


# Testing models of increasing complexity to develop ecosystem-informed fisheries advice

Matthew D. Robertson<sup>1</sup>  | Noel G. Cadigan<sup>1</sup>  | Paul M. Regular<sup>2</sup>  |  
Mariano Koen-Alonso<sup>2</sup>  | Frédéric Cyr<sup>2</sup>  | Fan Zhang<sup>3</sup>  | Tyler D. Eddy<sup>1</sup> 

<sup>1</sup>Centre for Fisheries Ecosystems

Research, Fisheries and Marine Institute of Memorial University of Newfoundland, St. John's, Newfoundland, Canada

<sup>2</sup>Fisheries and Oceans Canada, Northwest Atlantic Fisheries Centre, St. John's, Newfoundland, Canada

<sup>3</sup>College of Marine Science, Shanghai Ocean University, Shanghai, China

## Correspondence

Matthew D. Robertson, Centre for Fisheries Ecosystems Research, Fisheries and Marine Institute of Memorial University of Newfoundland, P.O. Box 4920, St. John's, NL A1C 5R3, Canada. Email: [matthew.robertson@mi.mun.ca](mailto:matthew.robertson@mi.mun.ca)

## Funding information

Ocean Choice International; Ocean Frontier Institute; Natural Sciences and Engineering Research Council of Canada; Vanier Canada Graduate Scholarship; Canada First Research Excellence Fund; Fisheries and Oceans Canada (DFO)

## Abstract

Despite continued calls for the application of ecosystem-based fisheries management, tactical fisheries management continues to be heavily reliant on single-species stock assessments. These stock assessments rarely quantitatively integrate the effects of ecosystem processes on fish stock productivity. This lack of integration is ultimately driven by the complexity of interactions between populations, ecosystems and fisheries, which produces uncertainty when defining which processes to include and how to include them. Models developed using a structured hypothesis testing framework would allow formalizing uncertainties while underscoring the importance of incorporating different population and ecosystem processes to explain non-stationary stock productivity. Here, we develop a conceptual framework for extending and comparing population dynamics models of increasing complexity. We illustrate the utility of the framework by investigating the population and ecosystem processes that most likely affected the differential recovery of two flatfish populations (American plaice and yellowtail flounder) on the Newfoundland Grand Banks over the past three decades. We found that yellowtail flounder population dynamics were primarily driven by recruitment variability, which was negatively affected by warmer climatological conditions, as indicated by an integrated regional climate index. Meanwhile, American plaice population dynamics were affected by a combination of temporal variability in recruitment and natural mortality, where natural mortality increased during colder than average conditions. By exploring hypotheses about the effects of population and ecosystem processes on population dynamics, this modelling framework will improve understanding about the drivers of shifts in population productivity while serving as a transparent and robust approach to support ecosystem-based fisheries management.

## KEYWORDS

ecosystem-based fisheries management, model comparison, multispecies model, Newfoundland Grand Banks, stock assessment

## 1 | INTRODUCTION

Despite continued calls for the application of ecosystem-based fisheries management (Denit, 2017; DFO, 2007; Koen-Alonso et al., 2019), tactical fisheries management continues to be

primarily based on single-species stock assessments that rarely quantitatively integrate the effects of ecosystem drivers on fish stock productivity (Skern-Mauritzen et al., 2016 but see Marshall et al., 2019; Pepin et al., 2022). Contrary to the historical population dynamics assumption of stationary productivity (i.e., lack of

temporal variability in the population's rate of growth), fish stock productivity varies over time in response to bottom-up (Regular et al., 2022; Szuwalski et al., 2015) and top-down processes (Baum & Worm, 2009; Neuenhoff et al., 2018). Given the increasing ability of state-space stock assessment models to estimate time-varying parameters (Punt et al., 2020; Stock & Miller, 2021), further exploration of the population and ecosystem processes (e.g., mean ocean climate, primary and secondary production) driving variability in stock productivity over time may yield improved management outcomes.

Ecosystem processes ultimately modify population productivity through effects on survival, growth and reproduction (reviewed in Kuparinen et al., 2012). While typical stock assessment models include components of survival (e.g., natural mortality), growth (e.g., size-at-age) and fecundity (e.g., maturity-at-age), they are often treated as fixed inputs and assumed to be stationary in some cases. The assumption of stationarity in survival, growth and fecundity limits our ability to assess the effects of ecosystem processes on population dynamics (Hamel et al., 2023; Lorenzen, 2016). When the assumption of stationary survival, growth and fecundity is violated, any variability will be misattributed to other population (e.g., fishing mortality) or observation processes (e.g., gear selectivity; Pope et al., 2021) that are estimated to vary over time. Such misattributions of process variability will affect management advice and our ability to identify relationships between population processes and ecosystem drivers. Therefore, misattribution of process variability may contribute to the well-known inability of estimated recruitment-environment correlations to persist over time (Haltuch et al., 2019; Myers, 1998).

The complexity of interactions between population dynamics, the environment and fisheries produces multiple plausible hypotheses for the dominant mechanisms of change. Examining models with different assumptions can help identify the relative importance of sources of uncertainty (Heneghan et al., 2021; Reum et al., 2021). While examining alternative hypotheses through model comparisons can become very complex or even intractable with large ecosystem models, Models of Intermediate Complexity for Ecosystem assessment (MICE) and Minimum-Realistic Models (MRM), which focus only on key species and processes to explain unaccounted variability in population dynamics (Collie et al., 2016; Plagányi, 2007; Plagányi et al., 2014), provide a sensible platform between single-species population dynamics models and larger ecosystem models for this type of exploration. However, the inherent complexity in the development of MICE, MRM and other ecosystem models often results in a disconnect in understanding how model formulations and hypotheses were tested to yield the final model. A similar problem exists for the selection of management strategy evaluation (MSE) operating models. MSE is increasingly advocated to provide ecosystem-informed management advice by using different operating models to investigate the performance of harvest control rules against ecosystem-based hypotheses (Goethel et al., 2022). However, MSE currently lack standardized methods for the selection of operating models, which makes it difficult to provide an open,

1.	INTRODUCTION	1
2.	MODELLING FRAMEWORK	2
3.	GRAND BANKS APPLICATION	3
3.1.	Overview	3
3.2.	Materials and methods	5
3.2.1.	Core population dynamics models	5
3.2.2.	Population processes	5
3.2.3.	Environmental processes	5
3.2.4.	Multispecies processes	6
3.2.5.	Model performance and estimation	6
3.2.6.	Recovery simulations	6
3.3.	Results	7
3.3.1.	Population processes	7
3.3.2.	Environmental processes	8
3.3.3.	Multispecies processes	10
3.3.4.	Best-performing models	10
3.3.5.	Recovery simulations	10
4.	DISCUSSION	10
4.1.	Grand banks application	11
4.2.	Providing ecosystem-informed fisheries management advice	13
	ACKNOWLEDGMENTS	14
	DATA AVAILABILITY STATEMENT	14
	REFERENCES	14

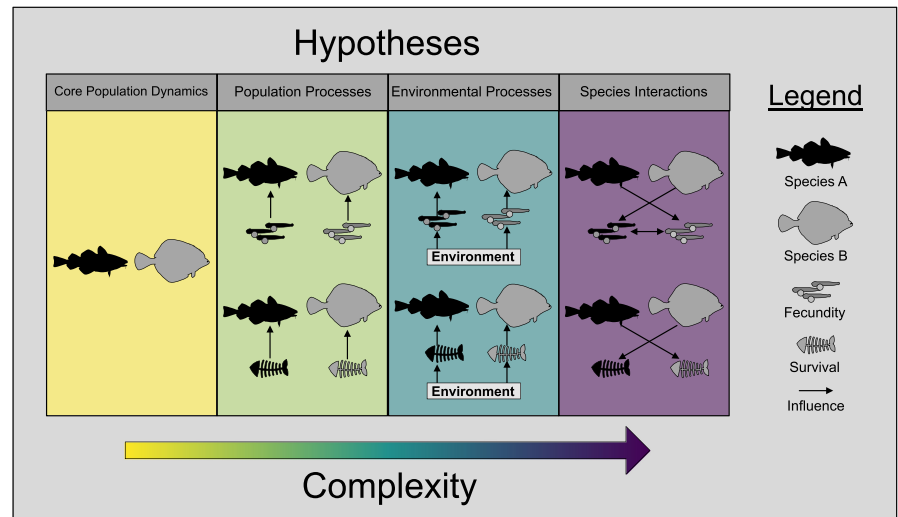
transparent process to the managers, scientists and stakeholders that are involved (Townsend et al., 2019).

Without a decision-making framework for model development and selection, plausible model options may go untested and the value of including the processes represented in them would remain unknown. Here, we develop a conceptual framework for extending and comparing population dynamics models of increasing complexity focused on explaining changes in productivity. This framework provides a structured approach for testing population and ecosystem hypotheses to identify whether proposed hypotheses of non-stationary productivity have data-based support within a population dynamics model. We will first provide an overview to illustrate the proposed framework and then provide an example application to two flatfish populations on the Newfoundland Grand Banks.

## 2 | MODELLING FRAMEWORK

Our proposed conceptual framework for testing and comparing models of increasing complexity involves systematically testing hypotheses about population and ecosystem processes that may modify population productivity (Figure 1). Once hypotheses for shifts in population productivity are identified, the first step of the modelling

**FIGURE 1** Conceptual diagram of the framework for extending and comparing population dynamics models of increasing complexity.



framework involves selecting a core population dynamics model. Ideally, core population dynamics models will match the model used for the current stock assessment, as it should be a rigorous and scientifically defensible description of the dynamics. However, multiple plausible population dynamics models may exist for a stock and different models may allow improved integration with available data or estimation of relevant population processes. Although most applications will involve adding complexity to one population dynamics model, adding complexity to multiple population dynamics models would allow further comparisons, which may better elucidate the true population and ecosystem mechanisms underlying time-varying productivity (Jardim et al., 2021). Regardless of the specific choice in core population dynamics model(s), the goal of the core model is to ensure that it estimates similar trends as those estimated by the current assessment model so that comparisons with current practice are meaningful. Fishery stakeholders and managers familiar with current practice may be reluctant to buy-in if the core modelling approach produces drastically different estimates from previously accepted models despite being driven by similar assumptions (Fulton, 2021; Townsend et al., 2019). Furthermore, adding complexity to a model that initially produces similar estimates to the current assessment model can portray limitations of current assumptions.

Once a core population dynamics model(s) is/are chosen, complexity is added to explore hypotheses regarding the mechanisms driving time-varying productivity. To accomplish this, models are modified to allow variability in population processes (e.g., natural mortality, age-at-maturity) that had previously been assumed known or stationary (Figure 1). Allowing population processes to vary (often using random effects and temporal correlation structures) creates an implicit link between population dynamics and the ecosystem (Thorson & Minto, 2015). The choice of which processes vary will depend on the hypothesized mechanisms driving time-varying productivity. By using a structured process, the selection of hypothesized mechanisms needs to be explicit, providing an opportunity to consider whether alternative hypotheses are missing. Furthermore, how the processes are modelled will depend on data availability and the processes themselves. For example, natural mortality rates

can be estimated by integrating tagging data, by separating natural mortality into predation and residual mortality components or via estimation as annual process deviations in a state-space model (reviewed in Maunder et al., 2023). The value of allowing a population process to vary over time will be assessed by examining model performance criteria (Section 3.2.5). If allowing a population process to vary improves model performance, then that model is used in subsequent analyses because that population process may have been influenced by environmental or multispecies processes. Each step of this modelling process is based on a combination of statistical model fit and ecological realism, such that resulting models are both interpretable and meaningful (Collie et al., 2016).

Environmental and multispecies processes can be included in models to explain population dynamics (Figure 1). Environmental processes should only be included if they maintain a clear mechanistic rationale for how they have influenced population dynamics (Collie et al., 2016; Plagányi et al., 2014). Environmental and multispecies processes can be included as linear covariates or modelled using functional forms, depending on ecological knowledge and data availability for the process of interest (Collie et al., 2016; Plagányi et al., 2022). Regardless of what process is included and how, model performance will be assessed to identify whether they helped account for previously unexplained variability.

