



HAL
open science

Testing novel methods for short-term forecasting of European glass eel recruitment

Mathilde Bénézech, Hilaire Drouineau, Anthony Acou, Agnes Bardonnnet, Clarisse Boulenger, Patrick Lambert, Etienne Prévost, Laurent Beaulaton

► **To cite this version:**

Mathilde Bénézech, Hilaire Drouineau, Anthony Acou, Agnes Bardonnnet, Clarisse Boulenger, et al.. Testing novel methods for short-term forecasting of European glass eel recruitment. Fisheries Research, 2024, 271, pp.106915. 10.1016/j.fishres.2023.106915 . hal-04402206

HAL Id: hal-04402206

<https://hal.science/hal-04402206>

Submitted on 13 May 2024

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

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

23 stock-recruitment (SR) relationship. When no SR relationship is available, recruitments can
24 be projected from past trends using methods such as the geometric means of the past
25 recruitments. This approach implicitly assumes stability in upcoming years and might not be
26 appropriate in the case of regime shifts. Recruitment forecasting is particularly critical when
27 predicting fishing opportunities for recruitment fisheries. The fishery of glass European eels
28 (*Anguilla anguilla*) is an example of a recruitment fishery. The French glass eel fishery,
29 which is currently guided by two models to forecast upcoming recruitments and derive fishing
30 opportunities, has failed to achieve management targets. This study develops new models that
31 forecast glass eel recruitments without assumptions about the SR relationship, and with
32 flexible assumptions about future recruitment evolutions. Tests based on multiple criteria
33 found the best performance in a random slope model, which provides flexibility to track
34 recruitment dynamics while balancing the interests of fisheries, management and
35 conservation. This model has wide potential for trend projections in natural resource
36 management, economics and other fields.

37 Keywords: Recruitment forecasting, Bayesian stock assessment, random walk, multicriteria,
38 European eel

39

40 1 Introduction

41 1.1 Recruitment: a key element in population dynamics

42

43 Understanding and quantifying the renewal of populations is important, especially for
44 exploited species, such as fish, for which renewal is a critical process in the sustainability of
45 the fishery. Among the many ecological processes, such as mortality (both natural and
46 anthropogenic) and biomass growth, the fishery population dynamics rely on recruitment,

47 i.e., offspring resulting from reproduction that become harvestable by the fishery (Mangel et
48 al., 2010). Fish recruitment is typically highly variable since it depends on both biotic (prey,
49 predators) and abiotic factors (temperature, currents) (Maunder and Thorson, 2019; Miller et
50 al., 2009; Myers, 1998; Subbey et al., 2014) and is sensitive to regime shifts (Ljunggren et
51 al., 2010; Mueter et al., 2007; Perretti et al., 2017). In this context, recruitment modeling,
52 and its temporal variation, plays a critical role in estimating future abundance and in setting
53 biological reference points that guide sustainable harvesting (Cury et al., 2014; Maunder and
54 Thorson, 2019; Subbey et al., 2014; Sydeman et al., 2018).

55

56 1.2 Forecasting short-term recruitment

57

58 Stock-recruitment (SR) relationships can be used to predict recruitments arising from
59 a given spawning biomass (Beverton and Holt, 1993; Hilborn and Stokes, 2010; Ricker,
60 1954). In practice, it is often difficult to discern the nature of an SR relationship and to
61 estimate its parameters (Myers, 1998; Stocker and Noakes, 1988; Sydeman et al., 2018).
62 Errors in the choice of such relationships can undermine the reliability of a fishery stock
63 assessments (Hilborn and Walters, 2013; Maunder and Deriso, 2003; Myers, 2001; Plagányi
64 et al., 2019).

65 When no clear SR relationship has been established for a stock (ICES, 2003),
66 forecasting recruitments generally requires making assumptions regarding recent and future
67 trends. The geometric mean of recent recruitments is often used in such short-term
68 recruitment forecasts (ICES, 2006; Needle, 2001). Environmental information (e.g., the use
69 of an upwelling index) or the empirical coupling of environmental and biotic information
70 can also improve forecasts (Needle, 2001). Since long-term temporal trends have been

71 observed for many stocks (Gascuel et al., 2016; STECF, 2022), the use of methods not
72 accounting for regime shifts could lead to potential biases in management recommendations.

73

74 1.3 Recruitment fishery

75

76 Assumptions about future recruitments are fundamental in the process of estimating
77 fishing opportunities. They are even more crucial for a fishery exploiting the young of the
78 year or juveniles, called a recruitment fishery, since fishing opportunity estimations directly
79 depend on the reliability of these assumptions.

80

81 1.4 Case study: the European eel (*Anguilla anguilla*) recruitment fishery

82

83 The glass eel fishery of the European eel (*Anguilla anguilla*) is an example of a
84 recruitment fishery (Briand et al., 2008; Dekker, 2003, 2002). The European eel is a
85 facultatively catadromous species that reproduces in a unique spawning ground in the
86 Sargasso Sea (Wright et al., 2022). European eel larvae are transported across the Atlantic
87 Ocean through oceanic currents and metamorphose into glass eels as they reach the
88 continental shelves from Norway to Morocco, including the Mediterranean and Baltic basins
89 (Tesch, 2003). The arrival of glass eels in the continental waters compels recruitment. The
90 eels enter the continental waters, turning into yellow eels, and colonize a large range of
91 habitats from saline to fresh water (Arai and Chino, 2012; Daverat et al., 2006). After a
92 growth phase, they metamorphose into silver eels and migrate back to the spawning ground
93 (Righton et al., 2016). Despite its large distribution area, the European eel is panmictic (Als
94 et al., 2011). The species is targeted by commercial and recreational fisheries throughout its
95 continental life stages (Dekker, 2019).

96 Glass eel recruitment declined by approximately 15% per year from 1980 to 2010,
97 remaining at a low level (Dekker, 2019; ICES, 2022; Moriarty, 1990). In 2000, the fishery
98 was declared outside safe biological limits due to impaired recruitment and threatened stock
99 renewal (ICES, 2001). The European eel was added to Appendix II of the Convention on
100 International Trade in Endangered Species of Wild Fauna and Flora in 2007 (CITES, 2007)
101 and has been listed as critically endangered by the International Union for Conservation of
102 Nature since 2008 (Freyhof and Kottelat, 2008). In 2007, the European Council adopted
103 Regulation No. 1100/2007, known as the Eel Regulation, requiring member states to
104 establish Eel Management Plans (EMPs) in their national waters that aimed at reducing
105 anthropogenic mortalities at all life stages (EU, 2007). The French EMP goal was to reduce
106 mortality by 40% (glass eel) or 30% (yellow and silver eels) in the first three years of the
107 plan, and by 60% by 2015 for each stage mentioned previously (the harvest rate being used
108 to quantify the mortality). To reduce the fishing mortality of glass eels, France enforced
109 measures including a national glass eel Total Allowable Catch (TAC) system for commercial
110 fisheries (Ministère de l'Écologie, de l'Énergie, du Développement durable et de
111 l'Aménagement du Territoire et al., 2010). In the absence of an established SR relationship,
112 the Scientific Committee uses two recruitment models based on past abundance indices to
113 forecast upcoming recruitments to support the determination of a glass eel TAC that will
114 produce the needed reduction in anthropogenic mortality (Beaulaton et al., 2020). However,
115 the management target has not been achieved since 2015, questioning the quality of these
116 short-term forecast models and their underlying assumptions.

117

118 1.5 Devising and comparing short-term recruitment forecast models

119

120 Devising recruitment models that are robust to departures from their underlying
121 assumptions is a key challenge (here, we used “robustness” as defined by Box (1979), i.e.,
122 the property that makes model outcomes insensitive to departures from ideal assumptions).
123 In the context of an undefined SR relationship and past recruitment knowledge, this paper
124 proposes two new recruitment models with flexible recruitment evolution assumptions to
125 produce robust short-term recruitment forecasts. Furthermore, this work presents a model
126 comparison strategy based on new criteria for assessing the implications of model outcomes
127 for conservation, management and exploitation. This work is undertaken through a
128 recruitment fishery case study: the European eel (*Anguilla anguilla*) recruitment fishery.

129

130 2 Material and Methods

131

132 In this work, we considered ourselves to be in the position of the Scientific Committee
133 of 2020, which met during summer 2020 and was in charge of providing advice for the 2020-
134 2021 fishing season.

135

136 2.1 Glass eel fishery data

137

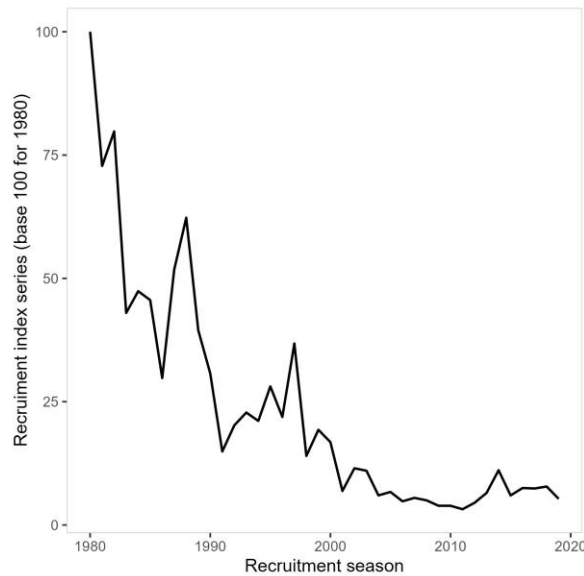
138 2.1.1 Recruitment time series

139

140 Using recruitment time series collected from widely distributed monitoring sites and
141 due to varying regional time series trends in these series, the Working Group on Eels
142 (WGEEL) uses a General Linear Model (GLM) to derive two glass eel recruitment indices,

143 the “North Sea” and “Elsewhere Europe” (i.e., Europe excluding the North Sea and Baltic
144 Sea) indices from 1980 to 2019 (ICES, 2022; ICES and EIFAC, 2010). The reconstructed
145 values are given in reference to the geometric mean of the period 1960-1979. While these
146 recruitment indices are considered to be good indicators of the population status, the complex
147 ecology of the species hinders the derivation of abundance indices at later life stages,
148 including spawning stock biomass (Drouineau et al., 2016; ICES, 2022). As such, while
149 recruitment trends are well documented, attempts to fit SR relationships have not been
150 successful (Dekker, 2004; ICES, 2013). Since the “Elsewhere Europe” index covers France
151 and includes several French time series (ICES, 2022), it is thought to describe past evolution
152 recruitments in France. Therefore, it was used as a recruitment estimate in this study (Fig. 1).
153 Since the Scientific Committee focuses on the 1980-2019 recruitment period to make its
154 forecast consistent with their model assumptions, the dataset is restricted to the post-1980
155 time range. The 1980 recruitment value is taken as a reference value and set at 100, so
156 upcoming annual recruitments are calculated as a percentage of this reference value. Since
157 glass eels reach continental French waters between November and April (ICES, 2020), the
158 annual recruitment is spread over two calendar years (e.g., glass eel arrivals between
159 November 2018 and April 2019, also called season 2019). For simplicity, season “ t ” refers to
160 the season spread over years $t-1$ and t .

161



162

163 **Fig. 1.** Recruitment estimates used in this study: the “Elsewhere Europe” recruitment index
 164 from 1980 to 2019 (ICES WGEEL 2019); it uses base 100 for the 1980 season.

165

166 **2.1.2 Landing data**

167

168 French glass eel landing data reported to the WGEEL from 1980 to 2020 were used
 169 (catches from 2009 were missing). The harvest rate corresponds to the annual ratio between
 170 landings data and recruitment. Since recruitment is a relative index (i.e., a dimensionless
 171 variable rather than an absolute estimate of abundance or biomass), the ratio of the landings
 172 data and the recruitment index gives a relative harvest rate index, with a mean reference
 173 harvest rate index calculated over the period 2004-2008 and defined as the 100 baseline
 174 (Beaulaton et al., 2020).

175

176 **2.2 Recruitment forecasting models**

177

178 2.2.1 General principle

179

180 In France, the Scientific Committee in charge of recommending TACs for the forthcoming
181 $t+1$ fishing season generally meets in early summer t . At this point, recruitment data are
182 available up to season $t-1$ (since fishing season t is not yet finished). Therefore, the Scientific
183 Committee needs to forecast recruitments for seasons t and $t+1$ to support fishing
184 opportunities for season $t+1$. To provide first insights on the recruitment level of the ongoing
185 season t , the Scientific Committee collects qualitative feedback from professional fishermen,
186 associations in charge of on-field scientific monitoring of diadromous fish, environmental
187 inspectors, and foreign scientists. This qualitative information allows building minimum and
188 maximum recruitment values for fishing season t as compared to the previous season, which
189 are then used as a censor interval in Bayesian forecasting models. In accordance with the
190 general pattern observed by the Scientific Committee, and since the 2020 season is
191 uncompleted, recruitment in 2020 is equivalent to recruitment in 2019 (e.g., 5.3) but with a
192 substantial uncertainty interval (e.g., 5.3 ± 3) (Beaulaton et al., 2020). Based on the WGEEL
193 “Elsewhere Europe” recruitment index from 1980 to 2019, and on the qualitative expert
194 feedback for 2020, the recruitment models are able to forecast the upcoming recruitment.

195

196 2.2.2 Existing glass eel recruitment models: single-trend and two-trend models

197

- 198 • Modeling framework

199

200 To forecast glass eel recruitment, two models fitted in a Bayesian framework have been
201 developed by the Scientific Committee and have been in use since 2014. Both models have a

202 yearly time step (Beaulaton et al., 2020). They are fitted to the WGEEL “Elsewhere Europe”
203 recruitment index from 1980 to 2019, considering the censor season for 2020.

