Mathematical Biosciences and Engineering

http://www.aimspress.com/journal/mbe

MBE, 22(9): 2434–2457.
DOI: 10.3934/mbe.2025089
Received: 15 April 2025
Revised: 31 May 2025
Accepted: 19 June 2025
Published: 22 July 2025

Research article

A spatial modeling approach for evaluating impacts of climate-driven species movement on biomass estimation methods

Benjamin A. Levy[1,2,*], **Christopher M. Legault**[3], **Timothy J. Miller**[3] **and Elizabeth N. Brooks**[3]

[1] Ocean Associates Inc. in support of National Marine Fisheries Service, Northeast Fisheries Science Center, VA 22207, USA

[2] Babson College, Mathematics, Analytics, Science, and Technology Division, MA 02457, USA

[3] National Marine Fisheries Service, Northeast Fisheries Science Center, MA 02543-1026, USA

* **Correspondence:** Email: blevy@babson.edu.

Abstract: Fishery stock assessments typically rely on biomass estimates derived from stratified random sampling, where a key assumption is a consistent spatial biomass distribution over time. However, climate-driven movements of marine species may be violating this assumption, potentially introducing biases into biomass estimates. To address this, we develop a spatially explicit data-driven mathematical modeling framework where species-specific movement is driven by environmental variables such as water temperature and geographic habitat preferences. To demonstrate this modeling approach we develop spatial simulations for three Atlantic fish species under several temperature scenarios and population trends. We then compute biomass estimates derived from the stratified random samples of the model output, and compare estimates derived from design-based stratified mean to those estimated from a spatio-temporal model-based approach that allows inclusion of environmental covariates. Our modeling framework produces spatial models that include climate-driven changes in biomass distributions, and resulting biomass estimates are impacted by species shifting their spatial densities over time. This framework has broad uses including evaluation of survey designs, management strategy evaluations, climate-driven biomass predictions, and conducting a rigorous statistical assessment for climate-induced bias of specific biomass estimation approaches.

Keywords: modeling framework; spatiotemporal heterogeneity; fishery resource assessment; design-based; model-based; population simulation; vector autoregressive spatio-temporal model (VAST)

1. Introduction

The continental shelf off the coast of the Northeast United States spans from the Outer Banks of North Carolina to the Gulf of Maine. The region covers over 250,000 km^2 of ocean, extending over 200 km from shore in the largest areas in New England to just 30 km off shore in the southern regions. This ecologically diverse region contains approximately 500 species of fish, making commercial fisheries an important part of local economies for centuries [1]. In 2019, New England fisheries produced $22 billion in sales, which supported over 200,000 jobs [2]. Maintaining a healthy ecosystem is therefore vital to the economic prosperity of the region.

The Northeast Fisheries Science Center (NEFSC) has conducted a bottom trawl survey since 1963 to support assessment and management of the fish and invertebrate populations in the region [3, 4]. The survey uses a stratified random design in which bottom trawl sampling takes place along the eastern continental shelf in strata that are defined by depth and latitude. The survey has created a rich time series of data with many uses including species-specific habitat identification, age and life history information, analysis of how environmental conditions influence species biomass, and estimating yearly species abundance and biomass trends to help inform stock assessments. The survey takes place twice each year, once in the spring and again in the fall. Most spatial analyses and projections of future distributions typically assume the proportion of fish available to the survey remains the same despite seasonal movements. For this reason, NOAA's survey design aims to sample each stratum in approximately the same 6 week period in each season.

Fisheries independent surveys collect data that are used to estimate indices of abundance to inform stock assessments about annual population trends. Stock assessment scientists choose from a number of approaches to obtain biomass estimates ranging from traditional design-based estimates to model-based estimates that vary in complexity. Design-based estimators rely on the sampling design and appropriately defined probabilities of inclusion for each sampling unit to ensure resulting biomass estimates represent population target quantities, on average. These methods typically account for spatial variation in samples by stratifying the sampling area. While it is possible to include environmental covariates with design-based methods [5], it is not a common practice. Model-based biomass estimates use statistical models to estimate the relationship between response variables (such as presence or biomass) and predictor variables (such as environmental factors). Model-based estimators can help account for changes in the spatial distribution of a given species to help overcome problems related to the sampling process. Common model-based approaches include generalized linear models (GLM), generalized additive model (GAM), and general linear mixed models (GLMM) [6]. Recently developed R packages such as VAST and sdmTMB have facilitated the use of advanced model-based approaches [7–11].

Due to a combination of climate change and shifts in ocean circulation, the Northeast United States continental shelf has experienced rapid warming in recent decades. At the same time, a number of recent studies have found a shift in spatial distributions of many species [12–14]. Since stock assessment models rely on accurate descriptions of population dynamics and recent patterns of spatial abundance, there is concern that rapid undocumented changes in spatial distributions of species could bias future stock assessments. More specifically, as fish populations shift their distributions over time, the proportion of the population that is available to be sampled by the survey gear (known as catchability) could change, altering the proportion of the population being sampled, which changes relationship between the index and the true population [15, 16]. Additionally, a species shifting its range beyond the survey area or outside of the assumed stock boundary could cause an abundance index to decrease while in reality the population could be stable or increasing. Existing research has focused on temperature as the driver of such changes [17] and evidence suggests that failing to account for the impact of climate-induced change can lead to management challenges [18]. In these scenarios, management strategy evaluations have shown that mischaracterization of stock status can lead to unintended overfishing, which can ultimately have detrimental ecologic and economic impacts [19]. We are therefore interested in developing a tool to analyze the impact of climate driven movement and distribution changes on the accuracy of biomass estimates derived from the NEFSC's ongoing bottom-trawl survey.

To test the ability of fisheries independent indices of abundance to track population trends under shifting environmental conditions, we construct spatial population simulations for fish in which movement depends on temperature, habitat, and spawning preferences. To consider the impact of climate change one can compare simulations that use a cyclical seasonal water temperature pattern to those where the average temperature increases over time. The impact of increasing water temperature can then be explored by analyzing yearly biomass estimates derived from stratified random samples of the simulated output and comparing them to the true biomass in the simulations. We demonstrate this approach by comparing biomass estimates derived with a stratified mean to estimates obtained from vector autoregressive spatio-temporal (VAST) models. The stratified mean is a design-based approach that calculates the stratum-specific mean biomass per tow, which are then weighted by stratum area to obtain an overall mean. VAST is a spatial delta-generalized linear mixed model that estimates both probability of positive tow and the catch rate for positive tows [10]. VAST also allows users to include covariate data to help explain changes in fish density and to account for changes in catchability. Covariates can be static (e.g., habitat preferences), or dynamic (e.g., temperature). We explore whether including environmental predictors leads to improved biomass estimates, which is particularly relevant as climate change progresses.

2. Methods

We build upon the R package *MixFishSim* to create realistic spatial population models [20]. *MixFishSim* is a discrete spatio-temporal simulation tool that allows users to model multiple species under varying environmental conditions. The package has discrete processes for growth, death, and recruitment. The default package allows users to build simulations using hypothetical temperature gradients on rectangular domains. While this approach was designed to evaluate fleet dynamics and management decisions in mixed fisheries, we demonstrate how to use the *MixFishSim* package to analyze the impact of climate change on our ability to track biomass trends.

2.1. Population simulations

We construct spatial population simulations for yellowtail flounder (*Limanda ferruginea*), Atlantic cod (*Gadus morhua*), and haddock (*Melanogrammus aeglefinus*) on Georges Bank, hereafter referred to as yellowtail, cod, and haddock. Movement of each species in our population simulations combine static species-specific habitat preferences with biologically-based temperature preferences. Dynamics are driven by a time series of temperature gradients to create simulated spatial data sets for each population where the true biomass is known. We simulate 2 temperature patterns: i) temperature gradients that repeat each year, creating data sets with cyclical spatial patterns. ii) temperature gradients that increase on average throughout the simulation, leading to spatial distributions that shift over time. We conduct stratified random sampling on our simulated data to mimic the bottom trawl survey and use the samples to compare the ability of contemporary index standardization methods to track population trends.