### 3 | GRAND BANK APPLICATION

#### 3.1 | Overview

To illustrate the modelling framework, we applied it to assess the population and ecosystem processes that affected the differential recovery of two flatfish populations on the Newfoundland Grand Banks (Northwest Atlantic Fisheries Organization [NAFO] Divisions 3LNO) over the past three decades. In response to prolonged intense fishing pressure and a period of anomalous environmental conditions (e.g., record low temperatures; Cyr & Galbraith, 2021), fish community biomass on the Newfoundland Grand Banks

collapsed in the early 1990s (Dempsey et al., 2017). Since this collapse, yellowtail flounder (*Limanda ferruginea*, Pleuronectidae) has recovered while American plaice (*Hippoglossoides platessoides*, Pleuronectidae) has remained at a low population level. While the magnitude of collapse for plaice was larger (reduced to 5% of biomass from mid-1980s) than for yellowtail (reduced to 50% of biomass from mid-1980s), it remains unknown whether this difference in magnitude is the sole reason why yellowtail recovered so rapidly (Brodie et al., 2010) while plaice did not. Plaice continue to exhibit limited recovery despite having relatively similar life-history characteristics to yellowtail (e.g., slow growth, late maturation, long lifespan), having no directed fishing pressure (Parsons et al., 2021; Wheeland et al., 2021) and several relatively strong recruitment events (Morgan & Brodie, 2001; Perreault et al., 2020; Walsh et al., 2004).

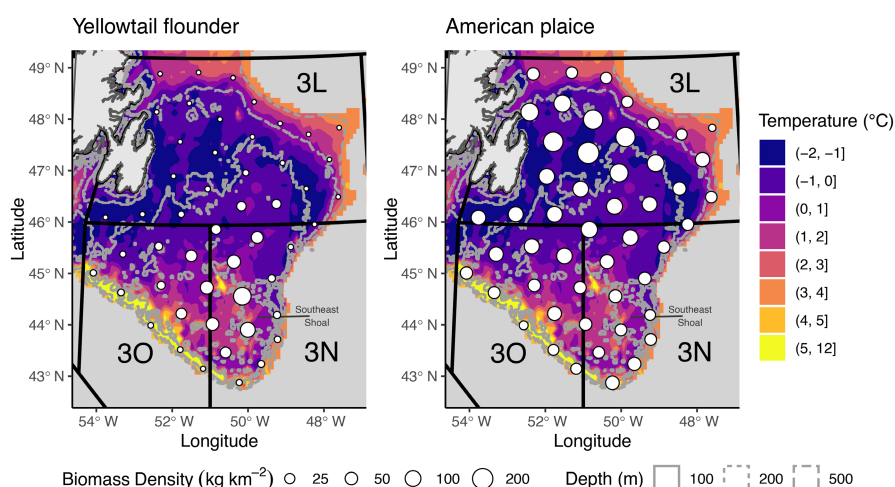
Changing environmental conditions in the 1990s likely contributed to groundfish population collapse by affecting recruitment and natural mortality rates ( $M$ ), although fishing pressure and climate likely also modified correlated life-history traits (e.g., maturity, growth and condition; Brodie et al., 2010; Halliday & Pinhorn, 2009; Parsons & Lear, 2001; Walsh et al., 2004). The primary hypothesis for the rapid recovery of yellowtail is that beneficial ecosystem conditions for juvenile survival permitted rapid population growth, despite relatively low spawning stock biomass (Brodie et al., 2010). In contrast, the dominant hypothesis for the lack of recovery of plaice stems from ecosystem-induced changes to natural mortality rates (Morgan et al., 2011; Perreault et al., 2020; Wheeland, 2021). These population processes may have been modified via several ecosystem processes, namely prey availability, climatic variability and competition.

Lack of prey can induce starvation mortality, as observed for Northern cod (Regular et al., 2022) and hypothesized for other local stocks (Cadigan et al., 2022), and could have affected the collapse and

recovery trajectories of plaice and yellowtail. These flounder species consume similar benthic (e.g., amphipods and crustaceans) and forage fish prey (e.g., sand lance [*Ammodytes dubius*, Ammodytidae] and capelin [*Mallotus villosus*, Osmeridae]) on the Grand Banks, although forage fish provide a larger contribution to the diet of plaice (Bruno et al., 2000; Gonzalez et al., 2006). The capelin population collapsed in the 1990s (DFO, 2019), and although pre-1990s estimates of sand lance population size do not exist, recent estimates indicate that their population has followed similar fluctuations to capelin in recent years (Robertson, Koen-Alonso, et al., 2022), indicating a potentially limited forage fish prey base.

Temperatures below physiological thresholds can induce lethal effects (Donaldson et al., 2008) and can limit foraging via physiological impacts (e.g., feeding cessation; Morgan, 1992) and reductions in spatial overlap between predators and prey (Hunsicker et al., 2013). Plaice and yellowtail historically had distinct spatial distributions, where yellowtail was distributed in shallow, warm waters in NAFO divisions 3NO and plaice was distributed throughout the Grand Banks, with the largest centres of biomass in deeper, colder waters to the north in NAFO division 3L (Figure 2; Robertson et al., 2021). Exposure to cold temperatures shifted the spatial distribution of plaice southwards (~200 km change in centre of gravity; Robertson et al., 2021) and yielded a greater spatial overlap with yellowtail and thorny skate (*Amblyraja radiata*, Rajidae). Thorny skate consume similar prey to yellowtail and plaice (Gonzalez et al., 2006) and, like yellowtail, are one of the few groundfish populations in the region that recovered following their collapse in the early 1990s (Sosebee et al., 2022). This shift in the distribution of plaice, coupled with similarities in diet and a limited availability of forage fish prey, may have exacerbated competition for habitat or food with yellowtail or thorny skate.

Given the dominant ecological hypotheses regarding the mechanisms responsible for the collapse and differential recovery of plaice



**FIGURE 2** Mean historical distribution (1985–1989) of NAFO Divisions 3LNO yellowtail flounder and American plaice overlaid on spring bottom-water temperatures from the same time-period. Circle size represents biomass density ( $\text{kg km}^{-2}$ ) estimates for each stock estimated using VAST models in Robertson et al. (2021). Spring bottom-water temperatures were derived from all available temperature profiles collected annually (e.g., DFO monitoring programmes, international oceanographic campaigns, Argo programme) between April and June on a regular  $0.1^\circ \times 0.1^\circ$  (latitudinal  $\times$  longitudinal) grid using a linear interpolation method introduced in Cyr et al. (2021).

and yellowtail populations on the Grand Banks, we apply our framework of extending and comparing population dynamics models to examine the effects of time-varying recruitment,  $M$ , bottom-water temperature, prey availability and competition. In doing so, we attempt to gain insight into the most likely population and ecosystem processes that affected the differential trajectories of population collapse and recovery.

## 3.2 | Materials and methods

### 3.2.1 | Core population dynamics models

The 2021 3LNO yellowtail flounder stock assessment model was a Bayesian stock production model, which has been used due to a lack of age-structured time series (Parsons et al., 2021). However, to facilitate more direct comparisons between the yellowtail and plaice models, we opted to model yellowtail population dynamics using a state-space parameterization of a Deriso-Schnute delay-difference model (Deriso, 1980; Schnute, 1987). This model estimates unstructured population dynamics but explicitly estimates recruitment and survival parameters. Our model estimated population dynamics for yellowtail from 1965 to 2017 and included annual catch data (1965–2017) and biomass estimates from five bottom-trawl research surveys (1971–2017; Parsons et al., 2021). A full model description can be found in Appendix S1.

The 2021 3LNO American plaice stock assessment model was conducted using an adaptive framework-virtual population analysis (ADAPT-VPA; Wheeland et al., 2021). However, this type of assessment cannot include uncertainty in the landings data, a key concern for this stock where there is evidence of substantial underreporting in historical landings (Dwyer et al., 2016). As a result, a state-space age-based model was recently developed for this stock (Perreault et al., 2020). We used a modified version of this model to estimate plaice dynamics. This model estimates population dynamics from 1960 to 2017 and includes data from research bottom-trawl surveys (1985–2017) and commercial landings (1960–2017). A full model description and rationale for modifications from the prior version of this model can be found in Appendix S2.

### 3.2.2 | Population processes

To identify whether recruitment and  $M$  varied over time (and/or with age), we compared two major model parameterizations for yellowtail and three for plaice. The parameterizations for yellowtail had (1) only time-varying recruitment deviations or (2) time-varying recruitment deviations and natural mortality deviations. Meanwhile, the plaice parameterizations included (1) only time-varying recruitment deviations, (2) time-varying recruitment deviations and natural mortality deviations or (3) time-varying recruitment deviations and time- and age-varying natural mortality deviations. We included parameterizations with both time-varying and time- and age-varying  $M$  because

limited data quality or resolution can prohibit improvements in model performance when  $M$  varies by age (i.e., overfitting, large parameter uncertainty) despite  $M$  being known to vary with body size and age (Lorenzen, 2022). We modelled recruitment and  $M$  deviations following a common approach (Maunder & Watters, 2003; O'Leary et al., 2019) and in a way that would permit covariate inclusion (Section 3.2.3). We modelled these time (or time- and age-) varying processes ( $X_y$ ) as,

$$X_y = f(x) \exp(\varepsilon_y), \quad (1)$$

$$\varepsilon_y \sim RW(\sigma), \quad (2)$$

where  $f(x)$  is the mean function of the individual process,  $\varepsilon_y$  are temporal deviations of that process,  $RW$  refers to a zero-mean Gaussian random walk, and  $\sigma$  is the standard deviation of the time-varying deviations. The mean function for  $M$  deviations was input as 0.2 to provide a baseline value and to allow comparability between models. For plaice parameterizations, we only considered models with  $M$  deviations for ages 5+ since a recent study identified that the influence of  $M$  deviations on juvenile plaice dynamics (ages 1–5 years) has been relatively small since the early 1990s (Robertson, Regular, & Cadigan, 2022). Additionally, to penalize  $M$  deviations (Punt et al., 2021), we included a zero-mean Gaussian prior with a standard deviation of 0.05 for  $\sigma$  in Equation (2). Meanwhile, the mean recruitment function for yellowtail was based on the number of mature fish in the population and a maximum juvenile survival rate, where the deviations essentially reduced survival from its maximum (see Appendix S1 for full description and equations). Finally, plaice mean recruitment was based on a Beverton–Holt stock-recruitment function (see Appendix S2 for full description and equations).

### 3.2.3 | Environmental processes

We examined parameterizations where environmental time series were input to explain recruitment or  $M$  deviations. To achieve this, we used the general formulation for integrating environmental time series into population dynamics models that was developed by Maunder and Watters (2003),

$$X_y = f(x) \exp(\varphi_0 + \varphi_1 I_y + \varepsilon_y), \quad (3)$$

where  $\varphi_0$  is a scaling parameter,  $I_y$  is the environmental time series, and  $\varphi_1$  relates the environmental time series to the process of interest. Data exploration indicated that relationships between environmental time series and process deviations were linear (see Appendix S3). By modelling the environmental drivers in this way, we were able to determine whether the population dynamics of a species are better described when environmental drivers are explicitly used to reduce the process deviations ( $\varepsilon_y$ ). Furthermore, by estimating autocorrelated  $\varepsilon_y$ , even with the inclusion of environmental drivers (Equation 2), this modelling approach does not rely on the relationship between population processes and environmental drivers to fully describe non-stationary processes, but instead allows these drivers to act as a component of the non-stationarity.



