Final Report

Scope of Work 7: Forage Indicator Development – Using Environmental Drivers to Assess Forage Status

Ryan Woodland, Edward Houde, Vyacheslav Lyubchich
University of Maryland Center for Environmental Science
Chesapeake Biological Laboratory
Solomons, MD 20688

UMCES Project: CBL2021-038CBT
Report Date: March 1, 2022
Invoice Period: January 15, 2022 – April 14, 2022
Progress Reporting Period:

Project highlights

- A suite of indicators was developed for Bay Anchovy (pelagic, forage fish) and Polychaetes (benthic, forage invertebrate) in Chesapeake Bay
- Indicator time series showed survey- and life-history dependent patterns within forage taxa, demonstrating the importance of considering multiple data sources when quantifying forage conditions
- Several variants of two climate-related indicators representative of water warming phenology in Chesapeake Bay (degree day) and sea surface temperature in the coastal Atlantic Ocean (Atlantic Multidecadal Oscillation) were developed and explored
- Forage-Climate modeling showed strong, non-linear relationships between forage and climate indicators
- Assignment of annual forage indicator value using a quantiles classification scheme based on abundance distribution terciles (High, Medium, Low) provided a potential option for management consideration

Statement of Problem

Forage species are a critical component in ecosystem-based fisheries management, linking lower trophic levels of food webs to economically and ecologically valuable predators. Recognizing this, the Forage Action Team of the Chesapeake Bay Program (CBP), a subunit of the Sustainable Fisheries Goal Implementation Team, committed to "…*develop an initial suite of indicators to assess the forage base in the Chesapeake Bay"* (FAT 2020). Building off outcomes of a Forage Workshop sponsored by the CBP's Scientific and Technical Advisory Committee (Ihde et al. 2015), recent research has documented: 1) diverse patterns in interannual abundance

of forage taxa, including stochastic stability and low-frequency cycles (Woodland et al. 2017); 2) spatiotemporal patterns in forage consumption by a suite of predators (Buchheister and Latour 2015, Buchheister and Houde 2016), and 3) forage-environment relationships shared across a range of taxonomically diverse forage groups (Woodland et al. 2021). Results have provided a foundation of knowledge to support development of a suite of taxonomic and climate-based forage indicators, informing Management Actions 1.1, 2.1, and 3.1 as described in the Forage Fish Logic and Action Plan (CBP, 2020). However, there remains a knowledge gap in understanding how best to quantify abundance and variability in specific taxonomic forage indicators and to understand how key forage population indicators covary with climate-based forage indicators, for example the 5°C degree-day indicator (DD; Woodland et al. 2017) or Atlantic Multidecadal Oscillation (AMO).

We addressed needs of the CBP as described in the request for proposals (RFP) by relating indices of abundance (Bay Anchovy *Anchoa mitchilli*) or biomass (Polychaetes) of two key forage taxa to two climate indicators (5°C DD index, AMO) in tributary and mainstem habitats of Chesapeake Bay. We accomplished this by 1) calculating and providing forage population indices, 2) further exploring new variants of the forage population indices (see below), and 3) relating forage population indices to forage climate indices.

This report summarizes our research on these objectives, providing the requested deliverables (*See Table Project Timeline, end of document*).

Table of Contents

Indicator selection and rationale

Forage indicators

Bay Anchovy (**Figure 1A**) has been identified as the most important forage fish in Chesapeake Bay in terms of its abundance and biomass consumed annually by predatory fish (Jung and Houde 2004a, Buchheister and Latour 2015, Ihde et al. 2015, Buchheister and Houde 2016, Woodland et al. 2017; **Figure 1A**). As a zooplanktivore, this species serves an important ecological role by linking planktonic food webs with higher trophic level predators that occupy pelagic, demersal, and benthic habitats (Baird and Ulanowicz 1989, Hartman and Brandt 1995, Buchheister and Latour 2015). High rates of natural mortality make Bay Anchovy essentially an annual species with annual mortality of the recruited stock exceeding 90% in Chesapeake Bay (Newberger and Houde 1995). Bay Anchovy is a batch spawner with individual females

spawning repeatedly, often daily, over the course of the summer (Luo and Musick 1991, Zastrow et al. 1991). Spawning takes place during late spring to summer months in Chesapeake Bay, with peak spawning activity typically occurring in mid-to-late July (Zastrow et al. 1991). Eggs hatch in ~1 day at typical summer water temperatures (Able and Fahay 1998) and young-of-the-year individuals grow and recruit during summer and fall months, overwinter, and mature during the following spring. Individuals become sexually mature at ~10–12 months of age. These "age-1" individuals dominate the adult spawning stock (Zastrow et al. 1991).

The life history of Bay Anchovy, particularly characteristics of an annual fish, suggests that recent and ongoing fisheries-independent surveys could support the development of several indicators for this forage species, based on relative abundance. Potential indicators include: (1) an adult spawning stock indicator based on catches during spring months prior to peak spawning, (2) an annual recruitment indicator based on catches of young-of-the-year individuals in the late summer and fall after most spawning has ended, and (3) a combined, aggregate index of relative abundance that includes adult and young-of-the year individuals. As proposed, these indices could provide useful, life-stage-specific information on trends in Bay Anchovy relative abundance, recruitment variability, and relationships between this forage fish and regional climate.

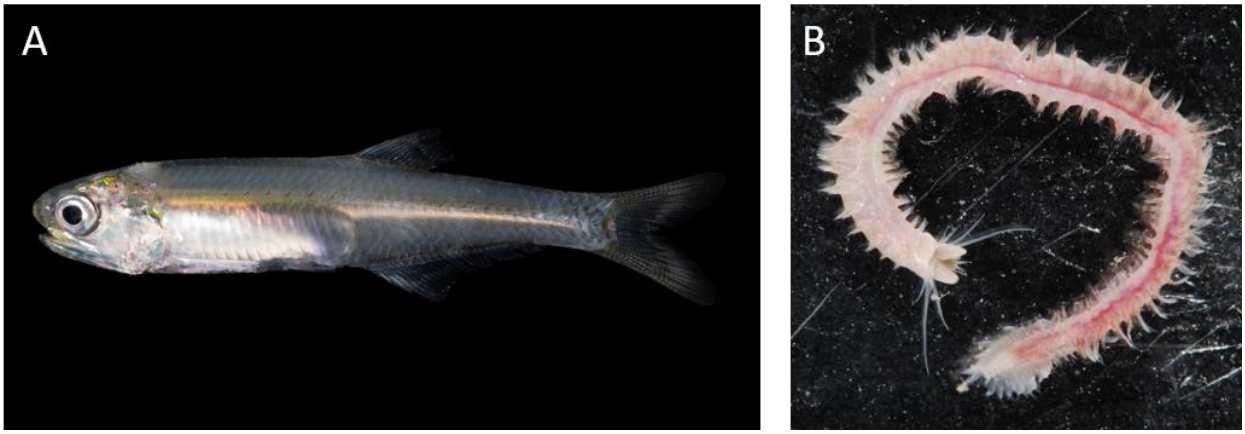

Figure 1. *Bay Anchovy (Anchoa mitchilli; A), and the Clam Worm (Alitta succinea; B), focal forage taxa for this study, collected from the Rhode River, MD. Photos: Robert Aguilar, Smithsonian Environmental Research Center (unaltered images; https://www.flickr.com/photos/serc_biodiversity/albums/with/72157642130185343).*

Among benthic invertebrates, marine annelid worms (Class Polychaeta) (Figure 1B) were identified as major prey for demersal and bentho-pelagic predators in Chesapeake Bay, for example, Atlantic Croaker (*Micropogonias undulatus*) and Striped Bass (*Morone saxatilis*) (Buchheister and Latour 2015, Ihde et al. 2015, Woodland et al. 2017). Diet data (2002–2015) from the Chesapeake Bay Multispecies Monitoring and Assessment Program (https://www.vims.edu/research/departments/fisheries/programs/mrg_oldwebsite/chesmmap/index.php) indicate that Bay Anchovy and an aggregate 'polychaete' taxonomic group contributed ~13% and 16%, respectively, to the total stomach contents by weight averaged across Atlantic Croaker, Striped Bass, White Perch (*Morone americana*), Spot (*Leiostomus xanthurus*), and

Summer Flounder (*Paralichthys dentatus*). Polychaetes are typically benthic although taxa within this class display a wide range of habitat associations, with some species building and residing within subsurface burrows or tubes at the sediment surface, while others move more freely about the surface of sediments or submerged structures (e.g., oyster reefs, seagrass beds, artificial structures; Lippson and Lippson 2006). The association of polychaetes with bottom habitats can expose them to episodes of bottom-water hypoxia; however, many polychaetes are physiologically tolerant of low oxygen conditions (Vaquer-Sunyer and Duarte 2008).

One of the most common taxonomic families of polychaetes encountered in the stomach contents of predatory fish in Chesapeake Bay is the Nereididae (http://www.vims.edu/fisheries/fishfood; this family includes the Clam Worm, *Alitta succinea*; **Figure 1B**). Nereids are abundant and widespread in Chesapeake Bay, ranging in occurrence from tidal freshwaters to the Bay mouth (Lippson and Lippson 2006, Testa et al. 2020). The abundance and trophic importance of Nereididae to Chesapeake Bay predators suggest that this taxonomic group is sufficiently abundant to support estimation of a biomass index. Annual indices of biomass per unit area for 1) an aggregate index of all polychaetes and 2) an index comprised solely of nereid polychaetes could provide complementary information on the status of polychaete forage for a wide range of fish predators in Chesapeake Bay.

Climate indicators

The Atlantic Multidecadal Oscillation (AMO) is an index derived from sea surface temperature in the 0-60°N latitude range of the Atlantic Ocean (Nye et al. 2014). As reviewed by Nye et al. (2014), the AMO displays a multidecadal cycle that oscillates between positive and negative phases with a period of ~ 60–70 years. In the mid-Atlantic region, positive AMO phases are associated with relatively warm, dry conditions while negative AMO phases are associated with cool, wet condition. The AMO is correlated with interannual patterns in the abundance of both fished species and forage groups, as well as assemblage changes in Chesapeake Bay's fish community (Wood and Austin 2009, Buchheister et al. 2013, Nye et al. 2014, Woodland et al. 2021). AMO data are available for download from the NOAA Physical Sciences Laboratory (https://psl.noaa.gov/data/timeseries/AMO/) as a detrended, standardized time-series (i.e., mean = 0 and unit variance).

Water temperature controls or modulates most biological and biogeochemical rate processes in aquatic ecosystems, including (but not limited to) metabolism and metabolic demand, reproductive phenology, primary and secondary production, and animal migrations. Water temperature, defined by degree day (DD; i.e., the mean daily water temperature residual above a minimum threshold temperature) has often been related to size-at-age or size-at-day in fishes or invertebrates, many of which have documented a positive effect between growth rates and DD-based indices (Ward and Stanford 1982, Bunnell and Miller 2005, Neuheimer and Taggart 2007, Humphrey et al. 2014). While positive biological rate responses are often associated with increased DD conditions, Woodland et al. (2021) documented an inverse relationship between the annual summertime abundances of many taxa within a diverse forage community in Chesapeake Bay and the rate of water warming during spring demonstrated by a 5°C DD phenology index. Subsequent research indicated a complex relationship between the 5°C DD

phenology index and benthic invertebrate biodiversity in Chesapeake Bay in which curvilinear interactions were identified between the phenology of water temperature DD and the cumulative amount of warming occurring during the first six months of the year (Woodland and Testa 2020).

Methods

Data sources

Data were obtained through direct download from online data hubs or by email correspondence with data managers (**Table 1**). Original versions of all obtained data have been archived and are available in a zipped file upon request. Minimally processed data have been aggregated into a single Excel workbook with each worksheet containing an individual data file. Metadata have been collated from each dataset, including standardized metadata provided during data downloads or by data managers. In some instances, additional metadata (e.g., date downloaded) have been added by the project team. The aggregated Excel file has been previously submitted to CBT as a deliverable for this project.

***Table 1**. Summary of data used to develop and calculate forage and climate indicators for this project.*

Data type/category	**Dates available**	**Source**	**Notes**
Forage data			
Polychaetes			
Sampling Events	1995-2019	Versar	MD and VA data
Biomass	1995-2019	Versar	MD and VA data
Bay Anchovy			
TIES/CHESFIMS	2001-2007	ChesFIMS/TIES	Mainstem data
Maryland DNR Juvenile Striped Bass Seine Survey	1966-2020	Eric Durrell – MD DNR	Email correspondence
Virginia DNR Juvenile Striped Bass Seine Survey	1967-2019	Troy Tuckey – VIMS	Email correspondence
VIMS Juvenile Fish and Blue Crab Trawl Survey	1988-2019	Troy Tuckey – VIMS	Email correspondence
Climate Data			
Regional climate			
AMO	1948-2020	NOAA	Monthly & yearly data; unsmoothed, short, detrended
Water Temperature			
Solomons SLIM2	2005-2020	NOAA NBDC	6-minute interval
CBL Pier	1938-2016	CBL Pier monitoring	Daily interval
York River YKTV2	2005-2020	NOAA NBDC	6-minute interval
VIMS Ferry Pier	1947-2003 (daily); 1986-2003 (6-minute)	W&M ScholarWorks W&M ScholarWorks	6-minute interval; daily
NERRS site – mouth of York River	1997-2021	NERRS Centralized Data Management Office	Data link broken

Polychaete survey

Polychaete presence-absence and biomass data were obtained from the benthic survey data collected as part of the USEPA Chesapeake Bay Long-term Benthic Monitoring and Assessment Program (1995-2019; baybenthos.versar.com/default.htm). Methods for the survey are published (e.g., Seitz et al. 2009), described in detail (Dauer 2011, Versar 2011), and are briefly summarized here. A total of 250 sampling stations are selected each year from a grid of potential sampling locations using a stratified random sampling plan that assigns 25 stations to each of 10 survey strata (**Figure 2**). The sampling is conducted during the summer from late July through September, but logistical and weather-related constraints lead to occasional sampling from early July to early October. A single deployment of a 440 cm^2 benthic Young Grab collects a sample of epibenthic and benthic fauna at each station. Samples are sieved through a 500-µm screen and organisms preserved in 10% formalin with Rose Bengal. All organisms are identified to the finest possible taxonomic resolution and then processed in a muffle furnace for 4 hrs at 500°C to obtain ash-free dry weight (AFDW). Ash-free dry weights of taxa are scaled to a normalized sampling area of 1 m^2 prior to analysis. We selected Polychaete taxa and their biomasses from the dataset and created two separate datasets, one that included all Polychaete taxa (Total Polychaetes) and one that included only nereid Polychaetes (Nereididae Polychaetes).

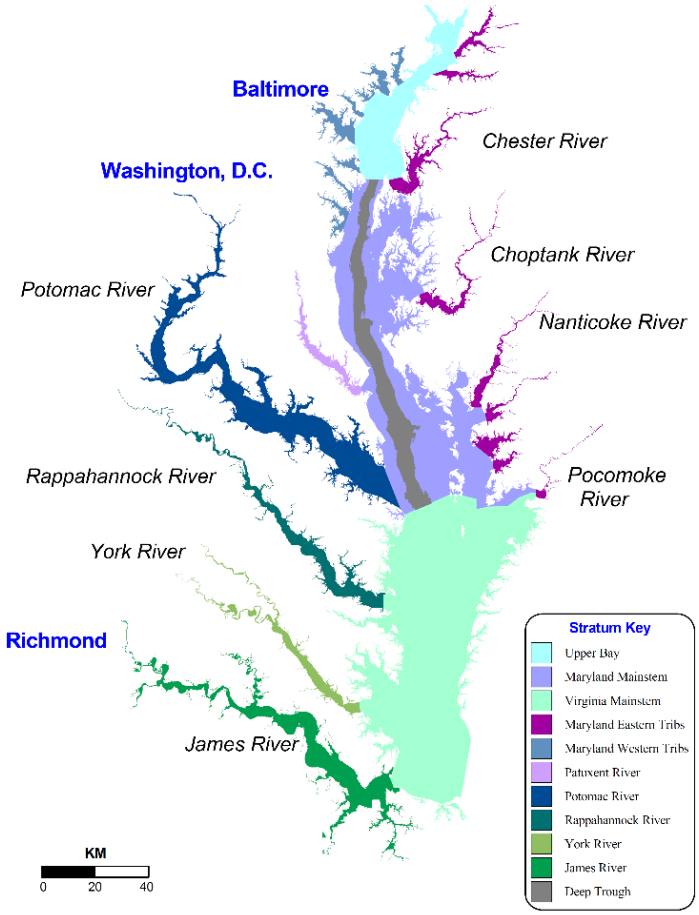

Figure 2. *Spatial distribution of survey strata for annual site selection by the Chesapeake Bay Program Benthic Monitoring Survey. Image from http://www.baybenthos.versar.com/docs/LTBStrata.pdf.*

Bay Anchovy surveys

A comprehensive, long-term trawl survey that samples smaller forage-fishes such as Bay Anchovy throughout the mainstem of Chesapeake Bay does not exist. To create a time-series we combined two historical surveys conducted by the University of Maryland Center for Environmental Science- - the Trophic Interactions in Estuarine Systems (TIES) and Chesapeake Bay Fishery-Independent Multispecies Survey (ChesFIMS) programs. Those programs were conducted throughout the mainstem bay (**Figure 3A**) in the period 1995–2007 using a standardized gear (18-m^2 mouth opening midwater trawl with 3-mm codend mesh) and deployment methods (20 min surface-to-bottom stepped tows [minimum depth = 4 m]) in the mainstem of Chesapeake Bay. Samples were collected each year in the upper (39°25′ –38°45′), middle (38°45′ – 37°55′) and lower (37°55′ – 37°05′) regions of the bay mainstem (Jung and Houde 2003, Jung and Houde 2005). Sampling was conducted seasonally during the spring (April-May), summer (July-August) and fall (September-October). All fish were identified to species and catch numbers (counts per 20-min tow) were estimated directly or estimated using volumetric subsampling. The TIES and ChesFIMS data were downloaded from the data portal hjort.cbl.umces.edu/chesfims.html and from survey records held by co-PI Houde.

Two seine-survey datasets were selected to estimate Bay Anchovy indices of relative abundance from bay tributaries in MD and VA: (1) the Maryland Department of Natural Resources Juvenile Striped Bass Seine Survey (MJS, 1959–2019; data provided by E. Durell personal communication) which samples the Potomac, Patuxent, Nanticoke, and Choptank rivers, as well as stations in upper Chesapeake Bay, and (2) the Virginia Institute of Marine Science Juvenile Striped Bass Seine Survey (VJS, 1968–1973 & 1980–2019) that is conducted in the Rappahannock, York and James rivers (**Figure 3B, C**). Both surveys deploy a beach seine perpendicular to shore to c. 1-1.2 m depth, quantify the actual length of deployment (if less than the full net length), then retrieve the seine in a quarter-arc sweep back to shore. Gear dimensions are identical in the two surveys (wooden brails, 30.5 m x 1.2 m beach seine, 6.4-mm bar mesh). The MD and VA surveys conduct duplicate seine hauls at each station; however, we used only the first haul in our analyses to avoid bias due to potential effects of repeated sampling.

Post-processed Bay Anchovy annual Spawning Stock and Recruitment indices were obtained from the VIMS Juvenile Fish and Blue Crab Trawl Survey for the VA portion of Chesapeake Bay's mainstem (https://www.vims.edu/research/departments/fisheries/programs/juvenile_surveys/index.php). Briefly, a monthly size threshold value was applied to the length-frequency information collected for Bay Anchovy to partition the catch data into spawning stock and recruits for index calculation (Colvocoresses and Geer 1991). Annual Spawning Stock and Recruit indices were calculated by VIMS colleagues by calculating stratum-specific means and variances and combining stratum-specific estimates using weights based on stratum area (see Tuckey and Fabrizio 2021 for additional information on trawl survey design, stratum area, and weighted index calculation method).

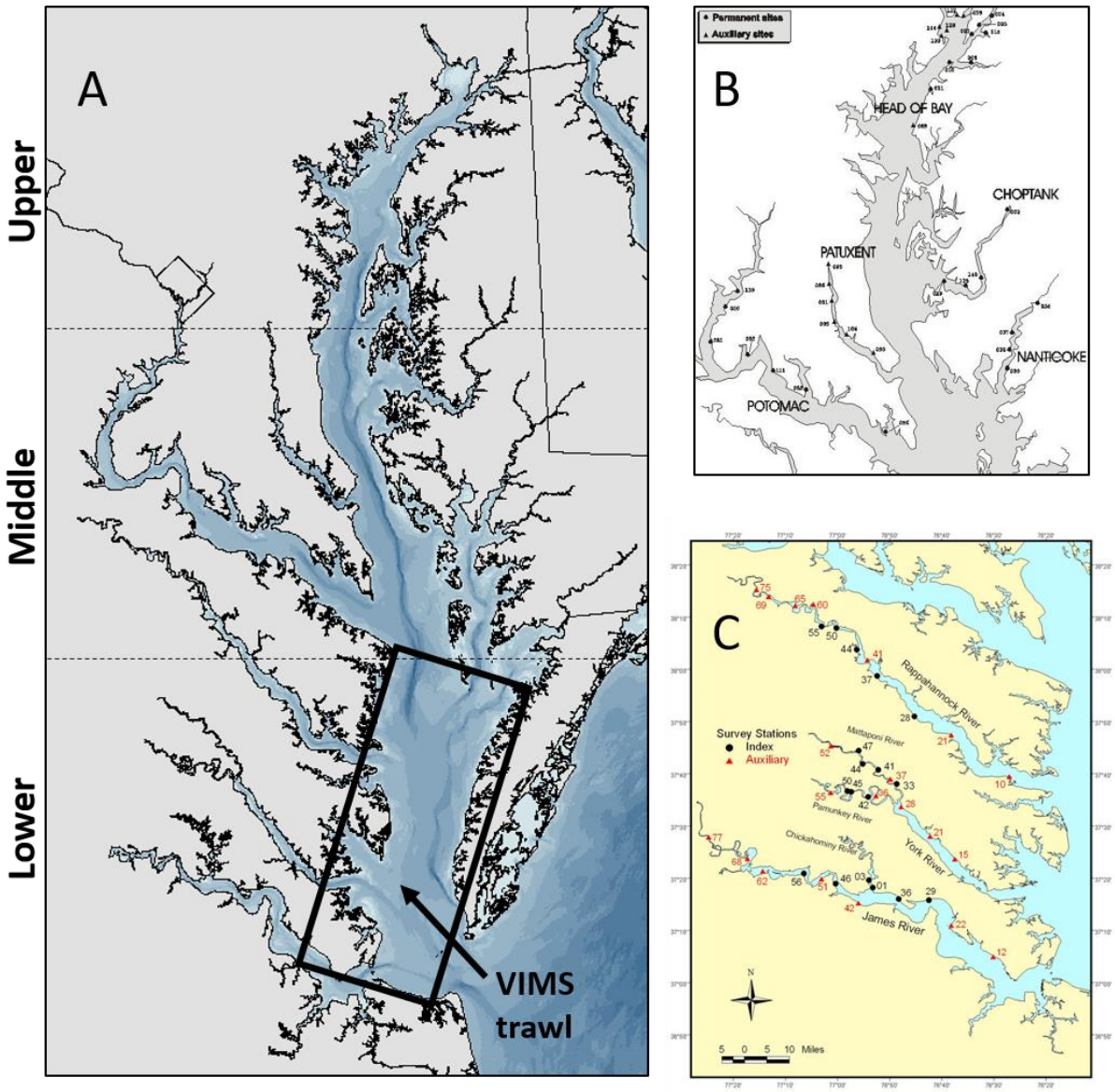

Figure 3. *Spatial distribution of survey extents and regions included in the TIES/ChesFIMS trawl dataset (A, region breaks for Upper, Middle, Lower areas given with dashed lines), the Maryland Department of Natural Resources Juvenile Striped Bass Seine Survey (B), and the Virginia Institute of Marine Science (VIMS) Juvenile Striped Bass Seine Survey (C). The sampling domain of an ancillary trawl dataset associated with the VIMS Juvenile Finfish and Blue Crab Survey (black rectangle, A) shown. Maps B and C downloaded from https://dnr.maryland.gov/fisheries/pages/striped-bass/juvenile-index.aspx and https://www.vims.edu/research/departments/fisheries/programs/juvenile_striped_bass/stations/index.php, respectively.*

Atlantic Multidecadal Oscillation

Detrended, monthly values (January to December) of the AMO indicator were obtained from 1988 to 2019. These AMO values were downloaded directly from the NOAA Physical Sciences Laboratory (NOAA/PSL, https://psl.noaa.gov/).