2.1.1. Study area

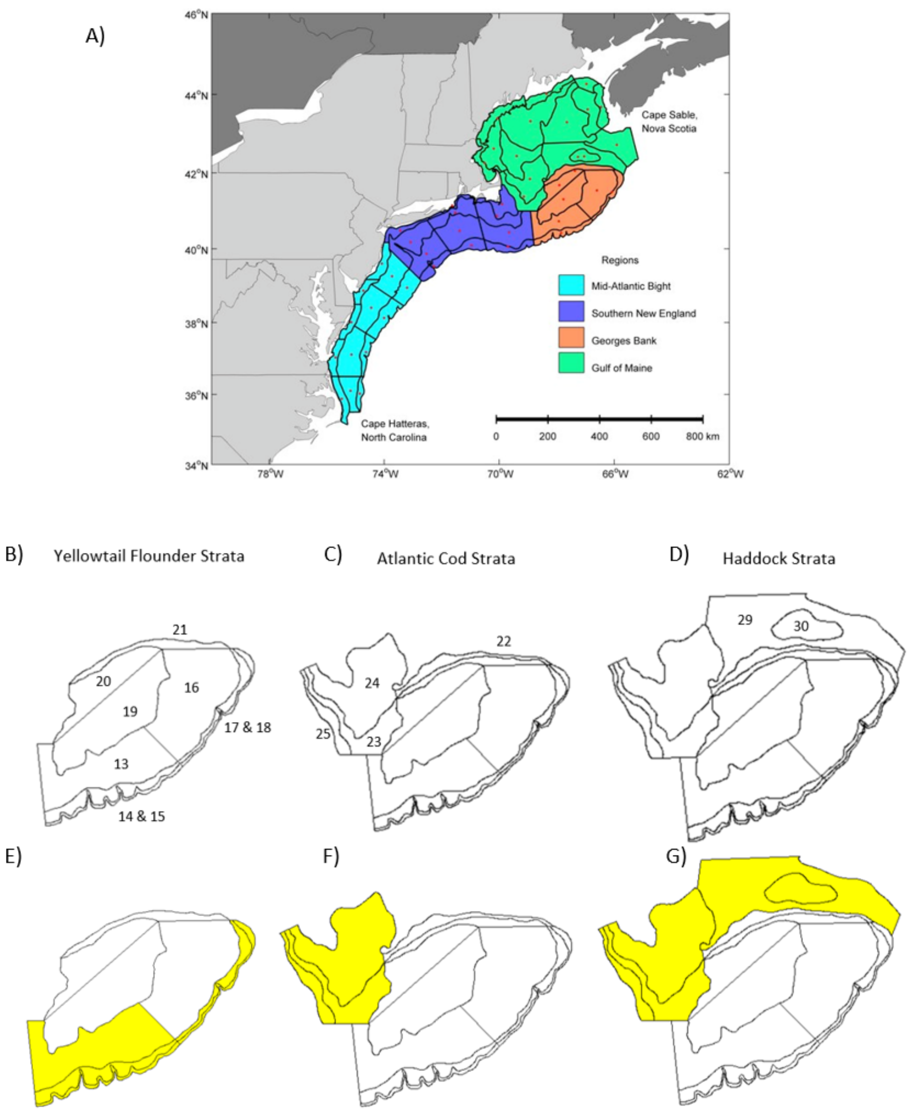

Figure 1. Strata inhabited by each species in our population simulations. All strata sampled by the NEFSC survey are shown plot A, with Georges Bank highlighted in orange (Walsh et al., 2015). Stratum numbers used by the NEFSC bottom trawl survey are shown in plots B, C, and D. Strata that are excluded from certain stratified random sampling scenarios are shown in yellow in plots E, F, and G.

Instead of using default options to generate a hypothetical rectangular domain, we obtained a shapefile for the 15 strata that comprise Georges Bank to use as our simulation environment. These strata reflect the current areas sampled by the ongoing bottom trawl survey. We chose a cell size that provided sufficient spatial detail while keeping the total number of cells manageable, allowing the simulations to run within a reasonable time frame. We discretized the domain into a raster with 88 rows and 144 columns, with each cell in our simulation domain representing approximately 8.7 km^2. A fish stock is considered to be a subpopulation of a species that has similar intrinsic parameters and is reproductively isolated from other members of the species. Each species in our simulations has multiple biologically distinct stocks along the northwest Atlantic coast resulting from local environmental conditions. As a result, the spatial domain for each stock is defined in a different number of strata on Georges Bank. The Georges Bank stock for haddock is defined using all 15 strata in the domain, the cod stock is defined using 13 strata, and the yellowtail stock is defined using 9 strata (Figure 1).

2.1.2. Dynamics and recruitment

The time step for our *MixFishSim* simulations is one week. *MixFishSim* uses a modified two-stage Deriso-Schnute delay difference equation that models the biomass in each cell in our study area [20]. Parameters in the population simulations account for growth of mature adults, mortality (natural and fishing), and the addition of new recruits. Recruitment is a function of the adult biomass that existed in the previous year and is added to the population incrementally throughout each species' predefined spawning period. Parameter inputs were either obtained from the literature or chosen to produce desired dynamics, and therefore, these results cannot be compared directly with stock assessment results (see Tables 1 and 2).

Table 1. Parameters used in all population simulations.

Parameter	Description	Unit	Yellowtail	Cod	Haddock	Source
$\rho = e^{-k}$	Ford's growth coefficient	wk^{-1}	4.48	4.43	4.49	[21]
W_R	Weight of fully recruited fish	kg	0.39	2.95	1.12	[22]
W_{R-1}	Weight of pre-recruit fish	kg	0.13	0.39	0.19	[22]
σ^2	Variance in recruited fish	kg^2	0.55	0.55	0.55	assumed
λ	Decay rate for movement	-	0.7	0.7	0.7	assumed
$Spwn_s$	Spawning weeks for species s	wk	9–12	8–13	11–14	[22]
Rec_s	Recruitment weeks for species s	wk	9–12	8–13	11–14	[22]

2.1.3. Movement

The package combines species-specific temperature tolerances with habitat preferences to drive the probability of movement from cell $c = I$ to cell $c = J$ in week $wk + 1$ using the formulation

$$Pr(c_{wk+1} = J | c_{wk} = I) = \frac{e^{-\lambda \cdot d_{I,J}} \cdot (Hab_{J,s}^2 \cdot Tol_{J,s,wk})}{\sum_{c=1}^{C} e^{-\lambda \cdot d} \cdot (Hab_{c,s}^2 \cdot Tol_{c,s,wk})}, \tag{2.1}$$

where C is the total number of cells in the domain,
$e^{-\lambda \cdot d_{I,J}}$ accounts for the distance d between cells I and J,

Table 2. Parameters used in population simulations for each scenario.

Parameter	Description	Unit	Yellowtail	Cod	Haddock
Constant Population					
M+F	Adjusted Mortality (Natural + Fishing)	1/yr	0.764	0.83	0.309
P0	Initial Biomass	kg	3.19×10^6	2.15×10^7	1.8×10^8
α	Max recruitment rate	kg	3.04×10^7	2.79×10^7	7.36×10^7
β	Recruitment half saturation value	kg	4.3×10^6	1.05×10^7	4.05×10^7
Decreasing Population					
M+F	Adjusted Mortality (Natural + Fishing)	1/yr	0.764	0.623	0.334
P0	Initial Biomass	kg	5×10^7	2.15×10^7	1.8×10^8
α	Max recruitment rate	kg	1.07×10^{15}	3.89×10^{11}	4.97×10^{11}
β	Recruitment half saturation value	kg	2.3×10^{11}	9.8×10^{11}	2.08×10^{15}
Increasing Population					
M+F	Adjusted Mortality (Natural + Fishing)	1/yr	0.564	0.372	0.134
P0	Initial Biomass	kg	3.19×10^6	2.15×10^7	1.8×10^8
α	Max recruitment rate	kg	4×10^7	4.5×10^7	1×10^8
β	Recruitment half saturation value	kg	4.3×10^7	6.28×10^7	4.05×10^8

$Hab_{J,s}$ is the static habitat value for species s in cell J, and

$Tol_{c,s,wk}$ is the value from normally distributed temperature tolerance for species s in cell c in week wk.

The *MixFishSim* package was designed to generate hypothetical temperature gradients and theoretical habitat preferences using Gaussian Random Fields. The following sections describe how we configured the habitat and temperature components to model our three species on Georges Bank.

2.1.4. Habitat input

The habitat input for each species in *MixFishSim* represents spatial preferences for each species that are time-invariant and related to static geographic features. We derived species-specific habitat preferences from niche models for each species using the *lrren* tool from the R package *envi* [23]. The *lrren* tool estimates an ecological niche using the relative risk function by relating presence/absence data to two covariate predictors. Since *lrren* relies on 2D kernel density estimates to approximate a species habitat, we weight our presence/absence values by the biomass sampled at each location. To capture recent preferences, we used bottom trawl point data from 2009–2021 as our presence/absence input by using a value of 0 for any tow that did not catch the given species and by representing a successful catch by the biomass (kg) of the given tow. We combined tow data from both the fall and spring surveys to obscure seasonal influences and instead allow the niche model to infer static habitat preferences independent of temperature. Depth and mean sediment size were used as our covariate predictors. Estimated depth for the region was obtained from FVCOM [24]. The mean sediment size raster was interpolated in ArcMap using the natural neighbor interpolation method using point data collected by the United States Geologic Survey (USGS) [25]. See Figure 2 for a visual representation of this process being applied to cod. Figure 3 depicts habitat preferences $Hab_{J,s}$ for each species s and cell J. Since the values in $Hab_{J,s}$ are required to be between 0 and 1, we rescaled the spatial estimates from *lrren* to fall between these bounds. The values are initially on log scale so we first exponentiate each value in $Hab_{J,s}$ before rescaling them to fall between 0 and 1 by letting

$$Hab_{J,s} = \frac{Hab_{J,s} - min(Hab_{J,s})}{max(Hab_{J,s}) - min(Hab_{J,s})}. \tag{2.2}$$

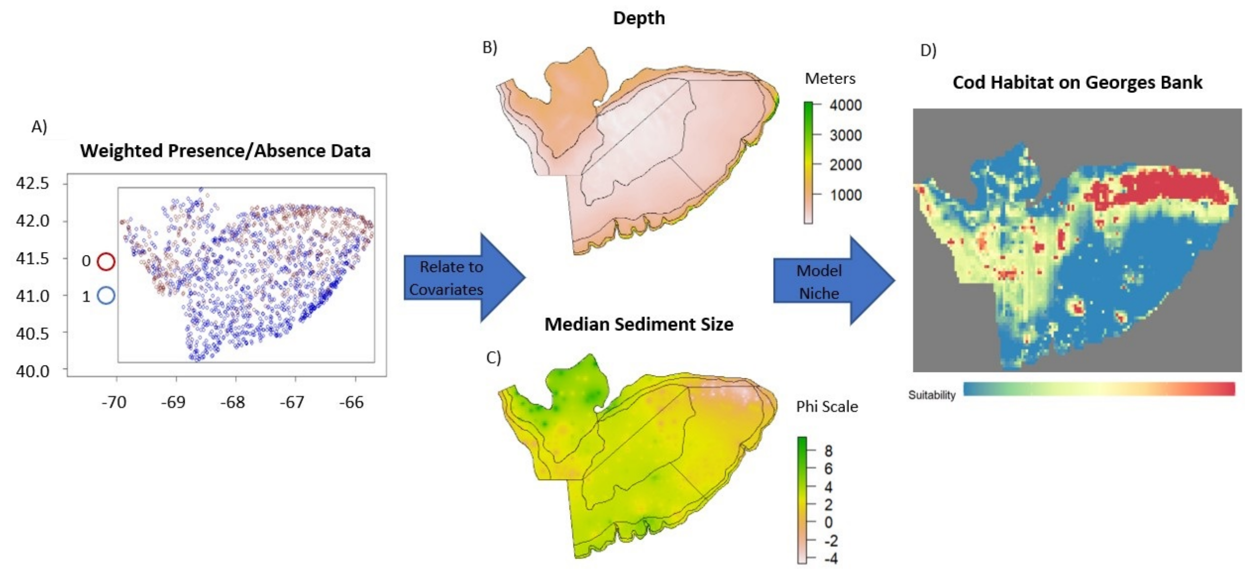

Figure 2. Visual representation of niche model for Atlantic cod. In plot A, tows with positive cod catch are shown in red while tows without cod are in blue. Positive tows are weighted by the biomass of the given catch. Plots B and C show our covariate predictors used in the *lrren* niche model- depth and median sediment size. Plot D shows the final niche model that has been rescaled between 0 and 1, where blue values are closer to 0 and red values are closer to 1.

