

Atlantic Cod VAST Model Technical Document

Katie Lankowicz

09 June, 2023

Contents

1 Abstract	1
2 Introduction	2
3 Methods	3
3.1 Survey Data	3
3.2 VAST	3
3.3 Model selection	14
3.4 Final model	17
4 Results	18
4.1 Spring Index	18
4.2 Fall Index	22
4.3 Habitat associations	27
4.4 Model diagnostics	30
5 Discussion	33
References	34

1 Abstract

The spatial distribution of the Atlantic cod (*Gadus morhua*) population is shaped by a suite of habitat and oceanographic variables. Additional complexity is added by differences in environmental preferences between life-history phases and spatially-separated stocks. Multiple bottom trawl survey programs have been developed to assess groundfish abundance in the Northeastern US continental shelf (NEUS), most notably the Northeast Fishery Science Center's twice-annual bottom trawl survey. Groundfish indices of abundance are typically calculated using data from these bottom trawl surveys. However, bottom trawls have reduced efficiency in complex bottom habitats that are highly suitable for Atlantic cod. It would be advantageous to calculate an index of cod abundance from a combination of bottom trawl survey data and data from other survey platforms to help bridge this gap. In this study, Vector Autoregressive Spatio-Temporal (VAST) models were used to combine data from survey programs using bottom trawls, bottom longlines, and video trawls to hindcast seasonal spatial density of three size classes of cod within the NEUS

from 1982-2021. Bottom habitat characteristics (sediment type, rugosity, and depth), bottom water temperature, and basin-averaged climate indices (North Atlantic Oscillation, Atlantic Multidecadal Oscillation) were included as covariates to improve predictions of population density. Indices of abundance were generated for each size class within the four spatially distinct stock areas of the cod population. Abundance of all size classes generally decreased within all stock areas.

2 Introduction

Atlantic cod (*Gadus morhua*) are an ecologically, economically, and culturally critical part of New England's groundfish fishery. However, the fishery has been under a disaster declaration since 2012 due to historically low abundance and rapidly declining stocks. Atlantic cod population assessments and management efforts are informed by a suite of bottom trawl surveys, perhaps most notably by the Northeast Fishery Science Center's (NEFSC) twice-annual bottom trawl survey. This survey has been an important tool to assess groundfish stocks from Cape Hatteras, NC to Nova Scotia since the early 1960s. However, bottom trawl surveys like the NEFSC bottom trawl are known to have reduced efficiency over complex bottom habitats such as cobble fields and rocky ledges (Grabowski et al. 2020). Most bottom trawl survey programs have low sampling effort within complex habitat areas due to risks to the equipment, and will instead focus on sampling in areas deeper than 18m and with soft and smooth bottom habitat.

The limitation of survey information within more complex areas is concerning to fishing industry stakeholders and fishery managers alike; several life history phases of cod are known to associate with complex bottom habitat areas. Juvenile cod have been found in higher densities over cobble and boulder habitats, likely as a refuge from predation Linner and Chen (2022). Industry stakeholders have also reported relatively high density of large cod within hard-bottom habitats, possibly indicating a density-dependent reduction in large cod spatial distribution and altered availability to bottom trawl surveys (McElroy et al. 2021). Stakeholders' observations of high cod density over complex bottom habitat have created the perception that cod abundance is much higher than what bottom trawl-informed assessments have suggested.

The objective of this project is to build a joint index of abundance and annual maps of spatial density for Atlantic cod using all relevant state and federal groundfish survey data. Creating a joint index of cod abundance using survey data spanning inshore, offshore, smooth, and complex bottom habitats would improve our understanding of cod spatial dynamics and demographics, which in turn would benefit management efforts and help rectify opposing perceptions of cod population size. These alternative data sources do exist, albeit typically with much shorter time series than the NEFSC bottom trawl survey. Many state and academic institutions conduct bottom trawl surveys that may help to fill data gaps for groundfish stocks in shallower inshore waters. The NEFSC has also launched a cooperative bottom longline survey to explore groundfish abundance and distribution in areas with high-complexity bottom habitat.

This research objective will be addressed using Vector Autoregressive Spatio-Temporal (VAST) models. VAST models estimate spatial density of multiple categories of a target organism conditioned on density covariates and controlling for catchability covariates. These estimates of spatial density can then be used in the calculation of indices of abundance. We will utilize multiple trawl surveys and a suite of habitat and environmental covariates to explore patterns in habitat use and create indices of abundance for Atlantic cod.

3 Methods

3.1 Survey Data

Eleven surveys of groundfish abundance are currently available for use. Overall spatial coverage runs along the coast from Lubec, Maine to Cape Hatteras, North Carolina and spans the four defined cod spatial stock areas: Eastern Gulf of Maine, Western Gulf of Maine, Georges Bank, and Southern New England. Overall temporal coverage spans from 1959 - 2022. Included surveys are:

- NEFSC Cooperative Bottom Longline Survey (longline)
- Eastern Gulf of Maine Sentinel Survey (jigging)
- SMAST Video Trawl Survey (combined trawl-camera)
- NEFSC Bottom Trawl Survey (trawl)
- DFO Bottom Trawl Survey (trawl)
- University of Rhode Island GSO Trawl Survey (trawl)
- Massachusetts DMF Industry-based survey (trawl)
- Massachusetts DMF Inshore Trawl Survey (trawl)
- Maine-New Hampshire Inshore Trawl Survey (trawl)
- Rhode Island DEM Trawl Survey (trawl)
- ASMFC Shrimp Trawl Survey (trawl)

These surveys utilize different gears and methods, have variable spatial extents, are completed in either annually or seasonally, and are of variable temporal length (as in, number of years in which the survey has occurred). All surveys report number of cod and total weight (kg) of cod caught per tow, and some surveys process all or a portion of the catch to provide further biological detail.

3.2 VAST

VAST models were used to estimate cod spatial density over time and create joint indices of abundance using all available groundfish survey data. VAST is a framework for implementing spatial delta-generalized linear mixed models (delta-GLMM) and can be manipulated to provide estimates for multiple categories of interest and spatial strata Thorson (2019). It is structured to utilize two linear predictors; the first linear predictor estimates encounter probability, and the second linear predictor estimates catch rates. The first linear predictor can be represented as

$$\rho_1(i) = \beta_1(c_i, t_i) + \omega_1^*(s_i, c_i) + \varepsilon_1^*(s_i, c_i, t_i) + \eta_1(v_i, c_i) + \nu_1(c_i, t_i)$$

where $\rho_1(i)$ is the predictor for observation i for category c_i at location s_i and time t_i . $\beta_1(c_i, t_i)$ represents temporal variation for each category and time, $\omega_1(s_i, c_i)$ represents spatial variation for each location and category, $\varepsilon_1(s_i, c_i, t_i)$ represents spatiotemporal variation for each location, category, and time, $\eta_1(v_i, c_i)$ represents vessel effects for each vessel and category, and $\nu_1(c_i, t_i)$ represents the effect of density covariates for each category and time. The second linear predictor is structured the same way. Both linear predictors incorporate fixed and random effects, and spatial and spatiotemporal variation are approximated as Gaussian Markov random fields Thorson (2019).

Implementation of VAST models requires several structural and data inclusion decisions, as outlined in Thorson (2019). Decisions used in this effort will be discussed in the following subsections.

3.2.1 Spatial, temporal, and spatiotemporal effects

Spatial, temporal, and spatiotemporal autocorrelation can be included in both linear predictors. A model selection process was used to justify the use of spatial and spatiotemporal random effects in the first and second linear predictors. The intercept for each linear predictor was defined as a fixed effect for each time step—this ensures independent estimates of abundance for each time step, which is most appropriate for creating abundance indices (Thorson 2019). Instead, a temporal component was estimated for the spatiotemporal components in both linear predictors. This is recommended for indices generated by multiple data sources that do not necessarily sample the same locations in every time step (Thorson 2019). Without this estimation, unrealistic “hot spots” may develop or be carried through the time series when this is inappropriate. The use of a temporal component for spatiotemporal components can also help interpolate density in unsampled time steps (Thorson 2019).

3.2.2 Spatial extent

The NEFSC bottom trawl survey is conducted through Cape Hatteras, North Carolina, but the southern edge of the Atlantic cod range likely does not extend to this point. For the purposes of this model, the spatial extent of interest will be American waters on the continental shelf from the northern edge of the Gulf of Maine through the mouth of the Chesapeake Bay (Fig. 1).

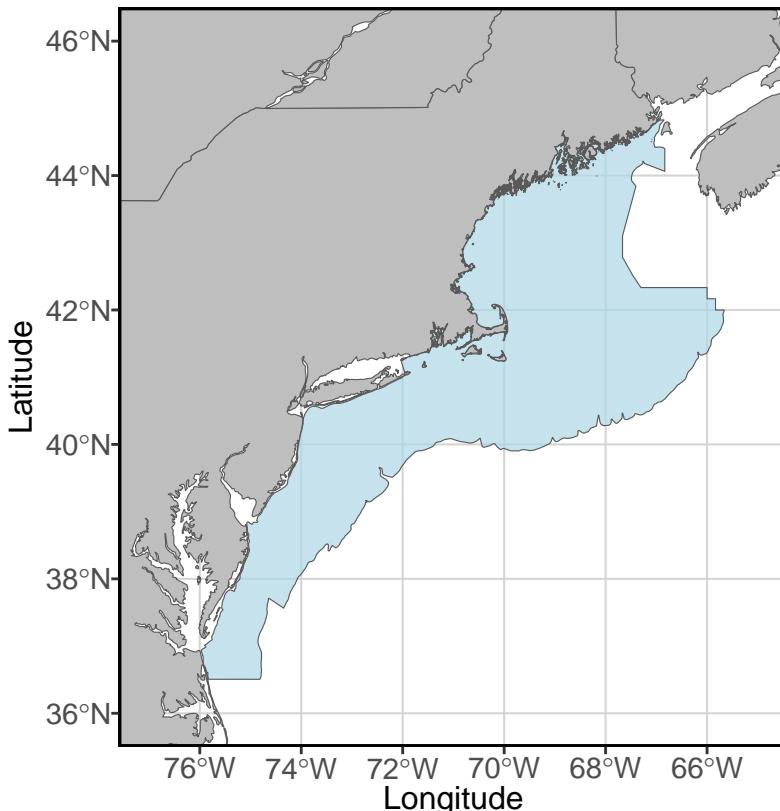


Figure 0: Spatial extent for cod density models

3.2.3 VAST spatial referencing

It has been said that selection of the number of knots and spatial smoothers in VAST is “an art.” For now, 200 knot locations will be used to make the SPDE grid on which VAST is run. Knot locations are determined by k-means clustering

of the data, and are meant to represent standardized locations optimally placed among all surveyed locations. The model then creates an extrapolation grid. This grid is built on an SPDE approximation to a Matérn correlation function. For this model, anisotropy was expected and therefore included, though its inclusion was assessed in the model selection process. The default number of grid cells (2000) was used in the model, which results in cells that are approximately 25km by 25km.

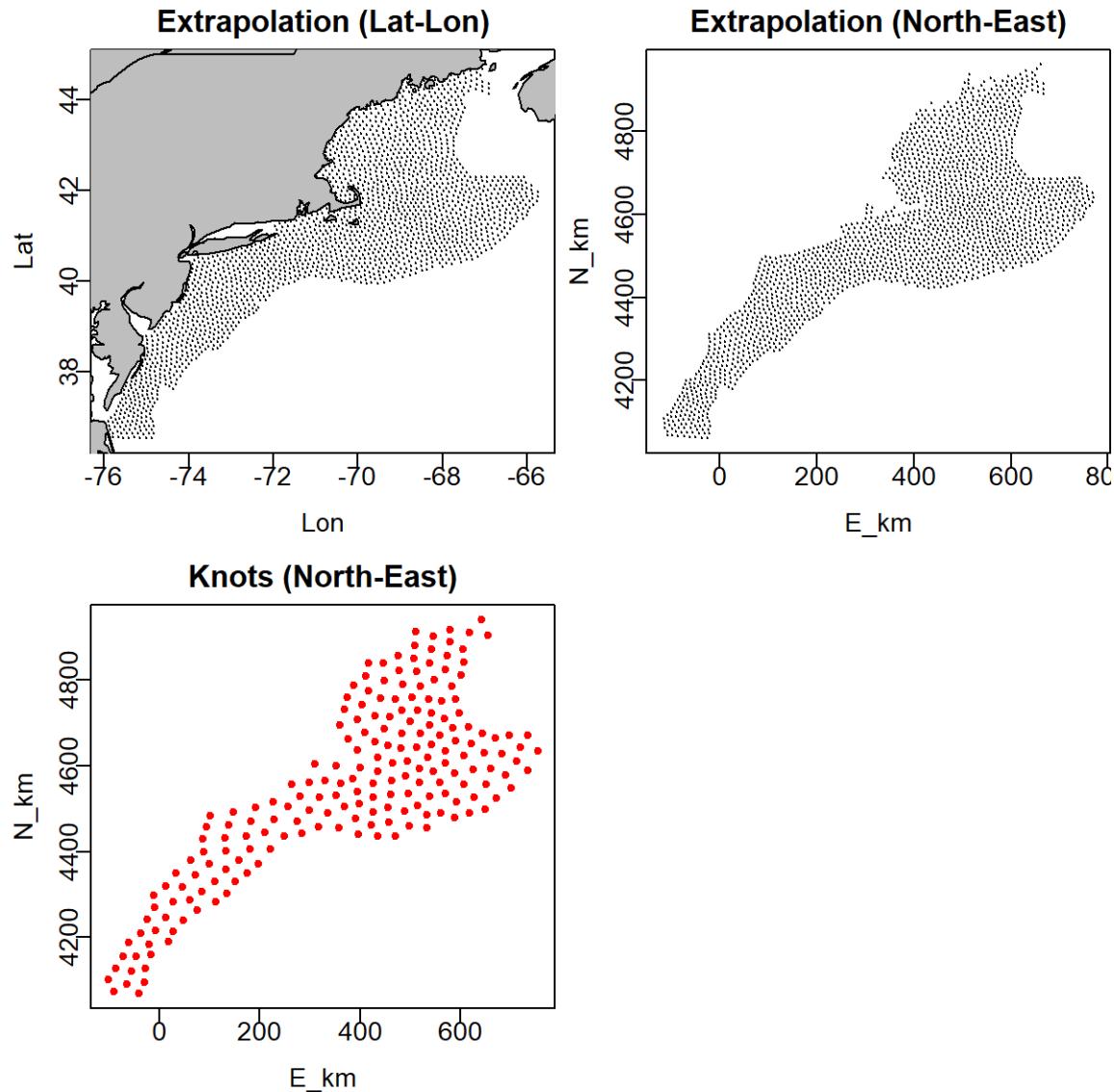


Figure 1: Location of extrapolation points and knots within the model spatial domain

3.2.4 Strata

There is evidence that cod spatial density varies among what are known to be distinct cod stock spatial areas (Linner and Chen 2022). The model will therefore be run with each stock area as separate spatial strata (Fig. 2). Metrics of spatial dynamics, indices of abundance, and habitat associations can later be described for each stock area.