204

- 205 • Introduction of the two existing recruitment models

206

207 This first recruitment model, called the single-trend (ST) model, postulates that
208 recruitment changes exponentially at a constant rate. Thus, recruitment is assumed to vary
209 linearly with a unique slope coefficient a in the log scale. Perturbations are autocorrelated in
210 time to reflect the fact that recruitment above the decreasing central trend is generally
211 followed by recruitment above the central trend (Beaulaton et al., 2020):

212

$$213 \quad \mathbf{1} \quad \text{For } t \text{ in } [1980; 2022], IR_t \sim Ln(\mu_{IR_t}, \sigma_{IR}^2)$$

214

$$\mu_{IR_t} = \mu_{IR_0} + a \cdot t + \rho \cdot \varepsilon_t$$

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

215

216 where IR_t is the recruitment index for year t that follows a lognormal distribution of mean μ_{IR_t}
217 and variance σ_{IR}^2 (log scale). μ_{IR_0} is an initialization parameter, a is the recruitment slope, ε_t is
218 the annual autocorrelated random disturbance and ρ is its autocorrelation coefficient.

219

220 Due to a possible change in trends in recent years (ICES, 2019), the Scientific Committee
221 proposed an alternative recruitment model in 2014. This model postulates a change in the
222 recruitment trend in 2012, as it is the first year in which the implementation of EMPs could
223 have had an effect on recruitment (Beaulaton et al., 2020). Two recruitment slopes a_1 (before
224 2012) and a_2 (since 2012) are considered. The two-trend (TT) model is defined as follows:

225

226 **2** For t in $[1980; 2022]$, $IR_t \sim Ln(\mu_{IR_t}, \sigma_{IR}^2)$

227 **For $t < 2012$ (time period 1)**, $\mu_{IR_t} = \mu_{IR_0} + a_1 \cdot t + \rho \cdot \varepsilon_{t_1}$

228 $\varepsilon_{t_1} = \log(IR_{t-1}) - [\mu_{IR_0} + a_1 \cdot (t - 1)]$

229

230 **For $t \geq 2012$ (time period 2)**, $\mu_{IR_t} = \mu_{IR_{2011}} + a_2 \cdot (t - 2011) + \rho \cdot \varepsilon_{t_2}$

$\varepsilon_{t_2} = \log(IR_{t-1}) - [\mu_{IR_{2011}} + a_2 \cdot (t - 2011)]$

231

232

233 The variables are similar to the ones in the ST model (Equation 1). Then, ε_{t_1} and ε_{t_2}
234 represent annual autocorrelated random disturbances depending on time periods (before or
235 after 2012), and ρ is the common autocorrelation coefficient. Models currently used to support
236 glass eel TAC recommendations are innovative because of their integration of a long-term
237 trend and a possible regime shift with the TT model. However, because of their rigid
238 underlying assumptions on recruitment future evolution, both model outcomes critically
239 depend on whether a regime shift occurred and whether it took place in 2012. Moreover, the
240 TT model is open only to a regime shift in a specific year and not to any other regime shift.

241

242 **2.2.3** Development of alternative glass eel recruitment models: the spike and slab model
243 and the random slope model

244

245 Two new models were developed to address the limitations of the existing models.

246

247 **2.2.3.1** *Spike and slab model*

248

249 In contrast to the TT model, the spike and slab (SS) model aims to objectively test for the
250 existence of a change in slope in 2012.

251

- 252 • The spike and slab method

253 The spike and slab method is a Bayesian selection technique to identify the most
254 relevant predictors of interest (George and McCulloch, 1993; Ishwaran and Rao, 2005; Titsias
255 and Lázaro-Gredilla, 2011). It relies on a specific mixture of priors based on a binary
256 indicator called γ , which quantifies, *a posteriori*, whether the variable of interest should be
257 included in the model (George and McCulloch, 1993).

258

- 259 • Spike and slab model formulation

260 The spike and slab prior appears suitable for objectively assessing the existence of a
261 change in the recruitment slope: it provides insights into whether a slope change actually took
262 place in 2012. Similarly to the TT model (Beaulaton et al., 2020), a potential slope change is
263 assumed to occur in 2012 in the SS model. A new variable $diff_a$ was introduced as the
264 difference between the two slopes a_2 and a_1 :

265

$$266 \quad \mathbf{3} \quad diff_a = a_2 - a_1$$

267

268 Instead of setting the priors to a_1 and a_2 (as in the TT model), we set the priors to a_1 and
269 $diff_a$:

270

$$271 \quad \mathbf{4} \quad diff_a \sim N(0, \sigma_{diff_a}^2)$$

$$272 \quad \sigma_{diff_a}^2 = c \cdot (1 - \gamma) + d \cdot \gamma \text{ with } c = 1 \cdot 10^{-4} \text{ and } d = 3 \cdot 10^{-2}$$

273

$$\gamma \sim Ber(0.5)$$

274

275 γ is a binary value; when set to 0, the $diff_a$ prior is narrow at approximately 0, so a_2 is
276 constrained to be similar to a_1 ; when set to 1, the resulting prior promotes $diff_a$ distinct from 0,
277 so a_2 is distinct from a_1 . Parameters c and d tune the mixture of the two normal distributions
278 of the different priors. Following Ishwaran and Rao (2005), they were chosen to provide
279 sufficiently distinct probability distributions to distinguish slope changes from slope
280 conservation. A sensitivity analysis is provided in the Supplementary Material. Posterior
281 distributions of γ and $diff_a$ indicate whether the slope change should be included (mean of the
282 γ posterior distribution ≈ 1 and $diff_a \neq 0$) or not (majority of $\gamma = 0$ and $diff_a$ close to 0). With
283 such priors, the SS model is the combination of the *a posteriori* distribution relative weights
284 of situations where $\gamma = 0$ (equivalent to the ST model) or $\gamma = 1$ (equivalent to the TT model),
285 and the ST and TT models are two subcases of the SS model.

286

287 2.2.3.2 *The random slope recruitment model*

288

289 The TT and SS models allow a change in the recruitment trend in a single specific year
290 or no changes at all. An alternative option is to consider the recruitment slope as a time-
291 varying random variable. Local linear trend models are widely used to model and predict
292 macroeconomic time series (Delle Monache and Harvey, 2012). In such models, the slope is
293 itself a time-varying random variable following a random walk. Therefore, the random slope
294 (RS) model is defined as follows:

295

296 5 For t in $[1980;2022]$, $IR_t \sim Ln(\mu_{IR_t}, \sigma_{IR}^2)$

297 $\mu_{IR_t} = \mu_{IR_{t-1}} + a_t + \eta_t$ with $\eta_t \sim N(0, \sigma_\eta^2)$

298 $a_t = a_{t-1} + \delta_t$ with $\delta_t \sim N(0, \sigma_\delta^2)$

299

300 where IR_t is still the recruitment index for year t that follows a log normal distribution of
301 parameters μ_{IR_t} (mean in the log scale) and $\sigma_{IR_t}^2$ (variance in the log scale). Then, μ_{IR_t} is
302 assumed to be centered around its previous value, plus a slope a_t and a random noise η_t . The
303 slope a_t itself follows a random walk, with random noise δ_t . It can be demonstrated that σ_δ
304 (variance of the noise of the mean changes) and σ_η (variance of the noise of the slope random
305 walk) are redundant, and as such, estimating both parameters can lead to identifiability issues
306 (Cole, 2020; Schnute, 1994). Here, we assumed that $q_\delta = (q_\eta / 2)^2$. By doing so, the model
307 is then equivalent to a commonly used double exponential smoothing (Harvey, 1986) and
308 yields robust predictions (Delle Monache and Harvey, 2012). The ST, TT and SS models are
309 special cases of this more generic RS model.

310

311 2.3 Bayesian inference and priors

312

313 The Bayesian models were fitted using JAGS (Plummer, 2003). We used the runjags
314 package as an interface between R and JAGS (Denwood, 2016). Uninformative priors were
315 used for most parameters (Table 1). Three chains were run in parallel for 200,000 iterations
316 with a thinning period of ten after a burn-in period of 200,000 iterations. Model convergence
317 was checked using Gelman tests (Gelman and Rubin, 1992) and with graphical verifications
318 by traceplots.

319

320 **Table 1.** Priors on the models (“ST”: Single-trend; “TT”: Two-trend; “SS”: Spike and slab;
321 “RS”: Random slope) parameters and the number of associated equations (“Eq.”).

	Variables	ST (Eq. 1)	TT (Eq. 2)	SS (Eq. 4)	RS (Eq. 5)
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], $\eta_t \sim N(0, \sigma_\eta^2)$
2	σ_η				Gamma(0.01,0.01)
6	a	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], $\delta_t \sim N(0, \sigma_\delta^2)$
12	σ_δ				$\sigma_\delta = (\sigma_\eta^2 / \sigma_{IR}) * 0.5$

322

323 2.4 Model comparison strategy

324

325 2.4.1 Deviance Information Criterion to assess model fitting quality

326

327 The Deviance Information Criterion (DIC), a Bayesian version or generalization of the
328 Akaike Information Criterion, quantifies the trade-off between goodness of fit and model
329 complexity (Spiegelhalter et al., 2002). The DIC is one of the most widely used criteria to
330 assess the performance of Bayesian models (Gelman et al., 2014). The DIC was calculated
331 through 200,000 iterations with a thinning period of ten for each model. Like the AIC, the

332 DIC assesses the capacity of the model to predict observed data and, as such, quantifies the
333 hindcasting capacity of the model. However, it does not assess the forecasting ability of a
334 model, which is necessary to underpin robust stock management advice (Kell et al., 2021).

335

336 2.4.2 Assessing model forecasting performance through a simulation exercise

337

338 Assessing the predictive capacity of the models requires a dataset in which the “true”
339 recruitment is known and can be compared to model forecasts. In this context, we carried out
340 a simulation exercise to generate artificial “surrogate” datasets.

341

- 342 • Generating surrogate recruitment time series

343

344 Three scenarios were used to simulate surrogate recruitment time series from 2020 to
345 2030. To simulate the declining recruitment in Scenario 1, we used the ST model to construct
346 a forward recruitment index series from 2020 to 2030. The TT model enabled the generation
347 of a forward recruitment time series for Scenario 3 to replicate an increasing recruitment trend
348 since 2012. Finally, Scenario 2 was an intermediate situation between Scenarios 1 and 3. The
349 future recruitment of Scenario 2 was simulated by the TT model with $diff_a$ set as the
350 difference between the $median(a_2)$ and $median(a_1)$ (both estimated by the TT model), which
351 equates to $diff_a$ close to 0. Thus, Scenario 2 roughly assumes stable recruitment. For each
352 scenario, 2,000 “observed” recruitment time series were generated (Fig. 2, Part 1).

353

- 354 • Fitting the model to the simulated recruitment time series

355 Each model was fitted to the simulated recruitment time series up to season t according to
356 the following iterative algorithm, mimicking the process of the Scientific Committee (Fig. 2,
357 Parts 2 and 3):

- 358 • For each Scenario s from 1 to 3
- 359 • For each time series i from 1 to 2,000
- 360 • For each model m from 1 to 4
- 361 • For each year t from 2020 to 2029
 - 362 ○ The combination of the WGEEL time series (from 1980 to 2019) and the
363 simulated recruitment from 2020 to $t-1$ is used as input data (for $t =$
364 2020, only the WGEEL time series is used).
 - 365 ○ The censor interval for recruitment t is set as the value at time $t-1 \pm 3$ (to
366 reproduce the management task).
 - 367 ○ The model m forecasts recruitment in $t+1$, and this forecasted value is
368 compared with the corresponding “observed” value (i.e., from the
369 simulated recruitment time series).

370

371 This process generates recruitment forecasts from 2021 to 2030. All the recruitment
372 models were fitted to the three recruitment scenarios. For this prospective analysis, one chain
373 was run with 5,000 iterations with a thinning period of ten, resulting in 500 samples (after a
374 burn-in period of 10,000 iterations for the ST, TT and SS models, and 200,000 iterations for
375 the RS model to ensure convergence, since it has many more parameters). These parameters
376 were applied to the three scenarios.

377

- 378 • Quantifying model forecasting performances

379

380 The forecasting performance of each model was evaluated by calculating Percent BIAS
381 (PBIAS) (Gupta et al., 1999). The PBIAS indicates whether the model tends to overestimate
382 (negative PBIAS) or underestimate (positive PBIAS) compared to “observed” recruitment
383 indices, where 0 is the optimal PBIAS value.