Degree Day – daily temperature site descriptions

Three central monitoring sites were selected based on spatial location and the availability of daily water temperature records. These sites included the Solomons, MD monitoring station (SLIM2; located along the mainstem of mid-Chesapeake Bay), the VIMS Ferry Pier monitoring location (Gloucester Point in the lower Bay, VA), and the Goodwin Islands monitoring station at the mouth of the York River (VA). There is a strong relationship between air and water temperatures in Chesapeake Bay (Ding and Elmore 2015) and Woodland et al. (Woodland et al. 2017) demonstrated a positive, linear relationship between paired air- and water-temperature data at SLIM2. Accordingly, air temperature data from SLIM2 also were downloaded to impute missing daily water temperatures.

Indicator calculations

Forage indicator modeling framework

Prior to conducting the statistical model selection process, two preliminary steps were taken during model development:

1. Assessment of data distributions – A visual analysis of the frequency distributions of Polychaete biomass and Bay Anchovy relative abundance data was conducted to identify the need for data transformations prior to indicator estimation.

2. Evaluation of data aggregation effects on final model estimates – This preliminary modeling step focused on testing whether or not model predictions and performance were affected by data aggregation. To test for aggregation effects we compared the statistical performance of models that used mean annual biomass or relative abundance for each regional stratum as observation-level data (aggregated prior to modeling) or models that used site-level observations within strata (disaggregated approach).

Following these preliminary steps, a suite of statistical models was evaluated, depending on the structure of the data, to identify an optimal estimation model for each forage indicator. These models included a generalized additive model (GAMs; Wood 2017), generalized linear model (GLM), delta-generalized linear model (Delta-GLM), delta-generalized additive model (Delta-GAM), and a random forest model (RF).

GAMs allow a non-normal distribution of the response variable, such as a zero-inflated Poisson distribution, to account for a large proportion of zeros in the data. The additive structure in GAMs allows smooth nonlinear relationships between variables to be modeled in a data-driven way. GAM-estimated curves can be used to identify thresholds in environmental conditions associated with biotic response variables. GAMs can be extended to include effects of autocorrelation that are often present in environmental time series (i.e., extend to models with mixed effects, GAMM). When a relationship is described best by a straight line, GAMs are automatically reduced to generalized linear models (GLMs).

The delta-GAM (or delta-GLM) is a two-step modeling approach, in which the first component models presence-absence, and second component uses a GAM (or GLM) only for non-zero

cases. A logistic model is used in the first component to model probability estimates of occurrence (all observations included) for a given forage taxon, and the second GAM (or GLM) component estimates relative abundance when present (Lynch et al. 2012). These two estimates are then combined to yield a single annual estimate of the biotic response variable.

The random forest (RF) approach was used as an alternative nonparametric method for modeling the indicator responses. RF is one of the best-performing machine learning techniques (Hastie et al. 2009), hence serves as a valuable benchmark for validating the results. RF is a set of regression trees, where each tree is trained on a bootstrap sample of the original data, and only a random subset of covariates is assessed at each tree split. The splits allow a RF to model non-smooth relationships (compared to only-smooth relationships in GLM and GAM), which can be visualized through partial dependence plots showing marginal changes in the response due to the changes in each covariate (all other covariates in the RF are fixed at their observed levels).

Each model applied to a forage indicator was fitted to a subset of data, then cross-validation applied to test model performance for out-of-sample predictions. For the Polychaete and Nereididae polychaete indicators, observations were weighted by the proportion of area they represent (area-weighted estimation; observations from larger strata were weighted more heavily than smaller strata). For the Bay Anchovy models using TIES/ChesFIMS data, observations were weighted by the proportional volume of water column habitat present in Upper, Middle, and Lower regions of Chesapeake Bay (Jung and Houde 2004a). Weighting was not applied in the fitting of the MD and VA seine survey Bay Anchovy indices. In the instances of data weighting during model fitting, the weighting was also used in the calculation of model evaluation metrics. Fitted data were compared to model estimates to evaluate out-of-sample model accuracy and select the best-performing models for each forage indicator. The selected model was used to fit the whole dataset to produce indicator estimates. The final results from the forage indicator modeling were annual estimates of biomass (Polychaetes) or relative abundance (Bay Anchovy) for each applicable spatial subunit based on the survey type and for the integrated Chesapeake Bay (weighted as described above; see *Appendix 1. Polychaete indicators* and *Appendix 2. Bay Anchovy indicators* for detailed output and associated r-code).

Climate indicator modeling

Atlantic Multidecadal Oscillation

Annual values of the AMO indicator were calculated by averaging monthly AMO values from January-December of each year (**Figure 4**). In addition to an annual AMO index spanning a standard calendar year, we undertook a preliminary data exploration to evaluate the effect of recalculating the AMO index to better reflect the interval of time between intervening forage sampling years. We did this for both the benthic survey as well as the fish surveys that we used to calculate Polychaete and Bay Anchovy indicators, respectively. The benthic survey is conducted annually during the summer months; however, the AMO is typically calculated for a given calendar year (January to December). This discrepancy means that traditionally calculated, annual AMO values include climate conditions (specifically sea-surface temperature conditions) present 1–5 months after benthic surveying was completed for a given year. Similarly, the intervals over which Bay Anchovy spawning stock and recruitment indicators are calculated align most closely with climate conditions present early or late in a given year. Thus, to explore

alternative AMO calculations, we explored the use of monthly AMO values that were averaged over the full range of possible monthly combinations during the July–December period of the previous year (i.e., lagged 1 year, hereafter AMO_L1) and during the January–July period of the sampling year (i.e., current year, hereafter AMO). Spearman correlation analysis was conducted to determine the optimal period(s) over which to relate the partial year AMO_L1 and AMO, or the standard January-December annual AMO to the forage indicators.

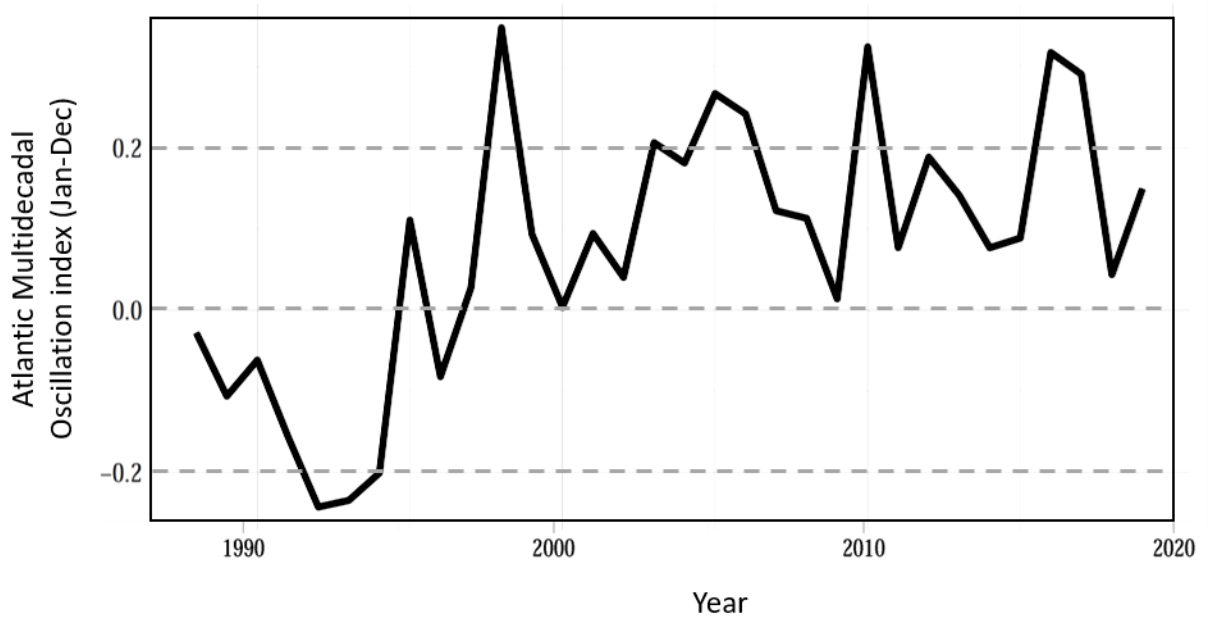

Figure 4. *Time-series of the annual Atlantic Multidecadal Oscillation indicator (AMO) calculated at 1-year calendar intervals from January–December.*

Degree Day

Daily water temperature records from all monitoring sites were included to calculate DD indices. Missing daily temperature records were reconstructed using the following protocol (see *Appendix 3. Forage-climate modeling* for detailed information):

1 – application of previously published regression models relating observed water temperatures at one location to water temperatures at the other location (Woodland et al. 2021) where at least one empirical daily measurement is present,

2 – use of an autoregressive model to impute temperature values for each station within 5 days of an empirical measurement (e.g., **Figure 5**), and

3 – for any remaining data gaps, application of a machine-learning algorithm (random forest) to estimate daily water temperature.

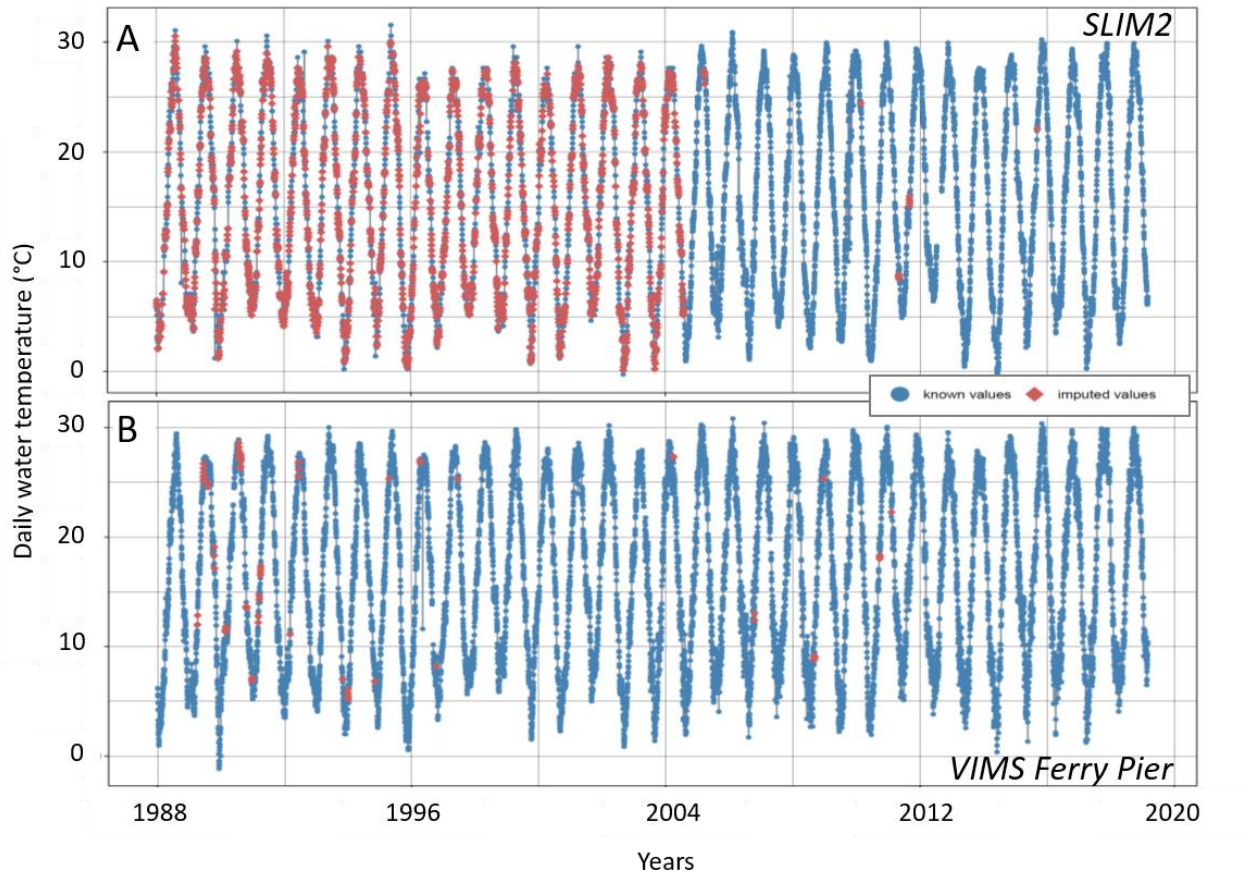

***Figure 5**. Time-series of observed (blue) and a partial representation of imputed (red; autoregressive model only) daily water temperature data for the lower Patuxent River (A; SLIM2 NOAA station) and the lower York River (B; VIMS Ferry Pier water temperature dataset).*

From the complete records of daily water temperature at the mouths of the Patuxent and York rivers, an average of these two datasets was calculated to yield a single, aggregated, mean daily water temperature time series. This aggregated, daily water temperature time series was used to calculate the 5°C DD temperature threshold index (DD5; **Figure 6**) following the method described by Woodland et al. (2021). In addition to the DD5 index, we calculated a similar index using 10°C as a DD temperature threshold to explore the potential utility of an alternative DD threshold (**Figure 6**). Briefly, the DD index derivation involves calculating the daily temperature anomaly as the difference between observed temperature and the specified threshold, either 5 or 10°C. Anomalies less than zero are assigned a value of 0 (i.e., threshold). For example, when calculating the DD5 indicator, a daily water temperature of 4°C would yield a daily anomaly = 0°C, but a daily water temperature of 8°C would yield a daily anomaly = 3°C. Starting on January 1 of each year, the daily anomalies are summed cumulatively until a total of ≥ 500 5°C DD has been accrued. The day of the year at which the 500 5°C DD occurs is the DD5 indicator value for that year. The DD10 indicator is calculated the same way, but assigning a 10°C threshold for the daily temperature anomaly calculation.

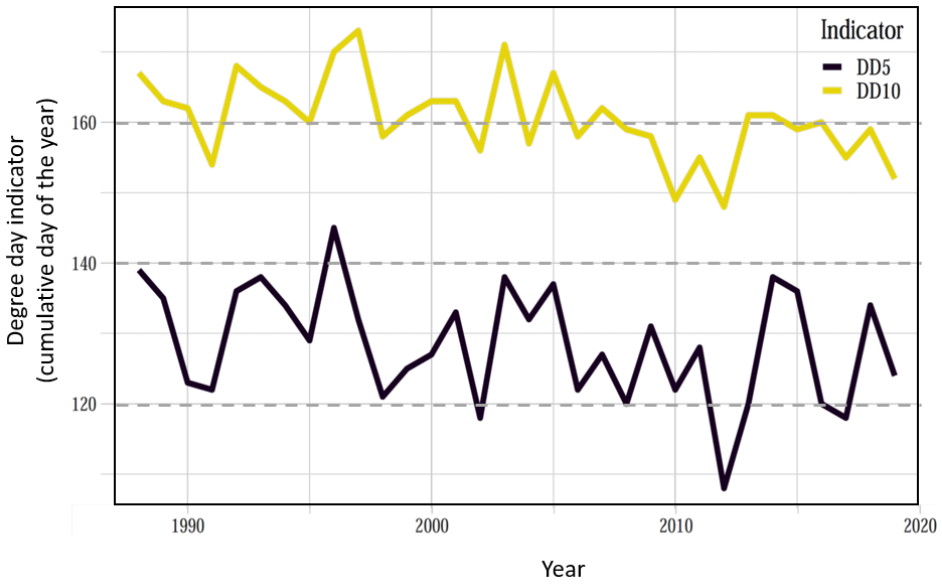

Figure 6. *Time-series of the 5°C (DD5) and 10°C (DD10) degree-day indicators*

Forage and climate indicator analyses

Indicator status

Several classification schemes were explored to assign annual forage and climate indicator observations into categories based on their relative value among the recent observations (i.e., within the time series available for this study). The number of categories was identified based on data distributions and following consultation with project partners and decision makers. In particular, we considered the potential application of several quantitative approaches for determining categories: 1) quantiles, including terciles and quartiles of observed data; and 2) semi-supervised *k*-means clustering, where the number of clusters *k* was selected to replicate the number of quantiles for an exploratory assessment (Bishop 2006, Manning et al. 2008). Finally, we applied a colorimetric scheme representing categories for indicator values, from red (Low) to yellow (Medium) to green (High).

Forage-Climate models

Three models were evaluated to relate forage indicators to climate predictors, including a GAM, GLM, and RF model. Each model was fitted to a subset of data, then leave-one-out cross-validation was used to test model performance for out-of-sample predictions.

For the Polychaete indicators, correlations for the AMO_L1 variants were strongest from midsummer–winter of the previous year (positively correlated) and correlations for AMO variants ranging from winter–midsummer of the current year (negatively correlated; **Figure 7**). These opposing correlations were our justification for including both AMO_L1(July–December) and AMO (January–July) as potential predictors in the models. In addition to the AMO_L1(July–December) and AMO (January–July) indices, both DD indicators and detrended versions of the DD indicators were evaluated as predictors. Unlike the AMO_L1 and AMO variables (which

were not correlated), separate model runs were conducted for each DD indicator to avoid multicollinearity due to presence of strong correlations (Spearman rank-order correlation, $r_S \geq 0.57$, $p << 0.05$) among DD5, DD10, and the detrended versions of these indicators.

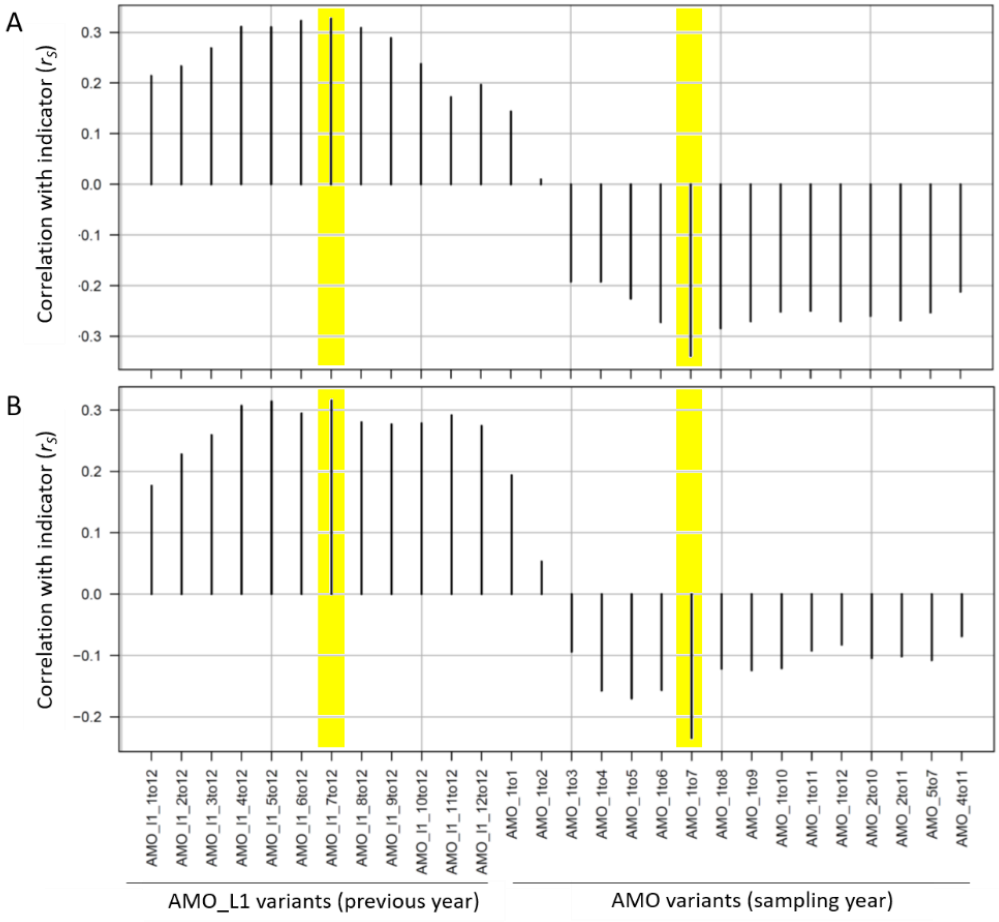

Figure 7. *Spearman rank-order correlations (rS) between the total Polychaete biomass forage indicator (A) and lagged and current sampling year AMO variants (AMO_L1, AMO), and between the Nereididae Polychaete biomass forage indicator (B) and the AMO_L1 and AMO variants. Numbers following the second underscore in AMO_L1 and AMO variants indicate months of the year over which a given variant was calculated. Highlighted correlations indicate AMO_L1 and AMO variants selected for inclusion in statistical models.*

For the Bay Anchovy indicators estimated using the TIES/ChesFIMS dataset, correlations of AMO_L1 variants were strongest from August–December for $BA_{Spawning}$, and from January–December for BA_{Total} (**Figure 8A, C).** AMO variants that correlated most strongly with $BA_{Spawning}$ and BA_{Total} were AMO January–February and AMO May–July, respectively. There was no strong relationship between AMO_L1 and $BA_{Recruits}$; therefore, two AMO variants were included in the $BA_{Recruits}$ climate model: AMO January, and AMO January–May. For the $BA_{Recruits}$ and BA_{Total} indicators that were estimated using the MD/VA Seine survey datasets, the strongest, negative correlations occurred for the AMO January–February index (**Figure 9**).