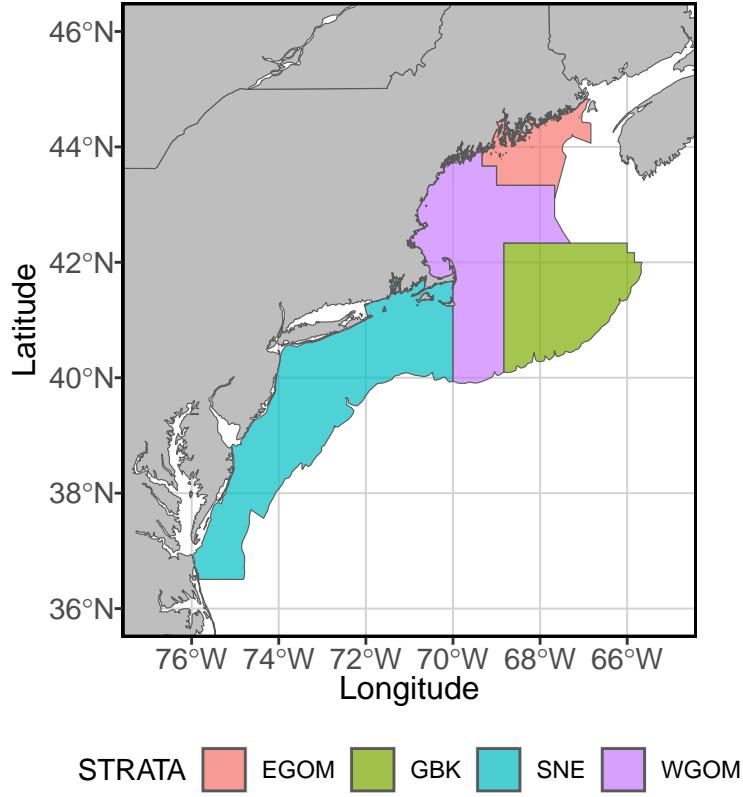


Figure 2: Cod stock areas within the model spatial domain

3.2.5 Temporal extent

Though some surveys were initiated in the 1950s and 60s, the time period of interest to this study begins in 1982. Some studies have provided full data from 2022, and others have not. For this reason, the modeled time period will be 1982 - 2021.

3.2.6 Seasonality

Time steps in the model will represent the spring and fall seasons of each year in the time series. Therefore, though there are 40 years of data there will be 80 time steps. This distinction is made to capture the seasonal spawning movements of adult cod. The spring season will be March through August of any year x . Fall will be September through December of year x and January and February of year $x + 1$. This is to ensure that the fall season is temporally continuous, rather than interrupted by the defined spring period. Categories Cod were modeled as three separate size classes, as it was expected that habitat utilization and spatial density would vary among size/ age groups. Small fish were defined as shorter than 39.1 cm total length or less than 0.58 kg. Medium fish were between 39.1 and 70.2 cm, or between 0.58 and 3.44 kg. Large fish were longer than 70.2 cm or heavier than 3.44 kg. This size structure roughly matches to ages 0-2 (pre-spawning), ages 2-5 (spawning), and ages 5+ (mature) cod, though this age structure was determined across all years and all spatial areas of the model domains and may be slightly different among time periods and locations. Uncommonly, fish without biological information could still be sorted into size categories. This typically occurred when the total weight of all fish caught in a tow was less than the maximum size limit of a singular small fish– in this instance, all fish were assumed to belong in the small size category. Fish unable to be assigned to a

size category could not be used in the model.

3.2.7 Data type, link functions, and distributions

VAST models can use biomass, abundance, or encounter data as the response variable to quantify presence at each sampling location. In this model, abundance data were used; not every survey provided biological data including length and weight for every fish sampled, and therefore not every fish could be sorted into a size category by weight as would be necessary for a biomass-data model. As recommended by model developers for abundance-data models, a lognormal-Poisson distribution was chosen for the first linear predictor and an alternative “Poisson-link delta-model” was chosen for the second linear predictor.

3.2.8 Area swept

VAST requires an input for the area swept for each tow. For surveys using bottom trawl methods, area swept was calculated as the distance covered by the tow multiplied by the wing spread. It is recommended that area swept be set to 1 for sampling gears with an unknown effective area swept. Initially, area swept for the bottom longline and Sentinel jigging surveys was set to 1. However, this created an issue of scaling and mixed units. Simple leave-one-out sensitivity tests were run to determine the influence of the bottom longline and Sentinel jigging surveys on the overall cod indices of abundance. Removing the Sentinel jigging survey had little impact on the modeled abundance of all three size classes of cod, and so it was excluded from further analysis. Removing the bottom longline survey reduced the abundance of medium and large cod by up to 50% in some years, and was therefore retained in further analyses. The description of the bottom longline motivation and methods in McElroy et al. (2019) states that it was developed to as closely as possible match the sampling effort of the NEFSC bottom trawl survey. For this reason, average area swept of the NEFSC bottom trawl survey was also used as the input for the area swept of the bottom longline survey.

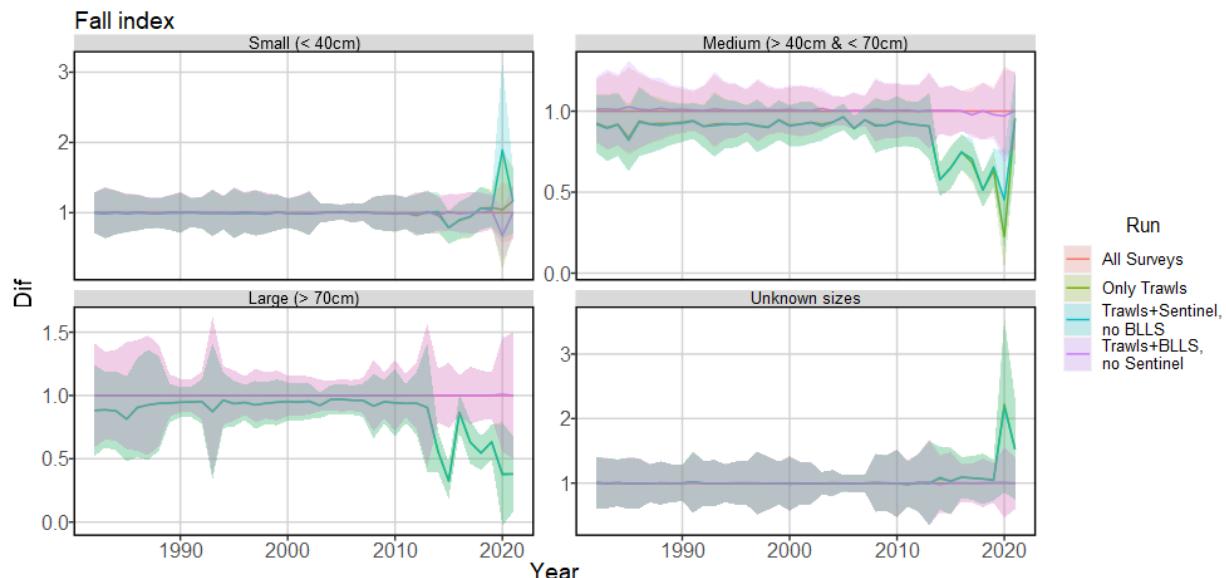


Figure 3: Proportional comparison of fall cod abundance by age group, using model runs with all available data as the control.

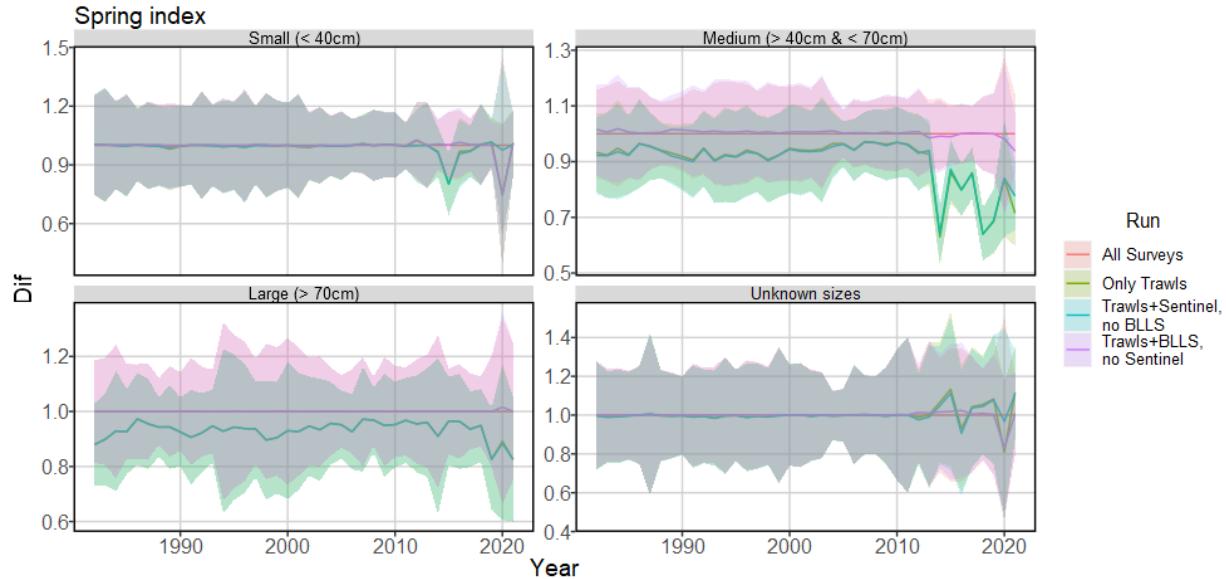


Figure 4: Proportional comparison of spring cod abundance by age group, using model runs with all available data as the control.

3.2.9 Vessel effects

Each survey used in this modeling effort has its own set of sampling protocols. It is highly likely that this variability also causes variability in catchability. This random variation in catchability among levels of a grouping variable is referred to as vessel effects in the VAST model structure. For this model, vessel effects were included. It is understood that multiple vessels were sometimes used to complete each survey, but some surveys did not specify the vessel used for each sampling event. Therefore, we simply used survey as the grouping variable rather than vessel. VAST models covariation in vessel effects with a factor model, where variation in catchability between groups is a random effect.

3.2.10 Density Covariates

VAST allows for the effects of both density and catchability covariates to be included in modeling efforts. Catchability covariates are processes one would expect to affect the ability to observe the target organism without necessarily affecting the distribution of the organism. Density covariates are processes that directly affect the distribution of the target organism, regardless of ability to observe it. Both covariates affect the catch rate of the target organism, but only density covariates are used to predict target organism density within the spatial domain. Therefore, VAST “controls for” catchability covariates and “conditions on” density covariates. VAST is unable to distinguish whether potential covariates should be treated as catchability or density covariates; this must be decided with theoretical insight from an analyst.

In this model, catchability covariates were not used. As mentioned previously, differences in sampling design were included as vessel effects.

The model will include the effects of several potential density covariates. The final model will only include density covariates with significant impact, as determined by a model selection process that will be discussed later in this document.

3.2.10.1 Sediment data Probability of mud, sand, gravel, cobble, and rock substrate existing in 1km by 1km grid cells through the entire continental shelf was provided by Brad Harris and Felipe Restrepo at Alaska Pacific University. These data are a expansion on the New England Fishery Management Council Swept Area Seabed Impact (SASI) model, which previously provided irregular polygons of distinct sediment grain types within the same spatial extent. Sediment observations from historical USGS sediment grabs and UMass Dartmouth's School for Marine Science and Technology (SMAST) video survey program are interpolated using an ordinary kriging approach to create the final 1km by 1km grid of sediment type probability.

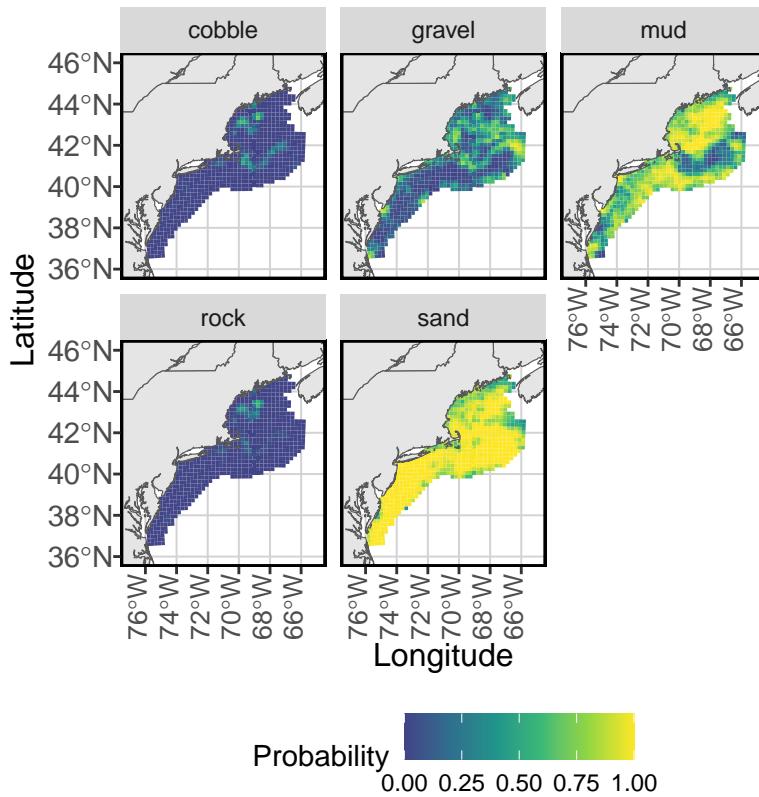


Figure 5: Likelihood of each of five sediment types across the n

3.2.10.2 Rugosity Rugosity data comes from a dataset created by Ariell Friedman (Friedman et al. 2012). Rugosity is a unitless measure of elevation change over horizontal distance, and is calculated for grid cells with a 15 arc-second resolution.

