# Seasonal and Annual Variation in the Extent of Suitable Habitats for Forage Fishes in Chesapeake Bay, 2000 - 2016

Mary C. Fabrizio[1], Troy D. Tuckey[1],

Aaron J. Bever[2], and Michael L. MacWilliams[2]

[1]Virginia Institute of Marine Science, Gloucester Point, VA 23062

[2]Anchor QEA, LLC, San Francisco, CA 94111

Prepared for NOAA Chesapeake Bay Office, Annapolis, MD 21403

November 2020

https://doi.org/10.25773/dyjy-mm73

VIMS | William & Mary
Virginia Institute of Marine Science

ANCHOR QEA

# Executive Summary

The sustained production of sufficient forage is critical to advancing ecosystem-based management in Chesapeake Bay. Yet factors that affect local abundances and habitat conditions necessary to support forage production remain largely unexplored. Here, we quantified suitable habitat in the Chesapeake Bay region for four key forage fishes: bay anchovy *Anchoa mitchilli*, juvenile spot *Leiostomus xanthurus*, juvenile weakfish *Cynoscion regalis*, and juvenile spotted hake *Urophycis regia*. We coupled information from 17 years of monthly fisheries surveys with hindcasts from a numerical model of dissolved oxygen (DO) conditions and a 3-D hydrodynamic model of the Bay that provided estimates of habitat conditions across 18 covariates of salinity, temperature, DO, depth, and current speed for the period 2000 to 2016. Sediment composition and distance to shore metrics were also considered. The hindcast covariates were subsampled at the times and locations of the fisheries surveys to provide dynamic habitat metrics that are not generally observed at the time of fish sampling (e.g., current velocity, salinity stratification). Hindcast covariates were also used to describe habitat conditions in areas of Chesapeake Bay that are not sampled routinely by fisheries-independent surveys such as the Potomac River and Mobjack Bay. Boosted regression trees were used to identify influential habitat covariates for each species, and these influential covariates were then used to construct habitat suitability models. Habitat suitability indices, which ranged between 0 (poor habitat) and 1 (superior habitat), were assigned to each location in the 3-D model grid for each season in 2000-2016. Based on the estimated habitat suitability index and using a GIS approach, we quantified suitable habitat (defined as habitats with a habitat suitability index $\geq$ 0.5) throughout the Chesapeake Bay and its tidal tributaries. Furthermore, we validated the modeling approach using out-of-sample observations from Mobjack Bay in 2010-2012.

Suitable seasonal habitat extents for forage species exhibited strong seasonal and annual signals reflecting temporal heterogeneity in habitat conditions in Chesapeake Bay. Current speed, water depth, and either temperature or dissolved oxygen were identified as important covariates for the four forage species we examined, and distance to shore was important for three of the four species; thus, suitable habitat conditions resulted from a complex interplay between water quality and the physical properties of the habitat. In our study, two species exhibited a relationship between relative abundance and extent of suitable habitats – juvenile spot in summer and bay anchovy in winter; as such, estimates of the minimum habitat area required to produce a desired abundance (or biomass) of forage fish can be used to establish quantitative habitat targets or spatial thresholds that may serve as spatial reference points for management. In an ecosystem-based approach, important habitats may be targeted for protection (e.g., by limiting fishing activities that may incidentally capture or injure forage fishes) or restoration (e.g., by improving water quality conditions), thereby ensuring production of sufficient forage for predators. In addition, the consequences of aquatic habitat alterations, whether due to climate change or physical disturbances can be investigated using projections of environmental conditions and habitat suitability in the region, though these projections will introduce additional uncertainty.

