Canadian Journal of Fisheries and Aquatic Sciences

OPEN ACCESS | Article

Age structure augments the predictive power of time series for fisheries and conservation

Tara E. Dolan[a], **Eric P. Palkovacs**[a], **Tanya L. Rogers**[b], **and Stephan B. Munch**[b,c]

[a]Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, CA 95060, USA; [b]Southwest Fisheries Science Center, National Oceanic and Atmospheric Administration, Santa Cruz, CA 95060, USA; [c]Department of Applied Mathematics, University of California, Santa Cruz, CA 95060, USA

Corresponding author: **Tara E. Dolan** (email: tara.dolan@mass.gov)

Abstract

Ecological forecasts are potentially of great value for managing fisheries and for stakeholders dependent on their long-term sustainability. Yet, most forecasting approaches are data-intensive, requiring information not just on the focal species but also on ecological interactions and the physical environment. Empirical dynamic modeling (EDM) is an equation-free approach to forecasting species' abundance using only data on past abundance, but the time series required for this approach must be long enough to reconstruct the dynamics of the system. This requirement is rarely met, especially for long-lived species. Here we used simulations and empirical data to demonstrate that incorporating time series from multiple age classes can improve our ability to forecast abundance compared to a single age class or index of total abundance. Including data from multiple age classes produced the greatest gains in forecast accuracy when time series were the shortest. Overall, our results show that the incorporation of age structure could allow EDM to be applied to many species for which relatively short time series would have previously been a limiting factor.

Key words: forecast, age structure, ecology, population dynamics, empirical dynamic modeling

Introduction

Predicting population changes is a critical component of sustainable fisheries management and conservation biology (Clark et al. 2001; Hobday et al. 2016; Payne et al. 2017; Jacox et al. 2020). Accurate forecasts can help managers develop strategies to mitigate population fluctuations caused by overfishing, climate change, and ecosystem interactions (Tommasi et al. 2017; Dietze et al. 2018; Miller et al. 2019; Pennekamp et al. 2019). Recent years have seen a rapid expansion in the ability of population models to predict spatial distributions and fish-stock recruitment (Subbey et al. 2014; Munch et al. 2018; Thorson 2019; Brodie et al. 2021). Yet, producing accurate forecasts remains a challenge because ecosystems are complex, high-dimensional systems, and nonlinear dynamics are common (Sugihara et al. 2011; Dakos et al. 2017), including in harvested species (Glaser et al. 2014*a*). Mechanistic models of ecosystems often require data on many system variables and their interactions (Jacox et al. 2020); data that are often not available (Dietze et al. 2018; Lasky et al. 2020; Munch et al. 2020). Even for correctly specified models, slight changes in model structure can lead to qualitatively different predictions (Wood and Thomas 1999; Walters et al. 2016).

Empirical dynamic modeling (EDM) is an alternative approach to classic mechanistic models that addresses some of these limitations. To account for uncertainty in ecological relationships, EDM uses nonparametric models that make few assumptions about the functional form or relationships, instead relying on available data to determine the relationships' shapes (DeAngelis and Yurek 2015; Chang et al. 2017; Munch et al. 2020). To account for missing variables, EDM uses delay embedding, in which time lags of one variable (delay embedding vectors) can stand in for missing dimensions and be used to make predictions (Takens 1981; Sugihara 1994; Ye et al. 2015*a*). The robustness of EDM and its utility to ecological forecasting have been demonstrated across many use cases (Ye et al. 2015; Deyle et al. 2018; Munch et al. 2020). Data-driven approaches like EDM can outperform even correctly specified mechanistic models (Perretti et al. 2013). Though only a single time series is needed, EDM has the flexibility to incorporate additional observations from the system, including environmental covariates (Deyle et al. 2013; Ye et al. 2015; Wasserman et al. 2022), additional species (Liu et al. 2012; Deyle et al. 2016, 2018; Ye and Sugihara 2016; Brias and Munch 2021), and spatial information (McNamara et al. 2019; Wang et al. 2020; Johnson et al. 2021).

One possible reason why EDM has not been widely adopted is that it requires long time series to produce accurate predictions. Specifically, the length of time series should be several multiples of the characteristic return time of the system (i.e., the full cycle of states the system make take on, such that past behavior can be used to predict future states, Munch et

al. 2020). Although this return time is difficult to know a priori, Munch et al. (2018) found that maturation age was a useful proxy when predicting fish recruitment, such that EDM models trained on time series with length 10 times the mean age at maturation recovered approximately 50% of variation in recruitment (Munch et al. 2018). Most population monitoring for fisheries and conservation in the United States began in the late 1960s[1] and many available surveys are even shorter (Melnychuk et al. 2021). Thus, time series for many species do not meet this minimum length guideline to apply EDM.

There are a few solutions to the challenge of short time series, many of which make use of additional, contemporaneously collected data. For example, data from other species in the ecosystem can be included as predictors (Hsieh et al. 2008; Ye and Sugihara 2016; Deyle et al. 2018; Kuriyama et al. 2020), or data from multiple populations in space can be leveraged (Glaser et al. 2014*b*; Johnson et al. 2021). Many population surveys collect information on age structure in addition to abundance. In light of the improvements obtained using spatially replicated time series or multispecies data, we hypothesize that multiple time series representing different age groups could be leveraged to improve forecasts, as is done in traditional data-rich stock assessments (Patterson et al. 2001; Botsford et al. 2011). However, age-structured data and spatial replicates have some fundamental differences. In the absence of dramatic fluctuations in fishing mortality, the assumption of constant natural mortality means that the ratio of abundance between one age class and another is roughly constant. In other words, each age-specific time series is a scaled copy. In reality, if maturation, survival, and growth are not constant (Gulland 1965; Hutchings 1993; Van der Veer et al. 2000; Secor 2007; White et al. 2014), then the time series of abundance from different ages will not maintain constant proportionality. Previous attempts to use additional information to improve prediction skill of short time series have emphasized that replicate groups (in this case, age classes) must be similar enough so that their dynamics are mutually informative but different enough to provide new information (Ye and Sugihara 2016; Munch et al. 2017; Johnson et al. 2021). We hypothesized that age-structured time series, in a sufficiently complex system, would contribute new information that will improve forecast skill relative to time series of a single age or total abundance.

There are several ways in which information can be shared across groups, including but not limited to random embedding, multiview approaches, and hierarchical approaches (Ye and Sugihara 2016; Munch et al. 2017; Ma et al. 2018). The latter, hierarchical approaches, allow information to be shared across datasets without assuming they are identical (Munch et al. 2017; Rogers and Munch 2020). In the case of age-structured time series, the hierarchical approach combines delay embedding vectors from multiple age classes to predict the future abundance of each age class and uses the data from multiple ages to improve estimation of the delay embedding model. The hierarchical model structure allows the degree of dynamic similarity between groups to be determined from the data and quantified in a parameter called dynamic correlation (ρ, Rogers and Munch 2020). Alternatively, one could predict the focal age class using lags of neighboring age classes, analogous to including abundance data from neighboring sites in spatial EDM (Johnson et al. 2021). Although in the spatial case the hierarchical approach was generally superior to ignoring spatial structure, it is not clear whether age structure will provide similar improvements.

The goal of our study was to examine the utility of age-structured data for nonlinear forecasting. First, we compared the performance of two approaches that incorporate age structure with that of two approaches that do not consider age structure. The approaches considering age structure were a *hierarchical* model, which allows for shared dynamics across single-age delay-embedding models, and a *mixed-age* model, in which lags from the focal and adjacent age classes are used as predictors. Approaches not considering age structure were a *single-age* model, in which only lags of the focal age class are used, and a *total abundance* model, in which abundance is summed across age classes and a delay embedding model is fit to total abundance. This analysis was done for three simulated datasets of varying complexity and two empirical datasets. Second, we used the three simulated datasets and the hierarchical method to explore exchangeability between the number of age classes and time series length. Specifically, we tested whether age structure enables us to make use of shorter time series for forecasting. Finally, we explored the use of the dynamic correlation (ρ) as an indicator of dynamic similarity across age classes and the potential for shared information across age classes to improve predictive performance. Our overall goal was to explore the utility of age structure information to improve the predictive power of time series, which could open the possibility of applying EDM to many more cases than currently feasible.

Materials and methods

We begin by describing the simulated and empirical datasets used in our analysis, and then present a description of our overall modeling framework (Gaussian process (GP) EDM). We then describe how different models, either standard models without age structure (single-age, total abundance) or models with age structure (hierarchical, mixed-age), can be formulated within this framework, whose performance we compare. Next, we describe our analysis exploring the relative utility of additional age classes versus longer time series for improving forecasts, and our methods for quantifying pairwise dynamic correlation between age classes.

Simulations

To test our age-structured EDM approaches, we developed three simulated, age-structured population datasets representing increasing levels of ecological complexity (Simulations I–III, Table 1). Simulation I was a one-species system where recruitment to age 0 was governed by the Ricker model (Ricker 1954). There were 20 age classes, all of which had constant survival and constant fecundity after maturation at age

[1] Azarovitz, T., Clark, S., Despres, L., and Byrne, C.J. 1997. The Northeast Fisheries Science Center bottom trawl survey program. ICES CM Doc No 1997/Y:33, Copenhagen, Denmark.

Table 1. Model details for simulated datasets.

