) et du recrutement $p(R^{obs}|D^R)$ (tacons 0+). Des échantillons de valeurs issues de ces distributions, exprimées en densité par m^2 d'équivalent radiers-rapides, sont disponibles sur le site https://github.com/ORE-DiaPFC/Abondance/tree/master/Nivelle, qui documente en outre à la

fois les données et les modèles d'observations. Quand S^{obs} et R^{obs} sont supposés connus sans erreur de mesure, ce sont les médianes a posteriori qui sont utilisées pour l'analyse de la relation SR.

1.2.4 Diagramme de risque et LC

La probabilité d'intérêt $p\left(R^{pred}\middle|S^{pred}, D^S, D^R\right)$ permet d'obtenir un diagramme de risque et in fine de déterminer la LC. Les prédictions du recrutement sont produites pour une gamme de stock S^{pred} correspondant, allant de 0 à 42 œufs/m² (tableau 1). Le pas choisi entre chaque valeur de stock est croissant par intervalle, le risque de faible recrutement variant de moins en moins rapidement quand le stock augmente. Les limites des intervalles des prédictions du stock S^{pred} correspondent aux quantiles à {0 ; 0,25 ; 0,50 ; 0,75 ; 1} des médianes a posteriori des stocks observés (S^{obs})

Tableau 1 : Gamme de stock (S^{pred}) utilisé pour les prédictions correspondant aux quantiles à {0 ; 0,25 ; 0,50 ; 0,75 ; 1} des médianes a posteriori des stock observés (S^{obs}). Le pas choisi entre chaque intervalle est croissant.

Intervalle S^{pred}	**Pas**
[0 ; 5,23]	0,1
[5,24 ;7,16]	0,2
[7,19 ; 10,81]	0,3
[10,84 ; 42]	0,5

Le seuil de faible recrutement choisi correspond à la moitié de la capacité d'accueil i.e. $\kappa/2$. Ce seuil est le même que celui utilisé au Canada et en Bretagne respectivement par (Lebot, 2021; Chaput et al., 2015). Le risque choisi que le recrutement prédit soit sous le seuil de faible recrutement est de 25 %, là aussi la même valeur que celle utilisée au Canada et en Bretagne.

1.2.5 Hypothèses alternatives de modélisation de la relation SR

Les différentes formes de courbe SR et de représentation de l'erreur de processus choisies pour l'étude correspondent à celle présentées dans la section 1.1.4. Deux versions de la distribution de mélange sont testées :

- La version générale (eq. 5) quand I_t varie de manière aléatoire le long de la série temporelle.
- Une version dégradée supposant l'indicatrice I_t est connue, avec I_t = 1 avant 1992, et I_t = 0 à partir de 1992. Cette dichotomie correspond à l'ouverture du barrage d'Olha qui a augmenté la surface colonisable par le saumon dans la Nivelle et potentiellement aussi modifié la qualité des habitats support du processus de recrutement, supposés de meilleure qualité en amont.

1.2.6 Choix des priors

La construction des priors sur α et κ suit la démarche présentée dans la section 1.1.5. Bien qu'il n'y ait pas eu d'enquête auprès de différents experts du saumon atlantique, des priors « à dire d'expert » ont été élaborés. L'objectif est avant tout ici d'évaluer la sensibilité des résultats des analyses SR aux choix de priors. L'intervalle plausible de valeurs pour α et κ est respectivement en échelle naturelle [0,01 ; 0,1] et [0,5 ;2]. Quatre niveaux d'informations sont testés sur ces deux paramètres allant de non informatif (loi uniforme) à très informatif. Trois niveaux d'informations sont construits selon des lois normales, en échelle logit pour α et en échelle log pour κ, en fonction de la probabilité cumulée associée a priori à cet intervalle

(tableau 2). Les probabilités cumulées p définissant le niveau d'information du prior sont 0.5,0.8,0.95. L'espérance et l'écart type des différents priors sont résumés dans le tableau 2.

Tableau 2 Espérances et écart-types des différents priors pour α et κ selon la probabilité cumulée associée.

Paramètre	Intervalle	Distribution a priori	Probabilité cumulée(p)	Espérance	Ecart-type	Information
α	[-4,60 ; -2,20] (Echelle logit)	Uniforme*		0,5	0,29	Non informatif
		Normale (En échelle logit)	0.5	-3,40	1,78	Peu informatif
			0.8		0,94	Informatif
			0.95		0,61	Très informatif
κ	[-3,00 ;0,69] (Echelle log)	Uniforme**		5	2,89	Non informatif
		Normale (En échelle log)	0.5	-1,15	1,72	Peu informatif
			0.8		0,90	Informatif
			0.95		0,58	Très informatif

** Loi uniforme sur l'intervalle [0 ; 1] en échelle naturelle. ** Loi uniforme sur l'intervalle [0 ; 10] en échelle naturelle.*

Concernant les autres paramètres des modèles pour lesquels il est plus difficile de formaliser un prior informatif à dire d'expert, des priors uniformes peu informatifs sont choisis par défaut (tableau 3).

Tableau 3 : Loi de probabilité choisi a priori (prior) pour les différents paramètres (θ).

Paramètre	Définition	Prior
γ	Paramètre de forme de la courbe SR avec la fonction de DS (eq. 3)	U(-1,0)
σ	Ecart-type du modèle SR en échelle logarithmique	U(0,10)
λ^{I_t}	Paramètre de la distribution de mélange (eq. 5) (Cas aléatoire) *	U(-10,0)
λ^{I_t}	Paramètre de la distribution de mélange (eq. 5) (Cas déterministe) **	U(-10,10)

p	Probabilité de l'indicatrice I_t d'être activé (eq. 6)	U(0,1)
a	Ordonnée à l'origine de la relation $\sigma(S_t)$ (eq. 7)	U(-10,10)
b	Coefficient directeur de la relation $\sigma(S_t)$ (eq. 7)	U(-10,10)
ρ	Coefficient d'autocorrélation (eq. 9)	U(-1,1)

** Cas aléatoire : λ^{I_t} est activé de manière aléatoire le long de la série temporelle selon la valeur l'indicatrice I_t suivant une loi de Bernoulli $I_t \sim Bern(p)$;*

*** Cas déterministe : λ^{I_t} est activé de manière déterministe selon la valeur de l'indicatrice I_t connue.*

1.3 Abréviations

Dans la suite du rapport des abréviations sont utilisées pour résumer les différentes combinaisons d'hypothèses sur le modèle SR et l'information a priori des paramètres. Elles sont résumées dans les tableaux 4 et 5.

Tableau 4 : Abréviations des hypothèses sur l'erreur de processus.

Hypothèses		Abréviations
Autocorrélation		***Ac***
Distribution de mélange de deux lois lognormale (LN)	Cas aléatoire ($I_t \sim Bern(p)$)	***al***
	Régime déterministe ($I_{1985-1991} = 1$ et $I_{1992-2021} = 0$)	***det***
Dépendance σ(S)		***dS***

Tableau 5 : Abréviations correspondant aux priors sur α et κ.

Paramètre	Probabilité	Abréviations
α	0.5 (Peu informatif)	***α0.5***
	0.8 (informatif)	***α0.8***
	0.95 (très informatif)	***α0.95***
κ	0.5 (Peu informatif)	***κ0.5***
	0.8 (informatif)	***κ0.8***
	0.95 (très informatif)	***κ0.95***

Si le prior sur α ou κ est uniforme il n'y pas d'abréviation particulière. Ainsi, par exemple ***α0.5_ κ0.5*** signifie un prior peu informatif sur α et κ et ***α0.5*** signifie prior peu informatif sur α et non informatif (uniforme) sur κ.

Pour les modèles SR, il n'y a pas d'abréviation pour l'hypothèse standard iid. Ainsi par exemple, ***al_Ac*** correspond à la combinaison de l'hypothèse d'une distribution de mélange de deux lois lognormales avec de l'autocorrélation, alors que ***al*** correspond à l'hypothèse iid avec une distribution de mélange de deux lois log-normales.

1.4 Calcul Bayésien

Pour chaque version de modèle testée, la distribution a posteriori jointe de toutes les grandeurs inconnues a été approximé par échantillonnage MCMC (Monte-Carlo Markov Chain) à l'aide du logiciel NIMBLE avec 3 chaînes MCMC. Chaque chaine est composée de 500 000 itérations après une phase initiale de 5000 itérations (burn-in). Les distributions de probabilité à posteriori (dpp) de chaque paramètre sont issus d'un sous- échantillonnage des chaines MCMC de 30000 itérations, 10000 itérations par chaîne (soit 1 itération sur 50 conservée). La convergence des chaînes MCMC a été vérifiée à l'aide du critère de Gelman et Rubin (< 1,1) et de la taille efficace (> 1000).

2 Résultats

Le jeu de données de la Nivelle montre un large contraste du stock (i.e dépose d'œufs) et aussi une forte variabilité du recrutement (figure 5). Les médianes des distributions de probabilités a posteriori (dpp) du stock sont comprises entre 3,31 œufs·m^{-2} (2008) à 41,55 œufs·m^{-2} (1990). Les R médians sont compris entre 0,041 tacon^{0+}·m^{-2} (2006) et 0,498 tacon^{0+}·m^{-2} (2021).

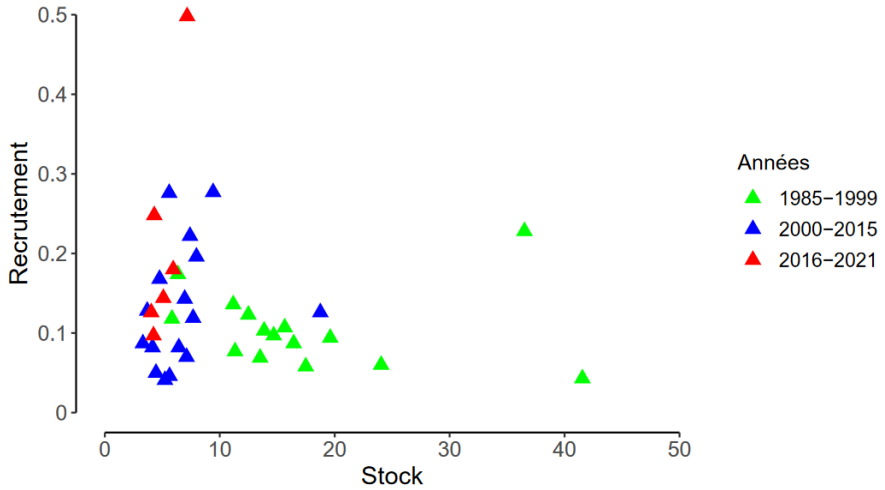

Figure 5 : Valeurs médianes des estimations de recrutement (R^{obs}) en fonction du stock (S^{obs}). Le stock est exprimé en œufs·m^{-2} et le recrutement en tacons^{0+}·m^{-2}.

2.1 Analyse de sensibilité aux priors et courbes SR

Lors de cette première étape nos deux paramètres clés sont le taux de survie maximale médian (α) et l'abondance maximale médiane de recrues (κ).

2.1.1 Vérification a posteriori

2.1.1.1 Comparaison entre les données répliquées et observées

La comparaison entre les données répliquées et observées ne révèle pas d'inadéquation entre le modèle et les données quel que soit le prior sur α et κ et la courbe SR choisis. Les valeurs des p-values sont légèrement supérieures à 0,5 pour la statistique du χ^2 et comprises entre 0,34 et 0,57 pour la statistique $\kappa/2$ (cf. ANNEXE 1.1).

