AGROCAMPUS OUEST

☐ CFR Angers ☑ CFR Rennes

Année universitaire : 2020 - 2021 Spécialité : Ingénieur Agronome Spécialisation (et option éventuelle) : Sciences Halieutiques et Aquacoles (Ressources et Ecosystèmes Aquatiques)	**Mémoire de fin d'études** ☒ d'ingénieur d'AGROCAMPUS OUEST (École nationale supérieure des sciences agronomiques, agroalimentaires, horticoles et du paysage), école interne de L'institut Agro (Institut national d'enseignement supérieur pour l'agriculture, l'alimentation et l'environnement) ☐ de master d'AGROCAMPUS OUEST (École nationale supérieure des sciences agronomiques, agroalimentaires, horticoles et du paysage), école interne de L'institut Agro (Institut national d'enseignement supérieur pour l'agriculture, l'alimentation et l'environnement) ☐ de Montpellier SupAgro (étudiant arrivé en M2) ☐ d'un autre établissement (étudiant arrivé en M2)

Development and comparison of tools to support the assessment of glass eels catch opportunities

Par : Mathilde BENEZECH

Soutenu à RENNES le 15/09/21

Devant le jury composé de :

Président : Etienne RIVOT

Maître de stage : Hilaire DROUINEAU

Enseignant référent : Etienne RIVOT

Autres membres du jury (Nom, Qualité)

Didier GASCUEL (Enseignant - chercheur Agrocampus Ouest)

Anthony ACOU (Jury extérieur)

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] :25/08/21
(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] **BENEZECH Mathilde**
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** : 25/08/21

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** : 15 Nov. 2021

(1) L'administration, les enseignants et les différents services de documentation d'AGROCAMPUS OUEST 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

ACKNOWLEDGMENTS

This internship was financed by the cooperation agreement Office Français de la Biodiversité - Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement - Institut Agro - Université de Pau et des Pays de l'Adour/E2S UPPA (Pôle Gestion des Migrateurs Amphihalins dans leur Environnement).

I would like to thank Hilaire Drouineau, my internship supervisor, for his supervision throughout my internship. Thank you for sharing your knowledge, always in an pedagogical way, and for your help to carry out this work. Thank you for your unfailing availability, and your sympathy.

I thank Anthony Acou, Agnès Bardonnet, Laurent Beaulaton, Clarisse Boulenger, Patrick Lambert and Etienne Prévost for their participation and investment during meetings, and for all the enriching discussions.

I would also like to thank all the members of the EABX Unit for these few months spent together. More specifically, thank you to Violette, my office mate, Marie and Chloé for the moments shared and the good mood, despite a special health context. Thanks to Chloé for her precious comments and to Violette for her tips.

I also thank Etienne Rivot, my Agrocampus Ouest referent teacher, for his availability during the whole internship and for his supervision. I would like to thank the entire pedagogical team for making this internship possible.

Finally, I would like to thank all the jury members for accepting and taking the time to evaluate my work.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AIC	Akaike Information Criterion
CITES	Convention of International Trade in Endangered Species
CRI	Conservation Risk Indicator
DIC	Deviance Information Criterion
EIFAAC	European Inland Fisheries Aquaculture Advisory Commission
EU	European Union
FRI	Fishery Risk Indicator
GEMAC	Glass Eel Model to Assess Compliance
GEREM	Glass Eel Recruitment Estimation Model
GFCM	General Fisheries Council of the Mediterranean
GLM	Generalized Linear Model
ICES	International Council for the Exploration of the Sea
IUCN	International Union for Conservation of Nature
TACs	Total Allowable Catches
WGEEL	Working Group on EEL

LIST OF FIGURES

LIST OF TABLES

LIST OF ANNEXES

1 Introduction

The European eel *Anguilla anguilla* (Linnaeus, 1758) is a temperate species. It is distributed over a 90,000 km^2 area, in continental waters from the Barents Sea (72°N, northern limit) to the southernmost limit of Morocco (30°N, southern limit), including the Mediterranean and Baltic basins (ICES, 2020a, p. 202; Moriarty & Dekker, 1997). As such, eels represent the most widespread fish in Europe (Dekker, 2003). Despite its large distribution area, the European eel is panmictic (i.e. homogeneous distribution of individuals and random reproduction) and consequently, it is considered as a single population (Als et al., 2011; Palm et al., 2009). Since the European eel's lifecycle is split between oceanic and continental environments (Figure 1), this species is catadromous (Tesch, 2003). Its reproduction occurs in the Sargasso Sea in Western Atlantic Ocean. After a larval drift of several months and thousands of kilometres through warm currents, eel larvae, called *Leptocephalae*, reach continental shelves of Europe and Northern Africa (Bonhommeau et al., 2010; Deelder, 1984; Dekker, 2000a; Drouineau et al., 2018; Tesch, 2003)) where they become glass eels. Then, glass eels penetrate continental waters, and progressively turn into pigmented yellow eels (Dekker 2000b). They remain in continental waters, from brackish to freshwater systems (Arai & Chino, 2012; Daverat et al., 2006), for their growth phase that can last from three to more than 30 years, depending on areas (Durif et al., 2020; Poole & Reynolds, 1996). Then, they metamorphose again into silver eels, and start a 6,000 km migration back to offshore spawning grounds (Arai & Chino, 2012; Bonhommeau et al., 2010; Dekker, 2000a; Tesch, 2003). Eels mature during the migration towards spawning grounds and presumably die after spawning (Béguer-Pon et al., 2015; Chang et al., 2016; Righton et al., 2016).

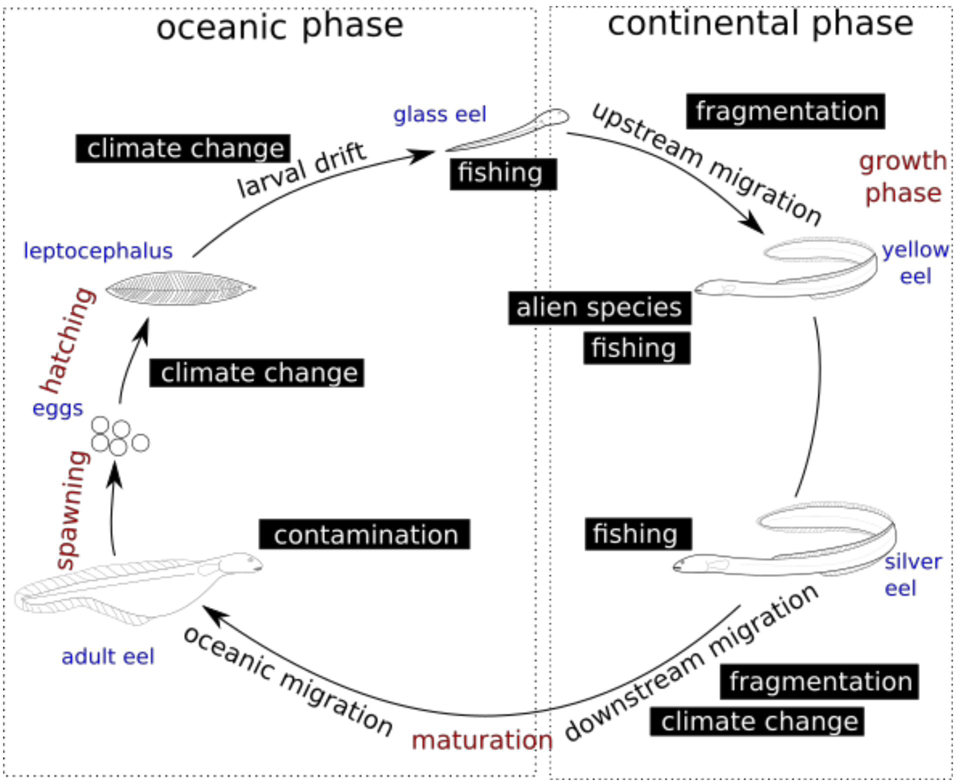


Figure 1. European eel life cycle with all life stages (in blue), biological mechanisms (in red) and pressures over the species (black labels) (according to Drouineau et al., 2018; modified).

Because of its complex life cycle (Figure 1), this species is submitted to various natural and anthropogenic pressures (Bonhommeau et al., 2010; Dekker, 2000a; Drouineau et al., 2018). For example, climate change is thought to have decreased the productivity in the Sargasso Sea reducing food availability for eel larvae. It may also have affected their migration (Arevalo et al., 2021; Bonhommeau et al., 2008). Parasites, such as *Anguillicola crassus*, is now widespread across Europe and can also affect eels ability to complete their spawning migration (ICES, 2020b; Kuwahara & Niimi, 1974). In addition to parasitism, contamination may also affect European eel fitness. While storing fat reserves in order to reproduce, eels tend to accumulate contaminants which can threaten migration success, fecundity and offspring survival (Tsukamoto & Kuroki, 2014). Changes in land use have greatly reduced available habitats such as wetlands (Feunteun, 2002). Settlement of anthropic structures such as dams impede upstream and downstream migrations while turbines can lead to direct mortality over silver eels during their downstream migration (Tsukamoto & Kuroki, 2014). Beyond the professional and recreational fisheries, overexploitation, poaching, and trafficking affect the state of the population (Dekker, 2000a; Drouineau et al., 2018). Concerning the fishery, it evolved from a local subsistence fishery to a larger scale commercial fishery in the early 1900s, eels becoming a luxury good (Dekker, 2019). The fishery has traditionally targeted all stages. Regarding glass eel, a fast development of the fishery took place in the 70s, in order to supply the eastern-Asian market (Crook & Shiraishi, 2015). In the early 2000s, the glass eel was the most valuable product landed in the French Bay of Biscay (Castelnaud, 2001). The glass eel fishery is mostly concentrated in the Bay of Biscay (87% of European glass eel fisheries) and to a lesser extent in the United Kingdom (Beaulaton & Briand, 2007; Dekker, 2000b). All of these factors, impacting eels throughout their life cycle, have participated in a cumulative way to the collapse of the population (Drouineau et al., 2018; van Ginneken & Maes, 2005).

Indeed, glass eel recruitment (i.e. the abundance of glass eel entering continental waters (Dekker, 2021; Drouineau et al., 2021)) started to decline by approximately 15% per year since the late 70s (Dekker, 2000a, 2019; ICES, 2018, 2020a; Moriarty, 1990). The stock was first declared “outside safe biological limits” in 1998 by the ICES (International Council for the Exploration of the Sea) (ICES, 2001; Tsukamoto & Kuroki, 2014) because the average recruitment was thought to be significantly lower than if the stock were at its full reproductive capacity. Therefore, the WGEEL (Working Group on EEL) recommended a recovery plan to reduce anthropogenic impacts in 2005 (ICES, 2005). In view of this, the European Council enforced a European Regulation in 2007 (EC No. 1100/2007), the Eel Regulation, which explicitly calls for a reduction of anthropogenic mortalities. More specifically, the Eel Regulation imposes that Member States implement Eel Management Plans in their national waters, to “permit with high probability the escapement to the sea of at least 40% of the silver eel biomass relative to the best estimate of escapement that would have existed if no anthropogenic influences had impacted the stock”. Concomitantly, the European eel was added to the Appendix II of CITES (Convention on International Trade in Endangered Species of Wild Fauna and Flora) in 2007 and, as a consequence, the European Union banned the export and the import of the species in 2010 (Nijman, 2015). Almost simultaneously, the IUCN (International Union for Conservation of Nature) listed the European eel as a critically endangered species in 2008 and its status has remained unchanged since then (Pike et al., 2020)

Following the Eel Regulation, France proposed its national Management Plan in 2008, which was accepted in 2010 by the European Commission. Since the Members States were responsible for selecting the most appropriate management measures in their own, this national plan aimed, among others, to reduce glass eels fishing mortality by 40% in three years (by 30% on yellow and silver eels), and then by 60% for all three fished life stages by 2015 (Ministère de l'Ecologie, de l'Energie, du Développement durable et de l'Aménagement du Territoire et al., 2010). In order to achieve the reduction of fishing mortality, France enforced measures such as the reduction of fishing seasons, the improvement of control on landings and reporting obligations, and the establishment of quotas for glass eels commercial fishery (Ministère de l'Ecologie, de l'Energie, du Développement durable et de l'Aménagement du Territoire et al., 2010). France is the only country operating with a glass eel quota system. Glass eels fishing quotas are set on a yearly basis and ensure that the exploitation rates are lower than 60% of the exploitation rates in the reference period 2004-2008 (Ministère de l'Ecologie, de l'Energie, du Développement durable et de l'Aménagement du Territoire et al., 2010). In this context, the French Ministry has established a Scientific Council in charge of making recommendations on the level of catches that would achieve the management target for the upcoming fishing season. The management target is an exploitation rate target, which is defined as the ratio of catches over absolute recruitment. Therefore, estimating an exploitation rate requires both catches and recruitment estimates (Beaulaton et al., 2020). However, the management target is defined relatively to an historical exploitation rate (the 2004-2008 reference period). As such, a relative recruitment index (at least since 2004/2008) can be enough. In addition to past trends in recruitment and landings, it is also necessary to predict future recruitment index to set up a glass eel quota. However, predicting recruitment is very complicated as recruitment is intrinsically random, since it depends on many factors including environmental conditions (e.g. temperature, ocean currents, food availability) (Bonhommeau et al., 2008; Drouineau et al., 2018; Myers, 1998; Subbey et al., 2014). In this context, the Scientific Council has developed two models to predict upcoming recruitments and to support the establishment of glass eel quota to achieve the management target. However, both recruitment models have limits, and the management target is no longer achieved since 2015 (Beaulaton et al., 2020). This raises questions about models' performances and their respective assumptions on past and future recruitment trends, which may lead to biases on predictions. Moreover, there is lack of relevant criteria to assess models' performance and of objective criteria to weight models' predictions. In this context, our objective is to improve the expertise on the glass eels catches question in addressing these limits by (1) proposing alternative recruitment models and (2) proposing new criteria to assess model performance in order to select a unique recruitment model.

First, we will present the two existing recruitment models and the previous models' comparison method in order to highlight their limits and their implications in terms of management. Then, we will introduce our two alternative recruitment models and we will detail a new model comparison strategy aiming to address all models' limits (Material and Methods section). Next, we will compare the four models' performance thanks to new comparison criteria (Results section). Finally, we will discuss the implications for management and the upcoming challenges for glass eel quota settlement (Discussion section).

2 Material and Methods

Catches, and consequently, quota setting are the result of two factors: (i) the exploitation rate and (ii) the abundance of the targeted species, which corresponds, here, to the quantity of glass eels penetrating into continental waters, so called recruitment (Dekker, 2021). As such, both factors are accounted to settle quotas (Beaulaton et al., 2020). In this internship, we exclusively focused on the evolution of recruitment and its prediction.