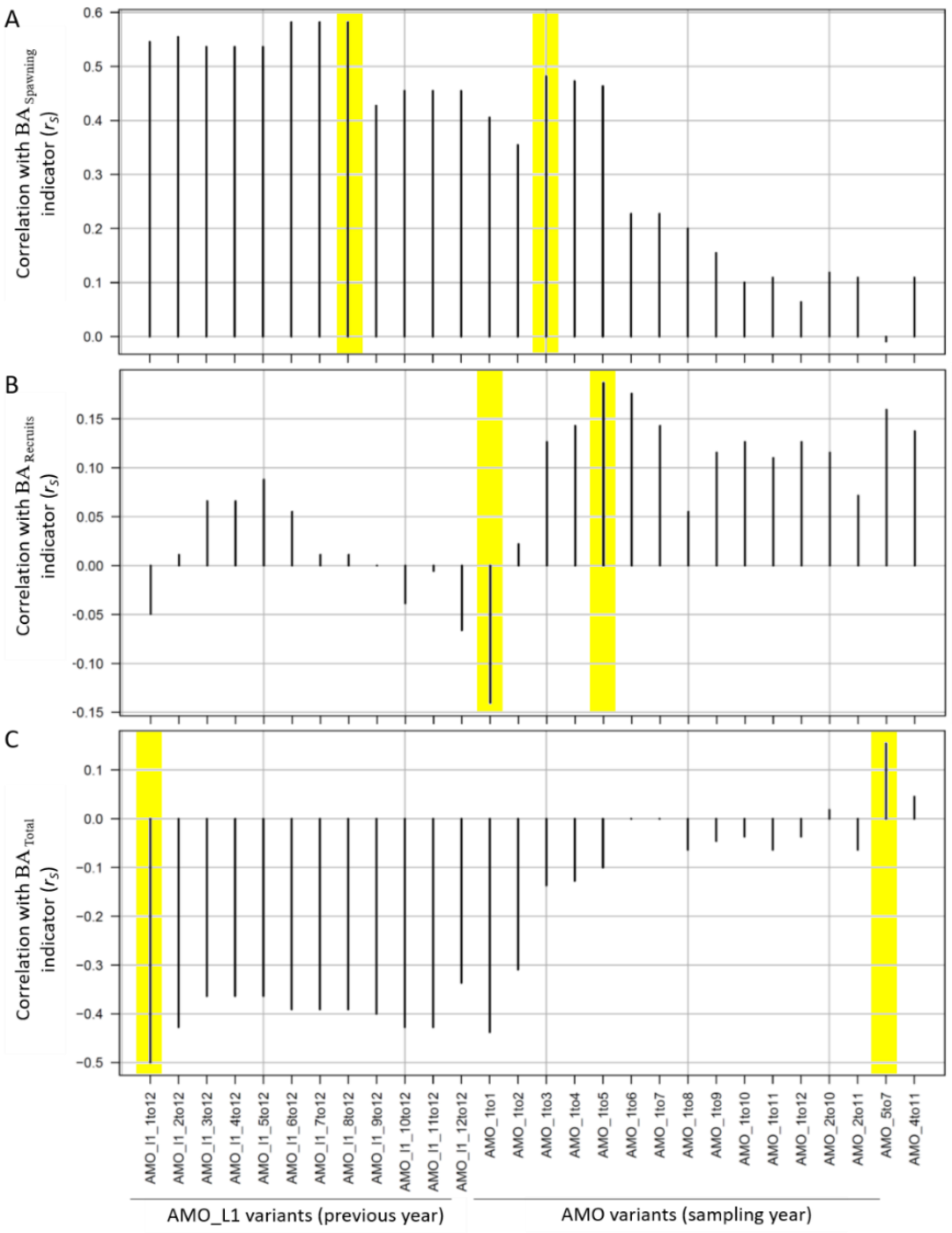

Figure 8. *Spearman rank-order correlations (r_S) between the TIES/ChesFIMS trawl Bay Anchovy spawning stock ($BA_{Spawning}$, A), recruits ($BA_{Recruits}$, B), and total population (BA_{Total}, C) indicators and lagged and current sampling year AMO variants (AMO_L1, AMO). Numbers following the second underscore in AMO_L1 and AMO variants indicate months of the year over which a given variant was calculated. Highlighted correlations indicate AMO_L1 and AMO variants selected for inclusion in statistical models.*

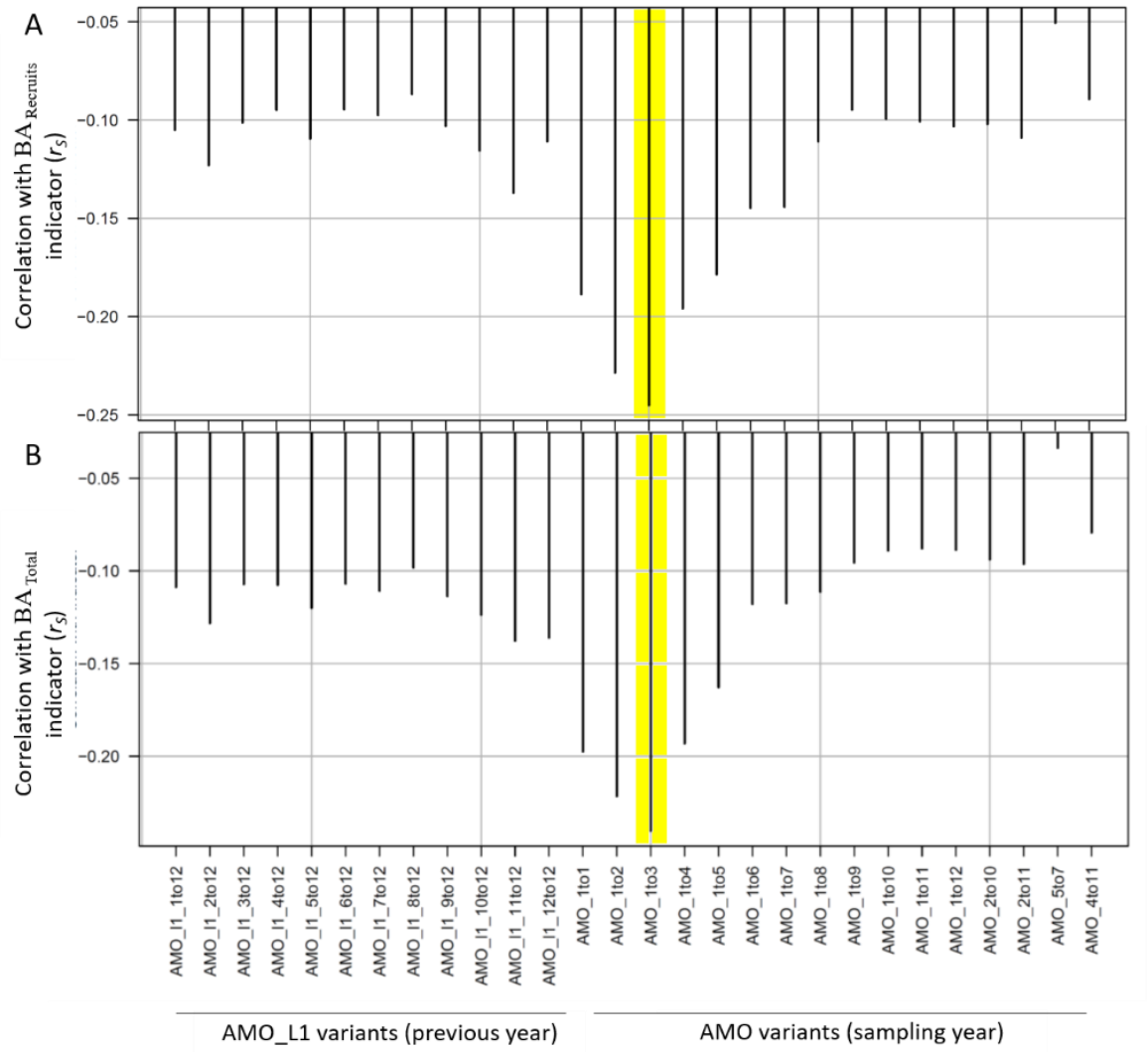

Figure 9. *Spearman rank-order correlations (r_S) between the Maryland & Virginia seine survey Bay Anchovy recruits ($BA_{Recruits}$, A) and total population (BA_{Total}, B) indicators and lagged and current sampling year AMO variants (AMO_L1, AMO). Numbers following the second underscore in AMO_L1 and AMO variants indicate months of the year over which a given variant was calculated. Highlighted correlations indicate AMO_L1 and AMO variants selected for inclusion in statistical models.*

Results

Forage indicators

Polychaetes

Total Polychaete and Nereididae Polychaete biomass distributions were strongly right-skewed, requiring $\log_e$-transformation to approach normality (**Figure 10**). Preliminary testing showed that model predictions and performance were better using disaggregated site-level observations of

Polychaetes biomass within strata rather than pre-aggregated, mean annual Polychaete biomass for each stratum. Therefore, all models reported and described below were fitted using individual site-level total or nereid Polychaete biomass (ash-free dry weight [AFDW], mg / m^2) as observational units.

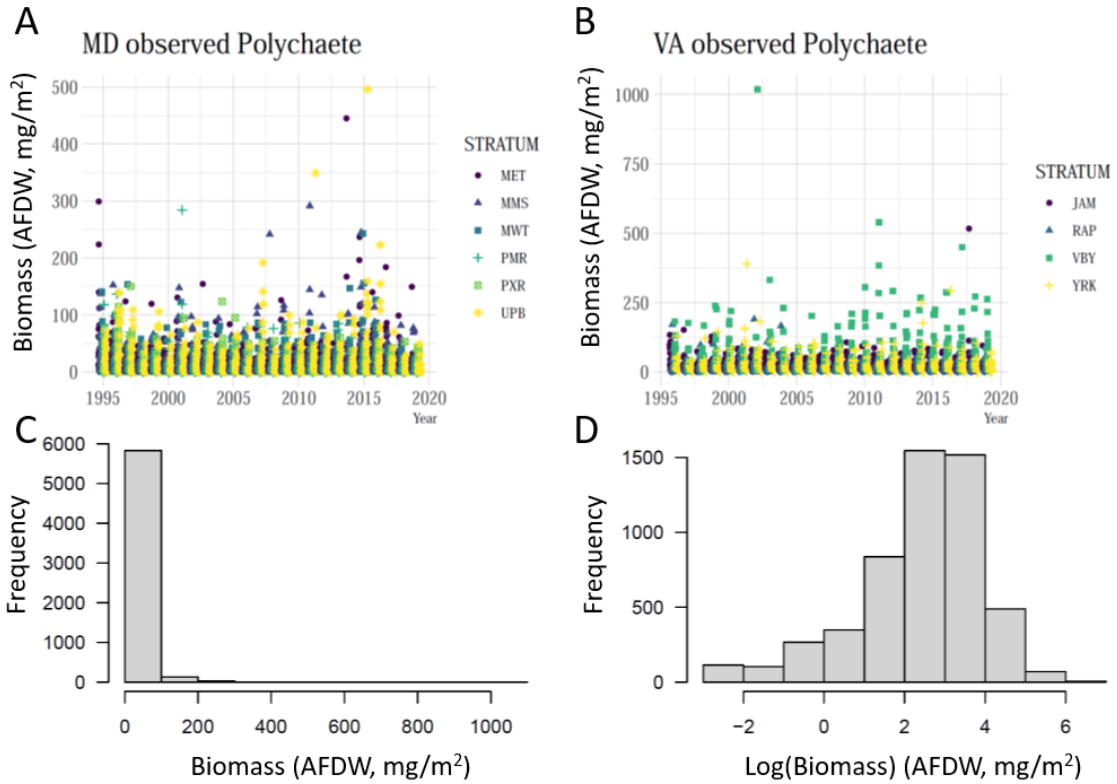

Figure 10. *Observed total Polychaete biomass at survey sites from sampling strata in the Maryland (MD; A), and Virginia (VA; B) portions of Chesapeake Bay. Total Polychaete biomass distributions shown on observed (C) and loge-transformed (D) scales. Sampling strata are: MET – Maryland Eastern Tributaries, MMS – Maryland Mainstem, MWT – Maryland Western Tributaries, PMR – Potomac River, PXR – Patuxent River, UPB – Upper Mainstem, JAM - James River, RAP – Rappahannock River, VBY – Virginia (Lower) Mainstem, YRK – York River.*

Results from the model fitting showed very similar performance by two models for each Polychaete index (**Table 2**). In the case of total Polychaetes, the RF and Delta-GAM approaches were the best performing models with lower associated model errors and higher R^2 values. Results from the Nereididae modelling were similar, with RF and GAM approaches having lower associated error metrics and higher coefficient of determination (R^2) values, although model performance was relatively poor for both models. Based on these results, the Delta-GAM was selected as the best model for the total Polychaete index and the GAM model was selected as the best model for the Nereididae Polychaete index.

Table 2. *Cross-validation (out-of-sample) model performance metrics for Total and Nereididae Polychaete biomass indices. Model error metrics include MAE (mean absolute error) and RMSE (root mean square error); R^2 is coefficient of determination (all metrics are computed using weights associated with survey sampling strata areas). Yellow highlighted cells associated with the best performing models for each performance metric, green cells indicate the second-best performance metric.*

		MAE	RMSE	R2
Total	GAM	25.54872	47.48895	0.1456766
	Delta-GLM	25.90639	47.95317	0.1301086
	Delta-GAM	25.35281	47.30858	0.1529223
	RF	24.98643	47.17995	0.1621094
Nereididae	GAM	4.672366	12.59160	0.02152202
	Delta-GLM	4.947856	12.96985	-0.06161526
	Delta-GAM	4.775525	12.85549	-0.02221711
	RF	4.713843	12.55976	0.02077619

Depth of the sampling site was a significant covariate in each of the final models with Polychaete biomass showing a curvilinear relationship with respect to sampling depth (**Figure 11**). In-sample predictions of Polychaete biomass showed close agreement for the total Polychaete biomass index in both geographically large and small strata (**Figure 12A**). In-sample predictions for the Nereididae index were less accurate, with the model showing a trend towards underestimating high observed biomasses (**Figure 12B**).

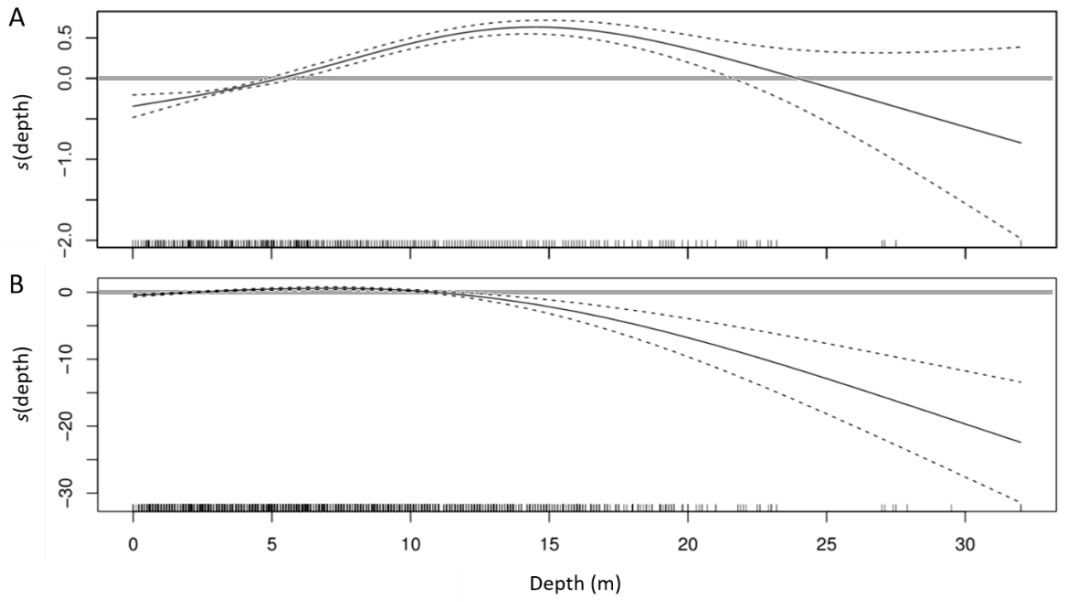

Figure 11. *Delta-GAM fitted smooth term for Total Polychaete (A) and GAM fitted smooth term of Nereididae (B) log(biomass) partial residuals to station depth. The inner tickmarks on x-axis denote observed depths; the y-axis is the value of the smoothed function of the loge(biomass) partial residuals across depths. Estimated degrees of freedom of the smooths are 3.38 and 1.97, respectively.*

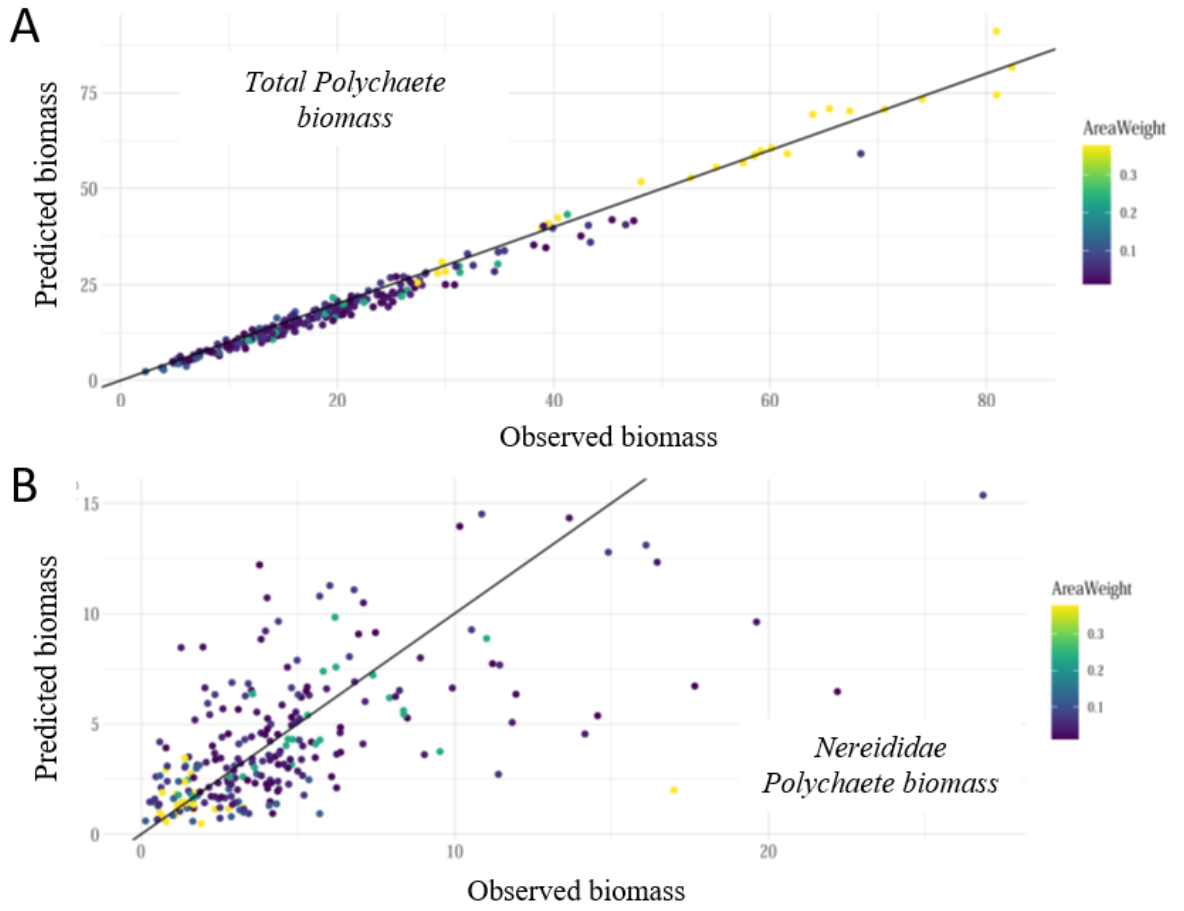

Figure 12. *Predicted versus observed biomass for the selected best model for Total Polychaete biomass (A) and Nereididae Polychaete biomass (B). Symbols are color-coded by statistical weighting based on geographical area of sampling strata. Straight black lines represent the case of ideal prediction accuracy.*

Final Polychaete indices based on the selected models show strong interannual variability across strata, similar to patterns observed in the raw data (**Figure 13**). Interannual fluctuations are less correlated among regional and system strata in the Total Polychaete index than in the Nereididae only index. This outcome likely is attributable to the wide range of environmental tolerances among all Polychaete taxa relative to tolerances of the Nereididae family) and their corresponding abilities to respond independently to annual environmental conditions in each stratum.

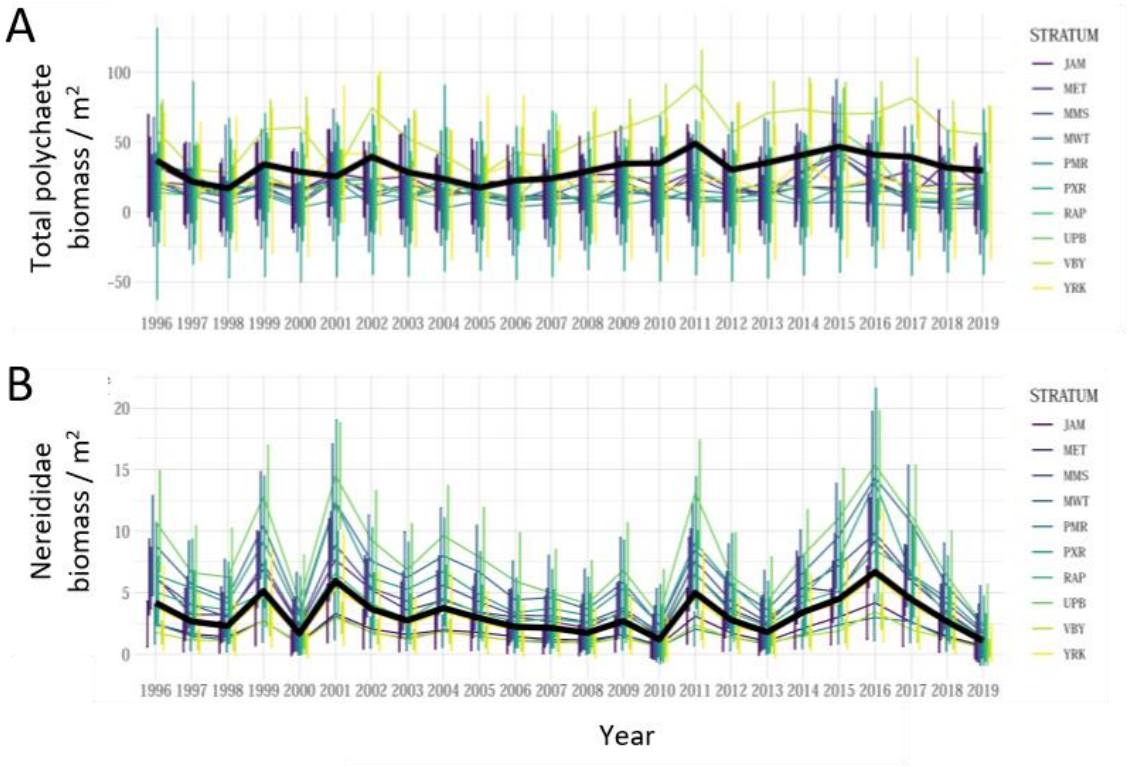

Figure 13. *Final time-series of Total Polychaete biomass (A) and Nereididae Polychaete biomass (B) from selected best models for each taxonomic group. Heavy black lines are area-weighted, Chesapeake Bay-wide annual index values.*

Our analysis of Total and Nereididae Polychaete biomass indicators did not lead us to reject the hypothesis of normality of the distributions (Shapiro-Wilk normality test $p = 0.90$ for total Polychaete and $p = 0.26$ for Nereididae biomass indicators; **Figure 14**). As such, quantiles may represent a useful approach for delineating qualitative ranges of the indicator values. Terciles were selected to provide three distinct biomass categories for each Polychaete indicator, with the upper, middle, and lower terciles including the upper, middle and lower 1/3 of the estimated biomasses from each dataset (**Figure 14**).

During the late 1990s and 2000s, the Total Polychaete indicator varied appreciably with annual values falling within each of the identified terciles (**Figure 14**). After 2007, the Total Polychaete indicator has been above the lower tercile and within the middle and upper tercile space since that year. Unlike the Total Polychaete indicator, the Nereididae indicator does not show an obvious trend over the analyzed period (**Figure 14**). There do appear to be two periods of high interannual variability in the time series, the first occurring from the mid-1990s to the early 2000s and the second occurring during the 2010s. The intervening period, consisting of the middle to late 2000s, was characterized by relatively stable, but declining, nereid biomass within the middle and lower terciles.