We considered including the role of (1) climatic variability on habitat availability and predator–prey overlap using the Newfoundland and Labrador (NL) climate index (NLCI; Cyr & Galbraith, 2021), mean normalized anomalies of the spring bottom-water temperature derived from multiple data sources (Cyr et al., 2021), and the mean normalized anomalies of the summer cold-intermediate layer (CIL) area over hydrographic sections on the NL shelf (Cyr & Galbraith, 2021), (2) prey availability using a time series of capelin biomass (Koen-Alonso et al., 2021) and northern sand lance abundance (Robertson, Koen-Alonso, et al., 2022) and (3) a potential competitor for food and habitat, thorny skate using a survey time series of estimated biomass for this species (Simpson et al., 2018).

To account for longer-term climatological effects and the period of influence on processes, climate time series were included differently for recruitment and  $M$  for both species. Yellowtail recruitment was tested with one-sided 5-year moving averages (i.e., averaging the 5 years prior to a given year) because recruitment was estimated to occur over a 5-year period (see Appendix S1). Plaice recruitment was tested against one-sided 3-year moving averages to account for limited catchability of age 1–2 plaice by the surveys (Morgan et al., 1998). Meanwhile,  $M$  for both species was tested against climate time series without a moving average, because  $M$  was most likely affected by the direct impacts of climate within a given year. The northern sand lance index was extended back to 1984 by combining estimates from separate nonlinear functional response models (Robertson, Koen-Alonso, et al., 2022) for Engels and Campelen research survey data. Every covariate was standardized using the standard score equation ( $X_{\text{standard}} = \frac{X - \bar{X}}{\sigma_x}$ ) prior to inclusion to improve model convergence and to determine whether longer time series could serve as proxies for correlated shorter time series (see Appendix S3). As a result, the included environmental covariates were reduced from six to five. The capelin time series were relatively short in comparison to other covariates and were highly negatively correlated with the NLCI, CIL area and bottom-water temperature so they were not directly tested.

### 3.2.4 | Multispecies processes

To determine whether any of the observed recruitment or  $M$  deviations were driven by competitive interactions with the other flatfish population, we compared several parameterizations of a multispecies model. This model included both population dynamics models in a joint likelihood where stock size or recruitment estimates of one species were used like an environmental time series (i.e.,  $I_y$ ) to affect recruitment or  $M$  process deviations of the other species using Equation (3). Specifically, the spawning stock biomass of each population affected either the recruitment or  $M$  deviations of the other stock to test for the possibility of competition and predation of juveniles by mature fish or direct competition for habitat or food between mature fish (e.g.,  $M_y^{\text{amp}} = 0.2 \exp(\varphi_0 + \varphi_1 \text{SSB}_y^{\text{yfl}} + \epsilon_y)$ ). Recruitment estimates were also used to explain the recruitment deviations of the other stock to test whether there may be juvenile habitat competition. Because

yellowtail recruitment deviations modelled spawning that occurred 5 years prior (see Appendix S1), the spawning stock biomass and recruitment of plaice were lagged 5 years. Parameterizations examined included one-way (i.e., effect of one stock on the other) and two-way (i.e., both stocks affected one another) interactions. Given that both population dynamics were included in a joint likelihood, stock size or recruitment estimates each year could be modified based on their effects on process deviations or based on the other stock's estimates if a two-way interaction was estimated.

### 3.2.5 | Model performance and estimation

Model performance was assessed using a combination of Akaike information criterion (AIC), Bayesian information criterion (BIC), examination of residuals and parameter uncertainty. AIC is a measure of prediction accuracy, while BIC measures goodness-of-fit (Sober, 2002). Meanwhile, parameter uncertainty can identify how the inclusion of a process affects model uncertainty (e.g., reduces survey index error), where a greater reduction in parameter uncertainty would indicate better performance. We first compared core population dynamics models with models that allowed population processes to vary to determine which formulation had the best performance (i.e., lowest AIC, BIC, residuals and parameter uncertainty). The best-performing model was then extended to include environmental and multispecies processes, and the performance of covariates was assessed in the same way as performance of models with added population processes. For simplicity, we selected a single best-performing model at each step, including the identification of a single best-performing covariate; however, this approach could include the identification of multiple models and/or allow multiple covariates in a single model. No covariates except the NLCI and CIL area existed for the entire time-period of the population dynamics models. To account for this, we conducted model selection with covariates only affecting process deviations from 1984 to 2017, when data existed for every covariate, to ensure that differing data lengths did not influence results (i.e., to ensure that AIC/BIC scores were meaningful). In consequence, covariate data from years prior to 1984 were excluded from model fitting. We used the Template Model Builder (TMB; Kristensen et al., 2016) package in R (R Core Team, 2018) to evaluate the negative logarithms of the marginal likelihoods (nll) of these models and the data, and to evaluate the nll gradients to improve estimation. Further, we used the R function `nlnmb()` to find the maximum likelihood estimates of model parameters.

### 3.2.6 | Recovery simulations

If models that included environmental or multispecies processes outperformed the best-performing population process models, we ran simulations to determine how recovery trajectories would have varied under alternative histories of the most important (i.e., best performing) environmental or multispecies process. These simulations gave an indication of the magnitude of difference in population

dynamics under different ecosystem conditions for our case study. Simulations began 2 years prior (1992) to fishing moratoria (1994) after the period of collapse had been officially recognized. All model parameters (including fishing mortality) were fixed at values from prior model runs so that variability in population dynamics was only driven by differences in ecosystem trajectories. The alternative histories involved simulating covariate time series ( $E_y$ ) based on a Gaussian random walk ( $e_y$ ) that was shifted with a constant ( $c$ ), such that,

$$e_y \sim \begin{cases} N(0, \sigma^2) & \text{if } y=0 \\ N(e_{y-1}, \sigma^2) & \text{if } y>0 \end{cases} \quad (4)$$

$$E_y = e_y + c. \quad (5)$$

Since all environmental covariates were standardized prior to inclusion in models,  $c$  was either set to  $-1$ ,  $0$  or  $1$  to represent low, average and high values for each covariate. Furthermore, the random walk standard deviation ( $\sigma$ ) was set to  $0.2$  to allow variability over time while minimizing overlap between alternative histories. Population dynamics were then simulated 1000 times, and the 20:80% and 10:90% quantiles of the simulations were visualized to identify the most common range of dynamics under each alternative history. Although it is likely that variability in recruitment and/or natural mortality would yield additional consequences for the population dynamics (e.g., modifying fishing history), the main goal of this simulation exercise was to illustrate the magnitude of the effect that environmental or multispecies processes may have had on the recovery of these stocks. This illustration may be useful because model comparisons between models that include estimated annual process deviations and covariates can produce very similar SSB,  $F$  and recruitment estimates (e.g., Correa et al., 2023).

### 3.3 | Results

#### 3.3.1 | Population processes

The best-performing models for both yellowtail and plaice included time-varying  $M$  deviations (YTFL2; AMPL2; Table 1). For

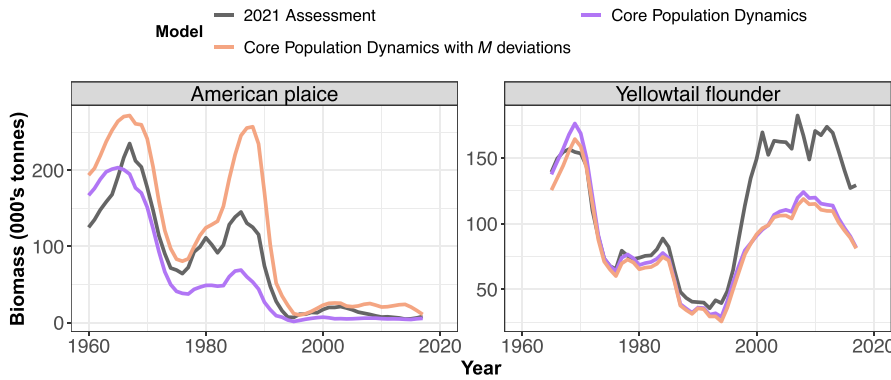
the yellowtail model, improvements to model criterion score and residuals were minor (e.g., mean reduction of survey index standard deviation estimates  $<0.01$ ) and the estimates of time-varying  $M$  deviations were negatively correlated with the recruitment deviations. As a result, further analyses for yellowtail were conducted with the model that did not estimate  $M$  deviations (YTFL1). For the plaice model, including time-varying  $M$  deviations improved the fit to the survey indices, catch proportions-at-age and landings data (see Appendix S2 for residual plots). Furthermore, including  $M$  deviations reduced all standard deviation estimates for catch proportions-at-age (reduction mean =  $0.38$ ) and reduced standard deviation estimates for survey indices for all (mean reduction =  $0.07$ ) but the oldest ages (age 14–15 years) which were approximately equal to estimates from the model without time-varying  $M$  (difference  $<0.01$ ). However, extending this further to include time- and age-varying  $M$  deviations had worse model performance than the model with only time-variations in  $M$ . This would indicate that although the data were informative about time-variations in  $M$ , there was not enough information to parse both time and age variation.

The best-performing models followed similar trends in spawning stock biomass (SSB) and biomass to the most recent stock assessments (Figure 3). The yellowtail models estimated a very similar trend to the 2021 stock assessment until the mid-1990s. After this point, both models estimated rapid increases followed by a plateau and a decline; however, both models estimated the magnitude of the increase to be smaller than the 2021 stock assessment (peak of 124 kt rather than a peak of 176 kt). This difference in estimated biomass appears to be driven by differing survey catchability estimates. The core plaice model (AMPL1) did not estimate as large of a peak in SSB in the late-1960s or in the mid-1980s as the 2021 stock assessment. Meanwhile, the model with time-varying  $M$  deviations (AMPL2) estimated a slightly larger peak in SSB in the late-1960s and a much larger peak in SSB in the mid-1980s. These differences in SSB are most likely driven by the increasing estimates of  $M$  in the 1980s–1990s. AMPL1 assumes that  $M$  is stationary at  $0.2$ , the current stock assessment inputs  $M$  as  $0.53$  from 1989 to 1996 for all ages (Wheeland et al., 2021), and AMPL2 estimates that  $M$  for fish ages 5+ is higher (mean =  $0.84$ ) during that period. The models account for this increase in  $M$  by modifying SSB to maintain landings

TABLE 1 Model comparisons for different formulations of both population dynamics models.

Name	$M$ dev.	nll	k	AIC	BIC	$\Delta$ AIC	$\Delta$ BIC
Yellowtail flounder							
YTFL1	Input	96	14	221	258	6	3
YTFL2	<b>Time-Varying</b>	<b>93</b>	<b>15</b>	<b>215</b>	<b>255</b>	<b>0</b>	<b>0</b>
American plaice							
AMPL1	Input	1126	55	2362	2622	335	331
AMPL2	<b>Time-Varying</b>	<b>958</b>	<b>56</b>	<b>2027</b>	<b>2291</b>	<b>0</b>	<b>0</b>
AMPL3	Time & Age-Varying	978	66	2087	2399	60	108

Note: nll represents the negative log-likelihood; k represents the number of parameters;  $\Delta$  represents the difference in model criterion score from the model with the lowest model criterion score. The bolded row indicates the model with the lowest AIC and BIC.



**FIGURE 3** Model comparisons between spawning stock biomass (American plaice) or biomass (yellowtail flounder) from the 2021 stock assessment (grey lines) to the core population dynamics model (purple lines; YTFL1 and AMPL1 from Table 1), and the core population dynamics model with M deviations (orange lines; AMPL2 and YTFL2 from Table 1).

**TABLE 2** Model comparisons between the two best-performing population process only models and models with environmental processes.