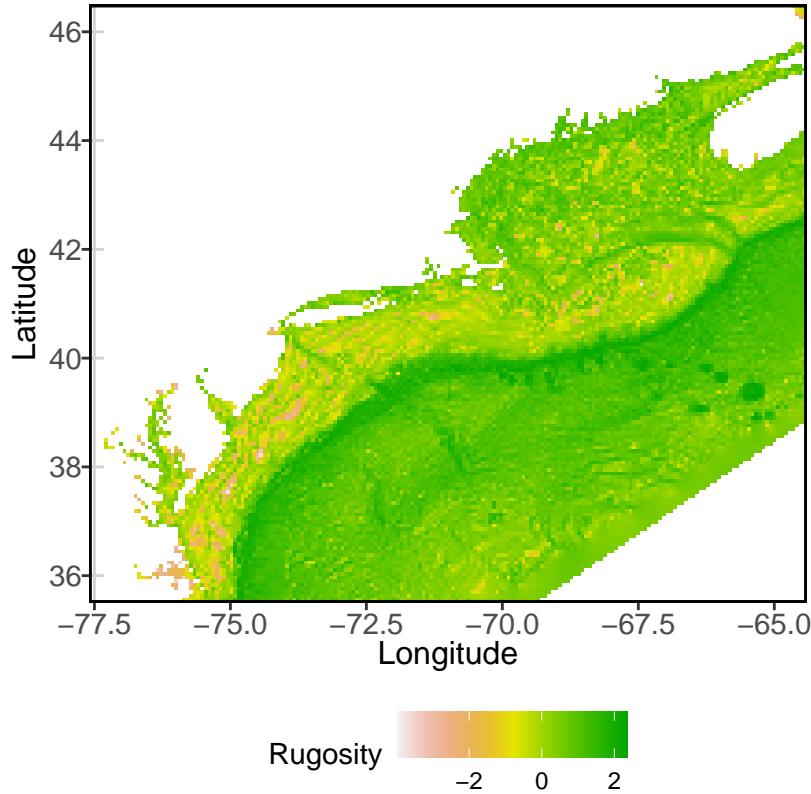


Figure 6: Rugosity within the model spatial domain

3.2.10.3 Bathymetry Most surveys recorded average depth at which the various gears were deployed. Some surveys did not provide information regarding gear depth. VAST models cannot have missing values in the density covariate inputs, so this inconsistency was an issue. Therefore, instead of using survey-reported values, depth at all survey locations was interpolated from GEBCO 15 arc-second bathymetry.

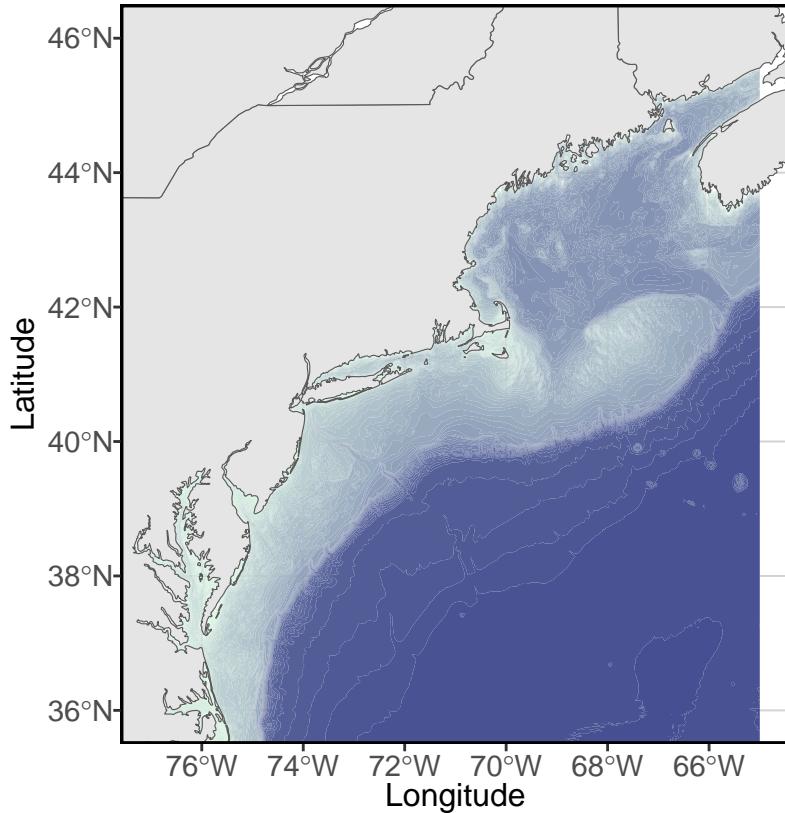


Figure 7: Bathymetry within the model spatial domain

3.2.10.4 Sea Surface Temperature The previous density covariates were temporally stationary. The following measures of temperature are temporally dynamic. For sea surface temperature, OISST values were pulled from NOAA data sources and extracted at observation locations. OISST values were compared to field measurements, when available, and the two were found to be generally similar. The following plot describes overall differences between field-measured SST and OISST when both measures were available for the same observation.

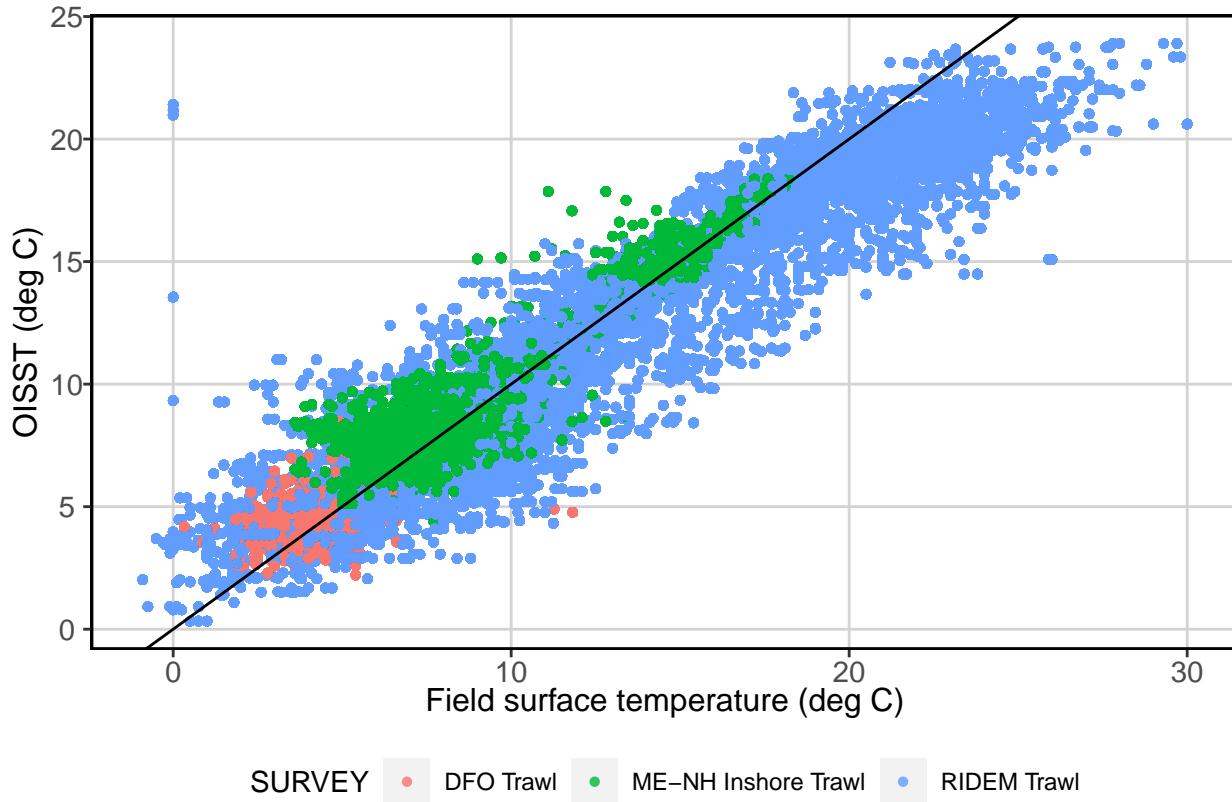


Figure 8: Comparison of field-measured SST to OISST

3.2.10.5 Bottom Temperature Bottom temperature data were provided by Du Pontavice et al. (2023). Bottom temperature within approximately 5' by 5' grid cells were calculated at a daily timestep for 1982 to 2020. Because cod observation data were collected through 2021, bottom temperature data for this year were simply copied from 2020. This may not be accurate, and it may be better to exclude 2021 data in future efforts. Further, the bottom temperature product did not extend to the inshore strata included in the model. To fix this issue, bottom temperature was interpolated to the inshore strata from the existing data.

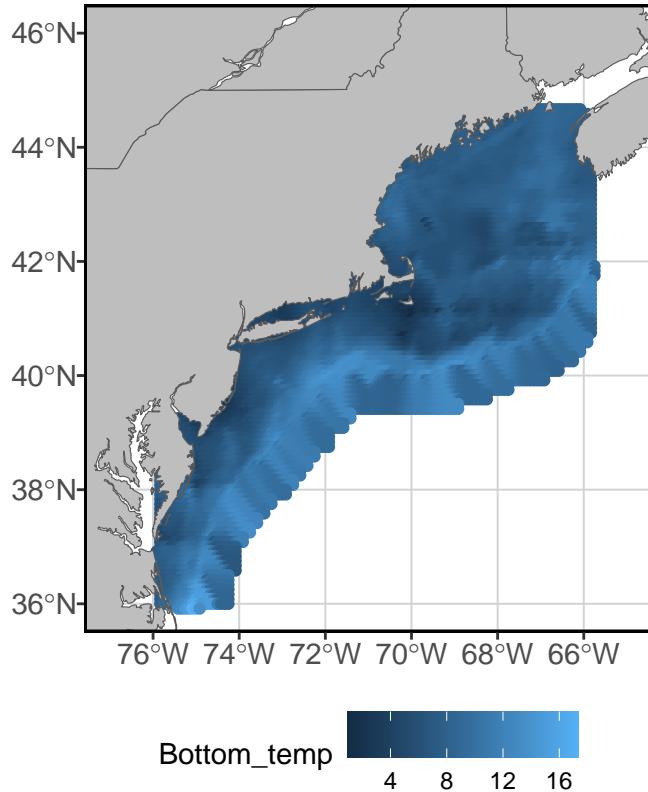


Figure 9: Bottom temperature within the model spatial domain

3.2.10.6 Climate Indices Several climate indices have been demonstrated to have a relationship to cod abundance and distribution. For this model, NAO and AMO data were used as climate indices and tested for inclusion in the final model. NAO data are provided at a daily timestep within the model, whereas AMO was provided at a monthly timestep. Both were basin-wide indices, meaning the value does not change between spatial locations within the region.

3.2.11 Density Covariate Collinearity

Potential density covariates were tested for collinearity. High correlations were found between two pairs: cobble and rock probability within 1km grid cells, and OISST and bottom temperature. Probability of rock and OISST were removed as covariates. For the first pair, rock was removed as bottom sample grabs (the bulk of sediment samples that make up the dataset) are more likely able to accurately characterize cobble distribution than larger boulders. For the second pair, OISST was removed as cod are groundfish and therefore more likely to be affected by bottom temperature than surface temperature. Once these two collinear relationships were addressed, the remainder of density covariates were not correlated and therefore were included in the model.

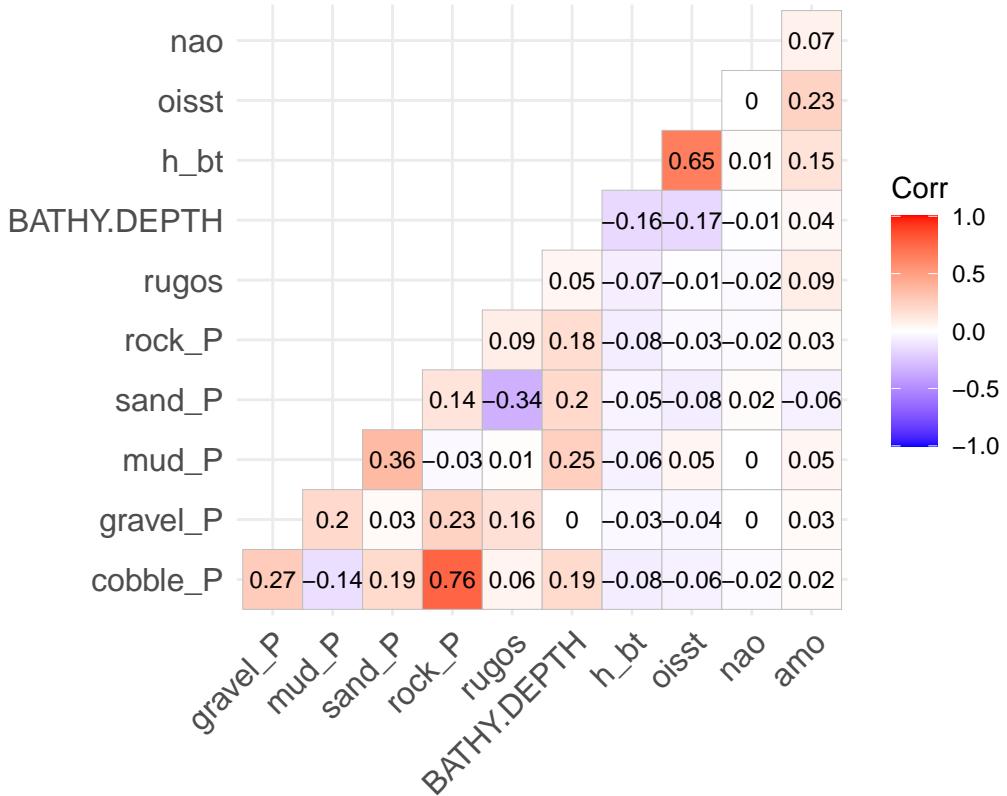


Figure 10: Density covariate correlation matrix

3.3 Model selection

The model selection process was completed in two steps. The first step compared models with and without anisotropy and/or spatial and spatiotemporal random effects in the first and second linear predictors. This follows the model selection process of Ng et al. (2021). The second step was to compare models with all possible combinations of the nine density covariates

3.3.1 Model settings selection

Model settings were explored using model selection. Models were tested for different values in the Field Configuration settings (testing the use of spatial and spatiotemporal variation in the linear predictors), anisotropy, and overdispersion. AIC was used to compare models and restricted maximum likelihood (REML) was used in model construction to make comparison via AIC possible (Zuur et al. 2009).

Model naming conventions are as follows:

- “alleffctson” - anisotropy on, spatial and spatiotemporal random effects in both linear predictors
- “noaniso” - anisotropy off, spatial and spatiotemporal random effects in both linear predictors
- “noomeps2” - anisotropy and spatial and spatiotemporal effects in linear predictor 1
- “noomeps2_noaniso” - anisotropy off, spatial and spatiotemporal effects in linear predictor 1
- “noomeps2_noeps1” - anisotropy and spatial effects in linear predictor 1
- “noomeps2_noeps1_noaniso” - anisotropy off, spatial effects in linear predictor 1

- “noomeps12” - anisotropy on, no spatial or spatiotemporal effects
- “noomeps12_noaniso” - anisotropy off, no spatial or spatiotemporal effects

The best model included anisotropy and spatial and spatiotemporal effects in both linear predictors.