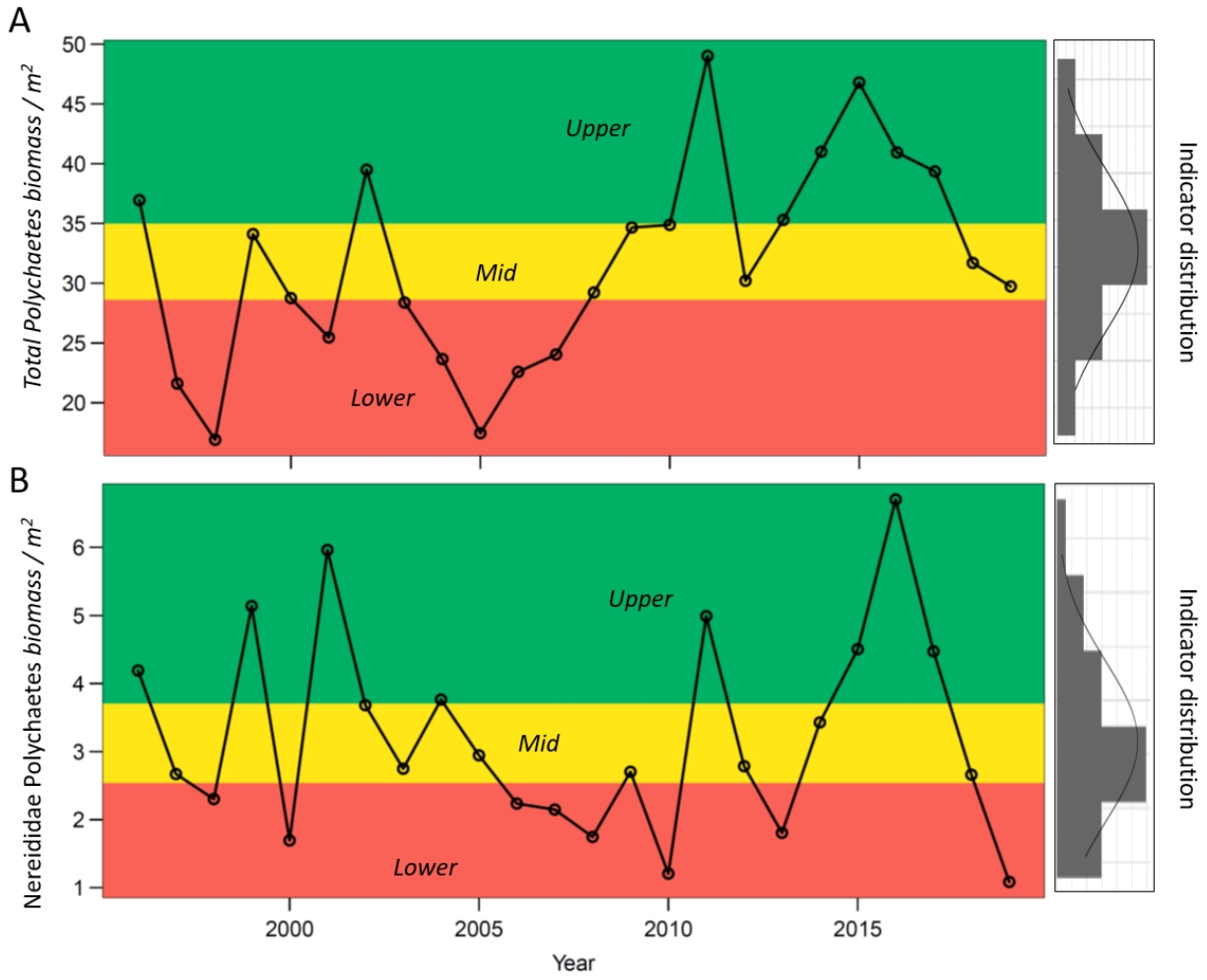

Figure 14. *Total Polychaete (A) and Nereididae (B) biomass indicators for the 1996–2019 time series with associated color-coded terciles representing upper (green), middle (yellow), and lower (red) thirds of the indicator distributions. To the right, frequency distributions with superimposed normal density curves are shown for each indicator.*

Bay Anchovy

We conducted an extensive evaluation of Bay Anchovy length data to assess the potential of using available length data to separate larger, mature individuals that comprise the spawning stock from smaller, young-of-the-year recruits produced in the same year. Bay Anchovy length data were available for 32 years in the standardized VA Juvenile Striped Bass Seine Survey (1988-2019). The MD Juvenile Striped Bass Seine Survey sampling sites recorded only minimum and maximum lengths of Bay Anchovy in survey years prior to 1991. Subsequently, no length data were collected in the MD seine survey. Accordingly, we initially focused on the

VA seine data to explore length thresholds that could define young-of-the-year (YOY) Bay Anchovy (i.e., recruits).

We elected to use the same method of separating Bay Anchovy into size-modes that was used by Jung and Houde (2004b). This method is based on representing the length distribution as a mixture of Gaussian distributions (Bhattacharya 1967). For our purposes we specified 2 clusters to represent recruits and spawning stock individuals, then fit the Gaussian distributions to aggregated monthly and total length distributions. Thresholds separating the modes were determined using two approaches. The first was a signal detection theory approach (TH) described by Little et al. (2013), and the second identified the lower 5% quantile of the adult mode ($TH_{5\%}$) to statistically assign smaller individuals to the recruit class (**Figure 15**). The separation methods yielded similar, although not identical, results. In the illustrated example, the signal detection approach indicated a threshold length (TH) of 55 mm while the 5% quantile approach ($TH_{5\%}$) indicated a threshold of 50 mm.

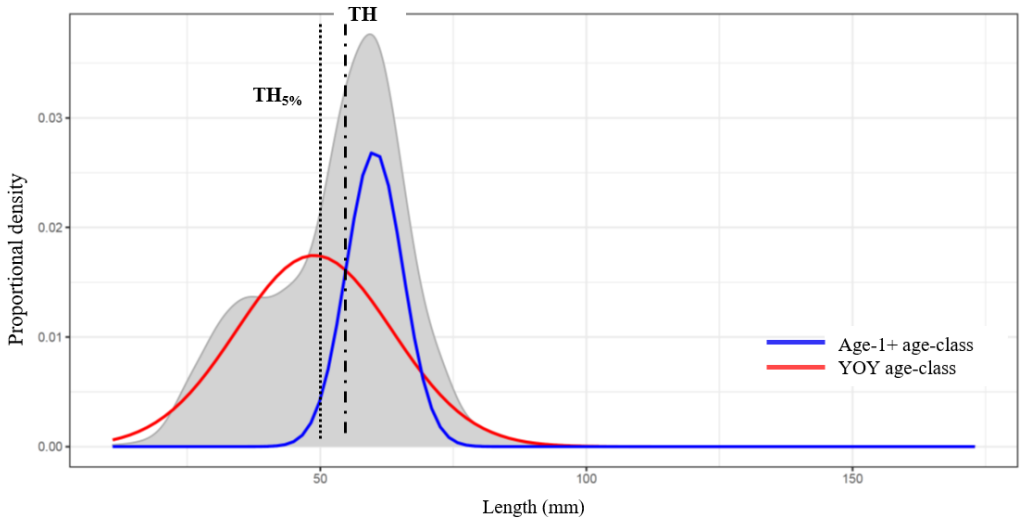

Figure 15. *Smoothed total size-structure of all Bay Anchovy measured in the VA Juvenile Striped Bass Seine Survey from 1988-2019 during June-July (shaded area). Gaussian distributions are fitted to aggregate size-structure. Examples of the recruit – spawning stock length thresholds for the two approaches are indicated (signal decomposition [TH]; 5% quantile of the adult mode [TH5%]).*

Based on our exploration and evaluation of these length-based threshold approaches for separating spawning stock from recruits (*see Appendix 2. Bay Anchovy indicators*), the evidence suggested that thresholds based on seasonal survey catches will likely yield better results (i.e., less uncertainty in life-stage classification). Further, the absence of complete length data records for the MD seine survey was an additional barrier to using length-based approaches. Based on the sampling months of the available trawl and seine survey datasets, the best estimates of recruit abundance were judged to come from August–October survey samples while estimates of spawning stock were best represented in April–May survey samples. A total Bay Anchovy index

(i.e., recruits plus spawning stock) was calculated from the April-October collections in the respective trawl and seine surveys.

TIES/CHESFIMS TRAWL SURVEY

In the combined TIES/ChesFIMS trawl survey, Bay Anchovy spawning stock ($BA_{Spawning}$), recruits ($BA_{Recruits}$), and total population (BA_{Total}), the catch frequency distributions were strongly right-skewed, requiring $\log_e$-transformation to approach normality (**Figure 16**).

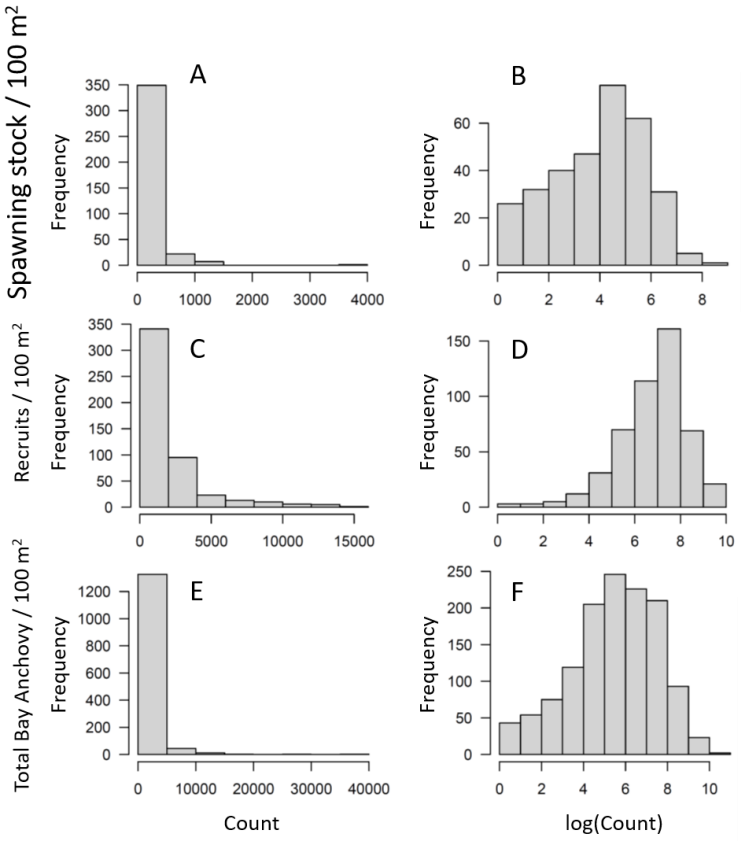

Figure 16. *Bay Anchovy spawning stock (A, B), recruits (C, D), and total population (E, F) counts (scaled to counts per 100 m2) distributions on observed (A, C, E) and loge-transformed (B, D, F) scales.*

Three models were evaluated for developing indices of relative abundance for $BA_{Spawning}$, $BA_{Recruits}$, and BA_{Total} in the TIES/ChesFIMS dataset. These included a GLM, Delta-GLM, and RF model. Model fitting and evaluation (e.g., cross-validation) followed the same protocols described above for Polychaetes. Observations were weighted by the habitat volume of the mainstem areas represented (volume-weighted estimation; observations from larger mainstem areas were weighted more heavily than smaller areas); the weighting was also applied in calculating model-evaluation metrics. Results from the model fitting indicated RF models outperformed GLM and Delta-GLM models for the Bay Anchovy life-stage groups (**Table 3**). In all cases, the RF model was associated with the lowest model errors and highest R^2 values.

Table 3. *Cross-validation (out-of-sample) model performance metrics for Bay Anchovy spawning stock (BASpawning), recruits (BARecruits), and total population (BATotal) indices. Model error metrics include MAE (mean absolute error) and RMSE (root mean square error); R2 is coefficient of determination. All metrics were computed using weights associated with mainstem region volumes. Yellow highlighted cells indicate the best-performing models for each metric; green cells indicate the second-best performance metric.*

		MAE	RMSE	R2
$BA_{Spawning}$	GLM	148.0490	255.3285	0.10733873
	Delta-GLM	146.5676	255.3279	0.08864828
	RF	140.5446	246.1819	0.15684995
$BA_{Recruits}$	GLM	1474.637	2212.391	0.05623064
	Delta-GLM	1475.005	2211.617	0.05701652
	RF	1284.825	1951.971	0.27424106
BA_{Total}	GLM	1078.382	2036.798	-0.01233780
	Delta-GLM	1078.387	2036.352	-0.01177806
	RF	1029.034	1963.187	0.05586577

In-sample predictions of Bay Anchovy relative abundances indicated good agreement with observed abundances (**Figure 17A**). In-sample predictions for the $BA_{Spawnign}$ and $BA_{Recruits}$ indices were somewhat less accurate than the BA_{Total} indices, with these models showing a slight tendency to overestimate catch at low observed values (**Figure 17B, C**). The BATotal showed no evidence of bias in terms of under- or overestimation of abundance at low or high values. Occurrences of model under- and over-estimation of relative abundances were most prevalent from the upper bay mainstem region that had the smallest area-weighting (**Figure 17A, B**).

Final Bay Anchovy abundance indices based on the RF models show strong interannual variability across regional strata, similar to patterns observed in the raw data (**Figure 18**). Rank-order of regional annual index values were relatively consistent within life stages but differed between life stages. The annual $BA_{Spawning}$ index was typically highest in the Lower Bay; the Middle Bay usually had intermediate annual index values, and the Upper Bay consistently had the lowest annual index values (**Figure 18A**). In contrast with the $BA_{Spawning}$ index, the $BA_{Recruits}$ index was highest in the Middle Bay, with the Upper and Lower Bay regions having similar index values that varied in their rank order from year to year (**Figure 18B**). Regional patterns in rank-order of the BA_{Total} index were a composite of the $BA_{Spawning}$ and $BA_{Recruits}$ indices, with different regions having highest index values in the bay mainstem across years (**Figure 18C).**

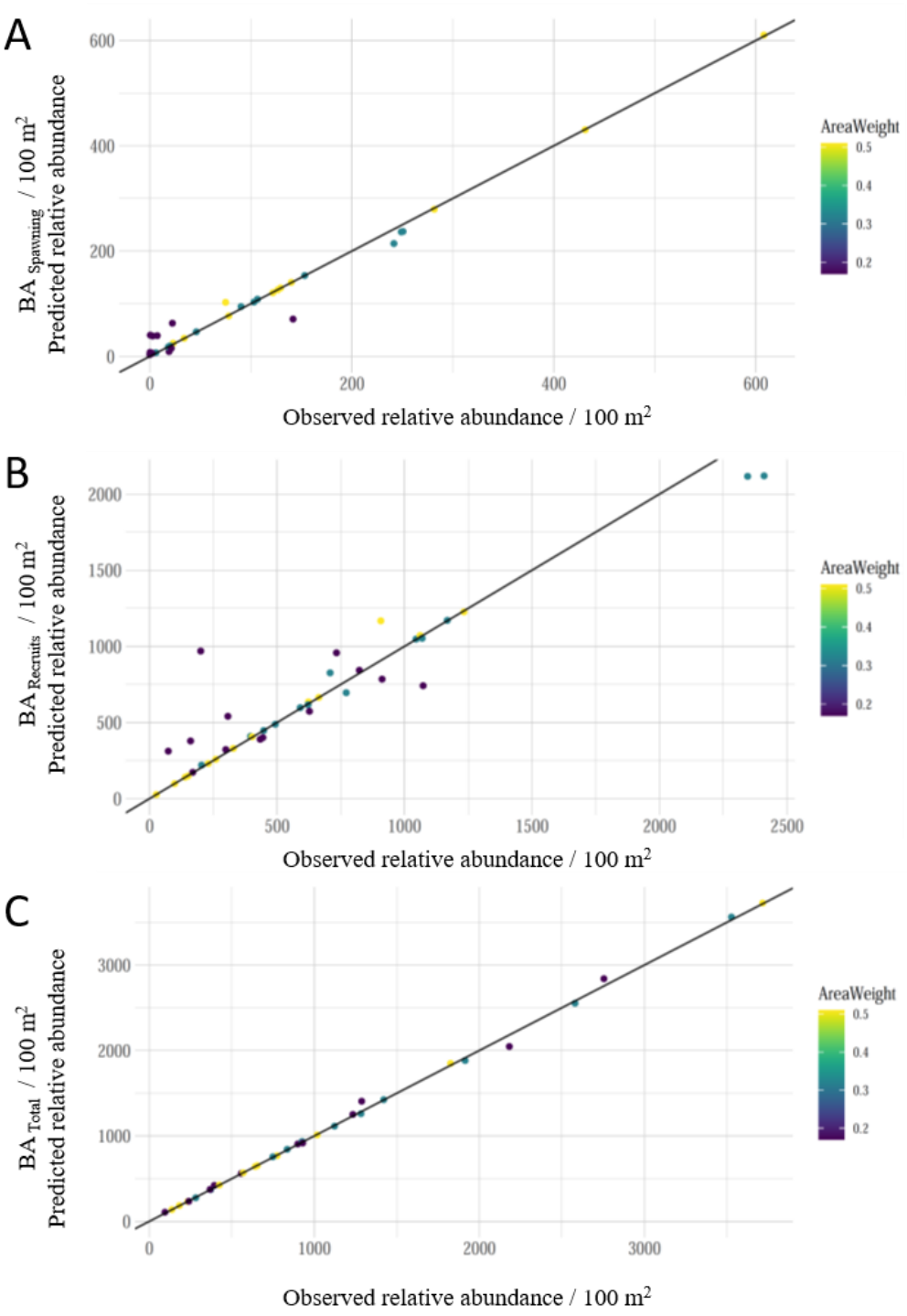

***Figure 17**. Predicted versus observed abundance for the selected best model for Bay Anchovy spawning stock ($BA_{Spawning}$; A), recruits ($BA_{Recruits}$; B), and total population (BA_{Total}; C). Symbols are color-coded by statistical weighting based on bathymetric volume of three mainstem regions. Straight black lines represent the case of ideal prediction accuracy.*

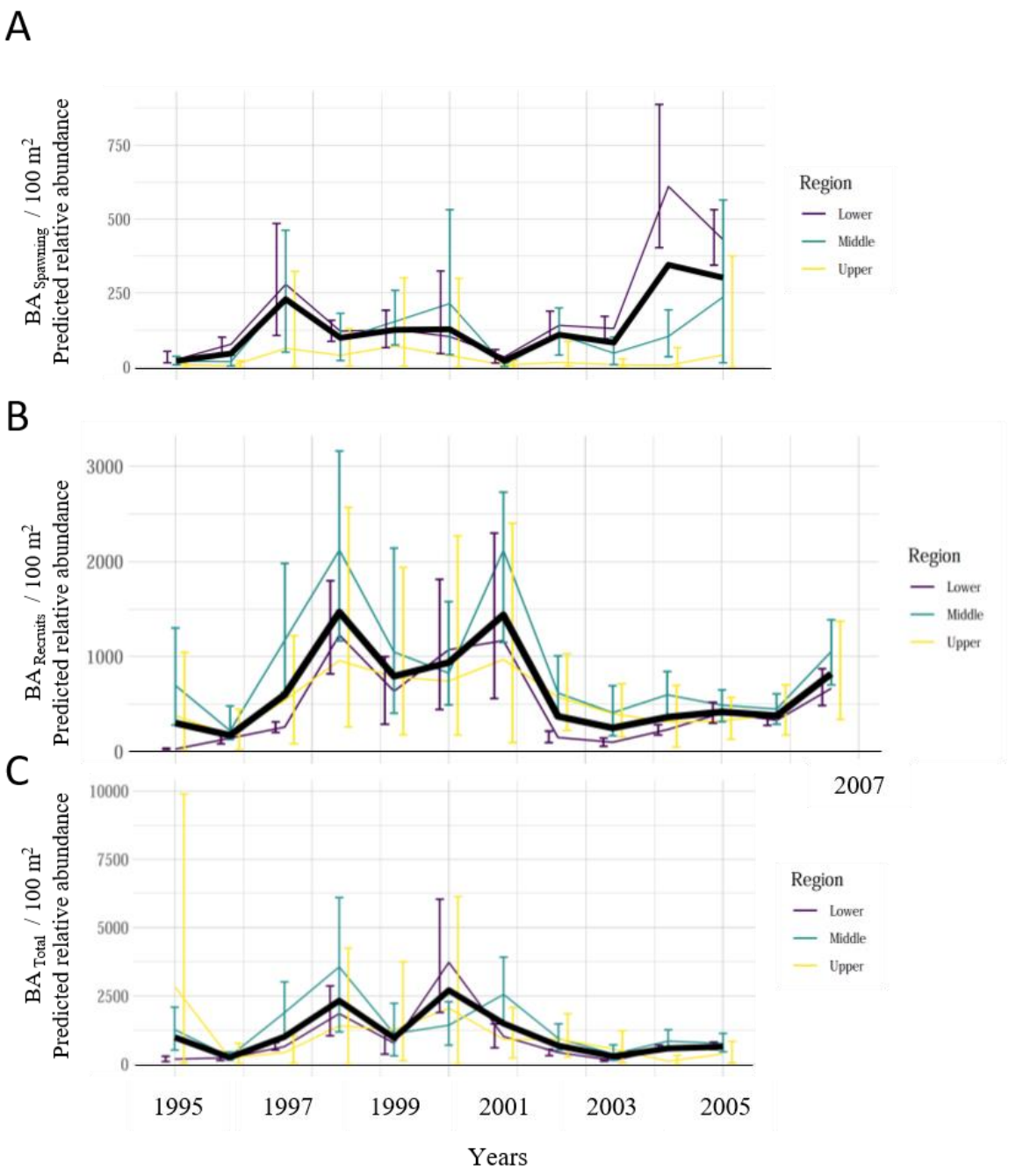

Figure 18. *Time-series of Bay Anchovy spawning stock ($BA_{Spawning}$; A), recruits ($BA_{Recruits}$; B), and total population (BA_{Total}; C) relative abundances. Because April-May sampling was not conducted by ChesFIMS in 2006 & 2007, estimation of $BA_{Spawning}$ and BA_{Total} indicator values was not possible for those years. Heavy black lines are the volume-weighted, Chesapeake Bay-wide annual index values.*

Our analysis of TIES/ChesFIMS $BA_{Spawning}$, $BA_{Recruits}$, and BA_{Total} relative abundances indicated general support for the assignment of indicator values to classes based on terciles. The

distribution of $BA_{Recruits}$ showed some evidence of non-normality (Shapiro-Wilk normality test $p = 0.03$) but both the $BA_{Spawning}$ and BA_{Total} distributions were normally distributed (Shapiro-Wilk normality test $p \geq 0.06$; **Figure 19**). To examine the effect of an alternative classification method on the $BA_{Recruits}$ indicator, *k*-means clustering was used to classify indicator values into $n = 3$ groups (**Figure 20**). Although thresholds for High, Medium, and Low group assignments were similar between the tercile and *k*-means approaches, the two approaches did result in slightly different classifications for 3 years – 1999, 2002, and 2004.