Name	Process	Covariate	nll	k	AIC	BIC	ΔAIC	ΔBIC
Yellowtail flounder								
YTFL1			96	14	221	258	21	16
<b>YTFLE1</b>	<b>Rec.</b>	<b>Climate index</b>	<b>84</b>	<b>16</b>	<b>200</b>	<b>242</b>	<b>0</b>	<b>0</b>
YTFLE2	Rec.	CIL	94	16	220	262	20	20
YTFLE3	Rec.	Bot. temp	96	16	225	267	25	25
YTFLE4	Rec.	Sand lance	96	16	224	266	24	24
YTFLE5	Rec.	Thorny Skate	95	16	222	264	22	22
American plaice								
AMPL2			958	56	2027	2291	20	11
AMPLE1	Rec.	Climate index	957	58	2029	2303	22	23
AMPLE2	Rec.	CIL	957	58	2031	2304	24	24
AMPLE3	Rec.	Bot. temp.	950	58	2016	2289	9	9
AMPLE4	Rec.	Sand lance	957	58	2031	2304	24	24
AMPLE5	Rec.	Thorny Skate	957	58	2031	2304	24	24
<b>AMPLE6</b>	<b>M</b>	<b>Climate index</b>	<b>945</b>	<b>58</b>	<b>2007</b>	<b>2280</b>	<b>0</b>	<b>0</b>
AMPLE7	M	CIL	953	58	2021	2295	14	15
AMPLE6	M	Bot. temp	954	58	2025	2298	18	18
AMPLE9	M	Sand lance	957	58	2031	2304	24	24
AMPLE10	M	Thorny skate	957	58	2030	2304	23	24

Note: nll represents the negative log-likelihood; k represents the number of parameters; Δ represents the difference in model criterion score from the model with the lowest model criterion score. The bolded row indicates the model with the lowest AIC and BIC.

estimates. Finally, AMPL2 estimated a higher, fluctuating SSB that declined rather than increased in the terminal years of the model (2016–2017).

### 3.3.2 | Environmental processes

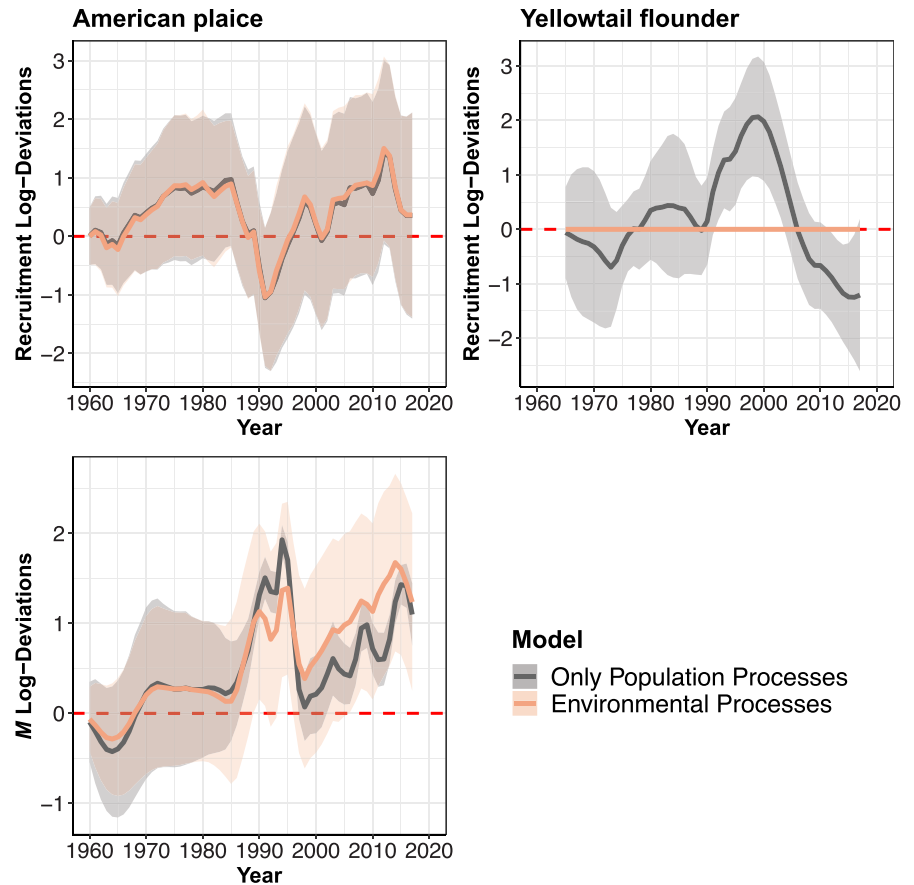
Both yellowtail and plaice models were improved by using environmental covariates to account for either recruitment or M deviations (Table 2; Appendices S1 and S2). The best-performing yellowtail model included the 5-year moving average NLCI to account for recruitment deviations (YTFLE1). Including this climate index completely accounted for recruitment deviations (recruitment deviation standard deviation in YTFL1 = 0.35, YTFLE1 < 0.001; Figure 4)

but also reduced model fit to survey indices (mean standard deviation estimate increase = 0.02; Figures S1.2, S1.3, S1.6 and S1.7 in Appendix S1). Furthermore, the landing residuals developed a slight pattern from the 1970s to early 1990s when the NLCI was used to modify recruitment deviations (Figures S1.5 and S1.9 in Appendix S1), although the magnitude of these trends is small (see Figures S1.4 and S1.8 in Appendix S1). Even though the NLCI only modified the deviations from 1984 onwards, the time series accounted for enough variability that recruitment deviations did not need to be estimated to yield similar population dynamics estimates to YTFL1.

The best-performing model for plaice included the NLCI to account for M deviations (AMPLE6; Table 2). Including the NLCI reduced the standard deviation estimate for the M deviations by 0.05



**FIGURE 4** Comparison of deviation parameters (recruitment and  $M$ ) between the best-performing population process models (orange lines; AMPL2, YTF1) and the best-performing environmental process models (grey lines; AMPLE6, YTFLE1). 95% confidence interval for estimates is shown as polygons.



**TABLE 3** Model comparisons for the multispecies formulations.

Name	Ampl rec.	Ampl $M$	Ytfl rec.	nll	k	AIC	BIC	$\Delta$ AIC	$\Delta$ BIC
<b>MS1</b>				<b>1035</b>	<b>70</b>	<b>2211</b>	<b>2549</b>	<b>0</b>	<b>0</b>
MS2	SSB			1035	72	2214	2562	3	13
MS3	Rec.			1035	72	2213	2561	2	12
MS4			SSB	1034	72	2213	2560	2	11
MS5			Rec.	1039	72	2223	2570	12	21
MS6		SSB		1227	72	2598	2946	387	397
MS7	SSB		SSB	1034	74	2216	2574	5	25
MS8	Rec.		Rec.	1033	74	2215	2572	4	23
MS9		SSB	SSB	1226	74	2600	2957	389	408
MS10		SSB	Rec.	1228	74	2605	2962	394	413

*Note:* The column headers represent the process (Rec. or  $M$ ) for a given species that was estimated to have a relationship with a covariate. Text within the Ampl. Rec, Ampl  $M$  and Ytfl Rec. columns indicates the covariate modelled to explain a particular process, where SSB refers to the other species spawning stock biomass while Rec. refers to the other species recruitment. nll represents the negative log-likelihood; k represents the number of parameters;  $\Delta$  represents the difference in model criterion score from the model with the lowest model criterion score. The bolded row indicates the model with the lowest AIC and BIC.

Abbreviations: Ampl, American plaice; Ytfl, yellowtail flounder.

and did not substantially affect recruitment deviations (Figure 4) or model fit to catch proportions-at-age or survey indices (standard deviation estimate differences  $<0.01$ ; Figures S2.11–S2.15, S2.18–S2.20, S2.21–S2.25, S2.28–S2.30, and S2.31–S2.35 in Appendix S2).  $M$  deviations in AMPLE6 were modified throughout the time series. The biggest change from the non-extended model (AMPL2) occurred from 1989 onwards, where the deviations were reduced in

the late 1980s – early 1990s as was the oscillating trend during the late 1990s – 2017. However, all changes remained within the 95% confidence intervals of both models. Finally, although none of the models that used covariates to account for recruitment deviations produced a better performance than AMPLE6, the model with 3-year moving average bottom-water temperatures (AMPLE3) produced the best performance of the recruitment deviation models.

### 3.3.3 | Multispecies processes

Multispecies formulations with competitive interactions did not produce improved model performance over the null multispecies model that did not include competitive interactions (Table 3). This lack of improvement indicates limited evidence that the dynamics of these two stocks are primarily driven by direct competitive interactions with one another.

### 3.3.4 | Best-performing models

Since there was no improvement in model performance with multispecies model formulations, the best-performing models for both species were those that included environmental processes (YTFLE1 and AMPLE6). The NLCI was estimated to have a negative effect on  $M$  and recruitment deviations for plaice and yellowtail, respectively (Figure 5). For yellowtail, a negative NLCI (i.e., colder conditions) promoted strong recruitment while a positive NLCI (i.e., warmer conditions) substantially reduced recruitment success. For plaice, the negative relationship between  $M$  and the NLCI indicated that  $M$  increased during colder conditions and decreased during warmer conditions.

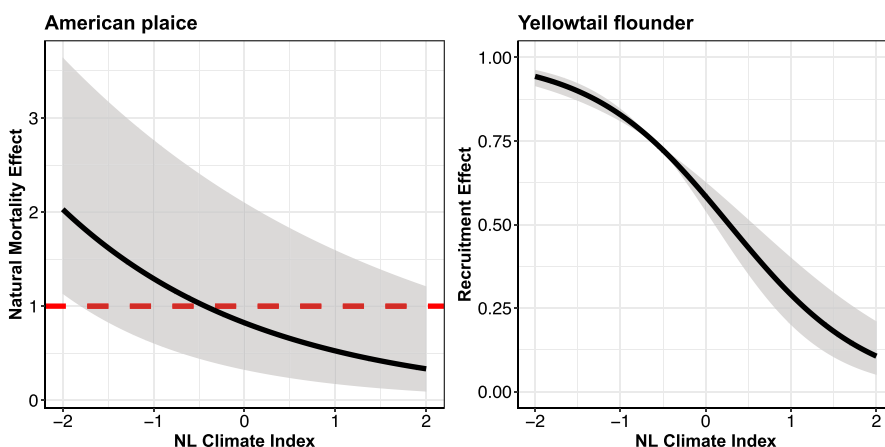
The best-performing models did not substantially modify population dynamics estimates from the best population process formulations (Figure 6). SSB, fishing mortality ( $F$ ) and recruitment estimates were only slightly modified for plaice between the model with only population processes (AMPL2) and the model that included environmental processes (AMPLE6), with the only observable differences occurring prior to 1990 and no estimates falling outside of the 95% confidence interval of the extended model. Meanwhile, yellowtail biomass,  $F$  and recruitment estimates were modified throughout the time series between the model with only population processes (YTFL1) and the best-performing model (YTFLE1). In general, the differences in estimates between models fell within the 95% confidence intervals of YTFLE1 in recent years. These changes were likely driven by the lack of estimated recruitment deviations.

### 3.3.5 | Recovery simulations

Simulations for AMPLE6 and YTFLE1 indicated that recovery trajectories for these stocks would have differed substantially with alternative histories for their most important environmental drivers (Figure 7). If environmental conditions had followed a warmer than average trajectory, yellowtail would have continued to decline after collapse, with most simulations indicating lower than current biomass. Average conditions could have yielded a wide range of population trajectories, with population levels below, equal to or higher than currently observed. Finally, simulations for cold conditions indicated that yellowtail biomass would most likely have been higher than current levels. It is worth noting that the magnitude of this increase is likely an artefact of the lack of carrying capacity term in the delay-difference formulation (Table S1 in Appendix S1). Meanwhile, the plaice simulation indicates that given warmer than average conditions, the plaice stock would not have collapsed and would have experienced an uncertain population trajectory from the early 2000s onwards, with most simulations yielding larger SSB than is currently estimated. However, with both average and cold conditions, plaice SSB would have declined and remained low.