Simulation	Process	Model	Parameters
I. One species: x *low complexity*	Recruitment	$r1 = be^{\phi\,(1\,-\,b)}e^{c\,\in}$	$\phi = \{0.001\}$ $c = \{1\}$ $\in \sim N(0,1)$
	Fecundity	$b = \sum_{i}^{j} \boldsymbol{x}_{i-1} I(b_0)$	$I(b_0) = \begin{cases} 0, \text{ if } j < 4 \\ 100, \text{ if } j \geq 4 \end{cases}$
	Survival	$S = 0.8$	
II. Two species: x_1, x_2 *medium complexity*	Recruitment	$r1 = b_1 e^{\,-\,\phi_1 b_1\,-\,\beta_1 x_2}e^{c_1\,\in}$ $r2 = x_2 e^{\,-\,\phi_2 x_2\,-\,\beta_2 x_1}e^{c_2\,\in}$	$\phi_{1,\,2} = \{0.001,\ 0.01\}$ $\beta_{1,\,2} = \{0.002,\ 0.001\}$ $c_{1,\,2} = \{1\}$ $\in \sim N(0,1)$ $N_{01,02} = \{1000\ \ 100\}$
	Fecundity	$b_1 = \sum_{i}^{j} \boldsymbol{x}_{1i-1} I(b_{0_1})$	$I(b_{0_1}) = \begin{cases} 0, \text{ if } j < 3 \\ 200, \text{ if } j \geq 3 \end{cases}$
		$b_2 = \sum_{i}^{j} \boldsymbol{x}_{2i-1} I(b_{0_2})$	$I(b_{0_2}) = \begin{cases} 0, \text{ if } j < 4 \\ 10, \text{ if } j \geq 4 \end{cases}$
	Survival	$S_{1\,ji} = \sum_{j} \boldsymbol{x}_{1i-1} e^{\frac{-\sum_{j} \boldsymbol{x}_{2i-1}}{\sum_{j} \boldsymbol{x}_{2_{i-1}} + \sum_{j} \boldsymbol{x}_{1i-1}}} +$ $I(p) \sum_{j-1}^{1} \boldsymbol{x}_{1i-1} e^{\frac{-\sum_{j-1}^{1} \boldsymbol{x}_{2i-1}}{\sum_{j-1}^{1} \boldsymbol{x}_{2_{i-1}} + \sum_{j-1}^{1} \boldsymbol{x}_{1_{i-1}}}}$ $S_{2_{ji}} = 0.8$	$I(p) = \begin{cases} 0, \text{ if } j < 20 \\ 1, \text{ if } j \geq 20 \end{cases}$
III. Two species: x_1, x_2 *high complexity*	Recruitment	$r_{2i+1} = \sum_{j}^{i} (b_1)\, e^{-\phi_2 \sum_{j}^{i} (b_2)} e^{c_2 \in}$ $r_{2i+1} = \sum_{j}^{i} (b_1)\, e^{-\phi_2 \sum_{j}^{i} (b_2)} e^{c_2 \in}$	$\phi_1 = 0.001$ $\phi_2 = 0.01$ $c_{1,\,2} = \{0.01\}$ $\in \sim N(0,0.2)$ $N_{01,02} = \{1000\,100\}$
	Fecundity	$b_1 = x_{1\,j,\,i-1} I(b_{0_1}) + I(q_1) \sum_{j} \mathrm{P}_{1\mathrm{k}}, \mathrm{x}_{2\mathrm{i}-1}$	$I(b_{0_1}) = \begin{cases} 0, \text{ if } j < 3 \\ 200, \text{ if } j \geq 3 \end{cases}$
		$b_2 = x_{2\,j,\,i-1} I(b_{0_2}) + I(q_2) \sum_{j} \mathrm{P}_{2\mathrm{k}}, \mathrm{x}_{1\mathrm{i}-1}$	$I(b_{0_2}) = \begin{cases} 0, \text{ if } j < 4 \\ 2, \text{ if } j \geq 4 \end{cases}$
			$I(q_1) = \begin{cases} 0, \text{ if } j < 3 \\ 0.8, \text{ if } j \geq 3 \end{cases}$
			$I(q_2) = \begin{cases} 0, \text{ if } j < 4 \\ 0.8, \text{ if } j \geq 4 \end{cases}$
			P_1 and P_2 are matrices of age-specific predation (Supplementary Tables S1 and S2)
	Survival	$x_{1\,ji} = x_{1\,j-1,\,i-1} e^{\sum_{j} \mathrm{P}_{1\mathrm{k}}, \mathrm{x}_{2\mathrm{i}-1}}$ $x_{2\,ji} = x_{2\,j-1,\,i-1} e^{\sum_{j} \mathrm{P}_{2\mathrm{k}}, \mathrm{x}_{1\mathrm{i}-1}}$	P_1 and P_2 are matrices of age-specific predation

4. Simulation II was a two-species simulation of a predator–prey system, where data were collected only for the prey species. Both species had 20 age classes and a "21 + group". Both species had Ricker recruitment to age 0 and constant fecundity after maturation. The predator matured at age 4 and the prey matured at age 3, and the prey species was more fecund than the predator. The predator had constant survival, whereas prey survival depended on predator abundance. Simulation III was equivalent to Simulation II, but the predator had age-specific consumption rates and consumption-

dependent fecundity (Supplementary Tables S1 and S2), and the prey consumed the eggs (age 0) of the predator at a rate that increased with prey age (Köster and Möllmann 2000). Thus, age-specific mortality rates were constant in Simulation I but varied with system state in Simulations II and III.

Simulations were run for 100 years after a 300-year burn-in period to remove transient dynamics. Simulations were initialized with a starting prey population of 1000 age-0 individuals, and in the case of two-species simulations, a starting predator population of 100 age-0 individuals. For both predator and prey, age-1 starting values were 0.8 times the age-0 starting value, age-2 starting values were 0.8 times age-1 values, and so forth. Stochasticity was added during the recruitment phase in the form of multiplicative random noise drawn from a normal distribution with a mean of 0 and standard deviation of 1. The resulting abundance time series was log-transformed before model fitting, which helped to resolve dynamics during periods of low abundance that would otherwise have little variance. The simulations displayed periodic dynamics, and parameters for the models were selected to ensure that oscillation periods were substantially less than the time series length (see Appendix A). Mean periods were 13.22 ± 1.62 years, 17.57 ± 9.61 years, and 22.74 ± 16.15 years for Simulations I–III, respectively (simulated time series by age class are shown in Fig. S1).

Empirical data

Data were obtained for two harvested fish species, Chinook salmon (*Oncorhynchus tshawytscha*) and striped bass (*Morone saxatilis*), both migratory, and a semelparous and an iteroparous species, respectively. Both datasets had annual samples over a 30-year period from 1986 to 2016. A time series length of 30 is often considered a reasonable minimum for successful use of EDM (Munch et al. 2018, 2020). Data were log transformed prior to analysis.

For Chinook salmon, age-specific escapement (abundance) indices for ages 2–5 were obtained from annual reports of the Klamath River Basin Fall Run[2]. Escapement was the number of fish that entered the Klamath River, survived migration up the river, and spawned in the wild or were taken into a hatchery for broodstock. Escapement sampling occurred during the spawning season, between September and December, typically peaking in November–December. Aging was accomplished through a combination of scale-aging and coded wire tag recoveries. Age-specific escapement indices were developed through cohort reconstruction[3].

For striped bass, age-specific indices of biomass per unit effort were obtained from the Connecticut Department of Energy and Environmental Protection (CTDEEP) Marine Fisheries Division's fisheries-independent Long Island Sound Trawl Survey. The survey is conducted in the spring (April to June) and fall (September to October) and follows a stratified random design. Striped bass were measured in all tows and converted to age-specific abundances using age-length keys by CTDEEP[4,5,6]. Striped bass age-length keys were developed by fitting the von Bertalanffy (1938) equation to length-at-age data[6]. Age classes spanned 1–14 years with a 15+ group. We used annual indices, taken from the publicly available stock assessment report. For each year, we multiplied the annual proportion of fish in each age class (corrected for selectivity at age) by the annual total biomass per tow to obtain an index of age-specific biomass per tow, assuming relatively constant mass at age. We also computed an age-specific abundance index per tow but obtained better forecast performance using biomass.

Gaussian process empirical dynamic modeling

We begin with a brief general introduction to EDM. Given a time series of a single variable, $x_0, \ldots, x_T$, EDM uses a flexible function approximation method to estimate a "delay-embedding map" of the form

$$\boldsymbol{x}_t = f\left(x_{t-\tau}, x_{t-2\tau}, \ldots, x_{t-E\tau}\right) \quad (1)$$

where E is the embedding dimension and τ is the time delay. Relevant covariates can also be included, when available. There are several ways to approximate this function from available data including the piecewise constant "Simplex" (Sugihara and May 1990), local weighted linear regression "S-Map" (Sugihara 1994), and splines (Ellner and Turchin 1995). In this study, we used GP regression (Poynor and Munch 2017) to approximate the unknown function f.

The GP is a continuous generalization of the multivariate normal distribution and represents a probability distribution for functions defined by a mean function μ and covariance function Σ. For age class a, our observations $y_{a,t}$ can be written as

$$y_{a,t} \sim N\left(f_a\left(\boldsymbol{x}_t\right), V_\epsilon\right) \quad (2)$$

where V_ϵ is process variance, and the dynamics can be written as

$$f_a \sim GP\left(\mu, \Sigma\right) \quad (3)$$

[2] Klamath River Technical Team. 2021. Klamath River Fall Chinook Salmon Age-Specific Escapement, River Harvest, and Run Size Estimates, 2020 Run. Pacific Fishery Management Council, Portland, OR. Available from https://www.pcouncil.org/documents/2021/02/klamath-river-fall-chinook-salmon-age-specific-escapement-river-harvest-and-run-size-estimates-2020-run-krtt-feb-15-2020.pdf/.

