FISH and FISHERIES WILEY

Temporary Allee effects among non-stationary recruitment dynamics in depleted gadid and flatfish populations

Maria Tirronen [ID] | Tommi Perälä | Anna Kuparinen [ID]

Department of Biological and Environmental Science, University of Jyväskylä, Jyväskylä, Finland

Correspondence
Maria Tirronen, Department of Biological and Environmental Science, University of Jyväskylä, PO Box 35, Jyväskylä 40015, Finland.
Email: maria.j.e.tirronen@jyu.fi

Abstract

Many considerably declined fish populations have not fully recovered despite reductions in fishing pressure. One of the possible causes of impaired recovery is the (demographic) Allee effect. To investigate whether low-abundance recruitment dynamics can switch between compensation and depensation, the latter implying the presence of the Allee effect, we analysed the stock–recruitment time series of 17 depleted cod-type and flatfish populations using a Bayesian change point model. The recruitment dynamics were represented with the sigmoidal Beverton–Holt and the Saila–Lorda stock–recruitment models, allowing the parameters of the models to shift at *a priori* unknown change points. Our synthesis study questions the common assumption that recruitment is stationary and compensatory and the high amount of scatteredness often present in stock–recruitment data is only due to random variation. When a moderate amount of such variation was assumed, stock–recruitment dynamics were best explained by a non-stationary model for 53% of the populations, which suggests that these populations exhibit temporal changes in the stock–recruitment relationship. For four populations, we found shifts between compensation and depensation, suggesting the presence of temporary Allee effects. However, the evidence of Allee effects was highly dependent on the priors of the stock–recruitment model parameters and the amount of random variation assumed. Nonetheless, detection of changes in low-abundance recruitment is essential in stock assessment since such changes affect the renewal ability of the population and, ultimately, its sustainable harvest limits.

KEYWORDS
change point model, compensation, depensation, regime shifts, stock–recruitment

1 | INTRODUCTION

The knowledge of how a depleted population recovers is essential for conservation biology and sustainable harvest of natural resources (Kuparinen et al., 2014). In aquatic systems, collapses and impaired recovery have been particularly well documented (Hutchings, 2015; Neubauer et al., 2013), often in the absence of substantial reductions in habitat carrying capacities. For example, the abundance of the northern Atlantic cod (*Gadus morhua*, Gadidae) was estimated to be 2% of the maximum estimated population size (N_{max}) in 1992 (Hutchings, 2015) and a moratorium on the Northern Cod fishery was declared. Despite the moratorium, the abundance

of the cod stayed below 10%N_{max} for the next 25 years (DFO, 2017; Hutchings, 2015). Similarly, the targeted and by-catch fishing mortality of south Newfoundland white hake (*Urophycis tenuis*, Phycidae) was greatly reduced in 1994, the year in which its abundance was estimated to be 6%N_{max}, but it stayed below 10%N_{max} for the following 15 years (Hutchings, 2015).

One of the possible causes of impaired recovery has been suggested (Neubauer et al., 2013) to be the (demographic) Allee effect, that is, the positive dependence between the *per capita* population growth rate and population size at low abundance (Courchamp et al., 1999). However, the statistical evidence of Allee effects across species and populations is still limited (Gregory et al., 2010; Kramer et al., 2009; Liermann & Hilborn, 2001), including fish populations (Hilborn et al., 2014; Hutchings, 2014; Liermann & Hilborn, 1997; Myers et al., 1995). Finding evidence of Allee effects may have been hindered by methodological and data limitations (Perälä & Kuparinen, 2017). For detecting Allee effects, data at low abundances are crucial as the lack of low-abundance data necessarily leads to large uncertainty about the low-abundance dynamics (Perälä & Kuparinen, 2017). Moreover, evidence of Allee effects can be obscured by dispersion of data due to collecting it at the species level instead of the population level (Perälä & Kuparinen, 2017). Also, at the population level, signals of Allee effects can be buried in noise which is often regarded as being caused by, e.g. the influences of climate or measurement error (Gregory et al., 2010). Yet, one unexplored aspect regarding Allee effects is whether the dependence between the *per capita* population growth rate and population size can change in time and Allee effects remain unnoticed with models not adjustable to such temporal changes. Indeed, it has been suggested that Allee effects can be strengthened by, that is, climate change and harvesting (Winter et al., 2020). More generally, population dynamics may change in time as part of an ecological regime shift, which incorporates persistent changes in the structure and function of the ecosystem (Conversi et al., 2015; Lees et al., 2006; Scheffer et al., 2001; Vasilakopoulos & Marshall, 2015). Many marine and freshwater ecosystems are known to have gone through such regime shifts in their dynamics (Conversi et al., 2015; Möllmann et al., 2015).

Studies exploring Allee effects in natural populations utilize fecundity and mortality data (Liermann & Hilborn, 2001). For marine fish populations, the [*per capita*] birth rate is essentially (Begon et al., 2007) the number of recruits (the number of young fish entering the fished population after surviving the larval phase) per the number of spawners often approximated from the spawning stock biomass (SSB). Here, the number of recruits can be defined as the number of individuals still alive at any specified time after the egg stage but usually, at the time of recruitment to a fishery (Hilborn & Walters, 1992). If the ratio of the number of recruits to SSB is reduced at low SSB (depensatory recruitment or depensation (Needle, 2001), the population exhibits the Allee effect. On the other hand, although a population exhibits compensatory recruitment (the ratio of the number of recruits to SSB is negatively dependent on SSB), the Allee

effect can still occur, if the *per capita* mortality rate is high at low abundance (Neuenhoff et al., 2019).

Knowing the relationship between the number of recruits and SSB is essential when projecting future fish population dynamics and, thus, models for this relationship, i.e. stock–recruitment (S-R) models, play an important role in fisheries management (Hilborn & Walters, 1992; Mangel et al., 2013; Needle, 2001). S-R models traditionally used in fisheries management represent compensatory and time-invariant recruitment dynamics, and the high amount of scatteredness often present in S-R data is explained by random variation from the average S-R relationship. However, compensatory recruitment dynamics of fish populations have been studied with non-stationary models: In the Bayesian framework, the recruitment dynamics of Japanese sardine were modelled with the Ricker model (Ricker, 1954), a hidden Markov model describing two states of productivity (Munch & Kottas, 2009). In addition, in the Bayesian framework, recruitment dynamics of fish populations have been modelled with the Ricker and Beverton–Holt (Beverton & Holt, 1957) models allowing an unknown number of change points at which the parameters of the S-R models may change (Perälä et al., 2017). All of the studied fish populations experienced shifts in their recruitment dynamics and such shifts cannot be captured by a model that assumes time-invariant parameters (Munch & Kottas, 2009; Perälä et al., 2017). Moreover, different statistical models were explored within the Ricker S-R framework including the addition of a stepwise change in the S-R relationship (Ottersen et al., 2013). In 27 of 38 North Atlantic fish populations studied, the S-R relationship was not constant but instead was better explained by a model allowing for a stepwise change (Ottersen et al., 2013). The results with

compensatory models suggest that abrupt changes in S-R relationships are common.

Commercially exploited fishes are a potential group of species to find productivity-related time series at diverse population abundances (Jensen et al., 2012), and thus, they provide an ideal case for studying temporal changes in reproductive dynamics. Furthermore, the study of Allee effects requires that population abundances have declined to the level at which possible Allee effects would manifest (Hutchings, 2014). Time-series data on such populations are most abundantly available in cod-type species and flatfishes. Thus, to investigate whether recruitment can switch between compensation and depensation, the latter implying the Allee effect, we analyse S-R time series of gadids and flatfish populations. Among all such populations available in the RAM Legacy Stock Assessment Database (RAM Database) (Ricard et al., 2012), we select the ones for which both recruit and SSB data are available and which have been depleted to a low abundance at some point during their monitoring period.

We model the S-R dynamics of the chosen fish populations by the sigmoidal Beverton–Holt (SBH) (Needle, 2001) and Saila–Lorda (SL) (Saila et al., 1988) models, which adjust to describing both compensatory and depensatory recruitment, and let the parameter values shift at *a priori* unknown change points. By using non-stationary models, we take an approach that is alternative to the common assumption of a high amount of random variation in S-R data, letting part of the variation be explained by temporal changes in the average S-R relationship. We demonstrate in the Bayesian framework that, when a moderate amount of random variation is assumed, drastic temporal changes in recruitment may be revealed and among them, signals of temporary Allee effects can be discovered. However, we notice that the evidence of Allee effects is highly dependent on the priors of the S-R model parameters and the amount of random variation assumed.

2 | METHODS AND DATA

2.1 | Change point model

We modelled S-R dynamics with a change point model that consists of an S-R model and an unknown number of change points at which the S-R model parameters change. The S-R model relates SSB and the number of recruits with a given lag, and we denote SSB in the years of recording by $S_1, S_2, \ldots, S_T$ and the corresponding numbers of recruits, adjusted to be for age-0 recruits, by $R_1, R_2, \ldots, R_T$. The change points divide the data into segments, $\rho = 1, 2, \ldots, I$, where I is the number of change points. We assumed that the time elapsed since the last change point, the run lengths, $r_t \in \{0, 1, \ldots, t-1\}$, $t = 1, 2, \ldots, T$ form a Markov chain.