# Introduction

Stock assessments rarely account for trophic interactions among aquatic predators and their prey, yet such interactions are recognized as critical to advancing ecosystem-based management. Indeed, the spatial distribution and abundance of prey may drive the distribution and abundance of their predators. For example, predators such as summer flounder use estuarine and coastal waters of the eastern US to feed on abundant prey during spring, summer, and fall (Murdy et al. 1997; Latour et al. 2008; Buchheister and Latour 2011), and striped bass undertake seasonal feeding migrations along the US coast to utilize abundant prey in non-natal estuaries (Mather et al. 2009). Migratory movements and trophic interactions of predators such as summer flounder and striped contribute to the connectivity between coastal and offshore ecosystems. Furthermore, the abundance of prey may affect the overall health and condition of predators that rely on the supply of sufficient prey for maintenance and growth. For example, long-term, broad-scale patterns in the productivity of the Chesapeake Bay is reflected as changes in body condition of piscivorous fishes, suggesting bottom-up effects on fish productivity (Latour et al. 2017). Low abundance of prey species, particularly prey fish, may affect the health of predators: for example, the severity of mycobacterium disease in Chesapeake Bay striped bass is associated with low prey abundance (Jacobs et al. 2009). Forage fishes often comprise important components of the diet of top predators and serve as a link between trophic levels; furthermore, the continued production of sufficient forage fish is recognized as critical to advancing ecosystem-based management in the Bay (CBP 2015, 2018; Ihde et al. 2015). Although feeding habits of many predators are well studied, the distribution and abundance of prey species that comprise the forage base of predators have received less attention (but see Arbeider et al. 2019 and Woodland et al. *in press*). In particular, the relationship between the abundance of forage species and the extent of their suitable habitats remains largely unexplored.

Static features of the environment, such as substrate type, are often used to characterize fish habitats because such features affect fish distributions and habitat use (Day et al. 1989; Fabrizio et al. 2013). Dynamic environmental conditions such as salinity, temperature, dissolved oxygen (DO), and depth also contribute to variations in the distribution and abundance of estuarine and coastal species. In the dynamic estuarine environment, conditions vary across multiple spatial scales and annual, seasonal, daily, and tidal scales. The range of dynamic conditions is large in temperate estuaries and often dictates the phenology and spatial distribution of fishes (Buchheister et al. 2013). For example, in river-dominated estuaries, river flow affects salinity and alters the extent of suitable habitats for juvenile fishes (Kostecki et al. 2010), many of which may serve as forage for predators. For ectotherms, temperature is a key determinant of habitat suitability because critical processes such as metabolic rates, movement, and growth are governed by temperature (Little et al. 2020). Seasonal changes in DO concentrations in estuarine and coastal waters may also shape the distribution and habitat use of fishes. In particular, abundance of fish is low in hypoxic (< 2 mg $O_2$/l) regions of estuaries and nearshore waters (Craig and Crowder 2005; Zhang et al. 2009; Buchheister et al. 2013; Glaspie et al. 2019), suggesting that fishes actively avoid hypoxic habitats. Low DO conditions are believed to limit the extent of suitable habitat for fishes, particularly during summer in estuarine and coastal systems that exhibit prolonged seasonal hypoxia.

Other habitat features, such as bottom-current velocities, water column stability, and salinity stratification, may contribute to the variation in the spatial distribution and abundance of fishes and other organisms (Manderson et al. 2011; Jenkins et al. 2015; Bever et al. 2016). For example, predictions of suitable habitats for the endangered Delta smelt *Hypomesus transpacificus* in the San

Francisco Estuary were improved when current velocities were considered in addition to salinity and turbidity conditions (Bever et al. 2016). Small-bodied fishes such as Delta smelt and anchovies may select areas of low current velocities presumably to maintain position while minimizing energy expenditure in a dynamic environment (Hatin et al. 2007). Species that use chemical cues to detect predators may avoid areas with high current velocities because such currents may interfere with their ability to escape from predators (Powers and Kittinger 2002). Conversely, estuarine habitats with higher current velocities may enhance feeding of some species by increasing the delivery of planktonic prey (Parsons et al. 2015). Other species may use tidal currents to assist in horizontal movements within the estuary (e.g., tidal-stream transport), and thus, current speed may be important for elucidating patterns in habitat use. Hydrodynamic complexity may therefore represent a key determinant of suitable habitats in estuarine systems (Bever et al. 2016). Indeed, hydrodynamic models have been used to estimate habitat volume for estuarine species using information on physiological tolerances and bioenergetics requirements (e.g., Schlenger et al. 2013). Outputs from such models have also been used to assess the effect of sea-level rise on fishes that depend on marsh habitats for juvenile growth and survival (Fulford et al. 2014). A few studies have coupled hydrodynamic models with fisheries surveys to assess the extent of suitable habitats (Le Pape et al. 2003; MacWilliams et al. 2016; Bever et al. 2016), but only one study (Le Pape et al. 2003) examined the relationship between abundance of fish and changes in the extent of suitable estuarine habitat.