2.1 Data overview

2.1.1 Recruitment time series

The French Eel Management Plan states that “l’objectif du plan de gestion est de réduire la mortalité par pêche […] par rapport à un niveau moyen calculé sur des années récentes (2005-2007 par exemple)”, i.e it implements an exploitation rate target (Ministère de l’Ecologie, de l’Energie, du Développement durable et de l’Aménagement du Territoire et al., 2010). Theoretically, an exploitation rate is defined as the ratio of catches and abundance (here recruitment). The absolute recruitment is not available for eels, but the target is defined with respect to a reference period, which was extended to 2004-2008 values (Beaulaton et al., 2020). Therefore, a relative abundance index can be used. As such, to assess the past trend in exploitation rates with respect to the management target, a relative recruitment and a catch data time series are required (at least since the reference period) (Beaulaton et al., 2020).

In France, most existing recruitment time series were fishery-based and therefore interrupted after the implementation of the quota system. However, recruitment trends in Europe proved to be spatially consistent historically. Hence, we used a recruitment index derived by the ICES/EIFAAC/GFCM WGEEL at the European scale and postulated that it was a good indicator of recruitment trend in France. This indicator relies on a GLM (Generalized Linear Model) fitted on 52 monitoring time series collected throughout Europe (ICES, 2021b). Since trends in the North Sea area are slightly different from trends elsewhere in Europe, a year:zone interaction is used in the ICES GLM, leading to two different indices, so called “North Sea” and “Elsewhere Europe”. In our analysis, we used the “Elsewhere Europe” index over the period 1980 to 2019 (Table 1) which covers France (Beaulaton et al., 2020). This recruitment index describes the past evolution of recruitment. Combined with catch data, it allows reconstructing the past trend of exploitation rate.

Table 1. Recruitment index series used (“Elsewhere Europe” series, ICES WGEEL 2019, base 100 for the 1979-1980 season). A year t corresponding to the t-1 - t recruitment season. Each row corresponds to a decade and each column corresponds to a year in each decade.

	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1980	100.0	72.8	79.8	43.0	47.4	45.6	29.8	51.8	62.3	39.5
1990	30.7	14.9	20.2	22.8	21.1	28.1	21.9	36.8	14.0	19.3
2000	16.8	6.9	11.5	11.0	6.0	6.7	4.8	5.5	5.0	3.9
2010	3.9	3.2	4.5	6.5	11.1	6.0	7.5	7.4	7.8	5.3

To set a quota, it is also necessary to have a model to predict the recruitment in upcoming years (Beaulaton et al., 2020). Therefore, different models were used to predict recruitment and to calculate quotas.

2.1.2 Generalities about the usual expertise procedure

In France, the Scientific Committee in charge of making recommendations for the quota for fishing season t+1 – t+2 (2020-2021 in this internship) generally meets in early summer of year t+1. On the other hand, the WGEEL (here 2020) only takes place in September so that the WGEEL for season t – t+1 is not available yet, and the last available recruitment index corresponds to season t-1 – t (here 2018-2019). Therefore, to estimate the TAC (Total Allowable Catches) for season t+2 (2020-2021), we need to predict recruitment for this season, but also for the fishing season t+1 (2019-2020). To address this issue, the Scientific Council collects feedback on the ongoing season (here 2019-2020) from professional fishermen, migratory associations, environmental inspectors, foreign scientists and the Council's monitoring. The feedback allows building a minimum and a maximum values for the recruitment, subsequently used as a censor interval in predictions models. For the season 2019-2020, the Scientific Council assumed that the recruitment was equivalent to the previous season (2018-2019, i.e. 5.3) but with a large uncertainty interval (2.3 and 8.3, i.e. 5.3 ± 3) (Beaulaton et al., 2020). Based on the WGEEL recruitment index series from season 1979-1980 to season 2018-2019 and on the "qualitative" expert feedback for season 2019-2020, the use of models allows predicting recruitment for season 2020-2021 (and also for season 2021-2022).

All models predicting glass eel recruitment were designed in the Bayesian framework. These models have a yearly timestep, as a season lasts one year. They were fitted on the WGEEL recruitment index series from 1979-1980 to 2018-2019 considering the censor season mentioned above for season 2019-2020. In order to facilitate the reading, when we mention "year t" or "season t", it refers to the season t-1 – t. As an example, when we mention 1980, it refers to the season 1979-1980.

2.2 Existing glass eel recruitment models: single-trend and two-trend models

2.2.1 Presentation of the two existing recruitment models

The single-trend model of recruitment was developed in 2012. The single-trend model postulates a decreasing exponential trend in recruitment. Thus, there is a unique recruitment slope coefficient a (in the log scale). In order to represent the fact that a year of recruitment above the decreasing central trend is generally followed by a year of recruitment above the central trend, the perturbations are auto-correlated in time (Beaulaton et al., 2020). The mathematical formulation of the single-trend model is given by:

$$\textbf{(1)}\ \text{For t in } [1980; 2022],\ IR_t \sim LN(\mu_{IR_t}, \sigma_{IR}^2)$$

$$\text{With } \mu_{IR_t} = \mu_{IR_0} + a.t + \rho.\varepsilon_t$$

$$\text{and } \varepsilon_t = \log(IR_{t-1}) - [\mu_{IR_0} + a.(t-1)]$$

where IR_t is the recruitment index for year t following a log normal distribution centred on the mean μ_{IR_t} and of variance σ_{IR}^2, representing the residual random noise normally, independently, and identically distributed (idd) over time around the mean. Then, in log scale, $\mu_{IR_0}$is an initialization parameter of the recruitment index, a is the recruitment slope, ε_t is the annual auto-correlated random disturbance and ρ its auto-correlation coefficient.

Due to a possible change of trends in recent years, the Scientific Council proposed a second recruitment model in 2014. This new model implements two distinct recruitment trends over time and a single disturbance structure. The year of change in recruitment trends is set in 2012, as it is the first year in which implementation of management plans could have had an effect on recruitment (Beaulaton et al., 2020). Since we consider two recruitment trends, there are two recruitment slopes a_1 (recruitment slope before 2012) and a_2 (recruitment slope after 2012). The two-trend model is defined as follows:

(2) For t in [1980; 2022], $IR_t \sim LN(\mu_{IR_t}, \sigma^2_{IR})$

For t < 2012 (time period 1), $\mu_{IR_t} = \mu_{IR_0} + a_1.t + \rho.\varepsilon_{t_1}$ and $\varepsilon_{t_1} = \log(IR_{t-1}) - [\mu_{IR_0} + a_1.(t-1)]$

For t ≥ 2012 (time period 2), $\mu_{IR_t} = \mu_{IR_{2011}} + a_2.(t-2011) + \rho.\varepsilon_{t_2}$ and $\varepsilon_{t_2} = \log(IR_{t-1}) - [\mu_{IR_{2011}} + a_2.(t-2011)]$

The description of variables is the same as for the single-trend model (see Equation 1 above). There are now two recruitment slopes a_1 and a_2 in log scale. Then, ε_{t_1} and ε_{t_2} represent annual auto-correlated random disturbances depending on time periods (before or after 2012) and ρ the auto-correlation coefficient. In this new model, a new variable $diff_a$ was introduced as the difference between the two slopes a_2 and a_1:

(3) $diff_a = a_2 - a_1$ (in log scale).

Since recruitment slopes are defined in the logarithmic scale, they correspond to yearly multiplicative coefficients of the recruitment in the linear scale. Indeed, exp(a), exp(a_1) and exp(a_2) correspond to the yearly rate of change of the recruitment (i.e. exp(a) = 1.15 means that in average, recruitment is 15% higher than recruitment of the previous year). As such, exp($diff_a$)=exp(a_2-a_1)=exp(a_2)/exp(a_1) corresponds to the relative change of yearly rates of change between time periods 1 and 2.

2.2.2 Comparison and limits of these two recruitment models

The single-trend model postulates that the decreasing trend has remained constant since the early 80s, as such it can hardly detect recent changes that may have occurred after the implementation of the Eel Management Plan. On the other hand, the two-trend model imposes a change of slope specifically in 2012. Therefore, these two models present different weaknesses. Moreover, the existence of two concurrent models providing two diagnoses raises the question of how to objectively "weigh" them. Having a model that internally estimates the existence of a change of slope in 2012 would be worthwhile. Second, both models postulate the existence of a limited number of trends and that the potential shift took place in 2012. It would be interesting to have a more flexible model that addresses the potential existence of more fluctuations in trends. To overcome these limitations, we explored alternative models.

2.3 Development of two alternative glass eels recruitment models

As previous models, both new models present a yearly timestep. They have a state-space structure meaning that (i) states (average recruitment) at timestep t is a Markovian process that depends on state at time t-1, and that (ii) observations (WGEEL index) are derived from the states at every timestep.

2.3.1 The spike and slab model

2.3.1.1 *The spike and slab method*

The spike and slab method is a Bayesian selection technique used to select the most relevant predictors of interest (Drouineau et al., 2017; George & McCulloch, 1993; Titsias & Lázaro-Gredilla, 2011). It is based on the construction of priors with two nested variances thanks to the introduction of a new indicator variable called γ (George & McCulloch, 1993). This variable γ is a binary variable whose value indicates whether the explanatory variable of interest (in our case the existence of a second slope) should be included in the model ($\gamma = 1$) or not ($\gamma = 0$). For example, let's denote β a parameter of interest, typically, a regression coefficient of an explanatory variable that we want to consider for inclusion in the model. The prior follows a normal distribution with mean 0 and variance as a function depending on γ and two coefficients *a* and *b*:

$$\textbf{(4)}\ \beta \sim \mathcal{N}(0, \sigma^2)$$
$$\sigma^2 = \mathrm{a}.(1-\gamma) + \mathrm{b}.\gamma \text{ with } \mathrm{a} < \mathrm{b}$$
$$\gamma \sim \mathrm{Ber}(0.5)$$

When γ equals 0, the variance is small so the prior for β is very narrow around 0 favouring posterior distributions centred around 0. As such, if β is indeed a regression coefficient, the variable of interest is not included in the model. On the other hand, when γ equals 1, the variance is large leading to a less informative prior, allowing for a wide range of values significantly different from 0 for β. As such, posterior distributions of γ and β will indicate whether the variable should be included (majority of $\gamma = 1$ and $\beta \neq 0$) or not (majority of $\gamma = 0$ and β close to 0). With such a prior, the resulting model nests both the model with and the model without the variable, with relative weights provided by γ.

2.3.1.2 *Recruitment model formulation*

The spike and slab prior appears relevant to assess objectively the existence of a change in recruitment slope. It would provide insights on whether a significant slope change took place in 2012 and provide weighted predictions in an objective way. As for the two-trend model (Beaulaton et al., 2020), we postulated that a potential slope change appears in season 2011-2012.

In the two-trend model, priors were assigned to a_1 and a_2 and the difference in slope $diff_a$ was defined as $a_2 - a_1$ (Equation 3). In the spike and slab model, the difference in slopes keeps the same biological meaning as in the two-trend model. However, instead of setting priors to a_1 and to a_2, we set prior to a_1 and to $diff_a$ and introduced a_2 as an expression of $diff_a$ and a_1 (Equation 5 below):

$$\textbf{(5)}\ \mathrm{diff_a} \sim \mathcal{N}(0, \sigma^2_{diffa})$$
$$\sigma^2_{diffa} = \mathrm{c}.(1-\gamma) + \mathrm{d}.\gamma \text{ with c = 0.0001 and d = 0.03}$$
$$\gamma \sim \mathcal{B}\mathrm{er}(0.5)$$

With a small variance σ^2_{diffa}, the difference of slope $diff_a$ is then restricted around 0, meaning that the slope of the second period a_2 is close to the slope of the first period a_1. Therefore, the slope change is rejected and the resulting model is almost similar to the single-trend model. On the other hand, a large variance σ^2_{diffa} increases the range of possible distinct values for $diff_a$. A significant change in slope recruitment is then possible and the resulting model is almost similar to the two-trend model. Basically, the spike and slab model is a combination of the single-trend and the two-trend models, making a weighted average of these two.

2.3.1.3 *Choosing variance's coefficients values*

Tuning the coefficients *c* and *d* in the expression of the variance σ^2_{diffa} (Equation 5) is a first important step and can be considered subjective (George & McCulloch, 1993). As mentioned earlier, the prior for $diff_a$ is bivariate with a narrow normal prior around 0 if the change of slope is rejected, and a wide normal prior (large variance) if the change of slope is accepted. Parameters *c* and *d* tune these two normal distributions. As such *c* (variance of the narrow normal) should be chosen so that the prior is restricted to values that are small enough to be considered as an insignificant change. Conversely, *d* (variance of the wide normal) should be large enough to allow significant change of slope (i.e. large $diff_a$) but small enough to eliminate impossible values (otherwise, the prior would be too flat leading to too small differences of density of probabilities between the two normal distributions). These priors will allow the model to select the relevant explanatory variables in a model in an autonomous way.

Remind that exp($diff_a$) can be interpreted as a rate of change between time periods 1 and 2. If there is no change in the recruitment trend, the multiplicative coefficient between exp(a_1) and exp(a_2) should be close to 1. As a first guess, we assumed that a variation of less than 2% around 1 is considered as a non-significant change in the recruitment trend. Just as an illustration, Dekker (2019) estimated that the recruitment dropped consistently by approximately 15% per year. This would correspond to exp(a_1) = 0.85 and to reject a shift for 0.85*0.98 = 0.83 ≤ exp(a_2) ≤ 0.85*1.02 = 0.87. Restricting a change of less than 2% between time periods 1 and 2 implies that $diff_a$ is restricted between log(0.98) = -0.02 and log(1.02) = 0.02. A normal distribution with standard deviation 0.02/1.96 = 0.01 would mean that 95% of $diff_a$ values are between -0.02 and 0.02. We therefore chose $c = (0.01)^2 = 1.10^{-4}$.

In recent years, the rate of decrease may have changed and the recruitment tends to be stable or slightly increasing according to the WGEEL. If the recruitment decreased by about 15% in the first period (Dekker, 2019), stabilising the trend would correspond to a change of 1/0.85 = 1.17. It is very unlikely that recruitment is now increasing as a rate similar to the former rate of decrease, so 1.15/0.85 = 1.35 appears to be a reasonable upper bound for a situation of significant change in the recruitment trend. This corresponds to a $diff_a$ of log(1.35) = 0.30. A normal distribution with standard deviation 0.30/1.96 = 0.15 means that 95% of $diff_a$ values are ranging from -0.30 to 0.30. As a result, we chose $d = (0.15)^2 = 3.10^{-2}$.

The values of parameters c and d have been chosen to provide sufficiently distinct probability distributions in order to distinguish situations of change in slope from situations of previous slope conservation. Indeed, situations where $\gamma = 0$ should not be accepted as a change of slope, and reciprocally. In order to evaluate model sensitivity to $diff_a$ prior, several pairs of values of c and d have been tested looking at the proportion of $\gamma = 1$ depending on these values.