The same underlying predictive model was assumed for the data in different segments so that only the parameters of the model vary between segments. The parameters of the underlying predictive model in different segments, denoted by $\boldsymbol{\eta}_\rho$, $\rho = 1, 2, \ldots, I$ were

assumed to be independent and identically distributed. The random variable $\boldsymbol{\eta}_t^{(r)}$, $t = 1, \ldots, T$ denotes the approximation of $\boldsymbol{\eta}_\rho$ at time t, inferred from a run length r_t. Given the segments ρ and the model parameters $\boldsymbol{\eta}_\rho$, we assumed R_t, $t = 1, \ldots, T$ to be independent of each other.

The Bayesian change point model consisted of the following conditional probability distributions:

$$p\left(\log(R_t) \,|\, S_t, \boldsymbol{\eta}_{t-1}^{(r)}\right), t = 1, \ldots, T, \tag{1}$$

$$p\left(\boldsymbol{\eta}_t^{(0)}\right), t = 1, \ldots, T, \tag{2}$$

$$p\left(r_t | r_{t-1}\right), t = 2, \ldots, T \text{ and} \tag{3}$$

$$p\left(r_1\right). \tag{4}$$

Above, the output distribution (1) is defined by the underlying predictive model, and (2) is the joint prior distribution of the model parameters. The transition probability (3) is the change point prior and (4) is the initial run length distribution.

2.2 | Underlying predictive model

For the simplicity of notation, we occasionally omit (r) and t, as well as ρ, in this and the following section. With this notation, the logarithm of R was modelled as a normally distributed random variable, $\log(R) \sim N\left(\mu, \sigma^2\right)$, with segment-wise constant standard deviation, σ and mean, $\mu = \log\left(\widehat{R}(S)\right) - \sigma^2/2$. With this model, R obeys a lognormal distribution with mean $\widehat{R}(S)$ and standard deviation proportional to the mean, $\sqrt{\exp\left(\sigma^2\right) - 1}\widehat{R}(S)$. This is a reasonable choice since there is a tendency for variability about the S-R curve to be higher at larger spawning stock sizes (Hilborn & Walters, 1992). The estimation errors of SSB and the number of recruits were not considered in our model.

The function $\widehat{R}$ describes the average relationship between the number of recruits and SSB in a segment. In this study, we modelled this relationship with two different models, the sigmoidal Beverton–Holt and Saila–Lorda S-R models (Needle, 2001; Saila et al., 1988). The SBH model (Figure 1) can be formulated as:

$$\widehat{R}(S) = \frac{R_\infty}{1 + (S_{50}/S)^c}, \tag{5}$$

so that $\boldsymbol{\eta} = (R_\infty, S_{50}, c)$, where R_∞ is the asymptotic maximum number of recruits, S_{50} is the SSB that produces half of R_∞ and the parameter $c > 0$ controls whether the recruitment dynamics are compensatory or depensatory. This model produces compensatory recruitment when $c \leq 1$ ($c = 1$ corresponds to the traditional Beverton–Holt model) and

(a) Sigmoidal Beverton-Holt
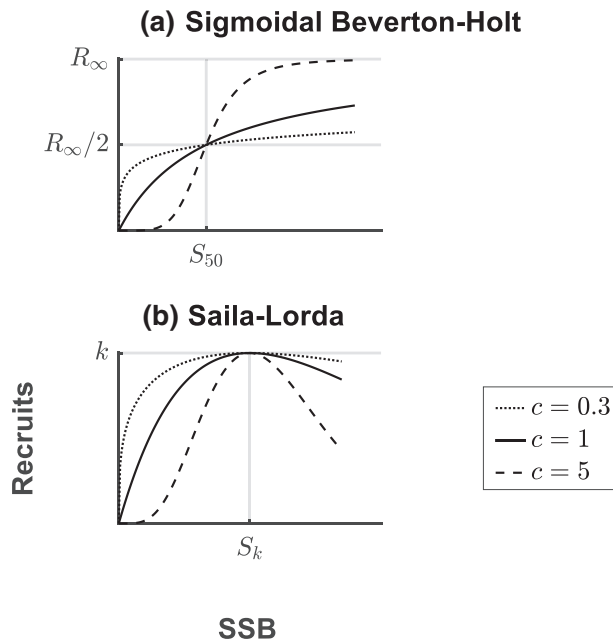
(b) Saila-Lorda

FIGURE 1 The sigmoidal Beverton–Holt (SBH) and Saila–Lorda (SL) stock–recruitment models describe the relationship between the number of recruits (vertical axis) and spawning stock biomass (SSB; horizontal axis). With SBH, recruitment monotonically increases with increasing SSB to the asymptote, R_∞. The amount of SSB that produces $R_\infty/2$ is denoted by S_{50}. SL is domed shaped; it is increasing when SSB $< S_k$ and decreasing when SSB $> S_k$, and its maximum, k, is obtained when SSB equals S_k. With their third parameter, c, SBH and SL allow the possibility of a convex region at low SSB representing depensatory recruitment (Iles, 1994). The models produce compensatory recruitment when $c \leq 1$ ($c = 1$ corresponds to the traditional Beverton–Holt and Ricker models) and depensatory recruitment when $c > 1$. Compensatory and depensatory recruitment are illustrated with $c = 0.3, 1$ and $c = 5$, respectively, while $R_\infty = k = 20,000$ and $S_{50} = S_k = 1,000$

depensatory recruitment when $c > 1$. The SL model (Figure 1) can be parameterized as follows:

$$\widehat{R}(S) = k \left(\frac{S}{S_k}\right)^c \exp\left(c\left(1 - S/S_k\right)\right), \tag{6}$$

so that $\eta = (k, S_k, c)$, where k is the maximum number of recruits, produced when $S = S_k$, and c denotes the depensation parameter as for SBH.

2.3 | Prior distributions for the underlying predictive model

A priori, we assumed the parameters of the underlying predictive model to be independent. For each parameter, we chose to use a weakly informative prior: a uniform distribution (Unif) or a mixture of uniform distributions. With such priors, we only constrained the parameter values to realistic ones but did not favour any particular

values in the chosen ranges. Moreover, we carried out sensitivity analysis in terms of the chosen ranges.

The priors of the parameters R_∞ (SBH) and k (SL) were set to $R_\infty, k \sim \text{Unif}\left(R^{\min}, a \cdot R^{\max}\right)$, with $R^{\min} = \min_{t=1,\dots,T} R_t$ and $R^{\max} = \max_{t=1,\dots,T} R_t$ (Table S9). In the empirical data, the coefficient a was set to 1 for populations the time series of which contained SSB values higher than SSB_{msy} (the amount of SSB that corresponds to maximum sustainable yield). For populations with all SSB data below SSB_{msy}, we set $a = 5$. Similarly, the prior distributions of the parameters S_{50} (SBH) and S_k (SL) were set to $S_{50}, S_k \sim \text{Unif}\left(S^{\min}, b \cdot S^{\max}\right)$ with $S^{\min} = \min_{t=1,\dots,T} S_t$ and $S^{\min} = \min_{t=1,\dots,T} S_t$ (Table S9), using the values $b = 1$ and $b = 5$ for populations which had some SSB data above SSB_{msy} or had all SSB data below SSB_{msy} respectively. The values set for a and b were arbitrary and we carried out a sensitivity analysis fitting the models for all populations with several different values of a and b. In method validation with simulated S-R data sets, we set $a = 3$ and $b = 3$.

The prior distribution of the depensation parameter c was set to a mixture of two uniform distributions that gives equal prior probabilities for depensatory and compensatory recruitment: $c \sim (1/2) \cdot \text{Unif}(0.05, 1) + (1/2) \cdot \text{Unif}(1, d)$. This is similar to the prior distribution of c used by Perälä and Kuparinen (2017), except that with the change point detection method used in this study, the density function does not need to be differentiable. In order to keep the value of c strictly positive, the support of the distribution was bounded from below by 0.05. For the upper bound of the support, we set $d = 5$ when the coefficients of the upper bounds of the other S-R model parameters were set to $a = 1$ or 5, and $b = 1$ or 5. In the sensitivity analysis, we also explored higher values for d in the empirical data sets. In method validation, we set $d = 5$.

The standard error σ of the underlying predictive model was also given a uniform prior distribution: $\sigma \sim \text{Unif}\left(\sigma_{reg}, 2\right)$. To see the impact of our assumptions about the amount of random variation present in the S-R dynamics on the results, we carried out a sensitivity analysis in the empirical data by testing the values 0.05, 0.1, 0.3, 0.5, 0.7 for σ_{reg}. In the simulated data sets, we tested the performance of the method in two cases: (i) $\sigma_{reg} = 0.05$ and (ii) σ_{reg} set to the actual value of σ used in simulation. The upper bound of the support was higher than the sample standard deviations of the empirical data sets. By setting the upper bound to a value this high, we did not force any segmentation in the data sets; with this prior, all variation in the data sets could be described by dispersion from the mean S-R curve of a stationary model. If any change points were to be detected, dividing the time series in segments would, presumably, decrease the value of σ. With the uniform prior, we also considered the possibility of the dispersion from the segment-wise mean S-R curve varying between segments.

2.4 | Change point prior

The change point prior was defined using a constant hazard function (Adams & MacKay, 2007):

$$p(r_t|r_{t-1}) = \begin{cases} \dfrac{1}{\lambda} & \text{if } r_t = 0 \\ 1 - \dfrac{1}{\lambda} & \text{if } r_t = r_{t-1} + 1 \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

for $t = 2, \ldots, T$. With this model, the prior probability of a change point is $1/\lambda$. We assumed that a change point occurred before the first data point, i.e. $p(r_1 = 0) = 1$. Moreover, we set a high prior probability for a change point in the empirical data by fitting the model to the data sets with $\lambda = 10$. In method validation, we also tested the impact of a lower prior probability ($\lambda = 50$) on the results.