2.1.1.2 Analyse des résidus standardisés

L'analyse graphique des résidus standardisés ϵ ne révèle pas d'influence du choix du prior sur α et κ (figure 6). La forme de courbe SR n'induit pas non plus de différence majeure, même si dans le détail elle semble légèrement influer sur l'ajustement de certains points (e.g. pour l'année 1987 avec la combinaison de prior très informatif, les résidus médians sont de 1,10 avec Beverton-Holt (BH), et de 2,23 avec Ricker).

Figure 6 : Résidus standardisés (ϵ) en fonction de l'année pour différentes combinaisons de priors et courbe SR. Les points rouges correspondent aux priors uniformes et les points bleus correspondent aux priors très informatifs.

L'hypothèse iid ne semble pas respectée : le coefficient d'autocorrélation d'ordre 1 des résidus montre qu'ils sont légèrement autocorrélés surtout pour BH (cf. ANNEXE 1.2). La borne inférieure à 95% du coefficient d'autocorrélation la plus faible est à 0,04 quelle que soit l'option de courbe ou de prior considérée. En valeur médiane, le coefficient d'autocorrélation est toujours supérieur à 0,27 pour la fonction de BH alors qu'il se situe aux alentours de 0,12 pour Ricker. Pour la fonction de Deriso-Schnute (DS) la valeur médiane est entre 0,19 et 0,24 selon les combinaisons de prior.

2.1.2 Critère d'information : WAIC

Les WAIC sont peu différents quelle que soit la combinaison de prior et de forme de courbe considérée (cf. ANNEXE 1.3). Avec la fonction de BH, le WAIC le plus faible (meilleur) est à -93 pour un prior uniforme sur le couple (α, κ). Pour la fonction de Ricker, le WAIC est autour de -88 peu importe la combinaison de priors. Les WAIC pour la fonction de DS sont compris entre -92 et -93 selon les priors choisis. La différence de 5 entre les WAIC les plus faibles et les plus élevés ne permet d'exclure aucun modèle, même si une légère prévalence des modèles avec courbes BH et DS semble se dégager.

2.1.3 Mise à jour et identification des modèles

La sensibilité au prior est testée sur le couple (α, κ). La sensibilité du taux de survie médian maximal des recrues α au prior varie selon la fonction SR choisie (figure 7). En effet pour BH, la dpp de α dépend essentiellement du prior choisi pour ce paramètre (cf. fig. 7). A l'inverse, les dpp de α sont précises et ne varient pas selon le prior dans le cas d'une courbe de Ricker. Concernant la fonction de DS la dpp de α dépend aussi en partie du prior, mais de manière moins évidente qu'avec BH.

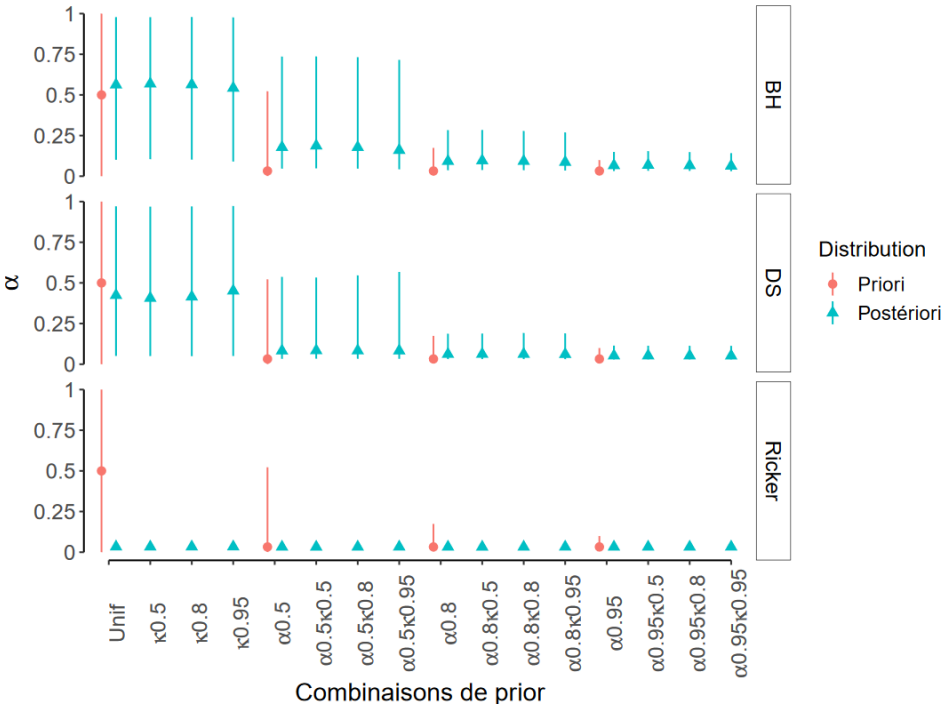

Figure 7 : Distributions marginales a priori en rouge et a posteriori en bleu du taux de survie médian maximal des recrues (α) en échelle naturelle en fonction des différentes combinaisons de prior.

La distribution de probabilité a posteriori du paramètre κ semble peu sensible au prior, quel que soit la forme de courbe SR considérée. L'actualisation a posteriori est toujours forte et peu dépendante de la combinaison de prior, particulièrement avec la fonction de Ricker (fig. 8). Cela se vérifie aussi avec la fonction de Deriso-Schnute.

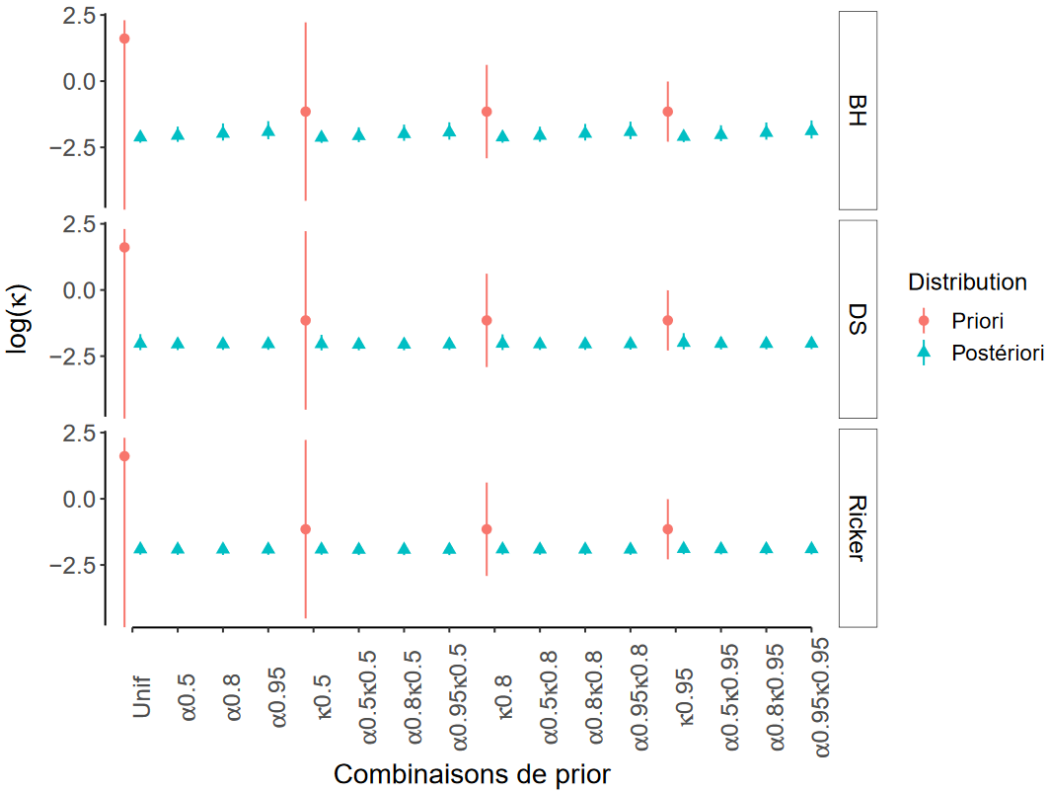

Figure 8 : Distributions marginales a priori en rouge et a posteriori en bleu du recrutement médian maximal (κ) en échelle logarithmique (log) en fonction des différentes combinaisons de priors.

Pour la fonction de BH, le prior très informatif sur α influence légèrement la dpp de κ. En échelle naturelle, l'intervalle à 95 % de κ est égal à [0,10 ; 0,15] avec l'hypothèse de prior uniforme sur le couple $(\alpha;\kappa)$ vs [0,11 ; 0,22] avec l'hypothèse de prior uniforme sur κ et très informatif sur α ($\alpha 0.95$). Ajouter un prior très informatif sur κ ne change pas cette tendance (ANNEXE 1.4).

Dans le cas de la courbe de DS, la dpp du paramètre de forme γ montre une faible mise à jour a posteriori. L'intervalle à 95 % de probabilité est égal à [-0.98 ; -0.37] pour la combinaison des deux priors uniformes sur (α,κ) et [-0,90 ; -0.07] pour la combinaison des deux priors très informatifs. (ANNEXE 1.5).

La dpp du paramètre de variabilité du recrutement σ est peu sensible aux formes de courbe et aux différentes combinaisons de priors (ANNEXE 1.6), avec une forte mise à jour a posteriori.

La matrices de corrélation a posteriori des paramètres révèle une relative confusion entre les deux paramètres clés α et κ dans le cas de la fonction de BH (ANNEXE 1.7). On observe une corrélation positive entre ces deux paramètres avec Ricker, qui est moins marqué pour DS.

2.1.4 Choix de combinaisons de priors et courbes SR

La difficulté à déterminer γ de manière précise sans amélioration notable des performances au sens des autres critères analysés conduit à ne pas retenir la courbe de Deriso-Schnute dans les analyses ultérieures. Pour les formes de courbes SR de Ricker et BH, le WAIC plaide plutôt en faveur de BH mais la qualité de la mise à jour a posteriori des paramètres penche plutôt en faveur de l'usage de Ricker. La différence entre BH et Ricker au sens du WAIC n'est cependant pas suffisamment marquée pour choisir l'une ou l'autre option. Malgré la dépendance au prior, l'option BH est beaucoup utilisée (Adkison, 2021) et recommandée pour l'étude du saumon atlantique (Pulkkinen and Mäntyniemi, 2013). Les dpp étant sensibles au prior, dans le cas de la courbe BH, des combinaisons de priors avec des niveaux d'information variés sont considérés dans la suite. Par souci de simplification, les priors uniformes sur α et κ, les plus irréalistes au regard des connaissances écologiques disponibles, sont éliminés. Trois combinaisons contrastées sont conservées : $\boldsymbol{\alpha 0.5_\kappa 0.5}$; $\boldsymbol{\alpha 0.8_\kappa 0.8}$; $\boldsymbol{\alpha 0.95_\kappa 0.95}$. Pour la suite, seule l'indication du chiffre « 0.5 », « 0.8 » ou « 0.95 » sera utilisée pour indiquer les priors sur α et κ utilisés.