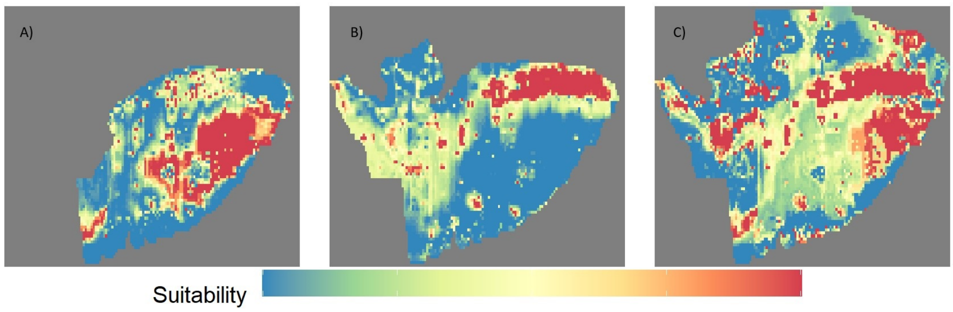

Figure 3. Static habitat preferences $Hab_{J,S}$ for each species s and cell J in our *MixFishSim* population simulations. From left to right: yellowtail flounder (A), Atlantic cod (B), and haddock (C). Values range from 0 (blue) to 1 (red).

2.1.5. Temperature input

The default assumption in *MixFishSim* is that each species has normally distributed temperature (°C) preferences $N(\mu, \sigma)$, with mean μ and standard deviation σ. While this is a simplifying assumption that may not exactly reflect the true distribution of temperature preferences for all Atlantic fish, this is a reasonable assumption given that there is evidence that at least Atlantic cod preferences are normally distributed [26]. Values were chosen by combining local occurrence information in the literature with temperatures recorded in the bottom trawl survey [22]. Yellowtail's occurrence data showed a distribution defined by $N(8.75, 4.25)$, while haddock and cod occurrence data can be represented by the distribution $N(9, 4)$. Weekly estimated temperature data for the region for 2012 was obtained from FVCOM [24] and used in *MixFishSim* for the $Tol_{c,s,wk}$ term in (1), where the value is the normally distributed temperature tolerance for species s in cell c in week wk. We chose to repeat temperature estimates for a single year rather than use data for consecutive years to reduce the number of factors impacting population dynamics while still incorporating real data. The 2012 data was chosen because it displayed an average temperature pattern that consistently oscillated between maximum and minimum temperature values, allowing for a smooth repeating yearly temperature pattern for the repeating temperature scenario. The 2012 temperature data was also transformed to create an oscillating pattern with a trend that increases 5 degrees Celsius on average over the duration of the simulation. We chose a 5 degree increase over a 20 year simulation to allow temperature change to have a meaningful impact on spatial patterns while remaining within reasonable computational limits in terms of the length of the simulation. Figure 4 depicts mean trends for the temperature scenarios used in our population simulations.

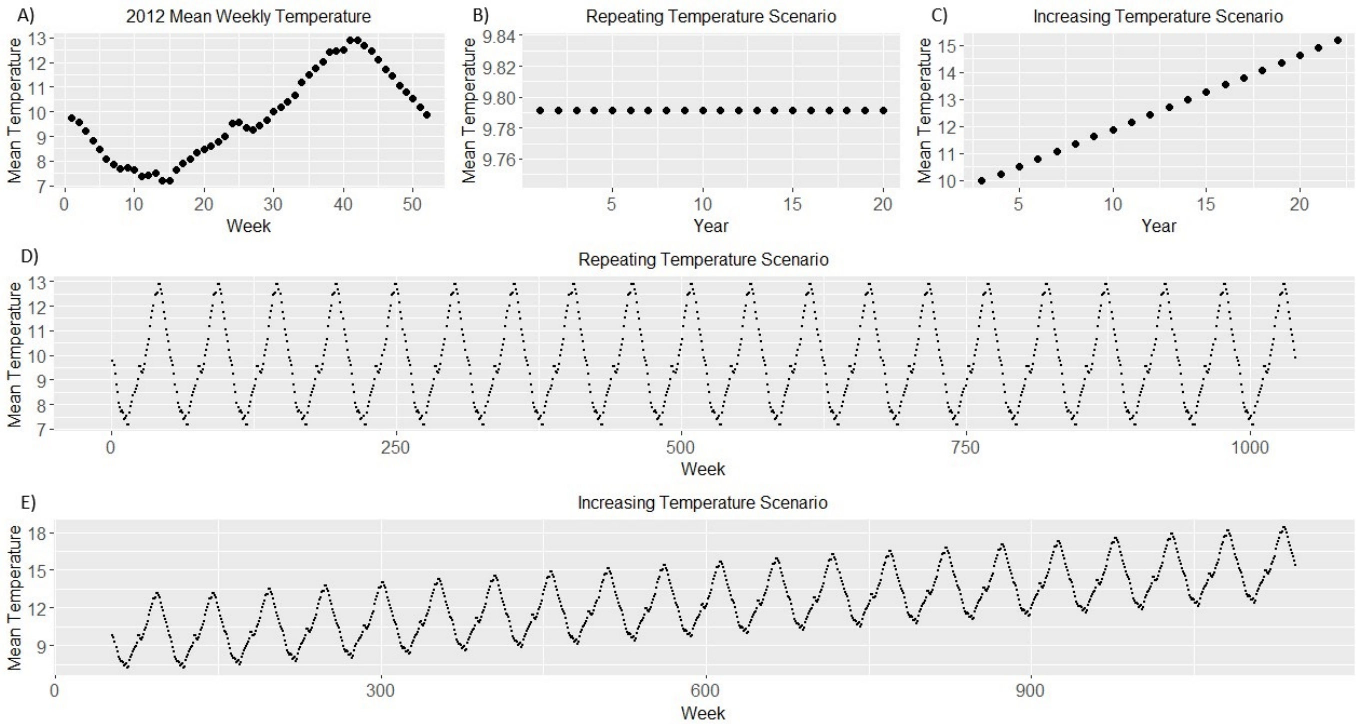

Figure 4. Average temperature trends used to drive movement in our *MixFishSim* population simulations. Plot A is the 2012 temperature data, which was copied to create the repeating weekly temperature trend in plot D. Plot E shows the analogous weekly temperature input used in the increasing temperature scenarios. Plots B and C show the average values of the two temperature trends on a yearly time scale.

In Eq (1), $Hab_{J,s}$ remains constant for each species for the duration of the simulation, except for spawning weeks during which each species-specific spawning area is exaggerated to force movement to that area. On the other hand, $Tol_{c,s,wk}$ changes each week with temperature fluctuations. Using a temperature gradient that repeats every 52 weeks produces the same spatial preferences in a given week each year, resulting in relatively consistent spatial biomass patterns. Scenarios in which the temperature increases over time create spatial preferences that evolve as the water warms, producing spatial biomass patterns that shift in a given week over the duration of the simulation.

2.1.6. Population simulation scenarios

To account for model initialization, we carry out 22 year simulations for each of our three species, but discard the first 2 years when conducting all analyses. For increasing temperature scenarios, the temperature gradient does not begin to increase until after the first 2 years have passed (thus, years 3 to 22 are plotted in all Figures). We consider various population trends in our simulations. Historically, cod has seen significant decline over the last 50 years while haddock has increased in abundance in recent years [22]. For this reason, we compare survey index estimates using stratified random sampling from decreasing population scenarios for cod and increasing population scenarios for haddock. To consider a wide range of possibilities, we include increasing, decreasing, and constant population trends for yellowtail. The specific population trends used in our analyses can be see in Figure S1. To analyze how indices may be impacted by changes in population size under different temperature trends, each of these population scenarios is simulated twice: first with an oscillating temperature gradient that repeats and second with a temperature gradient that increases roughly 5 degrees Celsius over the duration of the 20 year simulation. As a result, we generated a total of 10 simulated spatial data sets.

2.1.7. Simulating bottom trawl survey

After simulating each species in its stock area with different population trajectories, we mimic the bottom trawl survey by conducting stratified random sampling in the set of strata associated with each stock. With the spring survey taking place either during (haddock), or immediately after (cod, yellowtail) spawning season, each species typically has a different spatial distribution in each season. We therefore sample each stratum twice per year in the same weeks in which the spring and fall surveys have historically sampled Georges Bank (weeks 13 and 14 in the spring and 37 and 38 in the fall). Cells within each stratum were randomly selected and the number of simulated tows each season reflects true target values for each stratum. Survey values themselves consist of the total biomass that exists in a given cell from our *MixFishSim* simulations. Similar to the actual survey, the resulting simulated tows contain a large number of samples that did not catch these species (i.e., zero biomass) as seen in a representative histogram of tows from a scenario for haddock (Figure S2).