Forage-fish management must not only be informed by knowledge of the characteristics of suitable habitat, but also by the dynamics of the extent of habitats that support these species. The largest estuary in the US, the Chesapeake Bay, provides a useful model in which to investigate the relationship between annual changes in forage abundance and the extent of suitable coastal habitats. The health and sustainability of iconic fisheries in this system depend on sufficient production and availability of forage as well as effective management and protection from anthropogenic degradation of habitats. With the exception of a single study (Woodland et al. *in press*), habitat conditions necessary to support forage production in this system remain largely unexplored.

In this study, we considered four forage fishes that are numerically dominant in the fish community of Chesapeake Bay (Tuckey and Fabrizio 2020): bay anchovy *Anchoa mitchilli*, juvenile spot *Leiostomus xanthurus*, juvenile spotted hake *Urophycis regia*, and juvenile weakfish *Cynoscion regalis*. We focus on forage fishes rather than invertebrates because of the availability of temporally and spatially rich data for these taxa in Chesapeake Bay. Small-bodied fishes such as bay anchovy and the juvenile stages of larger species are important components of the diets of resident and transient predators in Chesapeake Bay (Buchheister and Latour 2011; Buchheister and Latour 2015). Together, the selected forage species are available year-round. Our objectives were to (1) quantify suitable habitats for forage species in Chesapeake Bay from 2000 to 2016, and (2) assess the relationship between the extent of suitable habitats and annual forage abundance. Because we selected taxonomically and ecologically disparate species, we expected that suitable habitats would be defined by habitat features that differed among species. If the extent of suitable habitats limits the production of forage fishes in Chesapeake Bay, then we would expect annual patterns in forage fish abundances to exhibit patterns similar to those for suitable habitats.

To address our objectives, we examined monthly catches of forage fishes from fisheries-independent surveys along with descriptors of habitat conditions. Because fish-habitat relationships are best derived from observations across broad spatial scales and long time periods (Gray et al. 2011; Lecours et al. 2015), we quantified these relationships for Chesapeake Bay and its subestuaries during the 17-year period, 2000-2016. We considered many covariates because we were unsure about which

ones may play a role in fish-habitat use, and because uncertainty in fish-habitat models may arise from the omission of covariates that limit habitat use (Cade et al. 2005). Furthermore, rather than considering only those habitat features measured at the time of fish sampling (i.e., bottom temperature, salinity, DO), we considered dynamic habitat conditions obtained as hindcasts from two numerical models, as well as several static features as covariates for habitat suitability modeling. Model covariates included DO, depth, and multiple salinity, temperature, and current speed covariates, as well as sediment composition and distance to shore. We applied a data-driven approach, boosted regression tree analysis (Elith et al. 2008), to select a subset of habitat covariates that were most influential in explaining fish relative abundance. Nonparametric suitability models using the histogram approach were then constructed using the selected influential covariates following the method in Tanaka and Chen (2015) and Guan et al. (2016). This nonparametric method ascribes higher suitability to conditions in which greater abundances of organisms are observed, and as such, are process-based models. Nonparametric suitability models have been applied to estimate annual changes in habitat suitability for American lobster (Tanaka and Chen 2015) and Atlantic cod (Guan et al. 2016) in the northwest Atlantic Ocean. However, habitat suitability modeling has not yet been applied to understand fish-habitat relationships in Atlantic coast estuaries, systems which likely exhibit less depth variation but greater tidal ranges and dynamic habitat conditions. Habitat suitability models for each of the four forage species were used to visualize and quantify seasonally suitable habitat throughout the Chesapeake Bay from 2000 to 2016. We examined seasonal habitat suitability because some of these species are seasonal migrants that use the Chesapeake Bay as a nursery. Finally, to assess the role of habitat area in driving forage fish abundance, we used nonparametric regressions to relate annual estimates of the extent of suitable habitat to annual baywide estimates of fish abundance.

# Methods

We developed an integrated modeling framework to couple information on the abundance of forage taxa with environmental conditions estimated from two numerical models of Chesapeake Bay. The primary data were monthly catches from fishery-independent surveys of forage fishes, hindcasts of dynamic environmental conditions (covariates describing salinity, temperature, current speed, depth, and dissolved oxygen conditions), and estimates of static habitat conditions (sediment composition and distance to shore).