[3] Mohr, M.S. 2006. The cohort reconstruction model for Klamath River Fall Chinook Salmon. National Marine Fisheries Service Southwest Fisheries Science Center, Santa Cruz, CA. Unpublished report.

[4] Atlantic States Marine Fisheries Commission, A.S.B.T.C. 2016. Atlantic striped bass stock assessment update 2016. Available from http://www.asmfc.org/species/atlantic-striped-bass.

[5] Northeast Fisheries Science Center (NEFSC). 2019. 66th Northeast Regional Stock Assessment Workshop (66th SAW) Assessment Summary Report. US Department of Commerce, Northeast Fisheries Science Center Reference Document 19-01; 40 p. Available from https://doi.org/10.25923/g6g5-ed18.

[6] Connecticut Department of Energy and Environmental Protection. 2020. Job 5: Long Island Sound Trawl Survey (LISTS). Cruise results from the 2019 spring and fall surveys. Available from https://portal.ct.gov/-/media/DEEP/fishing/fisheries_management/2020-Long-Island-Sound-Trawl-Survey.pdf.

Thus, the state of each age class at time t is a function f_a of past observations (delay embedding vectors $\boldsymbol{x_t} = \{x_{t-\tau}, x_{t-2\tau}, \ldots, x_{t-E\tau}\}$), and the unknown map ($f_a$) is drawn from a GP informed by the data.

If modeling total abundance or a single age class, there will only be one f_a and the mean function μ is typically taken to be 0. In hierarchical GP time-delay embedding (Munch et al. 2017), there will be one f_a for each age class, and the mean function μ is shared across age classes. The shared μ is modeled as a second GP with a mean of 0 and covariance C.

$$\mu \sim GP(0, C) \tag{4}$$

A modification of the squared exponential covariance function is used for C

$$C(\mathbf{x}, \mathbf{x}') = \sigma^2 \left\{\delta_{a,a'} + \rho\left(1 - \delta_{a,a'}\right)\right\} \prod_{i=1}^{E} \exp \times \left\{-\phi_i \left(x_i - x_i'\right)^2\right\} \tag{5}$$

where σ^2 is the pointwise prior variance, ϕ_i are hyperparameters representing the inverse length scales, and $\delta_{a,a'}$ is the Kroenecker delta function (1 if age class $a = a'$ and 0 otherwise). The hyperparameter ρ is the dynamic correlation between the reconstructed maps for the different age classes (Munch et al. 2017; Rogers and Munch 2020). The functions (f_a, for $a = 1,\ldots,n$) are identical when $\rho = 1$ and independent when $\rho = 0$. Thus, ρ is a measure of dynamic similarity.

To select the embedding parameters E and τ for a given dataset, we evaluated model fit over a grid of values, constraining to feasible combinations ($E\tau$ cannot be greater than the smallest number of observations in an age class) and selected the combination that produced the highest out-of-sample R^2. We used leave-one-out R^2 as our measure of out-of-sample performance for embedding parameter selection and model performance evaluation, unless otherwise stated. For any given combination of E and τ, the model fitting procedure used automatic relevance determination to select relevant lags (MacKay and Neal 1994; Munch et al. 2017; Johnson et al. 2021). Models were fit using the package GPEDM (Munch and Rogers 2021) in R version 4.1.1 (R Core Team 2021).

Comparison of models with and without age structure

We compared the performance of hierarchical age-structured models and mixed-age models to models fit separately to each age class and to total abundance. The *hierarchical model* was fit to a dataset containing all age classes and was formulated as described above:

$$\begin{aligned} y_{a,t} &\sim N\left(f_a\left(\mathbf{x}_{a,t}\right), V_\epsilon\right) \\ f_a &\sim GP(\mu, \Sigma) \\ \mu &\sim GP(0, C) \end{aligned} \tag{6}$$

where $\mathbf{x}_{a,t}$ represents a delay embedding vector of lags from age class a. All data were scaled within age class to mean 0 and variance 1 prior to model fitting, and predictions were backtransformed to the original scale. This model produced a forecast of abundance for each age class, and from their sum, an overall prediction for total abundance (referred to in this paper as the "hierarchical aggregate").

Another way to use age structure to improve forecasts, the *mixed-age model*, combined delay embedding vectors of the focal age class (a) with lags of neighboring ages ($a - 1$, $a + 1$). The exceptions were $a = 1$, which only used lags of $a = 1$ and $a = 2$, and $a = a_{\max}$, which only used lags of $a = a_{\max} - 1$ and $a = a_{\max}$.

$$\begin{aligned} y_{a,t} &\sim N\left(f_a\left(\mathbf{x}_{a,t}, \mathbf{x}_{a-1,t}, \mathbf{x}_{a+1,t}\right), V_\epsilon\right) \\ f_a &\sim GP(0, \Sigma) \end{aligned} \tag{7}$$

In the *single-age model*, each age class a is modeled separately with no shared information from other age classes:

$$\begin{aligned} y_{a,t} &\sim N\left(f_a\left(\mathbf{x}_{a,t}\right), V_\epsilon\right) \\ f_a &\sim GP(0, \Sigma) \end{aligned} \tag{8}$$

The *total abundance* model was analogous, but predicted (and used lags of) total abundance summed across age classes. For each simulated dataset (I–III), we fit each model type to the first 30 years of data from 20 age classes. We fit models to 100 replicates for each simulation, which varied in recruitment noise. We evaluated potential values of E from 1 to 10 and τ from 1 to 3.

We also fit models to both empirical datasets, using all 30 years of data. For the empirical data, we evaluated potential values of E from 1 to 10 and τ from 1 to 3. To compare performance among the hierarchical, mixed-age, and single-age models, R^2 values were computed for each age class. To allow for a fairer comparison of the total abundance model (which only produces predictions for total abundance) with the age-structured approaches, we also computed an R^2 value for the hierarchical aggregate prediction from the hierarchical model. Because data were standardized within age classes prior to model fitting, and R^2 values represent predictability relative to the existing variance, the original scale of the data (e.g., relative versus absolute abundance) should not affect the results.

Time series length versus number of age classes

We next explored the tradeoff between number of age classes and length of time series. For each of the three simulated datasets, we fit hierarchical models to subsets with between 5 and 30 years of data (starting with the first year) and between 1 and 20 age classes, in all combinations. Rather than using leave-one-out, performance was evaluated on a testing set, which was the 10 years of data immediately following the training set. The value of E was set at the square root of the time series length, rounded to the nearest whole number (a general rule for short time series following Munch et al. (2020). We tried values of τ ranging from 1 to 3, settling on a value of $\tau = 1$, because it performed best on average across combinations. This process was repeated for 100 replicate simulation runs.

Dynamic correlation

If age-specific mortality rate is constant and the only source of nonlinearity is in recruitment, then all age classes should have identical dynamics, and in a hierarchical model, the estimated dynamic correlation (ρ) among age classes should be near 1. However, if age-specific mortality rate is not constant (e.g., dependent on system state), then ρ should be less than 1. This indicates that the dynamics for each age class deviate from one another, and less information is shared across age classes. Importantly, the dynamic correlation quantifies similarity in the intrinsic dynamics of age classes independently of whether their abundance is correlated through time (as would be measured by Pearson correlation). The hierarchical model estimates a single ρ across all age classes, but we can obtain finer resolution estimates of dynamic similarity by estimating ρ pairwise between age classes. For each simulated dataset and both empirical datasets, we fit hierarchical models to all pairs of age classes and estimated their pairwise ρ. The length of time series was held constant at 30 years, $\tau = 1$ was used, and a maximum embedding dimension E was set to 5, the nearest whole number that is the square root of time series length. The ρ values were compared to the temporal (Pearson) correlation (r) for each pair of age classes, in which the younger of the two age classes was lagged by the difference between age classes, e.g., $r(x_{a=1,t-1}, x_{a=2,t})$. For the simulated time series, we used data from all 100 years, and for each pair obtained mean pairwise ρ values from 100 replicates. For the empirical datasets, we obtained pairwise ρ values using all 30 years of data. We calculated the mean and standard deviation of pairwise ρ and r values across all pairs in each dataset.

Results

In the cases tested here, including age structure generally improved the predictive power of time series. The degree of improvement depended on the method used, the number of age classes, time series length, and the degree of dynamical correlation between age classes. We found that age structure information can improve the predictive power of time series, which could open the possibility of applying EDM to many more cases than currently feasible.

Comparison of models with and without age structure (simulations)

All models showed high predictive skill in all three simulated datasets (mean R^2 values for each age class were all > 0.8; Fig. 1, see Fig. S2 for variability across replicate simulations). Median R^2 values across age classes were higher for the two models incorporating age structure than for the single-age model in all simulations. The hierarchical and mixed-age models had similar median performance, but mixed-age model performance was more variable among age classes in Simulations II and III. The single-age model showed the lowest predictive skill and the greatest variability in performance among age classes in Simulations I and III. See Fig. S1 for fits of the hierarchical model to the simulated datasets.