2.2 Analyse de sensibilité aux hypothèses de l'erreur de processus

2.2.1 Vérification à posteriori

2.2.1.1 Comparaison entre les données répliquées et observées

La comparaison entre les données répliquées et observées ne révèle pas d'inadéquation entre le modèle et les données quelle que soit l'hypothèse faite sur l'erreur de processus. Les valeurs des p-values sont comprises entre [0.43 ;0.61] pour la statistique du χ^2 et comprises entre [0.37 ; 0.57] pour la proportion de recrutements inférieurs à $\kappa/2$ (ANNEXE 2.1).

2.2.1.2 Analyse des résidus standardisées

Quelle que soit l'hypothèse les résidus semblent répartis de manière homogène autour de 0 tout en étant essentiellement compris dans l'intervalle [-1.96 ; 1.96] (ANNEXE 2.2).

La figure 9 montre la dpp du coefficient d'autocorrélation pour les résidus supposés iid seulement pour une combinaison de prior (0.8) car les autres combinaisons de prior donnent des résultats similaires (ANNEXE 2.3). L'autocorrélation des résidus pour les modèles sans hypothèse Ac est généralement positive (avec des probabilités que le coefficient soit positif entre 0,8-1 selon les hypothèses). L'autocorrélation des résidus est plus forte pour BH (médianes des dpp supérieures à 0,23) que pour Ricker (médianes des dpp sont comprises entre 0,10 et 0,20) avec les hypothèses sans autocorrélation.

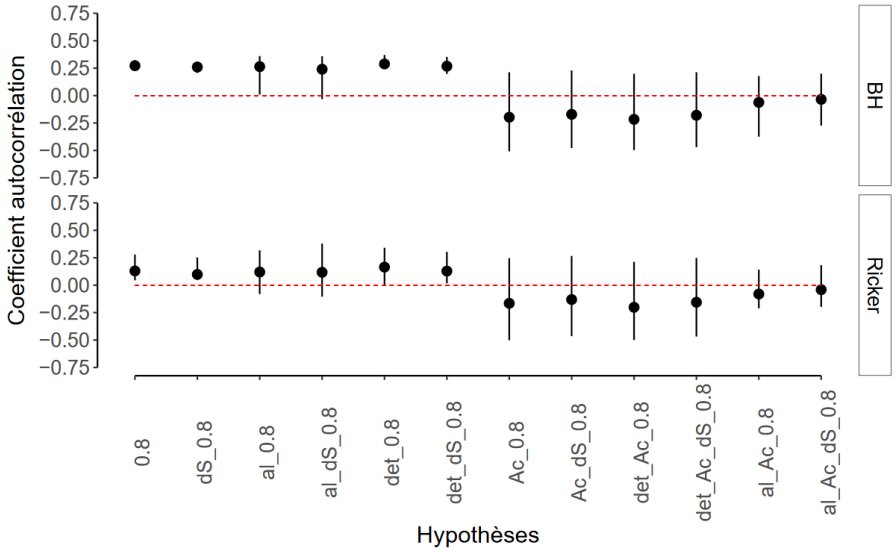

Figure 9 : Distributions de probabilité a posteriori du coefficient de d'autocorrélation des résidus iid standardisés, ϵ (eq. 14) pour les hypothèses sans autocorrélation et δ (eq. 15) pour les hypothèses avec autocorrélation. Les points et les lignes matérialisent respectivement la médiane et l'intervalle à 95% de probabilité des distributions.

La prise en compte explicite de l'autocorrélation dans le modèle (ajout de l'hypothèse Ac), corrige cette inadéquation entre modèle et données, particulièrement pour les combinaisons des hypothèses al et Ac, avec ou sans dépendance de la variabilité du recrutement en fonction du stock. Dans ces cas, les intervalles à 95 % du coefficient de d'autocorrélation des résidus supposés iid sont les plus resserrés pour Ricker, avec des valeurs médianes proches de 0 (cf. fig. 9).

2.2.1.3 Vérification à posteriori des effets

Pour les modèles incluant explicitement de l'autocorrélation, la probabilité a posteriori que ρ soit supérieur à 0 est marquée (>0.99) pour la combinaison d'hypothèses al_Ac et al_Ac_dS avec des médianes supérieures à 0,78 (ANNEXE 2.4). En revanche, aucune configuration testée n'identifie a posteriori le paramètre b comme étant supérieur ou inférieur à 0 avec une probabilité élevée (ANNEXE 2.5). Il semble donc que l'hypothèse d'une dépendance entre le stock et la variabilité du recrutement ne puisse être formellement validée par les données.

Pour les hypothèses det et det_dS avec un double régime déterministe, on observe, particulièrement avec Ricker, que l'effet du paramètre λ est faible avec une probabilité d'être supérieure à 1 comprise entre 0,50 et 0,15. Concernant la forme de BH, le paramètre λ a une probabilité d'être supérieur à 1 inférieure à 0,2 peu importe la combinaison d'hypothèses. Cela ne valide pas au regard des données l'existence de deux régimes distincts de recrutement avant et après l'ouverture du barrage de la partie amont du bassin de la Nivelle (au-dessus du barrage d'Olha ; ANNEXE 2.6).

2.2.2 Critère d'information : WAIC

Le critère WAIC montre de grandes différences en fonction des modèles (figure 10). La combinaison des hypothèses avec autocorrélation et un double régime aléatoire de recrutement conduit à des WAIC significativement inférieurs. Dans cette configuration, l'utilisation d'une courbe de Ricker est clairement favorisée (le WAIC du meilleur modèle avec courbe de Ricker est inférieur de 35 au meilleur avec une courbe BH). Les « meilleurs » modèles correspondent à la combinaison de l'hypothèse al_Ac avec la fonction de Ricker, sans que des différences notables soient détectées selon les priors retenus pour alpha et kappa. L'adjonction de l'hypothèse dS ou det n'apporte pas d'avantage supplémentaire, au contraire (fig. 10).

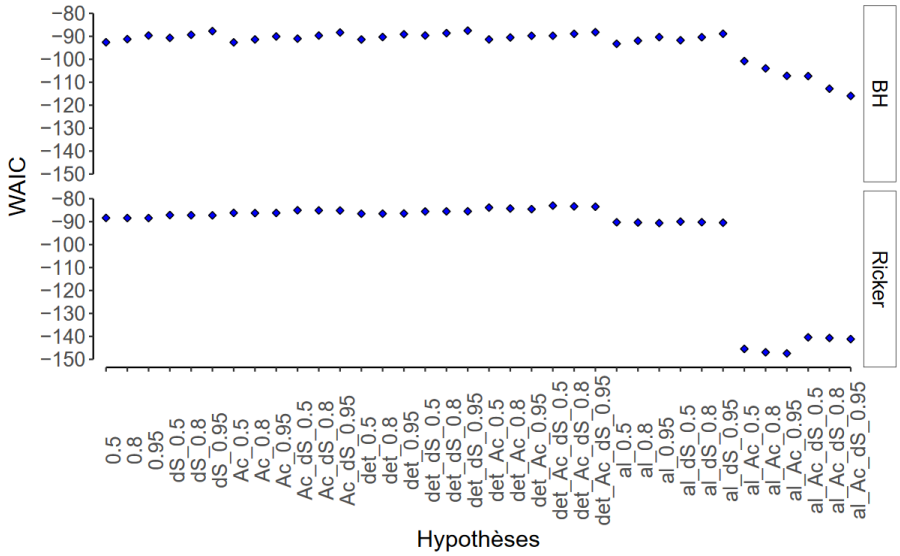

Figure 10 : WAIC en fonction des différentes combinaisons d'hypothèses et courbes SR.

2.2.3 Mise à jour et identification des modèles

Le paramètre α est actualisé a posteriori sans influence marquée du prior avec la fonction de Ricker, contrairement à la fonction de BH pour laquelle l'estimation a posteriori reste dépendante du prior (fig. 11). Pour les deux fonctions, l'hypothèse al et les combinaisons associées augmentent l'incertitude a posteriori de alpha, surtout avec les priors peu informatifs (fig. 11). Les valeurs médianes en échelle naturelle pour Ricker sont proches de 0,03, par contre pour BH les valeurs médianes varient fortement selon le prior sur α.

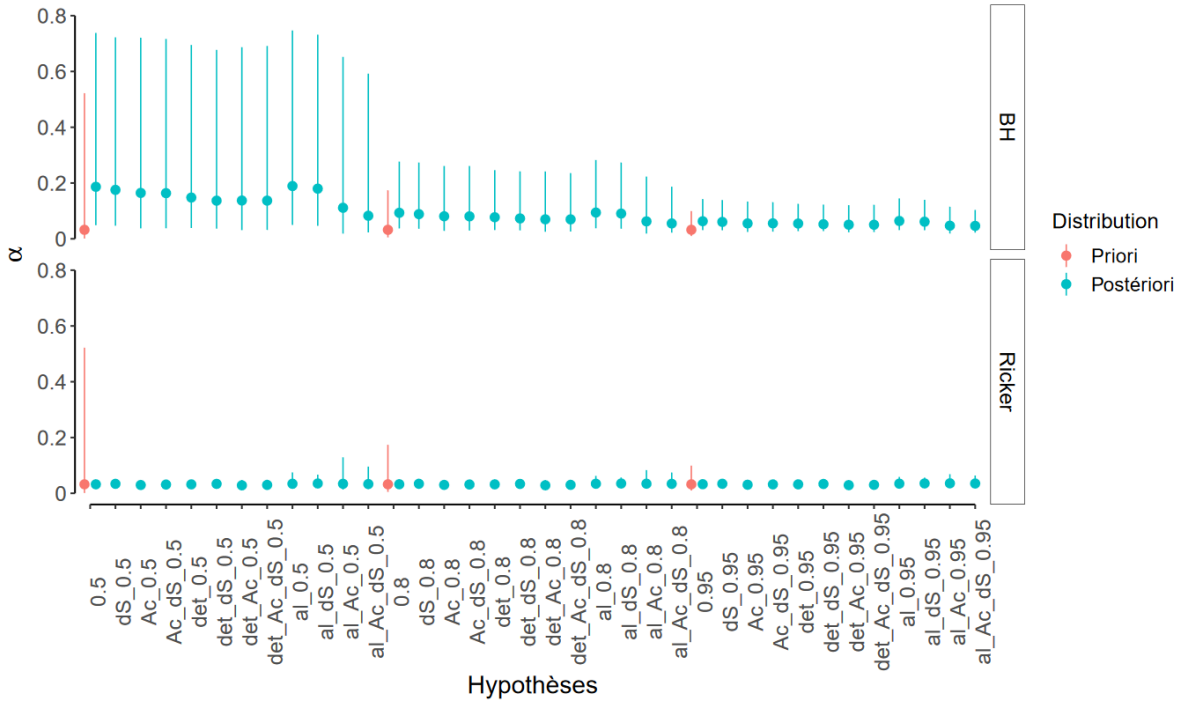

Figure 11 : Distributions marginales a priori en rouge et a posteriori en bleu du taux de survie des recrues médian maximal (α) en échelle naturelle. Les points et les lignes matérialisent respectivement la médiane et l'intervalle à 95% de probabilité des distributions.

La figure 12 présente seulement la combinaison de prior 0.8 (informatif), des résultats similaires étant obtenus pour les autres priors (cf. annexe 2.7). La mise à jour a posteriori du paramètre κ est notablement moins bonne dès lors que les hypothèses de mélange al ou det son combinées avec de l'autocorrélation Ac. Dans ces configurations, les estimations de κ sont aussi plus élevées en échelle naturelle. (ANNEXE 2.8). Par exemple, la médiane de la dpp de κ avec Ricker sous l'hypothèse 0.8 est égale 0,15 tacons^{0+}·m^{-2}, alors que pour l'hypothèse al_Ac_0.8 la médiane vaut 0,20 tacons^{0+}·m^{-2}.