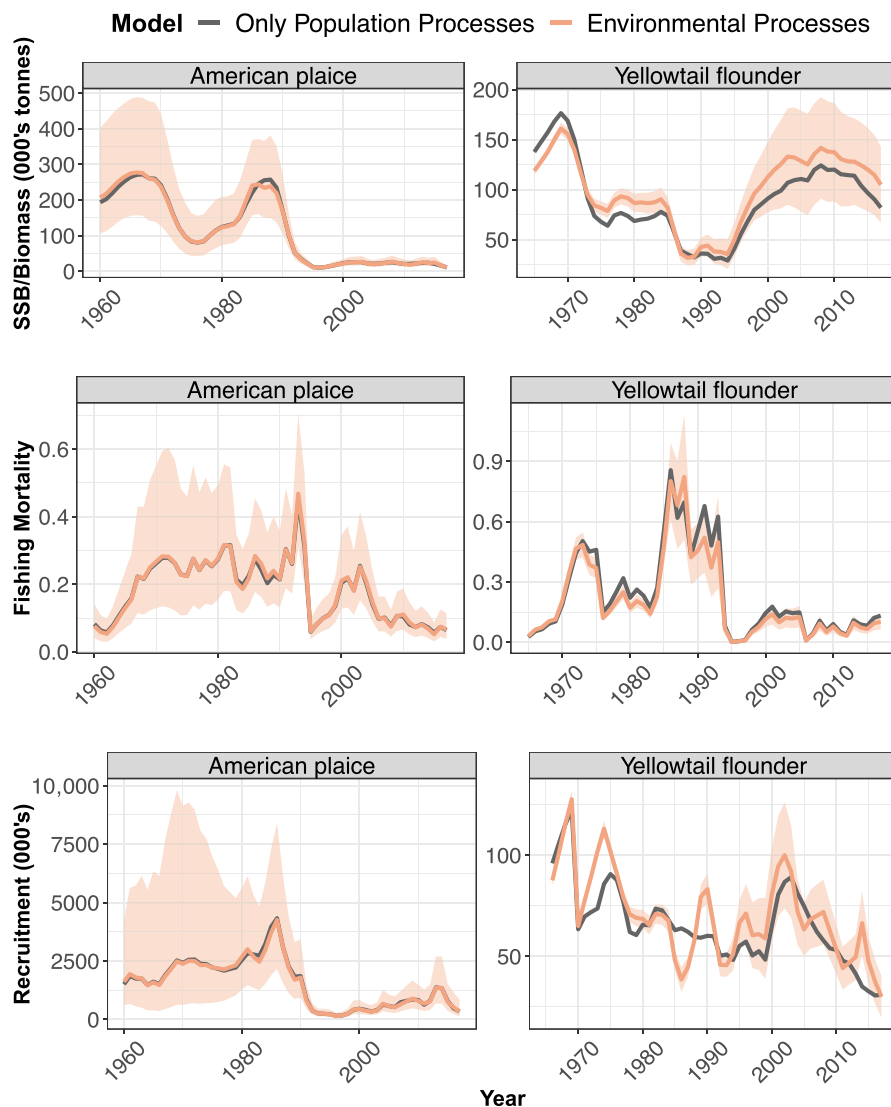
## 4 | DISCUSSION

Testing models of increasing complexity can help identify the most important processes influencing population dynamics. Single-species population dynamics models are generally limited to attributing changes in populations to shifts in recruitment or fishing mortality; however, there are other processes (e.g., growth, natural mortality) that may be more important for describing fish population dynamics (Rice, 2011). The framework proposed here allows modelers to explore hypotheses related to shifts in other population vital rates and interactions with the environment or other species in a structured and ordered fashion, permitting a sequential exploration of multiple hypotheses. Further, by exploring hypotheses via their impact on population dynamics, this modelling framework explicitly focuses on understanding drivers in relation to their consequences



**FIGURE 5** Estimates of the effects (black lines) and 95% confidence intervals (grey polygons) of environmental covariates for AMPLE6 and YTFLE1. The red dashed line in the American plaice panel represents 1, which would indicate no effect on  $M$  since the effect is multiplicative.

**FIGURE 6** Comparison of SSB/ biomass, fishing mortality and recruitment estimates between the best-performing population process models (grey lines; AMPL2, YTFL1) and the best-performing models with environmental processes (orange lines; AMPL6, YTFLE1). 95% confidence interval for estimates from best-performing extended models is shown as orange polygons. Estimates of SSB are shown for American plaice while yellowtail flounder estimates are for biomass. Fishing mortality estimates for American plaice are averaged from ages 9 to 14 years.



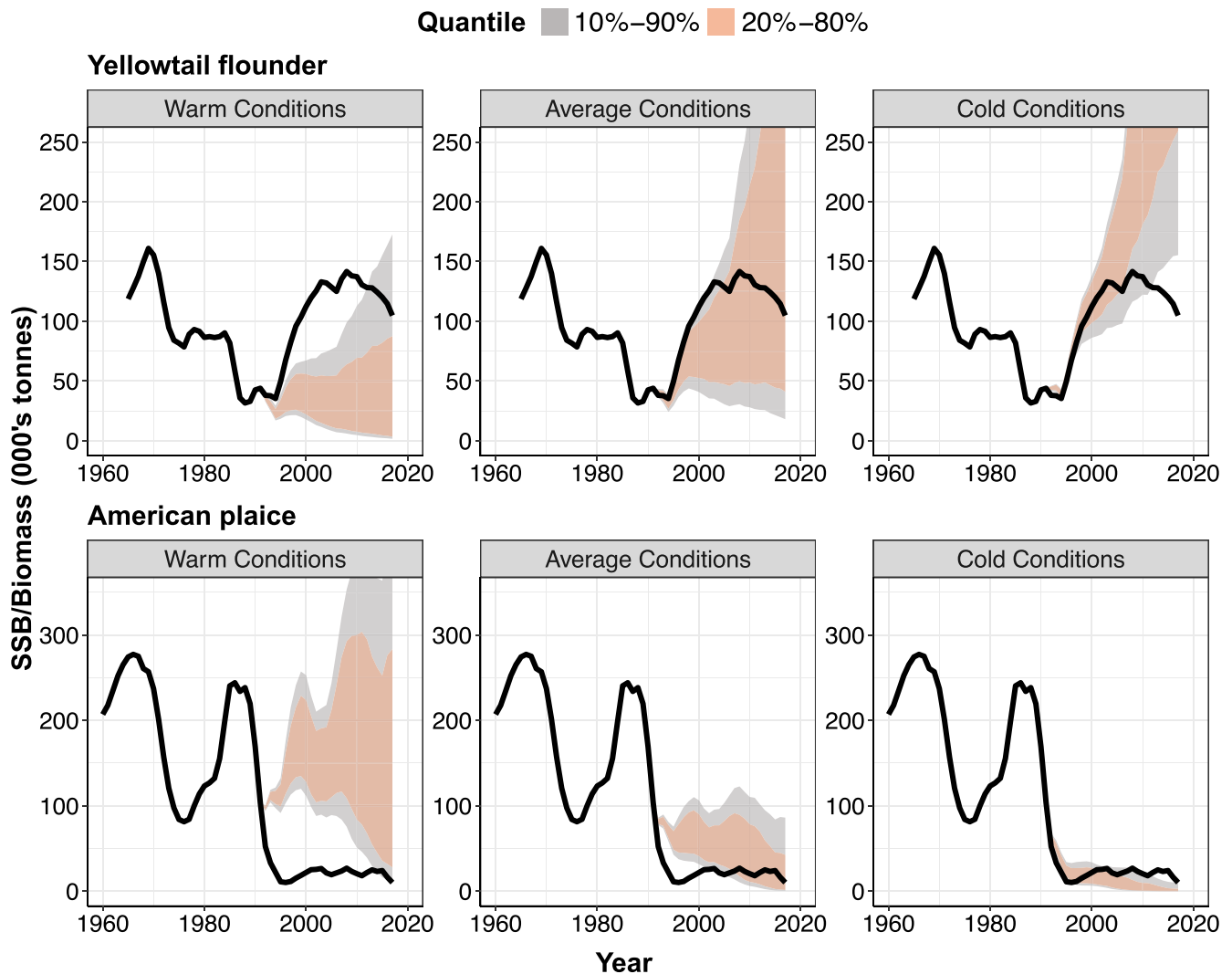
on population productivity, making it particularly relevant in the context of stock assessment. For example, by providing further evidence that plaice natural mortality has varied through time and impacted their recovery, our case study identified that changes in the ecosystem are at least partly responsible for this population's decreased productivity. This finding corroborates findings for other stocks in the region, where natural mortality impacted population collapse in the 1990s (Cadigan, 2015; Cadigan et al., 2022; Regular et al., 2022), thereby providing improved understanding about ecosystem impacts on fisheries.

#### 4.1 | Grand bank application

We sought to determine the utility of the proposed conceptual framework for testing models of increasing complexity to identify the population and ecosystem processes that affected the population dynamics for two flatfish populations on the Newfoundland Grand Banks. We observed that yellowtail population dynamics were primarily driven by recruitment variability, which was negatively

affected by warmer environmental conditions. Meanwhile, plaice population dynamics were affected by a combination of temporal variability in recruitment and  $M$ , where  $M$  increased during colder than average conditions. Simulations based on these best-performing models indicated that if there had been a different trajectory for the environmental drivers that affected these stocks, they may have exhibited drastically different pathways to recovery following their collapses in the early 1990s.

Multiple population processes contributed to time-varying productivity of flatfish on the Grand Banks. For example, our models indicate that plaice population collapse occurred in response to a dramatic decline in recruitment and increase in  $M$  and  $F$  in the late 1980s–early 1990s. However, these models indicated that recruitment deviations have been at or above-average since the late 1990s. Large pulses of pre-recruits (aged 0–5) have been identified in the recent stock assessment models as well (Wheeland et al., 2018, 2021). This indicates that population dynamics since the collapse have been primarily affected by the increasing fluctuations in  $M$  (and to some extent  $F$  in the early 2000s) rather than poor recruitment. Estimated increases in  $M$  corroborate past explorations



**FIGURE 7** Simulations of how population dynamics would have differed if their environmental drivers followed different trajectories from 1992 onwards. The polygons represent the 20%–80% (orange) and 10%–90% quantiles (grey) of dynamics from 1000 simulations. The different trajectories are based on shifting the random walk for the environmental covariates in each simulation with a constant. Since environmental covariates were scaled,  $-1$  was used for lower values (left panels),  $0$  was used for average values (middle panels) and  $1$  was used higher values (right panels).

of retrospective patterns for this stock, which indicated that an  $M=0.2$  assumption is an underestimate in recent years (Perreault et al., 2020; Wheeland, 2021). Although separating relative changes in  $M$ ,  $F$  and catchability is notoriously difficult, we followed best practice guidelines and estimated  $M$  with a prior (Punt et al., 2021). It seems increasingly likely that changes in plaice adult  $M$  have affected stock collapse and recovery; however, specifying different model configurations may influence the trend and magnitude of  $M$  estimates.

The NLCI accounted for variability in population processes for both yellowtail and plaice. Given that the Grand Banks are the northernmost extent of the range of yellowtail, the negative relationship between recruitment and the NLCI may indicate that cold conditions are an ecological indicator of beneficial recruitment habitat or reduced juvenile predation or competition. For example, a study on yellowtail in the Middle Atlantic Bight (Sullivan

et al., 2005) identified a strong negative correlation between recruitment and temperature and discussed that this relationship may be related to the match/mismatch hypothesis (Cushing, 1975). In this case, colder temperatures delay phytoplankton and zooplankton blooms, which may then better align zooplankton availability with the timing of initial feeding for yellowtail larvae. Phytoplankton bloom timing on the Grand Banks is negatively correlated with the NLCI (Cyr et al., 2023), indicating that yellowtail recruitment on the Grand Banks may also be affected by match/mismatch of larvae and their prey. Meanwhile, the timing of increases in plaice  $M$  that are explained by the NLCI matches the timing of the shift in plaice spatial distribution following exposure to the coldest temperatures in the late 1980s–early 1990s (Robertson et al., 2021). Given that plaice on the Grand Banks are relatively sedentary, a large migration in response to cold temperatures is unlikely (Morgan, 1996; Pitt, 1969). Therefore, it is

possible that the observed distributional shift represents a mass mortality event for the stock. However, the exact mechanisms regarding past and continued high levels of  $M$  remain uncertain.