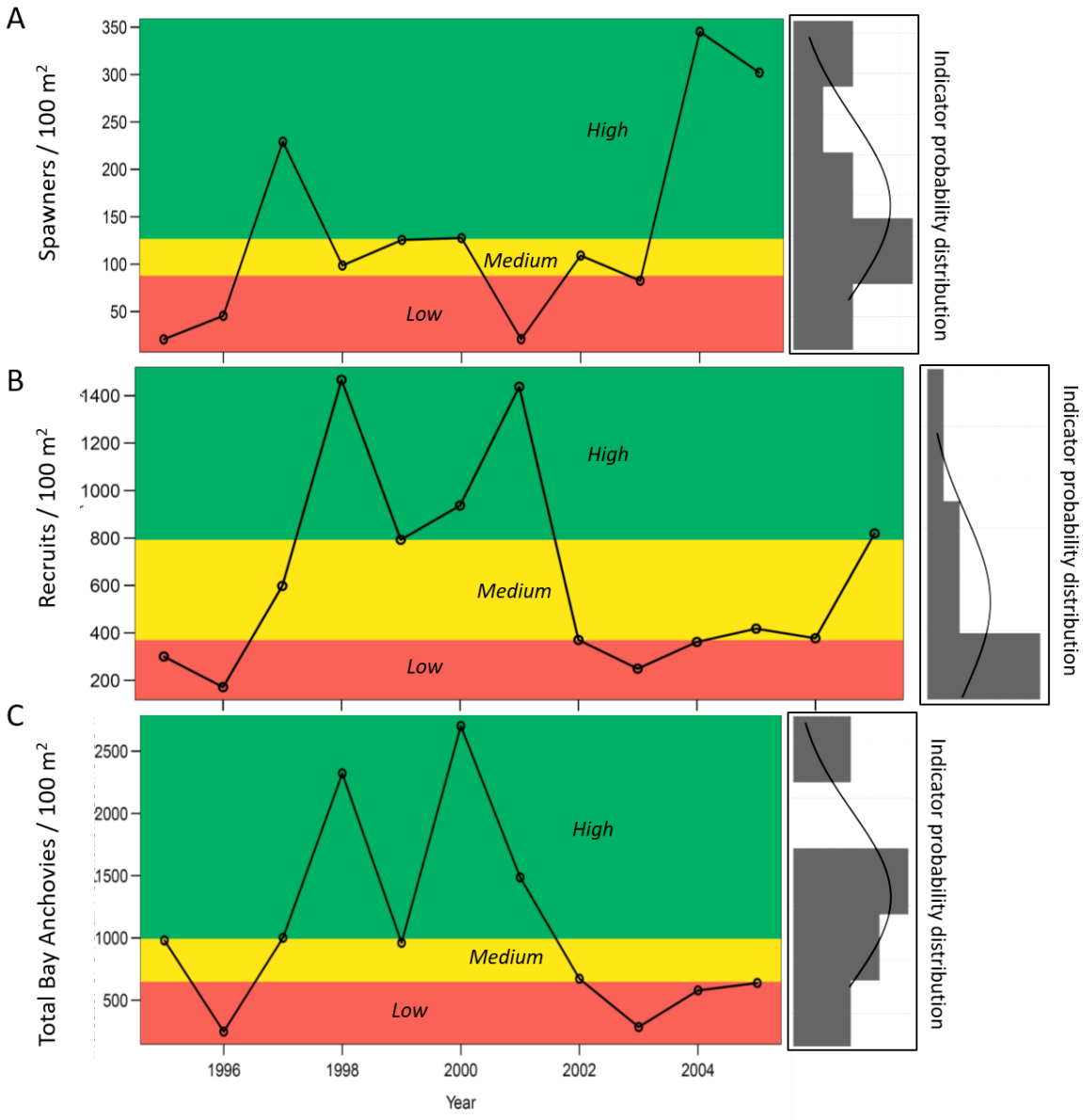

Figure 19. *Bay Anchovy Spawning Stock (Spawners, A), Recruits (B) and Total (C) indicators for the 1995–2005 time series (Spawners, Total) and the 1995 – 2007 time series (Recruits) with associated color-coded terciles representing High (green), Medium (yellow), and Low (red) values of the indicator distributions. To the right, probability distributions with superimposed normal density curves are shown for each indicator.*

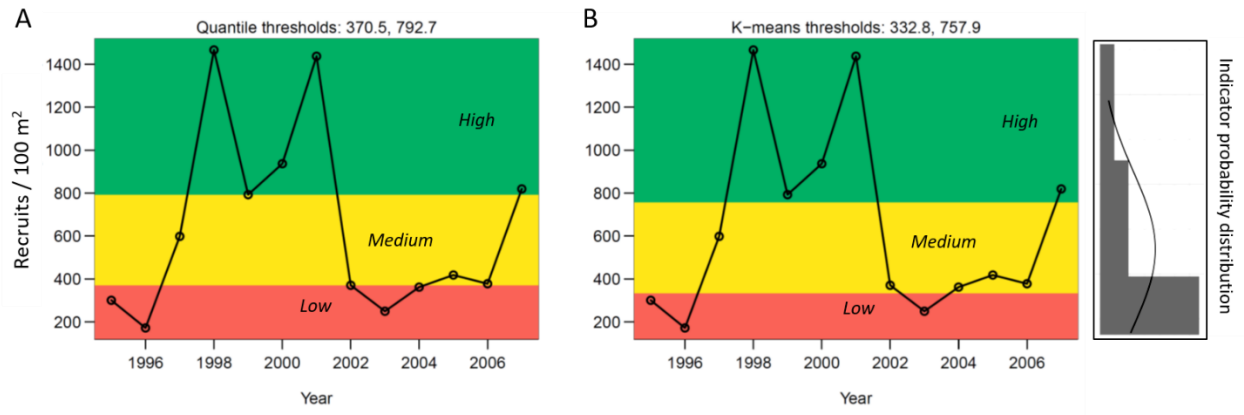

Figure 20. *Bay Anchovy Recruits indicator for the 1995–2007 time series with associated color-coded terciles (A) and k-means clustering groups (B) representing High (green), Medium (yellow), and Low (red) values of the indicator distributions. To the right, the probability distribution with superimposed normal density curve is shown for the indicator.*

Higher than average $BA_{Recruits}$ values were observed from 1998 to 2001 in the TIES/ChesFIMS indicator series (**Figure 20**). These years followed a high $BA_{Spawning}$ year in 1997, with $BA_{Spawning}$ remaining in the High to Medium indicator value range from 1998 to 2000 before dropping to one of the lowest observed values of the time series in 2001. Recruitment remained Low to Medium throughout the remainder of the time series, until 2007 when $BA_{Recruits}$ increased to a High ranking in the final year of the survey. Interannual patterns of the BA_{Total} indicator were similar to $BA_{Recruits}$, in part because of the relatively high abundances and contributions of YOY individuals to the BA_{Total} index.

MD/VA SEINE SURVEY

In the combined MD and VA seine surveys, the $BA_{Recruits}$ and BA_{Total} catch-frequency distributions were strongly right-skewed, requiring $\log_e$-transformation to approach normality (**Figure 21**).

Three statistical models were evaluated to develop indices of relative abundance for $BA_{Recruits}$ and BA_{Total} in the MD and VA seine survey datasets. These included a GLM, Delta-GLM, and RF model. Model fitting and evaluation (e.g., cross-validation) followed the same protocols described above for Polychaetes and for TIES/ChesFIMS Bay Anchovy. A categorical identifier for State (MD, VA) and a State × Year interaction term were included in the models, allowing model fitting to use data from both seine surveys while also allowing State-specific model output. Results from the model fitting indicated that the RF models outperformed GLM and Delta-GLM models for the two Bay Anchovy life-stage groups (**Table 4**). In both cases, the RF model was associated with the lowest model errors. All model R^2 values were negative for the $BA_{Recruits}$ models while the highest (and only positive) R^2 value for the BA_{Total} data was associated with the RF model.

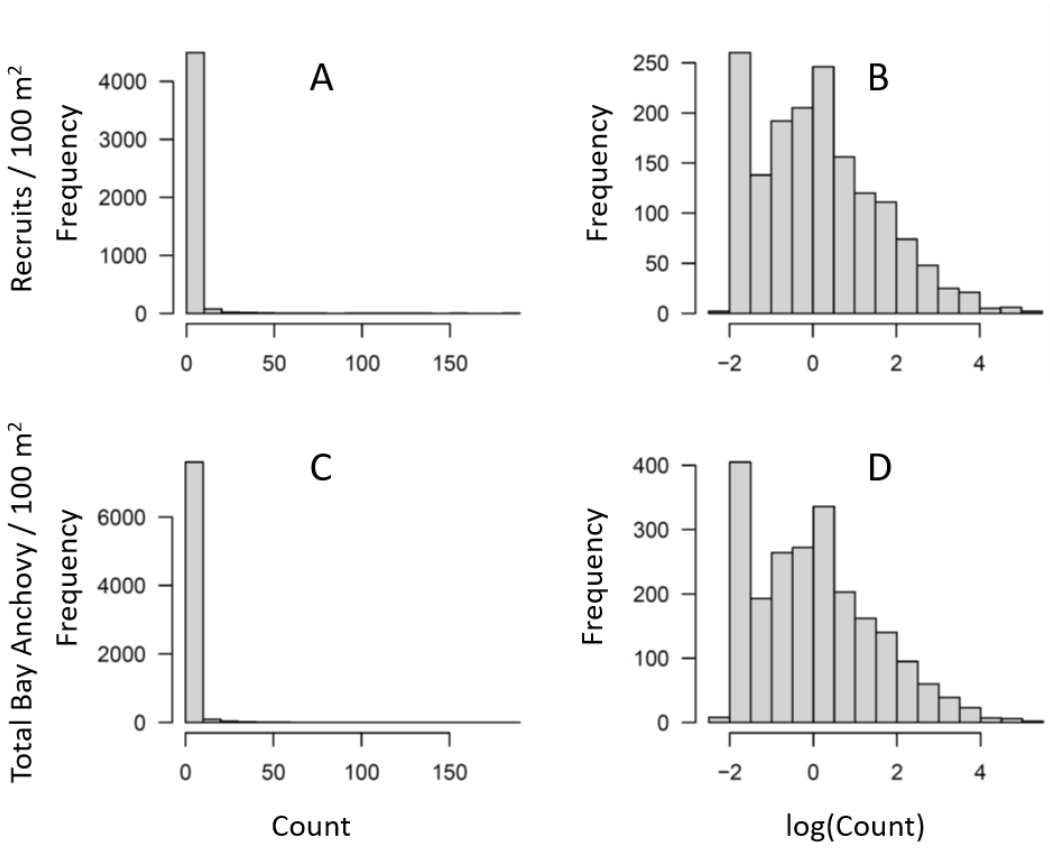

Figure 21. *Frequency distributions of observed Bay Anchovy recruit (A, B) and total (C, D) counts per 100 m^2 on the observed (A, C) and loge-transformed (B, D) scales.*

Table 4. *Cross-validation (out-of-sample) model performance metrics for Bay Anchovy recruits ($BA_{Recruits}$) and total population (BA_{Total}) indices. Model error metrics include MAE (mean absolute error) and RMSE (root mean square error); R^2 is the coefficient of determination. All metrics were computed using weights associated with mainstem region volumes. Yellow highlighted cells indicate the best-performing models for each metric; green cells indicate the second-best performance metric.*

		MAE	RMSE	R2
$BA_{Recruits}$	GLM	2.146887	6.568838	-0.004021980
	Delta-GLM	2.144938	6.566233	-0.002889325
	RF	2.100419	6.562114	-0.003235273
BA_{Total}	GLM	1.660756	5.410478	-0.000880687
	Delta-GLM	1.659923	5.409403	-0.000476068
	RF	1.616987	5.388567	0.007814419

In-sample predictions for the RF models applied to the $BA_{Recruits}$ and BA_{Total} indices were accurate, with the fitted models suggesting some evidence of overfitting (**Figure 22**). There were no obvious trends in under- or over-estimation.

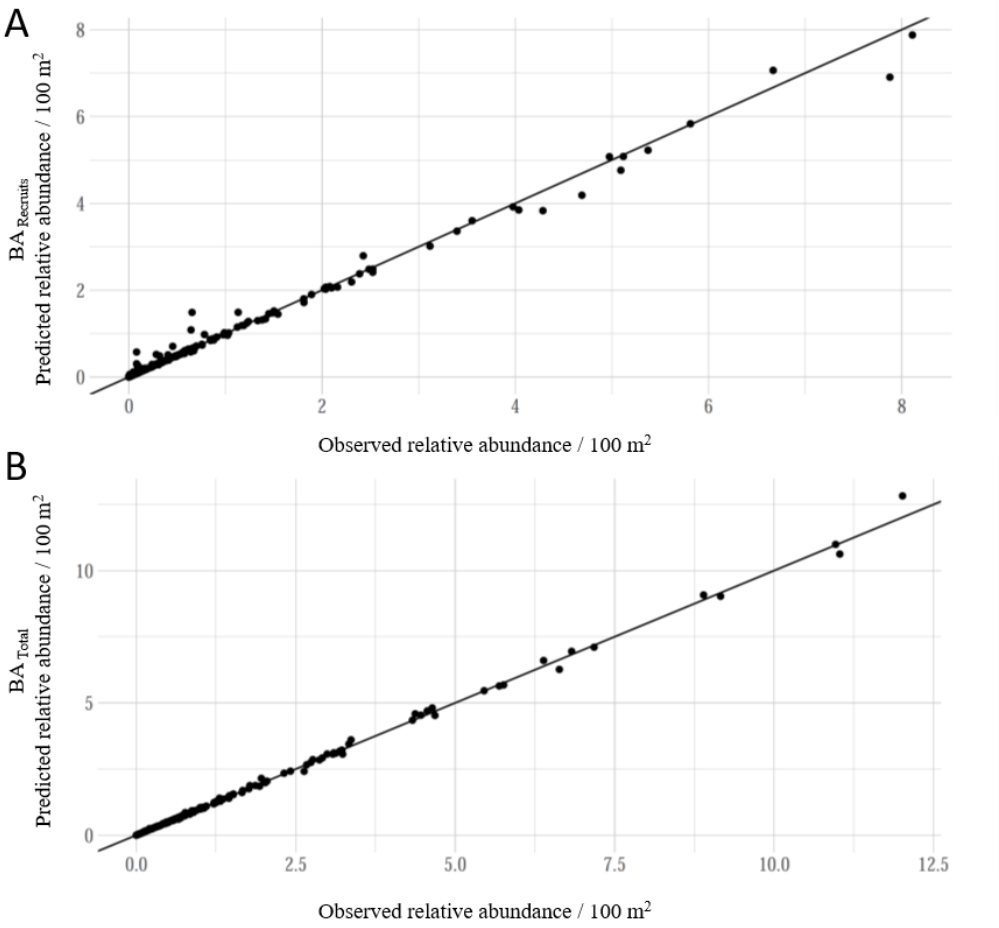

Figure 22. *Predicted versus observed abundance (counts per 100 m^2) for the selected best model for Bay Anchovy recruits ($BA_{Recruits}$; A), and total population (BA_{Total}; B). Maryland and Virginia seine survey data are included in the estimation of these indices. Straight black lines represent the case of ideal prediction accuracy.*

Final Bay Anchovy abundance (/100 m^2) indices based on the RF models show strong interannual variability between states (**Figure 23**). The rank order of annual $BA_{Recruits}$ and BA_{Total} index values from MD and VA shifted over time but MD estimates (and associated 95% CIs) were often substantially higher than the VA estimates. Within states, there was a strong correlation in interannual patterns and rank order between the $BA_{Recruits}$ and BA_{Total} indices (**Figure 23).**

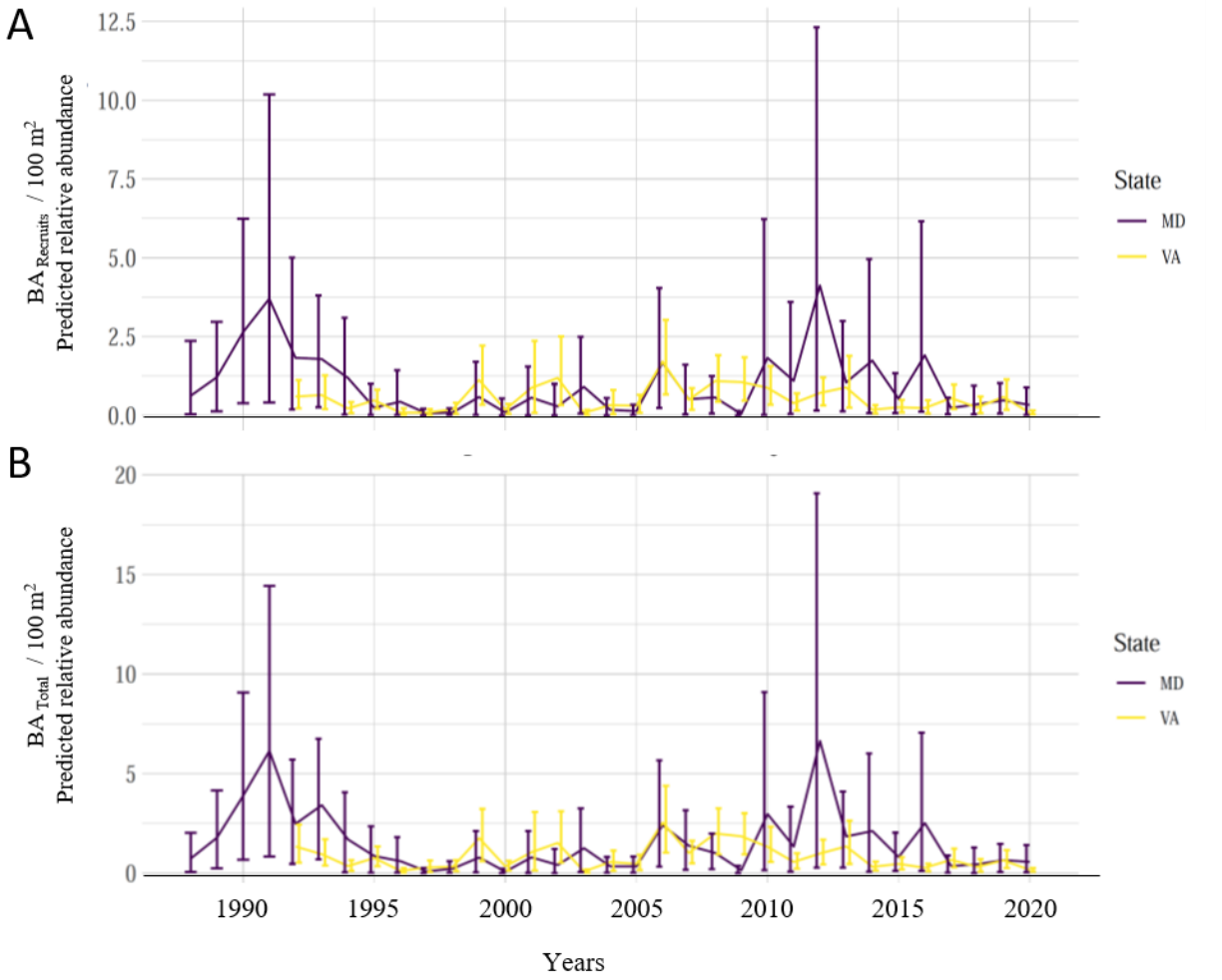

Figure 23. *Time-series of Bay Anchovy recruits ($BA_{Recruits}$; A), and total population (BA_{Total}; B) relative abundances (counts per 100 m^2 ± 95% CI) from the Maryland (MD) and Virginia (VA) seine surveys. Estimates of BA indices are excluded from the 1988-1991 VA time-series because data on the offshore extension of the seine net were not available in those years.*

The MD and VA seine survey distributions of $BA_{Recruits}$ and BA_{Total} were not normally distributed (Shapiro-Wilk normality tests $p \leq 0.018$); therefore, all seine-based indicators were classified using *k*-means clustering (**Figure 24**). In MD, the $BA_{Recruits}$ indicator peaked during two periods, 1990–1991 and in 2012. These peak periods were preceded and succeeded by several years of moderate and variable $BA_{Recruits}$ values, with an extended period of low recruit indicator values from 1995 to 2005. The MD BA_{Total} indicator time series was similar to the recruitment indicator due to absence of spring months sampling and the resulting dominant effect of recruitment during late summer on BA_{Total} estimates.

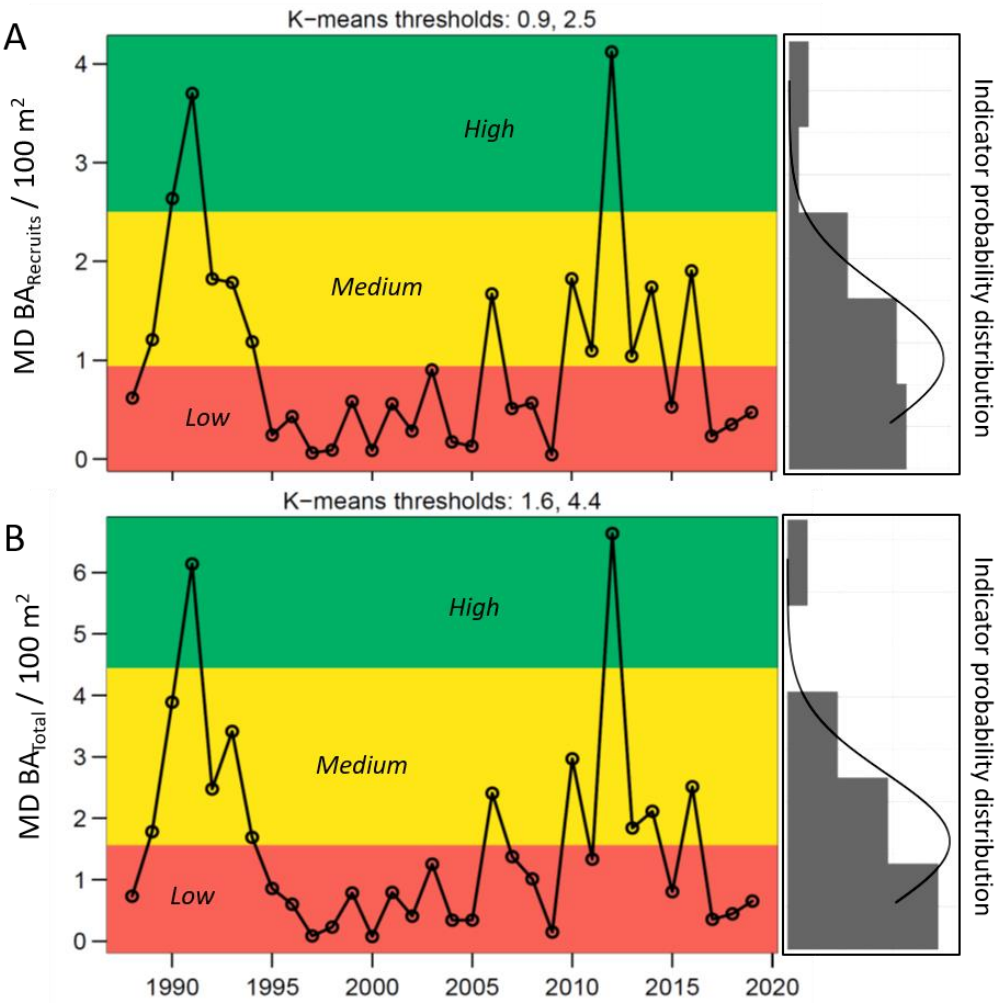

Figure 24. *Maryland Bay Anchovy recruits (MD BA$_{Recruits}$; A), and total population (MD BA$_{Total}$; B) indicators from the seine survey for the 1988–2019 time series with associated k-means clustering groups representing High (green), Medium (yellow), and Low (red) values of the indicator distributions. To the right, probability distributions with superimposed normal density curves are shown for each indicator.*

In VA, the seine time series for the BA$_{Recruits}$ indicator was less structured than the corresponding MD time series (**Figure 25**). The VA BA$_{Recruits}$ indicator was classified as High in 1999, 2006, 2008, and 2009. Years with high BA$_{Recruits}$ values were interspersed with years of Medium and Low classifications. In several instances, High recruitment indicator years were immediately followed by Low or Medium recruitment indicator years (e.g., 1999-2000, 2002-2003, 2006-2007). Early in the time series, the VA BA$_{Recruits}$ indicator declined from moderate to consistently Low for five years until 1999. In the final years of the time series, the BA$_{Recruits}$ indicator also was consistently Low. As in the MD seine indicators, the VA BA$_{Total}$ time series was similar to the VA BA$_{Recruits}$ time series because spring sampling that would have primarily captured spawners was not included in the time series.