In terms of predicting total abundance, the total abundance model had equivalent predictive skill as the hierarchical aggregate prediction (Fig. 2). In the hierarchical model, performance across all age-specific predictions was also comparable to hierarchical aggregate prediction (Fig. 2). Predictive skill was more variable for the hierarchical aggregate than for the total abundance model or across age-specific predictions, likely due to variations in how age-specific prediction errors combined.

Comparison of models with and without age structure (empirical data)

The mixed-age model had higher median performance than the single-age model and the hierarchical model for both species (Fig. 1). However, there was high variability in performance for individual age classes. In Chinook salmon, for instance, the mixed-age model worked very well for age classes 3 and 4 ($R^2 > 0.75$), but similarly to the hierarchical and single-age model for ages 2 and 5. (Supplementary Fig. S2). Most Chinook salmon in this population return to spawn at ages 3 and 4; so these classes are the most abundant. The mixed-age approach resulted in a higher mean R^2 for striped bass (mean = 0.663; SD = 0.231) than for Chinook salmon (mean = 0.541; SD = 0.39); however, the hierarchical and single-age models had better performance for Chinook salmon (hierarchical mean = 0.297 and SD = 0.119; single-age mean = 0.271 and SD = 0.152) than striped bass (hierarchical mean = 0.158 and SD = 0.158; single-age mean = 0.206 and SD = 0.326).

For both species, the hierarchical aggregate outperformed the total abundance model at predicting total abundance (Chinook salmon hierarchical aggregate $R^2 = 0.76$, total abundance $R^2 = 0.54$; striped bass: hierarchical aggregate $R^2 = 0.56$, total abundance $R^2 = 0.50$). Observed and predicted time series for each age class (from the hierarchical model) and total abundance (from the total abundance model and the hierarchical aggregate) are shown in Fig. 3. For both species, the standard deviation of the model predictions was greater for the hierarchical aggregate than for the total abundance model (Fig. 3). The best hierarchical model for Chinook salmon had $E = 6$ and $\tau = 1$, and for striped bass had $E = 5$ and $\tau = 1$ (Supplementary Fig. S3).

Time series length versus number of age classes

We compared the performance of models using all combinations of 5–30 years of data and 1–20 age classes and found that the marginal benefit of each additional age class was greatest for short time series (Figs. 4 and S4). For time series longer than 10–13 years, including more than one age class did not substantially increase forecast skill. However, for shorter time series, increasing the number of age classes typically improved mean R^2. For example, with 8 years of data, mean R^2 increased from 0.02 ± 0.62 with one age class to 0.80 ± 0.16 using four age classes in Simulation III.

Dynamic correlation

Mean pairwise dynamic correlation was lower than the mean Pearson correlation in simulated datasets (Table 2; see

Fig. 1. Prediction skill (measured as out-of-sample R^2) for Simulations I–III and two empirical datasets (KRFC = Klamath River Fall Chinook salmon, SB = Long Island Sound striped bass) using three different modeling approaches. The gray points represent the R^2 for each individual age class (for simulations, the mean R^2 across 100 replicates). Boxplots show the distribution of R^2 values across age classes for each approach. Note: Plots for Simulations II and III have vertical axis breaks.

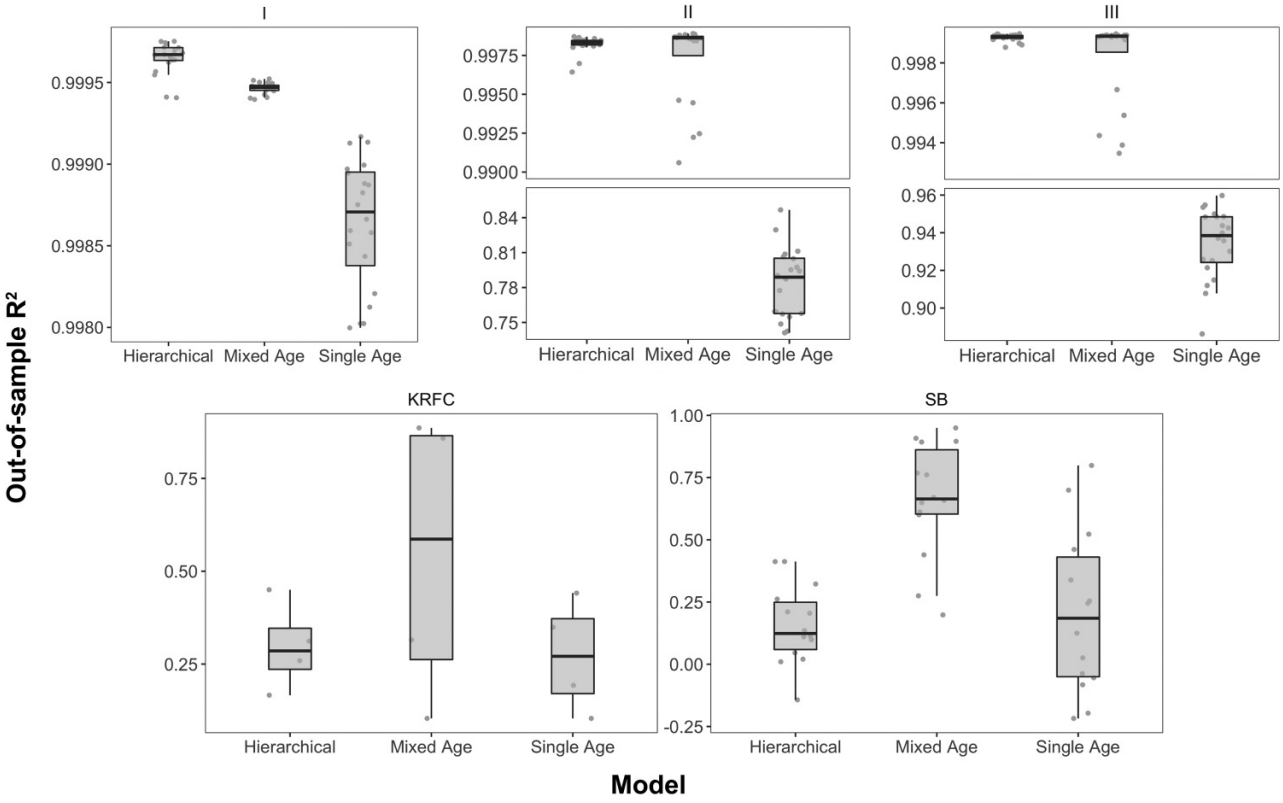


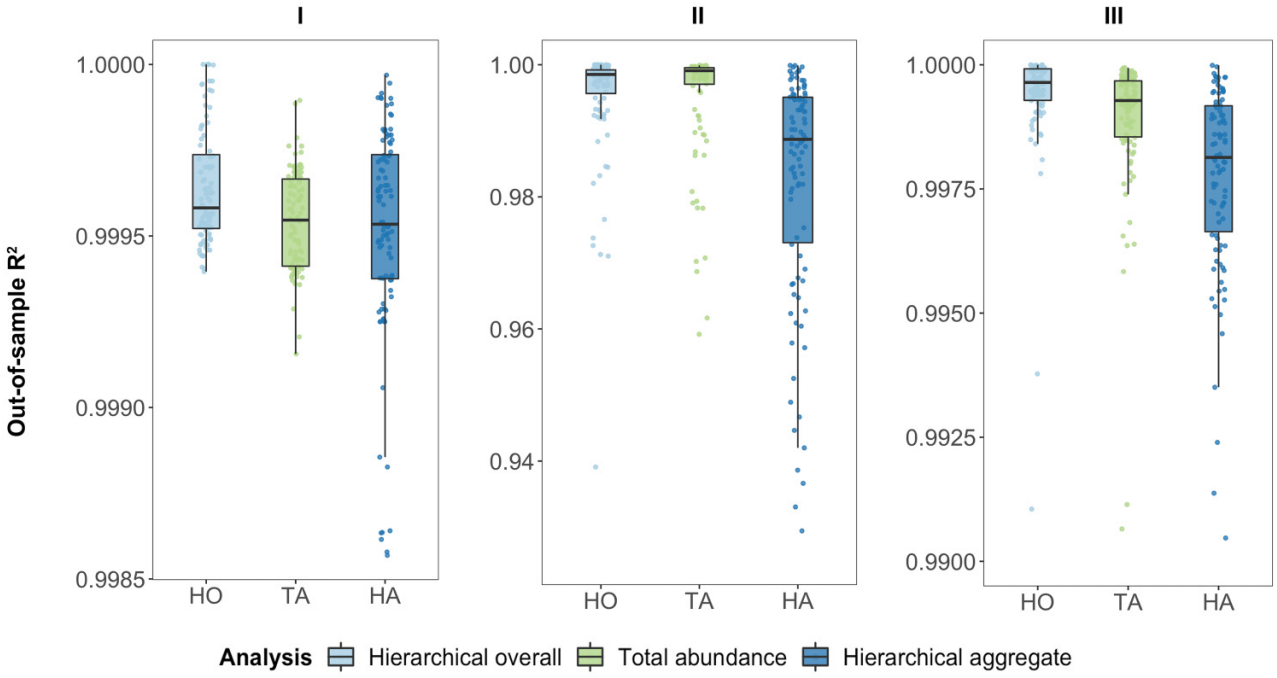
Fig. S5 for individual pairs). In contrast, for the empirical datasets, dynamic correlation was higher than Pearson correlation (Table 2 and Fig. S6). For the simulated datasets, dynamic correlation values decreased as simulation complexity increased, and the discrepancy between mean pairwise Pearson correlation and mean pairwise dynamic correlation increased. In Simulation I, pairwise dynamic correlations were all very close to 1, whereas in Simulation III, pairwise dynamic correlations were more varied.