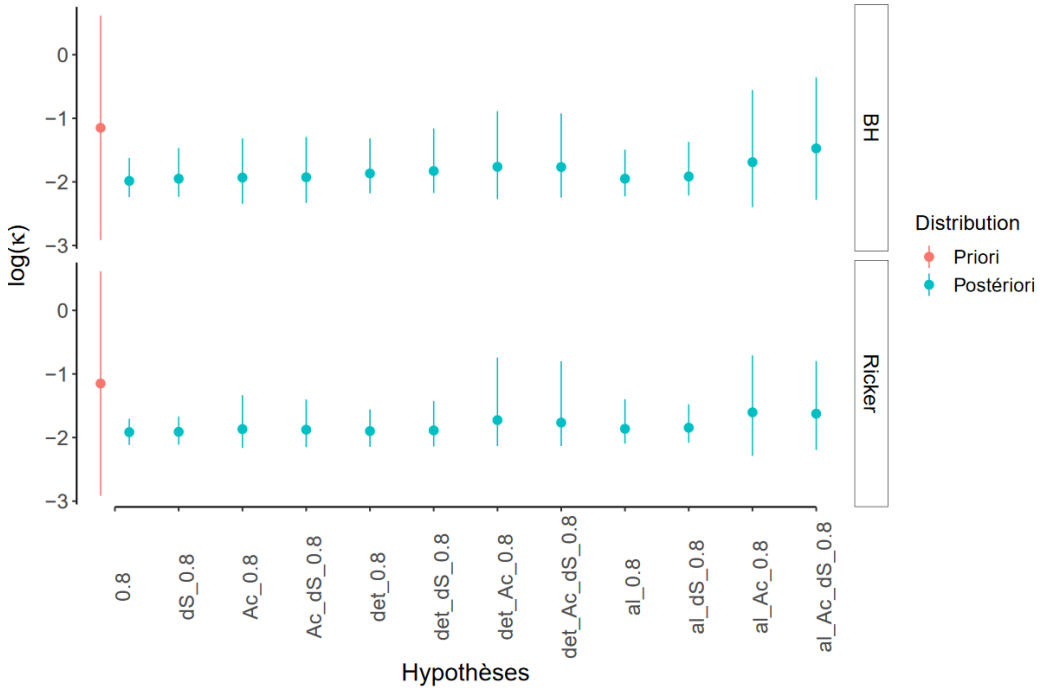

Figure 12 : Distributions marginales a posteriori en rouge et a priori en bleu du recrutement médian maximal (κ) en échelle log. Les points et les lignes matérialisent respectivement la médiane et l'intervalle à 95% de probabilité des distributions.

Le paramètre λ n'est pas toujours estimé a posteriori avec précision (fig. 13 et ANNEXE 2.9). Les mises à jour a posteriori les plus marquées sont obtenues avec la fonction de Ricker combinée aux hypothèses al_Ac et al_Ac_dS. Avec les mêmes hypothèses, l'actualisation est moins évidente pour une forme BH. Concernant les hypothèses det et les combinaisons associées, la médiane est centrée sur 0, particulièrement avec Ricker. L'effet de λ est faible sous l'hypothèse d'un régime déterministe.

Le paramètre p, i.e. la probabilité que l'indicatrice s'active dans le cas d'un régime de mélange aléatoire, semble avoir un comportement similaire à λ. Les mises à jour a posteriori les plus marquées sont obtenues avec la fonction de Ricker et les hypothèses al_Ac et al_Ac_dS. (ANNEXE 2.10).

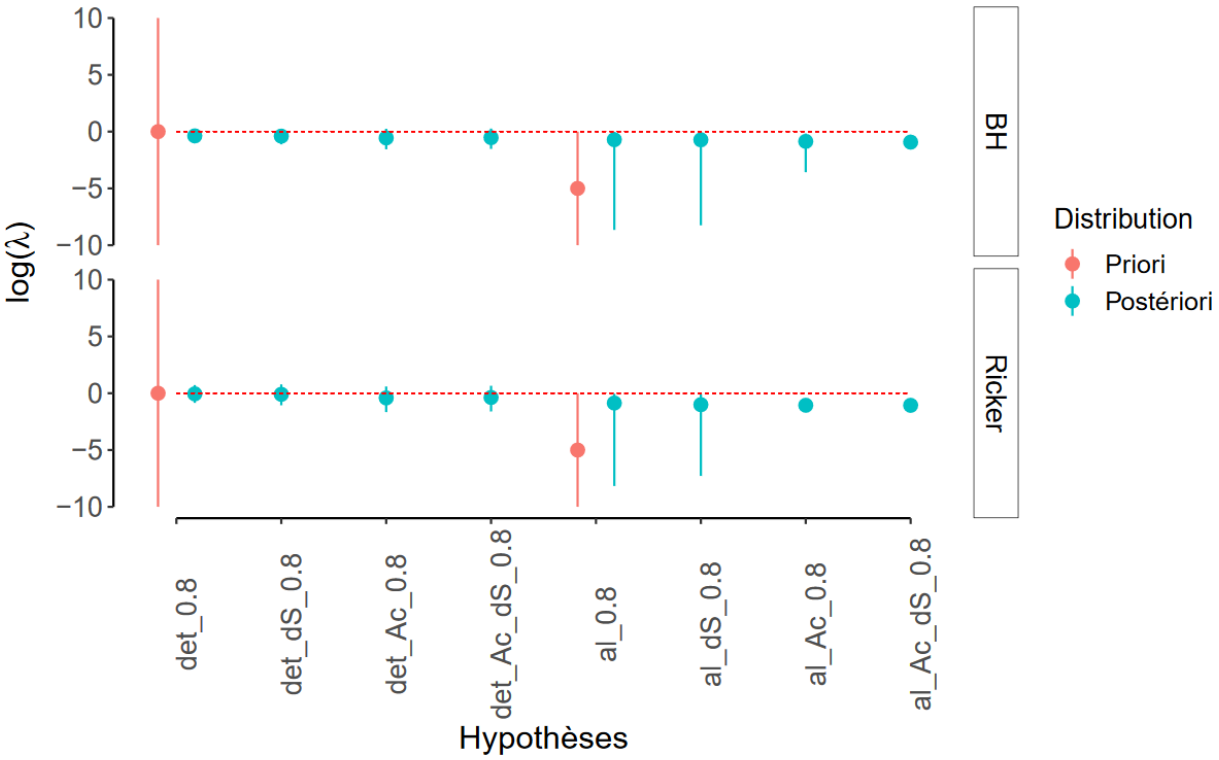

Figure 13 : Distributions marginales a priori en rouge et a posteriori en bleu du paramètre lié au double régime de recrutement λ en échelle log. Le premier prior correspond à l'hypothèse déterministe (det), le second à l'hypothèse aléatoire (al). Les points et les lignes verticales matérialisent respectivement la médiane et l'intervalle à 95% de probabilité des distributions.

La corrélation a posteriori du couple (α, κ) dépend des hypothèses et des fonctions utilisées. La corrélation entre les deux paramètres augmente fortement avec l'hypothèse al et ses combinaisons couplées à Ricker, avec des valeurs de corrélations d'environ 0.8-0.9 (ANNEXE 2.11). Pour BH au contraire avec l'hypothèse al et ses combinaisons la corrélation a posteriori du couple (α, κ) est presque nulle.

2.2.4 Choix des différentes hypothèses sur l'erreur de processus

Suite à la vérification a posteriori de l'effets du paramètre b, les modèles avec l'hypothèse de dépendance entre l'écart-type et le stock sont écartés. Concernant la distribution de mélange, la version déterministe ne semble pas adaptée du fait d'une incapacité à distinguer les deux périodes considérées (pré vs post ouverture du barrage d'Olha).

La combinaison des hypothèses al et Ac avec une courbe de Ricker se détache nettement avec à la fois un WAIC très inférieur à toutes les autres options et une meilleure actualisation a posteriori des paramètres spécifiques de ces hypothèses (p, λ et ρ). Ce modèle conduit cependant à des estimations a posteriori moins précises du couple (α, κ), avec une relative confusion entre ces deux paramètres (forte corrélation a posteriori). Malgré cela il semble que ce modèle permet de mieux capturer la stochasticité du processus du recrutement.

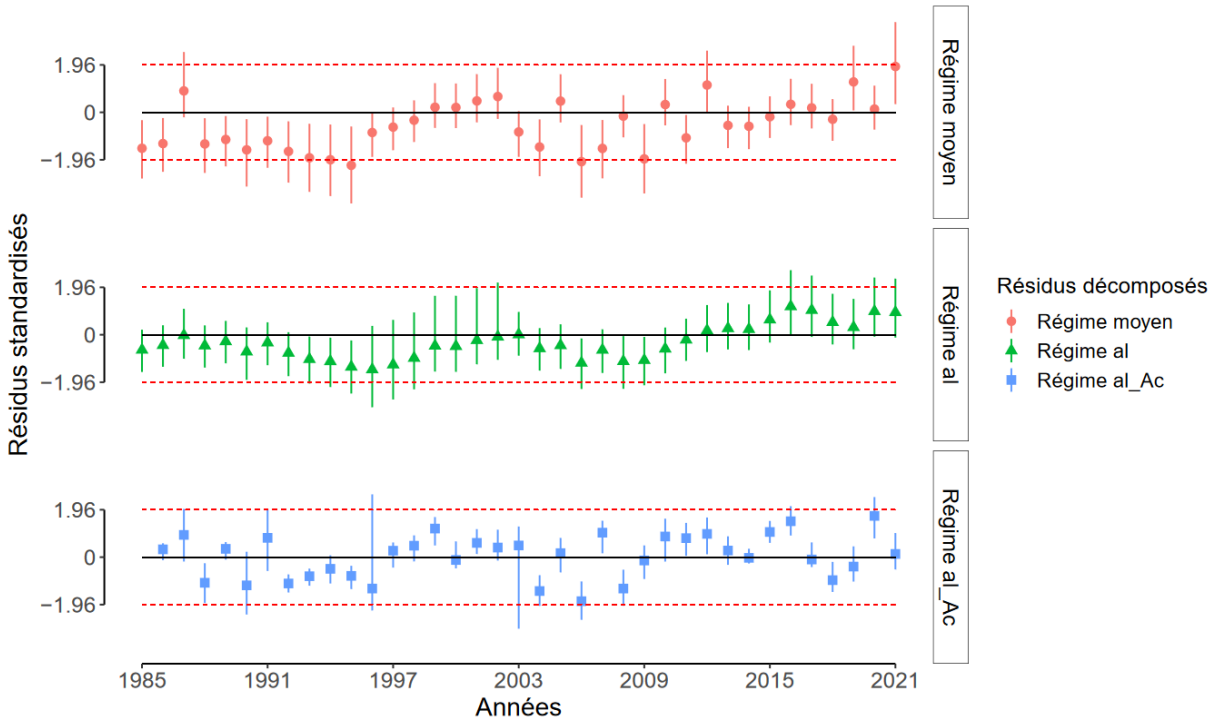

Figure 14 : Résidus standardisés avec l'hypothèse al_Ac_0.95 et la fonction de Ricker, décomposés par étape. Régime moyen = résidus standardisés mais seulement par rapport à la courbe moyenne de recrutement. Régime al = prise en compte du double régime aléatoire mais pas de l'autocorrélation. Régime al_Ac = Résidus δ de l'hypothèse al_Ac_0.95 qui prennent en compte l'autocorrélation et le double régime de recrutement.