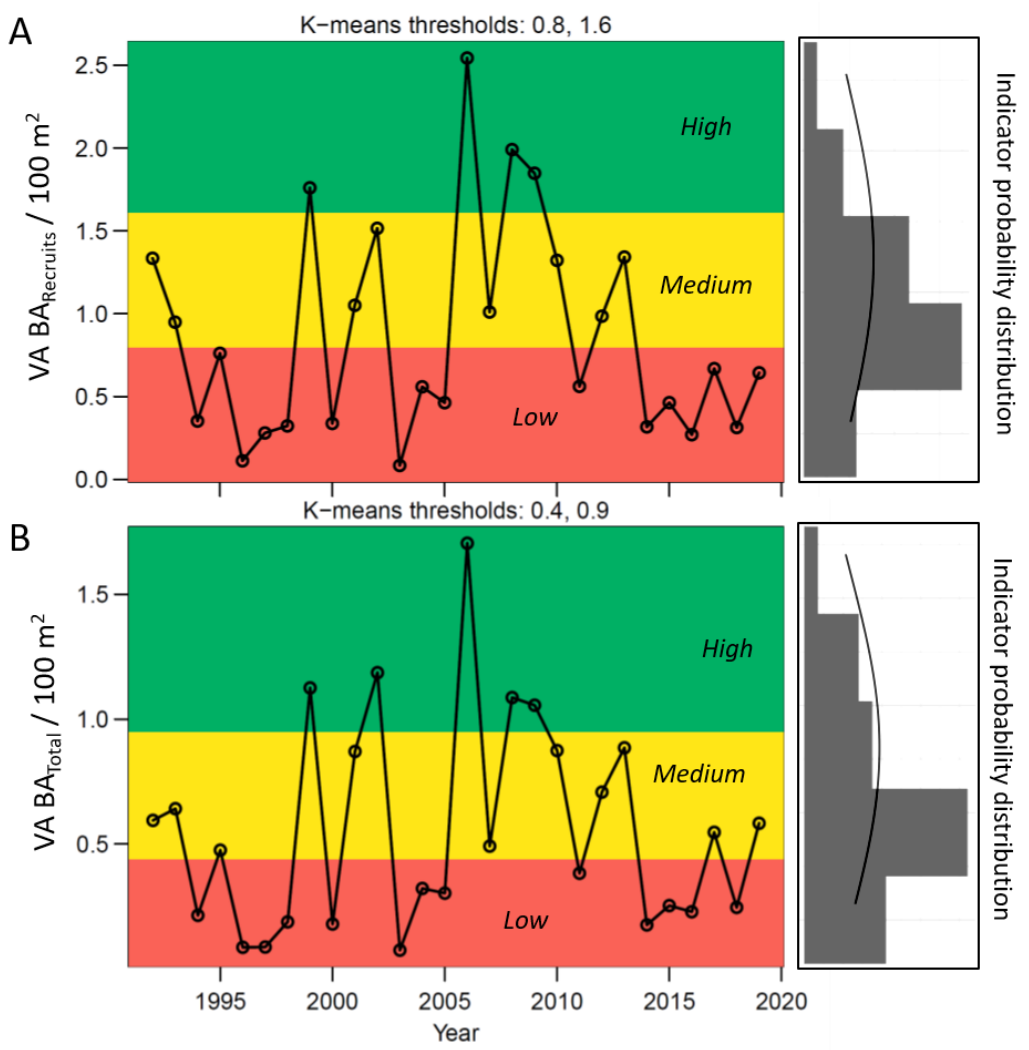

Figure 25. *Virginia Bay Anchovy recruits (VA $BA_{Recruits}$; A), and total population (VA BA_{Total}; B) indicators from the seine survey for the 1991–2019 time series with associated k-means clustering groups representing High (green), Medium (yellow), and Low (red) values of the indicator distributions. Net extension data from 1988-1991 were unavailable, making area-corrected estimates of BA abundance not possible for those years. To the right, probability distributions with superimposed normal density curves are shown for each indicator.*

VIMS TRAWL SURVEY

Annual indices of Bay Anchovy spawning stock and recruitment provided by project collaborators for the Lower Chesapeake Bay mainstem showed strong year-to-year variability (**Figure 26**). Our index based on the VIMS Spawning Stock dat (median = 5.7 ± 4.5 SD) was approximately 6-fold lower than the VIMS Recruit Index (31.0 ± 65.5) and the two indices were not correlated (r_P = 0.07, p > 0.05). Peak index years for Spawning Stock and Recruit indices differed, with a maximum Spawning Stock index occurring in 2004, and a maximum Recruit index occurring in 2010 (**Figure 26**).

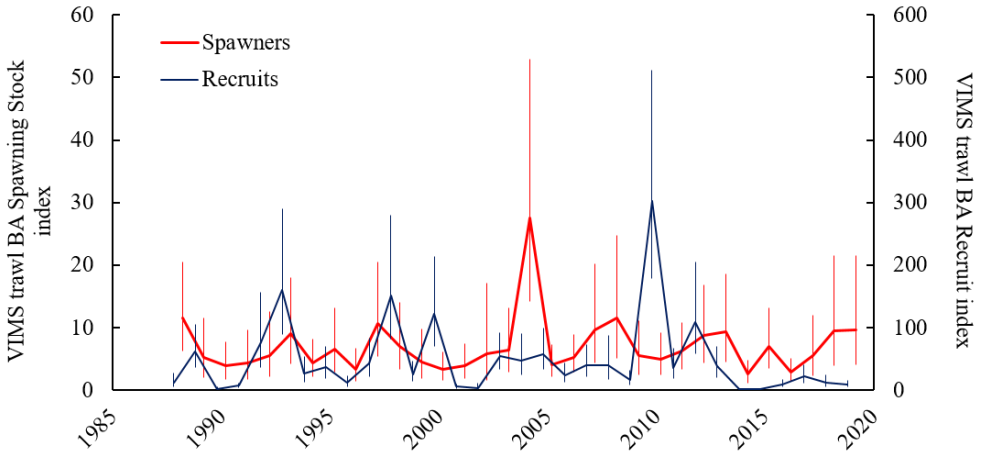

Figure 26. *Time-series of annual Bay Anchovy spawning stock (Spawners, primary y-axis) and recruit (Recruits, secondary y-axis) indices (± 95% CI), derived from the Virginia Institute of Marine Science (VIMS) trawl survey data in the Lower Chesapeake Bay mainstem.*

The VIMS Trawl Spawning Stock index was positively correlated with the lower Bay TIES/ChesFIMS $BA_{Spawning}$ index (n =11, r_P = 0.71, p = 0.014); however, the VIMS trawl Recruit index was not correlated with the lower Bay TIES/ChesFIMS $BA_{Recruit}$ index (n =11, r_P = 0.49, p = 0.126; **Figure 27**). The 2005 datum for the lower Bay TIES/ChesFIMS $BA_{Spawning}$ index (= 405.5) and the VIMS Spawning Stock index value (= 4.09) was an obvious outlier (**Figure 27B**). Removing the 2005 outlier increased the correlation between the two Spawning Stock indices to r_P = 0.97 (n = 10, p < 0.0001). The 2004 spawning cohort was the most abundant in both the TIES/ChesFIMS lower Bay survey and VIMS Spawning Stock index (**Figure 27B**).

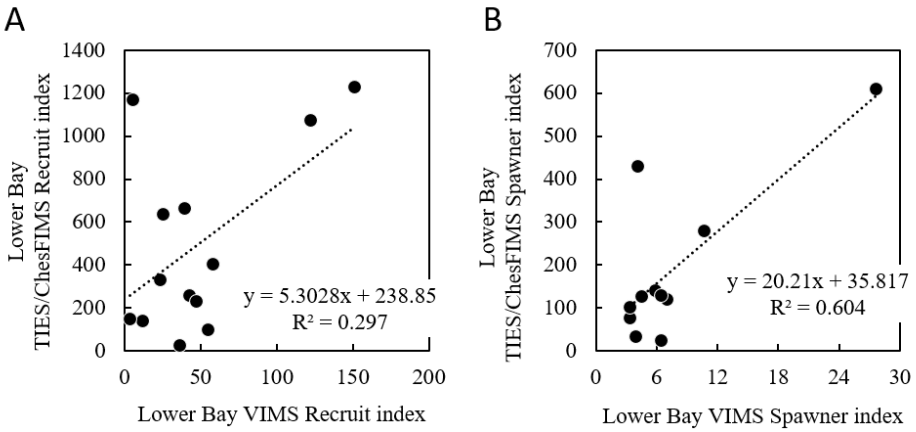

Figure 27. *Bay Anchovy abundance indices from the TIES/ChesFIMS survey (y-axes) plotted on the VIMS trawl indices (x-axes) for Recruits (A) and Spawning Stock (B). TIES/ChesFIMS indices plotted here correspond to the lower Chesapeake Bay region only. Dotted regression lines, linear regression equations, and associated R^2 values are provided for reference*

Ratios of recruits to spawning stock (R/SS) based on the TIES/ChesFIMS data were calculated to evaluate the utility of the R/SS metric as a potential indicator of overall reproductive success. At the basin-scale, R/SS values ranged from 1.0–68.7 recruits/spawner during the 1995-2005 interval (**Figure 28A**). Ratio values differed consistently among the three mainstem regions, with the Upper Bay > Middle Bay > Lower Bay in all years (**Figure 28B**). A comparable range of R/SS values was observed in the longer time series of the VIMS trawl survey (**Figure 28C**). The R/SS time series in both the TIES/ChesFIMS and the VIMS trawl surveys indicated several years of high recruit indices relative to spawner indices during the late 1990s to early 2000s interval. The VIMS trawl survey data indicate other clusters of years with elevated R/SS in the lower Bay during the early 1990s and 2000s, with a notable peak in 2010.

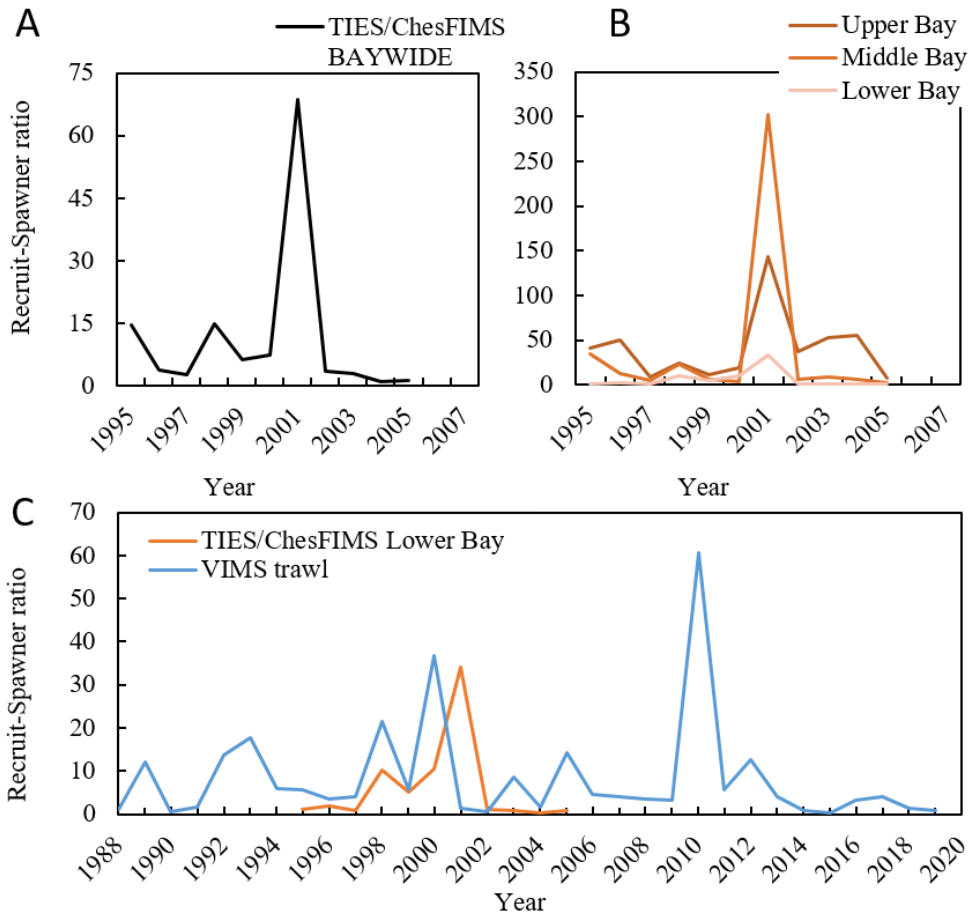

Figure 28. *Bay Anchovy Recruit/Spawning Stock (R/SS) index ratios from the TIES/ChesFIMS survey and the VIMS trawl survey. Annual Recruit-Spawner ratios from the TIES/ChesFIMS survey for the entire Chesapeake Bay (A) and regionally for the Upper, Middle, and Lower Bay (B). Recruit-Spawner ratios for the full 1988-2019 VIMS trawl survey and the Lower Bay TIES/ChesFIMS survey (1995-2005) were not correlated (C; $r_P = 0.08$, $p > 0.05$).*

Climate indicators

Atlantic Multidecadal Oscillation

Over the 1988–2019 interval, the annual AMO indicator was initially in a negative phase through the mid-1990s (**Figure 4**). The AMO oscillated between positive and negative during the late 1990s before entering and remaining in a positive phase during the remainder of the time series. Despite the consistently positive phase of the AMO over most of the time series, there was substantial variability in the indicator value (i.e., ranging ~0.01–0.35).

Degree Day index

Visual inspection of both DD indicators suggests a declining trend over time (**Figure 6**). Linear regression indicated a negative trend over time in the DD10 indicator ($F_{df\,(\text{num, den})\,=\,1,\,30} = 10.2$, $R^2 = 0.23$, $p = 0.003$) but not the DD5 indicator ($F_{1,\,30} = 3.91$, $R^2 = 0.09$, $p = 0.057$). The failure to detect a trend in the DD5 indicator differs from previous findings (Woodland et al. 2021) and almost certainly resulted from the focus of our present analysis on the short temporal interval of 1988–2019 (lower sample size and, hence, lower statistical power) relative to the > 60-year time series analyzed previously. Pearson product-moment correlation analysis indicated the DD5 and DD10 indicators were positively correlated ($r_P = 0.72$, $n = 32$, $p < 0.05$).

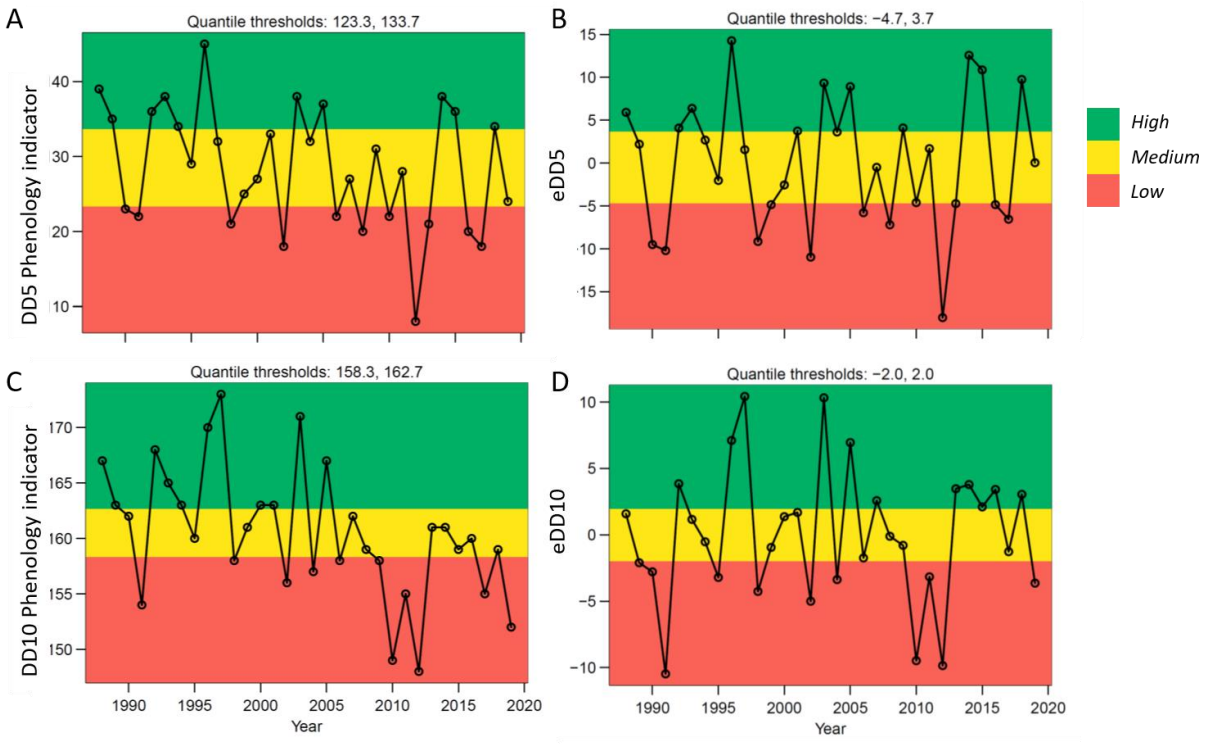

Figure 29. *Time series of the 5°C degree day indicator (DD5, A), detrended DD5 (eDD5, B),10°C degree day (DD10, C), and detrended DD10 (eDD10, D) from 1988-2019 in Chesapeake Bay with associated tercile-based classifications representing High (green), Medium (yellow), and Low (red) values of the indicator distributions.*

Tercile-based classifications were calculated for each variant of the DD indicators (including DD5, DD10, and the corresponding detrended versions of each: eDD5, eDD10; **Figure 29**). Despite an apparent qualitative decline in DD5 indicator values over the 1988–2019 interval, years with Low indicator values were present early in the time series and years with High indicator values were present late in the time series (**Figure 29A**). Patterns in the DD10 indicator were less evenly distributed through time, particularly late in the time series, with no years attaining a ranking of High after 2005 (**Figure 29C**). Detrended versions of DD5 and DD10 were variable through time with no immediately obvious patterns in years assigned to High, Medium or Low rankings (**Figure 29B, D**).

Forage-Climate models

Polychaetes

Results from the model fitting indicated that the GAM, using the DD10 indicator, performed best for the total Polychaetes, with the lowest associated model errors and highest R^2 value (**Table 5**). Results from the Nereididae modeling did not indicate strong associations with climate indicators, with the DD5 GLM having the lowest associated error metrics. All Nereididae indicator models resulted in negative out-of-sample (prediction) R^2 values. All Polychaete indicator climate models that included detrended versions of the DD variables performed poorly (see *Appendix 3. Forage-climate modeling*). Based on the results, we selected the DD10 GAM as the best version of the total Polychaete indicator climate model. No additional interpretation is provided for the Nereididae-Climate model because of the absence of interpretable relationships between the Nereididae and climate indicators.

Table 5. *Cross-validation (out-of-sample) model performance metrics for Total and Nereididae Polychaete Indicator climate models. Results associated with each DD indicator variant are indicated by the column header suffix (DD5, DD10). All detrended DD indicator variant models provided low predictive power and are not shown here (see Appendix 3. Forage-climate modeling for results). Model error metrics include MAE (mean absolute error) and RMSE (root mean square error); R^2 is the coefficient of determination. Yellow highlighted cells indicate the best-performing model.*

		MAE_DD5	RMSE_DD5	R2_DD5	MAE_DD10	RMSE_DD10	R2_DD10
Total	GLM	6.415	7.877	0.116	6.281	7.600	0.177
	GAM	6.688	8.075	0.071	5.915	7.432	0.213
	RF	7.074	9.024	-0.160	6.524	8.226	0.036
		MAE_DD5	RMSE_DD5	R2_DD5	MAE_DD10	RMSE_DD10	R2_DD10
Nereididae	GLM	1.268	1.580	-0.184	1.301	1.631	-0.262
	GAM	1.276	1.598	-0.211	1.717	2.300	-1.508
	RF	1.306	1.621	-0.247	1.327	1.621	-0.247

Within the DD10 GAM, total Polychaete biomass had a positive relationship with AMO_L1, with values increasing at low to intermediate levels of AMO_L1 (**Figure 30A**). Increased uncertainty at the highest AMO_L1 values resulted in confidence intervals that overlap 0 at AMO_L1 ≥ 0.35. There was a significant interaction between DD10 and AMO, with a positive relationship between total Polychaete biomass and DD10 at low AMO values but a negative relationship at high AMO values (**Figure 30B**). These relationships also can be visualized as 3-dimensional response surfaces (**Figure 31A–C**).

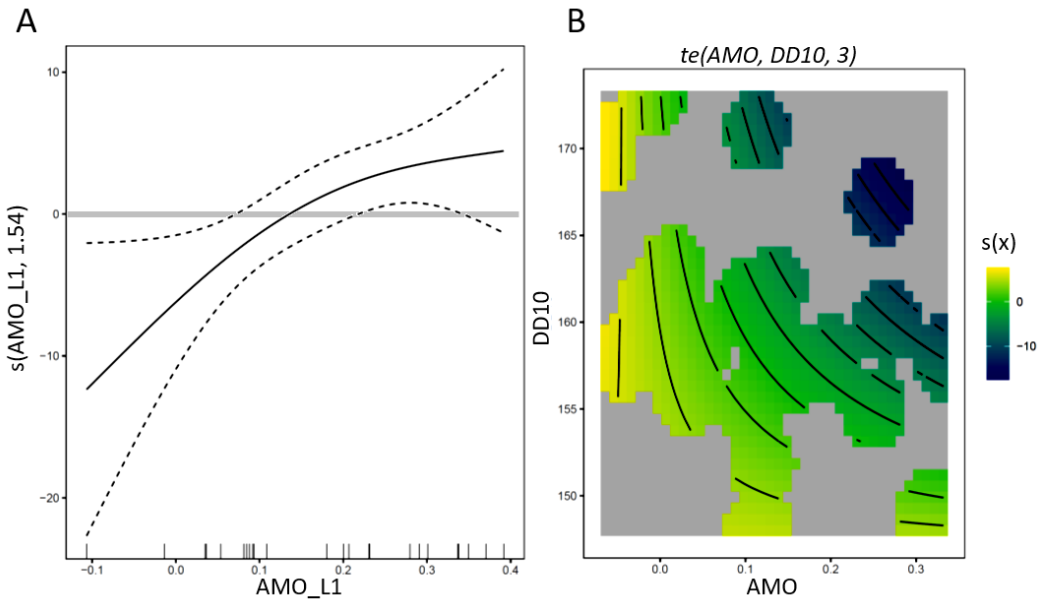

Figure 30. *Partial predictor of total Polychaete indicator as a function of AMO_L1 (A) and a contoured heat map showing the relationship of the total Polychaete indicator to the tensor interaction between DD10 and AMO (B).*

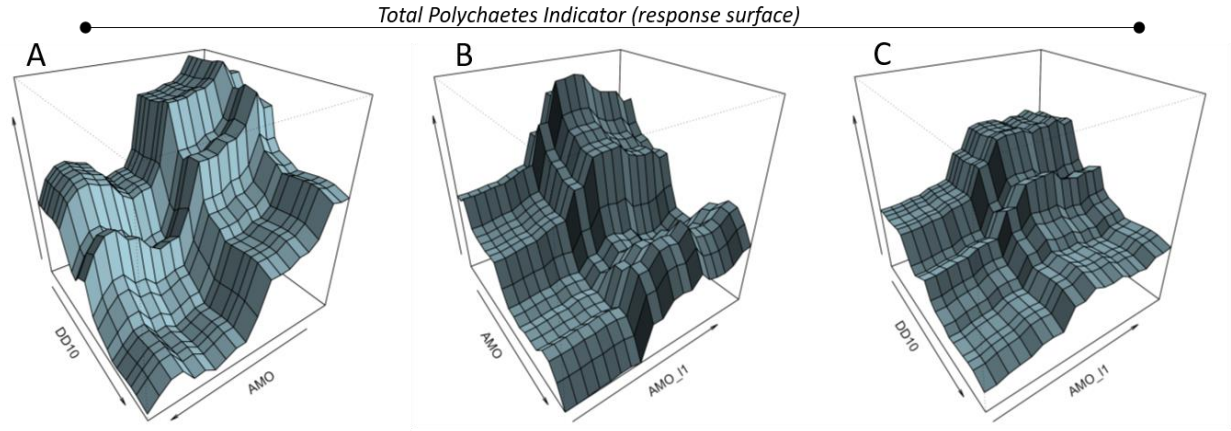

Figure 31. *Response surfaces for Total Polychaete indicator climate model based on the Random Forest model variant, showing relationships among total Polychaete biomass and: DD10 & AMO (A), AMO & AMO_L1 (B), DD10 & AMO_L1 (C).*

Bay Anchovy

Model-fitting results for climate factors indicated that the RF and GLM models performed the best among the various modeling approaches for both the TIES/ChesFIMS trawl BA indicators and the MD/VA seine BA indicators (**Table 6, Table 7)**. All TIES/ChesFIMS indicator models had negative out-of-sample R^2 values, including the best performing models. The MD/VA seine-survey indicator models performed better, with R^2 values of 0.23 and 0.25 for $BA_{Recruits}$ and BA_{Total}, respectively (**Table 7**). The best performing DD variant differed among BA indicators, surveys, and models with the DD5 or eDD5 variable showing the best relationships with four of the five best performing models. The DD10 variable was most strongly related to the TIES/ChesFIMS BA_{Total} indicator.