2.5 | Bayesian inference

To infer the run lengths and the model parameters, we applied the Bayesian online change point detection method (BOCPD; Adams & MacKay, 2007), combined with simulation-based filtering (Liu & West, 2001; Perälä et al., 2017). The method processes data in a sequential manner, updating estimates for the parameter values and computing posterior probabilities of run lengths after each data point. The run length posterior probabilities were used for computing smoothed run length probabilities (Perälä et al., 2017; Appendix S1: Section 1), i.e. run length probabilities in retrospect, given the whole data. The smoothed run length probabilities are not affected by single outliers, and based on them, the data sets were divided into segments. In this, we maximized the product of the smoothed run length probabilities over all possible segmentations (Appendix S1: Section 1.2), obtaining the most likely segmentation (MLS) (Perälä et al., 2017) of each data set. In MLS, we set the maximum number of segments to five in each data set and considered only segments that consisted of at least 3 years, except at the beginning and the end of the time series, since the data cannot inform when the first segment started or the last one ended. Although we focused on MLS, inspection of alternative segmentations of the data is also easy with the chosen approach.

For parameter estimates, we considered the conditional probability distribution of the S-R model parameters given the data belonging to a given segment at time t (the run-specific posterior distribution of the model parameters) (Appendix S1: Section 1.3). We were mainly interested in the marginal posterior distribution of the depensation parameter c and, related to it, we studied the posterior probability of depensation inferred at the end of the segments, denoted by $p_\rho(D)$, $\rho = 1, \ldots, l_s$, where l_s is the number of segments in the given segmentation. This probability simply answers the question of what is the probability that the depensation parameter has a value greater than 1, indicating depensatory recruitment at the end of the segment. In this, we also considered the median of the marginal posterior distribution of c at the end of the segments. When the posterior probability of depensation and the posterior median indicated depensation at the end of a given segment, the results implied that recruitment was depensatory during that particular time period. Moreover, although we considered parameter estimates according

to a certain segmentation, the chosen approach also provides full parameter posterior distributions which account for the uncertainty related to the location of the change points.

When combined with simulation-based filtering, BOCPD is a stochastic method. Thus, there was variation between inferred S-R dynamics when fitting was repeated by using a different set of random samples in the filter, although the priors were not changed. This variation was decreased when the number of samples was increased, and also more informative priors could make the results more stable. To estimate the robustness of the results obtained in the empirical data with a chosen number of samples in the filter, we first repeated the model fitting twice using different samples. For the populations for which evidence of depensation was found in this initial trial, we repeated the model fitting once more so that, in total, we obtained three replicates for each model and each prior setting for these populations. Initially, we fitted the models using 10^6 samples in the filter. If there were considerable differences in the results obtained by different samples, we increased the number of samples in the filter to 5×10^6 and obtained results with sufficiently low variation.

2.6 | Empirical data

Of the various fish populations in RAM Database, we focused on cod-type species and flatfishes, extracting the populations that could potentially exhibit the Allee effect according to a suggested reference point (Hutchings, 2014). More specifically, we chose populations which met the Allee effect abundance threshold (Hutchings, 2014), i.e. SSB had been below $0.26\,\text{SSB}_{msy}$ in some year during the monitoring period. All populations that met such criterion did not have both recruit and SSB data available, and in total, we studied 17 populations (Table 1). Also, in RAM Database, the recruit time series are adjusted to be for age-0 recruits when that age data are available, and for some populations, S-R data were partly missing at the beginning or end of the time series. Such years were not included in the analysis and consequently, one of the chosen time series (Yellowtail flounder Southern New England/Mid-Atlantic) did not include the data of the only year in which SSB had been below the Allee effect abundance threshold. However, the preceding recorded value of SSB was very close to the threshold, with respect to the range of SSB values in the data, and thus, we regarded this population to have data at low abundance. Moreover, one of the populations (Petrale sole Pacific Coast) had data from 1876 onwards but at the beginning of the time series, there were long periods at which recruitment was constant. These early records were not regarded as reliable estimates and we chose to analyse only data from 1940 onwards, the latest estimates being from 2016. Overall, the length of the analysed time series ranged from 7 to 76 years.

2.7 | Data simulation for method validation

To validate our method, we simulated data that resemble S-R time series of depleted populations which have changes in their

TABLE 1 The names of the 17 cod-type or flatfish populations studied, the number of segments as well as the years of change points according to the most likely segmentation (MLS) using different values of σ_{reg}, and the time periods included in the analysis

Population name\σ_{reg}	The number of segments and the years of change points in MLS			Time period
	0.3	0.5	0.7	
Atlantic halibut (*Hippoglossus*, Pleuronectidae) Scotian Shelf and Southern Grand Banks	2–3 (84,05)	1	1	1970–2013
Atlantic cod (*Gadus morhua*, Gadidae) NAFO 2J3KL	1	1	1	1992–2011
Atlantic cod Georges Bank	1	1	1	1978–2009
Atlantic cod Gulf of Maine	1	1	1	2007–2013
Atlantic cod West of Scotland	1	1	1	1981–2016
Flathead sole (*Hippoglossoides elassodon*, Pleuronectidae) Bering Sea and Aleutian Islands	2 (10)	2 (10)	1	1977–2011
Northern rock sole (*Lepidopsetta polyxystra*, Pleuronectidae) Eastern Bering Sea and Aleutian Islands	4 (92,00,06)	2–3 (00,06–07)	2 (07)	1975–2014
Japanese flounder (*Paralichthys olivaceus*, Paralichthyidae) East China Sea	1	1	1	1986–2012
Japanese flounder Sea of Japan North	1	1	1	1999–2013
Petrale sole (*Eopsetta jordani*, Pleuronectidae) Pacific Coast	1	1	1	1940–2015
Summer flounder (*Paralichthys dentatus*, Paralichthyidae) Mid-Atlantic Coast	2 (89)	1	1	1982–2014
Winter flounder (*Pseudopleuronectes americanus*, Pleuronectidae) Southern New England/Mid-Atlantic	2 (98–99)	2 (98)	1	1981–2010
Witch flounder (*Glyptocephalus cynoglossus*, Pleuronectidae) NAFO–5Y	2–3 (88,00–01)	1	1	1982–2011
Yellowtail flounder (*Limanda ferruginea*, Pleuronectidae) Cape Cod/Gulf of Maine	1	1	1	1985–2013
Yellowtail flounder Georges Bank	2 (81–82)	1	1	1973–2007
Yellowtail flounder Southern New England/Mid-Atlantic	3–4 (87,89–90,11)	3 (89,11)	2–3 (89,13)	1973–2013
Yellowfin sole (*Limanda aspera*, Pleuronectidae) Bering Sea and Aleutian Islands	3 (66,84)	1	1	1954–2014

low-abundance recruitment dynamics. In this, we generated data sets with two segments, the first one representing depensatory recruitment dynamics and the second one compensation. This appeared to be the most common order in the empirical data sets in which we found changes between compensation and depensation. The lengths of the depensation and compensation segments were sampled uniformly from the values 3 to 8 and 10 to 12 respectively. Compared to the empirical data sets for which we found depensation evidence, the test data sets were simpler as they resembled only those parts of the empirical time series in which changes between compensation and depensation were found.

We generated S-R data using the SBH and SL models. Assuming a considerable difference in low-abundance recruitment between the segments, we sampled the value of the depensation parameter from the intervals $[2, 3]$ and $[0.2, 1]$ in the segments of depensation and compensation respectively. Values of other S-R parameters were sampled from ranges similar to the values inferred for the empirical

data sets with depensation evidence (Appendix S1: Section 1.4). While the value of S_{50} (or S_k) was kept constant, the value of R_∞ (or k) was set higher in the depensation segment than in the compensation segment, taking into account the possibility of co-operation moving the carrying capacity upwards (Lidicker, 2010).

SSB data were generated such that the SSB values were evenly scattered in the segments and concentrated on small values (Appendix S1: Section 1.4). Moreover, we set the lowest SSB values in the depensation segments to $0.05 \cdot S_{50}$ (or $0.05 \cdot S_k$). In most of the test data sets, we generated SSB values to appear in a random order. However, we also tested the performance of the method in data sets in which the segment-wise SSB values appeared in the order of magnitude. With the sampled SSB data and parameter values, recruit data were generated from the lognormal distribution described above using $\sigma = 0.1, 0.3, 0.5, 0.7$. In this, we required that the generated recruit data were evenly scattered around the mean S-R curves (Appendix S1: Section 1.4). For a given value of σ, we generated 50 data sets for each S-R model.

3 | RESULTS

3.1 | Method validation

The accuracy of the simulation-based filtering method (Liu & West, 2001) in parameter inference played a major role in how accurately parameter values were estimated for the change point model. In the simulated time series, a lack of data and their scatteredness could affect the estimation accuracy of the filter. Overall, estimates were more accurate in the compensation segments than in the depensation segments (Appendix S1: Tables S1, S2, S5, S6), when considered according to the actual segmentations. This reflects the availability of data; the compensation segments were set longer than the depensation segments and covered a wider range of SSB values. Moreover, when dispersion in the data sets increased, the accuracy in parameter estimation decreased. Regularization by setting σ_{reg} to the actual value of σ used in simulation improved the accuracy of the filter in parameter estimation considerably. Particularly, for the depensation parameter, approximation was decent (the root mean square deviations between the posterior medians and the true values were ≤ 0.36; Appendix S1: Tables S5–S6) for $\sigma \leq 0.5$ for both models. When the SSB values appeared in the order of magnitude, the accuracy in parameter estimation was not consistently higher or lower, compared to the unordered data (Table S4).