Discussion

For generating accurate predictions, EDM has many advantages over classic mechanistic models. However, the requirement for long time series has been an impediment to its broad application in fisheries management and conservation biology. Age-structured data may be an untapped resource for improving predictive skill of empirical dynamic models, enabling researchers to make use of shorter time series for forecasting population abundances. We found that, in most cases, using age structure in a hierarchical or mixed-age modeling approach improved predictive skill of models, both in empirical and simulated data. Our results are consistent with theory, which suggests that lags of additional age classes can stand in for unknown state variables (Munch et al. 2017) analogously to incorporating spatial information (Johnson et al. 2021) or information from additional species (Ye and Sugihara 2016).

The hierarchical approach outperformed the mixed-age approach and single-age approach when the dynamics of the different age classes were relatively similar (dynamic correlation values are high, as in Simulation I). When some (but not all) age classes have similar dynamics, we might expect a mixed-age approach to perform best. If dynamics are very dissimilar across age classes, then single-age approaches may perform better than hierarchical or mixed-age approaches, provided time series are sufficiently long. In our simulated datasets with 30 timepoints, which all had relatively high dynamic correlation values, hierarchical and mixed-age models had greater forecast performance than single-age models. The difference in performance between single-age and age-structured models was higher for the more complex simulations (II and III), although these differences were relatively small, and all models tested had relatively high R^2 values. This lack of dramatic differences in model performance with increasing model complexity contrasts with our initial expectations. However, in retrospect, this result is entirely consistent with theory on time-delay embedding: since they are part of the same dynamical system, the attractors for different age classes cannot diverge too dramatically; so information from multiple ages can improve performance even in more complex systems with variable natural mortality. This result is

Fig. 2. Comparison of model performancefrom Simulations I–III using the hierarchical model and total abundance model. Out-of-sample R^2 are calculated from the total abundance model ("TA"), the summed predictions of the hierarchical model (hierarchical aggregate, "HA"), and all age-specific predictions for the hierarchical model (hierarchical overall, "HO"). The points represent values from each simulation run. Note: The hierarchical aggregate category contained four data points below the *y*-axis minimum limit in panel II and nine data points below the minimum in panel III.



consistent with other EDM studies, which found that concatenating delay vectors from different species in the same system increases performance (Hsieh et al. 2008).

More revealing are the results from the empirical data. Dynamic correlations were lower and more heterogeneous among age classes, and mixed-age models tended to perform best. There are many reasons why age-specific mortality may not be constant in real systems, resulting in different dynamical behavior among age classes of the same species. Processes such as recruitment, growth, ecosystem interactions, and mortality are often time-varying and stage-specific (Hutchings 1993; Bjørnstad et al. 2004; Secor 2007; Botsford et al. 2011; Rouyer et al. 2011; White et al. 2014). If mortality is constant and the dominant nonlinearities in a dataset come from recruitment variability, we might expect the hierarchical approach to work best, which treats age classes as temporal replicates. In systems with more complicated dynamics, it might be better to include mixed ages. In a practical setting, where system complexity is unknown, trying multiple approaches may be useful. The strength of both the hierarchical and mixed-age approaches are that age structure allows researchers to use shorter time series in nonlinear forecasting. Hierarchical models and use of the dynamic correlation might be useful not only for developing forecasting models but could be used as a test of whether the constant mortality assumption is appropriate for a given species, as is often assumed in stock assessments.

The relationship between dynamic correlation between ages and forecast skill was evident in the empirical datasets. Striped bass had lower forecast skill and lower dynamic correlation values than Chinook salmon. Lower forecast skill could be related to insufficient time series length relative to generation time. Previous results indicate that time series 10× the maturation age are often required to obtain high-quality forecasts with EDM (Munch et al. 2018). Consistent with this general rule, the 30-year time series was 10–15× the maturation age of Chinook salmon[7] (2–3 years, O'Farrell et al. 2013), and EDM captured more than 50% of the variation. In contrast, maturation age for female striped bass is ~5 years (Dew 1988; Berlinsky and Specker 1991; Berlinsky et al. 1995); so the 30-year time series was only 6× the maturation age. Striped bass are also iteroparous and attain a maximum age of >30 years (Secor 2000*a*), whereas Chinook salmon are semelparous with a maximum age of 9 years, though most reproduce at ages 3 or 4 (O'Farrell et al. 2013). Semelparity lends itself to simple time-delay models, while iteroparity necessarily involves additional delay-embedding vectors, unless it is assumed (somewhat unrealistically) that all age classes have a constant survival and fecundity.

Sampling of multiple subpopulations might also contribute to more divergent dynamics in striped bass. While

[7] Gough, S., and Williamson, S. 2012. Fall Chinook Salmon run characteristics and escapement for the main-stem Klamath River, 2001-2010. U.S. Fish and Wildlife Service Arcata Fish and WIldlife Office, Arcata, CA. Available from https://kbifrm.psmfc.org/wp-content/uploads/2017/01/Gough-et-al_2010_0207_Fall-Chinook-Run-Characteristics-and-Escapement.pdf.

Fig. 3. Observed and predicted time series for empirical data: (*a*) Striped bass biomass per unit effort (kg), and (*b*) Chinook salmon escapement. Black points are observed data, the line is the model prediction (mean value), and the shaded area is the standard deviation for the prediction. Panels with "age" in the title show the leave-one-out predictions for each age class from the hierarchical model for that species. The "aggregate" panel shows observed and predicted total abundance, where predictions are the sum of predicted abundance for each age class from the hierarchical model. The "total abundance" panel shows observed and predicted total abundance from the non-age-structured total abundance model.

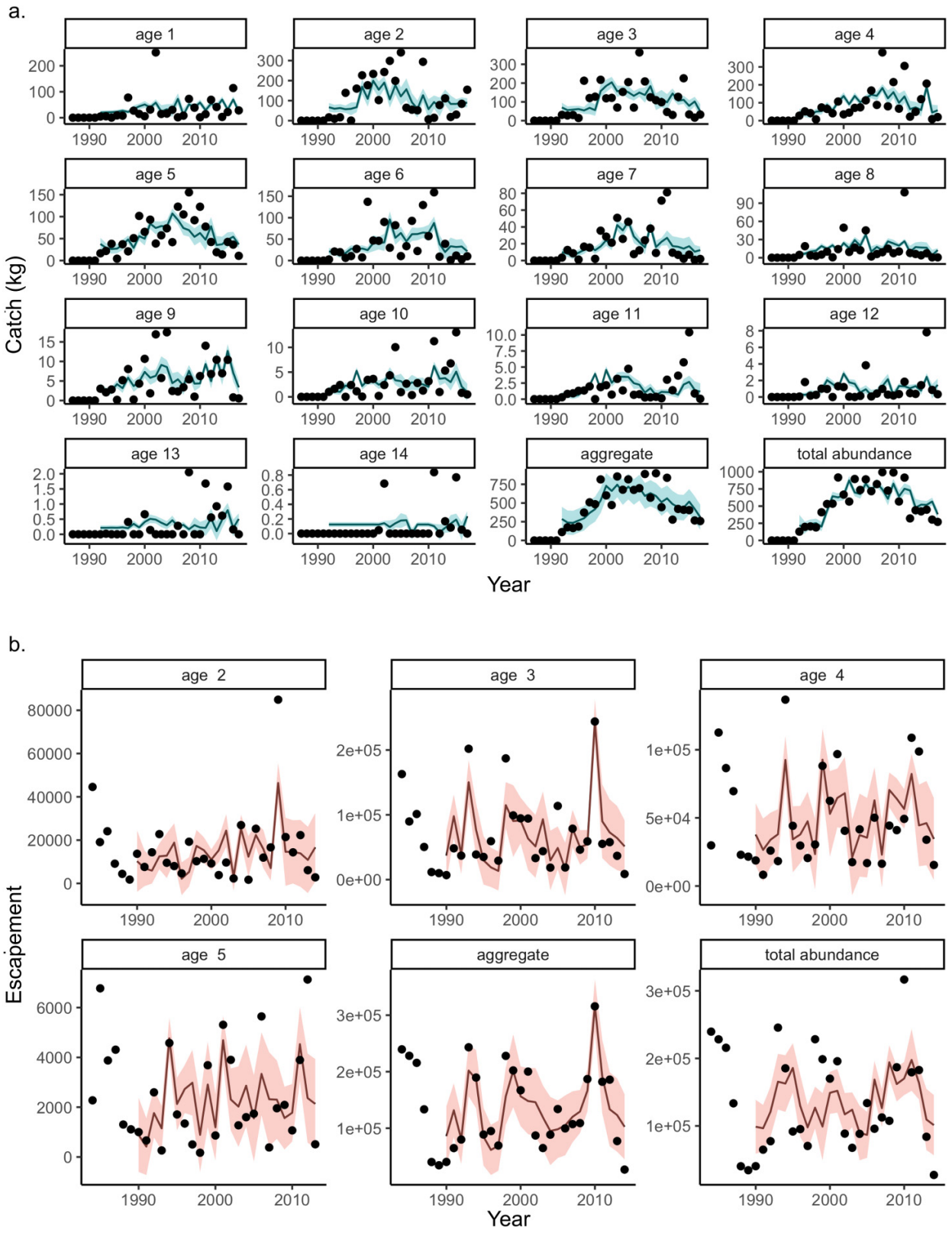


Fig. 4. Prediction skill (measured as out-of-sample R^2) for each combination of time series length and number of age classes from Simulations I–III (means over 100 replicate simulations). Age classes 1–20 and time series length 5–15 years are shown. For time series length 5–30 years, see Supplementary Fig. S4.