## Fisheries surveys and estimation of relative abundance

Geo-referenced catches of forage fishes were obtained from two bottom-trawl surveys: the Virginia Institute of Marine Science Juvenile Fish Trawl Survey (hereafter, Virginia survey) and the Maryland DNR Blue Crab Summer Trawl Survey (hereafter, Maryland survey). The sampling domain of the Virginia survey includes waters greater than 1.2 m depth throughout Virginia tidal waters of the Chesapeake Bay and its major tributaries (James, York, and Rappahannock rivers; Figure 1A). Each month, from January to December, the Virginia survey sampled fishes from 111 stations selected from a random stratified survey design (Table 1). A 30’ semi-balloon bottom trawl was deployed for 5 minutes at each site; protocol details are available in Tuckey and Fabrizio (2016). In addition, the Virginia survey sampled Mobjack Bay, a large subregion of the Chesapeake Bay system, in summer 2010-2012 using a stratified sampling design. Fisheries observations from Mobjack Bay were considered for external validation of the habitat suitability models. The Maryland survey is primarily a shallow-water survey

(mean depth=2.1 m; all sites < 5.5 m deep) that samples fishes from fixed sites in tributaries and sounds of the Maryland portion of the Chesapeake Bay (Figure 1B). A 16' semi-balloon otter trawl is towed for 6 minutes at each site. In 2000 and 2001, sampling was conducted monthly from May through October at 37 sites in the Chester River, Choptank River, Eastern Bay, Patuxent River, Pocomoke Sound, and Tangier Sound. In 2002 and thereafter, 16 additional sites were sampled in Fishing Bay, the Little Choptank River, and the Nanticoke River (57 sites total; Table 1; Figure 1B). No sampling occurred in Maryland waters in May 2006.

***Species and seasons*** -- We considered juvenile (age-0) spot, weakfish, and spotted hake as forage for piscivorous species in Chesapeake Bay; all life stages of bay anchovy were considered forage. Because the Virginia survey encounters multiple age classes of the targeted species, we applied monthly threshold values of fork length (bay anchovy only, mm) or total length (mm) from the Virginia survey to enumerate the catch of age-0 spot, weakfish, and spotted hake (Tuckey and Fabrizio 2016); spot, weakfish, and spotted hake captured by the Maryland survey were assumed to be predominantly juvenile fish, as no length measurements were available for these species. This assumption was reasonable because of the small size of the gear and shallow depths sampled by the Maryland trawl; shallow habitats such as these are predominantly used by small-bodied fishes (e.g., Blaber and Blaber 1980; Ruiz et al. 1993).

For each trawl tow, we expressed the relative index of abundance as the catch per unit effort (CPUE), where effort was estimated by the area swept by the net. Area swept ($km^2$) was calculated as the product of the width of the net opening (3.19 m for the Virginia survey, 2.85 m for the Maryland survey) and the length of the tow (km) estimated as the geodetic distance between the GPS coordinates recorded at the beginning and end of each tow. To ensure that CPUE represented relative abundance, catches from only those seasons in which individuals were available to the gear were considered: juvenile spotted hake in spring (March, April, May); juvenile spot in summer (June, July, August) and fall (September, October, November); juvenile weakfish in summer and fall; and bay anchovy in summer, fall, and winter (December, January, February). Note that no sampling was completed in Maryland in winter and thus, the bay anchovy CPUE index in winter was based on catches from Virginia waters only.

***Baywide indices of relative abundance*** – We performed a formal data-level integration (*sensu* Fletcher et al. 2019) of the Virginia and Maryland surveys to obtain a single estimate of scaled relative abundance that reflected the abundance of each species throughout the Chesapeake Bay. A Bayesian hierarchical method (Conn 2010) was used to estimate baywide relative abundance using species-specific CPUEs (standardized to a mean of 1.0 across the 17 years) from the two trawl surveys. The Conn (2010) method extracts a single annual index to represent the pattern exhibited by the multiple indices under the assumption that component indices are subject to process error (from variation in catchability, spatial distribution, etc.) and sampling error (i.e., within-survey variance; Conn 2010). The coefficient of variation is used to weight the individual data sources (Conn 2010). Simulations using this approach indicate good performance under a number of scenarios, including violation of assumptions (Conn 2010). Annual baywide indices of relative abundance and their associated 95% credible intervals were estimated for the season of interest for each forage species. We used WinBugs accessed through an R script (R Core Team 2019) to perform these calculations. All hierarchical baywide indices were inspected graphically using SAS® software to confirm that the baywide index reasonably captured the dynamics observed by the Maryland and Virginia surveys.