Table 1: (#tab:settings selection)Table 1: Model selection step 1 outcomes

Model name	delta AIC	Fixed coefficients	Random coefficients	AIC	converged
alleffectson	0.000000e+00	9	41790	60128.32	likely
noaniso	2.407274e-01	7	41790	60128.56	likely
noomeps2_noaniso	1.720459e+03	4	20973	61848.78	likely
noomeps2	1.720765e+03	6	20973	61849.08	likely
noomeps2_noeps1	3.448233e+03	5	413	63576.55	likely
noomeps2_noeps1_noaniso	3.451895e+03	3	413	63580.21	likely
noomeps12	1.283181e+04	1	156	72960.13	likely
noomeps12_noaniso	1.283181e+04	1	156	72960.13	likely

3.3.2 Density covariate selection

Density covariates were selected for inclusion into the final model by running a subset of the original observation data (last 5 years of NEFSC bottom trawl data) with all possible combinations of variables as spatial density inputs. The four sediment likelihood variables (cobble, gravel, sand, and mud) were never separated. Models were then compared using AIC. Maximum likelihood was used to calculate AIC, as the spatial and spatiotemporal random effects did not change between models.

Model naming conventions are as follows:

- “seds” - all sediment likelihood variables (cobble, gravel, sand, and mud)
- “rugos” - rugosity
- “bathy” - bathymetry
- “bottom” - bottom temperature
- “nao” - daily North Atlantic Oscillation index
- “amo” - monthly North Atlantic Oscillation index

The model with the lowest AIC was the model that included spatial density terms for all potential covariates.

Table 2: Table 2a: Model selection step 2 outcomes

Model name	delta AIC	Fixed coefficients	Random coefficients	AIC
sedsrugsbathybottomnaomo	0.0000000	47	5632	31827.73
sedsrugsbathybottommamo	0.2068173	45	5632	31827.94
sedsrugsbathybottomnao	0.7153067	45	5632	31828.44
sedsrugsbathybottom	1.4555360	43	5632	31829.18
sedsrugsbathynaoamo	24.2643819	45	5632	31851.99
sedsrugsbathyamo	24.6986828	43	5632	31852.43
sedsbathybottomnaomo	29.2902682	45	5632	31857.02
sedsbathybottommamo	29.6062683	43	5632	31857.34
sedsbathybottomnao	30.1573439	43	5632	31857.89
sedsrugsbathynao	30.4622757	43	5632	31858.19
sedsbathybottom	31.0567709	41	5632	31858.79
sedsrugsbathy	31.8155721	41	5632	31859.54
sedsbathynaoamo	52.4577598	43	5632	31880.19
sedsbathyamo	53.0261733	41	5632	31880.76
sedsbathynao	58.7423222	41	5632	31886.47
sedsbathy	60.2936041	39	5632	31888.02
rugosbathybottommamo	62.3041648	37	5632	31890.03
rugosbathybottomnaomo	63.0030174	39	5632	31890.73
rugosbathybottom	63.9483351	35	5632	31891.68
rugosbathybottomnao	64.1454668	37	5632	31891.87
bathybottommamo	84.2510758	35	5632	31911.98
bathybottomnaomo	84.7596911	37	5632	31912.49
rugosbathyamo	85.4995550	35	5632	31913.23
rugosbathynaoamo	85.9832816	37	5632	31913.71
bathybottomnao	86.0862582	35	5632	31913.82
bathybottom	86.1266561	33	5632	31913.86
rugosbathynao	92.5446440	35	5632	31920.27
rugosbathy	92.9041665	33	5632	31920.63
bathyamo	106.5076105	33	5632	31934.24
bathynaoamo	106.7738774	35	5632	31934.50
bathynao	113.4711624	33	5632	31941.20
bathy	114.1115364	31	5632	31941.84
sedsrugsbottomnao	249.0554839	43	5632	32076.78
sedsrugsbottomnaomo	249.7008586	45	5632	32077.43
sedsrugsbottom	250.3797688	41	5632	32078.11

Table 3: Table 2b: Model selection step 2 outcomes

Model name	delta AIC	Fixed coefficients	Random coefficients	AIC
sedsrugosbottomamo	250.5617	43	5632	32078.29
sedsbottomnao	269.8470	41	5632	32097.58
sedsbottomnaoamo	270.2860	43	5632	32098.01
sedsbottomamo	271.1444	41	5632	32098.87
sedsbottom	271.2075	39	5632	32098.94
sedsrugosnaoamo	295.7166	43	5632	32123.45
sedsrugosamo	297.0664	41	5632	32124.80
sedsrugosnao	301.6872	41	5632	32129.42
sedsrugos	304.0014	39	5632	32131.73
sedsnaoamo	314.6790	41	5632	32142.41
sedsamo	316.0543	39	5632	32143.78
sedsnao	320.8598	39	5632	32148.59
seds	323.2514	37	5632	32150.98
rugosbottom	406.3812	33	5632	32234.11
rugosbottomnao	406.5037	35	5632	32234.23
rugosbottomamo	407.0379	35	5632	32234.77
rugosbottomnaoamo	407.5113	37	5632	32235.24
bottomnao	423.8988	33	5632	32251.63
bottom	423.9570	31	5632	32251.69
bottomamo	424.3848	33	5632	32252.11
bottomnaoamo	424.7147	35	5632	32252.44
rugosamo	460.7255	33	5632	32288.45
rugosnaoamo	460.7792	35	5632	32288.51
rugosnao	466.5960	33	5632	32294.32
rugos	467.3466	31	5632	32295.08
naoamo	476.5718	33	5632	32304.30
amo	476.6922	31	5632	32304.42
nao	482.6383	31	5632	32310.37
base	483.6222	29	5632	32311.35

3.4 Final model

After the model selection process, the best model included anisotropy, spatial, and spatiotemporal random effects in both linear predictors and density covariates for four classes of sediment types, rugosity, bathymetry, bottom temperature, NAO index, and AMO index. The model was structured with 200 knots and a Matern process spatial smoother to create 2000 grid cells of approximately 25km by 25km to extrapolate spatial density of three size classes of Atlantic cod in both the spring and fall seasons of each year from 1982 - 2021. The final model was also run with bias correction turned on, which uses the “epsilon method” to ensure that the mean and variation of generated indices of abundance are not biased due to their transformation by a nonlinear function in the modeling process (Thorson and Kristensen 2016).

4 Results

4.1 Spring Index

Large cod had the lowest abundance within the study area of all three size classes, and abundance for this group decreased over time. This result held for indices calculated within each of the four stock areas as well. Large cod were most abundant in the Georges Bank stock area. The center of gravity for large cod tended to be within the Georges Bank stock area and shifted slightly offshore over the time series. Effective area occupied decreased dramatically over time.

Medium cod had higher abundance than large cod, but also had a decreasing abundance trend over time. Similar to large cod, this result held when separating indices of abundance into each of the stock areas. Medium cod were most abundant in the Western Gulf of Maine stock area. The center of gravity for medium cod was typically on the border between the Western Gulf of Maine stock area and the Georges Bank stock area, and moved slightly offshore over time. Effective area occupied decreased over time.

Small cod had high variability through the time series, and seemed to have a jump in recruitment at around 2003. There is no clear trend in small cod abundance over time, though there may have been a slight increasing trend since 2003. Small cod are most abundant in the Southern New England stock area. The center of gravity for small cod was typically in either Southern New England or Western Gulf of Maine waters around Cape Cod, though occasional strong abundance years for coastal Maine pulled the center of gravity dramatically northwards. There is no clear trend for changes in effective area occupied over time.

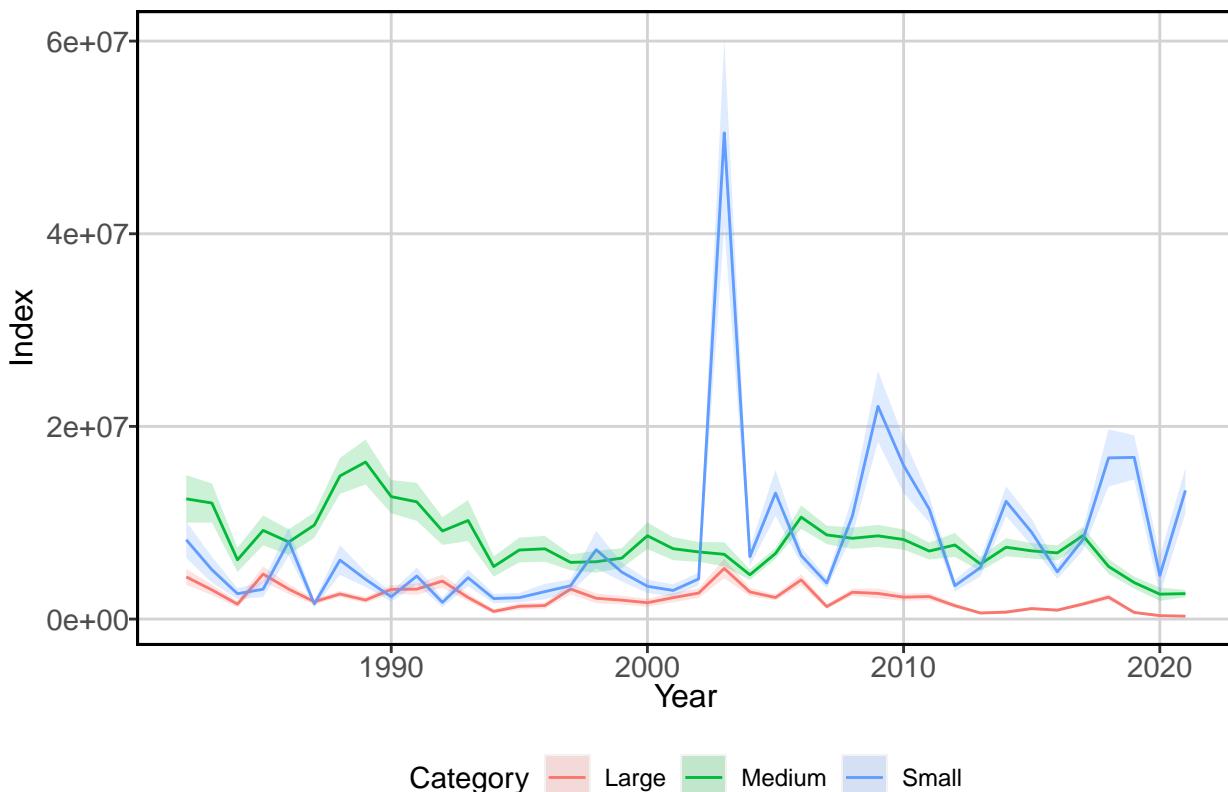


Figure 11: Spring indices of abundance for all size classes

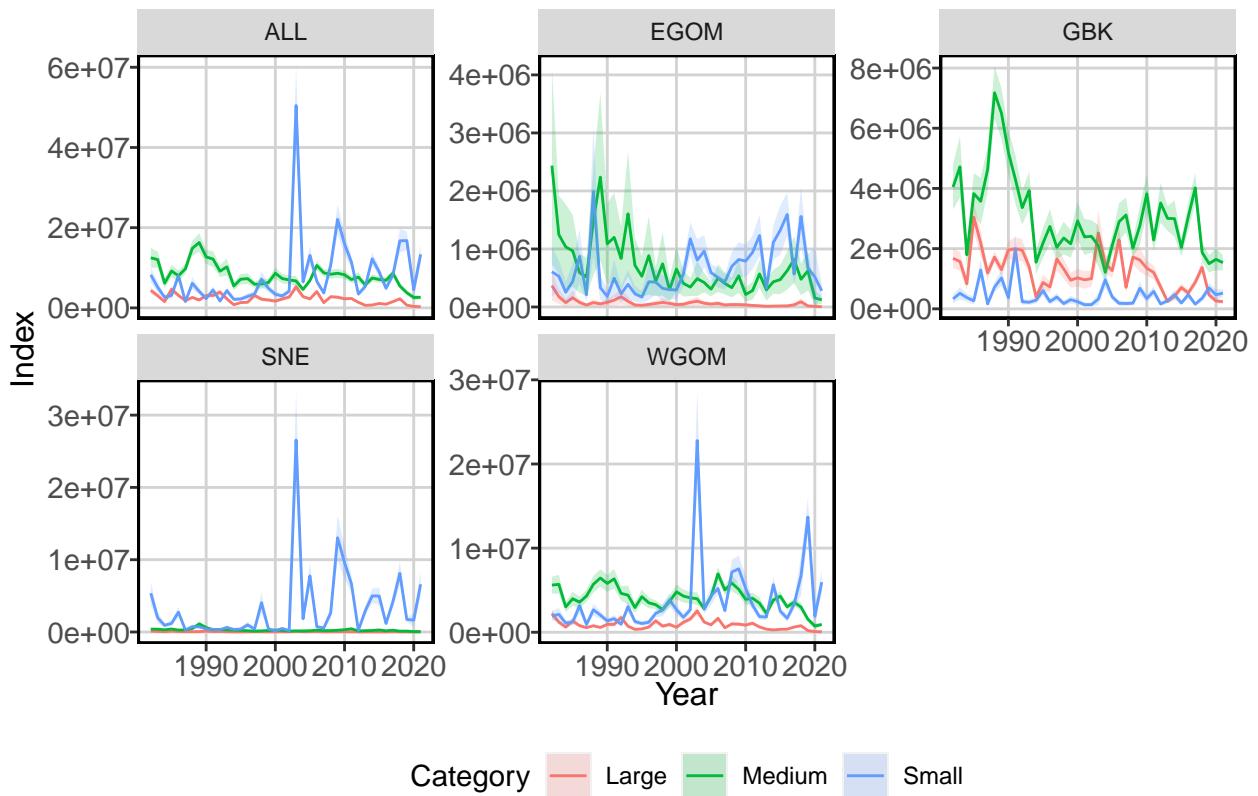


Figure 12: Spring indices of abundance within each stock area for all size classes