Comparing alternative representations of population dynamics can highlight potential sources of uncertainty for fisheries management advice. Here, the biomass estimates from our models differed from the most recent stock assessment model estimates for both species. The delay-difference formulation of yellowtail population dynamics estimated that the magnitude of biomass recovery was lower than was estimated using the surplus production stock assessment model (Parsons et al., 2021; but see Zhang & Cadigan, 2022). This difference appears to be derived from differences in catchability estimates. The current stock assessment estimates that biomass is near maximum sustainable yield ( $B_{MSY}$ ), and therefore, any decreases in magnitude of recent biomass may affect current stock status and harvest advice. Meanwhile, estimating time-varying  $M$  in the plaice stock assessment model yielded different estimates of SSB and  $F$  from the current stock assessment model (Wheeland et al., 2021). Although the differences in SSB are unlikely to affect stock status, the high levels of  $M$  for adults may change the context for allowable levels of bycatch (Shelton & Morgan, 2006).

## 4.2 | Providing ecosystem-informed fisheries management advice

This modelling framework may assist in the determination of reference points that guide fishery management decisions. When stocks exhibit non-stationary productivity, reference points may need to be modified to ensure a match between their productivity state and the allowable levels of harvest (Berger, 2019; Zhang et al., 2021). In fact, assuming fixed values for life-history parameters limits the ways that data can inform reference points and can promote a false sense of precision (Mangel et al., 2013; O'Leary et al., 2020). However, a major barrier to implementation of so-called 'dynamic reference points' has been uncertainty about whether a change in ecosystem or stock productivity has occurred and is lasting (Eddy et al., 2023). By estimating non-stationary patterns in productivity that are linked to ecosystem drivers, this framework could be used to identify when implementing dynamic reference points may be appropriate. For example, fishing mortality reference points could be scaled based on ecosystem conditions (i.e.,  $F_{ECO}$ ; Howell et al., 2021) that were identified by examining models of increasing complexity. This would allow the incorporation of ecosystem information into catch advice without needing to modify current assessment models. Additionally, this framework lends itself to comparisons between static and dynamic reference points to allow managers to evaluate how assumptions made in stock assessments impact scientific advice.

By exploring alternative hypotheses about drivers that have influenced population dynamics, this framework provides a straightforward process for identifying potential operating models for MSE. Given that MSE often includes two sets of operating models, reference sets (i.e., most plausible hypotheses) and robustness sets (i.e.,

less likely but still possible hypotheses), model performance could guide the assignment of model plausibility (Punt et al., 2016). This process often uses expert judgement alone, and therefore, having a formalized model exploration approach to provide evidence for the plausibility of a given operating model may increase transparency and credibility. In this way, models that did not have the best performance could still be included as operating models to ensure that management decisions are robust to various sources of uncertainty.

Extending models to include new processes can produce worse predictive performance even when model selection criteria (e.g., AIC, model residuals) indicate that including new processes improves model fit (Collie et al., 2016; Trijoulet et al., 2019). Model validation via hindcasts and retrospective analyses can directly examine the predictive skill of models to understand whether overfitting is a problem and if a model can provide the short-term predictions necessary for tactical fisheries management (Breivik et al., 2023; Kell et al., 2021). Given that this modelling framework relies on model selection criteria, model validation will serve as an important, complementary process. For example, prior analyses identified that extending models to allow new population processes to vary over time reduced retrospective patterns but those extensions translated to worse management advice if the varying process was not correct (Szuwalski et al., 2018). Therefore, complimenting the modelling framework described here with model validation tools will allow practitioners to identify if extended models can produce reliable predictions while simultaneously improving certainty that the chosen time-varying process is correct.

Including environmental drivers in fisheries management advice often requires determining whether evidence of a relationship between an environmental variable and population productivity crosses a sufficient threshold (Klaer et al., 2015; Link et al., 2021). Requiring some threshold level of evidence may ensure that fisheries management does not continually shift baselines of productivity and thereby produce inappropriate reference points (Fulton et al., 2022; Schijns & Pauly, 2021). However, determining the required threshold of evidence and how it may be represented to inform fisheries management decisions likely cannot be generalized across systems (Fu et al., 2020; Hillebrand et al., 2020). As a result, there have been increasing calls for incorporating environmental considerations as a potential source of uncertainty that may affect the risks associated with management decisions. Various approaches for addressing environmental uncertainty have been developed, including the use of ensemble models, using multispecies estimates of predation mortality as inputs in single-species models, climate change conditioned advice and using environmentally-driven operating models in MSE (Brodziak & Piner, 2010; Duplisea et al., 2021; ICES, 2022; Jardim et al., 2021; Punt et al., 2016). The modelling framework developed here adds an approach for directly comparing a range of hypothesized sources of uncertainty on current population status. These sources of uncertainty can then be used to evaluate their potential effects on management outcomes. In general, this framework can be used to broadly consider the implications of



management decisions in the context of environmental variability without requiring the definition of a threshold of evidence for the environmental drivers.

## ACKNOWLEDGEMENTS

Research funding to M.D.R. was provided by the Vanier Canada Graduate Scholarship. Research funding to N.G.C. was provided by the Ocean Choice International Industry Research Chair programme at the Marine Institute of Memorial University of Newfoundland. Research funding to N.G.C. was provided by the Ocean Frontier Institute, through an award from the Canada First Research Excellence Fund. Research funding to T.D.E. was provided by the Natural Sciences and Engineering Research Council of Canada (NSERC) as well as Fisheries and Oceans Canada (DFO). We thank the Northwest Atlantic Fisheries Centre, Fisheries and Oceans Canada, and the many people involved in the collection and processing of these data. We also thank Gavin Fay, Paul Snelgrove and Jonathan Fisher for their helpful comments on an earlier version of this manuscript as part of M.D.R.'s PhD thesis examination committee.

## DATA AVAILABILITY STATEMENT

The data analysed during the case study are available from Fisheries and Oceans Canada at The Northwest Atlantic Fisheries Centre but restrictions apply to the availability of these data, which were used due to collaboration with Fisheries and Oceans Canada scientists, and are not publicly available. Data are, however, available from the authors associated with Fisheries and Oceans Canada and with permission of Fisheries and Oceans Canada.

## ORCID

Matthew D. Robertson  <https://orcid.org/0000-0003-3252-9956>

Noel G. Cadigan  <https://orcid.org/0000-0002-5922-1014>

Paul M. Regular  <https://orcid.org/0000-0003-0318-2615>

Frédéric Cyr  <https://orcid.org/0000-0002-1581-7502>

Fan Zhang  <https://orcid.org/0000-0003-0214-1790>

Tyler D. Eddy  <https://orcid.org/0000-0002-2833-9407>

## REFERENCES

- Baum, J. K., & Worm, B. (2009). Cascading top-down effects of changing oceanic predator abundances. *Journal of Animal Ecology*, 78(4), 699–714. <https://doi.org/10.1111/j.1365-2656.2009.01531.x>
- Berger, A. M. (2019). Character of temporal variability in stock productivity influences the utility of dynamic reference points. *Fisheries Research*, 217, 185–197. <https://doi.org/10.1016/j.fishres.2018.11.028>
- Breivik, O. N. N., Aldrin, M., Fuglebakk, E., & Nielsen, A. (2023). Detecting significant retrospective patterns in state space fish stock assessment. *Canadian Journal of Fisheries and Aquatic Sciences*, 10, 1–10. <https://doi.org/10.1139/cjfas-2022-0250>
- Brodie, W., Walsh, S. J., & Parsons, D. M. (2010). An evaluation of the collapse and recovery of the yellowtail flounder (*Limanda ferruginea*) fishery on the grand Bank. *ICES Journal of Marine Science*, 67(9), 1887–1895. <https://doi.org/10.1093/icesjms/fsq121>
- Brodziak, J., & Piner, K. (2010). Model averaging and probable status of north pacific striped marlin, *Tetrapturus audax*. *Canadian Journal of Fisheries and Aquatic Sciences*, 67(5), 793–805. <https://doi.org/10.1139/F10-029>
- Bruno, I., Costas, G., González, C., & Paz, X. (2000). Feeding chronology of yellowtail flounder (*Limanda ferruginea*) and American plaice (*Hippoglossoides platessoides*) on grand Bank (NAFO division 3N). *NAFO SCS*, 33, 103–116.
- Cadigan, N. G. (2015). A state-space stock assessment model for northern cod, including under-reported catches and variable natural mortality rates. *Canadian Journal of Fisheries and Aquatic Sciences*, 73(2), 296–308. <https://doi.org/10.1139/cjfas-2015-0047>
- Cadigan, N. G., Robertson, M. D., Nirmalkanna, K., & Zheng, N. (2022). The complex relationship between weight and length of Atlantic cod off the south coast of Newfoundland. *Canadian Journal of Fisheries and Aquatic Sciences*, 79(11), 1798–1819. <https://doi.org/10.1139/cjfas-2021-0325>
- Collie, J. S., Botsford, L. W., Hastings, A., Kaplan, I. C., Largier, J. L., Livingston, P. A., Plagányi, É., Rose, K. A., Wells, B. K., & Werner, F. E. (2016). Ecosystem models for fisheries management: Finding the sweet spot. *Fish and Fisheries*, 17(1), 101–125. <https://doi.org/10.1111/faf.12093>
- Correa, G. M., Monnahan, C. C., Sullivan, J. Y., Thorson, J. T., & Punt, A. E. (2023). Modelling time-varying growth in state-space stock assessments. *ICES Journal of Marine Science*, 80(7), 2036–2049. <https://doi.org/10.1093/icesjms/fsad133>
- Cushing, D. (1975). *Marine ecology and fisheries*. Cambridge University Press.
- Cyr, F., & Galbraith, P. S. (2021). A climate index for the Newfoundland and Labrador shelf. *Earth System Science Data*, 13, 1807–1828. <https://doi.org/10.5194/essd-13-1807-2021>
- Cyr, F., Lewis, K., Bélanger, D., Regular, P., Clay, S., & Devred, E. (2023). Physical controls and ecological implications of the timing of the spring phytoplankton bloom on the Newfoundland and Labrador shelf. *Limnology and Oceanography Letters*. <https://doi.org/10.1002/lol2.10347>
- Cyr, F., Snook, S., Bishop, C., Galbraith, P. S., Chen, N., & Han, G. (2021). *Physical oceanographic conditions on the Newfoundland and Labrador shelf during 2021*. In DFO Can. Sci. Advis. Sec. Res. Doc: Vol. 2022/040.
- Dempsey, D. P., Koen-Alonso, M., Gentleman, W. C., & Pepin, P. (2017). Compilation and discussion of driver, pressure, and state indicators for the grand Bank ecosystem, Northwest Atlantic. *Ecological Indicators*, 75, 331–339. <https://doi.org/10.1016/j.ecolind.2016.12.011>
- Denit, K. (2017). *Ecosystem-based fisheries management road map*. National Marine Fisheries Service Procedure (vol. 01, issues 120–01).
- Deriso, R. (1980). Harvesting strategies and parameter estimation for an age-structured model. *Canadian Journal of Fisheries and Aquatic Sciences*, 37, 268–282. <https://doi.org/10.1139/f80-034>
- DFO. (2007). *A new ecosystem science framework in support of integrated management*.
- DFO. (2019). *Assessment of 2J3KL capelin in 2018*. CSAS, 048, 1–19.
- Donaldson, M. R., Cooke, S. J., Patterson, D. A., & Macdonald, J. S. (2008). Cold shock and fish. *Journal of Fish Biology*, 73, 1491–1530. <https://doi.org/10.1111/j.1095-8649.2008.02061.x>
- Duplisea, D. E., Roux, M., Hunter, K. L., & Rice, J. (2021). Fish harvesting advice under climate change: A risk-equivalent empirical approach. *PLoS ONE*, 16(2), e0239503. <https://doi.org/10.1371/journal.pone.0239503>
- Dwyer, K. S., Rideout, R., Ings, D., Power, D., Morgan, J., Brodie, B., & Healey, B. P. (2016). *Assessment of American plaice in div. 3LNO*. NAFO SCS, 16/30(N6573), 1–19.
- Eddy, T. D., Duplisea, D. E., Robertson, M. D., Ruiz-Díaz, R., Solberg, C. A., & Zhang, F. (2023). Barriers to implementation of dynamic reference points in fisheries management. *Facets*, 8, 1–10. <https://doi.org/10.1139/facets-2022-0216>