Underlying assumptions in all index methods is that individual random samples combine to accurately represent true biomass by consistently sampling the same proportion of the population over time. These assumptions can be questioned given enough noise in the sampling process and/or climate change causing a population to move into previously uninhabited strata. To simulate the impact of noise, we compare index estimates after adding noise to our samples versus those using the true sampling values. In scenarios where we add noise to our samples, log-normal deviations with a

CV of 0.35 were added to the annual "true" survey catches. The implication of a given population shifting its distribution into new habitat outside of the normal survey area is that the entire geographic extent of the population is not sampled, and the fraction that is sampled is not constant. We simulate the effect of populations moving into new habitat by comparing index estimates using samples from all strata inhabited by each species on Georges Bank (called full coverage) to those that only sample a subset of the domain for each species (called partial coverage). The strata to exclude for each species were chosen by reviewing how the distribution of each species evolved in our increasing temperature scenarios and removing some strata where the proportion of a given population was changing over time. Figure 1 shows all strata inhabited by each species as well as those that are removed from certain biomass estimates using the spatial shifting trends shown in Figures 5 and S3. The yellow regions in Figure 1 depict the strata that were not sampled for each species in the partial coverage scenarios. We then use the biomass collected from our samples in contemporary biomass index methods to estimate yearly seasonal population trends.

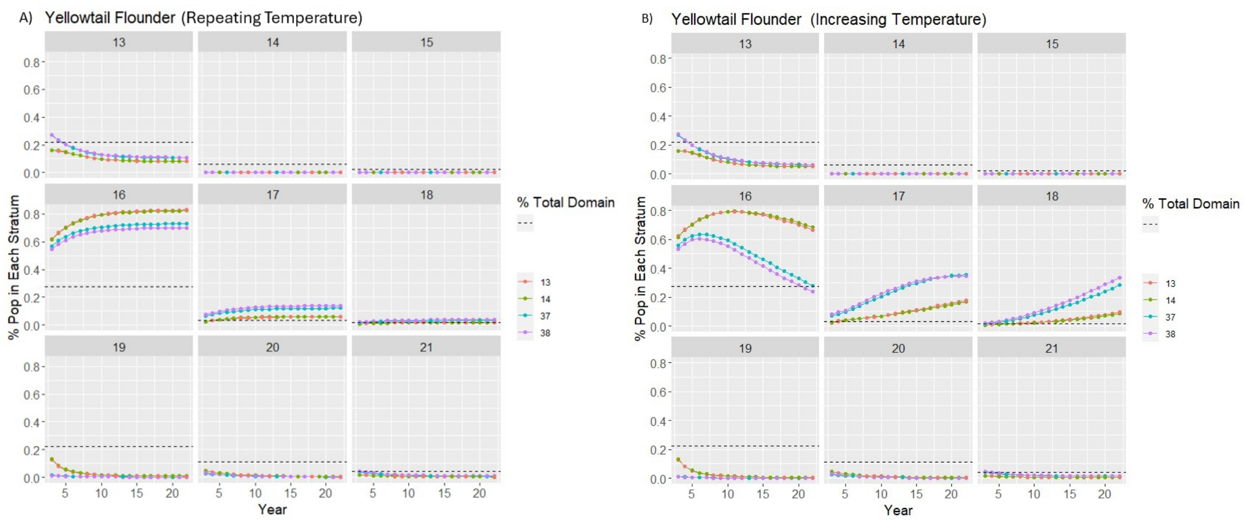

Figure 5. Percent of yellowtail flounder in each stratum during survey weeks (13, 14 in spring, 37, 38 in fall) in our spatial simulations. All repeating temperature scenarios follow the patterns in plot A on the left while increasing temperature scenarios follow the patterns in plot B on the right. See Figure 1 for a spatial reference of the Georges Bank strata. The black dotted line represents the percent of the entire domain that the given stratum covers, which also approximately corresponds to the allocation of tows to each stratum. Similar plots for cod and haddock are shown in Figure S2.

2.1.8. Computing biomass indices

The design-based method we used is the stratified mean, which divides the inhabited domain into M disjoint strata based on depth and latitude. The number of samples taken from each stratum N_i for $i = 1...M$ is proportional to the given area. The stratified mean biomass $\mu^D_{s,y}$ for a given season s and year y can then be calculated as the weighted average

$$\mu_{s,y}^{D} = \sum_{i=1}^{M} W_i \frac{\sum_{k=1}^{N_i} y_{k,i}}{N_i}, \tag{2.3}$$

where $W_i = \frac{A_i}{\sum_{j=1}^{N} A_j}$, A_i is the area of stratum i, and $y_{k,i}$ represents the biomass in sample k of stratum i in season s and year y. These calculations are quick and easy to compute relative to the model-based VAST estimates, but they are also more sensitive to the tow values for the given year.

We compare stratified mean estimates to those derived from the model-based VAST approach. VAST models both encounter probability $p_1(i)$ and positive catch rate $p_2(i)$ for each observation i as linear predictors using a spatial delta-generalized linear mixed model that can be represented by

$$f(p_*(i)) = \beta_*(t_i) + \omega_*(s_i) + \epsilon_*(s_i, t_i) + \upsilon(s_i, t_i) + \xi(i), \tag{2.4}$$

for $* = 1, 2$, where $\beta_*(t_i)$ models temporal variation at time t for observation i, $\omega_*(s_i)$ represents spatial variation at location s for observation i, $\epsilon_*(s_i, t_i)$ models spatio-temporal variation, $\upsilon(s_i, t_i)$ represents catchability, and $\xi(i)$ models the covariate response for observation i. The covariate term is only used in certain scenarios and the catchability term was removed since values from our survey contain the exact biomass value from the given cell in our *MixFishSim* simulations. More details related to each component are further described in [10]. VAST models require numerous user inputs to determine how the linear predictors will be conditioned and solved, and individual models can often take on the order of hours to converge, especially when including covariate information.

Using stratified random samples of our *MixFishSim* simulations, we follow the advice given in [10] and [27] to build VAST models that estimate biomass in these spatial populations on Georges Bank. In addition to exploring different link functions and assumed distributions, our VAST model-building process involved testing the impact of including spatial and/or spatio-temporal variation in our models, considering varying number of knots in our mesh, and testing different forms of temporal correlation. We analyzed how covariates can impact biomass estimates by fitting VAST models with and without covariates. We considered the driving factors of movement in our *MixFishSim* simulations as covariates- the dynamic temperature values and static habitat values. We ultimately decided to provide the most information possible to the model by including exact values for both temperature and habitat covariates for both linear predictors to consider a best-case scenario. VAST models that included covariates used a linear combination of second degree polynomials for habitat and temperature to approximate species-specific covariate responses (Table S1).

Through this process, and in consultation with the VAST package creator, we compared the performance of two model setting configurations in our VAST models as shown in Table S1. VAST settings A are more typical for stock assessment applications (personal communication with James T. Thorson, January 2022), while VAST settings B were chosen after reviewing the information provided in the “seasonal model” section of VAST’s online Github tutorial found at https://james-thorson-noaa.github.io/docs/tutorials/seasonal-model/. Both settings use a Tweedie model and the difference between the two settings can be seen in the Rho Configuration, where RhoBeta2=0 in settings A while RhoBeta2=3 in settings B. The difference amounts to intercepts for the biomass predictor being a fixed effect (RhoBeta2=0) versus being constant among years as a fixed effect (RhoBeta2=3). We focus our discussion on VAST results derived from settings A. Performance metrics for settings B are shown to demonstrate how the settings used in model-based approaches impact model performance.

Each biomass index scenario we consider is a combination of specific population trends for each species, differing temperature scenarios, altering seasons, and sampling possibilities (noise, strata, and covariates), resulting in a large number of scenario combinations to consider. Table 3 show the choices that define each scenario, resulting in 400 unique combinations.

Table 3. Population simulation configurations, each simulated over a 22 year period.

Item	Description	Variants	Cumulative Models
Species: Population Trend	Yellowtail: Increasing, Stable, Decreasing Cod: Decreasing Haddock: Increasing	5	5
Temperature values	Repeating, Increasing	2	10
Survey Season	Spring, Fall	2	20
Sample noise	Yes, No	2	40
Survey coverage	Full, Partial	2	80
Biomass Estimation Method	Stratified mean VAST_A_NoCov, VAST_A_WithCov VAST_B_NoCov, VAST_B_WithCov	5	400

2.1.9. Biomass index performance metrics

We evaluated the performance of our biomass indices by analyzing accuracy of biomass predictions and by comparing how well biomass estimates follow true biomass trends.

Knowing the true population values in our simulations allowed us to assess the accuracy of each estimation method by computing the absolute relative error as given by

$$E_{biomass} = \frac{\sum_{y=3}^{22} |M_y - I_y|}{\sum_{y=3}^{22} M_y}, \tag{2.5}$$

where M_y is the true biomass in year y for the given scenario and I_y is the corresponding biomass estimate provided by either the stratified mean or VAST. We refer to this metric as biomass error, which allows us to evaluate the accuracy of each biomass estimate relative to the true biomass in our *MixFishSim* population simulations, where smaller biomass error values represent more accurate biomass estimates.

Given species-specific differences in catchability and the varying nature of the survey itself, the accuracy of the trend of biomass estimates is often more relevant to a stock assessment than the specific values of the estimates. With this in mind, the underlying assumption is that the biomass estimate for species s is proportional to the true biomass by a scaling factor q_s^*. We evaluate the accuracy of the trend of each biomass estimate in the following way.