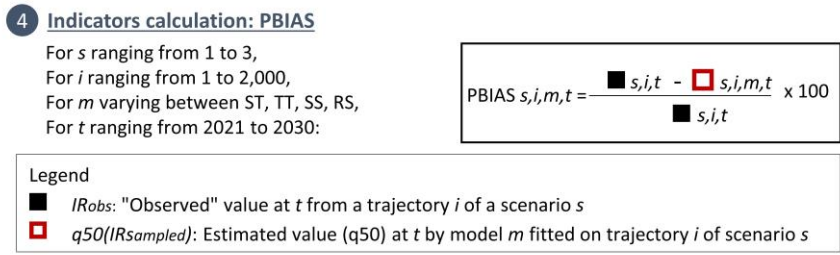
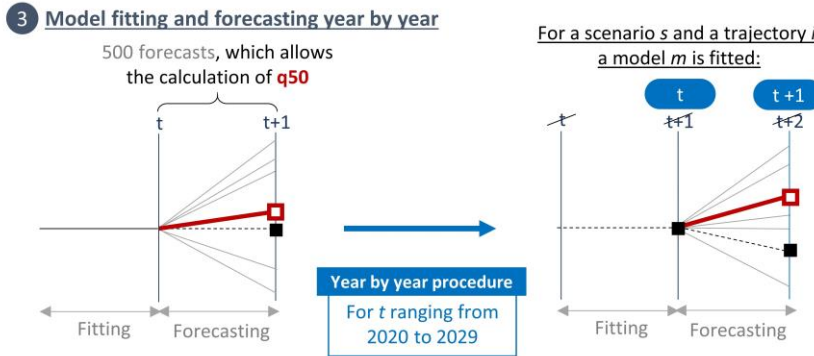
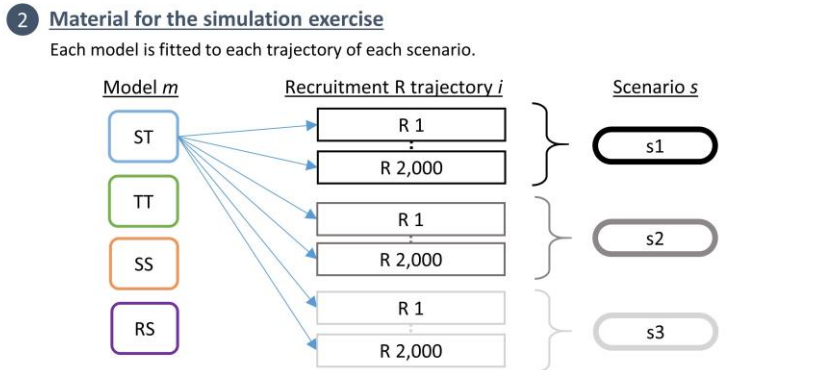
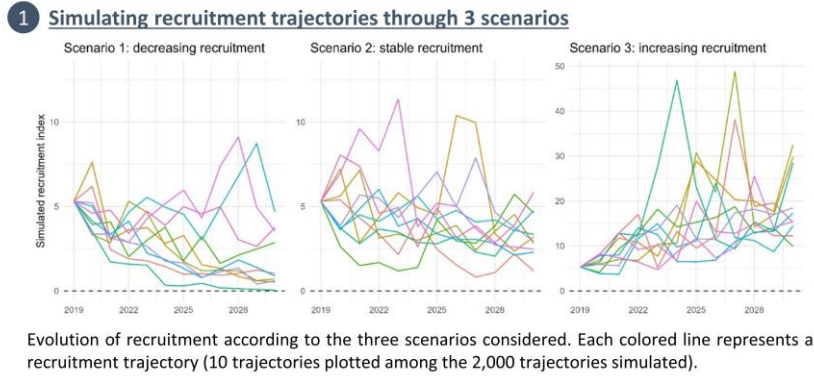
384 For a Scenario s (1:3), a simulated recruitment index series i (1: 2,000), a model m (1:4),
385 and a year t (2021:2030), an annual PBIAS is calculated by comparing the yearly “observed”
386 recruitment IR_{obs} with the median of the posterior distribution estimated by the forecasting
387 model $q50(IR_{sampled})$. We therefore obtain 20,000 PBIAS values (2,000 simulated recruitment
388 trajectories*10 years) per model for each scenario (i.e., 20,000*4*3 PBIAS values *in fine*)
389 (Fig. 2). Therefore, PBIAS is calculated as follows:

390

$$PBIAS_{s,m,i,t} = \frac{IR_{obs\ s,i,t} - q50(IR_{sampled\ s,i,m,t})}{IR_{obs\ s,i,t}} \times 100$$

391

392



393
394
395
396
397
398
399

Fig. 2. Diagram showing the quantification of the forecasting performance through the calculation of PBIAS and its different steps (from 1 to 4).

- Quantification of the risk for exploitation, resource and management: three comparison criteria

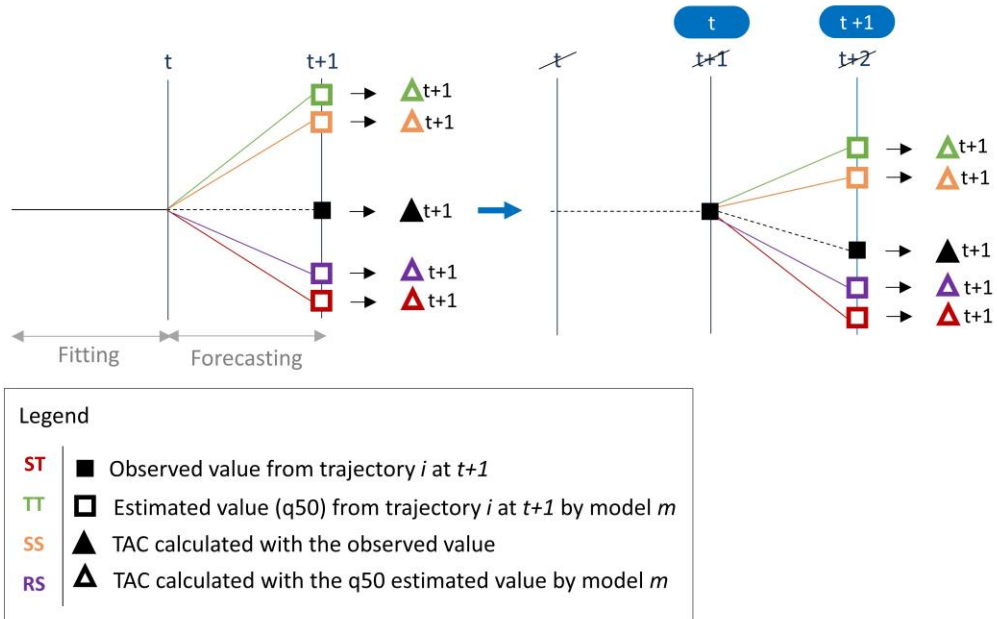
400 For each model fit (i.e., a combination of a simulated time series and a recruitment
401 model), we set the annual TAC by multiplying the target harvest rate by the median of the
402 posterior distribution of the forecasted recruitment $q50(IR_{sampled})$. Tests were also carried out
403 with quantiles of 25% and 75% and are presented in the Supplementary Material. We also
404 calculated the corresponding catch at the management target (i.e., the target harvest rate
405 multiplied by the “observed” recruitment IR_{obs}) (Fig. 3, Step 4a).

406 The application of recruitment forecasting models affects species conservation, the glass
407 eel fishery, and compliance with management objectives. We developed quantitative criteria
408 to compare model performance with respect to these effects (Fig. 3, Step 4b). For each
409 scenario, each model and each simulated recruitment time series, we computed the annual
410 average predicted catch TAC over the period 2021-2030, representing the Fishery criterion.
411 Then, the proportion of years in which the management target was met from 2021 to 2030
412 (years with predicted catch TAC < catch at the management target) was used as the
413 Management criterion. The quantity of glass eels saved if a given model is chosen,
414 representing the Conservation criterion, is computed as the annual mean of the difference
415 between the highest predicted TAC among the models and the predicted TAC of the
416 considered model over the period 2021-2030.

417

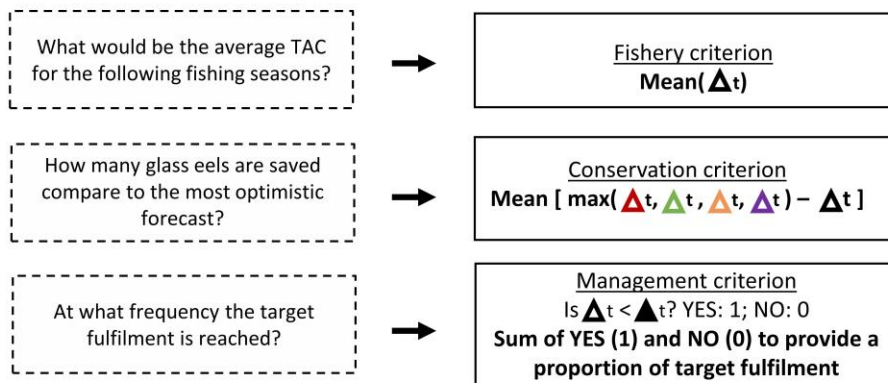
4a **Indicators calculation: Total Allowable Catch (TAC) year by year calculation**

For a scenario s and for a trajectory i :
 For t ranging from 2020 to 2029:



4b **Indicators calculation: comparison criteria**

For a scenario s , for a model m and for t ranging from 2021 to 2030:



418

419 **Fig. 3.** Diagram showing the quantification of the forecasting performance through the
 420 simulation exercise: Total Allowable Catch calculation (Step 4a) and comparison criteria
 421 calculation (Step 4b) (see Fig. 2 for previous steps).

422

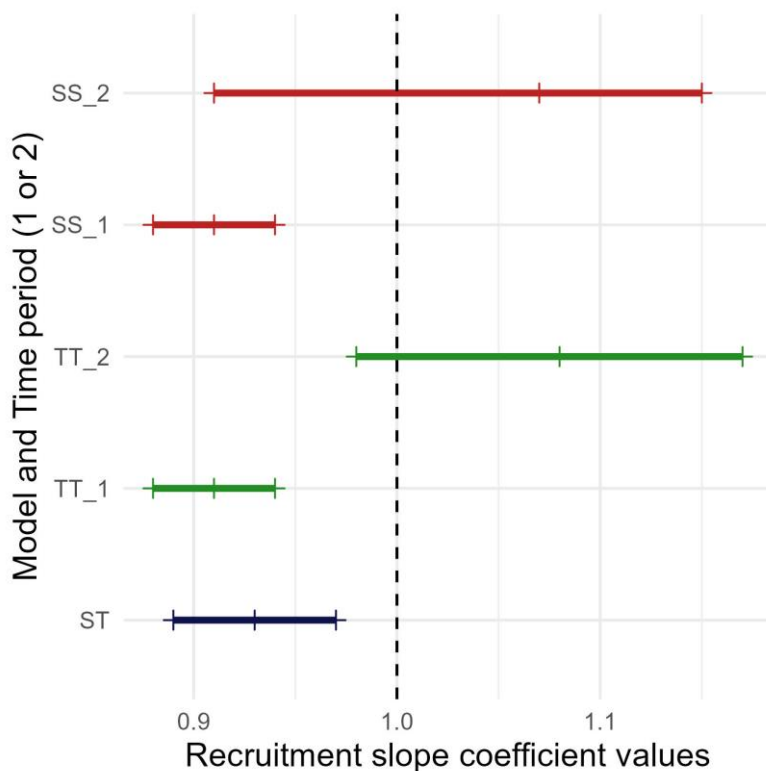
423 **3 Results**

424

425 3.1 Evolution of the recruitment slope over time depending on models

426

427 When comparing strictly declining recruitment slopes (i.e., coefficient distributions <
428 1), the ST model tends to estimate the lightest declining trend compared to the TT and SS
429 declining trends (the ST median recruitment slope is superior to other medians; Fig. 4). For
430 both the TT and SS models, recruitment declines during the first period (a_1 strictly inferior to
431 1), while the recruitment trend evolution during the second period is unclear (since 1 is within
432 the 95% credibility interval of a_2 and corresponds to stable recruitment). Recruitment slopes
433 a_1 and a_2 are distinct for the TT model, while the 95% confidence intervals of the two time
434 periods overlap for the SS model. The TT and SS models also differ by a larger amplitude of
435 the 95% credibility interval for the SS model compared to the TT model during time period 2.
436 These results arise from the fact that the SS model is constructed as a mixture of cases
437 rejecting or considering a change in slope between the two time periods, with 7.6% of cases
438 not considering a change in the recruitment slope (according to δ posterior distribution).

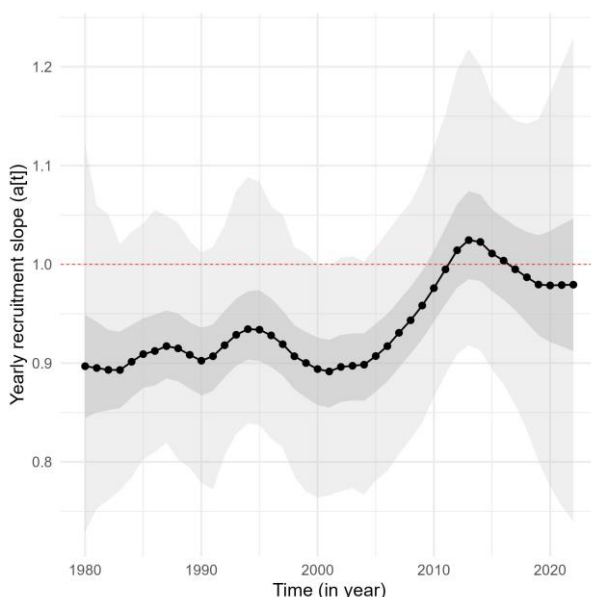


439

440 **Fig. 4.** Annual recruitment slope coefficients a (in linear scale). The letters and colors account
441 for the recruitment models. The numbers “1” and “2” account for time periods (1: 1980-2011;
442 2: 2012-2022; when no time period is assigned, the value is estimated through the entire time
443 period, i.e., 1980 to 2022). Dashes represent q2.5, q50 and q97.5.