- Fu, C., Xu, Y., Grüss, A., Bundy, A., Shannon, L., Heymans, J. J., Halouani, G., Akoglu, E., Lynam, C. P., Coll, M., Fulton, E. A., Velez, L., & Shin, Y. J. (2020). Responses of ecological indicators to fishing pressure under environmental change: Exploring non-linearity and thresholds. *ICES Journal of Marine Science*, 77(4), 1516–1531. <https://doi.org/10.1093/icesjms/fsz182>
- Fulton, E. A. (2021). Opportunities to improve ecosystem-based fisheries management by recognizing and overcoming path dependency and cognitive bias. *Fish and Fisheries*, 22(2), 428–448. <https://doi.org/10.1111/faf.12537>
- Fulton, E. A., Sainsbury, K., Noranartragoon, P., Leadbitter, D., Staples, D. J., Porobic, J., Ye, Y., Phoonsawat, R., & Kulanujaree, N. (2022). Shifting baselines and deciding on the desirable form of multispecies maximum sustainable yield. *ICES Journal of Marine Science*, 79(7), 2138–2154. <https://doi.org/10.1093/icesjms/fsac150>
- Goethel, D. R., Omori, K. L., Punt, A. E., Lynch, P. D., Berger, A. M., De Moor, C. L., Plagányi, É. E., Cope, J. M., Dowling, N. A., Mcgarvey, R., Preece, A. L., Thorson, J. T., Chaloupka, M., Gaichas, S., Gilman, E., Hesp, S. A., Longo, C., Yao, N., & Methot, R. D. (2022). Oceans of plenty? Challenges, advancements, and future directions for the provision of evidence – Based fisheries management advice. *Reviews in Fish Biology and Fisheries*, 33, 375–410. <https://doi.org/10.1007/s11160-022-09726-7>
- Gonzalez, C., Paz, X., Roman, E., & Hermida, M. (2006). *Feeding habits of fish species distributed on the grand Bank (NAFO divisions 3NO, 2002–2005)*. NAFO SCR, 06(31), 22.
- Halliday, R. G., & Pinhorn, A. T. (2009). The roles of fishing and environmental change in the decline of Northwest Atlantic groundfish populations in the early 1990s. *Fisheries Research*, 97(3), 163–182. <https://doi.org/10.1016/j.fishres.2009.02.004>
- Haltuch, M. A., Brooks, E. N., Brodziak, J., Devine, J. A., Johnson, K. F., Klibansky, N., Nash, R. D. M., Payne, M. R., Shertzer, K. W., Subbey, S., & Wells, B. K. (2019). Unraveling the recruitment problem: A review of environmentally-informed forecasting and management strategy evaluation. *Fisheries Research*, 217, 198–216. <https://doi.org/10.1016/j.fishres.2018.12.016>
- Hamel, O. S., Ianelli, J. N., Maunder, M. N., & Punt, A. E. (2023). Natural mortality: Theory, estimation and application in fishery stock assessment models. *Fisheries Research*, 261, 106638. <https://doi.org/10.1016/j.fishres.2023.106638>
- Heneghan, R. F., Galbraith, E., Blanchard, J. L., Harrison, C., Barrier, N., Bulman, C., Cheung, W., Coll, M., Eddy, T. D., Erasquin-Extramiana, M., Everett, J. D., Fernandes-Salvador, J. A., Gascuel, D., Guiet, J., Maury, O., Palacios-Abrantes, J., Petrik, C. M., du Pontavice, H., Richardson, A. J., ... Tittensor, D. P. (2021). Disentangling diverse responses to climate change among global marine ecosystem models. *Progress in Oceanography*, 198, 102659. <https://doi.org/10.1016/j.pcean.2021.102659>
- Hillebrand, H., Donohue, I., Harpole, W. S., Hodapp, D., Kucera, M., Lewandowska, A. M., Merder, J., Montoya, J. M., & Freund, J. A. (2020). Thresholds for ecological responses to global change do not emerge from empirical data. *Nature Ecology & Evolution*, 4(11), 1502–1509. <https://doi.org/10.1038/s41559-020-1256-9>
- Howell, D., Schueller, A. M., Bentley, J. W., Buchheister, A., Chagaris, D., Cieri, M., Drew, K., Lundy, M. G., Pedreschi, D., Reid, D. G., & Townsend, H. (2021). Combining ecosystem and single-species modeling to provide ecosystem-based fisheries management advice within current management systems. *Frontiers in Marine Science*, 7, 607831. <https://doi.org/10.3389/fmars.2020.607831>
- Hunsicker, M. E., Ciannelli, L., Bailey, K. M., Zador, S., & Stige, L. C. (2013). Climate and demography dictate the strength of predator-prey overlap in a subarctic marine ecosystem. *PLoS ONE*, 8(6), e66025. <https://doi.org/10.1371/journal.pone.0066025>
- ICES. (2022). Working group on the assessment of demersal stocks in the North Sea and Skagerrak (WGNSSK). *ICES Scientific Reports*, 4, 43. 1367 pp. <https://doi.org/10.17895/ices.pub.19786285>
- Jardim, E., Azevedo, M., Brodziak, J., Brooks, E. N., Johnson, K. F., Klibansky, N., Millar, C. P., Minto, C., Mosqueira, I., Nash, R. D. M., Vasilakopoulos, P., & Wells, B. K. (2021). Operationalizing ensemble models for scientific advice to fisheries management. *ICES Journal of Marine Science*, 78(4), 1209–1216. <https://doi.org/10.1093/icesjms/fsab010>
- Kell, L. T., Sharma, R., Kitakado, T., Winker, H., Mosqueira, I., Cardinale, M., & Fu, D. (2021). Validation of stock assessment methods: Is it me or my model talking? *ICES Journal of Marine Science*, 78(6), 2244–2255. <https://doi.org/10.1093/icesjms/fsab104>
- Klaer, N. L., O'Boyle, R. N., Deroba, J. J., Wayte, S. E., Little, L. R., Alade, L. A., & Rago, P. J. (2015). How much evidence is required for acceptance of productivity regime shifts in fish stock assessments: Are we letting managers off the hook? *Fisheries Research*, 168, 49–55. <https://doi.org/10.1016/j.fishres.2015.03.021>
- Koen-Alonso, M., Lindstrøm, U., Cuff, A., & Pennino, M. G. (2021). Comparative modeling of cod-capelin dynamics in the Newfoundland-Labrador shelves and Barents Sea ecosystems. *Frontiers in Marine Science*, 8, 579946. <https://doi.org/10.3389/fmars.2021.579946>
- Koen-Alonso, M., Pepin, P., Fogarty, M. J., Kenny, A., & Kenchington, E. (2019). The Northwest Atlantic fisheries organization roadmap for the development and implementation of an ecosystem approach to fisheries: Structure, state of development, and challenges. *Marine Policy*, 100, 342–352. <https://doi.org/10.1016/j.marpol.2018.11.025>
- Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H., & Bell, B. (2016). TMB: Automatic differentiation and Laplace approximation. *Journal of Statistical Software*, 70(5), 21. <https://doi.org/10.18637/jss.v070.i05>
- Kuparinen, A., Mäntyniemi, S., Hutchings, J. A., & Kuikka, S. (2012). Increasing biological realism of fisheries stock assessment: Towards hierarchical Bayesian methods. *Environmental Reviews*, 20(2), 135–151. <https://doi.org/10.1139/a2012-006>
- Link, J. S., Karp, M., Lynch, P., Morrison, W., & Peterson, J. (2021). Proposed business rules to incorporate climate-induced changes in fisheries management. *ICES Journal of Marine Science*, 78(10), 3562–3580. <https://doi.org/10.1093/icesjms/fsab219>
- Lorenzen, K. (2016). Toward a new paradigm for growth modeling in fisheries stock assessments: Embracing plasticity and its consequences. *Fisheries Research*, 180, 4–22. <https://doi.org/10.1016/j.fishres.2016.01.006>
- Lorenzen, K. (2022). Size- and age-dependent natural mortality in fish populations: Biology, models, implications, and a generalized length-inverse mortality paradigm. *Fisheries Research*, 255, 106454. <https://doi.org/10.1016/j.fishres.2022.106454>
- Mangel, M., MacCall, A. D., 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(6), 930–940. <https://doi.org/10.1139/cjfas-2012-0372>
- Marshall, K. N., Jensen, O. P., Koehn, L. E., Levin, P. S., & Essington, T. E. (2019). Inclusion of ecosystem information in US fish stock assessments suggests progress toward ecosystem-based fisheries management. *ICES Journal of Marine Science*, 76(1), 1–9. <https://doi.org/10.1093/icesjms/fsy152>
- Maunder, M. N., Hamel, O. S., Lee, H., Piner, K. R., Cope, J. M., Kapur, M. S., Punt, E., Ianelli, J. N., Castillo-jord, C., & Methot, R. D. (2023). A review of estimation methods for natural mortality and their performance in the context of fishery stock assessment. *Fisheries Research*, 257(4), 106489. <https://doi.org/10.1016/j.fishres.2022.106489>
- Maunder, M. N., & Watters, G. M. (2003). A general framework for integrating environmental time series into stock assessment models: Model description, simulation testing, and example article. *Fisheries Bulletin*, 101, 89–99.