Moreover, the prior probability of a change point may have a considerable impact on the accuracy of change point detection by MLS and thus, on the accuracy of parameter estimation for the change point model. Often, the change points in the simulated time series were not detected exactly but with a lead or delay. When the prior probability of a change point was set high ($\lambda = 10$), most of the change points (96%–98%) were detected with the maximum difference of two data points between the true and inferred change points for all values of σ for SBH, using $\sigma_{reg} = 0.05$. For SL, the proportion of detected change points was lower (88%–92%) for $\sigma = 0.5, 0.7$. Nonetheless, when $\lambda = 10$ and $\sigma_{reg} = 0.05$, the method produced a considerable amount of false change points, in addition to the actual ones (Appendix S1: Tables S1, S2). For $\sigma = 0.1$, almost all of them occurred at the beginning of the time series where the segment length was not limited, but for larger values of σ, falsely inferred change points could appear in the middle of the compensation and depensation segments. Setting the prior probability of a change point higher ($\lambda = 50$) decreased the number of falsely inferred change points while it did not considerably decrease the number of actual change points found (Appendix S1: Table S3). However, regularization by setting σ_{reg} to the actual value of σ had the most remarkable impact on the accuracy of change point detection. For SBH, all of the change points were detected with the maximum difference of three data points between the actual and inferred ones, and no false alarms of change points were produced even for $\sigma = 0.5$, when $\lambda = 10$. For SL, the results were similar for $\sigma \leq 0.3$, but 8% of the change points remained unnoticed and also false alarms were given (6% of the all change points detected) for $\sigma = 0.5$. When the data were highly scattered ($\sigma = 0.7$), a considerable proportion of detected change points

(8%–12%) were false alarms, despite regularization by the actual value of σ. Also, many of the change points remained undetected (8% for SBH, 24% for SL).

To correctly identify changes between compensation and depensation, the accuracy of the methods in inferring the type of low-abundance recruitment is most essential. Considering the possibility of inaccuracy in estimation of the depensation parameter, we regarded the posterior of c as indicating depensation only when the posterior probability of depensation was high (≥ 0.8) and the posterior median was over 1.3. Applying these criteria, the filter performed, overall, almost perfectly in detecting depensation for SBH for $\sigma \leq 0.5$ and for SL for $\sigma \leq 0.3$, when using $\sigma_{reg} = 0.05$ (Appendix S1: Tables S1, S2). For larger values of σ, the filter performed decently for SBH, but was prone to inferring depensation for SL. Also, when the performance of the filter was tested only on short segments consisting of three data points, the inference of the type of recruitment could be considerably less accurate, depending on the availability of low-abundance data (Table S7). Furthermore, the inaccuracy of change point detection could affect the inference. Nevertheless, for $\sigma \leq 0.3$ for both of the S-R models, the fitted change point models correctly implied depensation for a large proportion (91%–99%) of the data points that actually represented depensation while producing false inference of depensation in 1%–5% of the data points that represented compensation, using $\lambda = 10$ and $\sigma_{reg} = 0.05$. While the method could fail to infer the type of recruitment correctly in falsely inferred segments within the actual segments (for SL) and also when the actual segmentation was found, in most cases, false inference of depensation was caused by a delay in the change point detection. In such a case, the data at the beginning of the compensation segment were located in such a way that they could be regarded as belonging to the depensation segment as well. *Vice versa*, some of the data points in the depensation segments remained undetected due to a lead in change point detection, or the posterior distributions otherwise not meeting the thresholds we set for depensation. For $\sigma = 0.5, 0.7$, the inference of the type of recruitment for the change point model was considerably less accurate for both S-R models.

Regularization by setting σ_{reg} to the actual value of σ improved the accuracy in inferring the type of low-abundance recruitment for the change point model, particularly for SBH for which 93%–100% of the data points representing depensation were detected and $\leq 1\%$ of the data points in the compensation segments were falsely inferred representing depensation for $\sigma \leq 0.5$. For SL, such regularization increased the proportion of detected depensation to 96%–100% for $\sigma \leq 0.5$ and decreased false inference of depensation to $\leq 3\%$ for $\sigma \leq 0.3$. However for SL, the proportion of false inference of depensation remained high (11%) for $\sigma = 0.5$. For both of the models, the inference of low-abundance recruitment was inaccurate for $\sigma = 0.7$, despite the regularization. For $\sigma \leq 0.5$, false inference of depensation was only caused by a delay in the change point detection for SBH. This also hold for SL except that, in a few data sets for SL, when $\sigma = 0.5$, the whole time series was inferred representing depensation, when the data points in different segments

were very close to each other. Depensation remained undetected in some of the data points due to a lead in change point detection, or the posterior distributions otherwise not clearly indicating depensation according to the chosen thresholds. Moreover, on short segments consisting only of three data points, the availability of low-abundance data had a considerable impact on the detection of depensation. While depensation could remain undetected in short segments, false inference of depensation did not occur, using the regularization (Appendix S1: Table S8).

3.2 | Change points in the empirical data

The minimum amount of random variation around the segment-specific mean S-R curves that we assumed to be present in the empirical data had an impact on the number of change points inferred by MLS. Naturally, when a high amount of random variation from the mean curves was assumed and σ_{reg} was set large, more changes in the S-R time series were explained by dispersion from the mean S-R relationship than with a smaller value of σ_{reg}. For example, for $\sigma_{reg} = 0.3$, the maximum likelihood segmentation consisted (Table 1) of more than one segment for nine populations (53% of the total) for both S-R models, while for $\sigma_{reg} = 0.5$ or 0.7, change points were inferred for four (24% of the total) or two populations respectively.

There were some differences in the inferred years of change points when different S-R models and values of σ_{reg} were used. Also, some differences appeared in the results when fitting was repeated using a different set of random samples. For some values of σ_{reg}, some of the change points were not detected with either of the S-R models or when fitting was repeated using other samples. However, when the number of change points matched, the difference in the results was not more than 1 year in most cases, including regularization by 0.3 – 0.7 (Appendix S1: Sections 2.2 and 2.3). Moreover, the prior probability of a change point was set high in the analysis of the empirical data sets and decreasing it would result in fewer change points inferred. Yet, the priors of the S-R parameters had an impact on the inferred recruitment dynamics and the change points detected.

All the change points in the empirical data found by MLS were not related to notable changes in the depensation parameter c. Since MLS does not distinguish between changes in the model parameters, a change point could be inferred based on a considerable shift in any of the model parameters. Also, although the parameters were *a priori* assumed independent, samples representing them could become correlated at the end of the segments (Appendix S1: Section 2.9) and more than one parameter could contribute to an inferred change in the S-R dynamics. Nonetheless, at some of the change points, there was a considerable shift in the marginal posterior distribution of the depensation parameter.

3.3 | Populations with evidence of depensation

For four populations (Northern rock sole Eastern Bering Sea and Aleutian Islands, Summer flounder Mid-Atlantic Coast, Yellowtail

flounder Cape Cod/Gulf of Maine and Yellowtail flounder Southern New England/Mid-Atlantic; Table 2, Figures 2 and 3), we found one segment for each in MLS at which the posterior probability of depensation was over 0.75 at the end of the segment for both S-R models, at least with some value of σ_{reg}, and at which the recorded values of SSB and recruits were relatively low. Previously, Perälä and Kuparinen (2017) reported strong evidence for depensatory recruitment dynamics, when the posterior probability of depensation was 0.74 or higher. Actually, in most of these segments in which our method predicted depensation, the posterior probability of depensation was ≥ 0.90. In these segments of suspected depensation, the medians of the posterior distributions of the depensation parameter were also high (> 1.8) at the end of the segments. Despite this, uncertainty about the value of the depensation parameter remained rather large in these segments. For the Northern rock sole, the methods actually inferred two depensation segments for SBH, but the latter segment did not include data at low SSB and thus, we did not regard the results to provide evidence of depensation in the last segment. Also, there were some populations, for which the method inferred high posterior probabilities of depensation in some time periods but such segments did not include low-abundance data (either SSB was not below or close to the Allee effect threshold or recruitment was higher than in the rest of the data), only one of the S-R models indicated depensation or the posterior predictive distributions did not clearly indicate a good fit to the data.