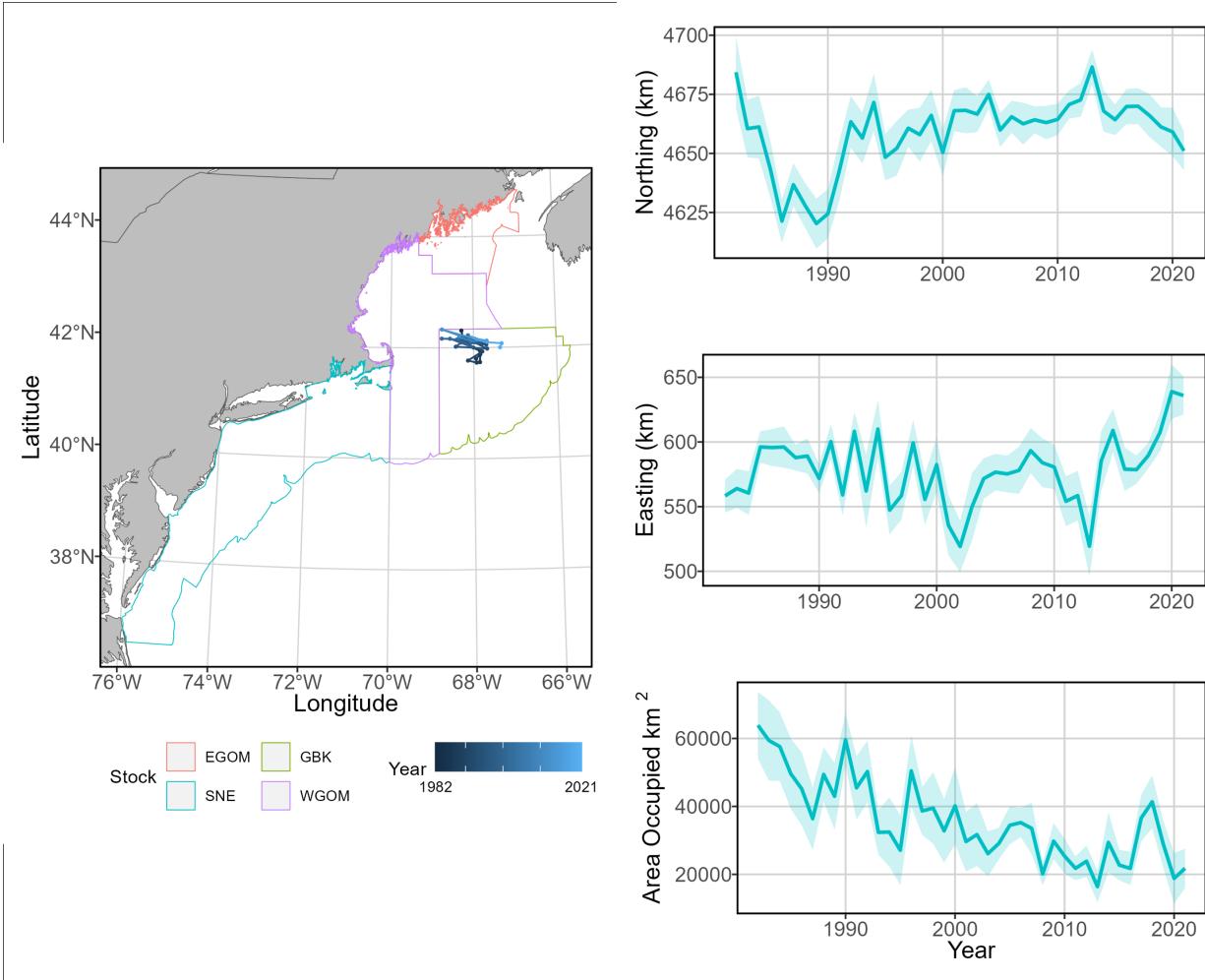


Figure 13: Large cod center of gravity and effective area occupied for all stock areas, spring seasons 1982 - 2021.

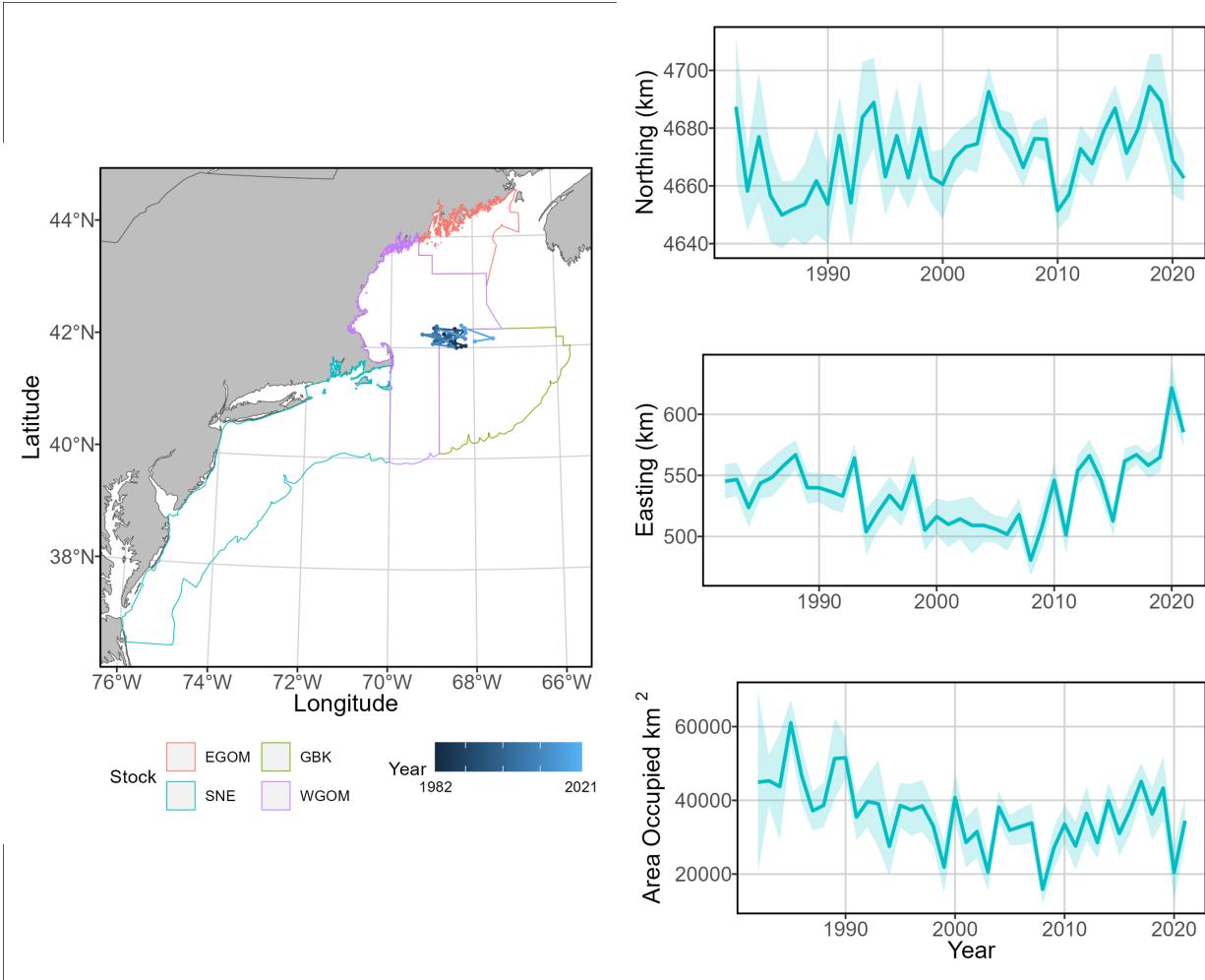


Figure 14: Medium cod center of gravity and effective area occupied for all stock areas, spring seasons 1982 - 2021.

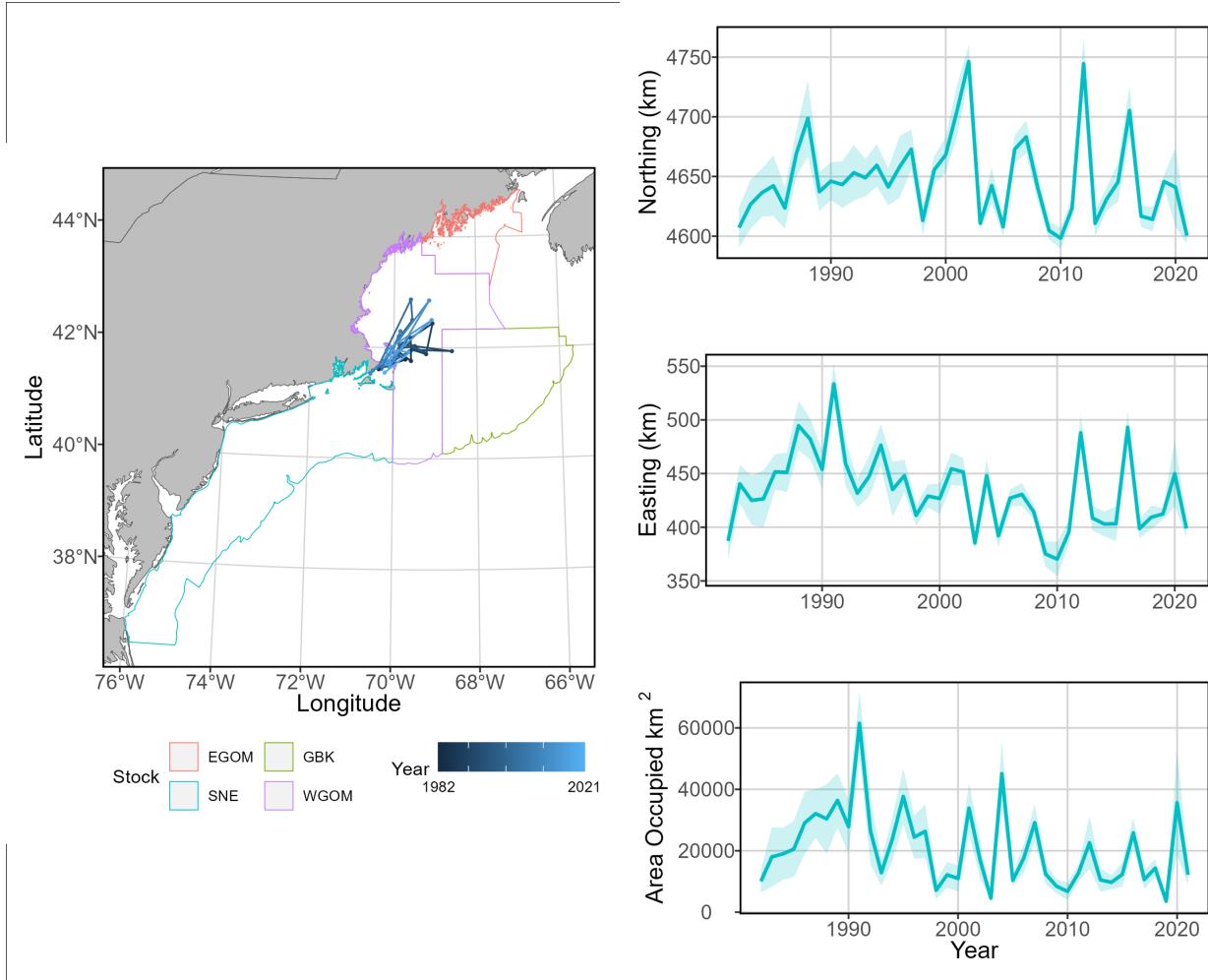


Figure 15: Small cod center of gravity and effective area occupied for all stock areas, spring seasons 1982 - 2021.

4.2 Fall Index

Large cod usually had the lowest abundance within the study area of all three size classes, and abundance for this group decreased over time. This result held for indices calculated within each of the four stock areas as well. Large cod were most abundant in the Georges Bank stock area. The center of gravity for large cod tended to be between the Georges Bank and Western Gulf of Maine stock areas and shifted slightly north over the time series. Effective area occupied decreased dramatically over time.

Medium cod had the highest fall abundance of all size classes, but also had a decreasing abundance trend over time. Similar to large cod, this result held when separating indices of abundance into each of the stock areas. Medium cod were most abundant in the Western Gulf of Maine stock area. The center of gravity for medium cod was typically on the border between the Western Gulf of Maine stock area and the Georges Bank stock area, and moved slightly north over time. Effective area occupied decreased over time.

Unlike the spring indices, small cod did not have high variation in the fall indices. Small cod abundance had a decreasing trend throughout the time but had a rapid fall in abundance in the Southern New England and Western Gulf of Maine stock areas starting around 2010. The center of gravity for small cod was in the Western Gulf of Maine stock area, and

there is no clear trend in movement over time. There is also no clear trend for changes in effective area occupied over time.