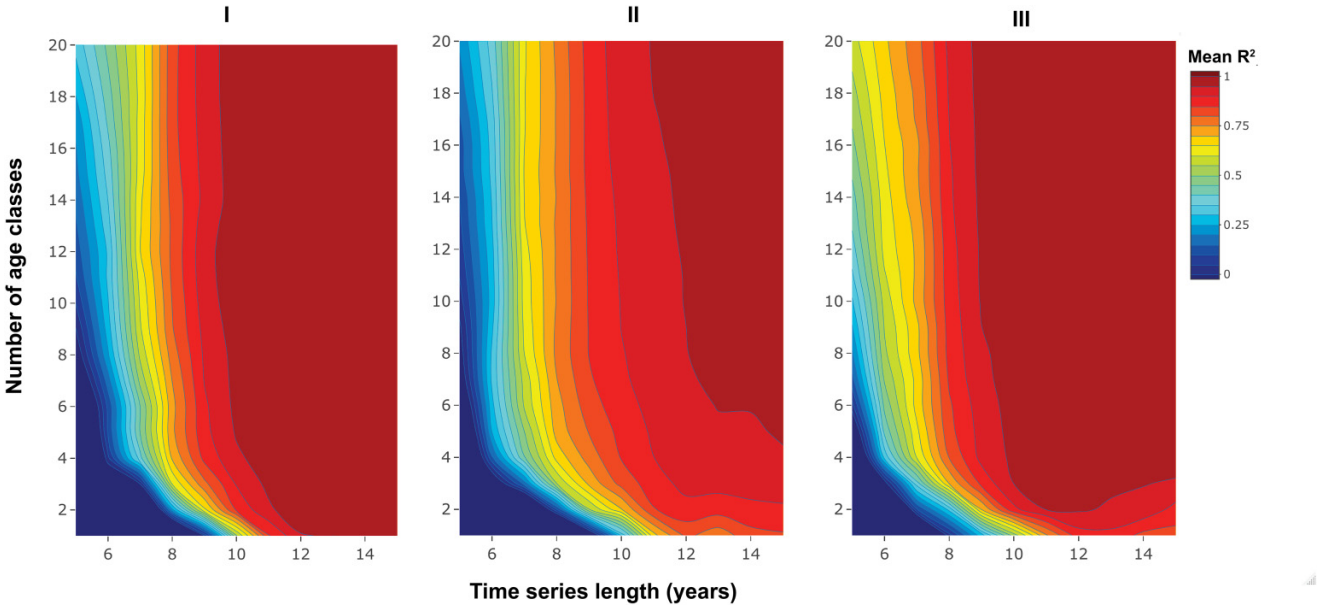


Table 2. Mean pairwise lagged Pearson correlation and mean pairwise dynamic correlation.

		Simulation I	Simulation II	Simulation III	Chinook salmon	Striped bass
Lagged Pearson correlation	Mean	1.00	0.970	0.993	0.623	0.236
	SD	0.00	0.0255	0.0120	0.150	0.256
Dynamic correlation	Mean	0.991	0.960	0.971	0.787	0.654
	SD	0.00688	0.0388	0.0294	0.131	0.156

Note: These values represent the mean of all pairwise combinations of age classes. For the simulated datasets, pairwise values were the averages across all simulation replicates. For all pairwise values, see Supplementary Figs. S5 and S6.

both Chinook salmon and striped bass are migratory species, striped bass in the Long Island Sound probably represents a mixed stock of individuals originating from across the east coast of the United States and Canada (Waldman and Fabrizio 1994; Wirgin et al. 1995; Callihan et al. 2015; Secor et al. 2020). Ocean salmon fisheries are also a mixed stock (Satterthwaite et al. 2015), but the Chinook salmon escapement index is composed of individuals intercepted within their natal river (Gough and Williamson 2012) and thus is likely more internally homogenous relative to striped bass. Additionally, whereas Chinook salmon sampling occurred during the spawning season after which nearly all natural and fishing mortality would have taken place, striped bass sampling occurred over a longer time period spanning two seasons, and thus natural and fishing mortality may have fluctuated over the course of data collection. Both species show intraspecific diversity in migration strategy (Secor et al. 2020; Apgar et al. 2021), but resident striped bass are unlikely to be intercepted by the trawl survey. Striped bass have suffered from recent age truncation and Chinook salmon from a decline in body size (Secor 2000*b*; Satterthwaite and Carlson 2015; Oke et al. 2020). These changes might potentially disrupt the dynamics between age classes, by shifting the age at which individuals are susceptible to size-dependent mortality and energetic constraints (de Roos et al. 2006; White et al. 2014; Oke et al. 2020).

Despite the advantages of including age structure in EDM models, it is not a direct substitute for additional years of data. Time series of sufficient length are still needed to cover the characteristic return time of the system. We found that increasing time series length went further toward improving forecast skill than adding additional age classes, underscoring the importance of collecting long time series. Leveraging age structure can improve forecasts, but the improvement is greatest for very short time series. However, even in this case, time series need to span any dominant periodicity in the dynamics. Provided data are sufficient, however, even models for short-lived species with few age classes, such as Chinook salmon, can benefit from incorporating age structure.

Conclusion

Forecasting populations is a challenging endeavor, but one that is critical to designing management and conservation targets to meet future challenges. The hierarchical approaches based on age structure demonstrated here help to overcome one of the major challenges with how EDM is typically applied, which is the demand for time series that are

much longer than those typically available. There is no one-size-fits-all approach to using hierarchical nonlinear models. For the best possible age-structured forecast, age classes must be similar enough so that their dynamics are mutually informative, but different enough to provide new information. This approach, when used judiciously, will expand the predictive power of available time series to a much larger set of cases.

Acknowledgements

We thank Michael O'Farrell and William Satterthwaite (NMFS SWFSC) for providing access to the Chinook salmon escapement data, and Kurt Gottschall (CTDEEP) for access to the striped bass data.

Article information

History dates

Received: 13 September 2022
Accepted: 9 January 2023
Accepted manuscript online: 18 January 2023
Version of record online: 22 February 2023

Copyright

Data availability

The simulated datasets, empirical datasets, generating code for the simulated datasets, and all analysis code are available in the hierAgeStrucEDM repository [DOI: https://github.com/tdolz/hierAgeStrucEDM]. Empirical datasets were compiled from publicly available records (see the "References" section), with help from NMFS SWFSC and CTDEEP (see the "Acknowledgements" section.

Author information

Author ORCIDs

Tara E. Dolan https://orcid.org/0000-0003-0937-0342
Eric P. Palkovacs https://orcid.org/0000-0002-5496-7263
Tanya L. Rogers https://orcid.org/0000-0003-1253-9903

Author notes

Present address for Tara E. Dolan is Massachusetts Division of Marine Fisheries, Cat Cove Marine Laboratory, 92 Fort Avenue, Salem, MA 01970, USA.

Author contributions

Conceptualization: TED, EPP, SBM
Data curation: TED
Formal analysis: TED
Funding acquisition: EPP, SBM
Investigation: TED
Methodology: TLR, SBM
Project administration: EPP, SBM
Resources: EPP, SBM
Software: TED, TLR, SBM
Supervision: EPP, SBM
Validation: TLR, SBM
Visualization: TED
Writing – original draft: TED
Writing – review & editing: TED, EPP, TLR, SBM

Competing interests

The authors declare there are no competing interests.

Funding information

This research was supported by the NOAA Quantitative Ecology and Socioeconomics Training (QUEST) Program and the Cooperative Institute for Marine, Earth, and Atmospheric Systems (CIMEAS).

Supplementary material

Supplementary data are available with the article at https://doi.org/10.1139/cjfas-2022-0219.

References

Apgar, T.M., Merz, J.E., Martin, B.T., and Palkovacs, E.P. 2021. Alternative migratory strategies are widespread in subyearling Chinook salmon. Ecol. Freshwater Fish, **30**(1): 125–139. doi:10.1111/eff.12570.

Berlinsky, D.L., and Specker, J.L. 1991. Changes in gonadal hormones during oocyte development in the striped bass, Morone saxatilis. Fish Physiol. Biochem. **9**(1): 51–62. doi:10.1007/BF01987611. PMID: 24214609.

Berlinsky, D.L., Fabrizio, M.C., O'Brien, J.F., and Specker, J.L. 1995. Age-at-maturity estimates for Atlantic coast female striped bass. Trans. Am. Fish. Soc. **124**(2): 207–215. doi:10.1577/1548-8659.

Bjørnstad, O.N., Nisbet, R.M., and Fromentin, J. 2004. Trends and cohort resonant effects in age-structured populations. J. Anim. Ecol. **73**(6): 1157–1167. doi:10.1111/j.0021-8790.