Although the posterior probabilities of depensation were high in the considered segments for the Northern rock sole and the three flounder populations, and those segments contained low-abundance data, some uncertainty surrounded the existence of depensation. For the yellowtail flounder Cape Cod/Gulf of Maine population, the segment of suspected depensatory recruitment consisted only of 4 years. Yet, although all the SSB values in the depensation segment were below the Allee effect abundance threshold of this population, recruitment was not at its minimum in this segment, compared to the rest of the data (Figure 2e). The evidence of depensation was strongly dependent on the recorded number of recruits in 1987 (Figure 2f). Although MLS grouped the data in 1987 into the first segment, the posterior probability that this data point started a new segment was high, and as it was not clear how long such segment would last, this data point could also be an outlier (Appendix S1: Figure S2a). Moreover, if the years 1987–1988 formed a segment, the posterior probability of depensation would still be high but such segment would have high recruitment (Appendix S1: Figure S3c). For the summer flounder, the depensation segment was 7 years long but the evidence of depensation was dependent on the data in 1988 (Figure 2d), which was the only year in the depensation segment in which SSB had been below the Allee effect abundance threshold. While recruitment was very low in 1988, the rest of the data in the suspected depensation segment corresponded to high or average recruitment compared to the rest of the data (Figure 2c). For the Northern rock sole (Figure 2a and b), the depensation segment consisted of 17 years, but there were only two SSB values (1975–1976) below the Allee effect abundance threshold. In 1974, recruitment had been low (3.4×10^5), but the corresponding value of SSB was

TABLE 2 The posterior probabilities of depensation and the medians of the marginal posterior distributions of the model parameters at the end of the segments according to the most likely segmentation (MLS) for one of the S-R models and a value of σ_{reg} used in regularization for the populations that were found to have evidence of depensation

Population name	Northern rock sole Eastern Bering Sea and Aleutian Islands	Summer flounder Mid-Atlantic Coast	Yellowtail flounder Cape Cod/Gulf of Maine	Yellowtail flounder Southern New England/Mid-Atlantic
Depensation probability	**0.92** (75–91) 0.25 (92–99) 0.55 (00–05) **0.82** (06–14)	**0.97** (82–88) 0.00 (89–10) 0.37 (11–14)	**0.91** (85–87) 0.00 (88–00) 0.00 (01–11) 0.08 (12–13)	0.27 (73–88) 0.00 (89–10) **0.89** (11–13)
Median of c	**1.96** (75–91) 0.56 (92–99) 1.23 (00–05) **3.2** (06–14)	**2.79** (82–88) 0.18 (89–10) 0.78 (11–14)	**2.35** (85–87) 0.28 (88–00) 0.23 (01–11) 0.47 (12–13)	0.62 (73–88) 0.22 (89–10) **2.69** (11–13)
Median of R_∞ (or k)	4.0e+6 (75–91) 1.9e+6 (92–99) 4.3e+6 (00–05) 2.1e+6 (06–14)	230e+6 (82–88) 92e+6 (89–10) 110e+6 (11–14)	68e+6 (85–87) 18e+6 (88–00) 8.8e+6 (01–11) 26e+6 (12–13)	83e+6 (73–88) 7.1e+6 (89–10) 66e+6 (11–13)
Median of S_{50} (or S_k)	140e+3 (75–91) 340e+3 (92–99) 350e+3 (00–05) 580e+3 (06–14)	29e+3 (82–88) 110e+3 (89–10) 130e+3 (11–14)	1.8e+3 (85–87) 6.3e+3 (88–00) 6.9e+3 (01–11) 6.8e+3 (12–13)	7.6e+3 (73–88) 11e+3 (89–10) 12e+3 (11–13)
Median of σ	0.41 (75–91) 0.39 (92–99) 0.39 (00–05) 0.57 (06–14)	0.24 (82–88) 0.17 (89–10) 0.36 (11–14)	0.24 (85–87) 0.18 (88–00) 0.13 (01–11) 0.13 (12–13)	0.90 (73–88) 0.58 (89–10) 0.68 (11–13)
S-R model	SBH	SBH	SBH	SL
Regularization (σ_{reg})	0.3	0.1	0.05	0.3

Note: The models were fitted using three different sets of random samples which resulted in some differences in the posteriors (Appendix S1: Section 2.4). The estimates are given for one case. The posterior distributions of the model parameters were approximated by 10³ samples. The statistics correspond to the results presented in Figures 2 and 3.

not recorded. Yet, although the posterior run length probabilities strongly suggested grouping the data in 1975–1991 into one segment, another possible segmentation of these data would have resulted in a model suggesting that non-stationary recruitment for this population was merely caused by changes in the asymptotic maximum number of recruits (Appendix S1: Figure S3a). For the yellowtail flounder Southern New England/Mid-Atlantic population (Figure 3), the depensation segment consisted only of 3 years (2011–2013) and in all of them, SSB had been above the Allee effect abundance threshold, although SSB in 2013 was close (893) to the threshold, considering the range of SSB values in the whole time series. The following value of SSB in 2014 was very low (243) and below the threshold, but the corresponding number of recruits was missing. Nonetheless, there could be a delay in the detection of the last change point which started the depensation segment. If this was the case, the depensation segment could be considerably longer (Appendix S1: Figures S2b and S3e).

The amount of depensation evidence that we found was dependent on the amount of random variation from the mean S-R curves we assumed to be present in the empirical data. Indeed, for most of the four populations with evidence of depensation, increasing the lower bound of the variance parameter, σ_{reg}, decreased the number of change points that were found and also decreased the evidence of depensation; either the fitted models began to clearly imply compensation, or the uncertainty of low-abundance recruitment increased (Appendix S1: Sections 2.4.2 and 2.7). For example, for the Northern rock sole, setting $\sigma_{reg} = 0.7$ resulted in data at the beginning of the time series showing stronger evidence for compensation than depensation for both S-R models, while setting $\sigma_{reg} = 0.1$ or 0.3 resulted in evidence for depensation. For the summer flounder and the yellowtail flounder Cape Cod/Gulf of Maine population, recruitment was inferred compensatory and stationary for $\sigma_{reg} > 0.3$ and $\sigma_{reg} > 0.1$ respectively. However, for the other yellowtail flounder population, the strongest evidence of depensation was found with $\sigma_{reg} = 0.3$.
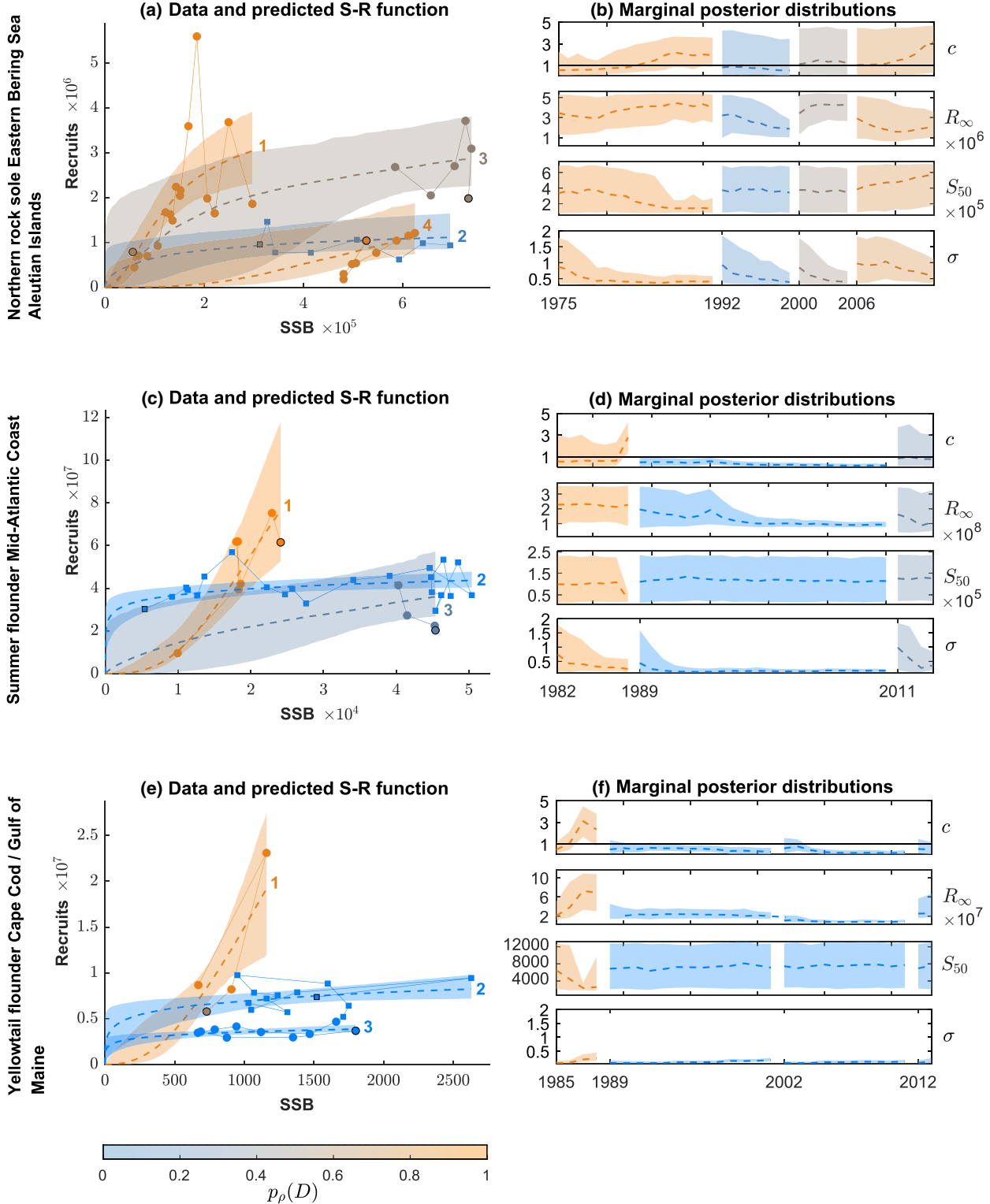
Moreover, the prior distributions of R_∞ and S_{50} (or k and S_k, respectively) had an impact on the inferred S-R dynamics and the amount of depensation evidence we found for some of the populations. For the Northern rock sole, the posterior probability of depensation at the first segment decreased when the upper bound of R_∞ (or k) was increased, particularly for SBH, but the posterior predictive distributions suggest that when the upper bound of the carrying capacity

was set very high, the methods did not fit the models to the data well (Appendix S1: Section 2.7.1). On the other hand, in the first and third segments, the posteriors of R_∞ were, to some extent, concentrated on values close to the upper bound of their support (Figure 2b), suggesting that the asymptotic maximum number of recruits can also be clearly higher than the maximum value recorded, although this population had data in various abundances of SSB. For the yellowtail flounder Cape Cod/Gulf of Maine population, overall, the posterior probability of depensation increased when the upper bounds were increased. However, for the summer flounder, the impact of the priors R_∞ (or k) and S_{50} (or S_k) on the inferred depensation probabilities was not considerable. For the yellowtail flounder Southern New England/Mid-Atlantic population, the data were very scattered and the results obtained by using different S-R models, samples in the filter and values of σ_{reg} varied considerably. Nonetheless, since this population had data at various abundances, the maximum recorded values for the number of recruits and SSB were regarded as decent upper bounds for the values of R_∞ (or k) and S_{50} (or S_k).