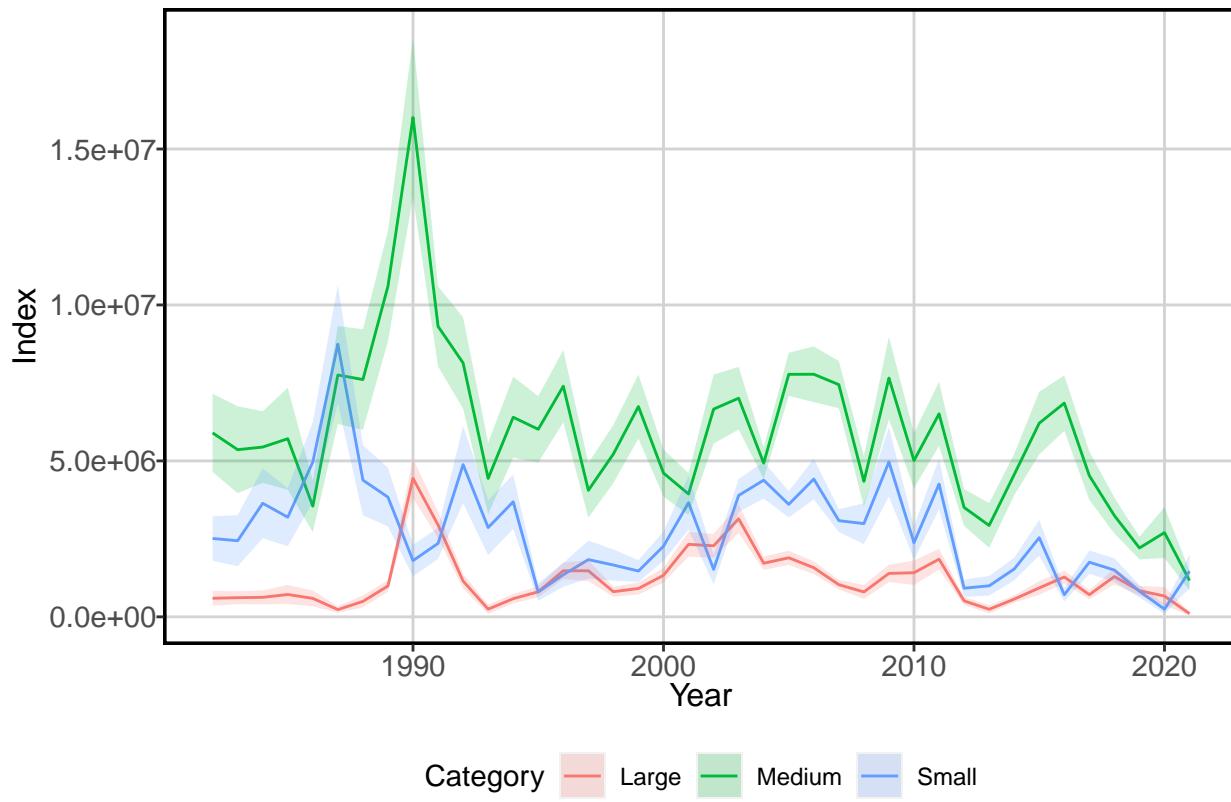


Figure 16: Fall indices of abundance for all size classes

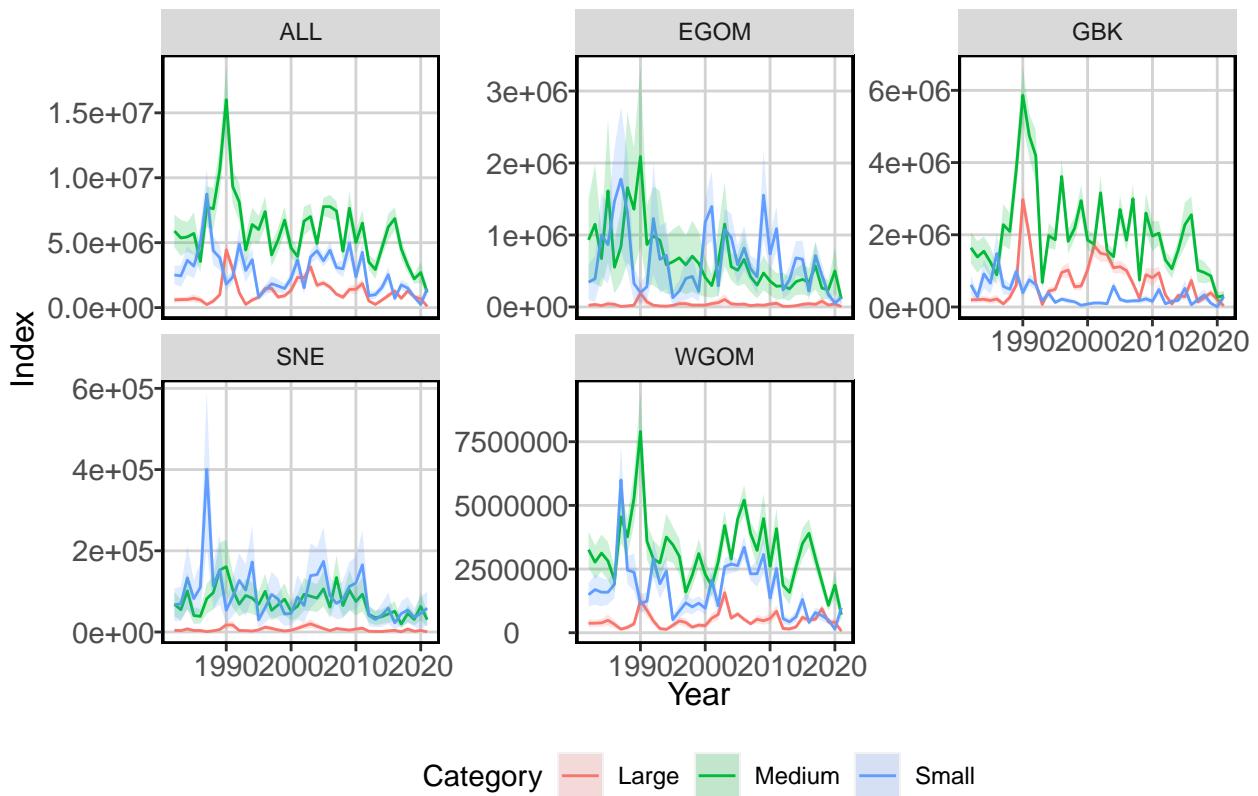


Figure 17: Fall indices of abundance within each stock area for each size class

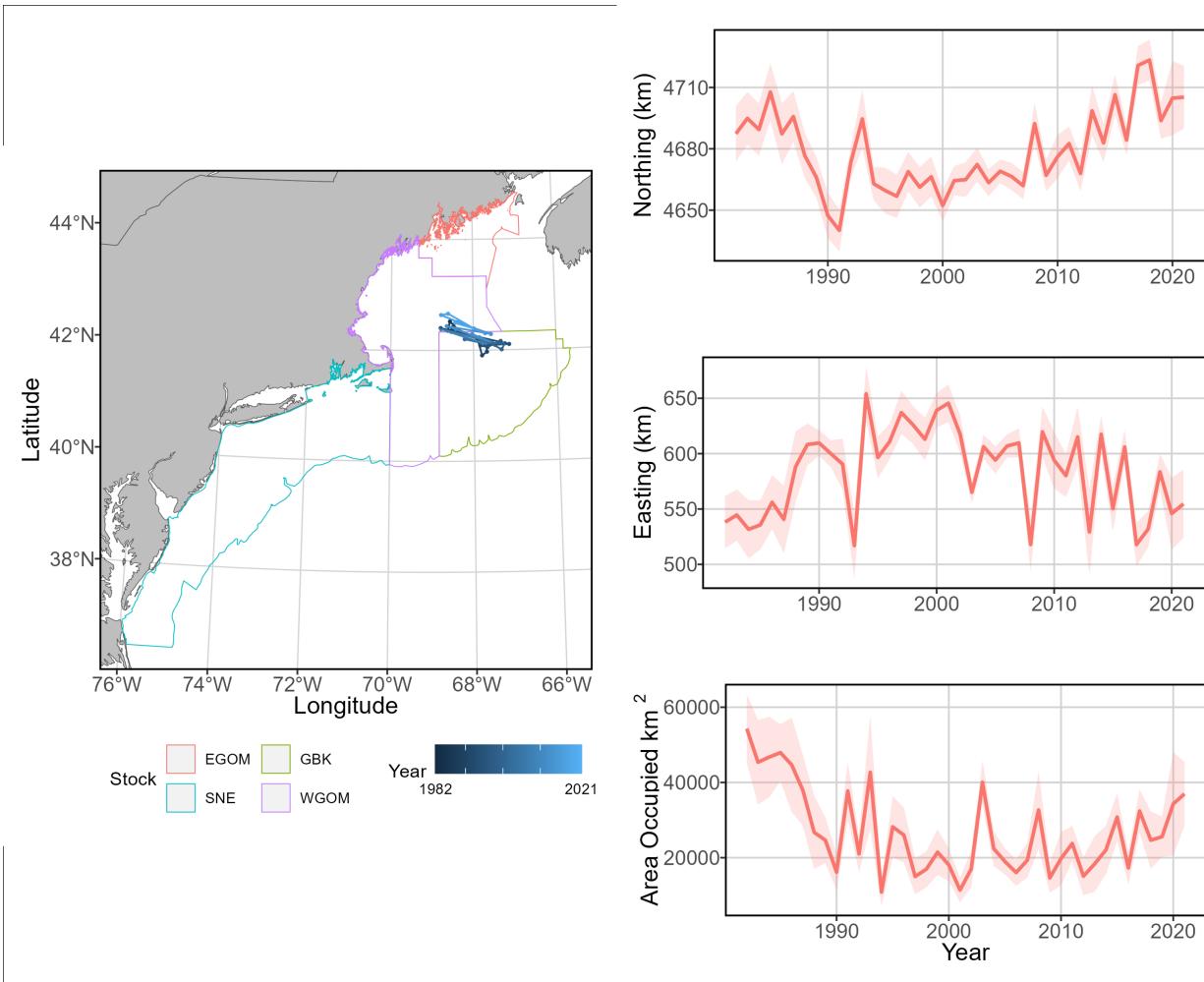


Figure 18: Large cod center of gravity and effective area occupied for all stock areas, fall seasons 1982 - 2021.

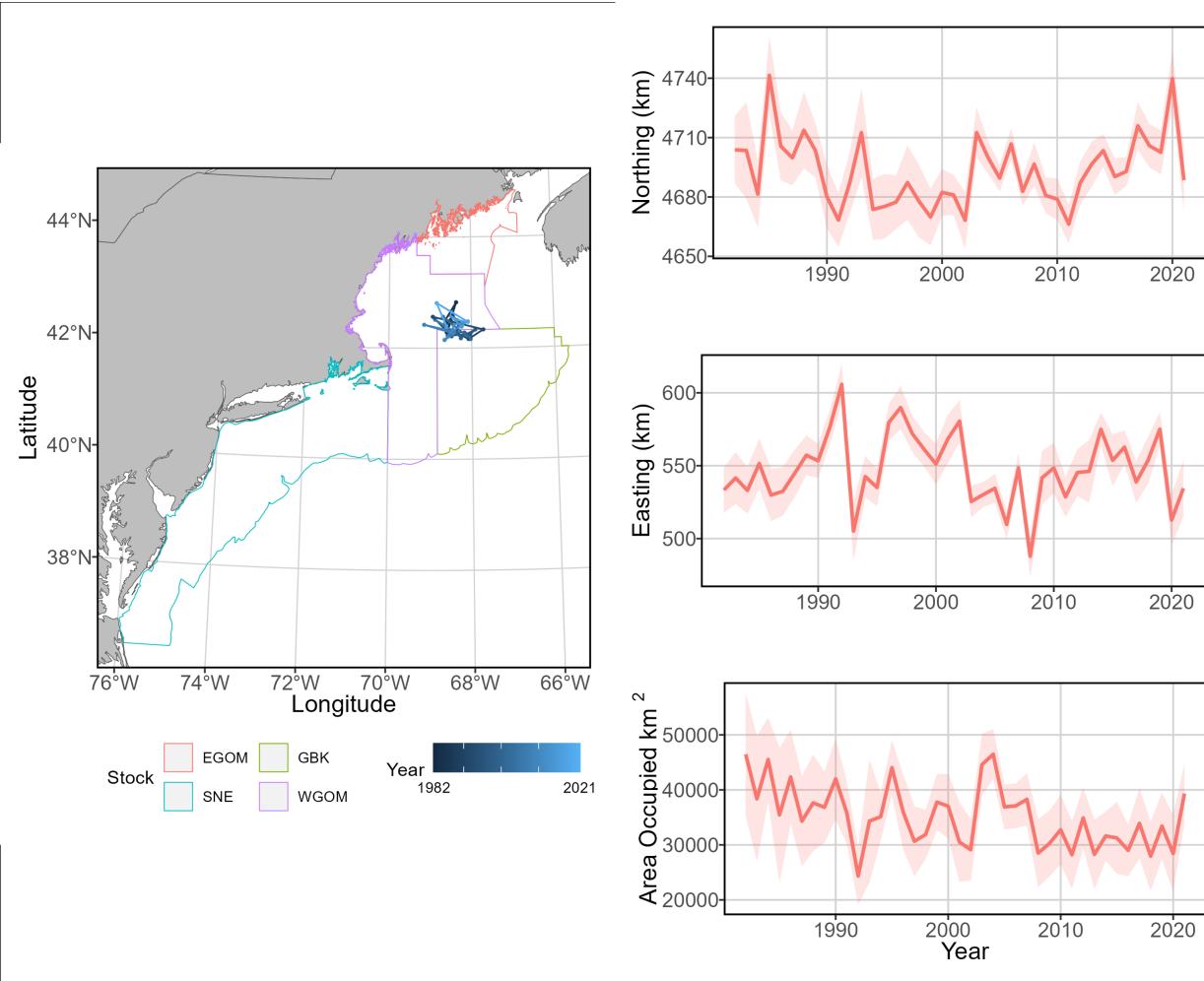


Figure 19: Medium cod center of gravity and effective area occupied for all stock areas, fall seasons 1982 - 2021.

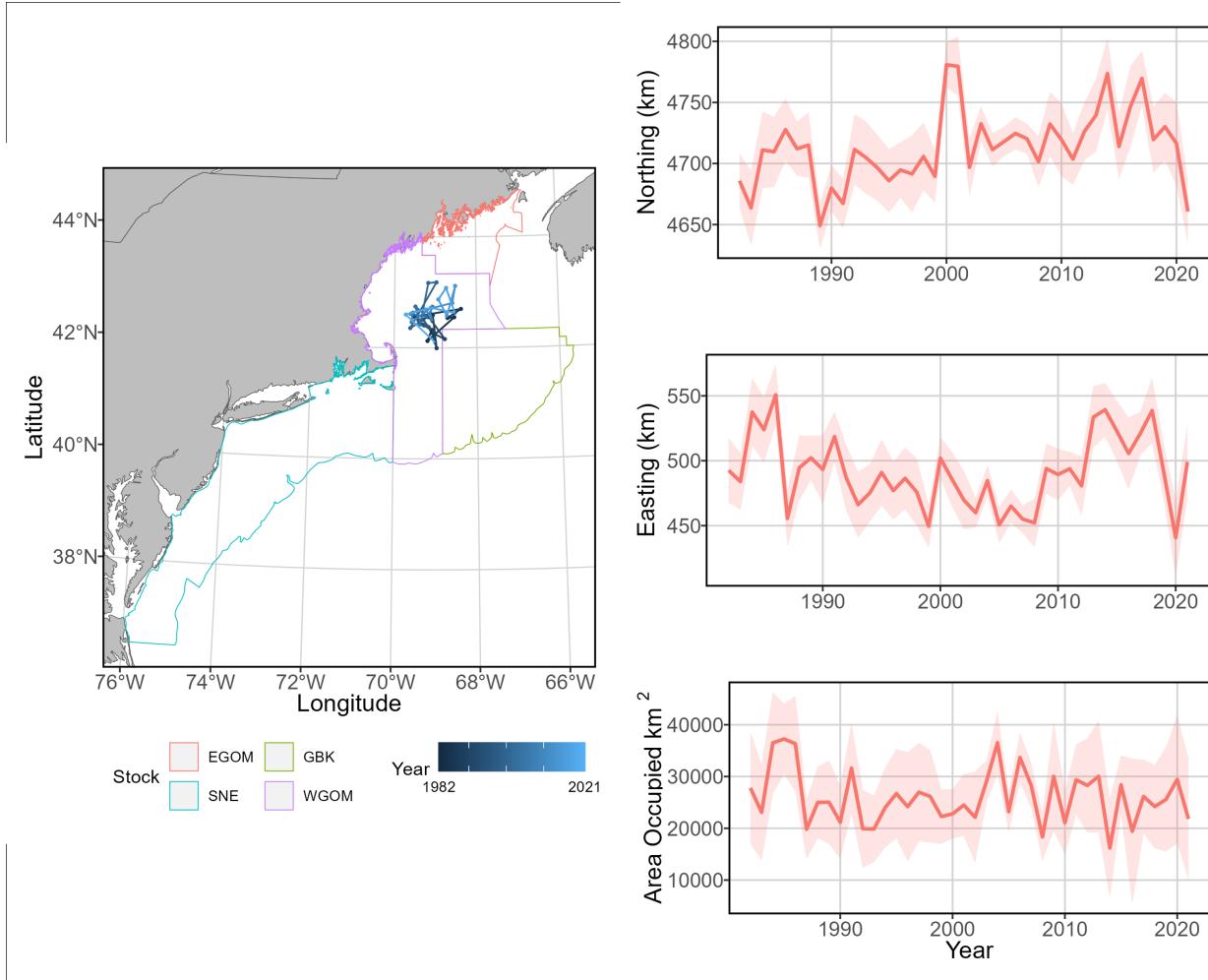


Figure 20: Small cod center of gravity and effective area occupied for all stock areas, fall seasons 1982 - 2021.