La figure 14 décompose en trois parties les résidus standardisés de l'hypothèse al_Ac_0.95 avec la fonction de Ricker. Le régime moyen correspond aux résidus calculés seulement par rapport à la courbe moyenne de recrutement. Cette série fait apparaître à la fois des périodes de fortes autocorrélation entre-coupées par des « à-coups » aléatoires. Quand on ajoute la distribution de mélange (Régime al) cela capture les « à-coups » du recrutement (fig. 14) : la série temporelle des résidus apparaît alors « lissée » et fortement auto-corrélée. L'ajout de l'autocorrélation permet de prendre en compte une partie de la dépendance temporelle des résidus (Régime al_Ac : résidus δ de l'hypothèse avec Ricker al_Ac_0.95) : les résidus sont plus indépendants séquentiellement.

Pour la suite des analyses, le choix a donc été fait de sélectionner la combinaison d'hypothèse al_Ac pour la fonction de Ricker. L'hypothèse iid combinée aux deux formes de courbes (i.e Ricker et BH) est également conservée pour avoir un point de comparaison avec des modèles simples consacrés par l'usage. Deux combinaisons de priors contrastées, peu informatif 0.5 et très informatif 0.95 sont également gardées pour chaque modèle pour évaluer l'éventuelle influence des priors.

2.3 Prise en compte des erreurs de mesures

L'effet des erreurs de mesure est évalué sur les six configurations retenues : BH iid, Ricker iid, Ricker Al-Ac sous deux spécifications de priors. La figure 15 représente les erreurs de mesure pour le stock et le recrutement.

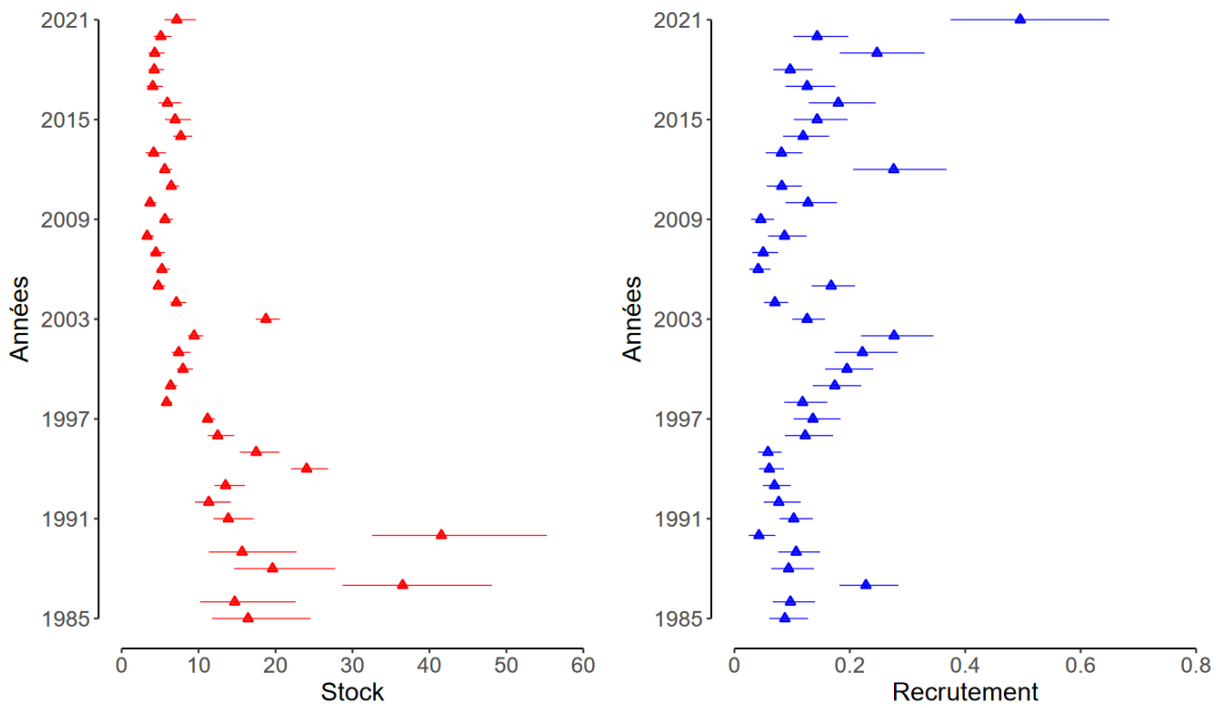

Figure 15 : Distributions de probabilité a posteriori du stock (en rouge) et du recrutement (en bleu). Les lignes et les points correspondent respectivement à l'intervalle à 95% des distributions et aux médianes.

Pour le stock, les premières années du jeu de données ont une incertitude importante ensuite cela se réduit avec le temps contrairement au recrutement où l'on observe une incertitude importante pour la dernière année en 2021.

2.3.1 Mise à jour et identification des modèles

L'effet des erreurs de mesure (EM) est relativement faible sur l'estimation des paramètres clés $(\alpha\,;\kappa)$ (ANNEXE 3.1) : une légère différence est visible seulement pour le recrutement médian maximal κ avec l'hypothèse al_Ac, peu importe le prior. La médiane a posteriori avec cette hypothèse et un prior peu informatif (0.5) est de 0,18 tacons^{0+}·m^{-2} avec EM et de 0,20 tacons^{0+}·m^{-2} sans EM (ANNEXE 3.2). Pour les autres paramètres du modèle plus complexe ac_Al, il y a une certaine influence des EM. L'estimation des paramètres liés aux deux régimes de recrutement λ et p est plus incertaine avec les EM que sans les EM (ANNEXE (3.3). Par exemple l'intervalle de la dpp à 95% pour λ avec un prior peu informatif (0.5) est [0,03 ; 0,80] avec EM alors que sans les EM l'intervalle est de [0,29 ; 0,40]. Le même comportement est observé avec le paramètre d'autocorrélation ρ (ANNEXE 3.3). La prise en compte des erreurs de mesure implique une plus grande difficulté à identifier le régime haut ou bas et implique une moins bonne mise à jour du paramètre d'autocorrélation.

2.3.2 Prédictions du recrutement

Pour construire les diagrammes de risque qui permettent de déterminer la LC, il faut d'abord obtenir la distribution de probabilité $p\left(R^{pred}\middle|S^{pred},D^{S},D^{R}\right)$. La figure 16 représente la médiane des prédictions du recrutement en fonction des différents modèles/priors choisis. Ces distributions prédictives diffèrent fortement selon le modèle utilisé et sont sensibles au prior uniquement dans le cas d'une courbe BH. En revanche, elles sont peu sensibles à la prise en compte des erreurs de mesure (fig. 16), sauf pour le modèle de Ricker couplé à l'hypothèse al_Ac. Dans ce cas, on remarque une différence des prédictions qui semblent légèrement plus faibles quand on prend en compte les erreurs de mesure. Le recrutement médian maximum

est de 0,18 tacons^{0+}·m^{-2} sans EM et de 0,16 tacons^{0+}·m^{-2} avec EM pour l'hypothèse al_Ac avec un prior très informatif (0.95).

Le fait que BH soit sensible au prior pour l'estimation du paramètre α (pente à l'origine) explique la différence de pente entre les deux courbes selon le prior. De plus le recrutement médian maximal, avec BH est plus faible pour un prior faiblement informatif qu'avec un prior informatif (fig. 16). L'hypothèse sur l'erreur de processus modifie aussi les prédictions, plutôt au niveau du recrutement médian maximal qui est plus grand avec Ricker et l'hypothèse al_Ac que l'hypothèse iid.

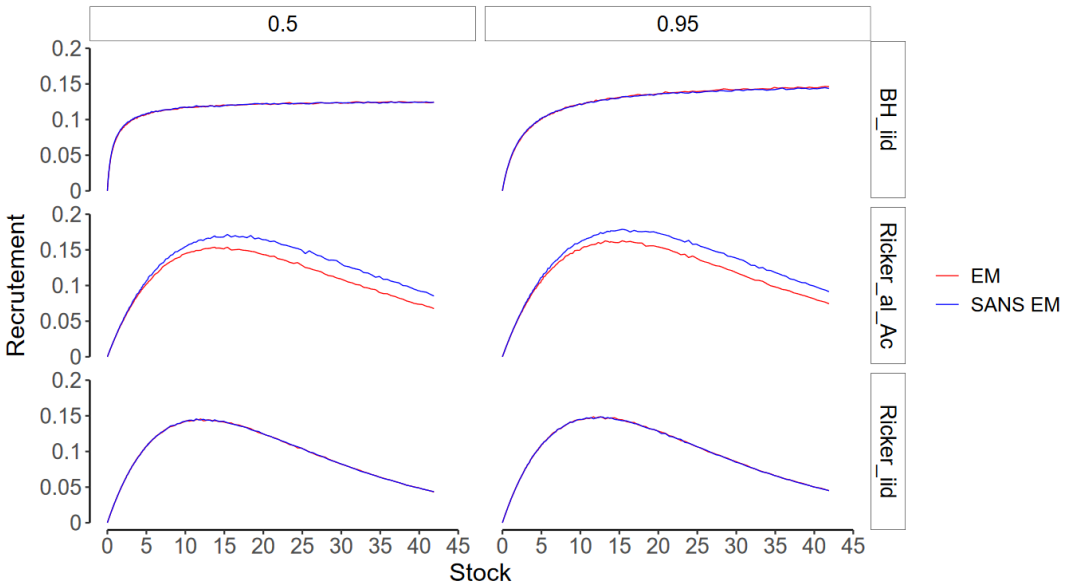

Figure 16 : Prédictions du recrutement (R^{pred}) en fonction du stock (S^{pred}). Les lignes pleines rouges et bleues correspondent respectivement aux médianes des distributions du recrutement prédit avec EM et sans EM.

L'examen des intervalles a posteriori à 90% des distributions de recrutement prédit révèle en outre une légère sous-dispersion des prédictions (relativement aux observations) pour chaque combinaison de priors avec l'hypothèse iid (ANNEXE 3.4).

2.3.3 Diagramme de risque et Détermination des LC's

Pour chaque combinaison d'hypothèse de modélisation et de prior, la courbe de risque de faible recrutement (i.e. $<\kappa/2$) est très peu sensible à la prise en compte des erreurs de mesure. La figure 17 représente les courbes de risques obtenus pour chaque combinaison d'hypothèses avec et sans erreurs de mesure. Globalement on observe une différence forte entre les trois modèles BH iid , Ricker iid et Ricker al_Ac. Dans le cas BH iid, la courbe de risque est en outre fortement influencée par le prior. Ces différences entre courbes de risques induisent elles-mêmes des LC's (pour un risque de 25%) assez notablement différentes.