2.3.2 Recruitment slope as a time-varying random variable

The three previous models only consider a single change of recruitment trend in a specific year, or no change at all. In order to enhance model flexibility, we postulated that recruitment slope is itself a time-varying random variable. Local linear trend models are largely used to model and predict macroeconomic time series (Delle Monache & Harvey, 2012). In such models, the slope follows a random walk from year t-1 to year t, as such, the slope can vary at any time. Therefore, the random slope model is defined as follows:

$$\textbf{(6)}\ \text{For t in } [1980;2022]\ ,\ IR_t \sim LN(\mu_{IR_t}\ ,\sigma_{IR}^2)$$
$$\mu_{IR_t} = \mu_{IR_{t-1}} + a_t + \eta_t \text{ with } \eta_t \sim \mathcal{N}(0, \sigma_\eta^2)$$
$$a_t = a_{t-1} + \delta_t \text{ with } \delta_t \sim \mathcal{N}(0, \sigma_{slope}^2)$$

where IR_t is still the recruitment index for year t following a log normal distribution centred on the mean μ_{IR_t} and of variance σ_{IR}^2, representing the residual random noise normally, independently, and identically distributed (idd) over time around the mean. Then, in log scale, the variable μ_{IR_t}, corresponding to the average value of the recruitment in year t, follows a random walk and depends on the average value of recruitment of the previous season t-1, on the recruitment slope a_t, which itself follows a random walk, and on η_t, which is the residual random noise relative to μ_{IR_t} normally, independently, and identically distributed (idd) over time with mean 0 and standard deviation σ_η. The recruitment slope a_t depends on the recruitment slope of the previous season and δ_t, which is the residual random noise relative to recruitment slope a_t normally, independently, and identically distributed (idd) over time with mean 0 and standard deviation σ_{slope}. No autocorrelation term was introduced in this model.

2.4 Bayesian inference and priors

The Bayesian models were fitted using JAGS (Plummer, 2003). We used the runjags package as an interface between R and JAGS (Denwood, 2016). Three chains were run in parallel with 60,000 iterations with a thinning period of nine (resulting in 20,000 samples per chain), after a burn-in period of 200,000 iterations. As the random slope model took longer to reach convergence, we set a thinning period of ten to decrease the autocorrelation between samples. Models' convergence was checked using the usual Gelman diagrams (Gelman & Rubin, 1992) and with graphical verifications thanks to density plots and traceplots.

Uninformative priors were used on most parameters (Table 2 below). The spike and slab priors were used according to lines 9 and 10 of Table 2. The priors specific to the random slope model are described in lines 11 and 12 of Table 2.

Table 2. Priors on parameters (all models included).

	Variable	Recruitment models: Single-trend	Two-trend	Spike and slab	Random slope
1	μ_{IR_0}	N(0,0.01)			
2	σ_{IR}	Gamma(0.01,0.01)			
3	ρ	U(-1,1)			
4	η	N(0,1)			For t in [1980 ;2022], $\eta_t \sim N(0,\sigma_\eta^2)$
5	σ_η				Gamma(1,1)
6	a_0	N(0,0.01)			N(0,0.01)
7	a_1		N(0,0.01)		
8	a_2		N(0,0.01)		
9	$diff_a$			N(0, $\sigma_{diff_a}{}^2$)	
10	γ			Ber(0.5)	
11	δ_t				For t in [1980 ;2022], $\delta_t \sim N(0, \sigma_{slope}^2)$
12	σ_{slope}				Gamma(1,1)

2.5 Models comparison strategy to keep only the most reliable model

We resulted with a set of four models. We now describe how we compared the performance of the different models.

2.5.1 DIC values comparison across models

The DIC (Deviance Information Criterion) is a model selection criterion allowing comparing models's performance. It quantifies the trade-off between both goodness of fit and model complexity represented by a measure of the number of effective parameters (Berg et al., 2004; Spiegelhalter et al., 2002, 2014). DIC is calculated as follows (Spiegelhalter et al., 2002):

$$\textbf{(7)}\ DIC = \overline{D} + p_D$$

As such, it quantifies the goodness of fit of the model, penalized by model complexity. However, the DIC only assesses the quality of the fit on historical data, but not the prediction ability of the model. Thus, relying on the DIC only is not sufficient to select a model that will subsequently allow establishing a quota. Therefore, it appears necessary to consider other metrics in order to evaluate the predictive capacity of each recruitment model. Hence, we implemented new models comparison strategy in order to keep only one recruitment model. To do so, we kept looking at DIC values and relied on new comparison criteria.

2.5.2 New metrics to quantify the predictive performance and robustness of models

2.5.2.1 *Credibility intervals' amplitude*

All models provided a posterior distribution for the recruitment in season t+1 – t+2 (here 2020-2021) for which the quota should be set. The width of the 95% credibility intervals was studied to assess prediction precision of each model. Indeed, in order to have the most accurate estimate on the recruitment of season t+1 – t+2 (and thus the most adapted quota to the state of the resource), a model should provide unbiased posterior distribution, but also ideally, narrow credibility intervals.

2.5.2.2 *Retrospective and simulation exercises*

Assessing the predictive capacity of the models required having a dataset in which the "real" recruitment is known and can be compared to models' predictions. In this context, we carried out two exercises:

- **A retrospective analysis**

In this exercise, we fitted all models on the WGEEL recruitment index up to season t, assuming that recruitment in season t+1 falls in an interval ranging from -3 to +3 around recruitment of season t (interval censorship). This mimicked the strategy used by the Scientific Council (Beaulaton et al., 2020). Then, we compared the model predictions for season t+2 with the time-corresponding "true" WGEEL value. We carried out this exercise for t ranging from 2013 to 2017, since the regime shift is supposed to take place in 2012 and the last WGEEL index is in 2019. For the retrospective analysis, three chains were run in parallel with 60,000 iterations with a thinning period of ten (resulting in 20,000 samples per chain), after a burn-in period of 200,000 iterations.

- **A simulation exercise**

Based on a similar strategy, we simulated surrogate time series of recruitment until 2030 with three different scenarios. Scenario 1 postulated that $diff_a = 0$. This corresponded to using the single-trend model to simulate future data. Therefore, the single-trend model was used to simulate 100 recruitment index series from 1980 to 2030. On the other hand, scenario 3 mimicked a situation in which the two-trend model was the most relevant. Thus, the two-trend model was used to simulate 100 recruitment index series from 1980 to 2030. Finally, scenario 2 was the intermediate situation between scenarios 1 and 3: we set $diff_a$ as the median of the median of a_1 and the median of a_2 (medians of a_1 and a_2 estimated with the two-trend model) to simulate 100 recruitment series. Then, following the retrospective analysis strategy, each model was fitted on a surrogate time series of recruitment up to season t. We assumed that recruitment in season t+1 falls in an interval ranging from -3 to +3 around recruitment in season t. We compared the model's predictions for season t+2 to the surrogate time series value. Therefore, we carried out this exercise for t ranging from 2017 to 2028. For this prospective analysis, three chains were run in parallel with 1,500 iterations with a thinning period of ten (resulting in 500 samples per chain), after a burn-in period of 200,000 iterations. For the three scenarios, we applied these parameters.

In both exercises, the convergence of models was checked using the usual Gelman diagrams (Gelman & Rubin, 1992) and with graphical verifications thanks to density plots and traceplots.

2.5.2.2.1 Evaluation of model prediction performance: the recruitment prediction error

We compared models' prediction performance by quantifying the discrepancy between recruitment predictions and "true" values thanks to a cross-validation (Vehtari et al. 2017) through a variable called the "prediction error". Regarding the retrospective analysis, for each year from 2015 to 2019, 60,000 iterations were made. Thus, 60,000 recruitment predictions were available per model per year. Each recruitment prediction was compared to the WGEEL value of the corresponding year. The difference between each predicted recruitment value and the WGEEL value was calculated and named as the "prediction error". If the prediction error was negative, the recruitment was underestimated (respectively, positive prediction error translated recruitment overestimation). The same principle was applied to the 2019 to 2030 simulation exercise except that, for computational reasons, the "yearly prediction error" was estimated as the difference between the median of the posterior distribution and the "true" value of the surrogate time series, instead of using the raw 1,500 iterations. The median of yearly predictions errors was then computed for each surrogate time series and model, as a criterion of overall prediction performance. As such, each model had 100 overall prediction errors for each scenario.

2.5.2.2.2 Evaluation of risks for both exploitation and resource: two comparison criteria

Then, we looked at the consequences regarding the risk of management target achievement and fishery losses by quantifying situations of over- and under-estimation of recruitment. We developed two quantitative criteria to lead a quantitative risk assessment (Aven & Renn, 2009). We defined risk notion as an "uncertainty about and severity of the consequences (or outcomes) of an activity with respect to something that humans value" (Aven & Vinnem, 2007) or, in other words, the probability of an event occurring and the stake associated with that occurrence (Hansson, 2004, 2005; Kermisch, 2012). We formulated these criteria thanks to the two following answers.

- **If recruitment is overestimated in year t+2, how often is it overestimated it if a certain model (among the four) is used to predict recruitment index?**

For each year and model, we computed the frequency of samples from the posterior distribution of recruitment that were greater than the corresponding WGEEL index value for the retrospective analysis or the value from a simulated surrogate time series for the simulation exercise. This frequency of recruitment overestimation represents how often recruitment is overestimated when using a certain model to predict it. It quantifies the risk of not achieving the management target. Therefore, we called this frequency of recruitment overestimation the "Conservation Risk Indicator" (CRI).

For the retrospective analysis, let $S_i(t,m)$ be a logical indicator with i ranging from 1 to the number of iterations (60,000 in this study) for the year t for one model m. If the difference between a recruitment prediction and the value of the WGEEL was positive, then the indicator $S_i(t,m)$ took the value 1, indicating a situation of overestimation of recruitment. Respectively, when the difference was negative, $S_i(t,m)$ took the value 0 and represented a situation of recruitment underestimation. We thus obtained a series of 1 and 0 of length 60,000. The following formula was then applied to obtain the frequency of overestimation of recruitment for one year t for one model m:

$$\textbf{(8)}\ \text{Overestimation frequency}(t,m) = \frac{\sum_{1}^{\text{number of iterations}} \text{Si}(t,m)}{\text{number of iterations}}$$

where number of iterations = 60,000.

Therefore, we obtained five frequencies of recruitment overestimation for each model for the retrospective analysis. We used a similar approach for the simulation exercise, except that we summarized the value by calculating the median of yearly overestimation frequencies for each model and surrogate time series. Therefore, for each model, we got 100 recruitment overestimation frequencies (one per surrogate time series).

- **What is the potential loss of catches if a certain model (among the four) is used to predict recruitment in year t+2?**

Underestimating recruitment would lead to implement too restrictive quotas and therefore generate potential catches losses for the fishery. Using the same approach as for the calculation of the recruitment prediction error (2.5.2.2.1), we computed the difference between predicted recruitment and the "true" value, except that if this difference was positive, we set the value at 0. Indeed, with a positive difference (i.e. an overestimated recruitment), there was no risk for the fishery to loose potential catches. This new indicator is called the "Fish Risk Indicator" (FRI).

In both retrospective and simulation exercises, we assessed the Conservation Risk Indicator (CRI) and the Fishery Risk Indicator (FRI) for each model to analyse the trade-offs between conservation and exploitation.

3 Results

3.1 The spike and slab model: validation of parameters *c* and *d* and sensitivity analysis

3.1.1 Prior and posterior distributions

Parameters c and d of the spike and slab prior should ensure that $\gamma = 1$ and $\gamma = 0$ lead to different posterior distributions. Figure 2 confirms that the posterior distribution is different from the prior distribution and that the conditional distribution of $diff_a$ with $\gamma = 1$ and $\gamma = 0$ are distinct.

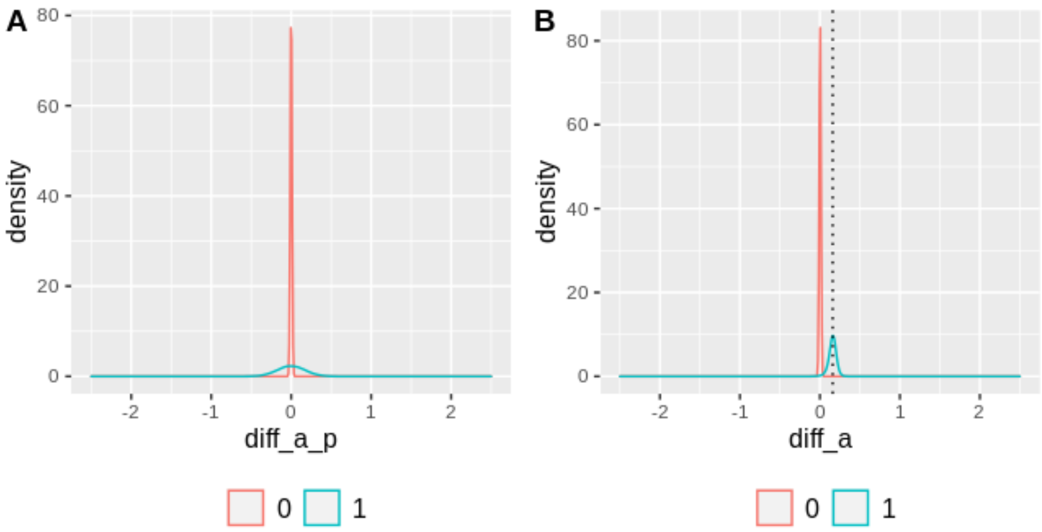


Figure 2. Density of probability for $diff_a$ (**A**: prior on the left; **B**: posterior on the right with vertical line x = median($diff_a$)) depending on γ values.

As recommended by Georges and McCulloch (1993), we also check that the intersection of the median value of $diff_a$ with the a priori distributions of $diff_a$ for $\gamma = 0$ and $\gamma = 1$ ensures an important difference (Figure 2).

3.1.2 Sensitivity analysis to *c* and *d*

Since the tuning of parameters *c* and *d* of the bimodal variance are partly subjective, we check the sensitivity of the model to these parameters by fitting the spike and slab model with different pairs of *c* and *d* values (Table 3).

Table 3. Values chosen and their meaning in terms of recruitment evolution to evaluate model sensibility to *c* and *d* choices (See section "*Choosing variance's coefficients values*" for details on the interpretation of parameters' value).

Tested values for c	**Percent of maximum decrease and maximum increase tolerated between two consecutive years**
4.10^{-8}	0.1%
5.10^{-6}	0.5%
2.10^{-5}	1%
1.10^{-4}	2%
2.10^{-4}	3%
Tested values for d	**Percent of maximum decrease and maximum increase tolerated between two consecutive years**
1.10^{-2}	20%
3.10^{-2}	35%
5.10^{-2}	55%

Then, we compare the proportion of $\gamma = 1$ in the γ posterior distribution (Figure 3). Whatever the values of *c* and *d*, the frequency of $\gamma = 1$ is greater than 90% (Figure 3). As such, the sensitivity of the model to *c* and *d* appears limited.