Botsford, L.W., Holland, M.D., Samhouri, J.F., White, J.W., and Hastings, A. 2011. Importance of age structure in models of the response of upper trophic levels to fishing and climate change. ICES J. Mar. Sci. **68**(6): 1270–1283. doi:10.1093/icesjms/fsr042.

Brias, A., and Munch, S.B. 2021. Ecosystem based multi-species management using empirical dynamic programming. Ecol. Model. **441**: 109423. doi:10.1016/j.ecolmodel.2020.109423.

Brodie, S., Abrahms, B., Bograd, S.J., Carroll, G., Hazen, E.L., Muhling, B.A., et al. 2021. Exploring timescales of predictability in species distributions. Ecography, **44**: 1–13. doi:10.1111/ecog.05504.

Callihan, J.L., Harris, J.E., and Hightower, J.E. 2015. Coastal migration and homing of Roanoke River striped bass. Mar. Coastal Fish. **7**(1): 301–315. doi:10.1080/19425120.2015.1057309.

Chang, C.-W., Ushio, M., and Hsieh, C. 2017. Empirical dynamic modeling for beginners. Ecol. Res. **32**(6): 785–796. doi:10.1007/s11284-017-1469-9.

Clark, J.S., Carpenter, S.R., Barber, M., Collins, S., Dobson, A., Foley, J.A., et al. 2001. Ecological forecasts: an emerging imperative. Science, **293**(5530): 657–660. doi:10.1126/science.293.5530.657.

Dakos, V., Glaser, S.M., Hsieh, C., and Sugihara, G. 2017. Elevated nonlinearity as an indicator of shifts in the dynamics of populations under stress. J. R. Soc. Interface, **14**(128): 20160845. doi:10.1098/rsif.2016.0845. PMID: 28250096.

de Roos, A.M., Boukal, D.S., and Persson, L. 2006. Evolutionary regime shifts in age and size at maturation of exploited fish stocks. Proc. R. Soc. B Biol. Sci. **273**(1596): 1873–1880. doi:10.1098/rspb.2006.3518.

DeAngelis, D.L., and Yurek, S. 2015. Equation-free modeling unravels the behavior of complex ecological systems. Proc. Natl. Acad. Sci. U.S.A. **112**(13): 3856–3857. doi:10.1073/pnas.1503154112. PMID: 25829536.

Dew, C.B. 1988. Biological characteristics of commercially caught Hudson River striped bass, 1973–1975. North Am. J. Fish. Manage. **8**(1): 75–83. doi:10.1577/1548-8675.

Deyle, E., Schueller, A.M., Ye, H., Pao, G.M., and Sugihara, G. 2018. Ecosystem-based forecasts of recruitment in two menhaden species. Fish Fish. **19**(5): 769–781. doi:10.1111/faf.12287.

Deyle, E.R., Fogarty, M., Hsieh, C., Kaufman, L., MacCall, A.D., Munch, S.B., et al. 2013. Predicting climate effects on Pacific sardine. Proc. Natl. Acad. Sci. U.S.A. **110**(16): 6430–6435. doi:10.1073/pnas.1215506110.

Deyle, E.R., May, R.M., Munch, S.B., and Sugihara, G. 2016. Tracking and forecasting ecosystem interactions in real time. Proc. R. Soc. B, **283**(1822): 20152258. doi:10.1098/rspb.2015.2258.

Dietze, M.C., Fox, A., Beck-Johnson, L.M., Betancourt, J.L., Hooten, M.B., Jarnevich, C.S., et al. 2018. Iterative near-term ecological forecasting: needs, opportunities, and challenges. Proc. Natl. Acad. Sci. U.S.A. **115**(7): 1424–1432. doi:10.1073/pnas.1710231115. PMID: 29382745.

Ellner, S., and Turchin, P. 1995. Chaos in a noisy world: new methods and evidence from time-series analysis. Am. Nat. **145**(3): 343–375. doi:10.1086/285744.

Glaser, S.M., Fogarty, M.J., Liu, H., Altman, I., Hsieh, C.-H., Kaufman, L., et al. 2014a. Complex dynamics may limit prediction in marine fisheries. Fish Fish. **15**(4): 616–633. doi:10.1111/faf.12037.

Glaser, S.M., Ye, H., and Sugihara, G. 2014b. A nonlinear, low data requirement model for producing spatially explicit fishery forecasts. Fish. Oceanogr. **23**(1): 45–53. doi:10.1111/fog.12042.

Gulland, J.A. 1965. Survival of youngest stages of fish and its relation to year-class strength. Spec. Publ. Int. Comm. NW Atl. Fish. **6**: 363–371.

Hobday, A.J., Spillman, C.M., Paige Eveson, J., and Hartog, J.R. 2016. Seasonal forecasting for decision support in marine fisheries and aquaculture. Fish. Oceanogr. **25**(S1): 45–56. doi:10.1111/fog.12083.

Hsieh, C., Anderson, C., and Sugihara, G. 2008. Extending nonlinear analysis to short ecological time series. Am. Nat. **171**(1): 71–80. doi:10.1086/524202. PMID: 18171152.

Hutchings, J.A. 1993. Adaptive life histories effected by age-specific survival and growth rate. Ecology, **74**(3): 673–684. doi:10.2307/1940795.

Jacox, M.G., Alexander, M.A., Siedlecki, S., Chen, K., Kwon, Y.-O., Brodie, S., et al. 2020. Seasonal-to-interannual prediction of North American coastal marine ecosystems: forecast methods, mechanisms of predictability, and priority developments. Prog. Oceanogr. **183**: 102307. doi:10.1016/j.pocean.2020.102307.

Johnson, B., Gomez, M., and Munch, S.B. 2021. Leveraging spatial information to forecast nonlinear ecological dynamics. Meth. Ecol. Evol. **12**(2): 266–279. doi:10.1111/2041-210X.13511.

Köster, F.W., and Möllmann, C. 2000. Trophodynamic control by clupeid predators on recruitment success in Baltic cod? ICES J. Mar. Sci. **57**(2): 310–323. doi:10.1006/jmsc.1999.0528.

Kuriyama, P.T., Sugihara, G., Thompson, A.R., and Semmens, B.X. 2020. Identification of shared spatial dynamics in temperature, salinity, and ichthyoplankton community diversity in the California current system with empirical dynamic modeling. Front. Mar. Sci. **870**. doi:10.3389/fmars.2020.557940.

Lasky, J.R., Hooten, M.B., and Adler, P.B. 2020. What processes must we understand to forecast regional-scale population dynamics? Proc. R. Soc. B Biol. Sci. **287**(1940): 20202219. doi:10.1098/rspb.2020.2219.

Liu, H., Fogarty, M.J., Glaser, S.M., Altman, I., Hsieh, C., Kaufman, L., et al. 2012. Nonlinear dynamic features and co-predictability of the Georges Bank fish community. Mar. Ecol. Prog. Ser. **464**: 195–207. doi:10.3354/meps09868.

Ma, H., Leng, S., Aihara, K., Lin, W., and Chen, L. 2018. Randomly distributed embedding making short-term high-dimensional data predictable. Proc. Natl. Acad. Sci. U.S.A. **115**(43): E9994–E10002. doi:10.1073/pnas.1802987115. PMID: 30297422.

MacKay, D.J., and Neal, R.M. 1994. Automatic relevance determination for neural networks. *In* Technical report in preparation. Cambridge University, Cambridge.

McNamara, D.E., Cortale, N., Edwards, C., Eynaud, Y., and Sandin, S.A. 2019. Insights into coral reef benthic dynamics from nonlinear spatial forecasting. J. R. Soc. Interface, **16**(153): 20190047. doi:10.1098/rsif.2019.0047. PMID: 30966951.

Melnychuk, M.C., Kurota, H., Mace, P.M., Pons, M., Minto, C., Osio, G.C., et al. 2021. Identifying management actions that promote sustainable fisheries. Nat. Sustain. **4**(5): 440–449. doi:10.1038/s41893-020-00668-1.

Miller, S., Rassweiler, A., Dee, L., Kleisner, K.M., Mangin, T., Oliveros-Ramos, R., et al. 2019. Optimal harvest responses to environmental forecasts depend on resource knowledge and how it can be used. Can. J. Fish. Aquat. Sci. **76**(9): 1495–1502. doi:10.1139/cjfas-2018-0283. PMID: 33353994.

Munch, S.B., and Rogers, T.L. 2021. GPEDM: Gaussian process regression for empirical dynamic modeling. R package version 0.0.0.9005.

Munch, S.B., Brias, A., Sugihara, G., and Rogers, T.L. 2020. Frequently asked questions about nonlinear dynamics and empirical dynamic modelling. ICES J. Mar. Sci. **77**(4): 1463–1479. doi:10.1093/icesjms/fsz209.

Munch, S.B., Giron-Nava, A., and Sugihara, G. 2018. Nonlinear dynamics and noise in fisheries recruitment: a global meta-analysis. Fish Fish. **19**(6): 964–973. doi:10.1111/faf.12304.

Munch, S.B., Poynor, V., and Arriaza, J.L. 2017. Circumventing structural uncertainty: a Bayesian perspective on nonlinear forecasting for ecology. Ecol. Complexity, **32**: 134–143. doi:10.1016/j.ecocom.2016.08.006.

O'Farrell, M.R., Mohr, M.S., Palmer-Zwahlen, M.L., and Grover, A.M. 2013. The Sacremento Index (SI). *In* NOAA Technical Memorandum. National Marine Fisheries Service Southwest Fisheries Science Center, San Diego, CA.