- Morgan, M. J. (1992). Low-temperature tolerance of American plaice in relation to declines in abundance. *Transactions of the American Fisheries Society*, 121, 399–402. [https://doi.org/10.1577/1548-8659\(1992\)121<0399:Itopi>2.3.co;2](https://doi.org/10.1577/1548-8659(1992)121<0399:Itopi>2.3.co;2)
- Morgan, M. J. (1996). Preliminary results of tagging experiments on American plaice in NAFO Divs. 3LNO. NAFO SCR Doc, 96/61(N2737), 13.
- Morgan, M. J., Bailey, J., Healey, B. P., Maddock Parsons, D., & Rideout, R. (2011). Recovery potential assessment of American plaice (*Hippoglossoides platessoides*) in Newfoundland and Labrador. In DFO Can. Sci. Advis. Sec. Res. Doc.: Vol. 2011/047.
- Morgan, M. J., & Brodie, W. (2001). An exploration of virtual population analyses for divisions 3LNO American plaice. NAFO SCR Doc, 01/4(N4638), 1–20.
- Morgan, M. J., Brodie, W., Bowering, W., Maddock Parsons, D., & Orr, D. (1998). Results of data conversions for American plaice in div. 3LNO from comparative fishing trials between the Engel otter trawl and the Campelen 1800 shrimp trawl. NAFO SCR Doc, 98/70(N3062), 10.
- Myers, R. A. (1998). When do environment – Recruitment correlations work? *Reviews in Fish Biology and Fisheries*, 8, 285–305. <https://doi.org/10.1023/A:1008828730759>
- Neuenhoff, R. D., Swain, D. P., Cox, S. P., McAllister, M. K., Trites, A. W., Walters, C. J., & Hammill, M. O. (2018). 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(1), 168–184. <https://doi.org/10.1139/cjfas-2017-0190>
- O'Leary, C. A., Miller, T. J., Thorson, J. T., & Nye, J. A. (2019). Understanding historical summer flounder (*Paralichthys dentatus*) abundance patterns through the incorporation of oceanography-dependent vital rates in Bayesian hierarchical models. *Canadian Journal of Fisheries and Aquatic Sciences*, 76, 1275–1294. <https://doi.org/10.1139/cjfas-2018-0092>
- O'Leary, C. A., Thorson, J. T., Miller, T. J., & Nye, J. A. (2020). Comparison of multiple approaches to calculate time-varying biological reference points in climate-linked population-dynamics models. *ICES Journal of Marine Science*, 77(3), 930–941. <https://doi.org/10.1093/icesjms/fsz215>
- Parsons, D., Rideout, R., & Rogers, R. (2021). 2021 Assessment of yellow-tail flounder in NAFO divisions 3LNO using a Stock production model in a Bayesian framework. NAFO SCR Doc, 21/018(7186), 48.
- Parsons, L. S., & Lear, W. H. (2001). Climate variability and marine ecosystem impacts: A North Atlantic perspective. *Progress in Oceanography*, 49, 167–188. [https://doi.org/10.1016/S0079-6611\(01\)00021-0](https://doi.org/10.1016/S0079-6611(01)00021-0)
- Pepin, P., King, J., Holt, C., Smith, H. G., Shackell, N., Hedges, K., & Bundy, A. (2022). Incorporating knowledge of changes in climatic, oceanographic and ecological conditions in Canadian stock assessments. *Fish and Fisheries*, 23, 1332–1346. <https://doi.org/10.1111/faf.12692>
- Perreault, A. M. J., Wheeland, L. J., & Morgan, M. J. (2020). A state-space stock assessment model for American plaice on the grand bank of Newfoundland. *Journal of Northwest Atlantic Fishery Science*, 51, 45–104. <https://doi.org/10.2960/J.v51.m727>
- Pitt, T. K. (1969). Migrations of American plaice on the grand Bank and in St. Mary's bay, 1954, 1959, and 1961. *Journal of the Fisheries Research Board of Canada*, 26, 1301–1319. <https://doi.org/10.1139/f69-115>
- Plagányi, É. E. (2007). *Models for an ecosystem approach to fisheries*.
- Plagányi, É. E., Blamey, L. K., Rogers, J. G. D., & Tulloch, V. J. D. (2022). Playing the detective: Using multispecies approaches to estimate natural mortality rates. *Fisheries Research*, 249, 106229. <https://doi.org/10.1016/j.fishres.2022.106229>
- Plagányi, É. E., Punt, A. E., Hillary, R., Morello, E. B., Thébaud, O., Hutton, T., Pillans, R. D., Thorson, J. T., Fulton, E. A., Smith, A. D. M., Smith, F., Bayliss, P., Haywood, M., Lyne, V., & Rothlisberg, P. C. (2014). Multispecies fisheries management and conservation: Tactical applications using models of intermediate complexity. *Fish and Fisheries*, 15(1), 1–22. <https://doi.org/10.1111/j.1467-2979.2012.00488.x>
- Pope, J. G., Gislason, H., Rice, J. C., & Daan, N. (2021). Scrabbling around for understanding of natural mortality. *Fisheries Research*, 240, 105952. <https://doi.org/10.1016/j.fishres.2021.105952>
- Punt, A. E., Butterworth, D. S., de Moor, C. L., De Oliveira, J. A. A., & Haddon, M. (2016). Management strategy evaluation: Best practices. *Fish and Fisheries*, 17(2), 303–334. <https://doi.org/10.1111/faf.12104>
- Punt, A. E., Castillo-Jordan, C., Hamel, O. S., Cope, J. M., Maunder, M. N., & Ianelli, J. N. (2021). Consequences of error in natural mortality and its estimation in stock assessment models. *Fisheries Research*, 233, 105759. <https://doi.org/10.1016/j.fishres.2020.105759>
- Punt, A. E., Dunn, A., Þór Elvarsson, B., Hampton, J., Hoyle, S. D., Maunder, M. N., Methot, R. D., & Nielsen, A. (2020). Essential features of the next-generation integrated fisheries stock assessment package: A perspective. *Fisheries Research*, 229, 105617. <https://doi.org/10.1016/j.fishres.2020.105617>
- R Core Team. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- Regular, P. M., Buren, A. D., Dwyer, K. S., Cadigan, N. G., Gregory, R. S., Koen-alonso, M., Rideout, R. M., Robertson, G. J., Robertson, M. D., Stenson, G. B., Wheeland, L. J., & Zhang, F. (2022). Indexing starvation mortality to assess its role in the population regulation of northern cod. *Fisheries Research*, 247, 106180. <https://doi.org/10.1016/j.fishres.2021.106180>
- Reum, J. C. P., Townsend, H., Gaichas, S., Sagarese, S., Kaplan, I. C., & Grüss, A. (2021). It's not the destination, It's the journey: Multispecies model ensembles for ecosystem approaches to fisheries management. *Frontiers in Marine Science*, 8, 631839. <https://doi.org/10.3389/fmars.2021.631839>
- Rice, J. (2011). Managing fisheries well: Delivering the promises of an ecosystem approach. *Fish and Fisheries*, 12(2), 209–231. <https://doi.org/10.1111/j.1467-2979.2011.00416.x>
- Robertson, M. D., Gao, J., Regular, P. M., Morgan, M. J., & Zhang, F. (2021). Lagged recovery of fish spatial distributions following a cold – water perturbation. *Scientific Reports*, 11(9513), 1–11. <https://doi.org/10.1038/s41598-021-89066-x>
- Robertson, M. D., Koen-Alonso, M., Regular, P., Cadigan, N., & Zhang, F. (2022). Accounting for a nonlinear functional response when estimating prey dynamics using predator diet data. *Methods in Ecology and Evolution*, 13(4), 880–893. <https://doi.org/10.1111/2041-210X.13795>
- Robertson, M. D., Regular, P. M., & Cadigan, N. G. (2022). Limited temporal variability in natural mortality for juvenile American plaice on the grand Bank of Newfoundland. *Journal of Northwest Atlantic Fishery Science*, 53, 47–56. <https://doi.org/10.2960/J.v53.m738>
- Schijns, R., & Pauly, D. (2021). Management implications of shifting baselines in fish stock assessments. *Fisheries Management and Ecology*, 29(2), 183–195. <https://doi.org/10.1111/fme.12511>
- Schnute, J. (1987). A general fishery model for a size-structured fish population. *Canadian Journal of Fisheries and Aquatic Sciences*, 44(5), 924–940. <https://doi.org/10.1139/f87-111>
- Shelton, P. A., & Morgan, M. J. (2006). Is by-catch mortality preventing the rebuilding of cod (*Gadus morhua*) and American plaice (*Hippoglossoides platessoides*) stocks on the grand Bank? *Journal of Northwest Atlantic Fishery Science*, 36, 1–17. <https://doi.org/10.2960/J.v36.m544>
- Simpson, M. R., Miri, C. M., & Collins, R. K. (2018). Assessment of thorny skate (*Amblyraja radiata* Donovan, 1808) in NAFO divisions 3LNO and subdivision 3Ps. NAFO SCR Doc, 18/027(N6811), 1–34.
- Skern-Mauritzen, M., Ottersen, G., Handegard, N. O., Huse, G., Dingsør, G. E., Stenseth, N. C., & Kjesbu, O. S. (2016). Ecosystem processes are rarely included in tactical fisheries management. *Fish and Fisheries*, 67(3), 165–175. <https://doi.org/10.1111/faf.12111>

- Sober, E. (2002). Instrumentalism, parsimony, and the Akaike framework. *Philosophy of Science*, 69(S3), 112–123. <https://doi.org/10.1086/341839>
- Sosebee, K. A., Simpson, M. R., & Miri, C. M. (2022). Assessment of thorny skate (*Amblyraja radiata* Donovan, 1808) in NAFO divisions 3LNO and subdivision 3Ps. NAFO SCR Doc, 22/026(N7299), 32.
- Stock, B. C., & Miller, T. J. (2021). The woods hole assessment model (WHAM): A general state-space assessment framework that incorporates time- and age-varying processes via random effects and links to environmental covariates. *Fisheries Research*, 240, 105967. <https://doi.org/10.1016/j.fishres.2021.105967>
- Sullivan, M. C., Cowen, R. K., & Steves, B. P. (2005). Evidence for atmosphere-ocean forcing of yellowtail flounder (*Limanda ferruginea*) recruitment in the middle Atlantic bight. *Fisheries Oceanography*, 14(5), 386–399. <https://doi.org/10.1111/j.1365-2419.2005.00343.x>
- Szuwalski, C. S., Ianelli, J. N., & Punt, A. E. (2018). Reducing retrospective patterns in stock assessment and impacts on management performance. *ICES Journal of Marine Science*, 75(2), 596–609. <https://doi.org/10.1093/icesjms/fsx159>
- Szuwalski, C. S., Vert-Pre, K. A., Punt, A. E., Branch, T. A., & Hilborn, R. (2015). Examining common assumptions about recruitment: A meta-analysis of recruitment dynamics for worldwide marine fisheries. *Fish and Fisheries*, 16(4), 633–648. <https://doi.org/10.1111/faf.12083>
- Thorson, J. T., & Minto, C. (2015). Mixed effects: A unifying framework for statistical modelling in fisheries biology. *ICES Journal of Marine Science*, 72(5), 1245–1256. <https://doi.org/10.1093/icesjms/fsu213>
- Townsend, H., Harvey, C. J., Davis, D., Zador, S. G., Gaichas, S., Weijerman, M., Hazen, E. L., & Kaplan, I. C. (2019). Progress on implementing ecosystem-based fisheries Management in the United States through the use of ecosystem models and analysis. *Frontiers in Marine Science*, 6(641), 1–17. <https://doi.org/10.3389/fmars.2019.00641>
- Trijoulet, V., Fay, G., & Miller, T. J. (2019). Performance of a state – Space multispecies model: What are the consequences of ignoring predation and process errors in stock assessments? *Journal of Applied Ecology*, 57(1), 121–135. <https://doi.org/10.1111/1365-2664.13515>
- Walsh, S. J., Simpson, M., & Morgan, M. J. (2004). Continental shelf nurseries and recruitment variability in American plaice and yellowtail flounder on the grand Bank: Insights into stock resiliency. *Journal of Sea Research*, 51(3–4), 271–286. <https://doi.org/10.1016/j.seares.2003.10.003>
- Wheeland, L. (2021). An exploration of the impact of natural mortality assumptions in a virtual population analysis for divisions 3LNO American plaice. NAFO SCR Doc, 21/025(N7193), 16.
- Wheeland, L., Dwyer, K., Kumar, R., Rideout, R., Perreault, A., & Rogers, B. (2021). Assessment of American plaice in div. 3LNO. NAFO SCR Doc, 21(035).
- Wheeland, L., Dwyer, K., Morgan, J., Rideout, R., & Rogers, R. (2018). Assessment of American plaice in div. 3LNO. NAFO SCS Doc, 18/039(N6829), 77.
- Zhang, F., & Cadigan, N. G. (2022). An age- and length-structured statistical catch-at-length model for hard-to- age fisheries stocks. *Fish and Fisheries*, 23, 1121–1135. <https://doi.org/10.1111/faf.12673>
- Zhang, F., Regular, P. M., Wheeland, L., Rideout, R. M., & Morgan, M. J. (2021). Accounting for non-stationary stock-recruitment relationships in the development of MSY-based reference points. *ICES Journal of Marine Science*, 78(6), 2233–2243. <https://doi.org/10.1093/icesjms/fsaa176>

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Robertson, M. D., Cadigan, N. G., Regular, P. M., Koen-Alonso, M., Cyr, F., Zhang, F., & Eddy, T. D. (2024). Testing models of increasing complexity to develop ecosystem-informed fisheries advice. *Fish and Fisheries*, 00, 1–17. <https://doi.org/10.1111/faf.12820>