Despite the sensitivity, the results obtained by MLS and setting the value of σ_{reg} low suggest that recruitment can change in time from depensation to compensation and particularly, *vice versa*. Evidence for the former was found in the data sets of three populations (the Northern rock sole, the summer flounder and the yellowtail flounder Cape Cod/Gulf of Maine population; Table 2, Figure 2), while only the yellowtail flounder Southern New England/Mid-Atlantic population showed evidence for a change from compensation to depensation (Table 2, Figure 3). All of these populations had at least one segment at which recruitment was inferred compensatory and which appeared immediately after or before the depensation segment, although for the Northern rock sole, such segment did not include data at low abundance.

For these four populations with evidence of temporary depensation, the inferred change points could also be related to considerable changes in other model parameters than the depensation parameter (Table 2). For the Northern rock sole, there were considerable shifts in the posterior medians of R_∞, which, approximately, alternated between two values, the higher one of them being twice as large as the lower one. However, in all of the segments, uncertainty about the value of R_∞ was large (Figure 2b). For the summer flounder, the posterior medians indicated a considerable drop in the value of R_∞ at the first change point (1989), followed by a smaller change at the second change point, but uncertainty about
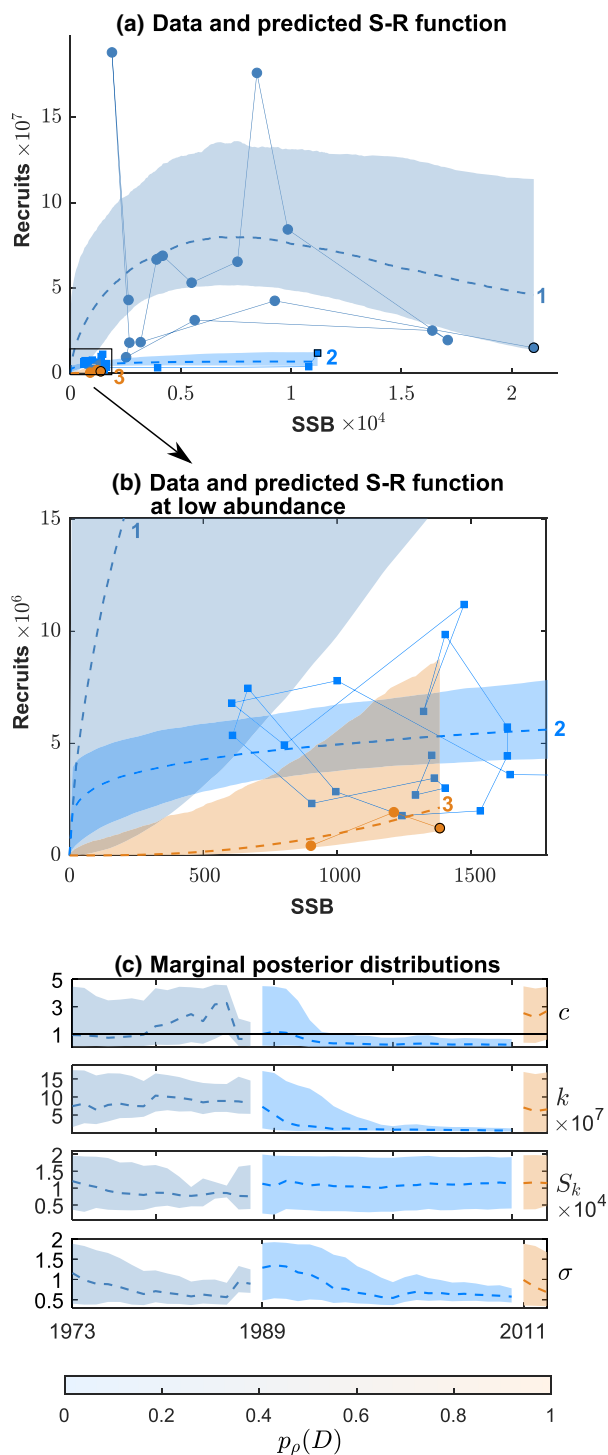
FIGURE 2 The 90% credible intervals of the stock–recruitment function $\widehat{R}$ (average recruitment; figures a, c and e) in each inferred segment with respect to spawning stock biomass (SSB), and the marginal posterior distributions of the model parameters (figures b, d and f) after each iteration step in the segments for populations with evidence of depensation for which recruitment changed from depensatory to compensatory. The method updates the parameter values at every time step when a new data point arrives and after going through the whole time series, finds its most likely segmentation. For approximation of $\widehat{R}$, we considered the posterior distributions of the model parameters at the end of the inferred segments. The segments are illustrated with colours scaled to the segment-wise posterior probability of depensation, $p_\rho(D)$, inferred at the end of the segments. The shaded areas and the dashed lines within them show the intervals between the 5th and the 95th percentiles and the medians respectively. The threshold $c = 1$ between compensation and depensation is plotted with a black line in figures b, d and f. The data (figures a, c and e) are shown with filled circles or squares connected with lines corresponding to the order of recording, and they are coloured according to $p_\rho(D)$. The first data points in the segments have black edges. In figure e, the fourth segment, consisting of only 2 years, is not presented

(a) Data and predicted S-R function
(b) Marginal posterior distributions

(c) Data and predicted S-R function
(d) Marginal posterior distributions

(e) Data and predicted S-R function
(f) Marginal posterior distributions

the value of R_∞ was large in the first and third segments (Figure 2d). Interestingly, similar changes were inferred for the yellowtail flounder populations in 1989, although large uncertainty about the parameter value was again present in their first segment (Figures 2f and 3c). For the yellowtail flounders, the changes in the posterior median of R_∞ (or k) were relatively largest. For all these four populations, there were also changes in the posterior median of S_{50} (or S_k) but in almost all inferred segments, uncertainty about the

parameter value was very large. The posterior medians of the variance parameter σ were, in most cases, close to the value used in regularization. However, the method inferred the high scatteredness in the data of the yellowtail flounder Southern New England/Mid-Atlantic population in 1973–1988 as random variation and produced a posterior concentrated on values clearly greater than σ_{reg} (Figure 3c). In the second segment, the posterior of σ was concentrated on smaller values.

Yellowtail flounder Southern New England / Mid-Atlantic

(a) Data and predicted S-R function

(b) Data and predicted S-R function at low abundance

(c) Marginal posterior distributions

FIGURE 3 The 90% credible interval of the S-R function $\hat{R}$ (average recruitment, figure a) and the marginal posterior distributions of the model parameters (figure c) for the yellowtail flounder Southern New England/Mid-Atlantic population, illustrated in the same way as for populations in Figure 2. For this population, recruitment changed from compensatory to depensatory. Figure b illustrates data at low abundance and reveals a short segment in which recruitment was inferred depensatory

As already noted, there was some synchrony in the inferred change points for the populations for which the evidence of depensation was found. Actually, the results strongly indicated that, for the summer flounder and the yellowtail flounder Southern New England/Mid-Atlantic populations, both of the inferred shifts in their recruitment dynamics appeared exactly at the same years. The first and the third change points of the yellowtail flounder Cape Cod/ Gulf of Maine population differed from these by 1 year at maximum. The second change point of this population differed from the change point inferred for the Northern rock sole in 2000 by 2 years at maximum. However, the difference between the other change points of the Northern rock sole and the change points of other three populations was at least 3 years when $\sigma_{reg} = 0.3$. When σ_{reg} was set a lower value, the difference could be smaller and the method could also infer a change point for the northern rock sole in 2010–2012 (Appendix S1: Tables S11, S12).

4 | DISCUSSION

The results of the present study suggest that low-abundance recruitment dynamics may change in time, but they also demonstrate that such changes can be hard to detect and their actual presence can be difficult to verify. Indeed, the results suggest that recruitment dynamics may change between compensation and depensation, but they also show that signals of temporary depensation in data can easily be interpreted merely as random variation from a compensatory mean S-R curve. The amount of random variation we assumed to be present in the data of the studied 17 fish populations had a considerable impact on the number of change points and the amount of evidence for depensation found in our Bayesian analysis. Overall, we found evidence of depensation for four (24%) populations and, for these populations, the evidence of depensation was not found across the whole monitoring period but in particular time periods when the amount of random variation was assumed moderate (Figures 2 and 3). The results also suggest non-stationary S-R dynamics for most (53%) of the populations studied, if the amount of random variation is assumed moderate, but some of the change points found were not related to changes between compensatory and depensatory recruitment. Since changes at low-abundance recruitment may affect population recovery potential and be related to ecosystem-level regime shifts, the key objective of the present study was to question the common assumption that recruitment is stationary and compensatory and the high amount of scatteredness often seen in S-R data is only due to random variation from a mean compensatory S-R curve.