444

445 For the RS model, the yearly varying random slope posterior distributions are mainly
446 (q50 and q75) below 1 (and strictly below 1 in 2001) until the mid-2000s, suggesting a
447 declining trend (Fig. 5). After a q50 increase between 2005 and 2013 in favor of an increase
448 in recruitment, there is no clear trend in the most recent years, with later posterior
449 distributions approximately 1. The mean and median recruitment slopes over the time series
450 are 0.9, representing a decrease of 8% in recruitment each year (on a linear scale).
451 Unsurprisingly, the forecasted slopes (for the 2020, 2021 and 2022 seasons) have larger
452 credibility intervals than the other models.



453

454 **Fig. 5.** Evolution of the annual recruitment slope over time according to the RS model. Black
455 line and dots: q50; the darker ribbon: the 50% credibility interval; the lighter ribbon: the 95%
456 credibility interval.

457

458 3.2 Model performance

459

460 3.2.1 Deviance Information Criterion

461

462 According to the DIC, the TT and SS models outperform the ST model (Table 2).

463 However, the RS model outperforms them despite its greater complexity.

464

465 **Table 2.** The Deviance Indicator Criterion value according to the recruitment model (“ST”:

466 Single-trend; “TT”: Two-trend; “SS”: Spike and slab; “RS”: Random slope)

Recruitment model	ST	TT	SS	RS
DIC	249.9	251.0	255.1	238.9

467

468

469 3.2.2 Error in recruitment estimations depending on model and scenario

470

471 The models mainly overestimate recruitment in Scenario 1 (The PBIAS values are

472 mostly negative) and underestimate recruitment in Scenario 3 (The PBIAS values are mostly

473 positive) (Fig. 6). The results are less obvious for Scenario 2; the ST and RS models

474 underestimate recruitment, while the TT and SS models overestimate recruitment.

475 Unsurprisingly, the ST and TT models perform particularly well in scenarios in which they

476 were used to simulate the data.

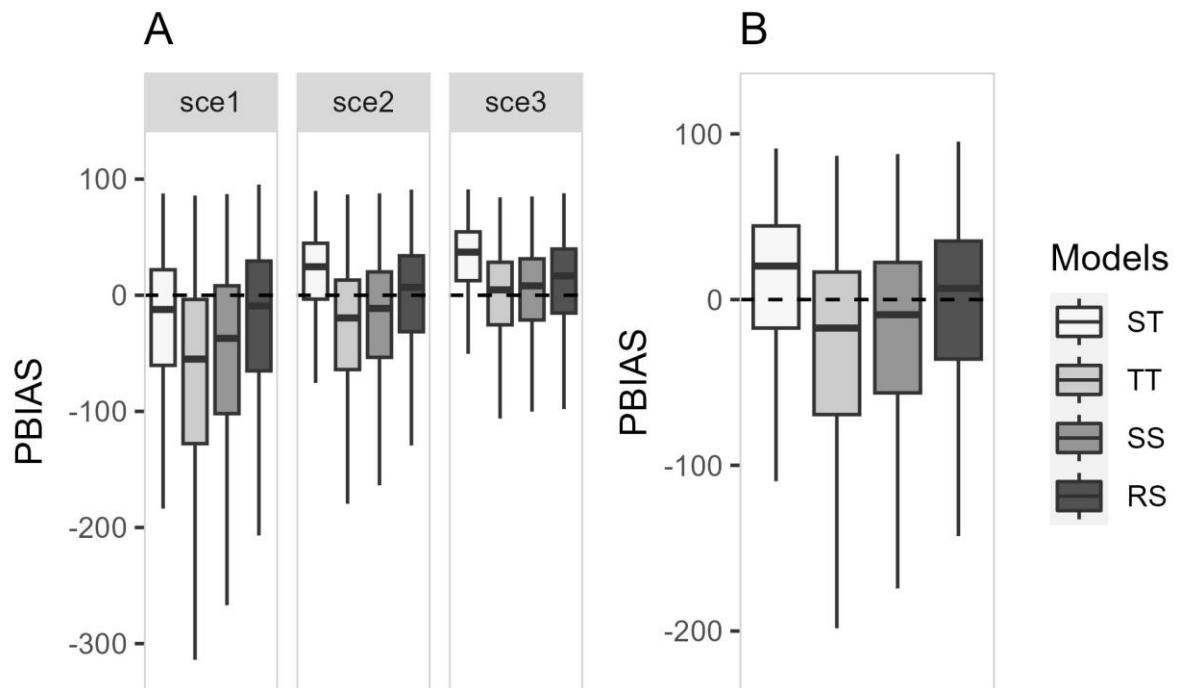
477 Interestingly, while the PBIAS distributions vary among models, and for each model

478 among the scenarios, the boxplot distribution pattern remains similar across the scenarios,

479 namely the ST model, the RS model, the SS model and the TT model in order of decreasing

480 PBIAS values. Across scenarios, the RS model appears to be the least biased model (The
481 PBIAS oscillating by approximately 0) (Fig. 6A) and underestimates recruitment by 6.7%
482 (median(PBIAS) over all the scenarios; closest value to 0 among models) (Fig. 6B).

483



484

485 **Fig. 6.** Boxplots presenting the PBIAS for the four models through the three scenarios (A) or
486 combining all scenarios (B) (“ST”: Single-trend; “TT”: Two-trend; “SS”: Spike and slab;
487 “RS”: Random slope). The outliers are not displayed for better readability but are considered
488 in the calculation of values (such as median) (strategy q50).

489

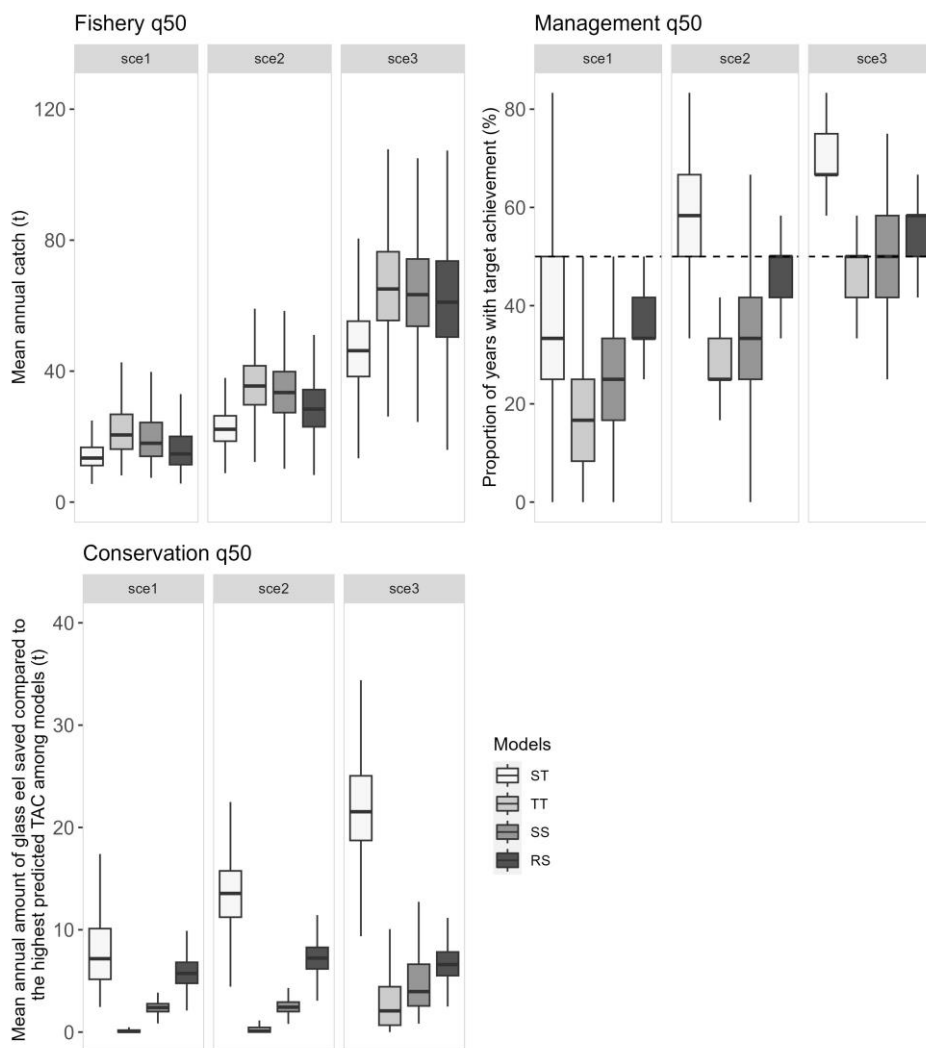
490 The RS model behaves in a similar way regardless of the scenarios during this 10-year
491 period, suggesting that the model is robust to varying recruitment patterns.

492

493 3.2.3 Quantification of the risk for fisheries, management and conservation

494

495 Fig. 7 displays the distributions of the Fishery, Conservation and Management criteria
 496 depending on the scenario and the model.



497
 498 **Fig. 7.** Boxplots presenting the three criteria “Fishery”, “Management” and “Conservation”
 499 across scenarios (“sce”) and among models (“ST”: Single-trend; “TT”: Two-trend; “SS”:
 500 Spike and slab; “RS”: Random slope) over the 10 years of simulation. The outliers are not
 501 displayed for better readability but are considered for building the graph and the calculated
 502 values (such as the median) (strategy q50).

503
 504 Unsurprisingly, since the ST model was used to simulate the recruitment series of
 505 Scenario 1, it performs well (i.e., limited potential catch loss when comparing mean annual

506 TAC to others, very good compliance to the management rule and many eels saved) in this
507 scenario. For Scenarios 2 and 3, the ST model reaches the management target with a high
508 frequency and provides a high quantity of glass eels saved at the cost of large reductions in
509 catches, resulting in the lowest mean annual catch.

510 Across scenarios and criteria, the TT and SS models have similar results when estimating
511 recruitment, and regarding the criteria comparison strategy. These models frequently miss the
512 management target and result in the lowest glass eel escapement, except for Scenario 3 (the
513 one in which the TT model was used to simulate the time series). On the other hand, these
514 models lead to the highest TAC (slightly higher for the TT model than for the SS model).

515 Considering a time period of 10 years, the RS model behaves in a similar way regardless
516 of the scenarios (i.e., intermediate mean TAC, proportion of years with target achievement
517 oscillating approximately 50%, and stable amount of glass eels saved), confirming its
518 robustness to different recruitment trends.

519

520 4 Discussion

521

522 Recruitment is a key question in population dynamics, especially for short-term
523 forecasting (Maunder and Thorson, 2019). This article provides insights into building and
524 comparing forecasting models. The RS model stands apart as the more robust method to
525 forecast short-term recruitment indices based on its bias quantification results and its
526 implications for management, exploitation and conservation.

527

528 4.1 Innovative recruitment forecasting models

529

530 • Model rationales

531

532 Spike and slab priors are widely used in Bayesian variable selection (Ishwaran and Rao,
533 2005; Titsias and Lázaro-Gredilla, 2011). The inclusion, or not, of a predictor is estimated
534 internally in the model, and the resulting predictions are a weighted average of the model with
535 and without the predictor (Titsias and Lázaro-Gredilla, 2011). As such, this approach appears
536 appropriate for testing the existence of a regime shift, objectively balancing the predictions of
537 the ST and TT models. However, the parametrization of a spike and slab prior remains
538 partially subjective (tuning the coefficients c and d in the expression of the variance) and
539 based on expertise to separate plausible values from impossible ones (George and McCulloch,
540 1993). A sensitivity analysis (see the Supplementary Material) demonstrated that our results
541 are not overly sensitive to the parameter values (this is not surprising, as situations with a
542 slope change largely dominate). However, this could change in the future, especially if
543 recruitment starts decreasing again and the weight of the ST model (i.e., the frequency of $\gamma =$
544 0) increases. Therefore, it would seem necessary to repeat the sensitivity analysis regularly
545 and update the parametrization of the prior in any future use of the SS model. Various slight
546 modifications have been proposed to help in the parametrization of the spike and slab priors
547 and to reduce the resultant uncertainty (George and McCulloch, 1993; Ishwaran and Rao,
548 2005). However, these modifications would not entirely remove the subjectivity of the
549 resulting model and the need to check its sensitivity.

550

551 The SS model imposes a change in the recruitment trend at a fixed and unique year.
552 This could be challenged by future data. One could imagine adding other years of trend
553 change, or at least modeling the year of the change as a random variable. However,
554 preliminary tests have shown that the resulting models fail to converge. In this framework, the

555 RS model is much more flexible. A random slope model appears suitable for resolving the
556 debate between a model with either a single or two slopes, and determining the year in which
557 a potential regime shift occurs. Indeed, random slope models have proven to provide robust
558 predictions, even in the presence of cyclic patterns or trend changes (Delle Monache and
559 Harvey, 2012).

560

- 561 • Model and data uncertainty handling

562 Following the Scientific Committee, we used the EIFAAC-ICES-GFCM WGEEL
563 “Elsewhere Europe” index as an indicator of French recruitment in our analysis. The use of
564 this indicator was needed given the limited number of ongoing French recruitment time series
565 (ICES, 2014). Since the index is partly derived from time series collected in regions other
566 than France, this might lead to bias (ICES, 2022). More importantly, the WGEEL data were
567 taken to be accurate and were known to be without error in the ST, TT and SS models. We
568 considered an estimation error in the RS model (Equation 5), which is an improvement.
569 However, the variance is assumed to be constant over time. The use of year-specific
570 estimation errors generated by the GLM would probably be more realistic.

571

572 4.2 Model comparison strategy

573

574 Inspired by the Management Strategy Evaluation approach (Smith, 1999), we developed
575 a multicriteria strategy to compare the performance of models.