In the TIES/ChesFIMS models, the DD5 variable was negatively associated with the $BA_{Spawning}$ and $BA_{Recruits}$ indicators (**Figure 32A&C, Figure 32D&F**). However, the relationship between DD5 and the $BA_{Spawning}$ indicator (**Figure 32A&C**) was not consistent, but suggested a modal or very weak relationship. Conversely, the $BA_{Recruits}$ indicator was negatively associated with DD5 across the range of modeled AMO and AMO_L1 values (**Figure 32A–C**). The BA_{Total} response surface showed a modal or threshold relationship with DD10, with high DD10 values associated with the lowest BA_{Total} values and low-to-intermediate DD10 values associated higher BA_{Total} values (**Figure 32G&I**). Relationships between AMO and AMO_L1 variables were unique to each BA indicator. For Bay Anchovy spawning stock, both AMO and AMO_L1 were positively associated with the $BA_{Spawning}$ indicator (**Figure 32A-C**). Among recruits, the $BA_{Recruits}$ indicator had a negative relationship with AMO_L1 but showed a curvilinear relationship with AMO (concave down; **Figure 32D-F**). The BA_{Total} response surface showed a variable, multi-modal relationship with AMO and negative relationship with AMO_L1 (**Figure 32G-I**).

In the MD/VA seine dataset, both the $BA_{Spawning}$ and the BA_{Total} indicators were strongly negatively associated with the eDD5 variable across the range of AMO values in both MD and VA systems (**Figure 33A, C, D, F**). Both BA indicators were also negatively associated with the AMO index (calculated January–March), with response surfaces showing evidence of a threshold response at intermediate AMO values (**Figure 33B&E**). Only small differences were observed in the modeled response-surface relationships between the $BA_{Spawning}$ and BA_{Total} indicators and the climate variables.

Table 6. *Cross-validation (out-of-sample) model performance metrics for Spawning Stock ($BA_{Spawning}$), Recruitment ($BA_{Recruits}$), and Total (BA_{Total}) Bay Anchovy TIES/ChesFIMS Baywide indicators from the climate predictors models. Results associated with each DD indicator variant are identified by the column header suffix (DD5, DD10, eDD5, eDD10). Model error metrics include MAE (mean absolute error) and RMSE (root mean square error); R^2 is the coefficient of determination. Yellow highlighted cells indicate the best-performing model for each indicator.*

$BA_{Spawning}$

	MAE_DD5	RMSE_DD5	R2_DD5	MAE_DD10	RMSE_DD10	R2_DD10
GLM	115.901	132.983	-0.622	111.999	129.740	-0.544
GAM	127.680	136.041	-0.698	126.472	149.353	-1.046
RF	105.071	116.131	-0.237	107.796	118.059	-0.279

	MAE_eDD5	RMSE_eDD5	R2_eDD5	MAE_eDD10	RMSE_eDD10	R2_eDD10
GLM	112.482	129.211	-0.531	109.766	125.839	-0.453
GAM	109.609	118.934	-0.298	118.182	140.057	-0.799
RF	107.907	118.592	-0.290	110.060	119.266	-0.305

$BA_{Recruits}$

	MAE_DD5	RMSE_DD5	R2_DD5	MAE_DD10	RMSE_DD10	R2_DD10
GLM	364.070	461.139	-0.245	360.160	487.296	-0.391
GAM	444.045	496.856	-0.446	563.948	697.964	-1.853
RF	431.048	492.398	-0.420	393.405	465.709	-0.270

	MAE_eDD5	RMSE_eDD5	R2_eDD5	MAE_eDD10	RMSE_eDD10	R2_eDD10
GLM	367.039	466.356	-0.274	369.264	500.430	-0.467
GAM	486.000	541.224	-0.715	580.855	733.580	-2.151
RF	451.659	504.041	-0.488	414.509	482.217	-0.362

BA_{Total}

	MAE_DD5	RMSE_DD5	R2_DD5	MAE_DD10	RMSE_DD10	R2_DD10
GLM	786.353	1016.508	-0.800	861.958	1010.735	-0.780
GAM	1135.860	1469.408	-2.762	1204.271	1364.753	-2.245
RF	710.221	864.605	-0.302	654.492	849.528	-0.257

	MAE_eDD5	RMSE_eDD5	R2_eDD5	MAE_eDD10	RMSE_eDD10	R2_eDD10
GLM	755.038	967.186	-0.630	855.379	1013.981	-0.791
GAM	1100.572	1361.160	-2.228	1120.757	1264.179	-1.785
RF	697.356	856.111	-0.277	659.914	859.657	-0.288

Table 7. *Cross-validation (out-of-sample) model performance metrics for Recruitment ($BA_{Recruits}$) and Total (BA_{Total}) Bay Anchovy combined MD/VA seine survey indicators with climate predictors models. Results associated with each DD indicator variant are identified by the column header suffix (DD5, DD10, eDD5, eDD10). Model error metrics include MAE (mean absolute error) and RMSE (root mean square error); R^2 is the coefficient of determination. Yellow highlighted cells indicate the best-performing model for each indicator.*

$BA_{Recruits}$

	MAE_DD5	RMSE_DD5	R2_DD5	MAE_DD10	RMSE_DD10	R2_DD10
GLM	0.539	0.719	0.226	0.575	0.760	0.135
GAM	0.634	0.987	-0.458	0.637	0.887	-0.178
RF	0.514	0.738	0.185	0.569	0.790	0.067
	MAE_eDD5	RMSE_eDD5	R2_eDD5	MAE_eDD10	RMSE_eDD10	R2_eDD10
GLM	0.531	0.710	0.246	0.559	0.713	0.240
GAM	0.614	0.907	-0.231	0.558	0.781	0.087
RF	0.533	0.753	0.152	0.530	0.725	0.215

BA_{Total}

	MAE_DD5	RMSE_DD5	R2_DD5	MAE_DD10	RMSE_DD10	R2_DD10
GLM	0.826	1.138	0.217	0.906	1.207	0.120
GAM	1.009	1.667	-0.679	0.950	1.401	-0.185
RF	0.787	1.174	0.167	0.867	1.244	0.065
	MAE_eDD5	RMSE_eDD5	R2_eDD5	MAE_eDD10	RMSE_eDD10	R2_eDD10
GLM	0.821	1.112	0.253	0.879	1.117	0.247
GAM	0.924	1.466	-0.298	0.869	1.231	0.085
RF	0.814	1.187	0.150	0.810	1.124	0.236

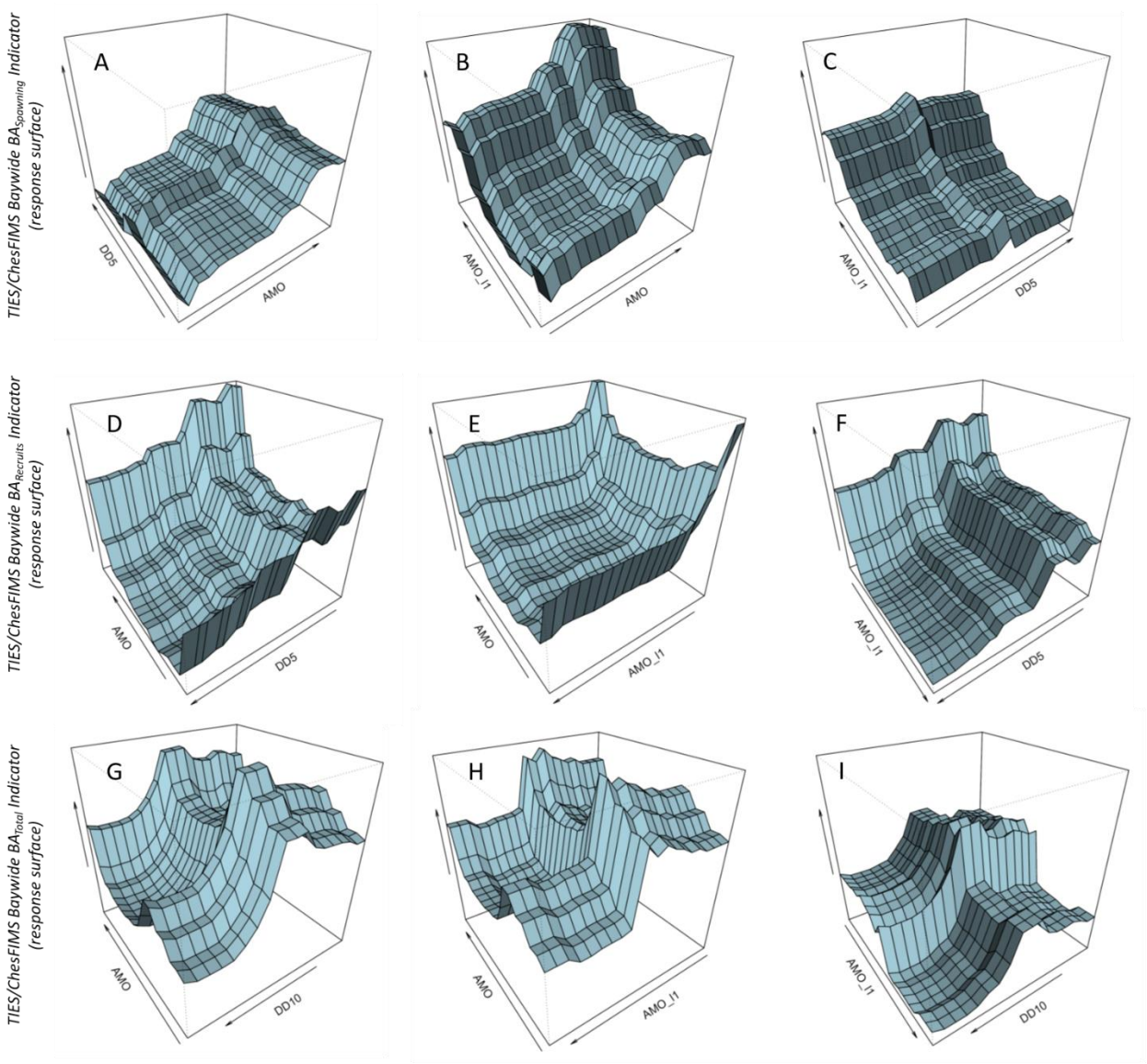

Figure 32. *Bay Anchovy Spawning Stock ($BA_{Spawning}$, A-C), Recruits ($BA_{Recruits}$, D-F), and Total (BA_{Total}) population TIES/ChesFIMS Baywide indicator climate model response surfaces from the Random Forest model variants ($BA_{Spawning}$, BA_{Total}) and GLM model variant ($BA_{Recruits}$). Surfaces show relationships among Bay Anchovy indicators and best-fitting AMO, AMO_L1, and DD predictors for each model.*

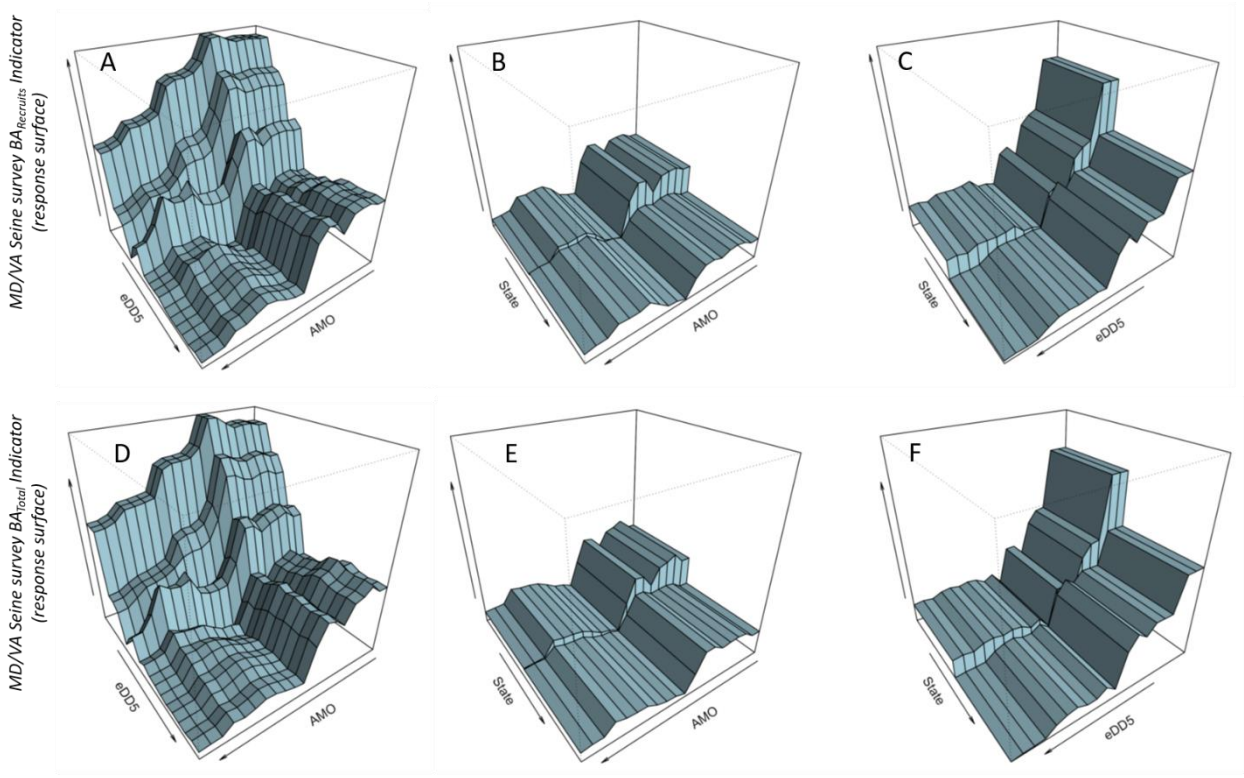

Figure 33. *Bay Anchovy Recruits ($BA_{Recruits}$, A–C) and Total (BA_{Total}, D–F) population Maryland and Virginia seine survey baywide indicator climate model response surfaces from the Random Forest model variants ($BA_{Recruits}$, BA_{Total}). Surfaces show relationships among Bay Anchovy indicators and best-fitting AMO, State, and DD predictors for each model.*

Discussion

Polychaetes

We were successful in calculating unique indices of biomass for the total Polychaete and Nereididae taxonomic groups. The two Polychaete indices showed different spatial and temporal patterns, as well as different relationships with the climate indicators. Spatial rank order of the Nereididae Polychaete index values was relatively stable over time, with certain regional strata consistently showing the highest index values (e.g., UPB, MWT, PXR), and others showing the lowest index values (e.g., VBY, PMR) **(Table 8**). The Total Polychaete index values for strata were less consistently ordered through time, with high and low index values observed for each stratum, depending on year. Despite this variability, the lower mainstem of Chesapeake Bay almost always showed the highest Total Polychaete index values across years (**Table 8**). The weaker correlation among strata in the Total Polychaete index than in the Nereididae index is likely due to the wide-range of environmental tolerances represented by all Polychaete taxa (relative to those that belong strictly to the Nereididae family) and their corresponding independent responses to annual conditions in each stratum.

An examination of polychaete-climate response surfaces indicates that the highest total Polychaete index values occur when DD10 is low, AMO is negative, and AMO_L1 is positive (**Figure 31**). Under these conditions, a 'good year' for Polychaetes would be preceded by a relatively warm and dry autumn, and a late winter-early spring in which Chesapeake Bay waters warm relatively rapidly despite experiencing relatively cool air temperatures and high precipitation. This interaction suggests a nuanced relationship between Polychaete biomass and climate conditions that depend on the timing of particular climate conditions, including water temperature and, potentially, freshwater inputs and salinity. Many Chesapeake Bay polychaete species have annual life cycles, with reproduction typically occurring during the spring, summer and fall months (Simpson 1962, Loi and Wilson 1979, Thompson and Schaffner 2001, Lippson and Lippson 2006). The number and type of larval stage(s) differ among polychaete species but many have pelagic, dispersive stages that settle to benthic habitats following a period of days to weeks. It is possible that climate conditions favoring reproductive success and juvenile survival (i.e., recruitment) could differ from climate conditions associated with high sub-adult and adult survival during the spring and early summer.

Table 8. *Mean and standard deviation of Total and Nereididae Polychaete indicators by sampling stratum: MET – Maryland Eastern Tributaries, MMS – Maryland Mainstem, MWT – Maryland Western Tributaries, PMR – Potomac River, PXR – Patuxent River, UPB – Upper Mainstem, JAM - James River, RAP – Rappahannock River, VBY – Virginia (Lower) Mainstem, YRK – York River.*

	Total Polychaete index (biomass / m^2)		**Nereididae Polychaete index (biomass / m^2)**	
Stratum	**Mean**	**Std.Dev.**	**Mean**	**Std.Dev.**
JAM	22.5	3.8	3.1	1.3
MET	20.0	6.0	4.3	1.4
MMS	20.2	6.0	5.1	1.7
MWT	17.2	4.9	5.8	1.9
PMR	10.9	3.9	2.7	1.0
PXR	14.3	2.9	5.8	1.8
RAP	17.4	2.3	4.8	1.7
UPB	22.3	7.3	6.5	2.7
VBY	55.0	11.1	2.1	2.2
YRK	28.0	2.6	4.1	0.9

The Nereididae Polychaete indicator climate models were less informative and response surfaces were difficult to interpret, suggesting different variables or local conditions could be more important factors affecting nereid biomass in Chesapeake Bay. Specifically, including sediment characteristics, river discharge, water temperature, dissolved oxygen conditions, salinity structure, or primary productivity metrics could provide additional explanatory power in future modeling efforts (Loi and Wilson 1979, Schaffner et al. 2001, Sturdivant et al. 2014, Woodland and Testa 2020). Nereids composed roughly 10-20% of the estimated Total Polychaete biomass

in the Bay and, while this is a substantial fraction of the biomass, many other polychaete taxa are present. Building on the present study by analyzing the fraction of Total Polychaete biomass that is not composed of nereid biomass (i.e., Total – Nereididae) could yield interesting results. At the family level, other polychaetes that are commonly consumed by predatory fish in Chesapeake Bay include Terebellidae, Glyceridae, Pectinariidae, and Capitellidae (Buchheister and Latour 2015). Additional research into the dynamics of these taxa and their relationships with the DD, AMO, and other climate variables could provide further information on how future climate conditions are likely to influence polychaete biomass in Chesapeake Bay. For example, capitellids could be particularly important to analyze because of their high tolerance of hypoxia and ability to rapidly recolonize benthic habitats following disturbance, which suggests they might serve as a benthic prey source when other benthos taxa are unavailable or scarce (Long and Seitz 2008).

Bay Anchovy

By aggregating data for specific months, we developed up to three separate annual indicators for Bay Anchovy from each available survey: spawning stock, recruitment, and total abundance. Each indicator is representative of a unique (or aggregate) life stage for Bay Anchovy in Chesapeake Bay. Among the primary datasets available in this study, the historical TIES/ChesFIMS survey provided the shortest time series (13 yrs) but was the only survey with sufficient data to estimate both spawning stock and recruitment, in addition to total abundance. The ongoing MD and VA seine surveys provided longer time series (~32 yrs) and supported estimation of separate recruitment and total abundance indicators but did not have sufficient spring-months sampling to support a unique estimate of spawning stock. The seine meshes (6.4 mm) in the VA and MD surveys allow escapement of a large fraction of subadult and recruiting anchovy but, nevertheless, these surveys are valuable assets for indicator development because of their long tenure and consistent survey methods. While not directly available for this study, a complementary trawl dataset for the VA portion of the Chesapeake Bay associated with the VIMS Juvenile Finfish and Blue Crab survey represents a potentially valuable source of additional Bay Anchovy data, as demonstrated by our minimal analysis of spawning stock and recruitment indicators derived from data provided to us by VIMS colleagues.

After a considerable effort to develop a useable size-based approach for separating spawning stock from recruits, we ultimately decided to separate life stages of Bay Anchovy primarily based on seasonal collections in the available surveys. This decision stemmed from issues of data availability and aspects of the life history of Bay Anchovy. Using seasonal collection data to separate spawning stock from recruiting individuals is a reasonable approach for a near-annual species (Newberger and Houde 1995) in which most individuals surveyed in spring months (April-May) are mature (>50 mm length) (Houde and Zastrow 1992; Zastrow et al. 1993) and nearly all individuals collected in late summer-early fall (August-October) are young-of-the-year recruits.

The lack of size data from the seine surveys, particularly the MD seine survey, was a major limitation in our ability to use size structure to statistically allocate catch proportions to life history stages. Without comparable length data from the MD seine survey, we were not able to

effectively use size structure data from the VA seine survey. Even when robust size distribution data were available, such as for the TIES/ChesFIMS dataset, the lack of a clear separation of length modes between spawning stock individuals and new recruits during mid-summer months made length-based cohort assignments uncertain. Despite the uncertainty, a comparison of the season-based TIES/ChesFIMS $BA_{Spawning}$ and $BA_{Recruits}$ indicators with the size-based VIMS Spawning Stock and Recruitment indicators suggested both methods yield similar relative results (**Figure 34A&C**). For example, both the TIES/ChesFIMS and the VIMS trawl surveys indicated high recruitment in 1998 and relatively high recruitment in 2000. Both surveys also indicated time-series high spawning stock abundance in 2004, as well as localized peaks in spawning stock in 1997. The similarities in patterns between the two surveys, particularly with regards to years with very high or very low indicator values, suggested that season-based and size-based approaches for separating Bay Anchovy life stages can yield similar results when applied to catches from similar gears and habitats.