## Habitat covariates

Habitat conditions throughout the Chesapeake Bay and its tributaries were obtained for 2000 through 2016; two static features and 22 dynamic habitat features (Table 2) were considered. An interpolated DO model and a hydrodynamic model (described below) were used to hindcast high-resolution estimates (spatially and temporally) of habitat conditions.

***Static habitat covariates*** - Sediment composition was expressed as the percent of fine sediment of the top layer of the seabed; this covariate provided fine-spatial scale information on a key feature of fish habitat (Kritzer et al. 2016). We also considered distance to shore (km) as a measure of the affinity of forage fishes to shorelines. Fringing marsh and other shallow water areas may provide resources that enhance survival and growth of forage fish (e.g., refuge from predators and provisioning of food; Manderson et al. 2004; França et al. 2009), and as such, this distance may influence fish habitat use. Distance to shore was calculated in a GIS using the straight-line distance between the sample site and the nearest shoreline.

***Bottom-water dissolved oxygen*** - DO concentrations (mg $O_2$/l) were predicted for bottom waters of the Chesapeake Bay and its tributaries from a numerical model (Du and Shen 2014) that was modified to include observations from monthly fisheries surveys, quarter-hourly records from Maryland data buoys (Maryland Eyes on the Bay), quarter-hourly records from the Virginia Estuarine and Coastal Observing System (VECOS), and monthly to bi-monthly surveys from the Chesapeake Bay Program's Water Quality Monitoring Program. We used the Du and Shen (2014) model to spatially interpolate bottom DO conditions using inverse-distance weighting; this horizontal interpolation allowed us to assign bottom DO values to each 1-$km^2$ grid cell. Daily interpolated DO values were estimated from monthly DO values for 2000 to 2016 by linear regression, where the observed daily change in DO was estimated by the slope of the corresponding regression for each grid cell.

To determine the validity of the interpolated bottom DO values, we ground-truthed model-based hindcasts with field observations from a subset of observations (2010 to 2012; n=4,604). Questionable hindcasts were those where the observed DO was less than or equal to 5 mg$O_2$/l (or 2 mg$O_2$/l) but the model-based estimate exceeded 5 mg$O_2$/l (or 2 mg$O_2$/l); about 1.0% (or 1.6%) of the estimated DO concentrations appeared questionable and were further cross-checked against field notes. Some of the discrepancies were due to malfunctioning DO meters (these 'observed' values were removed and the interpolation was repeated); other discrepancies were due to inconsistent deployment of DO probes (probe may have been more than 1 meter above the seabed; these values were removed from consideration). In one case, we found that the 1-$km^2$ resolution of the model could not capture localized low DO conditions: we interpolated DO to be greater than 5.0 mg $O_2$/l, but the on-site, instantaneous measurement was 0.27 mg $O_2$/l. Field notes indicated that all shrimp, crabs, and fishes captured at this site were dead. Because adjacent stations were normoxic, we concluded that a highly localized hypoxic event had been encountered by the survey. Thus, small-scale instantaneous measurements of environmental conditions will not match exactly the model-based interpolated values. Nevertheless, at least 98% of hindcasts from this model were reasonably accurate.

***Temperature, salinity, depth, and current speed covariates*** - Estimates of bottom temperature, bottom salinity, time-varying depth, and current speed were obtained from a three-dimensional model of the Chesapeake Bay developed using the UnTRIM hydrodynamic model (Casulli and Zanolli 2002; Casulli and Zanolli 2005). The UnTRIM hydrodynamic model was applied previously to large estuaries, such as Chesapeake Bay (Shen et al. 2006; Sisson et al. 2010; Wang et al. 2015) and San Francisco Bay

(Cheng and Casulli 2002; Bever and MacWilliams 2013; MacWilliams et al. 2015; Bever et al. 2016), and is well suited to perform hydrodynamic, water level, salinity, and temperature modeling in the Bay.