4.3 Habitat associations

After modeling, VAST extrapolation grid cells were stratified into hard, mixed, and soft bottom types using k-means clustering. Soft-bottom grid cells most likely contained only mud or sand, mixed-bottom cells contained similar likelihoods of mud, sand, and gravel, and hard-bottom grid cells likely contained cobble. The resulting map of sediment “hardness” was plotted with the aggregate spatial density of each cod size class for the spring and fall time series.

Small cod were highly concentrated in nearshore waters around Martha’s Vineyard and Narragansett Bay during the spring season, and more concentrated in nearshore waters around Boston Harbor in the fall. They were not strongly associated with any sediment type, though areas of high abundance in the fall were typically in areas of mixed sediment. Medium cod were more generally spread throughout the Gulf of Maine, and were typically found in highest density in areas of mixed sediment type. Large cod were highly dense in much smaller spatial areas. In the spring, this density was concentrated at the eastern end of Georges Bank and within areas of mixed sediment type. In the fall, large cod density was highest in the waters around the Western Gulf of Maine closed area and within habitats of either mixed or hard sediment type.

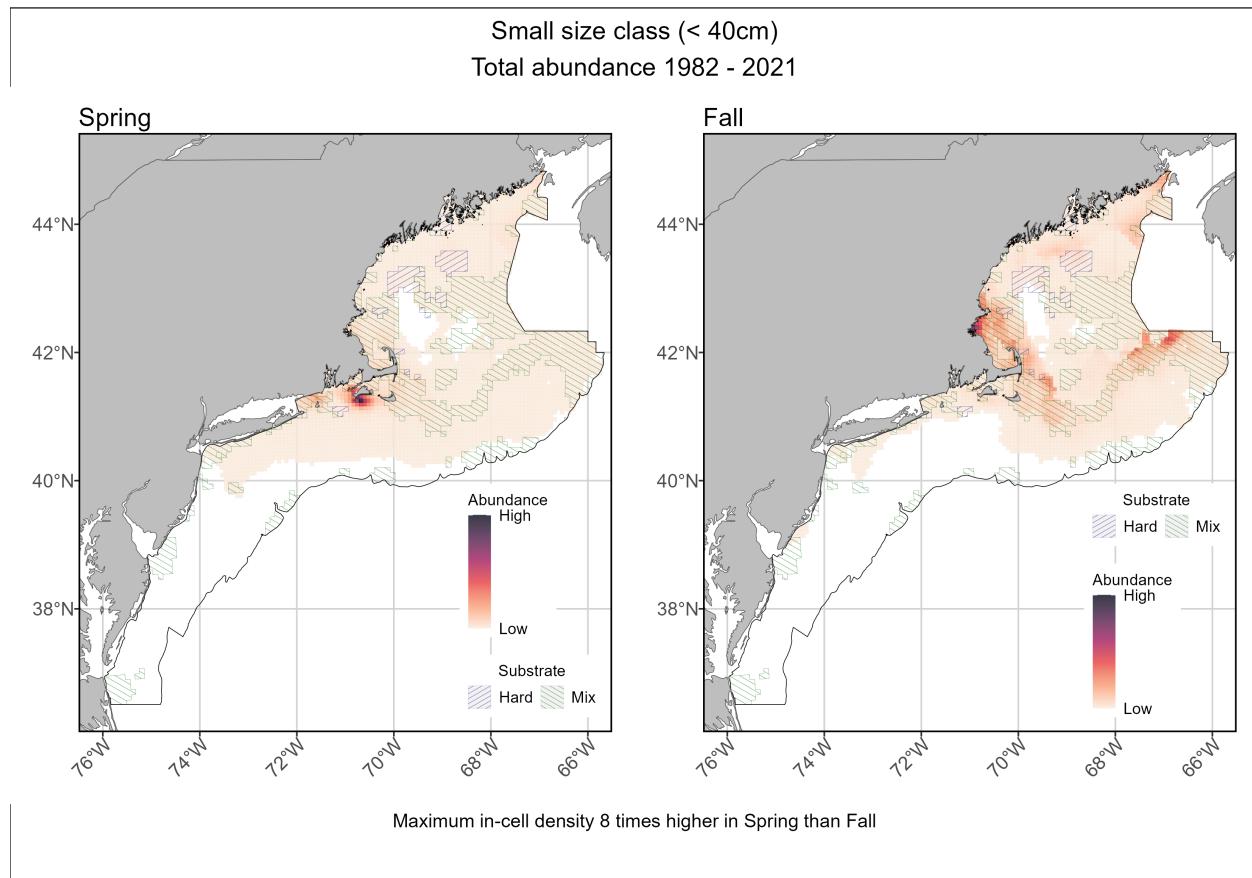


Figure 21: Small cod aggregate spatial density split by season, 1982 - 2021.

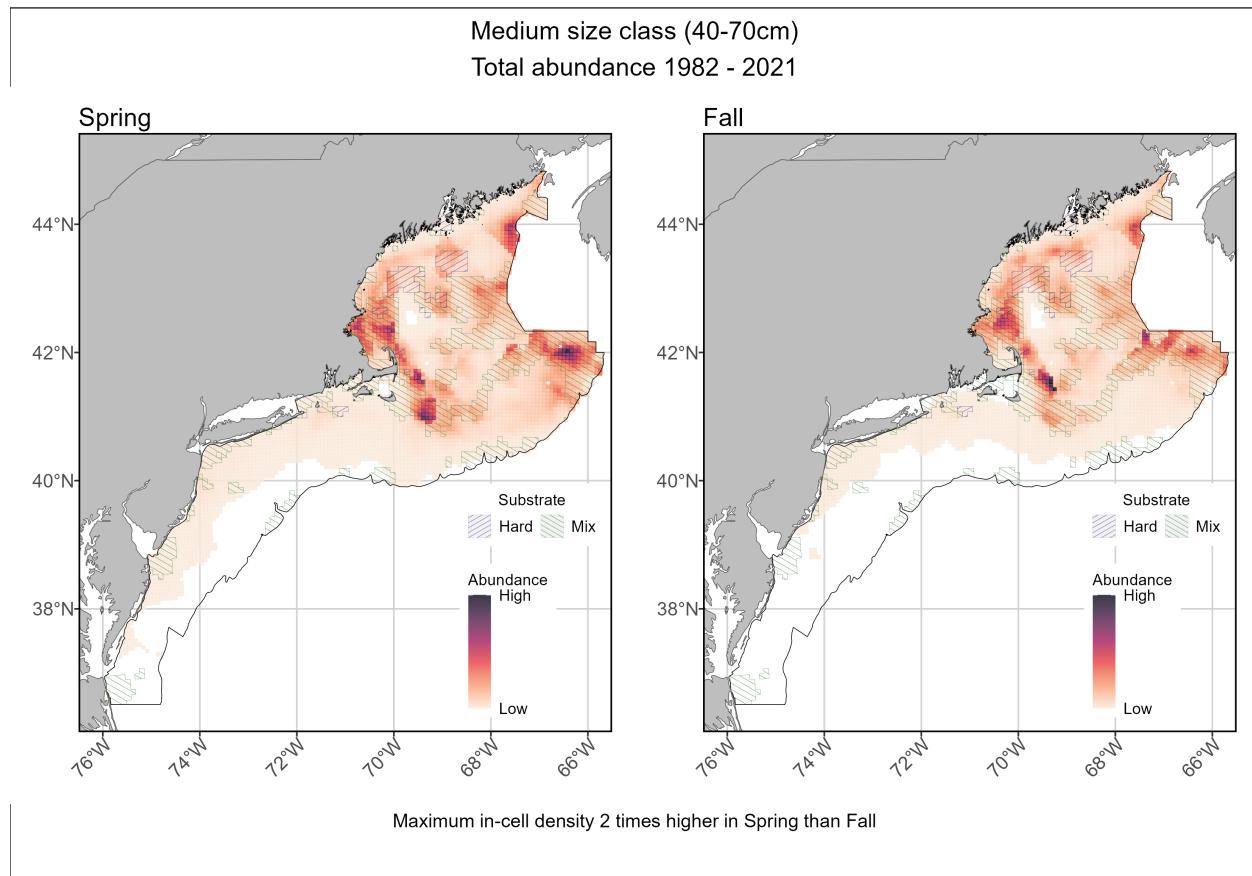


Figure 22: Medium cod aggregate spatial density split by season, 1982 - 2021.

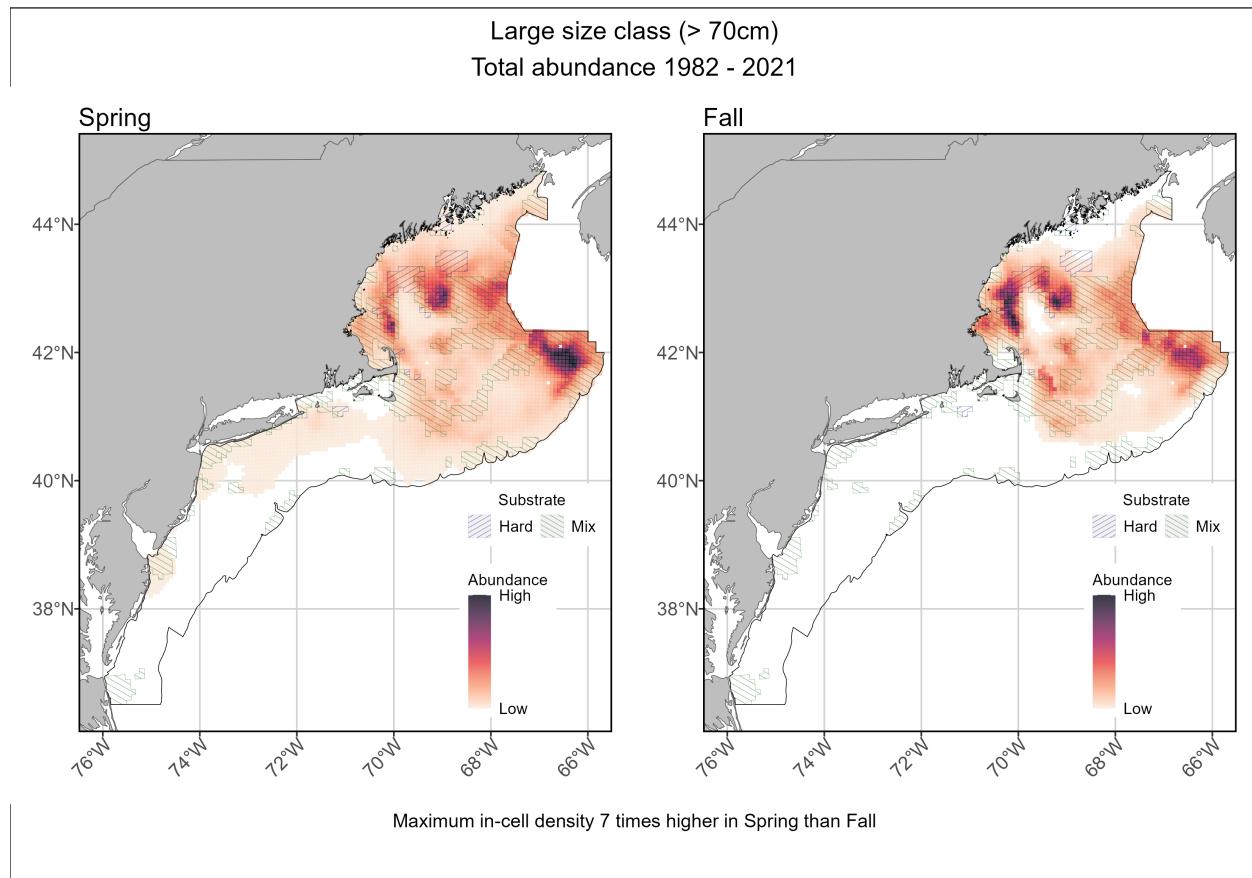


Figure 23: Large cod aggregate spatial density split by season, 1982 - 2021.

4.4 Model diagnostics

The effect of anisotropy was stronger in the first linear predictor than the second, meaning that probability of encountering cod was similar for longer distances along a northeast-to-southwest axis than the abundance of cod along the same axis.

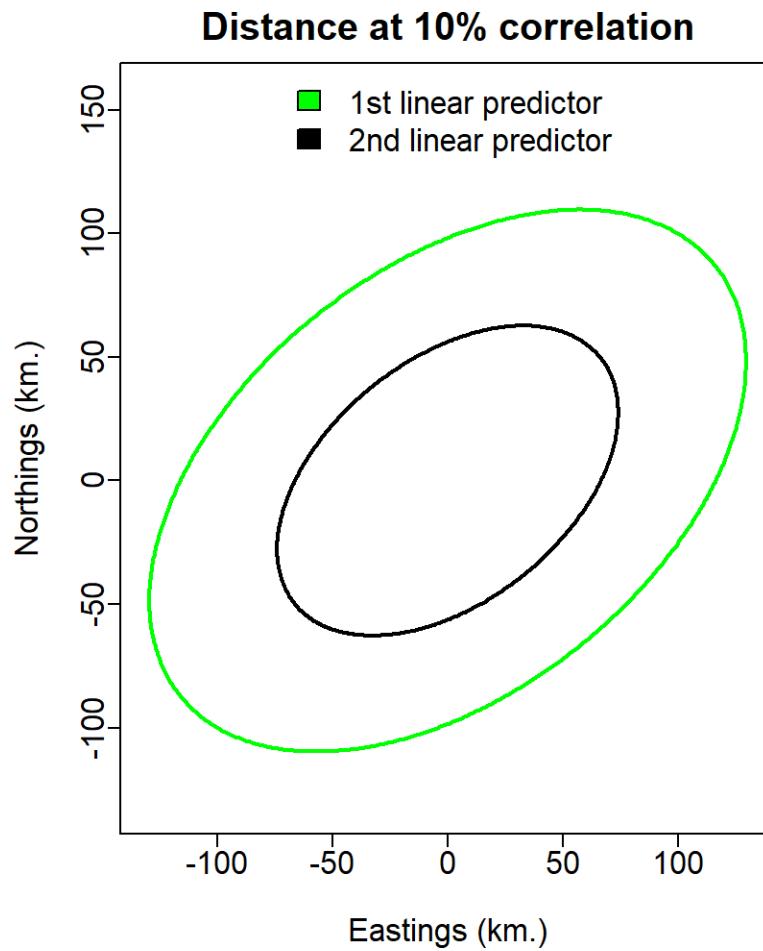


Figure 24: Anisotropic effect in the first and second linear predictors

Residual plots did not indicate any violation of assumptions in the construction or implementation of the model. Mapping residuals within the model spatial domain did not highlight any spatial area as having a consistently poor fit.