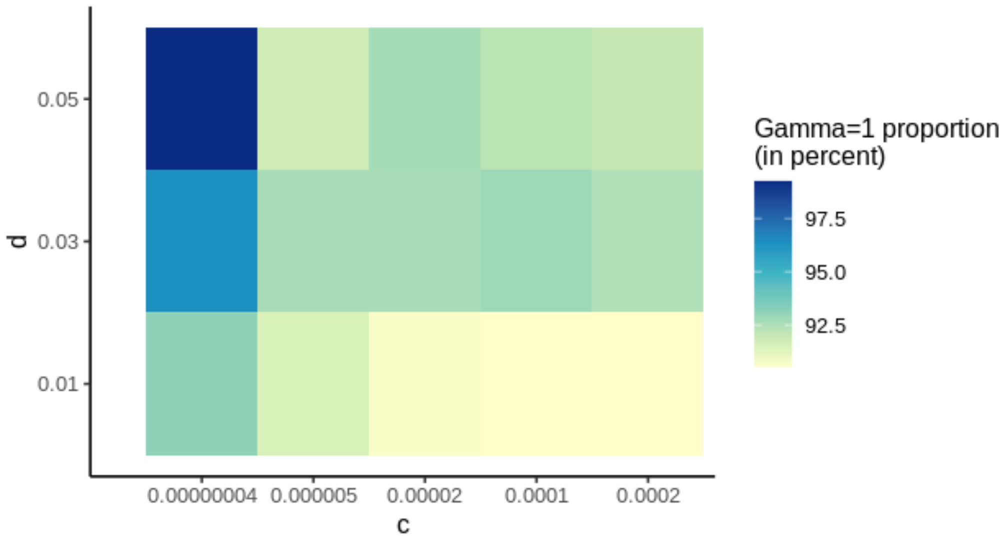


Figure 3. Sensitivity matrix of the model to values of *c* and *d* represented by the proportion of cases where γ is 1 (in percent).

3.2 The random slope model: problem of convergence

There is a risk of over parametrization with the random slope model. Indeed, the Gelman statistics and the density plots of posterior distributions confirms that some variables have not converged (see for example Figure 4 below).

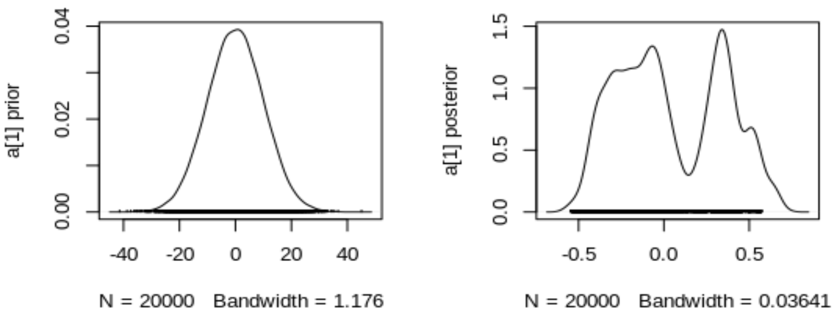


Figure 4. Prior (on the left) and posterior (on the right) probability distributions of exp(a[1])).

The Gelman diagnostics show that the parameters a_t has not converged for the first timesteps, but that convergence is achieved since season 1982-1983 and is complete for recent years, suggesting that it has limited impacts on predictions.

3.3 Models' comparison strategy

3.3.1 Evolution of recruitment predictions over time depending on model choice

3.3.1.1 *Evolution of recruitment slope over time depending on models*

We look at recruitment slopes depending on model (Table 4). The single-trend model estimates the lowest declining trend per year (5.82% per year). For both two-trend and spike and slab models, the slope in the second time period is not significantly positive. However, recruitment slopes a_1 and a_2 are significantly different for the two-trend model, while the difference is not statistically significant for the spike and slab model. Indeed, the two-trend model considers a change in slope ranging from an increase of 6% to 30% of recruitment slope, whereas the spike and slab model provides a change in recruitment slope ranging from a no-change to an increase of 27%. Moreover, an other difference between the two-trend model and the spike and slab model lies on the amplitude of the 95% credibility intervals of a_2 (0.16 and 0.21 respectively). The larger amplitude of a_2 95% credibility interval in the spike and slab model is due to the mix of cases rejecting or considering a change in slope.

Table 4. Recruitment slope coefficient (median value, 95% confidence interval and median value in linear scale translated in percent) and $diff_a$ (median value and 95% confidence interval in both log scale and linear scale). Values in bold are in linear scale, others are in log scale.

Models	Recruitment slope			$diff_a$
	a	a_1	a_2	
Single-trend	-0.06 [-0.11; -0.02] **~ decrease of 5.82% per year**			
Two-trend		-0.09 [-0.13; -0.06] **~ decrease of 8.61% per year**	0.07 [-0.01; 0.15] **~ increase of 7.25% per year**	0.17 [0.06; 0.26] **1.18 [1.06; 1.30]**
Spike and slab		-0.09 [-0.12; -0.05] **~ decrease of 8.61% per year**	0.06 [-0.08; 0.13] **~ increase of 6.18% per year**	0.15 [-0.004; 0.24] **1.16 [1.00; 1.27]**
Random slope	See Figure 5 below			

The yearly-varying random slope structure of the random slope model leads to slope that are never significantly different from 0 (Figure 5). The median recruitment slope value for predictions is -0.19 (in log scale, i.e. a decrease of 17% per year). Unsurprisingly, predicted slopes (for seasons 2020, 2021, and 2022) have larger credibility intervals than the other slopes. Since recruitment slopes for all models considered are never significantly different from 0 over the 1980-2022 time series, we cannot conclude about their respective trends.

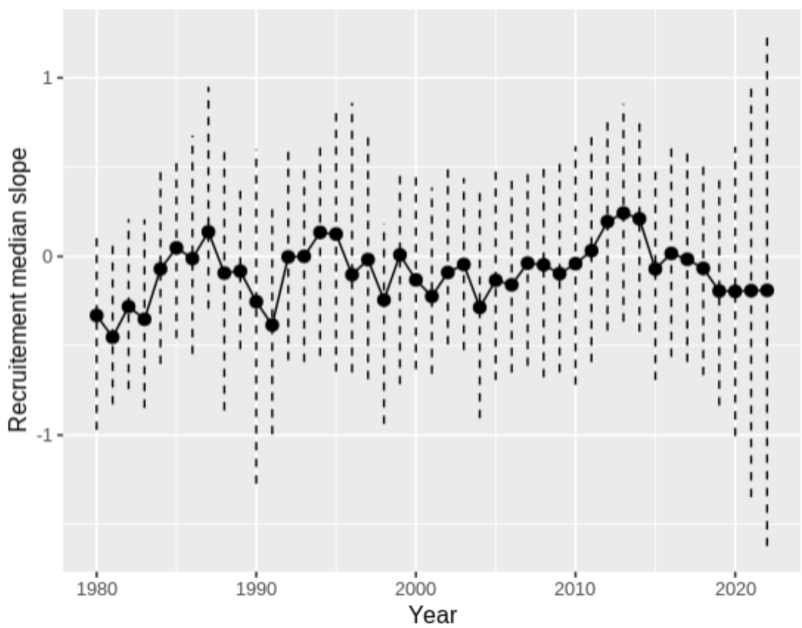


Figure 5. Evolution of the recruitment slope over time according to the random slope model (dotted bars represent the 2.5% and 97.5% quantiles).

3.3.1.2 *Recruitment evolution overtime depend on models*

As mentioned earlier and according to Figure 6, the single-trend model predicts a decreasing trend whereas the two-trend model predicts an increasing trend. Unsurprisingly because of the high frequency of $\gamma = 1$, predictions for the spike and slab model are similar to the ones provided by the two-trend model in terms of recruitment trend and of credibility intervals' amplitude. The random slope model predicts a decreasing trend, as the single-trend model, but with larger credibility intervals that encompass all models' credibility intervals. Credibility intervals are similar for the three other models.

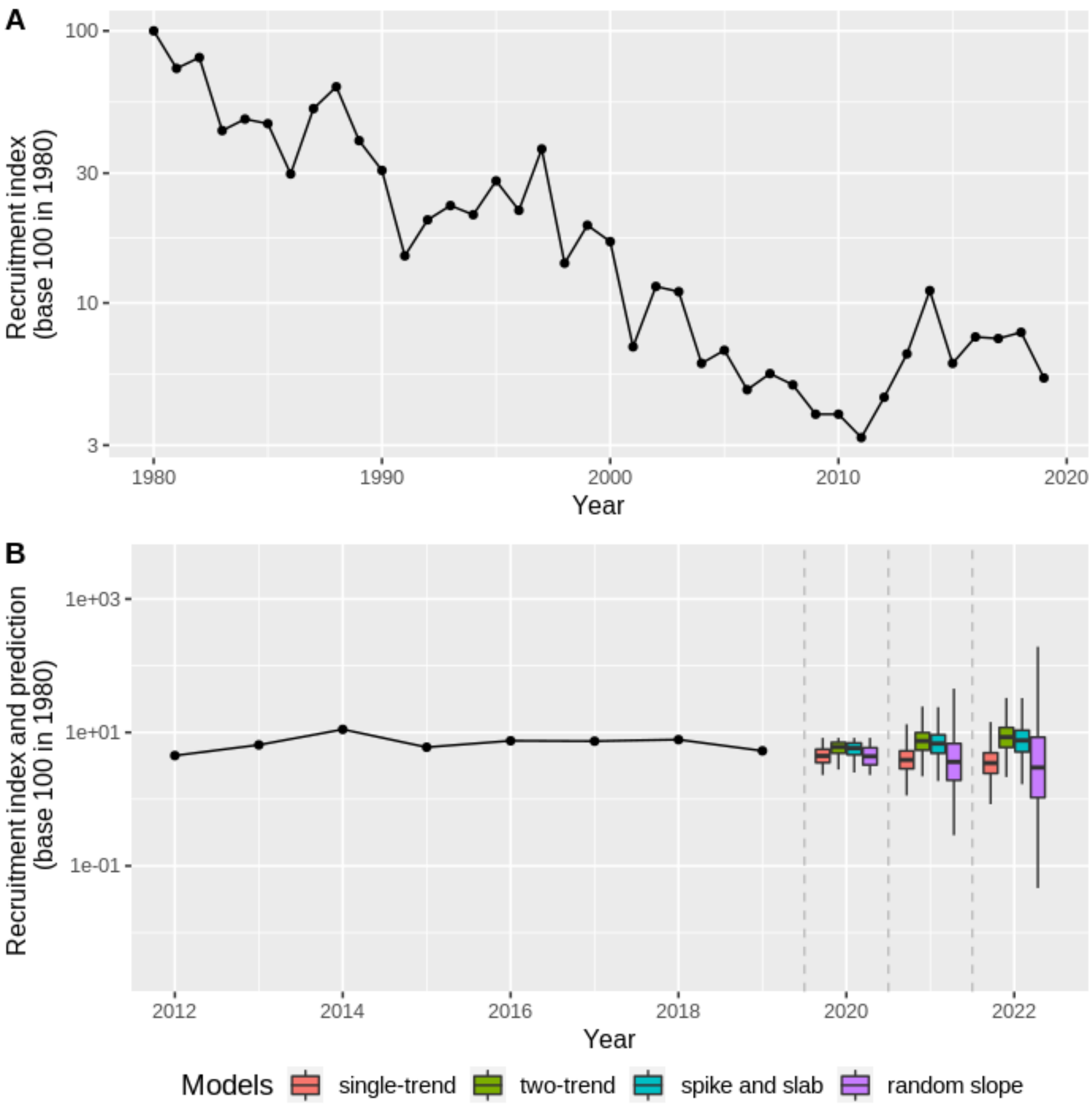


Figure 6. Trends in recruitment from 1979-1980 to 2018-2019 (A), and recruitment index since 2011-2012 followed by recruitment predictions distributions for the three following seasons for the four models (B). Boxplots show median, 1st and 3rd quartiles. Whiskers are 1.5 times the interquartile range. Log 10 scales are used to facilitate the reading.

3.3.2 Deviance Information Criterion

In order to compare the quality of the fit of models on data, we look at the DIC values for all models (Table 5).

Table 5. Deviance information criterion (DIC) calculated for each model with 1,000 iterations.

	Recruitment models			
	Single-trend	Two-trend	Spike and slab	Random slope
Mean Deviance	250.1	243.1	243.3	175.5
Penalized Deviance (DIC)	255.4	249.8	250.1	237.4

The single trend model, the two-trend model, and the spike and slab model have similar DIC values (255.4; 249.8; 250.1 respectively) according to Table 5. The best model from a statistical perspective is the random slope model (DIC = 237.4) despite its larger complexity.

3.3.3 Retrospective analysis from 2015 to 2019

3.3.3.1 *Distribution of prediction error depending on models*

In the retrospective analysis from 2015 to 2019, we first compared differences between models' predictions and corresponding values of the WGEEL (Figure 7).

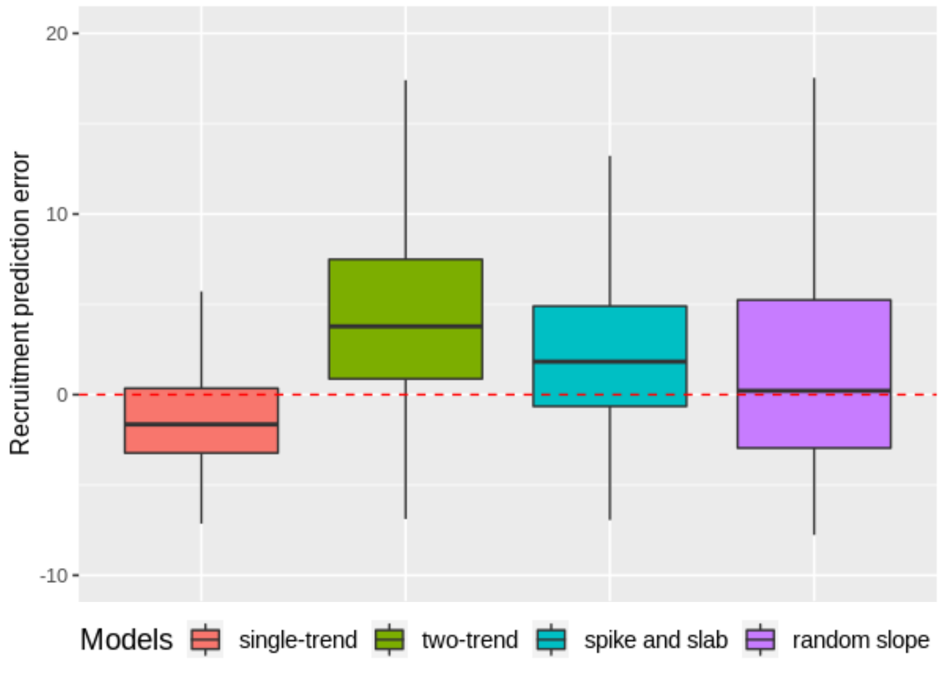


Figure 7. Boxplots presenting recruitment prediction error for the four recruitment models through the retrospective analysis from 2015 to 2019.

According to Figure 7, the single-trend model mainly underestimates recruitment in recent years compared to the ICES WGEEL recruitment index (median prediction error = -1.64 < 0, corresponding to 31% in absolute value of the latest 2019 WGEEL recruitment index i.e.