576

- 577 • Bias in model estimation

578 First, we explored biases in recruitment forecasts, since systematic biases would result
579 in systematic overestimations or underestimations of fishing opportunities. In our simulations,

580 the ST model tends to underestimate recruitment, while the TT and SS models tend to
581 overestimate recruitment. Due to its enhanced flexibility, the RS model provides the least
582 biased recruitment estimates across the three recruitment scenarios. As the RS model behaves
583 the same way across scenarios, it appears to be the least sensitive to possible changes in future
584 recruitment trends and the most robust, as defined by Box (1979).

585

- 586 • Model implications for European eel fishery, management and conservation

587 Inspired by the work of Lebot (2021) and Bevacqua (2007), we developed three criteria
588 to assess and illustrate the trade-off between mean annual catch, management target
589 achievement and quantity of glass eel saved. The ST model favors conservation by frequently
590 achieving the management target but at the cost of lower TACs. On the other hand, both the
591 TT and SS models promote the highest TACs but with a high risk of overshooting the
592 management target and compromising conservation. The RS model provides the best trade-off
593 since it shows intermediate responses for all considered sectors compared to other models.
594 Moreover, the outcomes of the RS model are minimally sensitive to the scenarios, unlike
595 other models.

596

597 4.3 Application to glass eel fishing opportunity evaluation

598

599 The current French evaluation of glass eel fishing opportunities relies on two
600 recruitment models (ST and TT) with contrasting recruitment forecasts (Beaulaton et al.,
601 2020). The Scientific Committee lacks a method for choosing between the two models, which
602 leads to some misunderstanding from stakeholders. The RS model appears to be a good
603 candidate for estimating glass eel fishing opportunities in France. Its use could overcome the
604 difficulty of having two recruitment models by providing a single recruitment forecast. By

605 doing so, it could improve the reliability of the assessment and make the decision-making
606 process more understandable to stakeholders. Moreover, the simulation strategy can be used
607 to carry out a long-term risk analysis that would support science-based management.

608

609 4.4 A generic approach

610

- 611 • Comparing abundance forecast outcomes

612

613 As scientific results can be difficult to convey (Lehuta et al., 2016; Pita et al., 2016), we
614 have proposed three explicit quantitative criteria with the aim of providing accessible
615 information to those who make management decisions. The model selection method presented
616 herein is not based on traditional statistical criteria, but on socioeconomic and environmental
617 criteria. Since managers generally have to deal with trade-offs among competing interests, the
618 multicriteria assessment appears to be an attractive decision-aiding tool that provides trade-off
619 overviews across multiple sectors of interest (exploitation, management and conservation in
620 this study) (Antunes et al., 2006; Röckmann et al., 2012). We do not pretend that the criteria
621 we propose here are the only ones possible; however, we aim to demonstrate our trade-off
622 exploration strategy. Indeed, implementing more relevant criteria, as well as adding a relative
623 weight to each criterion to obtain a final score per model could arise from a co-construction
624 involving all stakeholders (e.g., managers, fishers, scientists, NGOs) involved in the
625 management of the European eel (Drouineau et al., 2021). This collaboration, as part of
626 participatory modeling, could help to enhance the involvement of each stakeholder in fishery
627 management, strengthening mutual trust and understanding (Drouineau et al., 2021;
628 Röckmann et al., 2012). The same applies to the scenarios (here, we used three contrasting
629 scenarios to simulate future recruitment time series, but these scenarios could be co-

630 constructed with stakeholders) and to quantiles of the posterior distribution of the forecasted
631 recruitment used to calculate TACs (here, we used strategies q25, q50 and q75).

632

- 633 • Index-based abundance short-term forecasts

634

635 Analytical stock assessments of European (*Anguilla anguilla*) and American (*Anguilla*
636 *rostrata*) eel have not been achieved despite much effort (ASMFC, 2023; ICES, 2019). As an
637 alternative to analytic stock assessment, abundance index methods are being looked at as a
638 means to generate harvest rules (e.g., the ITarget method, trend indicators) (ICES, 2023;
639 Legault et al., 2023). With eel abundance indices playing a central role in supporting
640 management advice and no viable alternative in sight, eel assessment scientists ought to be
641 looking carefully at the best options to use abundance time series to guide harvest
642 management. Of the modeling approaches tested here, the RS model best accounts for the
643 recruitment variability. We therefore recommend its use for recruitment modeling and
644 forecasting to guide European glass eel management advice.

645 The index-based methods presented in this paper could be applied to other eel life
646 stages, and to other species. More generally, these methods could be easily implemented to
647 support management advice when an analytical stock assessment cannot be conducted (e.g.,
648 rejected assessments, data-limited stock) (NEFSC, 2020; Punt et al., 2020) and could also be
649 used in artisanal and traditional coastal recruitment fisheries (e.g., the sardine fry fishery in
650 Italy) (STECF, 2023a, 2023b). Although this work is conducted at the recruitment fishery
651 scale, these index-based forecasting methods could conceivably be applied to a wide range of
652 scientific and economic fields.

653

654 [CRediT authorship contribution statement](#)

655 **Mathilde Benezech:** Conceptualization, Methodology, Software, Formal analysis, Writing –
656 original draft, Writing –review and editing; **Hilaire Drouineau:** Conceptualization,
657 Methodology, Software, Formal analysis, Writing –review and editing, Project administration,
658 Supervision; **Anthony Acou:** Conceptualization, Writing –review and editing; **Agnès**
659 **Bardonnet:** Conceptualization, Writing –review and editing; **Clarisse Boulenger:**
660 Conceptualization, Writing –review and editing; **Patrick Lambert:** Conceptualization,
661 Writing –review and editing; **Etienne Prévost:** Conceptualization, Software, Formal analysis,
662 Writing –review and editing; **Laurent Beaulaton:** Conceptualization, Methodology, Writing
663 –review and editing, Project administration, Supervision.

664

665 [Declaration of Competing Interest](#)

666 The authors declare that they have no known competing financial interests or personal
667 relationships that could have appeared to influence the work reported in this paper.

668

669 [Acknowledgements](#)

670 This work was supported by the cooperation agreement OFB - INRAE - Institut Agro -
671 Université de Pau et des Pays de l'Adour/E2S UPPA (Management of Diadromous Fish in
672 Their Environment) and the Nouvelle Aquitaine Region.

673 [References](#)

674

675 Als, T.D., Hansen, M.M., Maes, G.E., Castonguay, M., Riemann, L., Aarestrup, K., Munk, P., Sparholt, H., Hanel,
676 R., Bernatchez, L., 2011. All roads lead to home: panmixia of European eel in the Sargasso Sea.
677 *Molecular Ecology* 20, 1333–1346. <https://doi.org/10.1111/j.1365-294X.2011.05011.x>
678 Antunes, P., Santos, R., Videira, N., 2006. Participatory decision making for sustainable development—the use
679 of mediated modelling techniques. *Land Use Policy, Resolving Environmental Conflicts: Combining*
680 *Participation and Multi-Criteria Analysis* 23, 44–52. <https://doi.org/10.1016/j.landusepol.2004.08.014>

681 Arai, T., Chino, N., 2012. Diverse migration strategy between freshwater and seawater habitats in the
682 freshwater eel genus *Anguilla*. *Journal of Fish Biology* 81, 442–455. <https://doi.org/10.1111/j.1095->
683 8649.2012.03353.x

684 ASMFC, 2023. Atlantic States Marine Fisheries Commission (ASMFC) - American Eel Benchmark Stock
685 Assessment and Peer Review Report.

686 Beaulaton, L., Bardonnnet, A., Prévost, E., Lambert, P., Drouineau, H., 2020. Estimations des possibilités de
687 captures totales d'anguilles de moins de 12 cm pour la saison 2020-2021 (Research Report). INRAE;
688 MNHN; AFB - Agence française pour la biodiversité.

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

692 Beverton, R.J.H., Holt, S.J., 1993. *On the Dynamics of Exploited Fish Populations*. Springer Netherlands,
693 Dordrecht. <https://doi.org/10.1007/978-94-011-2106-4>

694 Box, G.E.P., 1979. Robustness in the Strategy of Scientific Model Building, in: Launer, R.L., Wilkinson, G.N.
695 (Eds.), *Robustness in Statistics*. Academic Press, pp. 201–236. <https://doi.org/10.1016/B978-0-12->
696 438150-6.50018-2

697 Briand, C., Bonhommeau, S., Castelnaud, G., Beaulaton, L., 2008. An appraisal of historical glass eel fisheries
698 and markets: landings, trade routes and future prospect for management 22.

699 CITES, 2007. Amendments to Appendices I and II of the Convention adopted by the Conference of the Parties
700 at its 14th meeting. Notification to the Parties. No. 2007/022. Convention on International Trade in
701 Endangered Species of Wild Fauna and Flora, The Hague, Netherlands.

702 Cole, D.J., 2020. *Parameter Redundancy and Identifiability*, 1st ed. Chapman and Hall/CRC, New York.
703 <https://doi.org/10.1201/9781315120003>

704 Cury, P.M., Fromentin, J.-M., Figuet, S., Bonhommeau, S., 2014. Resolving Hjort's Dilemma: How Is Recruitment
705 Related to Spawning Stock Biomass in Marine Fish? *Oceanography* 27, 42–47.

706 Daverat, F., Limburg, K., Thibault, I., Shiao, J., Dodson, J., Caron, F., Tzeng, W., Iizuka, Y., Wickström, H., 2006.
707 Phenotypic plasticity of habitat use by three temperate eel species, *Anguilla anguilla*, *A. japonica* and
708 *A. rostrata*. *Mar. Ecol. Prog. Ser.* 308, 231–241. <https://doi.org/10.3354/meps308231>

709 Dekker, W., 2019. The history of commercial fisheries for European eel commenced only a century ago. *Fish*
710 *Manag Ecol* 26, 6–19. <https://doi.org/10.1111/fme.12302>

711 Dekker, W., 2004. Slipping through our hands : population dynamics of the European eel.

712 Dekker, W., 2003. On the distribution of the European eel *Anguilla anguilla* and its fisheries. *Can. J. Fish. Aquat.*
713 *Sci.* 60, 787–799. <https://doi.org/10.1139/f03-066>

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

717 Delle Monache, D., Harvey, A., 2012. The effect of misspecification in models for extracting trends and cycles.
718 6th Eurostat Colloquium on Modern Tools for Business Cycle Analysis: the lessons from global
719 economic crisis, held in Luxembourg, 26th - 29th September 2010 1–48.
720 <https://doi.org/10.2901/1977-3331.2011.013>

721 Denwood, M.J., 2016. *runjags* : An R Package Providing Interface Utilities, Model Templates, Parallel Computing
722 Methods and Additional Distributions for MCMC Models in JAGS. *J. Stat. Soft.* 71.
723 <https://doi.org/10.18637/jss.v071.i09>

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

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

731 EU, 2007. Council regulation No 1100/2007 establishing measures for the recovery of the stock of European
732 eel. *Official Journal of the European Union* L248, 17 23.

733 Freyhof, J., Kottelat, M., 2008. *Anguilla anguilla*. The IUCN Red List of Threatened Species 2008:
734 e.T60344A12354180. IUCN Red List of Threatened Species.

735 Gascuel, D., Coll, M., Fox, C., Guénette, S., Guitton, J., Kenny, A., Knittweis, L., Nielsen, J.R., Piet, G., Raid, T.,
736 Travers-Trolet, M., Shephard, S., 2016. Fishing impact and environmental status in European seas: a

737 diagnosis from stock assessments and ecosystem indicators. *Fish and Fisheries* 17, 31–55.
738 <https://doi.org/10.1111/faf.12090>

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

741 Gelman, A., Rubin, D.B., 1992. Inference from Iterative Simulation Using Multiple Sequences. *Statist. Sci.* 7.
742 <https://doi.org/10.1214/ss/1177011136>

743 George, E.I., McCulloch, R.E., 1993. Variable Selection Via Gibbs Sampling. *Journal of the American Statistical*
744 *Association* 88, 881–889. <https://doi.org/10.2307/2290777>

745 Gupta, H., Sorooshian, S., Yapo, P., 1999. Status of Automatic Calibration for Hydrologic Models: Comparison
746 With Multilevel Expert Calibration. *Journal of Hydrologic Engineering - J HYDROL ENG* 4.
747 [https://doi.org/10.1061/\(ASCE\)1084-0699\(1999\)4:2\(135\)](https://doi.org/10.1061/(ASCE)1084-0699(1999)4:2(135))

748 Harvey, A.C., 1986. Analysis and Generalisation of a Multivariate Exponential Smoothing Model. *Management*
749 *Science* 32, 374–380. <https://doi.org/10.1287/mnsc.32.3.374>

750 Hilborn, R., Stokes, K., 2010. Defining Overfished Stocks: Have We Lost the Plot? *Fisheries* 35, 113–120.
751 <https://doi.org/10.1577/1548-8446-35.3.113>

752 Hilborn, R., Walters, C.J., 2013. *Quantitative fisheries stock assessment: choice, dynamics and uncertainty.*
753 Springer Science & Business Media.