First we compute q_s^* for species s in a given scenario as the ratio between the sum of all true biomass estimates on the log scale divided by the sum of estimates from VAST or the stratified mean on the log scale

$$q_s^* = \frac{\sum_{y=3}^{22} log(M_y)}{\sum_{y=3}^{22} log(I_y)}. \tag{2.6}$$

We then apply q_s^* to the log of the estimate to align the biomass estimate with the true value, and calculate the trend error of the resulting scaled estimate

$$E_{trend} = \frac{\sum_{y=3}^{22} |\frac{q_s^* * log(I_y) - log(M_y)}{log(M_y)}|}{20}, \tag{2.7}$$

where M_y is the true biomass in year y for the given scenario and I_y is the corresponding biomass estimate provided by either the stratified mean or VAST. We refer to this metric as the the trend error. The value of the trend error quantifies the accuracy of the trend for each biomass estimate relative to the trend of the true biomass in our *MixFishSim* population simulations, with smaller values of E_{trend} representing biomass estimates with trends that more closely match the true biomass.

3. Results

3.1. MixFishSim spatial population simulations

We were able to create changing distribution patterns in our *MixFishSim* simulations using an oscillating temperature gradient that increased on average over time. For example, yellowtail flounder in the repeating temperature scenarios was found in the shallow stratum 16 in much higher proportion than its area in both seasons due to the combination of spawning location and habitat and temperature preferences (Figure 5(a)). In contrast, in the increasing temperature scenarios, yellowtail flounder reduced abundance somewhat in stratum 16 in the spring during spawning season (weeks 13 and 14) but were driven to the deeper cooler waters of strata 17 and 18 in the fall (weeks 37 and 38; Figure 5(b)). These results reflect the hypothesis that spawning locations are a result of the physical habitat more than the temperature. Similar movements to deeper and cooler waters in the increasing temperature scenarios in the fall were generated for cod and haddock (Figure S2). Note that despite removing the first two years of simulations, there was still some movement among strata in the first few years as the simulations stabilized. The temperature driven impacts are seen in the later years of the simulations.

3.2. Error values

Representative examples of species-specific biomass trends and biomass estimates are shown in Figures S4–S6. Figures 6 and 7 compare biomass error and trend error values computed via Eqs (2.5) and (2.7). The specific values used to create Figures 6 and 7 are provided in Tables S2–S8. A lower value indicates better performance in both metrics and since the biomass error cannot be directly compared to trend error values, we will focus on percent changes in each metric between different scenarios and inputs.

There is a consistent ordering of biomass error values across the three estimation approaches-stratified mean (SM), VAST without covariates (NC), and VAST with covariates (WC) (Figure 6). VAST models that include covariate information provided the lowest average biomass error and trend error values among the three estimate approaches. Although the stratified mean tended to have a lower average biomass error than VAST models that did not include covariates across the *MixFishSim* scenarios, the order was reversed for trend error values. This implies that although stratified mean estimates tended to provided more accurate approximations of the true biomass in our *MixFishSim* simulations, VAST without covariates more closely matched the true biomass trend. Simulations with an increasing temperature gradient tended to produce higher biomass and trend error values compared to the corresponding scenario with a repeating temperature gradient, with the one exception being the stratified mean estimation of the decreasing population trend. The magnitude of the error bars in this figure differ due to differences in estimation methods, seasons, and sampling coverage. VAST estimates without covariates tend to be most sensitive to these differences while VAST models that include covariates provides the most consistent estimates between scenarios. An exception to this general trend is that the variance of the stratified mean biomass error results are similar to the estimates with covariates in the increasing temperature scenarios.

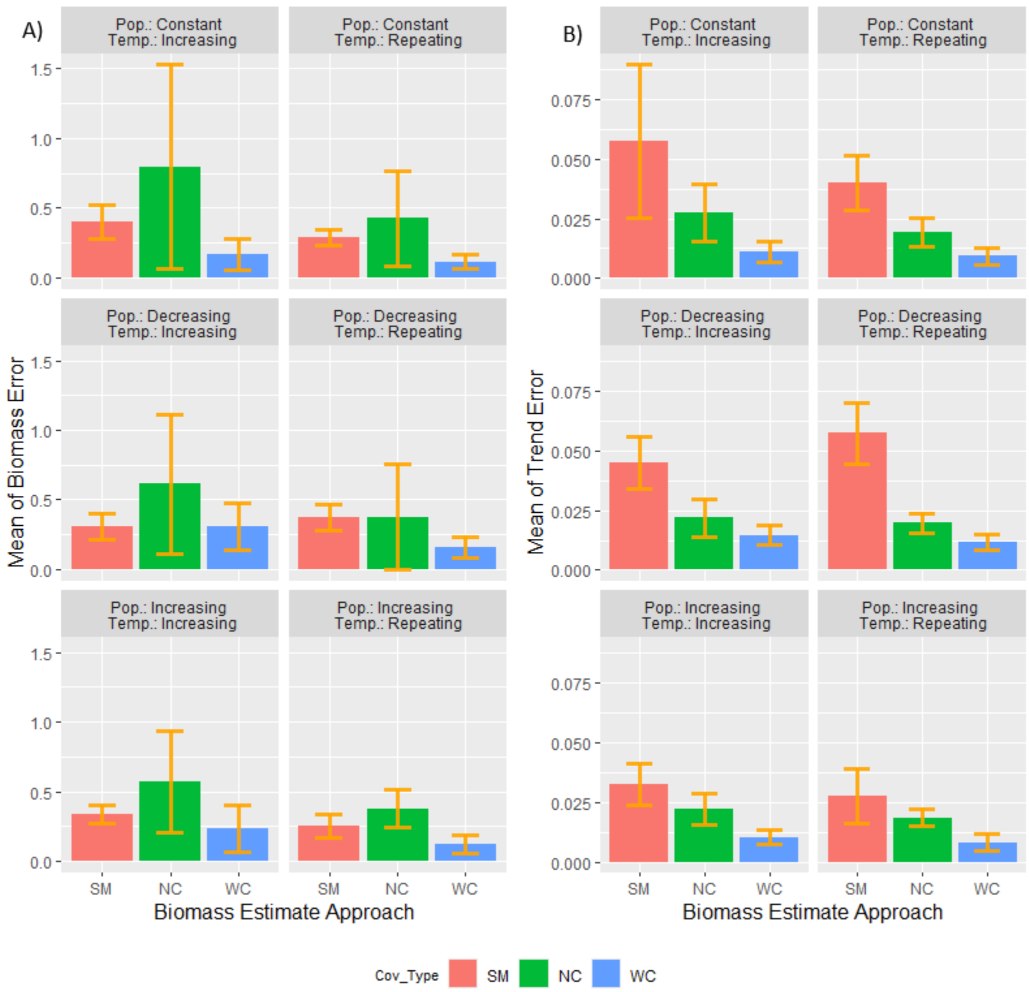

Figure 6. Comparing mean biomass error (plot A) and mean trend error (plot B) values for each *MixFishSim* biomass and temperature scenario, and each biomass estimation method-stratified mean (SM), VAST without covariates (NC), and VAST with covariates (WC). Biomass trend and temperature scenarios are described in the Methods section.

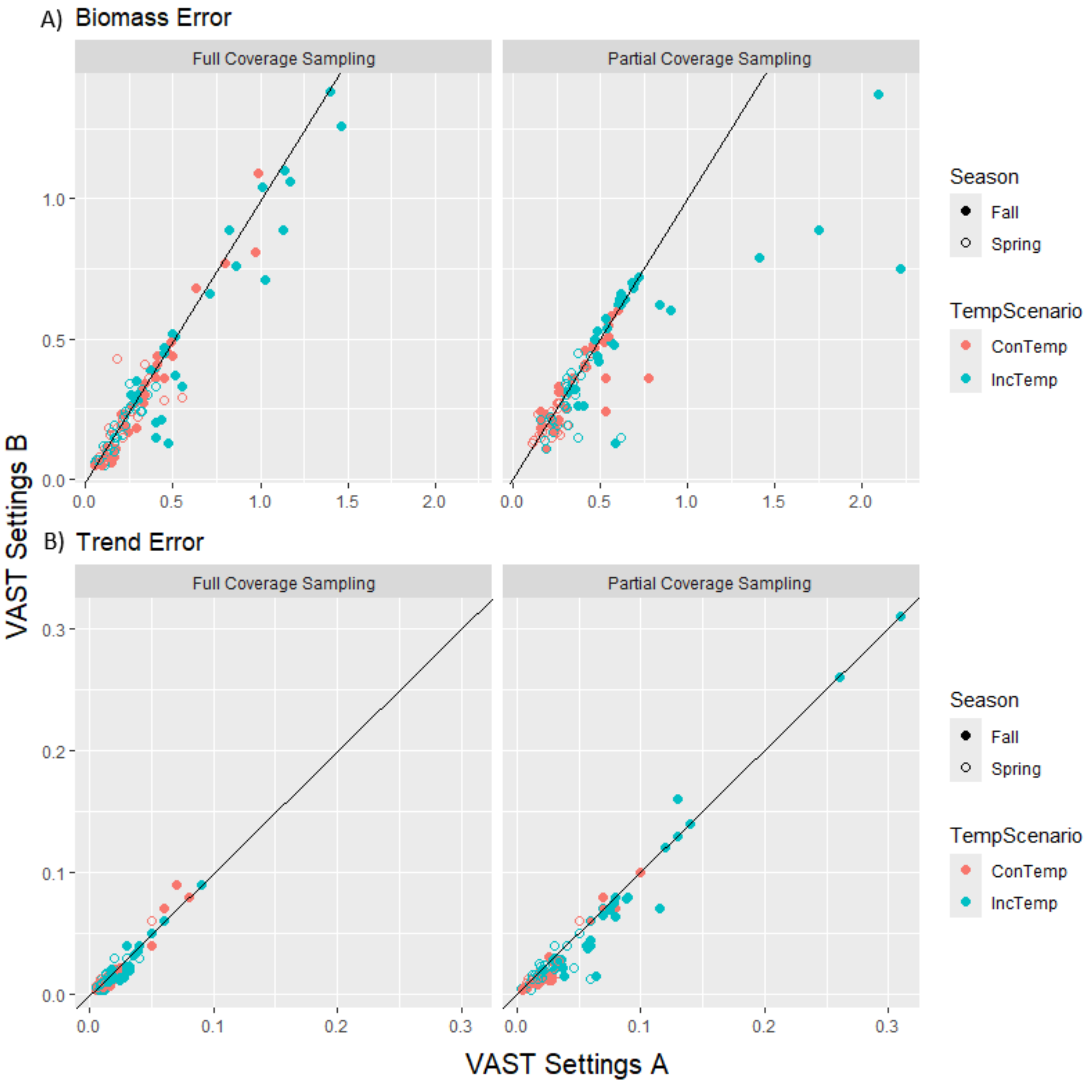

Figure 7. Comparing mean biomass error (plot A) and mean trend error (plot B) values for the two VAST setting configurations we tried- Settings A and Settings B. The values are further faceted by sampling coverage. A 1 : 1 line is shown for reference.