Pour le modèle BH iid avec EM et le prior 0.5, la LC est égale à 2.8 oeufs·m^{-2}, alors qu'avec le prior 0.95 la LC est égale à 8,1 oeufs·m^{-2}. Avec Ricker sous l'hypothèse iid avec EM, la LC est comprise entre 5.20 (prior 0.5) à 5.25 (prior 0.95) oeufs·m^{-2}, alors que pour l'hypothèse al_Ac avec EM, la LC elle est comprise entre 12.89 (prior 0.5) et 13,89 (prior 0.95) oeufs·m^{-2}. À noter aussi que sous cette hypothèse (i.e al_Ac), 25 % est la valeur de risque la plus faible que l'on peut maitriser avec une définition du faible recrutement égale à $\kappa/2$.

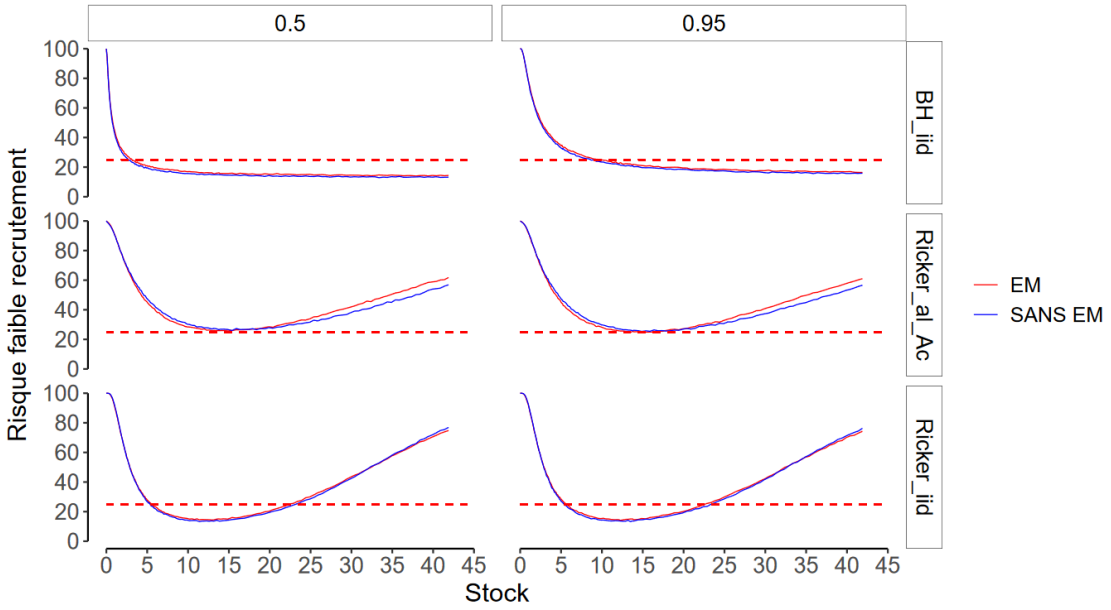

Figure 17 : Diagrammes de risque qui représentent l'évolution du risque de faible recrutement (%) en fonction d'une gamme de stock donnée. Les lignes bleues correspondent aux hypothèses sans erreur de mesure, les lignes rouges correspondent aux hypothèses avec erreurs de mesure.

2.4 Analyse rétrospective du statut de conservation

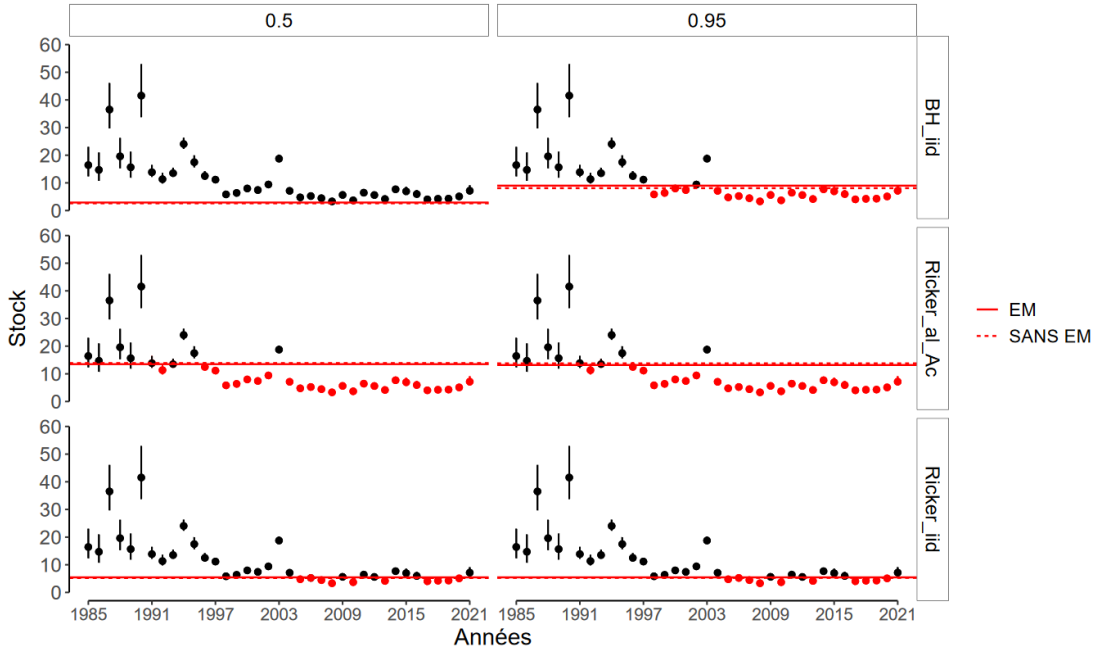

Figure 18 : Analyse rétrospective des estimations du stock de saumon sur la série temporelle (1985-2021) en fonction des LC's obtenues selon l'option de modélisation et de prior retenu.

La série temporelle du stock estimée rétrospectivement sur la Nivelle peut être comparée aux différentes LC's selon l'option de modélisation et de prior retenue (fig. 18). Pour les hypothèses al_Ac préférentiellement retenues, le stock est inférieur à la LC de façon pratiquement continue (sauf en 2003) depuis la seconde moitié des années 90 et on observe seulement 9 années (le long de la série temporelle) où le stock est supérieur à ces LC's. Ce résultat donne une vision très négative du statut de conservation de la population de saumon atlantique de la Nivelle depuis plus de 25 ans. Des résultats très différents sont obtenus sous

des options de modélisation plus simples et plus classiques. Le recours au modèle de Ricker iid conduit à considérer que le stock fluctue autour de sa LC depuis 25 ans. Avec un modèle BH iid et un prior peu informatif 0.5, qui pourrait être vu comme une option standard à utiliser en première intention, le stock est toujours au-dessus de sa LC depuis le début de la série observée. L'utilisation d'un prior plus informatif 0.95, mais aussi plus réaliste, conduit à une vision radicalement opposée, le stock étant essentiellement en dessous de la LC depuis 25 ans.

3 Discussion

3.1 Approche générale

Cette étude porte sur la représentation des différentes composantes de la relation SR (i.e forme de courbe SR, erreur de processus, probabilité à priori des paramètres $p(\theta)$, erreurs de mesure) et comment ces différentes représentations peuvent avoir un impact sur la détermination de point de référence biologique, ici la limite de conservation. Ceci constitue une avancée significative par rapport à beaucoup d'autres études sur la relation SR pour l'établissement de point de référence biologique qui ne vont s'intéresser qu'à une composante (i.e erreur de mesure, erreur de processus, forme de courbe, prior) (Hennemuth et al., 1980; Michielsens and McAllister, 2004; Pulkkinen and Mäntyniemi, 2013; Walters and Ludwig, 1981).

La combinaison des hypothèses pour les différentes composantes implique rapidement un nombre important d'options combinées à tester. Dans notre cas, la démarche séquentielle proposée permet de réduire significativement le nombre d'options testées (576 au total dans notre cas) tout en explorant des hypothèses variées. Au-delà de son utilité pratique, elle présente aussi des limites.

Tout d'abord elle amène à faire des choix à chaque étape qui peuvent s'avérer problématiques à l'étape suivante du fait d'interactions entre les différentes hypothèses. Ainsi, une hypothèse qui n'est pas forcément adaptée dans certaines configurations et serait écartée peut s'avérer par la suite pertinente quand elle est combinée avec d'autres hypothèses. Dans notre cas d'étude, l'hypothèse de Ricker qui ne semblait pas être la plus pertinente avec des erreurs de processus iid, devient ensuite l'option à privilégier avec une représentation différente de ces dernières (i.e hypothèse al_Ac). Il convient donc de ne pas être trop strict dans la sélection opérée à chaque étape pour minimiser le risque d'éliminer hâtivement des hypothèses dont la pertinence ne se révèleraient qu'en interaction avec d'autres considérées ultérieurement. À contrario une hypothèse qui semble pertinente i.e al_Ac avec Ricker, peut s'avérer moins judicieuse lors de la prise en compte des erreurs de mesure. D'où l'importance de vérifier à chaque étape de la démarche séquentielle la justesse des hypothèses retenues.

Les critères dits d'information type « WAIC » sont très pratiques car ils permettent rapidement d'avoir un classement des différents modèles testés. Ils ont cependant leurs limites (Gelman et al., 2014). En particulier, ils ne donnent qu'une indication relative des performances des modèles dans le cadre restreint du jeu de modèles soumis à comparaison. Il est donc important, comme proposé ici, de faire précéder l'utilisation de critères d'information d'une étape de vérification a posteriori d'adéquation entre le modèle et les données. Ignorer cette étape fait prendre le risque d'opérer un classement entre des modèles qui présentent tous des problèmes préoccupants d'inadéquation. D'autres méthodes de comparaisons de modèle peuvent être utilisées, la validation croisée « leave-one-out » (LOO) étant l'une des plus

robuste (Gelman et al., 2014). Elle peut cependant être lourde à mettre en œuvre, même si Vehtari et al. (2017) propose une méthode d'implémentation plus simple.

La méthode numérique utilisée pour prendre en compte les erreurs de mesures, inspirée de Rivot et al. (2001), présente l'avantage de tenir compte de la structure de covariation de ces dernières. En effet, elle utilise un tirage de jeux données possibles dans leur distribution a posteriori $p(R^{obs}|D^R)$ et $p(S^{obs}|D^S)$ issue des modèles d'observations. Elle est en ce sens préférable aux approches de pseudo-vraisemblance mise en œuvre dans d'autres études (Olmos et al., 2019; Michielsens et al., 2008). Elle est aussi plus facile à mettre en œuvre qu'une approche plus rigoureuse mais aussi plus complexe d'intégration explicite et complète des modèles d'observations dans une approche de modélisation à espace d'état (e.g. Parent and Rivot, 2012; Rivot et al., 2004). Cette simplification est importante dans notre cas d'étude car les modèles d'observation du stock et du recrutement sont eux-mêmes complexes (Brun, 2011). Mais notre approche numérique présente l'inconvénient par la suite de ne pas permettre de mettre en œuvre les techniques de sélection de modèle utilisées aux étapes précédentes, en particulier les critères d'information type WAIC. Dans notre cas d'étude, cet inconvénient est sans doute limité car la prise en compte des erreurs de mesure n'a pas d'effet notable.