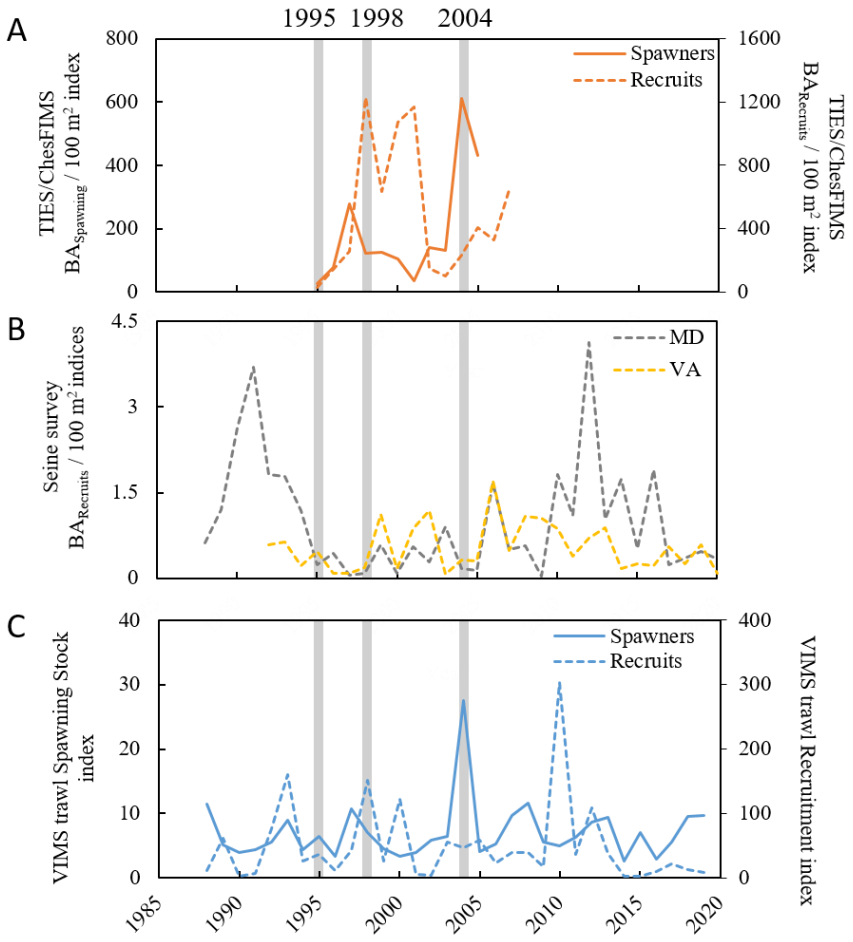

Figure 34. *Bay Anchovy Spawning Stock ($BA_{Spawning}$ [Spawners]) and Recruitment ($BA_{Recruits}$ [Recruits]) indices from the TIES/ChesFIMS survey Lower Chesapeake Bay only (A), $BA_{Recruits}$ indices from the Maryland (MD) and Virginia seine surveys (B), and the VIMS trawl Spawner and Recruits indices from Lower Chesapeake Bay. Gray vertical bars indicate years of maximum and minimum Spawning stock and Recruitment indicator values in the TIES/ChesFIMS survey.*

It is interesting to note that comparing the TIES/ChesFIMS and VIMS trawl-based recruitment indicators with the MD/VA seine-based recruitment indicators suggested an inverse relationship (**Figure 34**). Years with indicators of high recruitment shared between the two trawl time series, such as 1998, 2000 and 2005, were associated with low values for seine-based recruitment indicators. Conversely, years in which the trawl indicators suggested a decline in recruitment from the previous year were years in which one or both of the seine surveys indicated increased recruitment (e.g., 1999, 2003, 2006). The offset patterns between the mainstem trawl survey indicators and the tributary-based seine indicators could result from year-specific differences in recruitment phenology and how recruitment in the mainstem bay and tributary habitats align with the design of trawl and seine surveys. Alternatively, the pattern could arise from spatial differences in Bay Anchovy spawning stock distribution (Jung and Houde 2004b) and (or) recruitment dynamics across years, presumably due to physical, environmental, or biological factors. Our inability to develop spawning stock indicators from the seine surveys makes it difficult to examine these processes in more detail but year-to-year differences in the rank-order of $BA_{Spawning}$ and $BA_{Recruits}$ indicators across regions in the TIES/ChesFIMS trawl-survey time series supports the proposal that centers of distribution in the Bay can shift annually for both life stages. Furthermore, recruitment of Bay Anchovy is a dynamic process in which spawning stock abundance varies across regions in response to environmental variability (Jung and Houde 2004b) and in which a typical up-bay migration of pre-recruits occurs that is modified by environmental variability (Kimura et al. 2000). These factors could explain interannual differences in regional abundances.

Modeling results suggested that Bay Anchovy recruitment and total population indicators were related to both AMO and DD variables. Relationships were consistent across surveys for both indicators (TIES/ChesFIMS, MD/VA seine), supporting results of the analysis. For these two life stage indicators, Bay Anchovy relative abundance was negatively associated with AMO values for the current year. The total population and recruitment indicators displayed negative relationships with DD indicators. These patterns are the same qualitative patterns that were observed for the Total Polychaetes – years in which waters warm relatively rapidly in the late winter but are associated with cooler air temperatures and higher precipitation are associated with high Bay Anchovy recruitment and total abundance later in the summer. A previous analysis investigating the relationship between DD5 and forage taxa in Chesapeake Bay showed that Bay Anchovy abundance was weakly, negatively associated with vernal water warming rate in the Upper Chesapeake Bay and tributaries (Woodland et al. (2021). Differences between that study and results in the present study are substantial. The DD and Bay Anchovy indicators were calculated differently and climate-environmental models were structured and fitted differently in the two studies. These differences likely underlie the dissimilar results. Specifically, Woodland et al. (2021) included a range of other potential covariates in their models, including springtime (chlorophyll-a concentration, river discharge volume) and summer bottom water conditions (dissolved oxygen concentration, salinity, water temperature), and treated the AMO as a single annual index rather than adopting the integrated, moving monthly aggregate approach we applied in this study.

In the case of the TIES/ChesFIMS $BA_{Spawning}$ indicator, there was evidence of a positive association with AMO_L1 index values representing late summer to winter of the previous year. That period corresponds to the months when young-of-the-year recruitment and growth are determined for individuals indexed by the $BA_{Spawning}$ indicator the following year. It is noteworthy that drier and warmer conditions during the previous year and following recruitment of Bay Anchovy were also positively associated with the Total Polychaete indicator. These findings point to higher secondary production among some Chesapeake Bay forage taxa in years following years with positive AMO monthly values during the late summer and autumn.

Application of forage and climate indicators

The indicators calculated in this study provide an opportunity for resource managers to begin tracking forage and forage-related climate conditions in Chesapeake Bay. Using the modeled estimates provided here, a range of useful tools becomes available for qualitative (e.g., graphical), semi-quantitative (e.g., tabulated), and quantitative (e.g., statistical approaches [correlations, regression]) assessment of forage and climate conditions. For example, an aggregated 'Forage Tracker' index for qualitative assessments could be calculated with the following steps:

a. assign ordinal values of 1, 2, or 3 each year to a given forage indicator, depending on whether the modeled estimate was associated with the Low, Medium, or High tercile for that indicator,

b. calculate the average of the annual ordinal values of two or more forage indicators of interest, yielding a Forage Tracker index that ranges from 1 (all forage indicators fall in the Low tercile for a given year) to 3 (all forage indicators fall in the High tercile for a given year), and

c. plot the Forage Tracker index over time to provide a heuristic tool for understanding forage availability in Chesapeake Bay.

An equivalent approach could be taken to develop an aggregate 'Climate Tracker' for forage in Chesapeake Bay. In this case, modeled relationships from the Forage-Climate models could be used to assign ordinal values of 1, 2 or 3 to tercile values of annual climate indicators (AMO, AMO_L1, DD), depending on whether or not the Forage-Climate models indicated a positive or negative relationship between response and predictor variables. Taking this approach, time series of climate conditions associated with Good, Intermediate or Poor forage indicator values can be developed and used for qualitative assessment or evaluation purposes. These forage and climate 'Trackers' are provided as examples of how both the biotic and climate indicators developed here can be adapted to provide useful tools for Chesapeake Bay managers. Further applications of the specific indicator time series of developed as part of this project could include more sophisticated analytical approaches such as multivariate analyses that identify common trends (autocorrelation) shared among multiple time series (e.g., Dynamic Factor Analysis; Zuur et al. 2003). Further applications of more derived aggregate forage and climate indices, such as the Tracker examples provided here, could also include using those aggregate indices as response variables in statistical models.

Conclusions

We developed a suite of biological and climate indicators that have potential for forage-based management applications, including tracking of annual status of key forage taxa in ecosystem-based approaches for Chesapeake Bay. A robust understanding of trends and variability in these indicators will certainly require additional evaluation of physical, environmental, and biological factors that drive forage abundance. Accordingly, the forage indicator-climate modeling conducted here should be considered a starting place to inform future research. Future research should focus on investigating additional forage taxa for indicator development, testing for autocorrelation among diverse forage indicators, identifying mechanisms underlying forage-climate relationships, and determining the value of incorporating alternative or novel data sources during forage indicator development, e.g., genetic/genomic data, remote sensing, sonar-based detection approaches.

Project Reporting: Activities in Support of Deliverables

Reporting Period 1

Quality Assurance Project Plan: A signed, fully approved version of the QAPP was obtained May 14, 2021. The QAPP documentation was uploaded to the CBT grants portal along with Project Report 1.

Analytical Framework presentation: A presentation of the proposed analytical framework, including the data, variables, models, and spatial/temporal scales that were to be used to assess the effects of environmental conditions on forage populations was provided to the Sustainable Fisheries GIT (SFGIT; B. Vogt) and Forage Action Team project leads (M. Bromillow, J. Shapiro). The presentation occurred at the SFGIT Spring 2021 meeting.

Compiled project data (XLSX): A compiled Excel workbook that included all data downloaded or obtained from data managers for this project. Data were provided in 'raw' format with minimal processing to maintain the original data structures. Metadata, including source information and URL (where available), were provided for each dataset.

Project Report 1: A written summary of efforts and progress made during the first reporting interval.

Reporting Period 2

Interim modeling/analysis report: Detailed documents that include R-code with in-depth descriptions of the purpose, function, and results of blocks of code. All code used up to that date were included in the report, spanning data importing, manipulation, and analysis. All model variants under consideration at the time were included in the report, as well as preferred model variants (as determined by model performance metrics detailed in the associated report). Code-based reports were provided as appendices and included: *Polychaete indicators.pdf*, *Anchovy length threshold appendix.pdf*.

Analytical Progress presentation: A verbal update on analytical progress made up to that date, including the models, model evaluation, and ongoing efforts that were to be used to assess the effects of environmental conditions on forage populations, was provided to the Sustainable Fisheries GIT Forage Action Team (FAT) during the Fall 2021 biannual meeting. Attendees included the SFGIT (B. Vogt) and FAT project leads (M. Bromilow, J. Shapiro) and general membership of the FAT.

Project Report 2: A written summary of efforts and progress made during the second reporting interval.

Reporting Period 3

Interim modeling/analysis report: Detailed document that included R-code with in-depth descriptions of the purpose, function, and results of blocks of code. All code used up to that date were included in this report or available through previous reports, spanning data importing, manipulation, and analysis. All model results were included in the report, as well as best-

performing model structures (as determined by model performance metrics detailed in this document). A code-based report was provided as an appendix: *Polychaete index appendix.pdf*.

Analytical Progress update: A verbal update on analytical progress made to date, including the models, model evaluation, and ongoing efforts to assess the effects of environmental conditions on forage populations, was provided in two separate presentations. These were 1) a Project Advisory Team Meeting in December 2021 and 2) a Sustainable Fisheries Goal Implementation Team (SFGIT) meeting in January 2022. Attendees at these meetings included the SFGIT (B. Vogt) and FAT project leads (M. Bromilow, J. Shapiro), regional scientific stakeholders (Julie Reichert-Nguyen, NOAA; Henry Legett, SERC), and general membership of the SFGIT.

Project Report 3: A written summary of efforts and progress made during the third reporting interval.

Reporting Period 4

Interim modeling/analysis report: Detailed documents that include R-code with in-depth descriptions of the purpose, function, and results of blocks of code. All code used to date are included in this report or available through previous reports, spanning data importing, manipulation, and analysis. All model results are included in the report, as well as best-performing model structures (as determined by model performance metrics detailed in this document). Three final, code-based draft reports are provided as appendices: *Appendix 1. Polychaete indicators.pdf, Appendix 2. Bay Anchovy indicators.pdf, Appendix 3. Forage-climate modeling.pdf*.

Final Report (DRAFT; Project Report 4): A written summary of methods, findings and conclusions for the project.

REFERENCES

Able, K. W., and M. P. Fahay. 1998. The first year of life of estuarine fishes in the Middle Atlantic Bight. Rutgers University Press, New Brunswick.

Baird, D., and R. E. Ulanowicz. 1989. The seasonal dynamics of the Chesapeake Bay ecosystem. Ecological Monographs **59**:329-364.

Bhattacharya, C. 1967. A simple method of resolution of a distribution into Gaussian components. Biometrics:115-135.

Bishop, C. M. 2006. Pattern recognition and machine learning. Springer, New York.

Buchheister, A., C. F. Bonzek, J. Gartland, and R. J. Latour. 2013. Patterns and drivers of the demersal fish community of Chesapeake Bay. Marine Ecology Progress Series **481**:161-180.

Buchheister, A., and E. D. Houde. 2016. Forage indicators and nutrional profiles for Chesapeake Bay fishes. Final Report to Chesapeake Bay Trust, Chesapeake Biological Laboratory, Final Report to Chesapeake Bay Trust, University of Maryland Center for Environmental Science.

Buchheister, A., and R. J. Latour. 2015. Diets and trophic-guild structure of a diverse fish assemblage in Chesapeake Bay, U.S.A. Journal of Fish Biology **86**:967-992.

Bunnell, D. B., and T. J. Miller. 2005. An individual-based modeling approach to spawning-potential per-recruit models: an application to blue crab (Callinectes sapidus) in Chesapeake Bay. Canadian Journal of Fisheries and Aquatic Sciences **62**:2560-2572.

Chesapeake Bay Program (CBP). 2020. Logic and Action Plan: Post-Quarterly Progress Meeting, Forage Fish Outcome – 2020-2021. Biennial Strategy Review System, Chesapeake Bay Program, March 2020.

Colvocoresses, J. A., and P. J. Geer. 1991. Estimation of relative juvenile abundance of recreationally important finfish in the Virginia portion of Chesapeake Bay. Virginia Institute of Marine Science, Gloucester Point, VA.

Dauer, D. M. 2011. Quality assurance/quality control plan: Benthic Monitoring Program of the lower Chesapeake Bay. Old Dominion University, Norfolk, VA.

Ding, H. Y., and A. J. Elmore. 2015. Spatio-temporal patterns in water surface temperature from Landsat time series data in the Chesapeake Bay, USA. Remote Sensing of Environment **168**:335-348.

Hartman, K. J., and S. B. Brandt. 1995. Trophic resource partitioning, diets, and growth of sympatric estuarine predators. Transactions of the American Fisheries Society **124**:520-537.

Hastie, T., R. Tibshirani, and J. H. Friedman. 2009. The elements of statistical learning : data mining, inference, and prediction. 2nd edition. Springer, New York, NY.

Humphrey, J., M. J. Wilberg, E. D. Houde, and M. C. Fabrizio. 2014. Effects of temperature on age-0 Atlantic menhaden growth in Chesapeake Bay. Transactions of the American Fisheries Society **143**:1255-1265.

Ihde, T. F., E. D. Houde, C. F. Bonzek, and E. Franke. 2015. Assessing the Chesapeake Bay forage base: Existing data and research priorities. Chesapeake Bay Program Scientific and Technical Advisory Committee, Edgewater, MD.

Jung, S., and E. D. Houde. 2003. Spatial and temporal variabilities of pelagic fish community structure and distribution in Chesapeake Bay, USA. Estuarine, Coastal and Shelf Science **58**:335-351.

Jung, S., and E. D. Houde. 2004a. Production of bay anchovy Anchoa mitchilli in Chesapeake Bay: application of size-based theory. Marine Ecology Progress Series **281**:217-232.

Jung, S., and E. D. Houde. 2004b. Recruitment and spawning-stock biomass distribution of bay anchovy (Anchoa mitchilli) in Chesapeake Bay. Fishery Bulletin **102**:63-77.

Jung, S., and E. D. Houde. 2005. Fish biomass size spectra in Chesapeake Bay. Estuaries **28**:226-240.

Kimura, R., D. H. Secor, E. D. Houde, and P. M. Piccoli. 2000. Up-estuary dispersal of young-of-the-year bay anchovy Anchoa mitchilli in the Chesapeake Bay: inferences from microprobe analysis of strontium in otoliths. Marine Ecology Progress Series **208**:217-227.

Lippson, A. J., and R. L. Lippson. 2006. Life in the Chesapeake Bay. 3rd edition. Johns Hopkins University Press, Baltimore.

Little, D. R., R. Oehmen, J. Dunn, K. Hird, and K. Kirsner. 2013. Fluency Profiling System: An automated system for analyzing the temporal properties of speech. Behavior research methods **45**:191-202.

Loi, T. n., and B. J. Wilson. 1979. Macroinfaunal structure and effects of thermal discharges in a mesohaline habitat of Chesapeake Bay, near a Nuclear Power Plant. Marine Biology **55**:3-16.

Long, W. C., and R. D. Seitz. 2008. Trophic interactions under stress: hypoxia enhances foraging in an estuarine food web. Marine Ecology Progress Series **362**:59-68.

Luo, J. G., and J. A. Musick. 1991. Reproductive biology of the bay anchovy in Chesapeake Bay. Transactions of the American Fisheries Society **120**:701-710.

Lynch, P. D., K. W. Shertzer, and R. J. Latour. 2012. Performance of methods used to estimate indices of abundance for highly migratory species. Fisheries Research **125**:27-39.

Manning, C., P. Raghavan, and H. Schütze. 2008. Introduction to information retrieval. Cambridge University Press, Cambridge University Press.

Neuheimer, A. B., and C. T. Taggart. 2007. The growing degree-day and fish size-at-age: the overlooked metric. Canadian Journal of Fisheries and Aquatic Sciences **64**:375-385.

Newberger, T., and E. Houde. 1995. Population biology of bay anchovy *Anchoa mitchilli* in the mid Chesapeake Bay. Marine Ecology Progress Series:25-37.

Nye, J. A., M. R. Baker, R. Bell, A. Kenny, K. H. Kilbourne, K. D. Friedland, E. Martino, M. M. Stachura, K. S. Van Houtan, and R. Wood. 2014. Ecosystem effects of the Atlantic Multidecadal Oscillation. Journal of Marine Systems **133**:103-116.

Schaffner, L. C., T. M. Dellapenna, E. K. Hinchey, C. T. Friedrichs, M. T. Neubauer, M. E. Smith, and S. A. Kuehl. 2001. Physical energy regimes, seabed dynamics and organism-sediment interactions along an estuarine gradient. Organism-sediment interactions:159-179.

Seitz, R. D., D. M. Dauer, R. J. Llansó, and W. C. Long. 2009. Broad-scale effects of hypoxia on benthic community structure in Chesapeake Bay, USA. Journal of Experimental Marine Biology and Ecology **381**:S4-S12.

Simpson, M. 1962. Reproduction of the polychaete *Glycera dibranchiata* at Solomons, Maryland. The Biological Bulletin **123**:396-411.

Sturdivant, S. K., R. J. Diaz, R. Llanso, and D. M. Dauer. 2014. Relationship between hypoxia and macrobenthic production in Chesapeake Bay. Estuaries and Coasts **37**:1219-1232.

Testa, J. M., J. Faganeli, M. Giani, M. J. Brush, C. De Vittor, W. R. Boynton, S. Covelli, R. J. Woodland, N. Kovač, and W. M. Kemp. 2020. Advances in our understanding of pelagic–benthic coupling. Pages 147-175 *in* T. C. Malone, A. Malej, and J. Faganeli, editors. Coastal Ecosystems in Transition: A Comparative Analysis of the Northern Adriatic and Chesapeake Bay. American Geophysical Union.

Thompson, M. L., and L. C. Schaffner. 2001. Population biology and secondary production of the suspension feeding polychaete *Chaetopterus cf. variopedatus*: Implications for benthic-pelagic coupling in lower Chesapeake Bay. Limnology and Oceanography **46**:1899-1907.

Tuckey, T. D., and M. C. Fabrizio. 2021. Estimating Relative Juvenile Abundance of Ecologically Important Finfish in the Virginia Portion of Chesapeake Bay (1 July 2020 – 30 June 2021). Virginia Institute of Marine Science, William & Mary, Gloucester Point, Virginia 23062.

Vaquer-Sunyer, R., and C. M. Duarte. 2008. Thresholds of hypoxia for marine biodiversity. Proceedings of the National Academy of Sciences of the United States of America **105**:15452-15457.

Versar, I. 2011. Chesapeake Bay Water Quality Monitoring Program long-term benthic monitoring and assessment component quality assurance project plan 2011-2012. Versar, Inc, Columbia, MD.

Ward, J. V., and J. A. Stanford. 1982. Thermal responses in the evolutionary ecology of aquatic insects. Annual Review of Entomology **27**:97-117.

Wood, R. J., and H. M. Austin. 2009. Synchronous multidecadal fish recruitment patterns in Chesapeake Bay, USA. Canadian Journal of Fisheries and Aquatic Sciences **66**:496-508.

Wood, S. N. 2017. Generalized additive models: an introduction with R. 2nd edition. Chapman and Hall/CRC, Chapman and Hall/CRC.

Woodland, R. J., A. Buchheister, R. J. Latour, C. Lozano, E. Houde, C. J. Sweetman, M. C. Fabrizio, and T. D. Tuckey. 2021. Environmental drivers of forage fishes and benthic invertebrates at multiple spatial scales in a large temperate estuary. Estuaries and Coasts **44**:921-938.

Woodland, R. J., E. D. Houde, A. Buchheister, R. J. Latour, C. Lozano, M. C. Fabrizio, T. Tuckey, and C. Sweetman. 2017. Environmental, spatial and temporal patterns in Chesapeake Bay forage population distributions and predator consumption. University of Maryland Center for Environmental Science Chesapeake Biological Laboratory, Solomons, MD.

Woodland, R. J., and J. M. Testa. 2020. Response of benthic biodiversity to climate-sensitive regional and local conditions in a complex estuarine system. Pages 87-122 *in* V. Lyubchich, Y. Gel, K. H. Kilbourne, T. J. Miller, N. K. Newlands, and A. B. Smith, editors. Evaluating Climate Change Impacts. CRC Press, 6000 Broken Sound Parkway NW, Suite 300, Boca Raton, FL 33487-2742.

Zastrow, C., E. Houde, and L. Morin. 1991. Spawning, fecundity, hatch-date frequency and young-of-the-year growth of bay anchovy Anchoa mitchilli in mid-Chesapeake Bay. Marine Ecology Progress Series **73**:161-171.

Zuur, A. F., R. J. Fryer, I. T. Jolliffe, R. Dekker, and J. J. Beukema. 2003. Estimating common trends in multivariate time series using dynamic factor analysis. Environmetrics **14**:665-685.

Project Timeline

Revised project reporting schedule, as approved by the Project Advisory Team during the Project Kickoff Meeting (4/19/21) and via email on (3/1/22). Highlighted rows indicate reporting deliverable for this project report.

Report # and Reporting Period	Project Deliverables	Approved revised Date of Delivery
QAPP deliverable (Reporting Period #1)	• Final (signed) QAPP in PDF format	**5/21/2021**
Report #1 (Reporting Period #1)	• Excel or Access database of all biological and environmental data and sources • Presentation and PDF of the proposed analytical framework • Progress report	**7/14/2021**
Report #2 (Reporting Period #2)	• R modeling/analysis script (code) and model outputs • Progress report	**10/14/2021**
Meeting deliverable (Reporting Period #3)	• Meet with the FAT and other CBP partners and stakeholders to discuss and coordinate indicator development options based on the results of the analyses	**10/15/2021-11/1/2021**
Report #3 (Reporting Period #3)	• R indicator script (code) and visualization outputs • Progress report	**1/14/2022**
Draft Report #4 (Reporting Period #4)	• Editable draft report, submitted to the GIT Lead and the FAT for review and feedback	**3/15/2022**
Report #4 (Reporting Period #4)	• Final report package, including editable database, the R files and PDFs of all R scripts and outputs for modeling/analysis and indicator development, and the final indicator graphics • Presentation of final project results	**4/14/2022**