Figure 7 plots biomass error and trend error values for both of the VAST settings we considered, further stratified by the sampling coverage. Recall that the partial coverage sampling scenario excludes certain strata from the simulated survey to explore how species moving into unsampled locations impacts estimates. The largest values in all plots are from fall samples, with the largest values being from scenarios with an increasing temperature gradient. Trend error values remained close to the 1 : 1 line, with two outliers seen in the Partial Coverage error values. The biomass error values do not have outliers in either sampling scenario, but display larger variation in error between settings, with settings B often performing better than settings A. This implies the two settings estimated the trend error similarly across all scenarios, including the two outliers where the trend error deviated from other values, while the biomass estimate was more sensitive to the VAST settings used.

Comparison of error across all estimates

The biomass and trend error values across all results produce symmetric distributions for stratified mean estimates and distributions that are skewed left for VAST estimates (row 1, Tables 4 and 5). With similar mean error values (0.38 compared to 0.39) and different coefficient of variation (CV) values (0.95 compared to 0.36), the distribution of model-based biomass error results cover a wider range of outcomes over the 160 scenarios. With respect to trend error values, model-based mean and median values were less than half the design-based values with similar CVs (0.77 versus 0.82) implying two symmetric distributions that have similar spread where model-based trend errors tend to be lower. Thus, across all scenarios, while the accuracy of the design-based and model-based biomass error estimates were similar, the stratified mean provided more precise biomass estimates, while VAST provided more accurate trend error.

Table 4. Summary biomass error results for different scenarios.

Row	Scenario	N VAST (SM)	Mean VAST Error	Mean SM Error	Median VAST Error	Median SM Error	VAST Error CV	SM Error CV
1	All results	160 (80)	0.38	0.39	0.26	0.37	0.95	0.36
2	W/ Covariates	80	0.25	N/A	0.21	N/A	0.59	N/A
	No Covariates	80	0.51	0.39	0.32	0.37	0.9	0.36
3	Full Coverage (FC)	80 (40)	0.35	0.32	0.21	0.32	0.99	0.28
	Partial Coverage (PC)	80 (40)	0.4	0.46	0.3	0.45	0.97	0.31
3a	FC W/ Covariates	40	0.19	N/A	0.16	N/A	0.71	N/A
	FC No Covariates	40	0.51	0.32	0.35	0.32	0.81	0.28
3b	PC W/ Covariates	40	0.3	N/A	0.25	N/A	0.48	N/A
	PC No Covariates	40	0.51	0.46	0.31	0.45	0.98	0.31
4	FC Repeating Temp (RT)	40 (20)	0.26	0.31	0.16	0.31	0.94	0.31
	FC Increasing Temp (IT)	40 (20)	0.44	0.34	0.26	0.32	0.91	0.26
4a	FC RT W/ Covariates	20	0.14	N/A	0.13	N/A	0.45	N/A
	FC RT No Covariates	20	0.39	0.31	0.33	0.31	0.72	0.31
4b	FC IT W/ Covariates	20	0.25	N/A	0.22	N/A	0.69	N/A
	FC IT No Covariates	20	0.63	0.34	0.53	0.32	0.76	0.26
5	FC Fall (Fa)	40 (20)	0.54	0.38	0.42	0.40	0.72	0.22
	FC Spring (Sp)	40 (20)	0.16	0.26	0.14	0.26	0.63	0.24
5a	FC Fa W/ Covariates	20	0.27	N/A	0.24	N/A	0.54	N/A
	FC Fa No Covariates	20	0.82	0.38	0.84	0.40	0.44	0.22
5b	FC Sp W/ Covariates	20	0.12	N/A	0.11	N/A	0.46	N/A
	FC Sp No Covariates	20	0.2	0.26	0.16	0.26	0.59	0.24

Impact of covariates across all estimates

A primary goal of this work is to analyze how including covariates in model-based estimates helps account for increasing water temperature, seasonal differences, and evolving spatial biomass distributions. Depending on the scenario, including covariates decreases average model-based biomass error value by 50–75% and shows improvement over the average stratified mean values by 25–55% (Table 4). Regardless of the scenario, the average stratified mean biomass error falls between the model-based estimates that use covariates and those that do not, which demonstrates the potential improvement covariates can provide while also highlighting the variability of the model-based approach with respect to the accuracy of the biomass estimates. This is in contrast to the trend error values, where the stratified mean trend errors are larger than all model-based averages, regardless of whether or not covariates were included (Table 5).

Including covariates in our model-based estimates decreased biomass CV values relative to those that did not use covariate values (0.59 versus 0.90). The difference in these values appears to be driven by results related to partial coverage (Table 4, row 3b) and full coverage sampling with a repeating temperature pattern (Table 4, row 4a). For comparison, each CV value for stratified mean biomass results were lower than every model-based CV, indicating consistent biomass estimates for the designed-based approach. With respect to trend error, the coefficient of variation values across all results were more similar between all three estimation methods (Table 5, row 2). The CV values remain closely grouped and relatively low (< 0.5) across all scenarios, except for a noticeably large value when covariates were used in model-based estimates with full coverage sampling and repeating temperature gradient (Table 5, row 4a).

Table 5. Summary trend error results for different scenarios.

Row	Scenario	N VAST (SM)	Mean VAST Error	Mean SM Error	Median VAST Error	Median SM Error	VAST Error CV	SM Error CV
1	All results	160 (80)	0.02	0.06	0.02	0.05	0.77	0.82
2	W/ Covariates	80	0.02	N/A	0.01	N/A	0.88	N/A
	No Covariates	80	0.03	0.06	0.02	0.05	0.62	0.82
3	Full Coverage (FC)	80 (40)	0.02	0.04	0.02	0.04	0.44	0.43
	Partial Coverage (PC)	80 (40)	0.03	0.07	0.02	0.05	0.71	0.87
3a	FC W/ Covariates	40	0.01	N/A	0.01	N/A	0.41	N/A
	FC No Covariates	40	0.02	0.04	0.02	0.04	0.3	0.43
3b	PC W/ Covariates	40	0.02	N/A	0.02	N/A	0.83	N/A
	PC No Covariates	40	0.03	0.07	0.03	0.05	0.66	0.87
4	FC Repeating Temp (RT)	40 (20)	0.01	0.04	0.01	0.05	0.45	0.41
	FC Increasing Temp (IC)	40 (20)	0.02	0.04	0.01	0.04	0.47	0.40
4a	FC RT W/ Covariates	20	0.01	N/A	0.01	N/A	1.73	N/A
	FC RT No Covariates	20	0.02	0.04	0.02	0.05	0.24	0.41
4b	FC IC W/ Covariates	20	0.01	N/A	0.01	N/A	0.37	N/A
	FC IC No Covariates	20	0.02	0.04	0.02	0.04	0.36	0.40
5	FC Fall (Fa)	40 (20)	0.02	0.05	0.02	0.05	0.44	0.39
	FC Spring (Sp)	40 (20)	0.01	0.03	0.01	0.03	0.42	0.30
5a	FC Fa W/ Covariates	20	0.01	N/A	0.01	N/A	0.32	N/A
	FC Fa No Covariates	20	0.03	0.05	0.03	0.05	0.28	0.39
5b	FC Sp W/ Covariates	20	0.01	N/A	0.01	N/A	0.22	N/A
	FC Sp No Covariates	20	0.02	0.03	0.02	0.03	0.17	0.30

Comparison of error between survey sampling scenarios

When we compared mean error results obtained from sampling all strata (full coverage) compared to partial coverage sampling, we saw a larger percent increase for the stratified mean relative to VAST with respect to the mean biomass error, mean trend error, and the variance of each value. Most notably, model-based results have CV values that are ˜300% the magnitude of designed-based CVs with mean biomass error values sensitive to whether or not covariate values are included in our VAST models (Table 4, rows 3a & 3b).

Both VAST and the stratified mean had % increase in average trend error values with partial coverage sampling of the domain (Table 5, row 3). Both VAST average trend error values were less than the lowest stratified mean average, and results displayed ˜60% decrease in error values when model-based estimates were used in place of the design-based approach in a given scenario (Table 5, rows 3, 3a and 3b). Unlike the biomass error results, CV values within a given coverage scenario were similar between model-based and design-based estimates.

Collectively, these results suggest that failing to sample the entire domain has a larger impact on our ability to track population trends compared to absolute biomass values. Additionally, partial coverage sampling impacts the precision of model-based biomass estimates more than design-based estimates.