754 ICES, 2023. Working Group on American Eel (WGAMEEL; outputs from 2022 meeting), ICES Scientific Reports.
755 ICES. <https://doi.org/10.17895/ICES.PUB.21762842>

756 ICES, 2022. Joint EIFAAC/ICES/GFCM Working Group on Eels (WGEEL), ICES Scientific Reports. ICES.
757 <https://doi.org/10.17895/ICES.PUB.20418840.V2>

758 ICES, 2020. Workshop on the temporal migration patterns of European eel (WKEELMIGRATION), ICES Scientific
759 Reports. ICES. <https://doi.org/10.17895/ices.pub.5993>

760 ICES, 2019. Joint EIFAAC/ICES/GFCM Working Group on Eels, ICES Scientific Reports. ICES.
761 <https://doi.org/10.17895/ices.pub.5545>

762 ICES, 2014. Report of the Joint GFCM/EIFAAC/ICES Working Group on Eels (WGEEL), ICES Expert Group reports
763 (until 2018). ICES. <https://doi.org/10.17895/ICES.PUB.19282928.V1>

764 ICES, 2013. Report of the Joint EIFAAC/ICES Working Group on Eels (WGEEL) (report), ICES Expert Group reports
765 (until 2018). ICES. <https://doi.org/10.17895/ices.pub.19282142.v1>

766 ICES, 2006. Report of the ICES Advisory Committee on Fisheries Management, Advisory Committee on the
767 Marine Environment, and Advisory Committee on Ecosystems. ICES Advice. Books 1–10.

768 ICES, 2003. Study Group on Precautionary Reference Points For Advice on Fishery Management (SGPRP), ICES
769 Expert Group reports (until 2018). ICES. <https://doi.org/10.17895/ices.pub.19265519>

770 ICES, 2001. Report of the ICES/EIFAC Working Group on Eels (WGEEL), ICES Expert Group reports (until 2018).
771 ICES. <https://doi.org/10.17895/ICES.PUB.19264436.V1>

772 ICES, EIFAC, 2010. Report of the 2010 session of the Joint EIFAC/ICES Working Group on Eels (WGEEL) and
773 Country Reports 2009–2010 (report). ICES Expert Group reports (until 2018).
774 <https://doi.org/10.17895/ices.pub.19280648.v1>

775 Ishwaran, H., Rao, J.S., 2005. Spike and slab variable selection: Frequentist and Bayesian strategies. *Ann. Statist.*
776 33. <https://doi.org/10.1214/009053604000001147>

777 Kell, L.T., Sharma, R., Kitakado, T., Winker, H., Mosqueira, I., Cardinale, M., Fu, D., 2021. Validation of stock
778 assessment methods: is it me or my model talking? *ICES Journal of Marine Science* 78, 2244–2255.
779 <https://doi.org/10.1093/icesjms/fsab104>

780 Lebot, C., 2021. Evaluation de l'impact de différents scénarii de gestion pour réguler l'exploitation par pêche du
781 saumon atlantique (Salma Salar).

782 Legault, C.M., Wiedenmann, J., Deroba, J.J., Fay, G., Miller, T.J., Brooks, E.N., Bell, R.J., Langan, J.A., Cournane,
783 J.M., Jones, A.W., Muffley, B., 2023. Data-rich but model-resistant: an evaluation of data-limited
784 methods to manage fisheries with failed age-based stock assessments. *Can. J. Fish. Aquat. Sci.* 80, 27–
785 42. <https://doi.org/10.1139/cjfas-2022-0045>

786 Lehuta, S., Girardin, R., Mahévas, S., Travers-Trolet, M., Vermard, Y., 2016. Reconciling complex system models
787 and fisheries advice: Practical examples and leads. *Aquat. Living Resour.* 29, 208.
788 <https://doi.org/10.1051/alr/2016022>

789 Ljunggren, L., Sandström, A., Bergström, U., Mattila, J., Lappalainen, A., Johansson, G., Sundblad, G., Casini, M.,
790 Kaljuste, O., Eriksson, B.K., 2010. Recruitment failure of coastal predatory fish in the Baltic Sea
791 coincident with an offshore ecosystem regime shift. *ICES Journal of Marine Science* 67, 1587–1595.
792 <https://doi.org/10.1093/icesjms/fsq109>

793 Mangel, M., Brodziak, J., DiNardo, G., 2010. Reproductive ecology and scientific inference of steepness: a
794 fundamental metric of population dynamics and strategic fisheries management. *Fish and Fisheries* 11,
795 89–104. <https://doi.org/10.1111/j.1467-2979.2009.00345.x>
796 Maunder, M.N., Deriso, R.B., 2003. Estimation of recruitment in catch-at-age models. *Can. J. Fish. Aquat. Sci.*
797 60, 1204–1216. <https://doi.org/10.1139/f03-104>
798 Maunder, M.N., Thorson, J.T., 2019. Modeling temporal variation in recruitment in fisheries stock assessment:
799 A review of theory and practice. *Fisheries Research, Recruitment: Theory, Estimation, and Application*
800 in *Fishery Stock Assessment Models* 217, 71–86. <https://doi.org/10.1016/j.fishres.2018.12.014>
801 Miller, M., Kimura, S., Friedland, K., Knights, B., Kim, H., Jellyman, D., Tsukamoto, K., 2009. Review of Ocean-
802 Atmospheric Factors in the Atlantic and Pacific Oceans Influencing Spawning and Recruitment of
803 Anguillid Eels. *American Fisheries Society Symposium* 69, 231–249.
804 Ministère de l'Écologie, de l'Énergie, du Développement durable et de l'Aménagement du Territoire, Onema,
805 Ministère de l'Agriculture et de la Pêche, 2010. Plan de gestion Anguille de la France - Application du
806 règlement (CE) n°1100/2007 du 18 septembre 2007 - Volet national.
807 Moriarty, C., 1990. European Catches of Elver of 1928-1988. *Internationale Revue der gesamten Hydrobiologie*
808 und *Hydrographie* 75, 701–706. <https://doi.org/10.1002/iroh.19900750603>
809 Mueter, F.J., Boldt, J.L., Megrey, B.A., Peterman, R.M., 2007. Recruitment and survival of Northeast Pacific
810 Ocean fish stocks: temporal trends, covariation, and regime shifts. *Can. J. Fish. Aquat. Sci.* 64, 911–
811 927. <https://doi.org/10.1139/f07-069>
812 Myers, R.A., 2001. Stock and recruitment: generalizations about maximum reproductive rate, density
813 dependence, and variability using meta-analytic approaches. *ICES Journal of Marine Science* 58, 937–
814 951. <https://doi.org/10.1006/jmsc.2001.1109>
815 Myers, R.A., 1998. When Do Environment–recruitment Correlations Work? *Reviews in Fish Biology and*
816 *Fisheries* 8, 285–305. <https://doi.org/10.1023/A:1008828730759>
817 Needle, C.L., 2001. Recruitment models: diagnosis and prognosis. *Reviews in Fish Biology and Fisheries* 11, 95–
818 111. <https://doi.org/10.1023/A:1015208017674>
819 NEFSC, 2020. Northeast Fisheries Science Center (NEFSC) - Draft Report of the Index Based Methods Working
820 Group. Woods Hole, Massachusetts.
821 Perretti, C., Fogarty, M., Friedland, K., Hare, J., Lucey, S., McBride, R., Miller, T., Morse, R., O'Brien, L., Pereira,
822 J., Smith, L., Wuenschel, M., 2017. Regime shifts in fish recruitment on the Northeast US Continental
823 Shelf. *Mar. Ecol. Prog. Ser.* 574, 1–11. <https://doi.org/10.3354/meps12183>
824 Pita, P., Fernández-Vidal, D., García-Galdo, J., Muíño, R., 2016. The use of the traditional ecological knowledge
825 of fishermen, cost-effective tools and participatory models in artisanal fisheries: Towards the co-
826 management of common octopus in Galicia (NW Spain). *Fisheries Research, The use of fishing vessels*
827 *as scientific platforms* 178, 4–12. <https://doi.org/10.1016/j.fishres.2015.07.021>
828 Plagányi, É.E., Haywood, M.D.E., Gorton, R.J., Siple, M.C., Deng, R.A., 2019. Management implications of
829 modelling fisheries recruitment. *Fisheries Research, Recruitment: Theory, Estimation, and Application*
830 in *Fishery Stock Assessment Models* 217, 169–184. <https://doi.org/10.1016/j.fishres.2019.03.007>
831 Plummer, M., 2003. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *Working*
832 *Papers* 8. <http://www.ci.tuwien.ac.at/Conferences/DSC-2003/Drafts/Plummer.pdf>
833 Punt, A.E., Tuck, G.N., Day, J., Canales, C.M., Cope, J.M., De Moor, C.L., De Oliveira, J.A.A., Dickey-Collas, M.,
834 Elvarsson, B.P., Haltuch, M.A., Hamel, O.S., Hicks, A.C., Legault, C.M., Lynch, P.D., Wilberg, M.J., 2020.
835 When are model-based stock assessments rejected for use in management and what happens then?
836 *Fisheries Research* 224, 105465. <https://doi.org/10.1016/j.fishres.2019.105465>
837 Ricker, W.E., 1954. Stock and Recruitment. *J. Fish. Res. Bd. Can.* 11, 559–623. <https://doi.org/10.1139/f54-039>
838 Righton, D., Westerberg, H., Feunteun, E., Økland, F., Gargan, P., Amilhat, E., Metcalfe, J., Lobon-Cervia, J.,
839 Sjöberg, N., Simon, J., Acou, A., Vedor, M., Walker, A., Trancart, T., Brämick, U., Aarestrup, K., 2016.
840 Empirical observations of the spawning migration of European eels: The long and dangerous road to
841 the Sargasso Sea. *Sci. Adv.* 2, e1501694. <https://doi.org/10.1126/sciadv.1501694>
842 Röckmann, C., Ulrich, C., Dreyer, M., Bell, E., Borodzicz, E., Haapasaaari, P., Hauge, K.H., Howell, D., Mäntyniemi,
843 S., Miller, D., Tserpes, G., Pastoors, M., 2012. The added value of participatory modelling in fisheries
844 management – what has been learnt? *Marine Policy* 36, 1072–1085.
845 <https://doi.org/10.1016/j.marpol.2012.02.027>
846 Schnute, J.T., 1994. A General Framework for Developing Sequential Fisheries Models. *Can. J. Fish. Aquat. Sci.*
847 51, 1676–1688. <https://doi.org/10.1139/f94-168>

848 Smith, A., 1999. Implementing effective fisheries-management systems – management strategy evaluation and
849 the Australian partnership approach. ICES Journal of Marine Science 56, 967–979.
850 <https://doi.org/10.1006/jmsc.1999.0540>

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

854 STECF, 2023a. Scientific, Technical and Economic Committee for Fisheries (STECF) - 73rd PLENARY REPORT
855 (STECF-PLN-23-02)., Publications Office of the European Union. Luxembourg.

856 STECF, 2023b. Scientific, Technical and Economic Committee for Fisheries (STECF) - 72nd plenary report (PLEN
857 22 03), Publications Office of the European Union. Luxembourg. [https://doi.org/10.2760/977664,](https://doi.org/10.2760/977664,JRC133392)
858 JRC133392

859 STECF, 2022. Scientific, Technical and Economic Committee for Fisheries (STECF) - Monitoring of the
860 performance of the Common Fisheries Policy (STECF Adhoc 22 01), Publications Office of the European
861 Union. Luxembourg. [https://doi.org/10.2760/566544, JRC129080](https://doi.org/10.2760/566544,JRC129080)

862 Stocker, M., Noakes, D.J., 1988. Evaluating Forecasting Procedures for Predicting Pacific Herring (*Clupea*
863 *harengus pallasii*) Recruitment in British Columbia. Can. J. Fish. Aquat. Sci. 45, 928–935.
864 <https://doi.org/10.1139/f88-114>

865 Subbey, S., Devine, J.A., Schaarschmidt, U., Nash, R.D.M., 2014. Modelling and forecasting stock–recruitment:
866 current and future perspectives. ICES Journal of Marine Science 71, 2307–2322.
867 <https://doi.org/10.1093/icesjms/fsu148>

868 Sydeman, W.J., García-Reyes, M., Szoboszlai, A.I., Thompson, S.A., Thayer, J.A., 2018. Forecasting herring
869 biomass using environmental and population parameters. Fisheries Research 205, 141–148.
870 <https://doi.org/10.1016/j.fishres.2018.04.020>

871 Tesch, F.W., 2003. The Eel, J.E. Thorpe. ed. Blackwell Science, Oxford, UK.

872 Titsias, M.K., Lázaro-Gredilla, M., 2011. Spike and Slab Variational Inference for Multi-Task and Multiple Kernel
873 Learning 9.

874 Wright, R.M., Piper, A.T., Aarestrup, K., Azevedo, J.M.N., Cowan, G., Don, A., Gollock, M., Rodriguez Ramallo, S.,
875 Velterop, R., Walker, A., Westerberg, H., Righton, D., 2022. First direct evidence of adult European eels
876 migrating to their breeding place in the Sargasso Sea. Sci Rep 12, 15362.
877 <https://doi.org/10.1038/s41598-022-19248-8>

878

879

880

881

882

883

Supplementary Material

Testing novel methods for short-term forecasting of European glass eel recruitment

Authors: Mathilde Benezech^{1,2}, Hilaire Drouineau^{1,2}, Anthony Acou^{3,2}, Agnès Bardonnnet^{4,2}, Clarisse Boulenger^{2,6}, Patrick Lambert^{1,2}, Etienne Prévost^{4,2}, Laurent Beaulaton^{2,5}

¹ INRAE, UR EABX, 50 avenue de Verdun, 33612 Cestas, France; mathilde.benezech@inrae.fr ; hilaire.drouineau@inrae.fr ; patrick.mh.lambert@inrae.fr

² Management of Diadromous Fish in Their Environment, OFB, INRAE, Institut Agro, UNIV PAU & PAYS ADOUR/E2S UPPA, 35042 Rennes, France

³ UAR OFB-CNRS-MNHN PatriNat, Station Marine du Museum National d'Histoire Naturelle, 35800 Dinard, France ; anthony.acou@ofb.gouv.fr

⁴ INRAE, UMR ECOBIOP, 64310 Saint-Pée-sur-Nivelle, France ; agnes.bardonnnet@inrae.fr ; etienne.prevost@inrae.fr

⁵ OFB, Service Conservation et Gestion Durable des Espèces Exploitées, U3E, 35042 Rennes, France ; laurent.beaulaton@ofb.gouv.fr

⁶ U3E, Unité Expérimentale d'écologie et d'écotoxicologie aquatique, INRAE, OFB, 35000 Rennes, France ; clarisse.boulenger@inrae.fr

1 The spike and slab model

1.1 Density of probability for $diff_a$

As recommended by Georges and McCulloch (1993), we 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 (Fig. 1). Fig 1 confirms that the posterior distribution is different from the prior distribution and that the conditional posterior distribution of $diff_a$ with $\gamma = 1$ and $\gamma = 0$ are distinct.