Occurrence of depensatory recruitment in some time periods implies that the population exhibits the Allee effect in that time period. The number (and proportion) of populations we found with evidence of depensatory recruitment dynamics and thus, Allee effects, was higher than what was found in previous studies (Liermann

& Hilborn, 1997; Myers et al., 1995), in which the evidence of Allee effects was found only for a few of the 128 fish populations studied. Our results are more similar to a more recent study by Perälä and Kuparinen (2017) who found statistical support for Allee effects for four of the nine herring populations they studied. Similar to their study, we restricted our analysis only to populations that had been depleted to low abundance at some point during the monitoring period and gave equal prior probabilities for depensatory and compensatory recruitment. These can be essential factors in finding evidence of Allee effects (Perälä & Kuparinen, 2017). In addition, our results demonstrate the role of detecting temporal changes in the population dynamics in discovering signals of Allee effects. Regarding temporal changes in general, our results are similar to the results by Ottersen et al. (2013), obtained by analysing recruitment dynamics of 38 commercially harvested fish stocks in the northern North Atlantic. The authors (Ottersen et al., 2013) found that the recruitment of approximately 70% of the studied fish populations was best explained by some kind of a change point model, although they assumed compensatory recruitment dynamics.

The Bayesian framework in which we chose to analyse the recruitment dynamics has its key strength in that if, *a priori*, neither compensation nor depensation is favoured and a clear choice between compensatory and depensatory recruitment cannot be made based on the data, uncertainty about the low-abundance recruitment remains large (Perälä & Kuparinen, 2017). The lack of low-abundance data is likely to result in such uncertainty. Thus, to be able to estimate the low-abundance recruitment dynamics, we only considered population which had been depleted to low abundance, but otherwise, we did not set any requirements for the data. However, our analysis pointed out that, to obtain accurate estimates for all of the S-R model parameters, data at high abundance may also be required. When high-abundance data were not available, results were often sensitive to the prior distribution of the (asymptotic) maximum number of recruits. On the other hand, since we used weakly informative priors, uncertainty about the high-abundance recruitment often remained large. Also, by requiring that the populations had been depleted to low abundance at some point of the monitoring period did not, naturally, guarantee that every inferred segment had low-abundance data, and in many cases, uncertainty about the low-abundance recruitment dynamics remained large. Yet, we did not consider how long the populations had stayed at low abundance, and the populations could have low-abundance data only in 1 or a few years. Overall, some of the depensation segments we found were short (3 or 4 years). However, when testing the performance of the method in simulated time series, the method could estimate the depensation parameter accurately enough for inferring the type of recruitment correctly in most of the cases where the segments were short but contained data at low abundance, when sufficient regularization was applied.

Moreover, the change point detection method we used is very sensitive to changes in sequential data and thus, it is a potential tool for detecting temporal changes in S-R data. However, in the simulated data, when the possible amount of random variation around

the segment-specific mean curves was bounded from below to minimum, the method could infer false change points along with the true ones and fit the S-R model to shorter segments of data than were actually present. This kind of over-fitting could naturally lead to inferring false patterns for the average S-R relationship, including patterns indicating depensation. Regularization by a higher value improved the accuracy of model fitting considerably, yielding sufficiently accurate change point detection and depensation inference for the sigmoidal Beverton–Holt model, except when the amount of random variation was very high in the simulated data sets. Despite regularization, there could be a lead or delay in the detection of the change point, but the chosen approach enables consideration of alternative segmentations as well. In a previous study of non-stationary recruitment using a change point model (Perälä et al., 2017), although the underlying predictive model was different, the authors also limited the amount of random variation that could be present in their model. In the present study, we did not aim to give directions about the appropriate amount of regularization but carried out a systematic and comprehensive analysis about the impact of different regularizations on model fitting on the empirical data. The depensation evidence was found by using a moderate amount of regularization, and it remains a question whether the inferred change points for the populations with evidence of depensation truly indicated changes in their average S-R relationships, and depensation really was present in these populations, or if the patterns resembling depensatory S-R curves were only caused by random variation from stationary S-R relationships. Indeed, if there are recruitment regimes present in the empirical data, what is the realistic amount of random variation in the segments? Naturally, when a stationary model is fitted to S-R data, the amount random variation around its mean curve can be very high.

Without considering any segmentation, it is also an essential question, how well the underlying predictive model can capture the recruitment dynamics of fish populations. Although the underlying predictive model for the S-R relationship was similar to traditional S-R models in their assumptions, these assumptions may not be realistic. Our model did not consider any dependence in consecutive SSB values. Also, we assumed that the yearly numbers of recruits are independent, given the parameter values. However, the order of magnitude that was present in some of the SSB data suggest that the value of SSB in a given year might depend on its previous values. Also, the strong positive autocorrelation of the residuals suggests against the independence assumptions (Appendix S1: Section 2.10). Considering such dependences could naturally yield to different kinds of results concerning changes in low-abundance recruitment. The limitations of the model in mimicking real-life data should also be kept in mind when regarding the method validation results. Nonetheless, modelling S-R dynamics by considering temporal dependence remains a topic for future research.

If temporary Allee effects existed, the reasons behind them would naturally be of interest. However, mechanisms underlying temporal changes in the recruitment was not addressed by the model in the present study and including such effects in the model remains

a topic for future research. A more holistic approach to the ecosystem could explain the partial synchrony in the results we obtained and help in judging whether the Allee effects really were present for the studied populations in the first place. Overall, recruitment is a complex process incorporating, at least, adult survival, adult fecundity, juvenile survival and juvenile growth (Begon et al., 2007) and thus, there are many possible drivers behind its temporal changes. For example, fishing may have a considerable impact on recruitment. For fish, size-selective harvesting reduces body size and its variance in the exploited population (Uusi-Heikkilä, 2020). All else being equal, populations with different age structure may differ in their recruitment such that a population of older, larger individuals has a higher recruitment than a population of younger, smaller individuals (Venturelli et al., 2008). Also, harvesting decays the genetic diversity of the exploited population and consequently, inbreeding may result in descendants with decreased fertility (Uusi-Heikkilä, 2020). Recruitment dynamics of a population at low abundance may also change in time due to changes in abundance of its forage populations (predators or competitors of its young). For example, Atlantic cod in the southern Gulf of St. Lawrence was at low abundance in the mid-1970s and 1990s but its recruitment rate differed considerably between these time periods (Swain & Sinclair, 2000). Remarkably high recruitment, and quick recovery, at the end of 1970s, coincided with low abundance of the populations preying on its larvae and eggs (Swain & Sinclair, 2000). Contrarily, the abundance of the forage populations was not reduced in the 1990s which coincides with low recruitment of the cod population (Swain & Sinclair, 2000). Ultimately, changes in the recruitment dynamics of a given population may result from an ecological regime shift that causes changes at the abundance of its forage populations. Moreover, although the mortality rates of the populations also have an impact on their growth rate, we only studied recruitment. Relatively modest increases in cod natural mortality by predation/competition have been shown sufficient to generate a demographic Allee effect (Kuparinen & Hutchings, 2014).

The present study suggests that reproduction may change in time. Most importantly, previously compensatory recruitment may change into depensation. Thus, our results suggest that, when estimating the recovery potential of a population, recruitment should not, *a priori*, be assumed compensatory nor stationary. With an S-R model not adjustable to changes in the recruitment dynamics, patterns in data indicating Allee effects in some time periods may remain unnoticed and be explained merely by random variation from a compensatory mean S-R curve. Taking into account the possibility of temporary Allee effects is especially important in fisheries management, in order to avoid over-exploitation and, eventually, extinction of populations depleted to low abundances (Hilborn & Walters, 1992; Kuparinen et al., 2014). In practice, stock- assessment predictions about the fish population's future development and its responses to alternative harvest strategies should be derived based on full posterior predictive distributions that incorporate uncertainties surrounding the prevailing regime and the depensatory/compensatory nature of the low-abundance dynamics (Perälä & Kuparinen, 2015). From the conservation point of view, the possibility of changes from compensation to depensation

is of importance as such changes may reshape the system response to small shifts in the environmental conditions, to human-induced disturbances and, consequently, evoke catastrophic transitions in the abundance of single populations (Scheffer et al., 2001; Winter et al., 2020), which may ultimately lead to ecological regime shifts concerning the whole ecosystem. Similarly, the risks of population collapses and prolonged recovery times can be highly elevated should a declined population experience a regime shift towards depensatory dynamics (e.g. Kuparinen et al., 2014).

DATA AVAILABILITY STATEMENT

All data that support the plots and other findings in this paper are available from the corresponding author upon reasonable request. Source data and code are available at: https://doi.org/10.5061/dryad.3xsj3txd5

ORCID

Maria Tirronen (iD) https://orcid.org/0000-0001-6052-3186
Anna Kuparinen (iD) https://orcid.org/0000-0002-7807-8946

REFERENCES

Adams, R. P., & MacKay, D. J. C. (2007). *Bayesian online change point detection [Technical report]*. University of Cambridge.