Comparison of error between temperature scenarios

To analyze the impact of a repeating temperature gradient compared to a temperature gradient that increases on average, we reduced the number of factors impacting each estimate by only considering estimates derived from full coverage sampling. With respect to the stratified mean, the temperature scenario had a small impact the average biomass or trend error values, or their corresponding CV (Tables 4 and 5, row 4). The average VAST error values had a noticably larger increase between temperature scenarios, especially with respect to mean biomass values.

The VAST biomass CV compared to the stratified mean CV in each table continue the trend of especially high values related to VAST biomass estimates, but comparable CV values related to trend error. The disparity between CV values is driven by whether or not covariates were used in these models (Tables 4 and 5, rows 4a and 4b). Thus, while the stratified mean is less sensitive to changing water temperature and provides precise estimates between temperature scenarios, including covariates provided the most accurate biomass and trend estimates under these conditions.

Comparison of error between seasons

To compare differences in error values by season, we again reduced the number of factors by only including scenarios that sampled from all strata. This resulted in 40 estimates per season for our model-based approach (VAST with and without covariates) and 20 estimates for the stratified mean. The average biomass error value for the model-based results increased 238% from the spring value of 0.16 to the fall average of 0.54, while the corresponding stratified mean biomass error values display a 46% increase between the two seasons (0.26 to 0.38) (Table 4, row 5). The notable increase in mean model-based values is driven by the difference between the mean fall value without covariates of 0.82 and the mean spring value with covariates of 0.12. The fact that the model-based results for these scenarios contain the largest biomass error value in the table (0.82) as well as the lowest value (0.12), highlights the value and variability of a model-based approach. Similar to other scenarios, VAST average trend error values both with and without covariates were again lower than the lowest stratified mean average. The trend results also reflect lower values in the spring compared to the fall.

4. Discussion

We have presented a modeling approach to create spatial datasets for real fish where movement is driven by species-specific geographic preferences combined with dynamic temperature gradients. We also demonstrated how these datasets can be used to analyze the impact of climate-driven movement on design-based and model-based biomass estimates derived from a fishery-independent survey. Due to computational constraints related to the model-based estimates, especially those that include environmental covariates, we estimate biomass from only a single replicate for each population and temperature scenario to demonstrate the potential of our modeling approach. Assessing bias requires multiple replicates to reliably determine whether observed differences between estimated and true biomass are systematic or random. With additional replicates, we could statistically analyze and quantify these biases, improving confidence in biomass estimation methods under various climate-driven scenarios.

There are a few additional limitations to this study worth highlighting. While using VAST with environmental covariates can be a useful way to account for shifting distributions under climate change, its effectiveness depends on having the right covariates and a good understanding of how they influence fish density. If the wrong variables are included, or their effects are misrepresented, the model can mislead rather than help a stock assessment. On the other hand, the more traditional design-based method of the stratified mean is simple and transparent. While it may not fully capture changing patterns in fish distribution, our results suggest it still performs reasonably well. We don't mean to suggest it should be abandoned, but rather that there may be situations where model-based approaches could offer improvements.

This study also used *MixFishSim*, which simulates biomass but not age structure. This allowed us to focus on spatial distribution and simplify the analysis, but it left out a key piece of information that matters in stock assessments. *MixFishSim* also assumes that each species' temperature preferences follow a normal distribution. This assumption made it possible to generate realistic spatial datasets with well-defined covariate relationships, which helped in testing how biomass estimation approaches perform under ideal conditions. However, real fish distributions may not follow such simplified patterns. As a result, VAST may have had an easier time detecting and modeling the effects of temperature and habitat covariates than it would in more complex, real-world scenarios. If species' environmental responses were more irregular or nonlinear, the biomass and/or trend error performance of covariate-based models like VAST could be different.

Our present analysis showed that both the design-based and model-based approaches are each capable of providing biomass estimates that approximate yearly biomass values as well as the overall biomass trend. VAST estimates that used covariates consistently provided the most accurate biomass and trend estimates compared to model-based estimates that did not use covariates and the design-based stratified mean. While VAST estimates without covariates included in the model tended to result in higher biomass error values compared to the stratified mean, the same estimates provided lower trend error values across all scenarios.

While we consistently see the lowest error values for both metrics when covariate information was included in the VAST models, there were also dozens of individual VAST estimates that had larger error values than the largest stratified mean error. As a result, VAST had the largest range in CV values for both metrics. These outcomes demonstrate the power of model-based estimates to provide biomass and trend estimates that are closer to the true values under certain conditions, while also highlighting how sensitive they are to the choices made regarding settings and data used in analyses. Users often rely on traditional model-building processes and diagnostic techniques such as Akaike information criterion (AIC) to guide their model selection choices, but successfully navigating the many decisions depends on the background of the given user and care should be taken in determining which diagnostic tools are used. For example, a recent simulation study showed that the existing diagnostic tools available in VAST sometimes guide users towards using settings that decrease the accuracy of the model [28]. The value of model-based approaches such as VAST are clear–they allow users to account for environmental variability and changes in catchability. However, the many consequential decisions that must be made in a given model and the lack of clear criteria for making these decisions makes it challenging to make these choices and even more difficult to evaluate whether the choices are correct.

Spatial factors contributed to many of the scenarios that produced larger error values seen in our estimates. For example, movement of fish from the shallow centrally located strata into the deeper, smaller, and less-sampled eastern strata contributed to VAST failing to accurately model the biomass of yellowtail without the addition of covariate information, the impact of which was exacerbated in the increasing temperature scenarios. While adding covariates can help inform the model and improve the estimate as seen with the yellowtail simulation results, covariates hindered the model in certain instances with cod and haddock, particularly when sampling did not span the full range of environmental conditions. This highlights a key limitation of relying on covariates in model-based approaches: when the covariate information in the data is incomplete, such as missing the coldest habitats for haddock, the model may extrapolate poorly, leading to misleading estimates. This underscores the importance of representative spatial coverage in surveys, especially when using covariate-driven models to inform stock assessments or project biomass into unsampled regions.

In our analysis of fitting a single biomass trend, the stratified mean had more consistent error metrics across scenarios compared to the model-based estimates, including under varying temperature gradients. This stability might make it preferable when precision and stability are priorities, even if it provides less accurate trend tracking compared to model-based methods. While there was a larger change in average error values between scenarios for VAST estimates that include covariates compared to the stratified mean, mean model-based estimates offered improved error metrics and reduced variance in error values in specific scenarios relative to the design-based and covariate-free approaches. This suggests that models with covariates can increase the precision of biomass and trend estimates under the right conditions. This emphasizes the value of covariate inclusion in model-based methods, but a full bias analysis would be required to confirm this preliminary results.

Biomass and trend errors were generally higher in scenarios with an increasing temperature gradient for all estimation methods. Results showed an especially large increase in VAST variance when covariates were excluded and these estimates provided higher average biomass error values than the stratified mean. Covariate inclusion in VAST was particularly beneficial under increasing temperature conditions, suggesting that model-based approaches that account for temperature variability can improve robustness in the face of dynamic environmental conditions.

As seen in Table S1, VAST models that included covariates used a linear combination of second degree polynomials for habitat and temperature to approximate species-specific covariate responses. We considered a best-case scenario by providing covariate information to VAST in the form of the exact habitat and temperature values that governed movement in our population simulations. This covariate information typically resulted in improved estimates, with 73% of VAST estimates with covariates providing a lower biomass error compared to the corresponding estimate without covariates, and 96% of VAST models with covariates resulting in a lower biomass error than the corresponding stratified mean estimate. Figure S7 shows covariate response plots estimated by VAST that reflect the true normally distributed temperature preferences used to drive MixFishSim movement for each species. In reality the covariates that influence a given species would be less clear and the collected values would contain sampling noise. One could test the impact of these differences by assuming the wrong covariate response function (e.g., linear vs polynomial) or by including each of the covariates individually. Including noise in the covariate data would further explore how robust the model-based estimates are to additional uncertainty in the covariate information.

Fall estimates, particularly for biomass, displayed substantially higher error values (up to a 238% increase in model-based estimates compared to spring), indicating a seasonal sensitivity that may reflect changing spatial distributions and abundances between seasons. For example, species in our *MixFishSim* simulations spawn during weeks 9–14, which coincides with the spring sampling weeks 13 and 14. Spawning in our population simulations concentrates each species in the spawning grounds towards the center of the domain where more samples occurred. The populations are more spread out during the fall survey in weeks 37 and 38 and since the water temperature is warmer during these months biomass shifts towards the deeper strata, which tend to be less sampled as they have a smaller area. These dynamics had a larger impact on VAST biomass error (3.38 times increase from spring to fall) compared to biomass and trend error for other estimates (1.5 time increase from spring to fall). The inclusion of covariates mitigated this variability to some extent, with spring estimates showing better accuracy and lower variance compared to fall estimates. This highlights the importance of including relevant biological and environmental drivers in the operating models that generate hypothetical biomass data when comparing design-based and model-based estimators.

Of the VAST models for which including covariates did not improve estimates, most provided a comparable error (e.g., 0.11 vs 0.13). The instances where including covariates provided a noticeable change in error took place in scenarios that included either an increasing temperature and/or reduced sampling domain. More specifically, when a reduced number of strata are sampled for haddock and cod, adding covariate information leads to an increase in biomass error for fall estimates. This is in contrast to our yellowtail results, for which adding covariate information to VAST models always decreased the biomass error in the resulting biomass estimates. Although strata were excluded for both species, the decline in performance for just cod and haddock can be explained by a failure to sample the full spectrum of temperature values where the species exists. For example, the strata excluded from sampling for yellowtail still allowed the survey to range across the species preferred values and thus provided a covariate response that improved the estimate (Figure 1). On the other hand, the specific strata that were excluded from sampling for haddock were located in the northern region of the domain, which means the covariate response for haddock was not informed by any data from the coldest temperature range. This could explain the failure to form a complete covariate response that not only degrades the estimate in the regions being sampled, but also confounds the ability to project estimates into an unsampled region. Figure S8 depicts the impact on biomass estimates of including a covariate response for haddock when both the average temperature and biomass were increasing over time, and biomass was projected into an unsampled strata. Projecting biomass into the unsampled strata provides the wrong magnitude of biomass and also a different trend (Figure S8).