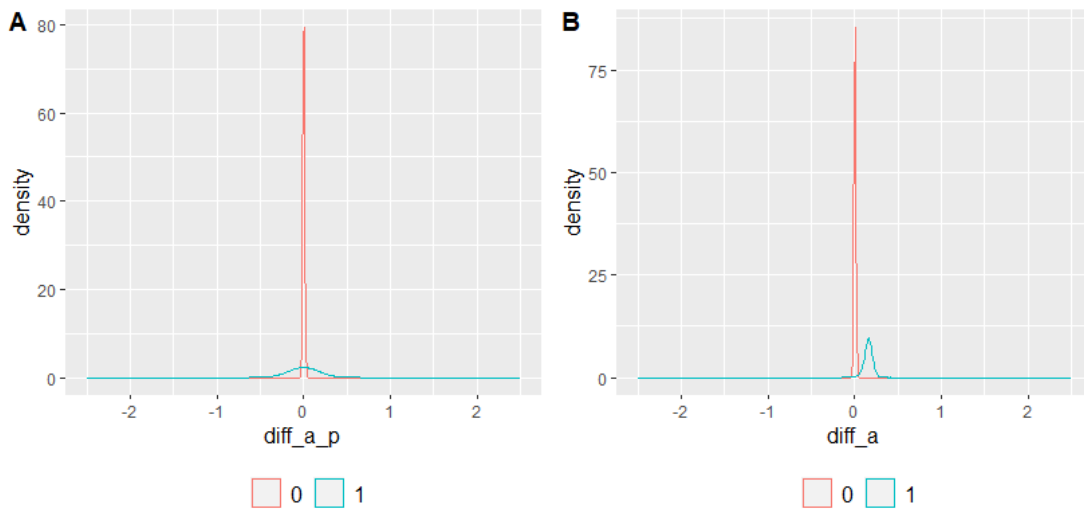


Fig. 1. - Density of probability of the the two normal distributions that define the prior of $diff_a$
(A) - marginal posterior distributions of $diff_a$ depending on the value of γ **B**

1.2 Sensitivity analysis to variance parameters c and d

1.2.1 Setting c and d parameters

Tuning the coefficients c and d in the expression of the variance $\sigma_{diff_a}^2$ (Eq. 4) can be considered subjective (George and 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 $\exp(\text{diff}_a)$ 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 $0.85 \cdot 0.98 = 0.83 \leq \exp(a_2) \leq 0.85 \cdot 1.02 = 0.87$. 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 $c = (\mathbf{0.01})^2 = \mathbf{1.10^{-4}}$.

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 $d = (\mathbf{0.15})^2 = \mathbf{3.10^{-2}}$.

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.

1.2.2 Sensitivity assessment

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 1).

Table 1. Values chosen and their meaning in terms of recruitment evolution to evaluate model sensibility to c and d choices.

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%

We compared the proportion of $\gamma = 1$ in the γ posterior distribution (Fig. 2). For the tested c and d values, the frequency of $\gamma = 1$ is greater than 90%. As such, the sensitivity of the model to c and d appears limited.

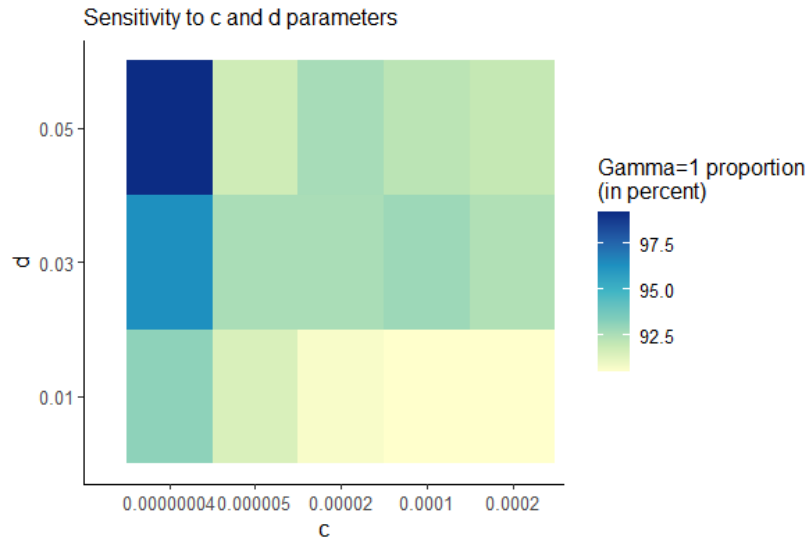


Fig. 2. Sensitivity matrix to values of c and d represented by the proportion of cases where γ is 1 (in percent).

2 Multi-criteria comparison approach

2.1 Retrospective exercise

The metrics to quantify models forecasting performances and the risk for exploitation, resource presented hereby are the initial versions of the ones presented in the article.

In this exercise, we fitted all models (ST, TT, SS and RS) 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. Three chains were run in parallel with 60,000 iterations with a thinning period of ten after a burn-in period of 200,000 iterations.

2.1.1 Material and methods

2.1.1.1 Recruitment prediction error formulation

We compared models forecasting performance by quantifying the discrepancy between recruitment forecasts and “true” values (the WGEEL time series) through a variable called the “Recruitment prediction error”. The difference between each predicted recruitment value and the WGEEL of the corresponding year 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).

2.1.1.2 Evaluation of risks for both exploitation and resource: development of two criteria

We developed two quantitative criteria to lead a quantitative risk assessment (Aven and Renn, 2009). We formulated these criteria thanks to the two following answers.

- **If recruitment is overestimated in year $t+2$, how often is it overestimated 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. 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).

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 :

$$(1) \text{ Overestimation frequency}(t, m) = \frac{\sum_1^{\text{number of iterations}} S_i(t,m)}{\text{number of iterations}}$$

where number of iterations = 60,000.

Therefore, we obtained five frequencies of recruitment overestimation for each model for the retrospective analysis.

- **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, 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).

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.

2.1.2 Results

2.1.2.1 Recruitment prediction error

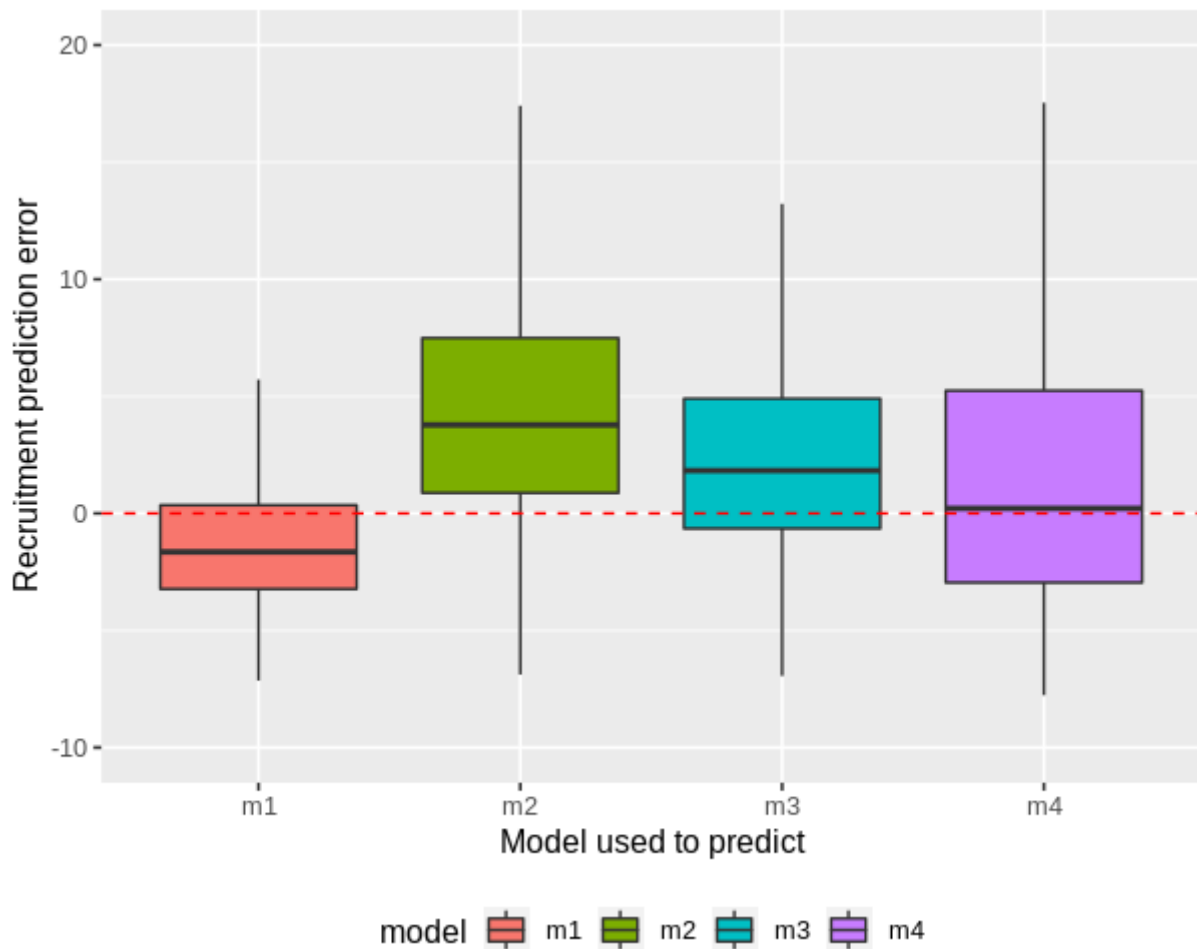


Fig. 3. Boxplots presenting recruitment prediction error for the four recruitment models through the retrospective analysis from 2015 to 2019 (m1: “ST” single-trend model; m2: “TT” two-trend model; m3: “SS” spike and slab model; m4: “RS” random slope model).

According to Figure 3, the ST model mainly underestimates recruitment compared to the ICES WGEEL recruitment index (median prediction error = $-1.64 < 0$). This is consistent with previous observations of the decreasing exponential trend in the model slope. On the other hand, the TT model tends to be over-optimistic compared to WGEEL recruitment index (median = $3.92 > 0$). With a median prediction error of 1.85, the SS model is the intermediate

between the ST model and the TT model. This result is consistent with the structure of the model, which is a mixture of the two models. Nevertheless, Figure 3 shows that the SS model mainly overestimates recruitment. The RS model appears to be the least biased model (median prediction error = 0.96).

2.1.2.2 CRI and FRI

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 4 illustrates the trade-off between conservation and fishery objectives.

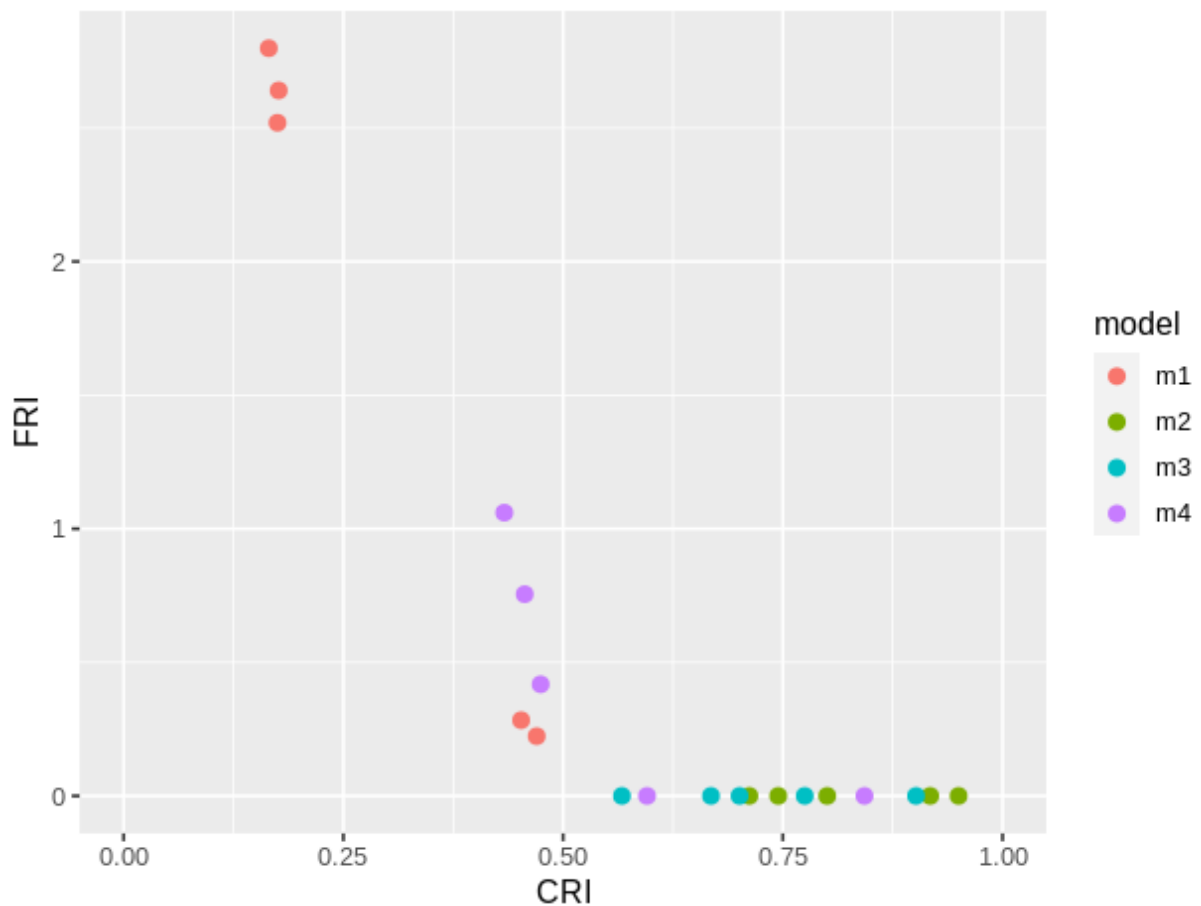


Fig. 4. Position of the four recruitment models when considering simultaneously the Fishery Risk Indicator (FRI) and the Conservation Risk Indicator (CRI) during the retrospective analysis from 2015 to 2019 (m1: “ST” single-trend model; m2: “TT” two-trend model; m3: “SS” spike and slab model; m4: “RS” random slope model).