5.3). This is consistent with previous observations of the decreasing exponential trend in the model slope. Therefore, using the single-trend model to support quota setting would lead to too low quotas and potential loss of catches for the fishery, but would favour management target achievement. On the other hand, the two-trend model tends to be over-optimistic compared to WGEEL recruitment index (median = 3.92, representing 74% of the WGEEL 2019 recruitment index). From a management point of view, the risk is to set a too high quota and to miss the management objective set by the national management plan. With a median prediction error of 1.85 (representing 35% of the 2019 WGEEL recruitment index), the spike and slab model is the intermediate between the single-trend model and the two-trend model. This result is consistent with the structure of the model, which is a mixture of the two models. Nevertheless, Figure 7 shows that the spike and slab mainly overestimates recruitment. The random slope model appears to be the least biased model (median prediction error = 0.96 corresponding to 18% of the 2019 WGEEL recruitment value).

Finally, the most biased models are the random slope model, the single-trend model, the spike and slab model and the two-trend model in ascending order (Figure 7). According to prediction error results, the random slope model appears to be the less risky model for both resource and exploitation.

3.3.3.2 *Criteria for conservation and exploitation*

Since the Conservation Risk Indicator (CRI) represents compliance or non-compliance with the conservation rule and the Fishery Risk Indicator (FRI) traduces the potential loss of catches for the commercial fishery, Figure 8 illustrates the trade-off between conservation and fishery objectives.

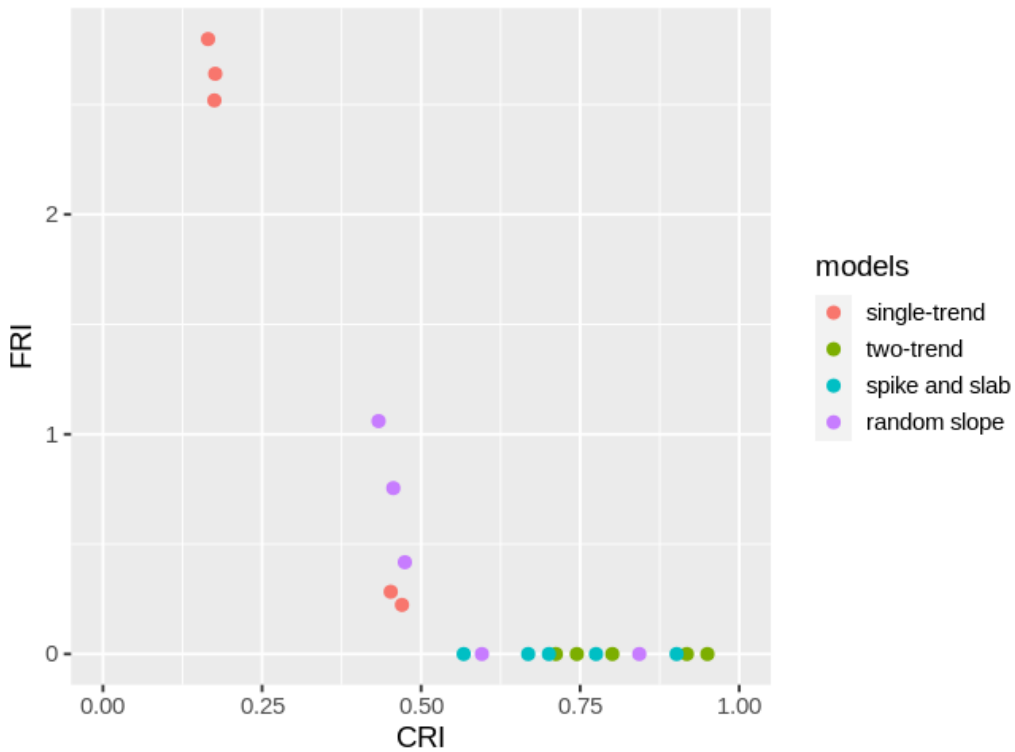


Figure 8. Position of the four recruitment models when considering simultaneously the Fishery Risk Indicator (FRI) and the Conservation Risk Indicator (CRI) during the retrospective analysis from 2015 to 2019.

An "ideal" model would prevent catches loss for the fishery and achieve the conservation target. However, it is impossible to always achieve both objectives. A model with a CRI of about 0.5 would randomly alternates between over- and under-estimation situations (since CRI is the recruitment overestimation frequency), as expected for a non-biased model. Ideally, a model should also have a FRI as small as possible (i.e. limiting too strong underestimation of recruitment and therefore restricting potential loss of catch). As such, the model that comes closest to this CRI = 0.5 and FRI = 0 point is the random slope model (Figure 8). Indeed, points relative to the random slope model are always close to the 50% overestimation frequency threshold and minimize catches loss compared to the single-trend model. Indeed, the single-trend favours compliance with the conservation rule, but at the expense of catches loss for the fishery. On the other hand, the two-trend and spike and slab models ensure limited catch losses but at the cost of a systematic recruitment overestimation.

3.3.4 Simulation exercise from 2019 to 2030

3.3.4.1 *Evolution of the simulated recruitment series*

As expected, the three scenarios provide different trends in recruitment over time: a decreasing trend (i.e. scenario 1), a stable trend (i.e. scenario 2) and an increasing trend (i.e. scenario 3) (Figure 9). However, it should be noted that the boxplots hide a large heterogeneity among individual simulated time series.

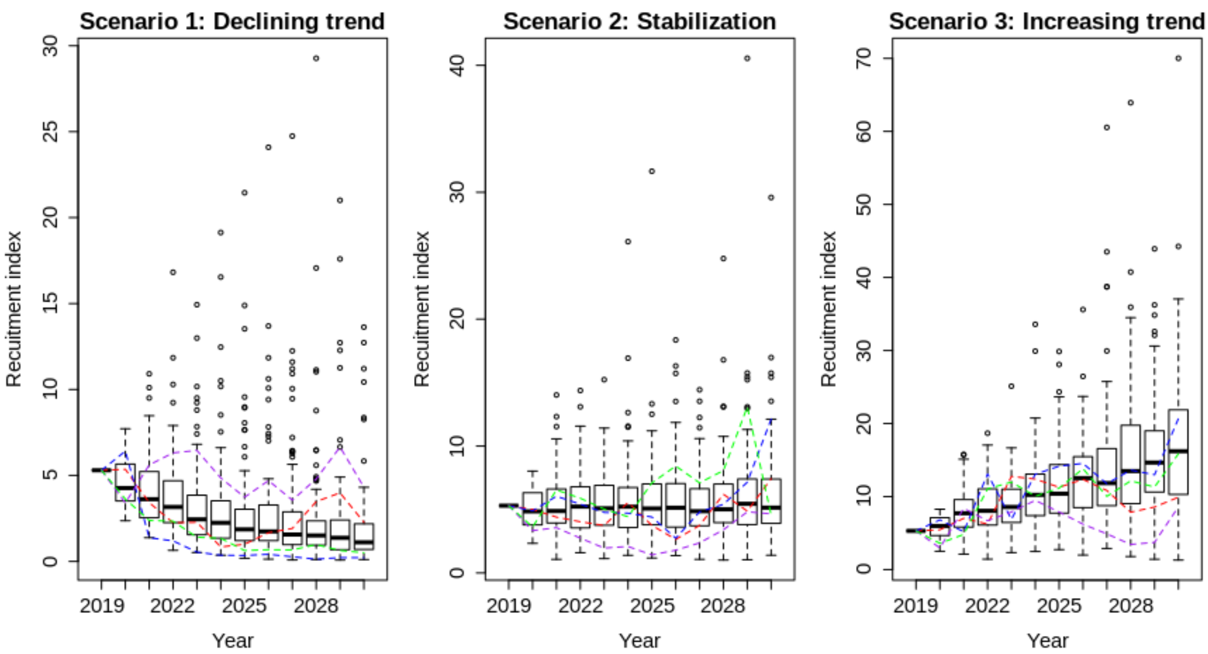


Figure 9. Evolution of the simulated recruitment index over time according to three scenarios (Scenario 1: Decrease in recruitment; Scenario 2: Stabilization of recruitment; Scenario 3: Increase in recruitment). The coloured lines (purple, red, green and blue) represent four time series of recruitment index out of the 100 total simulated in order to highlight the heterogeneity among individual trajectories.

3.3.4.2 *Error in recruitment estimations depending on model and scenario*

The recruitment prediction error varies among scenario and models, but the order among models remains almost similar across scenarios (Figure 10).

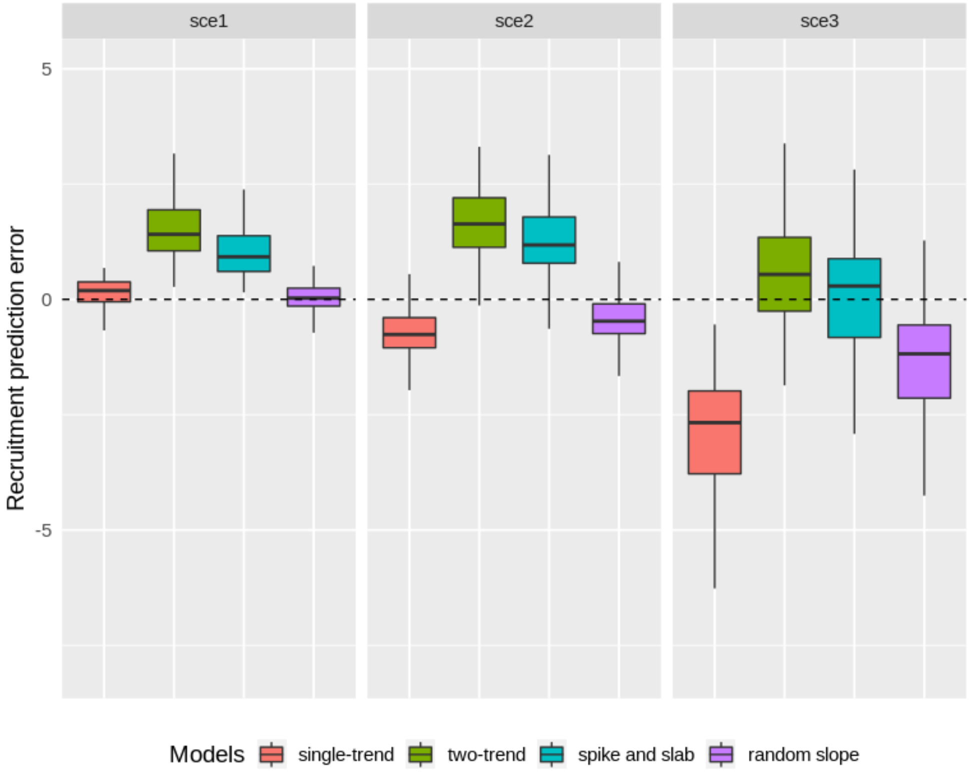


Figure 10. Boxplots presenting prediction error for the four models through the three recruitment evolution scenarios.

Regarding the single-trend model, recruitment prediction errors are close to 0 with scenario 1 (Figure 10). However, the single-trend model leads to increasing underestimation prediction errors with scenario 2 and 3. Indeed, the median of recruitment error prediction varies from 0.19 to -0.76 from scenario 1 to 2. These errors corresponds to 4% and 14% of the 2019 WGEEL recruitment index (i.e. 5.3). Then, this error decreases to -2.71 (represents 41% of the 2019 WGEEL recruitment index) in scenario 3. On average, the single-trend model underestimates recruitment by 1.1 (21% of the 2019 WGEEL recruitment index).

While the two-trend and the spike and slab models provide the best predictions with scenario 3, they strongly overestimate recruitment with scenarios 1 and 2. On average, over the three scenarios, the two-trend model overestimates recruitment by 1.2 (22% of the 2019 WGEEL recruitment index) and the spike and slab model overestimates recruitment by 0.79 (15% of the 2019 WGEEL recruitment index).

The error in recruitment predictions remains relatively similar among scenarios for the random slope model (median error of 0.27; -0.47; -1.18 for scenarios 1, 2 and 3 respectively). Although the prediction error is larger for scenario 3 and represents 22% of the final 2019 recruitment index, this error leads to an underestimation of recruitment. Therefore, there is no risk for the resource and no risk to set too optimistic quotas, favouring the achievement of the management target. On average, the error associated to recruitment estimation by the random slope model is -0.46 so it represents 9% of the 2019 WGEEL recruitment index. The random slope model appears to be more robust facing the different scenarios.

As a result, across all scenarios, the prediction error of the random slope model is the lowest among the four models and in most cases does not generate a risk of overestimating the resource.

3.3.4.3 *Integrating conservation and fishery criteria*

As discussed earlier in the retrospective analysis, the optimal model has a CRI of about 0.5 and a FRI as small as possible. As such, for scenario 1, the single-trend model and the random slope model are the two best models looking at Figure 11. However, the random slope appears a bit better than the single-trend model, even though it was used to simulate the data. The two-trend and spike and slab models ensure limited catch losses (as FRI is around 0) costing a high risk for conservation.

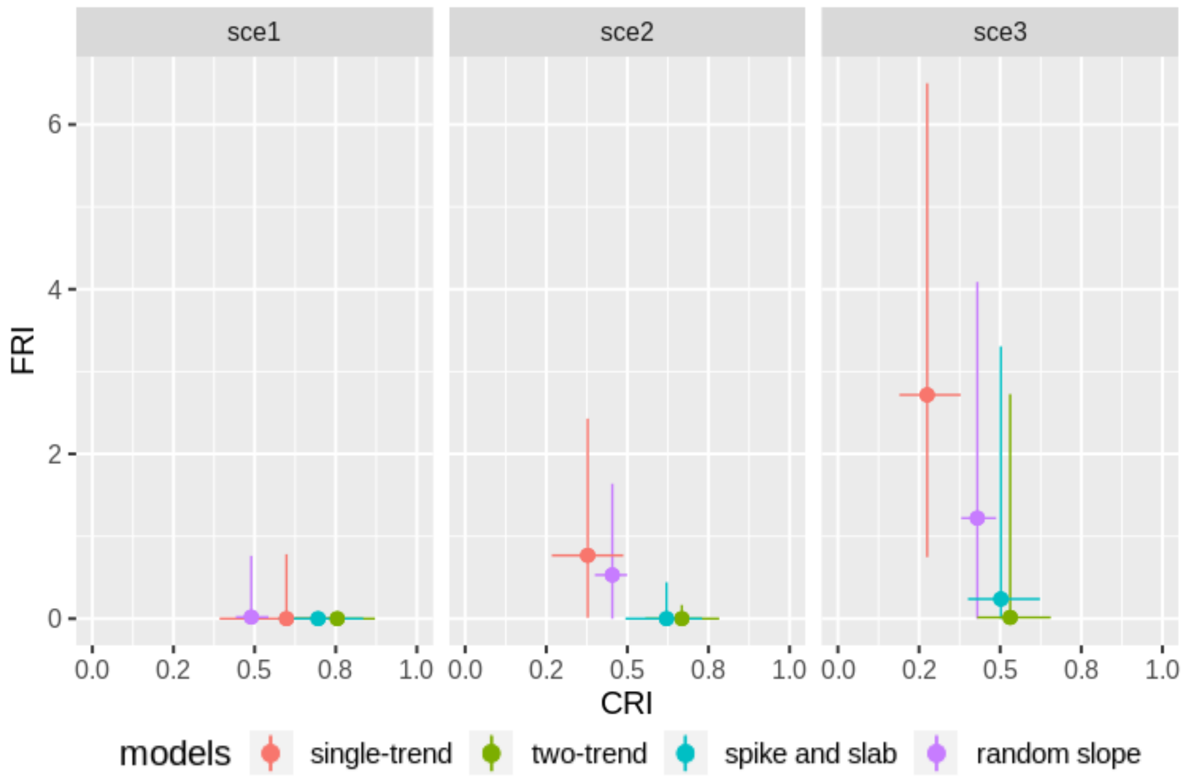


Figure 11. Distribution of the models when considering simultaneously the fishery risk indicator (FRI) and the conservation risk indicator (CRI) according to the three scenarios of simulation of recruitment from 2019 to 2030. The error bars represent the 2.5% and 97.5% quantiles. Each point represents the median values for both indicators (recruitment overestimation frequency and error in recruitment estimate when recruitment is underestimated, i.e. risk of catch loss) and "sce" stands for "scenario".