Oke, K.B., Cunningham, C.J., Westley, P.A.H., Baskett, M.L., Carlson, S.M., Clark, J., et al. 2020. Recent declines in salmon body size impact ecosystems and fisheries. Nat. Commun. **11**(1): 1–13. doi:10.1038/s41467-020-17726-zPMID: 31911652.

Patterson, K., Cook, R., Darby, C., Gavaris, S., Kell, L., Lewy, P., et al. 2001. Estimating uncertainty in fish stock assessment and forecasting. Fish Fish. **2**(2): 125–157. doi:10.1046/j.1467-2960.2001.00042.x.

Payne, M.R., Hobday, A.J., MacKenzie, B.R., Tommasi, D., Dempsey, D.P., Fässler, S.M.M., et al. 2017. Lessons from the first generation of marine ecological forecast products. Front. Mar. Sci. **4**: 289. doi:10.3389/fmars.2017.00289.

Pennekamp, F., Iles, A.C., Garland, J., Brennan, G., Brose, U., Gaedke, U., et al. 2019. The intrinsic predictability of ecological time series and its potential to guide forecasting. Ecol. Monogr. **89**(2): e01359. doi:10.1002/ecm.1359.

Perretti, C.T., Munch, S.B., and Sugihara, G. 2013. Model-free forecasting outperforms the correct mechanistic model for simulated and experimental data. Proc. Natl. Acad. Sci. U.S.A. **110**(13): 5253–5257. doi:10.1073/pnas.1216076110.

Poynor, V., and Munch, S. 2017. Combining functional data with hierarchical Gaussian process models. Environ. Ecol. Stat. **24**(2): 175–199. doi:10.1007/s10651-017-0366-2.

R Core Team. 2021. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from www.R-project.org.

Ricker, W.E. 1954. Stock and recruitment. J. Fish. Board Can. **11**(5): 559–623. doi:10.1139/f54-039.

Rogers, T.L., and Munch, S.B. 2020. Hidden similarities in the dynamics of a weakly synchronous marine metapopulation. Proc. Natl. Acad. Sci. U.S.A. **117**(1): 479–485. doi:10.1073/pnas.1910964117. PMID: 31871191.

Rouyer, T., Ottersen, G., Durant, J.M., Hidalgo, M., Hjermann, D.Ø., Persson, J., et al. 2011. Shifting dynamic forces in fish stock fluctuations triggered by age truncation? Global Change Biol. **17**(10): 3046–3057. doi:10.1111/j.1365-2486.2011.02443.x.

Satterthwaite, W.H., and Carlson, S.M. 2015. Weakening portfolio effect strength in a hatchery-supplemented Chinook salmon population complex. Can. J. Fish. Aquat. Sci. **72**: 1860–1875. doi:10.1139/cjfas-2015-0169.

Satterthwaite, W.H., Ciancio, J., Crandall, E., Palmer-Zwahlen, M.L., Grover, A.M., O'Farrell, M.R., et al. 2015. Stock composition and ocean spatial distribution inference from California recreational Chinook salmon fisheries using genetic stock identification. Fish. Res. **170**: 166–178. doi:10.1016/j.fishres.2015.06.001.

Secor, D. 2000a. Longevity and resilience of Chesapeake Bay striped bass. ICES J. Mar. Sci. **57**(4): 808–815. doi:10.1006/jmsc.2000.0560.

Secor, D.H. 2000b. Spawning in the nick of time? Effect of adult demographics on spawning behaviour and recruitment in Chesapeake Bay striped bass. ICES J. Mar. Sci. **57**(2): 403–411. doi:10.1006/jmsc.1999.0520.

Secor, D.H. 2007. The year-class phenomenon and the storage effect in marine fishes. J. Sea Res. **57**(2–3): 91–103. doi:10.1016/j.seares.2006.09.004.

Secor, D.H., O'Brien, M.H.P., Gahagan, B.I., Watterson, J.C., and Fox, D.A. 2020. Differential migration in Chesapeake Bay striped bass. PLoS One, **15**(5): e0233103. doi:10.1371/journal.pone.0233103.

Shi, J.Q., Murray-Smith, R., and Titterington, D.M. 2005. Hierarchical Gaussian process mixtures for regression. Stat. Comput. **15**(1): 31–41. doi:10.1007/s11222-005-4787-7.

Subbey, S., Devine, J.A., Schaarschmidt, U., and Nash, R.D.M. 2014. Modelling and forecasting stock–recruitment: current and future perspectives. ICES J. Mar. Sci. **71**(8): 2307–2322. doi:10.1093/icesjms/fsu148.

Sugihara, G. 1994. Nonlinear forecasting for the classification of natural time series. Philos. Trans.: Phys. Sci. Eng. **348**(1688): 477–495.

Sugihara, G., and May, R.M. 1990. Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. Nature, **344**(6268): 734–741. doi:10.1038/344734a0. PMID: 2330029.

Sugihara, G., Beddington, J., Hsieh, C., Deyle, E., Fogarty, M., Glaser, S.M., et al. 2011. Are exploited fish populations stable? Proc. Natl. Acad. Sci. U.S.A., **108**(48): E1224–E1225. doi:10.1073/pnas.1112033108.

Takens, F. 1981. Detecting strange attractors in turbulence. Dynamical Systems and Turbulence.,R and, L.S. Springer, New York. pp. 336–381

Thorson, J.T. 2019. Forecast skill for predicting distribution shifts: a retrospective experiment for marine fishes in the Eastern Bering Sea. Fish Fish. **20**(1): 159–173. doi:10.1111/faf.12330.

Tommasi, D., Stock, C.A., Hobday, A.J., Methot, R., Kaplan, I.C., Eveson, J.P., et al. 2017. Managing living marine resources in a dynamic environment: the role of seasonal to decadal climate forecasts. Prog. Oceanogr. **152**: 15–49. doi:10.1016/j.pocean.2016.12.011.

Van der Veer, H.W., Berghahn, R., Miller, J.M., and Rijnsdorp, A.D. 2000. Recruitment in flatfish, with special emphasis on North Atlantic species: progress made by the Flatfish Symposia. ICES J. Mar. Sci. **57**(2): 202–215. doi:10.1006/jmsc.1999.0523.

von Bertalanffy, L. 1938. A quantitative theory of organic growth (inquiries on Growth Laws. II). Hum. Biol. **10**(2): 181–213. Available from https://www.jstor.org/stable/41447359.

Waldman, J.R., and Fabrizio, M.C. 1994. Problems of stock definition in estimating relative contributions of Atlantic striped bass to the coastal fishery. Trans. Am. Fish. Soc. **123**(5): 766–778. doi:10.1577/1548-8659.

Walters, C., Christensen, V., Fulton, B., Smith, A.D.M., and Hilborn, R. 2016. Predictions from simple predator-prey theory about impacts of harvesting forage fishes. Ecol. Model. **337**: 272–280. doi:10.1016/j.ecolmodel.2016.07.014.

Wang, J.-Y., Kuo, T.-C., and Hsieh, C. 2020. Causal effects of population dynamics and environmental changes on spatial variability of marine fishes. Nat. Commun. **11**(1): 2635. doi:10.1038/s41467-020-16456-6.

Wasserman, B.A., Rogers, T.L., Munch, S.B., and Palkovacs, E.P. 2022. Applying empirical dynamic modeling to distinguish abiotic and biotic drivers of population fluctuations in sympatric fishes. Limnol. Oceanogr. **67**(S1):S403–S415. doi:10.1002/lno.12042.

White, J.W., Botsford, L.W., Hastings, A., and Holland, M.D. 2014. Stochastic models reveal conditions for cyclic dominance in sockeye salmon populations. Ecol. Monogr. **84**(1): 69–90. doi:10.1890/12-1796.1.

Wirgin, I., Pedersen, M., Maceda, S., Jessop, B., Courtenay, S., and Waldman, J.R. 1995. Mixed-stock analysis of striped bass in two rivers of the Bay of Fundy as revealed by mitochondrial DMA. Can. J. Fish. Aquat. Sci. **52**(5): 961–970. doi:10.1139/f95-095.

Wood, S.N., and Thomas, M.B. 1999. Super–sensitivity to structure in biological models. Proc. R. Soc. London Ser. B: Biol. Sci. **266**(1419): 565–570. doi:10.1098/rspb.1999.0673.

Ye, H., and Sugihara, G. 2016. Information leverage in interconnected ecosystems: Overcoming the curse of dimensionality. Science, **353**(6302): 922–925. doi:10.1126/science.aag0863.

Ye, H., Beamish, R.J., Glaser, S.M., Grant, S.C.H., Hsieh, C., Richards, L.J., et al. 2015. Equation-free mechanistic ecosystem forecasting using empirical dynamic modeling. Proc. Natl. Acad. Sci. U.S.A. **112**(13): E1569–E1576. doi:10.1073/pnas.1417063112.

Appendix A. Additional details on simulation construction

The mean period of each simulation was measured as the number of years in between peaks, where the minimum height of a potential peak was one-tenth of the maximum within the first 10 years after burn-in. Simulations II and III were constrained to limit the mean period across age classes to less than 10 years. We ran 1000 runs of each simulation and kept the first 100 runs that produced populations that did not crash before the burn-in period began, and that met the aforementioned criteria for limited period. These constraints were established to shorten the period of oscillation, providing enough variance within a reasonable time period for the model to fit the data.