Estimation methods that produce large variation between yearly estimates can potentially lead to changes in catch limits that do not correspond to the true population trend, which could have a compounding effect. For example, a large, increasing biomass estimate when the population has actually decreased and is fairly low could potentially lead to a windfall catch limit that further reduces the total biomass available the following year. A second overestimate the following year could then have a detrimental impact by reducing the population even further. Conversely, an overly smoothed estimator could miss true signals of change in the population and delay needed management response to either sudden increases or decreases in the population. Our *MixFishSim* population simulations assumed a constant total mortality that includes both fishing and natural death, and therefore will not account for impacts of such management decisions. This type of question can be best explored with a management strategy evaluation.

For monitoring population trends, we found that model-based methods with covariate information that actually drives spatial distribution and is measured without error appear advantageous, especially in seasonally or environmentally variable conditions. Since this is rarely achievable in practice, design-based methods may provide a more stable baseline in less variable conditions or when coverage is consistent. With enough computational power, one can use this modeling framework to rigorously quantify and analyze biases in biomass estimation methods, evaluate alternative survey designs, project climate-driven spatial distribution shifts, and ultimately strengthen fisheries management decisions in the face of ongoing environmental change.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

The authors would like to thank Joe Langan for helping to develop the proposal that funded this work. Funding was provided by the Northeast Fisheries and Climate program at the Northeast Fisheries Science Center. Thank you to Jim Thorson for consulting with us about VAST settings and implementation. We thank Chris Cacciapaglia and Charles Adams for the time they spent with us sharing their experiences using VAST. We appreciate the help from Robyn Linner for helping to extract FVCOM temperature data and are thankful for David Chevrier for meeting with us to clarify questions related to some of our GIS data.

Data and code availability

All data and code used in this work are available at https://github.com/Blevy2/READ-PDB-blevy2-MFS2.

Conflict of interest

The authors declare there is no conflict of interest.

References

1. *State of the Ecosystem Reports for the Northeast U.S. Shelf*, 2024. Avaiable from: https://www.fisheries.noaa.gov/new-england-mid-atlantic/ecosystems/state-ecosystem-reports-northeast-us-shelf.
2. National Marine Fisheries Service (NMFS), Fisheries Economics of the United States, 2022, *U.S. Department of Commerce, NOAA Tech. Memo*, 2024.

3. T. R. Azarovitz, A brief historical review of the Woods Hole Laboratory trawl survey time series, (1981), 62–67.

4. P. J. Politis, J. K. Galbraith, P. Kostovick, R. W. Brown, Northeast Fisheries Science Center bottom trawl survey protocols for the NOAA Ship Henry B. Bigelow, *Natl. Oceanic Atmos. Adm.*, (2014). http://doi.org/10.7289/V5C53HVS

5. J. C. Deville, C. E. Särndal, Calibration estimators in survey sampling, *J. Am. stat. Assoc.*, **87** (1992), 376–382.

6. S. D. Hoyle, R. A. Campbell, N. D. Ducharme-Barth, A. Grüss, B. R. Moore, J. T. Thorson, et al., Catch per unit effort modelling for stock assessment: A summary of good practices, *Fish. Res.*, **269** (2024), 106860. https://doi.org/10.1016/j.fishres.2023.106860

7. K. D. Baker, S. C. Anderson, D. R. J. Mullowney, W. Walkusz, K. R. Skanes, Moving away from a scale mismatch: Spatiotemporal modelling of striped shrimp (Pandalus montagui) density in Canada‘s subarctic, *Fish. Res.*, **270** (2024), 106898. https://doi.org/10.1016/j.fishres.2023.106898

8. A. C. Hansell, M. C. McManus, Integrating fisheries independent surveys to account for the spatiotemporal dynamics of spiny dogfish (Squalus acanthias) in US waters of the northwest Atlantic, *Fish. Res.*, **281** (2025), 107173. https://doi.org/10.1016/j.fishres.2024.107173

9. A. Grüss, R. L. O‘Driscoll, J. T. Thorson, J. R. McKenzie, S. L. Ballara, A. R. Charsley, Impacts of different types of data integration on the predictions of spatio-temporal models: A fishery application and simulation experiment, *Fish. Res.*, **284** (2025), 107321. https://doi.org/10.1016/j.fishres.2025.107321

10. J. T. Thorson, Guidance for decisions using the vector autoregressive spatio-temporal (VAST) package in stock, ecosystem, habitat and climate assessments, *Fish. Res.*, **210** (2019), 143–161. https://doi.org/10.1016/j.fishres.2018.10.013

11. J. T. Thorson, S. C. Anderson, P. Goddard, C. N. Rooper, TinyVAST: R package with an expressive interface to specify lagged and simultaneous effects in multivariate spatio-temporal models, *Global Ecol. Biogeography*, **34** (2025), e70035. https://doi.org/10.1111/geb.70035

12. J. A. BiogeographyNye, J. S. Link, J. A. Hare, W. J. Overholtz, Changing spatial distribution of fish stocks in relation to climate and population size on the Northeast United States continental shelf, *Mar. Ecol. Prog. Ser.*, **393** (2009), 111–129. https://doi.org/10.3354/meps08220

13. K. M. Kleisner, M. J. Fogarty, S. Fogarty, J. A. Fogarty, S. Moret, C. T. Perretti, et al., Marine species distribution shifts on the US Northeast Continental Shelf under continued ocean warming, *Prog. Oceanogr.*, **153** (2017), 24–36. https://doi.org/10.1016/j.pocean.2017.04.001

14. M. E. Henderson, K. E. Mills, A. C. Thomas, A. J. Pershing, J. A. Nye, Effects of spring onset and summer duration on fish species distribution and biomass along the Northeast United States continental shelf, *Rev. Fish Biol. Fish.*, **27** (2017), 411–424. https://doi.org/10.1007/s11160-017-9487-9

15. F. Arreguín-Sánchez, Catchability: A key parameter for fish stock assessment, *Rev. Fish Biol. Fish.*, **6** (1996), 221–242. https://doi.org/10.1007/BF00182344

16. J. A. Langan, G. Puggioni, C. A. Oviatt, M. E. Henderson, J. S. Collie, Climate alters the migration phenology of coastal marine species, *Mar. Ecol. Prog. Ser.*, **660** (2021), 1–18. https://doi.org/10.3354/meps13612
17. E. S. Klein, S. L. Smith, J. P. Kritzer, Effects of climate change on four New England groundfish species, *Rev. Fish Biol. Fish.*, **27** (2017), 317–338. https://doi.org/10.1007/s11160-016-9444-z
18. L. Kerr, M. Barajas, J. Wiedenmann, Coherence and potential drivers of stock assessment uncertainty in Northeast US groundfish stocks, *ICES J. Mar. Sci.*, **79** (2022), 2217–2230. https://doi.org/10.1093/icesjms/fsac140
19. M. D. Mazur, J. Jesse, S. X. Cadrin, S. B. Truesdell, L. Kerr, Consequences of ignoring climate impacts on New England groundfish stock assessment and management, *Fish. Res.*, **262** (2023), 106652. https://doi.org/10.1016/j.fishres.2023.106652
20. P. J. Dolder, C. Minto, J. Guarini, J. J. Poos, Highly resolved spatiotemporal simulations for exploring mixed fishery dynamics, *Ecol. Modell.*, **424** (2020), 109000. https://doi.org/10.1016/j.ecolmodel.2020.109000
21. J. T. Thorson, Predicting recruitment density dependence and intrinsic growth rate for all fishes worldwide using a data-integrated life-history model, *Fish Fish.*, **21** (2020), 237–251. https://doi.org/10.1111/faf.12427
22. Management track assessments fall 2022, *Northeast Fisheries Science Center (U.S.)*, 2022. Avaiable from: https://doi.org/10.25923/380j-t283.
23. I. D. Buller, Envi: Environmental interpolation using spatial kernel density estimation, 2022. Available from: https://github.com/lance-waller-lab/envi.
24. C. Chen, R. C. Beardsley, G. Cowles, An unstructured grid, finite-volume coastal ocean model: FVCOM user manual, *Oceanography*, **19** (2006), 78. https://doi.org/10.5670/OCEANOG.2006.92
25. L. J. Poppe, K. Y. McMullen, S. J. Williams, V. F. Paskevich, USGS east-coast sediment analysis: Procedures, database, and gis data, *US Geological Survey Open-File Report*, 2005. https://doi.org/10.3133/ofr20051001
26. P. Lafrance, M. Castonguay, D. Chabot, C. Audet, Ontogenetic changes in temperature preference of Atlantic cod, *J. Fish Biol.*, **66** (2005), 553–567. https://doi.org/10.1111/j.0022-1112.2005.00623.x
27. Vector autoregressive spatio-temporal (VAST) model wiki page, 2019. Available from: https://github.com/James-Thorson-NOAA/VAST/wiki.
28. C. Cacciapaglia, E. N. Brooks, C. F. Adams, C. M. Legault, C. T. Perretti, D. Hart, Developing workflow and diagnostics for model selection of a vector autoregressive spatiotemporal (VAST) model in comparison to design-based indices, *Fish. Res.*, **275** (2024), 107009. https://doi.org/10.1016/j.fishres.2024.107009

AIMS Press