For scenario 2, the single-trend model provides slightly over-pessimistic predictions that lead to potential catch loss, while ensuring the achievement of the conservation target. The situation is similar with the random slope but slightly better. On the other hand, the two-trend and spike and slab models provide over-optimistic predictions, as such, there is limited risk of catch loss but increased risk of not achieving the management target.

Looking at scenario 3, the two-trend and spike and slab models have a range of possible CRI greater than 0.5. Therefore, both present a risk for the resource, while presenting a limited FRI. Both single-trend and random slope models present a low risk for the resource (CRI below 0.5). However, as the single-trend model presents a CRI around 0.25 and the highest FRI, it leads to substantial loss for the fishery, when the random slope model performs slightly better in terms of FRI. Finally, for any scenario in this simulation exercise, the model that comes closest to the point CRI = 0.5 and FRI = 0 is the random slope model.

4 Discussion

Within the framework of the French Eel Management Plan (Anonymous, 2018; Ministère de l'Ecologie, de l'Energie, du Développement durable et de l'Aménagement du Territoire et al., 2010; Onema et al., 2010), this internship aimed to propose new methods improving the tools used to support glass eel quotas setting. Among others, such tools require models that predict glass eel recruitment. Nevertheless, recruitment is particularly stochastic and, therefore, very hard to predict in fisheries science (Subbey et al., 2014). Indeed, the recent failure in achieving the requested 60% reduction in glass eel fishery mortality (Beaulaton et al., 2020) raised questions about the suitability of the existing tools. In this context, we proposed models to overcome previous models' limits and we suggested a strategy to compare models' performance.

4.1 Contributions and limitations of the new models

4.1.1 Position in relation to existing models in the literature

4.1.1.1 *Previous models and their limitations*

Several estimates of fishing mortality or exploitation rates for eel fisheries have been proposed in the past (Adam et al., 2008; Aranburu et al., 2016; Beaulaton & Briand, 2007; Briand et al., 2003, 2012; Bru et al., 2009; Dekker, 2000b). Dekker proposed in 2000 a procrustean model at the population scale, which was based on strong assumptions such as (i) silver eel catch rates and life cycle duration similarity over the whole distribution area, (ii) error-free knowledge of catches and (iii) a stable state for the fishery (Dekker, 2000b). Then, various projects focusing on glass eels have been developed to assess recruitment and exploitation rates at the local scale. As an example, the GEMAC model (Beaulaton & Briand, 2007) allowed the estimation of glass eel exploitation rate and recruitment but at a catchment scale only. However, results can hardly be extrapolated to larger spatial scales because of the fractal dimension of the population (Dekker, 2000a). Subsequently, the GEREM model (Drouineau, Briand, et al., 2016) allowed determining an absolute glass eel exploitation rate over the period 1980-2011 at the French territory scale, then throughout European range (Bornarel et al., 2018), exceeding the local scale and the fractal distribution of the European eel limitations. However, this model has a very simple structure through an unbiased random walk, which makes it unsuitable for predictions. Moreover, it is not updated regularly nowadays. As such, there was a need of new models specifically designed for predictions at the French scale to set quotas.

4.1.1.2 *Single-trend and two-trend models and their strong assumptions*

In this context, the single-trend model was developed in 2012, followed by the two-trend model in 2014 (Beaulaton et al., 2020). The single-trend and two-trend models rely on strong assumptions about the steadiness of recruitment trend over time. As recruitment is a highly stochastic ecological process (Subbey et al., 2014), assuming stable trend over long time periods seems inappropriate (Bogstad et al., 2000; Kimoto et al., 2007). For the European eel, while the decreasing trend had remained stable for many years (Dekker 2000b), assuming a stable long-terms trend for prediction may be unsuitable since of the recent modifications in the recruitment signal and because of restoration attempts (Beaulaton et al., 2020; Dekker, 2000a, 2008).

4.1.2 Spike and slab model

Spike and slab priors are widely used in Bayesian variable selection (Ishwaran & Rao, 2005; Titsias & Lázaro-Gredilla, 2011). The inclusion or not of a predictor is estimated internally in the model and the resulting predictions are a weighted average of the model with and without the predictor (Titsias & Lázaro-Gredilla, 2011). As such, it appears as a relevant solution to test the existence of a regime shift, as a way to objectively balance the predictions of single-trend and two-trend models. It also allows testing different years, distinct from 2012, as the year of change in recruitment trends. However, the parametrization of a spike and slab prior remains partially subjective and based on expertise to separate plausible from impossible values. Here, we checked that outcoming predictions were not too sensitive to this parametrization. This limited sensitivity is not surprising as situations with a slope change (i.e. $\gamma = 1$) highly dominate. However, this can change in the future, especially if recruitment starts decreasing and the weight of the single-trend model (i.e. the frequency of $\gamma = 0$) increases. Therefore, it would seem necessary to regularly perform the sensitivity analysis and to update the parametrization of the prior, if the spike and slab model is used in the future. Different slight modifications have been proposed to help in the parametrization of the spike and slab priors and to reduce resultant uncertainty (George & McCulloch, 1993; Ishwaran & Rao, 2005) that can be tested to improve the model. However, it would not entirely remove the subjectivity in the resulting model and the need to check the sensitivity.

4.1.3 Random slope model

A random slope model appears to be suitable to overcome the debate between a model with either a single or two slopes, and the year in which a potential regime shift occurs. Indeed, such models have proved to provide robust predictions, even in the presence of cyclic patterns or trend changes (Delle Monache & Harvey, 2012). In fact, it delivers promising results in our study. However, we noticed some convergence problems on the first timesteps. This convergence problem is due to a partial redundancy between the slope and the mean random walks, especially in a situation in which their variances are both free parameters (Equation 6). Delle Monarchy and Harvey (2012) listed solutions to address this problem. For example, they propose to constrain the ratio of the variances of these two random walks. Their example showed that, by setting this ratio, the random slope model can be made equivalent to a double exponential smoothing (Harvey, 1986) and yields robust predictions (Delle Monache & Harvey, 2012). Indeed, preliminary tests were made with such a random slope model with fixed ratio of variances (not presented here) and they confirmed that they address the convergence issues and yield smaller credibility intervals.

4.2 Models comparison strategy

Currently, the Scientific Committee uses two models that yields contrasted predictions, with no objective criteria to weight their results, except the use of the DIC (Beaulaton et al., 2020) which do not assess models' predictive ability. In this context, it appeared necessary to develop criteria evaluating models' performance. Therefore, to assess models' predictive capacity and its robustness, we set up an original models' comparison strategy through an innovative retrospective and simulation exercises.

4.2.1 Recruitment error prediction

In order to assess model's predictive ability, we were interested in the recruitment prediction error. In both retrospective and simulation exercises, the random slope model provides the lowest prediction error, and therefore the smaller risk for the resource and the fishery.

It would be interesting to look at the variability of the recruitment prediction error over time and to see if this recruitment prediction error is specifically high or low for certain years in the data set. This would allow us to investigate the causes of this error variability for these specific years. In addition, the prediction error is expected to decrease with the lengthening of the time series of recruitment index, if the trend remains stable over a given number of years. It would be worthwhile comparing the rates of decrease in recruitment prediction error of the four models. It should be noted that, in order to simplify the format and lengths of the outputs related to the simulation exercise, we worked on the medians of prediction errors. It would have been interesting to work directly on the complete distributions of the variables. Nevertheless, during the retrospective analysis, the work was carried out on both medians and complete distributions (presented in this report). The results were similar, suggesting that the results would not change much in the simulation exercise when studying complete distributions.

4.2.2 Resource and fishery criteria

Inspired by the work of Lebot (2021) and Bevacqua (2007), we developed two criteria to assess and illustrate the trade-off between management target achievement and loss of potential catches for the fishery (Bevacqua et al., 2007; Lebot, 2021). In the retrospective analysis, the single-trend model proves to ensure management target achievement but leads to severe catch losses. On the other hand, both spike and slab and two-trend models limit fishery losses but with a high risk of overshooting the management target. The random slope model appears to provide an intermediate solution because of its non-biased model predictions. While the trade-offs were contrasted depending on simulation scenarios for other models, the random slope model proved to be more robust across all scenarios. In the future, these two criteria could be co-constructed with the actors of the glass eel fishery sector (managers, fishermen, scientists...). This could help to enhance involvement of each stakeholder in the fishery management, strengthening trust and understanding between them (Drouineau et al., 2021).

4.2.3 Choice of an unique recruitment model

When using several models for making predictions, there are different strategies: either making multi-model averaging using, for example, the Bayesian model averaging method (Drouineau, Lobry, et al., 2016; Duan et al., 2007; Hoeting et al., 1999; Raftery et al., 2005) or choosing the most appropriate model. In a way, the spike and slab model provides a multi-model average of the single-trend and two-trend models (Titsias & Lázaro-Gredilla, 2011). It would perhaps be possible to make a model averaging of the spike and slab model and the random slope model, but the spike and slab model can already be seen as nested within the random slope model. Moreover, model averaging is often based on criterion such as AIC (Buckland et al., 1997; Burnham et al., 2002; Symonds & Moussalli, 2011) that, here, would be related to fit quality rather than to robust predictive capacity. Here, we rather chose to develop criteria to select the most appropriate model and to illustrate to managers the trade-off between conservation and exploitation. These criteria show that choosing one model would clearly favour either conservation or exploitation to the detriment of the other, except the random slope model that ensures a more balanced trade-off.

4.3 Missing points from a methodological point of view

4.3.1 Reliability of WGEEL data

Of the 52 European eel recruitment time series used by the WGEEL as indices, only ten are from scientific surveys while the other are fishery-based. Fishery data may report catches inaccurately and with bias and as it is subjected to variable external influences (markets, regulations and technology) (Aranburu et al., 2016; ICES, 2013). As advised by the WGEEL (ICES, 2013), it would be desirable to increase the number of recruitment series collected during scientific monitoring.

Moreover, in the single-trend, two-trend and spike and slab models, WGEEL data is considered accurate and known without error. In the random slope model, we assume that there is an estimation error, which is an improvement. However, the variance is assumed constant over time, while we could probably use the estimation errors provided by the GLM to have variances by year.

4.3.2 Exploitation rate

In this work, we were only interested in modelling the recruitment index. We could also have looked at the exploitation rate, which is another important factor of the catch process. Indeed, the Scientific Committee has developed a model to account for the variability in exploitation rates, and potentially, for the effect of the decrease in fishermen reduction (Beaulaton et al., 2020). Nevertheless, since this exploitation rate model is used independently of the recruitment model, it would have not modified our results. It would still be interesting to work on this aspect in the future.

4.3.3 Censor influence

When the Scientific Committee meets to make recommendations for quotas for the fishing season t+2, the WGEEL index value for fishing season t+1 is not yet available (considering the last quantitative data is relative to season t). As such, the Scientific Committee uses *ad hoc* expert feedbacks to build a censor interval for recruitment value in season t+1. Clearly, it would be important in the future to assess the sensitivity of model prediction to this censor interval. Moreover, it would be worthwhile implementing a more rigorous method to collect and integrate preliminary feedbacks on the ongoing fishing season t+1. While collecting data from all over Europe is not possible, as the recruitment in Northern Europe would end too late because of the spatial pattern in recruitment seasonality (ICES, 2020c), it may be possible to collect preliminary data from Southern Europe. This year, commercial fishermen have implemented questionnaires to collect qualitative feedbacks from fishermen. This interesting initiative can be used in the future. However, it might be possible to move one step further and to implement quantitative pre-indices based on the French observations (fishery, scientific monitoring). Moreover, working on such indices would be the opportunity to build new recruitment time series in France, which are currently scarce, since most of them were interrupted after the implementation of the Eel Management Plan.

4.3.4 Silver eel escapement

Because this work aimed to respond to questions about glass eel quota setting in the context of the French Eel Management Plan, we only focused on recruitment. However, it could be interesting to evaluate the consequences of varying catches levels in terms of escapement, to check the outcomes on the achievement of 40% escapement target enforced by the European Eel Regulation. Clearly, this is not straightforward for eel because of its complex life cycle, including density-dependent mortality (Bevacqua et al., 2011), but it would inform on whether the 40% escapement target can be achieved with a 60% reduction of glass eel fishing mortality.

This is even more important because the current recruitment is so low that it is unsure whether it is possible to achieve 40% escapement of silver eels, even without any fishing mortality. Especially since measures on other pressures (e.g. obstacles to migration, contamination) take a long time to implement and their outcomes can be delayed for many years, and it is hardly possible to mitigate some other pressures at least on the short term (e.g. climate change, parasitism) (Drouineau et al., 2018).

4.4 Perspectives for the stock management

4.4.1 Annual glass eel quota?

We have set our work in the context of the framework of the Ministry's request, where TACs are fixed annually. However, due to the variability of recruitment and the difficulty of predicting it, we could imagine and test alternative strategies. Simulating data out to 2030 may allow new management approaches to be tested for the future. For example, a fixed quota over a few years could be considered. Indeed, it is now possible to test management methods where the quota is set every three years for example. We could then see if this leads to a lower or equivalent quality of management. As such, though the criteria we have proposed were not validated, the approach we have developed is generic enough to be applied to test a wide variety of management framework. This opens the field of possibilities in terms of management methods.

4.4.2 Difficulties of managing this widespread species

As the stock is widespread and panmictic (Als et al., 2011; Palm et al., 2009), it is hard to quantify the relevance of setting quotas in France. Indeed, to date, the heterogeneity in data availability, methods, life history traits and environmental conditions impair the development of models to assess the stock at a population scale (Amilhat et al., 2008; Charrier et al., 2012) (Dekker, 2000a; Drouineau et al., 2021; Feunteun, 2002). However, the ICES has identified this lack and promoted the need for methods to integrate local scales into global dynamics and has put in place a roadmap to achieve this goal through a data call (ICES, 2021a, 2021b). If such tools are developed, it would be possible to assess the impact of local quotas on overall escapement and to compare the effects of local fisheries with other anthropogenic pressures. As such, it would help measures prioritization and support the restoration of the species.