L'approche générale proposée est réplicable sur d'autres jeux de données. Une telle extension est à considérer, ne serait-ce que pour mieux en apprécier les avantages et limites. Cela pourrait permettre en outre d'avoir des résultats à comparer avec ceux de notre étude de cas pour des situations contrastées. Si le domaine d'application de l'approche proposée est potentiellement très large, une extension de la présente étude aux autres jeux de données homologues collectés dans le cadre de l'ORE DiaPFC (https://www6.inrae.fr/diapfc/) est directement envisageable. Elle viendrait utilement compléter les résultats nécessairement limités avec un seul cas d'étude. En particulier, dans le cas de la Nivelle, on dispose d'une série temporelle importante avec un fort contraste de stock et de recrutement, qui peut expliquer en partie que la prise en compte des erreurs de mesure a un effet réduit. Ignorer les erreurs de mesure avec des jeux de données moins longs et moins contrastés peut avoir des effets plus marqués (Walters and Hilborn, 1992; Walters and Ludwig, 1981)

3.2 Hypothèses écologiques

Le modèle le plus adapté semble in fine résulter de la combinaison d'une courbe de Ricker avec l'hypothèse al_Ac. Cette hypothèse suppose des variations aléatoires du recrutement couplant une autocorrélation d'ordre 1 avec un double régime de recrutement. Cette structure particulière, et originale par rapport à la littérature abondante concernant l'analyse des relations SR (Adkison, 2021), essaye de prendre en compte la variabilité environnementale difficilement explicable à l'aide de covariables directement intégrées dans le modèle (Adkison, 2021; Maunder and Thorson, 2019). Dans notre cas d'étude, elle permet à la fois de tenir compte de forts à-coups du processus de recrutement (fluctuations entre régimes haut et bas) combinés avec une forte autocorrélation temporelle entre ces à-coups. Les processus écologiques à l'œuvre qui induisent cette structure de variabilité du recrutement restent à investiguer. Leur dépendance vis-à-vis de facteurs hydro-climatiques serait d'un intérêt particulier pour anticiper l'évolution de la stochasticité de la dynamique des populations de saumon face au changement climatique en cours.

Pour la Nivelle, un modèle avec la fonction Ricker est favorisé. Les fonctions SR reposent sur différentes hypothèses écologiques concernant le taux de mortalité instantané des recrues (cf. section 1.1.4.1). Selon Pulkkinen and Mäntyniemi (2013), la fonction de BH est plus adaptée pour décrire les mécanismes écologiques sous-jacent chez le saumon atlantique. D'après Walters and Juanes (1993), la survie des juvéniles serait dépendante en

grande partie de leur capacité à se nourrir efficacement sans subir la prédation. La « foraging arena theory » (Walters and Korman, 1999) postule que les juvéniles passeraient la plupart de leur temps dans des refuges à l'abri des prédateurs et iraient se nourrir à proximité de ces abris le plus rapidement possible. Quand le nombre de juvéniles dans un espace donné est trop important, cela implique de devoir partir plus loin et plus longtemps pour se nourrir, d'où une prédation plus importante. Cette théorie peut aussi justifier l'utilisation de courbe SR de type BH (Walters and Korman, 1999). Néanmoins, Walters and Juanes (1993) notent que des formes proches de Ricker peuvent être obtenues avec certaines variations d'hypothèses en matière de compromis entre alimentation et risque de prédation. Il est également possible que la forme de la densité dépendance de la survie (i.e. BH ou Ricker) change au cours du cycle de vie, et implique donc différentes formes de courbes intégrées sur la phase œufs-tacons0+. Un modèle SR séquentiel (i.e pour différents stades entre œufs-tacons0+) peut permettre d'explorer cette hypothèse (Brooks and Powers, 2007) .

3.3 Diagramme de risque et LC

Les différentes représentations de l'erreur de processus, le choix du prior et la forme de la courbe ont un impact sur les prédictions du recrutement et in fine sur la LC. La valeur de la LC est radicalement différente avec l'hypothèse al_Ac et la fonction de Ricker (fig. 17) par rapport à l'hypothèses standard iid avec les fonctions Ricker et BH. Les erreurs de mesure dans notre cas n'ont pas d'effet sur la courbe de risque et donc sur la LC. Elles impactent tout de même l'estimation de certains paramètres et aussi les prédictions du recrutement, d'où l'importance d'évaluer leur effet sur la détermination des LC's. Comme souligné précédemment, ce point est d'une importance particulière pour des séries SR plus courtes et moins contrastées que celle de la Nivelle.

Avec la fonction de BH, le choix de prior, surtout sur α qui est notoirement difficile à estimer dans ce cas, peut aussi avoir un effet sur la LC. Sous l'hypothèse BH, une attention particulière doit être portée à la détermination du prior sur le taux de survie médian maximal des recrues α. Deux approches principales sont possibles, i.e. à dire d'expert (O'Hagan et al., 2006) ou via une méta-analyse (Avci, 2017; Pulkkinen and Mäntyniemi, 2013). L'usage d'un prior peu informatif, qui pourrait être considéré « naïvement » comme recommandé en première intention, est à proscrire si il conduit à donner du crédit a posteriori à des valeurs surestimées de la survie des recrues pour de faibles dépose d'œufs.

Avec la fonction de Ricker, la LC dépend de la manière dont on estime à la fois les paramètres α et κ. Sous l'hypothèse d'erreurs de processus privilégiée al_Ac, α et κ sont estimés légèrement plus élevés que sous l'hypothèse iid (principalement pour κ). L'incertitude des estimations de α et κ s'accroit également sous l'hypothèse al_Ac, avec une certaine confusion entre ces deux paramètres. Il en résulte des courbes de risques dont la décroissance initiale est moins forte et le minimum est plus élevé sous l'hypothèse al_Ac relativement à l'hypothèse iid. In fine cela conduit à des LC beaucoup plus élevées avec le modèle al_Ac. On notera en outre que dans ce cas, la LC correspond au risque de faible recrutement le plus faible que l'on puisse maitriser au regard des données disponibles.

Les LC's obtenues avec l'hypothèse « iid » correspondent aux valeurs obtenues pour d'autres rivières, en particulier celles produites par Lebot (2021) en Bretagne. Dans ce cas, les LC's sont également définies en fonction du même objectif de maitrise du risque de faible recrutement (i.e. 25% de probabilité d'être inférieur à $\kappa/2$), et au moyen d'une vaste méta-analyse portant sur 18 populations et un modèle SR de type BH avec des erreurs iid selon une hypothèse al. Les LC's obtenues au Canada, sont calculées à partir d'une vaste méta-analyse portant sur 14 rivières et d'un modèle SR de type Ricker ou BH (Chaput et al., 2015). L'ajout de covariables est aussi testé. Elles sont aussi définies selon le même objectif de maitrise du

risque de faible recrutement, elles vont de 0,9 à 14 oeufs·m^{-2} pour Ricker et entre 0,7 et 6,6 oeufs·m^{-2} pour BH pour les modèles sans covariable. Les résultats obtenus correspondent donc aussi à ces valeurs obtenues au Canada.

Pour la Nivelle, (Dumas and Prouzet, 2003) proposent une analyse de la relation SR selon un modèle de Ricker avec erreurs iid. La présente étude est une extension notable de cette analyse initiale, par l'ampleur de la série temporelle considérée, réduite à 10 cohortes (1991-2000) dans le cas de Dumas and Prouzet (2003), la diversité des modèles considérés et la prise en compte des erreurs de mesure.

3.4 Statut de conservation de la Nivelle

Les LC's variant en fonction du modèle choisi, le diagnostic de conservation de la population de saumon de la Nivelle change également. Ainsi, alors que sous l'option standard « par défaut » d'un modèle BH avec erreurs iid et des priors peu informatifs le stock est au-dessus de la LC sur l'ensemble la série temporelle étudiée, l'évaluation est radicalement inverse pour les 25 dernières années sous le modèle privilégié combinant une forme de Ricker et des erreurs de processus al_Ac. Le statut de conservation de la population de la Nivelle apparait donc comme préoccupant. Dumas and Prouzet, (2003) proposaient un diagnostic similaire, mais selon une argumentation assez différente (i.e. sans référence explicite à une LC) et surtout sur la base d'un jeu de données beaucoup plus réduit.

Bien que traitant un cas particulier, la population de saumon de la Nivelle, cette étude met en lumière l'importance de modéliser en amont de façon aussi adéquate que possible la stochasticité du recrutement pour déterminer des LC's fondées sur une maitrise du risque de faible recrutement. Les modèles standards les plus simples peuvent conduire à des évaluations du statut de conservation exagérément optimistes. Le recours à des modèles plus complexes, tel que celui combinant forme de Ricker et erreurs de processus « al_Ac », requiert sans doute cependant de disposer de séries SR suffisamment longues et avec des niveaux de stocks contrastés. L'évaluation du statut de conservation de populations de saumon fondés sur des LC's déterminées sur des séries courtes et des modèles simples peut-être problématique et doit être considérée avec la plus grande prudence.

Références bibliographiques

Adkison, M.D., 2021. A Review of Salmon Spawner-Recruitment Analysis: The Central Role of the Data and Its Impact on Management Strategy. Rev. Fish. Sci. Aquac. 0, 1–37. https://doi.org/10.1080/23308249.2021.1972086

Avci, E., 2017. Using Informative Prior from Meta-Analysis in Bayesian Approach. J. Data Sci. 15, 575–588. https://doi.org/10.6339/JDS.201710_15(4).00001

Beverton, R.J.H., Holt, S.J., 1957. On the Dynamics of Exploited Fish Populations, Chapman and Hall. ed.

Brooks, E.N., Powers, J.E., 2007. Generalized compensation in stock-recruit functions: properties and implications for management. ICES J. Mar. Sci. 64, 413–424. https://doi.org/10.1093/icesjms/fsl046

Brun, M., 2011. Aide à la décision pour la conservation des populations de saumon atlantique (Salmo salar L.) (These de doctorat). Pau.

Chaput, G., Prévost, E., Dempson, B., Dionne, M., Jones, R., Levy, A., Robertson, M., Veinott, G., 2015. Hierarchical Bayesian modelling of Atlantic salmon egg to smolt time series from monitored rivers of Eastern Canada to define and transport reference points (Research Report). Fisheries and Oceans Canada.

Dumas, J., Prouzet, P., 2003. Variability of demographic parameters and population dynamics of Atlantic salmon Salmo salar L. in a south-west French river. ICES J. Mar. Sci. 60, 356–370. https://doi.org/10.1016/S1054-3139(03)00003-1

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2003. Bayesian Data Analysis, Second Edition. CRC Press.

Gelman, A., Hwang, J., Vehtari, A., 2014. Understanding predictive information criteria for Bayesian models. Stat. Comput. 24, 997–1016. https://doi.org/10.1007/s11222-013-9416-2

Gelman, A., Simpson, D., Betancourt, M., 2017. The Prior Can Often Only Be Understood in the Context of the Likelihood. Entropy 19, 555. https://doi.org/10.3390/e19100555

Hennemuth, R.C., Palmer, J.E., Brown, B.E., 1980. A Statistical Description of Recruitment in Eighteen Selected Fish Stocks. J. Northwest Atl. Fish. Sci. 1, 101–111. https://doi.org/10.2960/J.v1.a10

Kuhnert, P.M., Martin, T.G., Griffiths, S.P., 2010. A guide to eliciting and using expert knowledge in Bayesian ecological models. Ecol. Lett. 13, 900–914. https://doi.org/10.1111/j.1461-0248.2010.01477.x

Lebot, C., 2021. Evaluation de l'impact de différents scenarii de gestion pour réguler l'exploitation par pêche du saumon atlantique (Salma Salar). [WWW Document]. http://www.theses.fr. URL http://www.theses.fr/s273973 (accessed 4.28.22).