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 RS model (Figure 4). Indeed, points relative to the RS model are always close to the 50% overestimation frequency threshold and minimize catches loss compared to the ST model. Indeed, the ST model favours compliance with the conservation rule, but at the expense of catches loss for the fishery. On the other hand, the TT and SS models ensure limited catch losses but at the cost of a systematic recruitment overestimation.

2.2 Simulation exercise

2.2.1 Evolution of simulated recruitment series

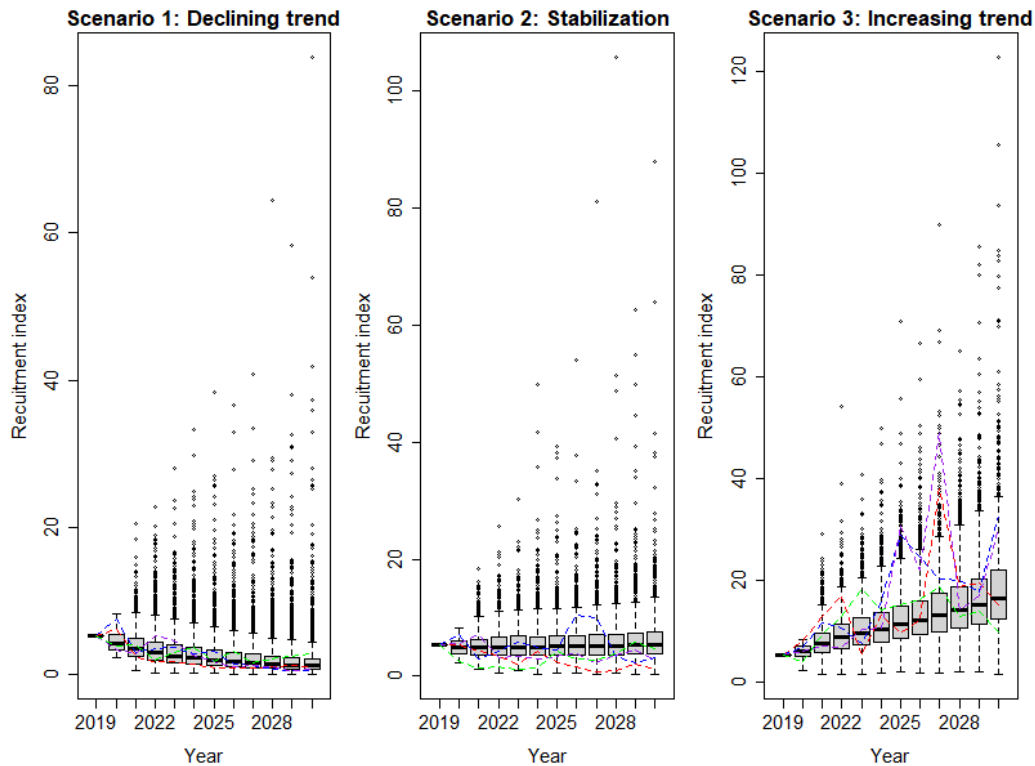


Fig. 5. Evolution of the simulated recruitment index over time according to the three scenarios (Scenario 1: decrease in recruitment; Scenario 2: stabilization of recruitment; Scenario 3: increase in recruitment). The coloured lines represent four time series of recruitment (of the total 2,000 simulated recruitment series) in order to highlight the heterogeneity among individual trajectories.

As expected, the 3 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 above). However, it should be noted that the boxplots hide a great deal of heterogeneity between individual simulated time series. By testing the model on such

contrasted trajectories relying on different assumptions, it is possible to assess the robustness of forecasting models as defined by Box (1979).

2.2.2 Other strategies when calculating quotas

In the main document, we assumed that the quota was set based on the median of the posterior prediction of predicted recruitment (strategy q50). Here, we show how the fishery, management and conservation criteria would be effected if we use the quantiles 25% (strategy q25) or 75% (strategy q75). Main results are not affected though there is a high probability of not achieving the management target with the strategy q75.

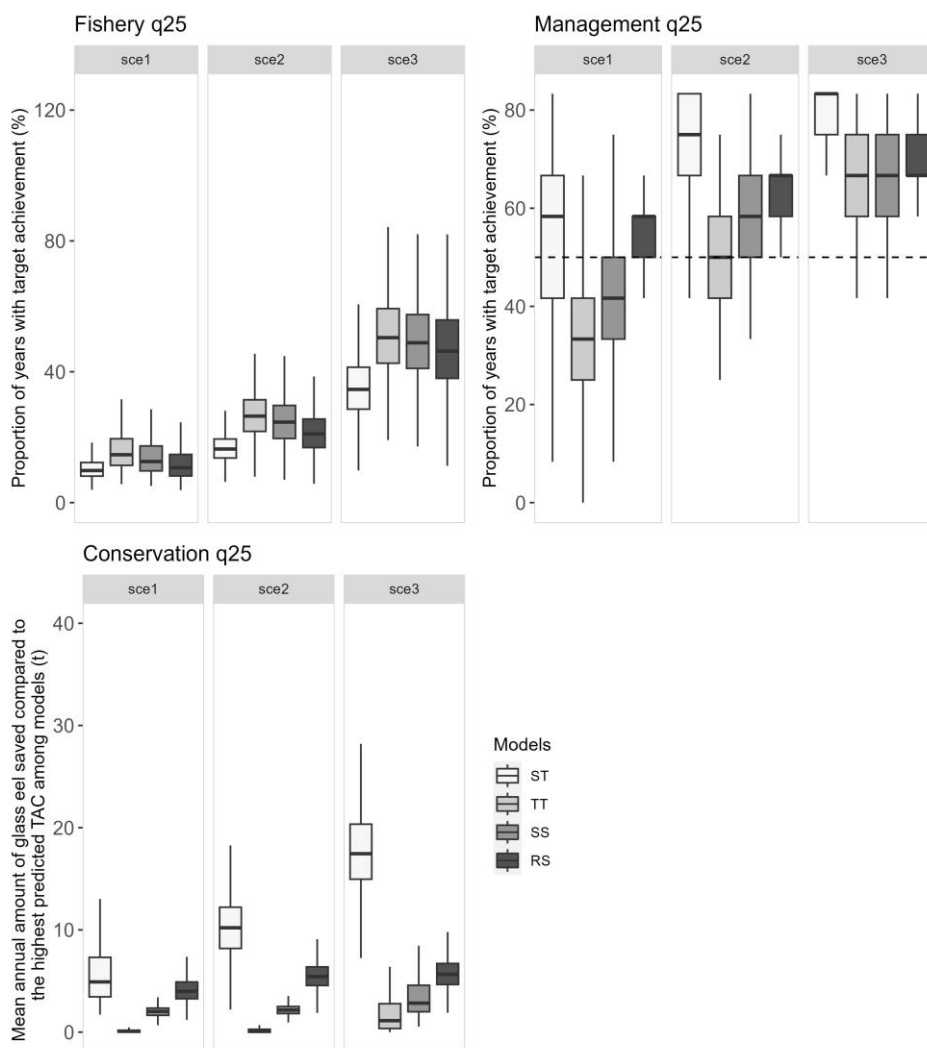


Fig. 6. - Boxplots presenting the three criteria “Fishery”, “Management” and “Conservation across scenario (“sce”) and among models (“ST”: single-trend; “TT”: two-trend; “SS”: spike and slab; “RS”: random slope) over the 10 years of simulation. Outliers not displayed for a better readability but considered for the graph and the calculated values (such as median) (strategy q25).

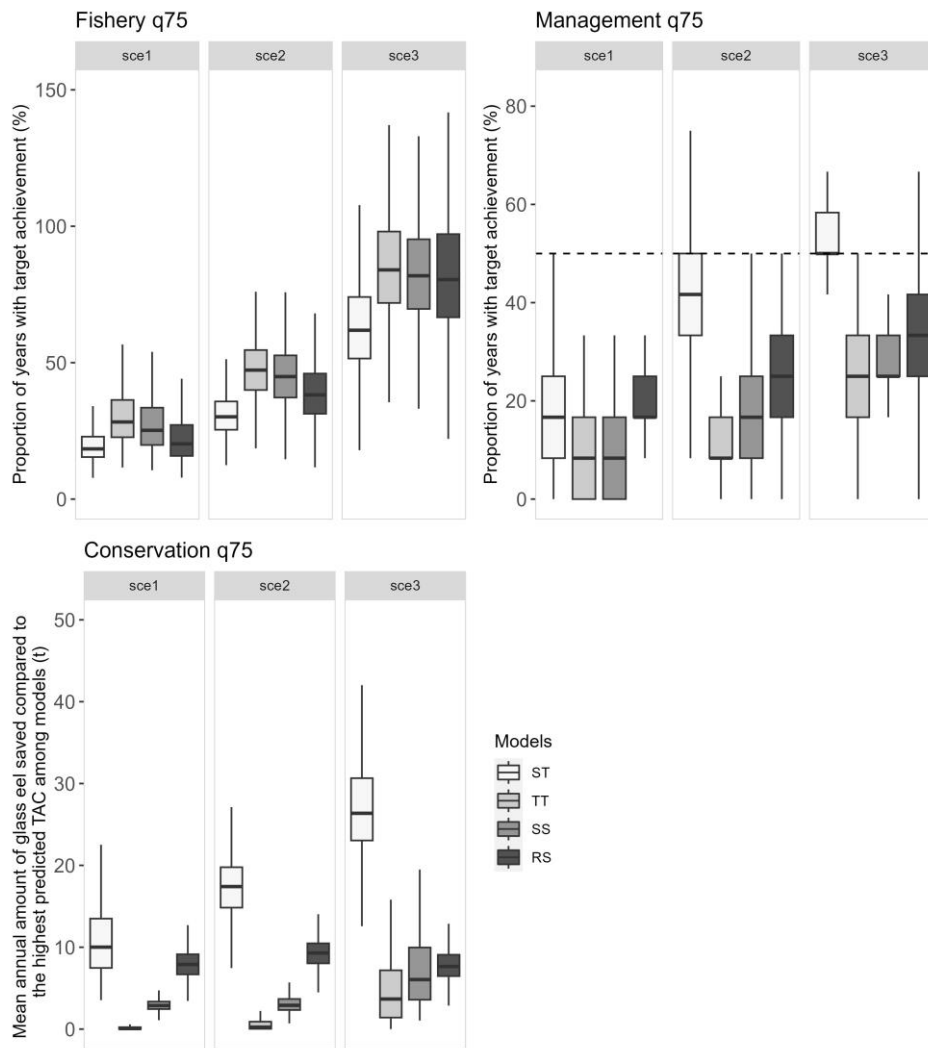


Fig. 7. - Boxplots presenting the three criteria “Fishery”, “Management” and “Conservation across scenario (“sce”) and among models (“ST”: single-trend; “TT”: two-trend; “SS”: spike and slab; “RS”: random slope) over the 10 years of simulation. Outliers not displayed for a better readability but considered for the graph and the calculated values (such as median) (strategy q75).

3 References

- 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, 587–600.
<https://doi.org/10.1111/j.1539-6924.2008.01175.x>
- 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). INRAE; MNHN; AFB - Agence française pour la biodiversité.

- Box, G.E.P., 1979. Robustness in the Strategy of Scientific Model Building, in: Launer, R.L., Wilkinson, G.N. (Eds.), *Robustness in Statistics*. Academic Press, pp. 201–236. <https://doi.org/10.1016/B978-0-12-438150-6.50018-2>
- Dekker, W., 2019. The history of commercial fisheries for European eel commenced only a century ago. *Fish Manag Ecol* 26, 6–19. <https://doi.org/10.1111/fme.12302>
- George, E.I., McCulloch, R.E., 1993. Variable Selection Via Gibbs Sampling. *Journal of the American Statistical Association* 88, 881–889. <https://doi.org/10.2307/2290777>