5 Conclusion

This work has proposed two alternative recruitment models based on different methods: one relying on a spike and slab prior and one considering the recruitment slope as a time-varying random variable. Then, three criteria evaluating predictive performance of models were set up such as (i) a recruitment prediction error, (ii) a Fishery Risk Criterion (with the potential loss of catches) and (iii) a Conservation Risk Criterion (with the frequency of recruitment overestimation). These criteria were used to compare models' performance through their prediction ability, and the random slope model appears to be promising. Indeed, the time-varying random slope model looks appropriate to address recruitment stochasticity. It provides limited prediction errors and do not favour a single criterion, among conservation or exploitation criteria, in detriment of the other one. In the future, it would be interesting to analyse the consequences of quota setting when using the random slope model for the achievement of the 40% escapement objective in France, and even, on the overall population dynamics.

REFERENCES

Adam, G., Feunteun, E., Prouzet, P., & Rigaud, C. (2008). *L'anguille européenne : Indicateurs d'abondance et de colonisation*. Quae.

Als, T. D., Hansen, M. M., Maes, G. E., Castonguay, M., Riemann, L., Aarestrup, K., Munk, P., Sparholt, H., Hanel, R., & Bernatchez, L. (2011). All roads lead to home : Panmixia of European eel in the Sargasso Sea: PANMIXIA OF EUROPEAN EEL IN THE SARGASSO SEA. *Molecular Ecology*, *20*(7), 1333‑1346. https://doi.org/10.1111/j.1365-294X.2011.05011.x

Amilhat, E., Farrugio, H., Lecomte-Finiger, R., Simon, G., & Sasal, P. (2008). Silver eel population size and escapement in a Mediterranean lagoon : Bages-Sigean, France. *Knowledge and Management of Aquatic Ecosystems*, *390‑391*, 05. https://doi.org/10.1051/kmae/2009005

Anonymous. (2018). *Plan de gestion anguille de la France. Rapport de mise en œuvre—Juin 2018. Article 9 du R (CE) n°1100/2007. République Française.* https://professionnels.ofb.fr/sites/default/files/pdf/RapportPGA2018.pdf

Arai, T., & Chino, N. (2012). Diverse migration strategy between freshwater and seawater habitats in the freshwater eel genus Anguilla. *Journal of Fish Biology*, *81*(2), 442‑455. https://doi.org/10.1111/j.1095-8649.2012.03353.x

Aranburu, A., Díaz, E., & Briand, C. (2016). Glass eel recruitment and exploitation in a South European estuary (Oria, Bay of Biscay). *ICES Journal of Marine Science*, *73*(1), 111‑121. https://doi.org/10.1093/icesjms/fsv116

Arevalo, E., Drouineau, H., Tétard, S., Durif, C. M. F., Diserud, O. H., Poole, W. R., & Maire, A. (2021). Joint temporal trends in river thermal and hydrological conditions can threaten the downstream migration of the critically endangered European eel. *Scientific Reports*, *11*(1), 1‑10. https://doi.org/10.1038/s41598-021-96302-x

Aven, T., & Renn, O. (2009). The Role of Quantitative Risk Assessments for Characterizing Risk and Uncertainty and Delineating Appropriate Risk Management Options, with Special Emphasis on Terrorism Risk. *Risk Analysis*, *29*(4), 587‑600. https://doi.org/10.1111/j.1539-6924.2008.01175.x

Aven, T., & Vinnem, J.-E. (2007). *Risk Management : With Applications from the Offshore Petroleum Industry*. Springer-Verlag. https://doi.org/10.1007/978-1-84628-653-7

Beaulaton, L., Bardonnet, A., Prévost, E., Lambert, P., & Drouineau, H. (2020). *Estimations des possibilités de captures totales d'anguilles de moins de 12 cm pour la saison 2020-2021* [Research Report]. Ministère de l'Ecologie, du Développement Durable et de l'Energie. https://www.logrami.fr/download/estimations-des-possibilites-de-captures-totales-danguilles-de-moins-de-12-cm-pour-la-saison-2020-2021/

Beaulaton, L., & Briand, C. (2007). Effect of management measures on glass eel escapement. *ICES Journal of Marine Science*, *64*(7), 1402‑1413. https://doi.org/10.1093/icesjms/fsm071

Béguer-Pon, M., Castonguay, M., Shan, S., Benchetrit, J., & Dodson, J. J. (2015). Direct observations of American eels migrating across the continental shelf to the Sargasso Sea. *Nature Communications*, *6*(1), 8705. https://doi.org/10.1038/ncomms9705

Berg, A., Meyer, R., & Yu, J. (2004). Deviance Information Criterion for Comparing Stochastic Volatility Models. *Journal of Business & Economic Statistics*, *22*(1), 107‑120. https://doi.org/10.1198/073500103288619430

Bevacqua, D., Melià, P., Crivelli, A. J., Gatto, M., & De Leo, G. A. (2007). Multi-objective assessment of conservation measures for the European eel (Anguilla anguilla) : An application to the Camargue lagoons. *ICES Journal of Marine Science*, *64*(7), 1483‑1490. https://doi.org/10.1093/icesjms/fsm126

Bevacqua, D., Melià, P., De Leo, G. A., & Gatto, M. (2011). Intra-specific scaling of natural mortality in fish : The paradigmatic case of the European eel. *Oecologia*, *165*(2), 333‑339. https://doi.org/10.1007/s00442-010-1727-9

Bogstad, B., Røttingen, I., Sandberg, P., & Tjelmeland, S. (2000). *The use of Medium-Term Forecasts in advice and management decisions for the stock of Norwegian spring spawning herring (Clupea harengus L.)*. https://imr.brage.unit.no/imr-xmlui/bitstream/handle/11250/100455/V0100.pdf?sequence=1

Bonhommeau, S., Castonguay, M., Rivot, E., Sabatié, R., & Le Pape, O. (2010). The duration of migration of Atlantic Anguilla larvae : Migration duration of Atlantic eel larvae. *Fish and Fisheries*, *11*(3), 289‑306. https://doi.org/10.1111/j.1467-2979.2010.00362.x

Bonhommeau, S., Chassot, E., & Rivot, E. (2008). Fluctuations in European eel (Anguilla anguilla) recruitment resulting from environmental changes in the Sargasso Sea. *Fisheries Oceanography*, *17*(1), 32‑44. https://doi.org/10.1111/j.1365-2419.2007.00453.x

Bornarel, V., Lambert, P., Briand, C., Antunes, C., Belpaire, C., Ciccotti, E., Diaz, E., Diserud, O., Doherty, D., Domingos, I., Evans, D., de Graaf, M., O'Leary, C., Pedersen, M., Poole, R., Walker, A., Wickström, H., Beaulaton, L., & Drouineau, H. (2018). Modelling the recruitment of European eel (Anguilla anguilla) throughout its European range. *ICES Journal of Marine Science*, *75*(2), 541‑552. https://doi.org/10.1093/icesjms/fsx180

Briand, C., Buard, E., & Postic-Puivif, A. (2012). *Modélisation de la capture de civelles dans l'estuaire de la Seudre par l'utilisation du modèle GEMAC. Traitement des données des saisons 2008-2009 et 2009-2010.* 29. https://www.migrateurs-charenteseudre.fr/IMG/pdf/rapport_gemac_seudre_2008-2010.pdf

Briand, C., Fatin, D., Fontenelle, G., & Feunteun, E. (2003). Estuarine and fluvial recruitment of the European glass eel, *Anguilla anguilla* , in an exploited Atlantic estuary : GLASS EEL RECRUITMENT AND EXPLOITATION. *Fisheries Management and Ecology*, *10*(6), 377‑384. https://doi.org/10.1111/j.1365-2400.2003.00354.x

Bru, N., Prouzet, P., & Lejeune, M. (2009). Daily and seasonal estimates of the recruitment and biomass of glass eels runs Anguilla anguilla and exploitation rates in the Adour open estuary (Southwestern France). *Aquatic Living Resources*, *22*(4), 509‑523. https://doi.org/10.1051/alr/2009050

Buckland, S. T., Burnham, K. P., & Augustin, N. H. (1997). Model Selection : An Integral Part of Inference. *Biometrics*, *53*(2), 603‑618. https://doi.org/10.2307/2533961

Burnham, K. P., Anderson, D. R., & Burnham, K. P. (2002). *Model selection and multimodel inference : A practical information-theoretic approach* (2nd ed). Springer.

Castelnaud, G. (2001). Localisation de la pêche, effectifs des pêcheurs et prodution des espèces amphihalines dans les fleuves français. *Bulletin Français de la Pêche et de la Pisciculture*, *357‑360*, 439‑460. https://doi.org/10.1051/kmae/2001060

Chang, Y.-L., Miyazawa, Y., & Béguer-Pon, M. (2016). Simulating the Oceanic Migration of Silver Japanese Eels. *PLOS ONE*, *11*(3), e0150187. https://doi.org/10.1371/journal.pone.0150187

Charrier, F., Mazel, V., Caraguel, J.-M., Abdallah, Y., Le Gurun, L., Legault, A., & Laffaille, P. (2012). Escapement of a silver-phase eel population, Anguilla anguilla, determined from fishery in a Mediterranean lagoon (Or, France). *ICES Journal of Marine Science*, *69*, pp.30-33. https://doi.org/10.1093/icesjms/fsr169

Crook, V., & Shiraishi, H. (2015). *Eel market dynamics : An analysis of Anguilla production, trade and consumption in East Asia*. https://doi.org/10.13140/RG.2.1.4426.2487

Daverat, F., Limburg, K., Thibault, I., Shiao, J., Dodson, J., Caron, F., Tzeng, W., Iizuka, Y., & Wickström, H. (2006). Phenotypic plasticity of habitat use by three temperate eel species, Anguilla anguilla, A. japonica and A. rostrata. *Marine Ecology Progress Series*, *308*, 231-241. https://doi.org/10.3354/meps308231

Deelder, C. L. (1984). *Synopsis of biological data on the eel Anguilla anguilla (Linnaeus, 1758)*. 82.

Dekker, W. (2000a). The fractal geometry of the European eel stock. *ICES Journal of Marine Science*, *57*(1), 109-121. https://doi.org/10.1006/jmsc.1999.0562

Dekker, W. (2000b). A Procrustean assessment of the European eel stock. *ICES Journal of Marine Science*, *57*(4), 938-947. https://doi.org/10.1006/jmsc.2000.0581

Dekker, W. (2003). On the distribution of the European eel Anguilla anguilla and its fisheries. *Canadian Journal of Fisheries and Aquatic Sciences*, *60*(7), 787-799. https://doi.org/10.1139/f03-066

Dekker, W. (2008). Coming to Grips with the Eel Stock Skip-Sliding Away. In *International Governance of Fisheries Ecosystems : Learning from the Past, Finding Solutions for the Future* (M.G Schechter, N.J. Leonard, Vol. 62, p. 335-355). American Fisheries Society.

Dekker, W. (2019). The history of commercial fisheries for European eel commenced only a century ago. *Fisheries Management and Ecology*, *26*(1), 6-19. https://doi.org/10.1111/fme.12302

Dekker, W. (2021). Monitoring of glass eel recruitment. Report of a working group funded by the European Union. Study contract no. 98/076 : Management of the European eel: Establishment of a recruitment monitoring system, glass eel: 3 volumes. *IJmuiden : RIVO, 2002. - (RIVO report ; C007/02)*.

Delle Monache, D., & Harvey, A. (2012). The effect of misspecification in models for extracting trends and cycles. *6th Eurostat Colloquium on Modern Tools for Business Cycle Analysis: The Lessons from Global Economic Crisis, Held in Luxembourg, 26th - 29th September 2010*, *6*, 1-48. https://doi.org/10.2901/1977-3331.2011.013

Denwood, M. J. (2016). runjags : An R Package Providing Interface Utilities, Model Templates, Parallel Computing Methods and Additional Distributions for MCMC Models in JAGS. *Journal of Statistical Software*, *71*(9). https://doi.org/10.18637/jss.v071.i09

Drouineau, H., Bau, F., Alric, A., Deligne, N., Gomes, P., & Sagnes, P. (2017). Silver eel downstream migration in fragmented rivers : Use of a Bayesian model to track movements triggering and duration. *Aquatic Living Resources*, *30*, 5. https://doi.org/10.1051/alr/2017003

Drouineau, H., Briand, C., Lambert, P., & Beaulaton, L. (2016). GEREM (Glass Eel Recruitment Estimation Model) : A model to estimate glass eel recruitment at different spatial scales. *Fisheries Research*, *174*, 68‑80. https://doi.org/10.1016/j.fishres.2015.09.003

Drouineau, H., Durif, C., Castonguay, M., Mateo, M., Rochard, E., Verreault, G., Yokouchi, K., & Lambert, P. (2018). Freshwater eels : A symbol of the effects of global change. *Fish and Fisheries*, *19*(5), 903‑930. https://doi.org/10.1111/faf.12300

Drouineau, H., Lobry, J., Bez, N., Travers‑Trolet, M., Vermard, Y., & Gascuel, D. (2016). The need for a protean fisheries science to address the degradation of exploited aquatic ecosystems. *Aquatic Living Resources*, *29*(2), E201. https://doi.org/10.1051/alr/2016021

Drouineau, H., Vanacker, M., Diaz, E., Mateo, M., Korta, M., Antunes, C., Delgado, C., Domingos, I., Zamora, L., Beaulaton, L., Lambert, P., & Briand, C. (2021). Incorporating Stakeholder Knowledge into a Complex Stock Assessment Model : The Case of Eel Recruitment. *Water*, *13*(9), 1136. https://doi.org/10.3390/w13091136

Duan, Q., Ajami, N. K., Gao, X., & Sorooshian, S. (2007). Multi‑model ensemble hydrologic prediction using Bayesian model averaging. *Advances in Water Resources*, *30*(5), 1371‑1386. https://doi.org/10.1016/j.advwatres.2006.11.014

Durif, C. M. F., Diserud, O. H., Sandlund, O. T., Thorstad, E. B., Poole, R., Bergesen, K., Escobar‑Lux, R. H., Shema, S., & Vøllestad, L. A. (2020). Age of European silver eels during a period of declining abundance in Norway. *Ecology and Evolution*, *10*(11), 4801‑4815. https://doi.org/10.1002/ece3.6234

Feunteun, E. (2002). Management and restoration of European eel population (Anguilla anguilla) : An impossible bargain. *Ecological Engineering*, *18*(5), 575‑591. https://doi.org/10.1016/S0925-8574(02)00021-6

Gelman, A., & Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, *7*(4). https://doi.org/10.1214/ss/1177011136

George, E. I., & McCulloch, R. E. (1993). Variable Selection Via Gibbs Sampling. *Journal of the American Statistical Association*, *88*(423), 881‑889. https://doi.org/10.2307/2290777

Hansson, S. O. (2004). Philosophical Perspectives on Risk. *Techné: Research in Philosophy and Technology*, *8*(1), 10-35. https://doi.org/10.5840/techne2004818

Hansson, S. O. (2005). The Epistemology of Technological Risk. *Techne*, *9*(2), 68-80.

Harvey, A. C. (1986). Analysis and Generalisation of a Multivariate Exponential Smoothing Model. *Management Science*, *32*(3), 374-380. https://doi.org/10.1287/mnsc.32.3.374

Hoeting, J. A., Madigan, D., Raftery, A. E., & Volinsky, C. T. (1999). *Bayesian Model Averaging : A Tutorial*. *Vol. 14*(No. 4), 382-417.