Begon, M., Townsend, C. R., & Harper, J. L. (2007). *Ecology: From individuals to ecosystems*, 4th edn. Blackwell Publishing.

Beverton, R. J. H., & Holt, S. J. (1957). *On the dynamics of exploited fish populations* (Vol. *19*). UK Ministry of Agriculture, Fisheries and Food.

Conversi, A., Dakos, V., Gårdmark, A., Ling, S., Folke, C., Mumby, P. J., Greene, C., Edwards, M., Blenckner, T., Casini, M., Pershing, A., & Möllmann, C. (2015). A Holistic view of Marine Regime shifts. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*, 20130279. https://doi.org/10.1098/rstb.2013.0279

Courchamp, F., Clutton-Brock, T., & Grenfell, B. (1999). Inverse density dependence and the Allee effect. *Trends in Ecology & Evolution*, *14*, 405–410. https://doi.org/10.1016/S0169-5347(99)01683-3

DFO (2017). *Nothern (NAFO Divs. 2J3KL) Cod Stock Update (No. 034)*. DFO Canadian Science Advisory Secretariat.

Gregory, S. D., Bradshaw, C. J. A., Brook, B. W., & Courchamp, F. (2010). Limited evidence for the demographic Allee effect from numerous species across taxa. *Ecology*, *91*, 2151–2161. https://doi.org/10.1890/09-1128.1

Hilborn, R., Hively, D. J., Jensen, O. P., & Branch, T. A. (2014). The dynamics of fish populations at low abundance and prospects for rebuilding and recovery. *ICES Journal of Marine Science*, *71*, 2141–2151. https://doi.org/10.1093/icesjms/fsu035

Hilborn, R., & Walters, C. J. (1992). *Quantitative fisheries stock assessment: Choice, dynamics and uncertainty*. Chapman and Hall.

Hutchings, J. A. (2014). Renaissance of a caveat: Allee effects in marine fish. *ICES Journal of Marine Science*, 71, 2152–2157. https://doi.org/10.1093/icesjms/fst179

Hutchings, J. A. (2015). Thresholds for impaired species recovery. *Proceedings of the Royal Society B: Biological Sciences*, 282, 20150654. https://doi.org/10.1098/rspb.2015.0654

Iles, T. C. (1994). A review of stock-recruitment relationships with reference to flatfish populations. *Netherlands Journal of Sea Research*, 32, 399–420. https://doi.org/10.1016/0077-7579(94)90017-5

Jensen, O. P., Branch, T. A., & Hilborn, R. (2012). Marine fisheries as ecological experiments. *Theoretical Ecology*, 5, 3–22. https://doi.org/10.1007/s12080-011-0146-9

Kramer, A., Dennis, B., Liebhold, A. M., & Drake, J. M. (2009). The evidence for Allee effects. *Population Ecology*, 51, 341–354. https://doi.org/10.1007/s10144-009-0152-6

Kuparinen, A., & Hutchings, J. A. (2014). Increased natural mortality at low abundance can generate an Allee effect in a marine fish. *Royal Society Open Science*, 1, 140075. https://doi.org/10.1098/rsos.140075

Kuparinen, A., Keith, D. M., & Hutchings, J. A. (2014). Allee effect and the uncertainty of population recovery. *Conservation Biology*, 28, 790–798. https://doi.org/10.1111/cobi.12216

Lees, K., Pitois, S., Scott, C., Frid, C., & Mackinson, S. (2006). Characterizing regime shifts in the marine environment. *Fish and Fisheries*, 7, 104–127. https://doi.org/10.1111/j.1467-2979.2006.00215.x

Lidicker, W. Z. (2010). The Allee effect: Its history and future importance. *The Open Ecology Journal*, 3, 71–82. https://doi.org/10.2174/1874213001003010071

Liermann, M., & Hilborn, R. (1997). Depensation in fish stocks: A hierarchic Bayesian meta-analysis. *Canadian Journal of Fisheries and Aquatic Sciences*, 54, 1976–1984. https://doi.org/10.1139/f97-105

Liermann, M., & Hilborn, R. (2001). Depensation: Evidence, models and implications. *Fish and Fisheries*, 2, 33–58. https://doi.org/10.1046/j.1467-2979.2001.00029.x

Liu, J., & West, M. A. (2001). Combined parameter and state estimation in simulation-based filtering. In A. Doucet, N. de Freitas & N. Gordon (eds), *Sequential Monte Carlo methods in practice* (pp. 197–223). Springer.

Mangel, M., MacCall, A., Brodziak, J., Dick, E. J., Forrest, R. E., Pourzand, R., & Ralston, S. (2013). A perspective on steepness, reference points, and stock assessment. *Canadian Journal of Fisheries and Aquatic Sciences*, 70, 930–940. https://doi.org/10.1139/cjfas-2012-0372

Möllmann, C., Folke, C., Edwards, M., & Conversi, A. (2015). Marine regime shifts around the globe: Theory, drivers and impacts. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370, 20130260. https://doi.org/10.1098/rstb.2013.0260

Munch, S., & Kottas, A. (2009). A Bayesian modeling approach for determining productivity regimes and their characteristics. *Ecological Applications*, 19, 527–537. https://doi.org/10.1890/07-2116.1

Myers, R. A., Barrowman, N. J., Hutchings, J. A., & Rosenberg, A. A. (1995). Population dynamics of exploited fish stocks at low population levels. *Science*, 269, 1106–1108. https://doi.org/10.1126/science.269.5227.1106

Needle, C. (2001). Recruitment models: Diagnosis and prognosis. *Reviews in Fish Biology and Fisheries*, 11, 95–111.

Neubauer, P., Jensen, O. P., Hutchings, J. A., & Baum, J. K. (2013). Resilience and recovery of overexploited marine populations. *Science*, 340, 347–349. https://doi.org/10.1126/science.1230441

Neuenhoff, R. D., Swain, D. P., Cox, S. P., McAllister, M. K., Trites, A. W., Walters, C. J., & Hammill, M. O. (2019). Continued decline of a collapsed population of Atlantic cod (*Gadus morhua*) due to predation-driven Allee effects. *Canadian Journal of Fisheries and Aquatic Sciences*, 76, 168–184. https://doi.org/10.1139/cjfas-2017-0190

Ottersen, G., Stige, L. C., Durant, J. M., Chan, K.-S., Rouyer, T. A., Drinkwater, K. F., & Stenseth, N. C. (2013). Temporal shifts in recruitment dynamics of North Atlantic fish stocks: Effects of spawning stock and temperature. *Marine Ecology Progress Series*, 480, 205–225. https://doi.org/10.3354/meps10249

Perälä, T., & Kuparinen, A. (2015). Detecting regime shifts in fish stock dynamics. *Canadian Journal of Fisheries and Aquatic Sciences*, 72, 1619–1628. https://doi.org/10.1139/cjfas-2014-0406

Perälä, T., & Kuparinen, A. (2017). Detection of Allee effects in marine fishes: Analytical biases generated by data availability and model selection. *Proceedings of the Royal Society B: Biological Sciences*, 284, 20171284. https://doi.org/10.1098/rspb.2017.1284

Perälä, T. A., Swain, D. P., & Kuparinen, A. (2017). Examining nonstationarity in the recruitment dynamics of fishes using Bayesian change point analysis. *Canadian Journal of Fisheries and Aquatic Sciences*, 74, 751–765. https://doi.org/10.1139/cjfas-2016-0177

Ricard, D., Minto, C., Jensen, O. P., & Baum, J. K. (2012). Examining the knowledge base and status of commercially exploited marine species with the RAM Legacy Stock Assessment Database: The RAM Legacy Stock Assessment Database. *Fish and Fisheries*, 13, 380–398. https://doi.org/10.1111/j.1467-2979.2011.00435.x

Ricker, W. E. (1954). Stock and Recruitment. *Journal of the Fisheries Research Board of Canada*, 11, 559–623. https://doi.org/10.1139/f54-039

Saila, S. B., Recksiek, C. W., & Prager, M. H. (1988). *Basic fishery science programs*. Elsevier.

Scheffer, M., Carpenter, S., Foley, J. A., Folke, C., & Walker, B. (2001). Catastrophic shifts in ecosystems. *Nature*, 413, 591–596. https://doi.org/10.1038/35098000

Swain, D. P., & Sinclair, A. F. (2000). Pelagic fishes and the cod recruitment dilemma in the Northwest Atlantic. *Canadian Journal of Fisheries and Aquatic Sciences*, 57, 1321–1325. https://doi.org/10.1139/f00-104

Uusi-Heikkilä, S. (2020). Implications of size-selective fisheries on sexual selection. *Evolutionary Applications*, 13, 1487–1500. https://doi.org/10.1111/eva.12988

Vasilakopoulos, P., & Marshall, C. T. (2015). Resilience and tipping points of an exploited fish population over six decades. *Global Change Biology*, 21, 1834–1847. https://doi.org/10.1111/gcb.12845

Venturelli, P. A., Shuter, B. J., & Murphy, C. A. (2008). Evidence for harvest-induced maternal influences on the reproductive rate of fish populations. *Proceedings of the Royal Society B: Biological Sciences*, 276, 919–924. https://doi.org/10.1098/rspb.2008.1507

Winter, A.-M., Richter, A., & Eikeset, A. M. (2020). Implications of Allee effects for fisheries management in a changing climate: Evidence from Atlantic cod. *Ecological Applications*, 30, 1–14. https://doi.org/10.1002/eap.1994

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.