Malavoi, J.-R., Souchon, Y., 2002. Description standardisée des principaux faciès d'écoulement observables en rivière: clé de détermination qualitative et mesures physiques. http://dx.doi.org/10.1051/kmae:2002040 365–366. https://doi.org/10.1051/kmae:2002040

Maunder, M.N., Thorson, J.T., 2019. Modeling temporal variation in recruitment in fisheries stock assessment: A review of theory and practice. Fish. Res., Recruitment: Theory, Estimation, and Application in Fishery Stock Assessment Models 217, 71–86. https://doi.org/10.1016/j.fishres.2018.12.014

Michielsens, C., McAllister, M., Kuikka, S., Mäntyniemi, S., Romakkaniemi, A., Pakarinen, T., Karlsson, L., Uusitalo, L., 2008. Combining multiple Bayesian data analyses in a sequential framework for quantitative fisheries stock assessment. Can. J. Fish. Aquat. Sci. 65, 962–974. https://doi.org/10.1139/F08-015

Michielsens, C.G., McAllister, M.K., 2004. A Bayesian hierarchical analysis of stock–recruit data: quantifying structural and parameter uncertainties. Can. J. Fish. Aquat. Sci. 61, 1032–1047. https://doi.org/10.1139/f04-048

Millar, R.B., 2002. Reference priors for Bayesian fisheries models. Can. J. Fish. Aquat. Sci. 59, 1492–1502. https://doi.org/10.1139/f02-108
NASCO, 2019. Report, State of North atlantic Salmon (Salmo salar).
NASCO, 2009. NASCO Guidelines for the Management of Salmon Fisheries. North Atlantic Salmon Conservation Organization (NASCO), Edinburgh, Scotland, UK.
O'Hagan, A., Buck, C.E., Daneshkhah, A., Eiser, J.R., Garthwaite, P.H., Jenkinson, D.J., Oakley, J.E., Rakow, T., 2006. Uncertain Judgements: Eliciting Experts' Probabilities. John Wiley & Sons.
Olmos, M., Massiot-Granier, F., Prévost, E., Chaput, G., Bradbury, I.R., Nevoux, M., Rivot, E., 2019. Evidence for spatial coherence in time trends of marine life history traits of Atlantic salmon in the North Atlantic. Fish Fish. 20, 322–342. https://doi.org/10.1111/faf.12345
Parent, E., Rivot, E., 2012. Nonlinear models for stock-recruitment analysis, in : Introduction to Hierarchical Bayesian Modeling for Ecological Data. pp. 145–168.
Peterman, R.M., 1981. Form of random variation in salmon smolt-to-adult relations and its influencd on production estimates. Can. J. Fish. Aquat. Sci. J. Can. Sci. Halieut. Aquat.
Prévost, E., Baglinière, J.-L.J.-L., 1994. Présentation et premiers éléments de mise au point d'une méthode simple d'évaluation du recrutement en juvéniles de saumon atlantique (Salmo salar) de l'année en eau courante. Cah. ORSTOM Pédologie 12, 11 p.
Pulkkinen, H., Mäntyniemi, S., 2013. Maximum survival of eggs as the key parameter of stock–recruit meta-analysis: accounting for parameter and structural uncertainty. Can. J. Fish. Aquat. Sci. 70, 527–533. https://doi.org/10.1139/cjfas-2012-0268
Ricker, W.E., 1954. Stock and Recruitment. J. Fish. Res. Board Can. 11, 559–623. https://doi.org/10.1139/f54-039
Rivot, E., Prévost, E., Parent, E., 2001. How robust are Bayesian posterior inferences based on a Ricker model with regards to measurement errors and prior assumptions about parameters? Can. J. Fish. Aquat. Sci. - CAN J Fish. AQUAT SCI 58, 2284–2297. https://doi.org/10.1139/cjfas-58-11-2284
Rivot, E., Prévost, E., Parent, E., Baglinière, J.L., 2004. A Bayesian state-space modelling framework for fitting a salmon stage-structured population dynamic model to multiple time series of field data. Ecol. Model. 179, 463–485. https://doi.org/10.1016/j.ecolmodel.2004.05.011
Royer, S.S., Prévost, E., 2016. Mise à jour et standardisation des séries chronologiques d'abondance du saumon atlantique sur les cours d'eau de l'ORE DiaPFC et la Bresle (report).
Schnute, J., 1985. A General Theory for Analysis of Catch and Effort Data. Can. J. Fish. Aquat. Sci. 42, 414–429. https://doi.org/10.1139/f85-057
Shelton, P.A., 1992. The Shape of Recruitment Distributions. Can. J. Fish. Aquat. Sci. 49, 1754–1761. https://doi.org/10.1139/f92-194
Subbey, S., Devine, J.A., Schaarschmidt, U., Nash, R.D.M., 2014. Modelling and forecasting stock–recruitment: current and future perspectives. ICES J. Mar. Sci. 71, 2307–2322. https://doi.org/10.1093/icesjms/fsu148
Thibault, M., 1994. Aperçu historique sur l'évolution des captures et des stocks, in: Le Saumon Atlantique. pp. 175–185.
Thorstad, E., Whoriskey, F., Rikardsen, A., Aarestrup, K., 2010. Aquatic Nomads: The Life and Migrations of the Atlantic Salmon. Atl. Salmon Ecol. 1–32. https://doi.org/10.1002/9781444327755.ch1
van de Schoot, R., Depaoli, S., King, R., Kramer, B., Märtens, K., Tadesse, M.G., Vannucci, M., Gelman, A., Veen, D., Willemsen, J., Yau, C., 2021. Bayesian statistics and modelling. Nat. Rev. Methods Primer 1, 1–26. https://doi.org/10.1038/s43586-020-00001-2
Van Dongen, S., 2006. Prior specification in Bayesian statistics: Three cautionary tales. J. Theor. Biol. 242, 90–100. https://doi.org/10.1016/j.jtbi.2006.02.002

Vehtari, A., Gelman, A., Gabry, J., 2017. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Stat. Comput. 27, 1413–1432. https://doi.org/10.1007/s11222-016-9696-4

Walters, C., Hilborn, J., 1992. Quantitative Fisheries Stock Assessment.

Walters, C., Korman, J., 1999. Linking recruitment to trophic factors: revisiting the Beverton--Holt recruitment model from a life history and multispecies perspective. Rev. Fish Biol. Fish. 9, 187–202. https://doi.org/10.1023/A:1008991021305

Walters, C.J., Juanes, F., 1993. Recruitment Limitation as a Consequence of Natural Selection for Use of Restricted Feeding Habitats and Predation Risk Taking by Juvenile Fishes. Can. J. Fish. Aquat. Sci. 50, 2058–2070. https://doi.org/10.1139/f93-229

Walters, C.J., Ludwig, D., 1981. Effects of Measurement Errors on the Assessment of Stock–Recruitment Relationships. Can. J. Fish. Aquat. Sci. 38, 704–710. https://doi.org/10.1139/f81-093

Watanabe, S., 2010. Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory. J. Mach. Learn. Res. 11, 3571–3594.

Wongnak, P., Bord, S., Donnet, S., Hoch, T., Beugnet, F., Chalvet-Monfray, K., 2022. A hierarchical Bayesian approach for incorporating expert opinions into parametric survival models: A case study of female Ixodes ricinus ticks exposed to various temperature and relative humidity conditions. Ecol. Model. 464, 109821. https://doi.org/10.1016/j.ecolmodel.2021.109821

L'INSTITUT agro Rennes Angers	Diplôme : Master 2 Spécialité SML-Biologie Parcours / option: SHA / Ressources et écosystèmes aquatiques Enseignant référent : Etienne Rivot

Auteur(s) : Benjamin FLAIS Date de naissance* : 29/11/1997	Organisme d'accueil : UMR ECOBIOP Adresse : 173, D918 route de Saint Jean Luz 64310 Saint Pée sur Nivelle Maître de stage : Etienne Prevost
Nb pages : 35 Annexe(s) : 24	
Année de soutenance : 2022	

Titre français

Détermination de limite de conservation du saumon Atlantique (*Salmo salar)* : analyse de sensibilité à partir du cas d'étude de la Nivelle (Pays Basque, France).

Titre anglais :

Determination of the conservation limit for Atlantic salmon (*Salmo salar*) : sensitivity analysis based on the case study of the Nivelle River (Basque Country, France).

Résumé (1600 caractères maximum) :

La limite de conservation (LC) est un point de référence biologique important pour le saumon atlantique (*Salmo salar*). Dans cette étude la LC correspond à l'abondance de géniteurs permettant de contrôler le risque de faible recrutement. La relation stock-recrutement (SR), au cœur de cette définition, peut être décomposée en plusieurs composantes : courbe SR, erreur de processus, connaissance a priori du modèle SR, erreur de mesure. L'objectif principal de ce travail est d'analyser la sensibilité des LC's à ces différentes composantes. Plusieurs types de courbes SR et hypothèses sur l'erreur de processus ainsi que différents niveaux de connaissance a priori du modèle SR sont analysés. Pour ce faire une démarche séquentielle est suivie afin de limiter le nombre de combinaisons à analyser. L'impact des erreurs de mesure sur la LC est analysé à la suite de cette analyse séquentielle sur une sélection préalable de combinaisons. Ce travail est mené au moyen d'une étude de cas : la population de saumon atlantique de la Nivelle (Pays Basque, France). Dans notre cas la LC peut radicalement changer selon la représentation des différentes composantes de la relation SR. La prise en compte des erreurs de mesure ne semble pas avoir un impact important sur les LC's. Enfin nous discutons de l'importance d'analyser de manière précise la stochasticité du recrutement pour déterminer des LC's fondées sur une maitrise du risque de faible recrutement.

Abstract (1600 caractères maximum) :

The conservation limit (CL) is an important biological reference point for Atlantic salmon (*Salmo salar*). In this study the CL is the spawning abundance that controls the risk of low recruitment. The stock-recruitment (SR) relationship, at the heart of this definition, can be decomposed into several components: SR curve, process error, a priori knowledge of the SR model, measurement error. The main objective of this work is to analyze the sensitivity of CL's to these different components. Several types of SR curves and assumptions on the process error as well as different levels of a priori knowledge of the SR model are analyzed. To do this, a sequential approach is followed in order to limit the number of combinations to analyze. The impact of the measurement errors on the LC is analyzed following this sequential analysis on a preliminary selection of combinations. This work is carried out by means of a case study: the Atlantic salmon population of the Nivelle (Basque Country, France). In our case the LC can change radically depending on the representation of the different components of the SR relationship. Factor in the measurement errors does not seem to have a significant impact on the CL's. Finally, we discuss the importance of accurately analyzing the stochasticity of recruitment to determine CL's based on a control of the risk of low recruitment.

Mots-clés : Saumon atlantique, Nivelle, Limite de conservation, Stock, Recrutement, Relation SR

Key Words: Atlantic salmon, Nivelle, Conservation limit, Stock, Recruitment, SR relationship

RÉPUBLIQUE FRANÇAISE

Liberté
Égalité
Fraternité

Avec le soutien financier de :

www.afbiodiversite.fr http://www.inrae.fr/ https://www.institut-agro-rennes-angers.fr/ www.univ-pau.fr