ICES. (2001). *Report of the ICES Advisory Committee on Fishery Management, 2000 Part 1*. https://doi.org/10.17895/ICES.PUB.5375

ICES. (2005). Report of the ICES/EIFAC Working Group on Eels (WGEEL). *ICES Scientific Reports*, *ICES CM 2005/I:01*. https://www.ices.dk/sites/pub/Publication%20Reports/Expert%20Group%20Report/acfm/2005/WGEEL/WGEEL05.pdf

ICES. (2013). Report of the Joint EIFAAC/ICES Working Group on Eels (WGEEL). *ICES Scientific Reports*, *ICES CM 2013/ACOM:18*, 851.

ICES. (2018). Report of the Joint EIFAAC/ICES/GFCM Working Group on Eels (WGEEL). *ICES Scientific Reports*, *ICES CM 2017/ACOM:15*, 99. https://doi.org/10.17895/ices.pub.5627

ICES. (2020a). Joint EIFAAC/ICES/GFCM Working Group on Eels. *ICES Scientific Reports*, *2*(85), 223. https://doi.org/10.17895/ICES.PUB.5982

ICES. (2020b). *Stock Annex for the European Eel*. 84. https://www.ices.dk/sites/pub/Publication%20Reports/Stock%20Annexes/2020/Anguilla_anguilla_SA.pdf

ICES. (2020c). Workshop on temporal migration patterns of European eel (WKEELMIGRATION). *ICES Scientific Reports*, *2*(25), 109. https://doi.org/10.17895/ICES.PUB.5993

ICES. (2021a). Third Workshop on Designing an Eel Data Call (WKEELDATA3). *ICES Scientific Reports*, *3*(50), 19. https://doi.org/10.17895/ICES.PUB.8140

ICES. (2021b). Workshop on the Future of Eel Advice. *ICES Scientific Reports*, *3*(13), 67. https://doi.org/10.17895/ICES.PUB.5988

Ishwaran, H., & Rao, J. S. (2005). Spike and slab variable selection : Frequentist and Bayesian strategies. *The Annals of Statistics*, *33*(2). https://doi.org/10.1214/009053604000001147

Kermisch, C. (2012). Vers une définition multidimensionnelle du risque. *VertigO - la revue électronique en sciences de l'environnement*, *12*(2). https://doi.org/10.4000/vertigo.12214

Kimoto, A., Mouri, T., & Matsuishi, T. (2007). Modelling stock–recruitment relationships to examine stock management policies. *ICES Journal of Marine Science*, *64*(5), 870‑877. https://doi.org/10.1093/icesjms/fsm054

Kuwahara, A., & Niimi, A. (1974). *Studies on a Nematode Parasitic in the Air Bladder of the Eel I. Description of Anguillicola crassa n. Sp. (Philometridea, Anguillicolidae)*. 5.

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

Ministère de l'Ecologie, de l'Energie, du Développement durable et de l'Aménagement du Territoire, Onema, & Ministère de l'Agriculture et de la Pêche. (2010). *Plan de gestion Anguille de la France—Application du règlement (CE) n°1100/2007 du 18 septembre 2007—Volet national. Ministère de l'Ecologie, de l'Energie, du Développement durable et de la Mer, en charge des Technologies vertes et des Négociations sur le climat, Onema, Ministère de l'Alimentation, de l'Agriculture et de la Pêche.* (p. 120). http://www.pays-de-la-loire.developpement-durable.gouv.fr/IMG/pdf/panational-1.pdf

Moriarty, C. (1990). European Catches of Elver of 1928-1988. *Internationale Revue Der Gesamten Hydrobiologie Und Hydrographie*, *75*(6), 701‑706. https://doi.org/10.1002/iroh.19900750603

Moriarty, C., & Dekker, W. (1997). Moriarty C. & W. Dekker (eds). Management of the European Eel. *Fisheries Bulletin (Dublin)*, *15*, 53‑56.

Myers, R. A. (1998). When Do Environment–recruitment Correlations Work? *Reviews in Fish Biology and Fisheries*, *8*(3), 285‑305. https://doi.org/10.1023/A:1008828730759

Nijman, V. (2015). CITES-listings, EU eel trade bans and the increase of export of tropical eels out of Indonesia. *Marine Policy*, *58*, 36‑41. https://doi.org/10.1016/j.marpol.2015.04.006

Onema, MEEDDM, & MAAP. (2010). *Sauvegarde de l'anguille : Le plan de gestion français*. https://professionnels.ofb.fr/sites/default/files/pdf/memento-sauvegarde-anguille-2.pdf

Palm, S., Dannewitz, J., Prestegaard, T., & Wickström, H. (2009). Panmixia in European eel revisited : No genetic difference between maturing adults from southern and northern Europe. *Heredity*, *103*(1), 82-89. https://doi.org/10.1038/hdy.2009.51

Pike, C., Crook, V., Jacoby, D., & Gollock, M. (2020). *Anguilla bicolor (amended version of 2019 assessment). The IUCN red list of threatened species 2020 : E.T166894A176494582. Https://www.iucnredlist.org/.* [Data set]. https://dx.doi.org/10.2305/IUCN.UK.2020-3.RLTS.T166894A176494582.en

Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *Working Papers*, 8. http://www.ci.tuwien.ac.at/Conferences/DSC-2003/Drafts/Plummer.pdf

Poole, W. R., & Reynolds, J. D. (1996). Growth rate and age at migration of Anguilla anguilla. *Journal of Fish Biology*, *48*(4), 633-642. https://doi.org/10.1111/j.1095-8649.1996.tb01458.x

Raftery, A. E., Gneiting, T., Balabdaoui, F., & Polakowski, M. (2005). Using Bayesian Model Averaging to Calibrate Forecast Ensembles. *Monthly Weather Review*, *133*(5), 1155-1174. https://doi.org/10.1175/MWR2906.1

Righton, D., Westerberg, H., Feunteun, E., Økland, F., Gargan, P., Amilhat, E., Metcalfe, J., Lobon-Cervia, J., Sjöberg, N., Simon, J., Acou, A., Vedor, M., Walker, A., Trancart, T., Brämick, U., & Aarestrup, K. (2016). Empirical observations of the spawning migration of European eels : The long and dangerous road to the Sargasso Sea. *Science Advances*, *2*(10), e1501694. https://doi.org/10.1126/sciadv.1501694

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Linde, A. V. D. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *64*(4), 583-639. https://doi.org/10.1111/1467-9868.00353

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linde, A. (2014). The deviance information criterion : 12 years on. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, *76*(3), 485-493.

Subbey, S., Devine, J. A., Schaarschmidt, U., & Nash, R. D. M. (2014). Modelling and forecasting stock–recruitment : Current and future perspectives. *ICES Journal of Marine Science*, *71*(8), 2307-2322. https://doi.org/10.1093/icesjms/fsu148

Symonds, M. R. E., & Moussalli, A. (2011). A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion. *Behavioral Ecology and Sociobiology*, *65*(1), 13-21. https://doi.org/10.1007/s00265-010-1037-6

Tesch, F. W. (2003). *The Eel* (J.E. Thorpe, Vol. 65). Blackwell Science. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.0022-1112.2004.00509a.x

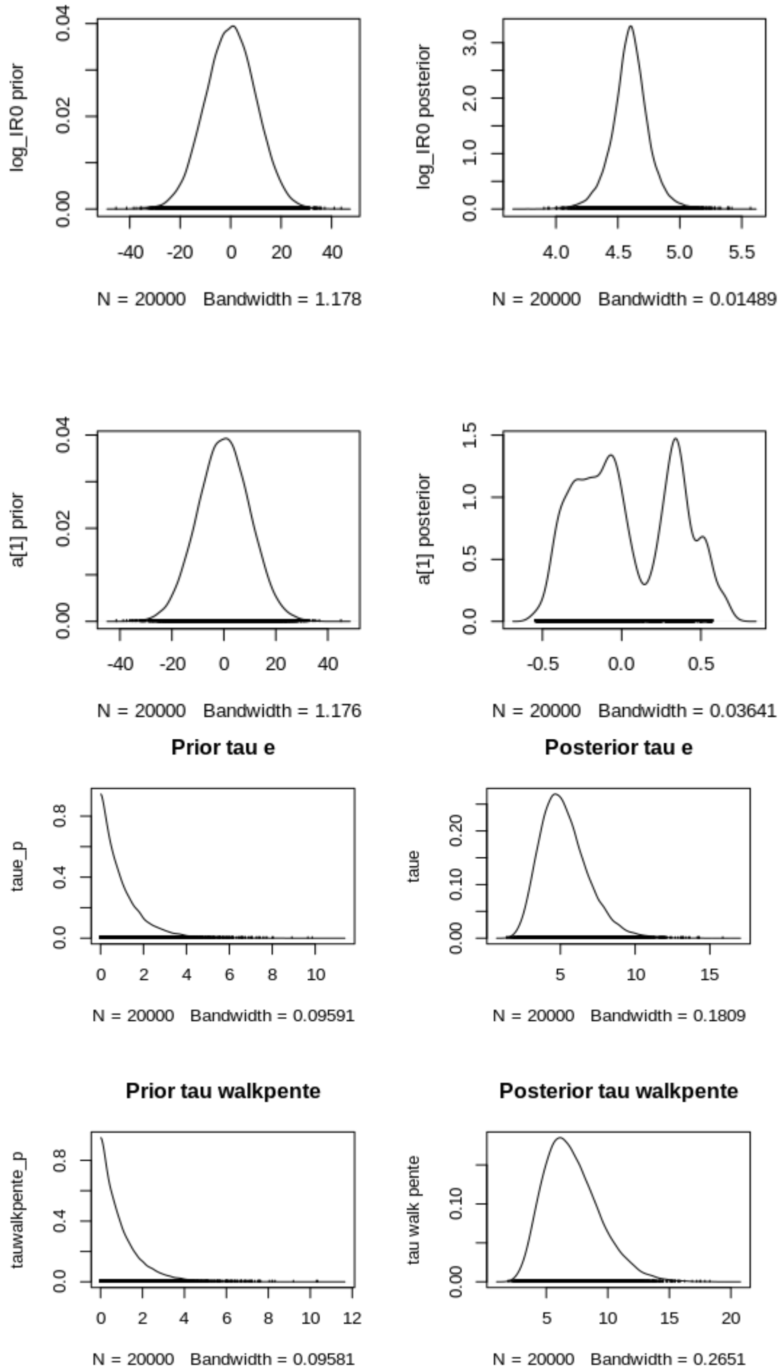
Titsias, M. K., & Lázaro-Gredilla, M. (2011). *Spike and Slab Variational Inference for Multi-Task and Multiple Kernel Learning*. 9.

Tsukamoto, K., & Kuroki, M. (Éds.). (2014). *Eels and Humans*. Springer Japan. https://doi.org/10.1007/978-4-431-54529-3

van Ginneken, V. J. T., & Maes, G. E. (2005). The European eel (Anguilla anguilla, Linnaeus), its Lifecycle, Evolution and Reproduction : A Literature Review. *Reviews in Fish Biology and Fisheries*, *15*(4), 367-398. https://doi.org/10.1007/s11160-006-0005-8

ANNEXES

Annex 1. Prior (on the left) and posterior (on the right) probability distributions of μ_{IR_0} (here "log_IR0"), the recruitment slope for the first timestep (i.e. season 1979-1980) *a[1]*, the standard deviation σ_η (here "taue"), and σ_{slope} (here "tauwalkpente") when using the random slope model.



Diplôme : Ingénieur Agronome

Spécialité : Sciences Halieutiques et Aquacoles

Spécialisation / option : Ressources et Ecosystèmes Aquatiques

Enseignant référent : Etienne RIVOT

Auteur(s) : Mathilde BENEZECH

Date de naissance* : 20/04/97

Nb pages : 28 Annexe(s) : 1

Année de soutenance : 2021

Organisme d'accueil : INRAE

Adresse : Unité EABX,

50, avenue de Verdun

33612 CESTAS Cedex

Maître de stage : Hilaire DROUINEAU

Titre français : Développement et comparaison d'outils d'aide à l'expertise sur les opportunités de capture de civelles

Titre anglais : Development and comparison of tools to support the assessment of glass eels catch opportunities

Résumé : Suite au déclin du recrutement de civelles depuis les années 70 et à la mise en œuvre du règlement anguille en 2007, la France a établi un plan de gestion en 2008 visant à réduire de 60% la mortalité par pêche des civelles. Un Conseil scientifique a alors développé des modèles de prédiction du recrutement pour aider à l'établissement des quotas de capture. Cependant, l'objectif de gestion n'est plus atteint depuis 2015, soulevant des questions sur la performance des modèles, et leurs hypothèses sur les tendances passées et futures du recrutement. Ainsi, cette étude vise à remédier aux limites des deux modèles actuels (1) en proposant des modèles alternatifs et (2) en proposant de nouveaux critères pour comparer la performance des modèles. Ainsi, un modèle basé sur un prior « spike and slab » et un sur une pente aléatoire ont été proposés pour assouplir les hypothèses sur les tendances du recrutement, et ont été ajustés sur l'indice d'abondance du WGEEL de 1980 à 2019. Pour comparer les modèles, l'erreur de prédiction du recrutement reposant sur la validation croisée ainsi que des critères quantifiant les risques pour la conservation et la pêcherie sont étudiés. Ces indicateurs sont mesurés lors d'exercices rétrospectif et prospectif. Le modèle à pente aléatoire s'avère être un bon compromis, présentant une erreur de prédiction du recrutement limitée et ne favorisant pas un seul critère parmi les critères d'exploitation et de conservation. A terme, il serait intéressant d'examiner l'impact des captures civelières sur l'échappement, qui est l'objectif de gestion du règlement.

Abstract: Following the decline in glass eel recruitment since the 70s and the implementation of the Eel Regulation in 2007, France implemented a Management Plan in 2008 aiming to reduce glass eel fishing mortality by 60%. Thus, a Scientific Council developed recruitment prediction models to support the establishment of catch quotas. However, since 2015, the management target is no longer achieved, raising questions about models' performance, and about assumptions on past and future recruitment trends. In this context, this study aims to address the limitations of the two current models by (1) proposing alternative models and (2) proposing new criteria to compare models' performance. Thus, a model based on a spike and slab prior and one on a random slope were proposed to overcome too strong assumptions on recruitment trends, and they were fitted to the WGEEL abundance index from 1980 to 2019. To compare models, recruitment prediction error based on cross-validation as well as criteria quantifying conservation and fishery risks are studied. These indicators are measured in both retrospective and prospective exercises. The random slope model proves to be a good compromise, providing a limited recruitment prediction error and not favouring a single criterion among the exploitation and conservation criteria. In the future, it would be worthwhile exploring the impact of glass eel catches on escapement, which is the Regulation Management target.

Mots-clés : Anguille européenne, surexploitation, recrutement, modèle d'évaluation bayésien

Key Words: European eel, overexploitation, recruitment, Bayesian assessment model