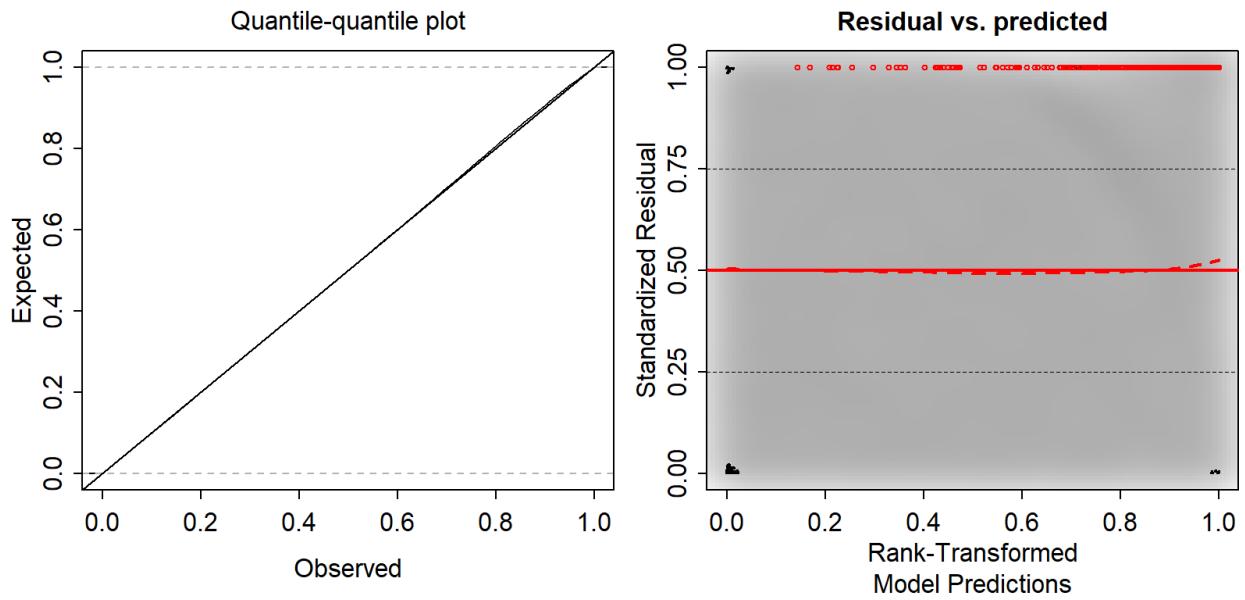


Figure 25: Quantile residuals of model run

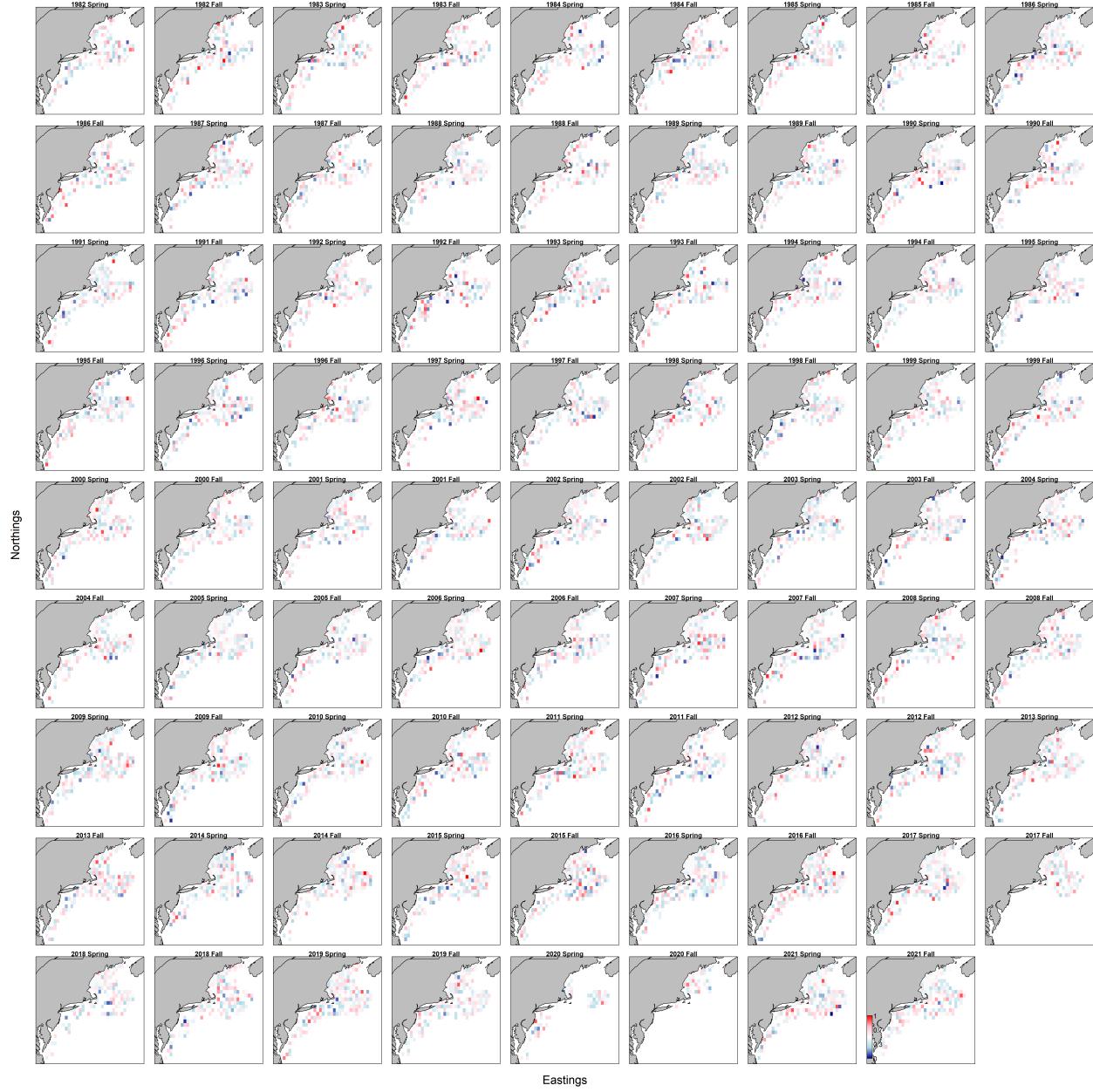


Figure 26: Quantile residuals of model run in spatial domain for every time step

5 Discussion

Currently, groundfish population assessments rely heavily on data gathered with bottom trawl surveys. Despite evidence of increased cod abundance in areas with complex bottom habitats, locations with large-grain sediments or high rugosity are avoided or sampled at low frequency with bottom trawls due to the risks of damaging equipment. This is a major point of contention for fishing industry members, who believe that limiting information from complex habitats will not accurately reflect cod spatial distribution or abundance (Grabowski et al. 2020). Industry members instead support sampling across all habitat types and utilizing dynamic stratification, in which areas of similar habitat characteristics and likelihood of encountering cod are grouped into strata for sampling and later inclusion in assessment

efforts. These strata cannot be built without accurate information on cod spatial density for all habitats within their spatial range.

VAST models can provide a flexible but robust framework to better assess groundfish populations across all habitat types within their spatial ranges. This is, of course, reliant on the collection of survey data in areas typically not well sampled by bottom trawls and by high-quality environmental and habitat data as density covariates. The model of cod spatial density created in this project utilized observation data from a suite of gear types and survey platforms, which bridged some of the data gaps inherent to models built with only bottom trawl survey data. The VAST modeling framework allowed for the use of density covariates to interpolate cod spatial density in non-sampled areas with known habitat and environmental characteristics. Further, VAST allowed for the estimation of spatial, temporal, and spatiotemporal correlation of cod abundance, as well as the inclusion of vessel effects, to facilitate the combination of observation data from multiple sources with varied protocols. This modeling process may be more favorable to industry stakeholders than models built with only bottom trawl observation data.

VAST models highlighted the variable spatial density of cod size classes across the four stock areas within the model spatial domain. There was a clear north-south directional gradient of abundance linked to size, where small cod were most abundant furthest south among the size classes and large cod were most abundant furthest north among the size classes. Further, VAST results highlighted the importance of relatively small areas within the stock areas to the size classes of cod. For example, the importance of Georges Bank to the large size class was highlighted; in the spring season, large cod abundance was highest in Georges Bank, driven by high spatial density at its eastern edge. For small cod, areas of high spring spatial density included the nearshore waters at the southwest edge of Martha's Vineyard.

Cod populations have continued to decline regardless of spatial area closures and massive reductions in fishing pressure. This trend is echoed in model results for all age classes in the fall time series, and all but small cod in the spring time series. The model does not support high emigration out of the model spatial domain as the cause of this population decline; stock centers of gravity for all size classes were typically well within the spatial bounds of the model. This indicates that factors other than fishing are playing a large role in preventing the recovery of the cod population. It is possible that reductions in fishing and spatial area closures are not protecting juveniles, which may be more important to cod population growth than adult survival and fecundity (Wright 2014). VAST results have shown variable distribution of cod by size class, with smaller fish preferring more southern, inshore waters with variable sediment types. Characterization of bottom habitat and subsequent modeling of length-specific habitat associations could help address this uncertainty and lead to more effective protective measures.

Stock centers of gravity and effective area occupied are variable with cod size but generally indicate a northward and offshore shift and contraction in spatial distribution. Range shifts towards more northerly and deeper water have been commonly seen in North American marine species (Fredston et al. 2021). For cod, this range shift may be related to warming bottom temperatures in areas historically occupied by cod, particularly in the inshore and southern areas of their spatial range. As cod populations move, it will be important to quantify the possible reduction in high-suitability habitats. This will require robust models of cod habitat associations built on cod observation data and high-quality habitat data throughout the cod spatial range.

References

- Du Pontavice, H., Z. Chen, and V. S. Saba. 2023. A high-resolution ocean bottom temperature product for the northeast U.S. Continental shelf marine ecosystem. *Progress in Oceanography* 210:102948.
- Fredston, A., M. Pinsky, R. L. Selden, C. Szewalski, J. T. Thorson, S. D. Gaines, and B. S. Halpern. 2021. Range edges of

- north american marine species are tracking temperature over decades. *Global Change Biology* 27:3145–3156.
- Friedman, A., O. Pizarro, S. B. Williams, and M. Johnson-Roberson. 2012. Multi-Scale Measures of Rugosity, Slope and Aspect from Benthic Stereo Image Reconstructions. *PLoS ONE* 7(12):e50440.
- Gotceitas, V., and J. A. Brown. 1993. Substrate selection by juvenile Atlantic cod (*Gadus morhua*): Effects of predation risk. *Oecologia* 93(1):31–37.
- Gotceitas, V., S. Fraser, and J. A. Brown. 1995. Habitat use by juvenile Atlantic cod (*Gadus morhua*) in the presence of an actively foraging and non-foraging predator. *Marine Biology* 123(3):421–430.
- Grabowski, J. H., S. X. Cadrin, V. Giacalone, T. Gouhier, L. A. Kerr, J. Odell, A. J. Pershing, G. D. Sherwood, K. D. E. Stokesbury, and G. C. Trussell. 2020. Strengths and Weaknesses of the Northeast Fisheries Science Center's Bottom Trawl Survey.
- Grabowski, J. H., C. W. Conroy, R. K. Gittman, J. T. Kelley, S. Sherman, G. D. Sherwood, and G. Wippelhauser. 2018. Habitat Associations of Juvenile Cod in Nearshore Waters. *Reviews in Fisheries Science & Aquaculture* 26(1):1–14.
- Linner, R. M., and Y. Chen. 2022. Implications of stock structure in understanding juvenile Atlantic cod (*Gadus morhua*) habitat suitability in the Gulf of Maine. *Regional Studies in Marine Science* 54:102473.
- McElroy, W. D., J. Blaylock, G. R. Shepherd, C. M. Legault, P. C. Nitschke, and K. A. Sosebee. 2021. Comparison of a bottom longline survey and a bottom trawl survey for 2 groundfish species in the Gulf of Maine to evaluate habitat-related availability of large fish. *Fishery Bulletin* 119(4):231–242.
- McElroy, W. D., L. O'Brien, J. Blaylock, M. H. Martin, P. J. Rago, J. J. Hoey, and V. A. Sheremet. 2019. Design, Implementation, and Results of a Cooperative Research Gulf of Maine Longline Survey, 2014–2017. {NOAA} {Technical} {Memo} NMFS-NE-249.
- Ng, E. L., J. J. Deroba, T. E. Essington, A. Grüss, B. E. Smith, and J. T. Thorson. 2021. Predator stomach contents can provide accurate indices of prey biomass. *ICES Journal of Marine Science* 78(3):1146–1159.
- Thorson, J. T. 2019. Guidance for decisions using the Vector Autoregressive Spatio-Temporal (VAST) package in stock, ecosystem, habitat and climate assessments. *Fisheries Research* 210:143–161.
- Thorson, J. T., and L. A. K. Barnett. 2017. Comparing estimates of abundance trends and distribution shifts using single- and multispecies models of fishes and biogenic habitat. *ICES Journal of Marine Science* 74(5):1311–1321.
- Thorson, J. T., R. Fonner, M. A. Haltuch, K. Ono, and H. Winker. 2017. Accounting for spatiotemporal variation and fisher targeting when estimating abundance from multispecies fishery data. *Canadian Journal of Fisheries and Aquatic Sciences* 74(11):1794–1807.
- Thorson, J. T., and K. Kristensen. 2016. Implementing a generic method for bias correction in statistical models using random effects, with spatial and population dynamics examples. *Fisheries Research* 175:66–74.
- Thorson, J. T., H. J. Skaug, K. Kristensen, A. O. Shelton, E. J. Ward, J. H. Harms, and J. A. Benante. 2015. The importance of spatial models for estimating the strength of density dependence. *Ecology* 96(5):1202–1212.
- Wright, P. J. 2014. Are there useful life history indicators of stock recovery rate in gadoids? *ICES Journal of Marine Science* 71(6):1393–1406.
- Zuur, A. F., E. N. Ieno, N. J. Walker, A. A. Saveliev, and G. M. Smith. 2009. Mixed effects models and extensions in ecology with R. Springer, New York.