The Chesapeake Bay model uses an unstructured grid that allows grid cell sizes to vary spatially to increase computational efficiency and to directly resolve the complex shoreline and bathymetry of the Chesapeake Bay and its tributaries (Figure 2). This model takes advantage of the grid flexibility allowed in an unstructured mesh by gradually varying grid cell sizes, beginning with large grid cells in the Atlantic Ocean and transitioning to finer grid resolution in the smaller channels of the tributaries and the northern portion of the estuary. This approach offers significant advantages in terms of numerical efficiency and accuracy, and allows for local grid refinement for detailed analysis of local hydrodynamics, while incorporating the overall hydrodynamics of the larger estuary in a single model. The Chesapeake Bay model uses fixed vertical layers (Z-grid) with a vertical grid resolution of 0.5 m to a depth of 40 m below zero North American Vertical Datum of 1988 (NAVD88) and a resolution of 1 m thereafter. The Federal Emergency Management Agency Region III bathymetric and topographic digital elevation model in NAVD88 was used to specify bathymetry throughout the model domain (Forte et al. 2011). Bathymetric data from the U.S. Army Corps of Engineers, Baltimore District, were used for many of the main navigation channels to best include the bathymetry of dredged navigation channels (USACE 2016).

Observed water levels from the National Oceanic and Atmospheric Administration stations at the Chesapeake Bay Bridge Tunnel (CBBT, station 8638863) and Reedy Point (station 8551910) were used to specify water levels at the Atlantic Ocean boundary and Delaware side of the C&D Canal, respectively (Figure 2). Salinity and water temperature at the ocean boundary were specified based on monthly climatology values from World Ocean Atlas 2013. Average daily salinity and water temperature observations from the U.S. Geological Survey (USGS) Delaware River at Reedy Island Jetty (01482800) station were used at the Delaware side of the C&D Canal.

River inflows included 12 tributaries (Figure 2), representing the majority of the freshwater flow into the Chesapeake Bay. USGS discharge data were scaled based on the ratio of the gauged area to the overall drainage area to specify the freshwater discharges for the model input following methods in Xu et al. (2012). Salinity of the inflows was set to zero practical salinity units (psu), except for the Susquehanna, which was set to 0.1 psu based on data from the Chesapeake Bay Interpretive Buoy System (CBIBS) Susquehanna location. Temperature of the inflows was set using observations from various data sources based on availability and proximity to the inflow locations.

Wind, evaporation, precipitation, air temperature, incoming solar radiation, and relative humidity were specified using 3-hourly gridded North American Regional Reanalysis (NARR) products (Figure 2). NARR incoming radiation includes the effect of cloud cover, so the cloud cover for the heat flux calculation was set to zero. To help minimize the NARR underestimation of wind speed over the Chesapeake Bay (Scully 2013), the NARR wind speed was scaled based on relationships developed between NARR wind speeds and observed wind speeds from the National Data Buoy Center. Wind forcing was applied at the water surface as a wind stress with the wind drag coefficient varied based on local wind speed according to the formulation of Large and Pond (1981).

The Chesapeake Bay model was used to simulate the 17-year period spanning 2000 through 2016 for this study. The model was initialized in August 1999, providing 4.5 months for the model to spin up before 1 January 2000. Vertical profile data collected as part of the EPA Water Quality Monitoring Program (WQMP) were used to specify initial conditions for salinity and water temperature.

Prior to application of the Chesapeake Bay model, the system was delineated into 142 contiguous polygons to facilitate development of fish habitat maps and estimation of habitat areas for specific regions (e.g., the Potomac River). Polygons were delineated based on the Chesapeake Bay model grid to ensure that calculation of polygon characteristics was directly related to the spatial distribution of modeled environmental conditions used to characterize fish habitat. Multiple polygons were delineated in the Bay and each of the larger tributaries in the longitudinal and lateral directions, to capture the longitudinal salinity gradients and lateral variability in depth. Each longitudinal reach was subdivided into three lateral polygons using a left shoal, channel, and right shoal approach. A depth of 9.1 m (used here as -9.1 m North American Vertical Datum of 1988) was used to distinguish between channel and shoal regions, based on graphical inspection of catch rates by depth for each of the forage species. Not all longitudinal reaches included a left shoal, channel, and right shoal polygon. Also, for some of the mainstem Bay reaches, a center shoal polygon was delineated because of multiple channels deeper than the 9.1 m threshold. Small embayments were delineated with a single longitudinal reach containing a shoal and a channel polygon (where channel depths occurred in the polygon). These embayments included Fishing Bay, the Little Choptank River, the lower Nanticoke/Wicomico rivers, Baltimore Harbor to Patapsco River, and the upper Bay flats and lower Susquehanna River.

Both static (constant in time) and dynamic (time-varying) environmental variables were calculated for use in habitat suitability models and for estimating suitable habitat area throughout Chesapeake Bay. A total of 23 variables was calculated from the Chesapeake Bay hydrodynamic model and another variable was extracted from the numerical model of bottom dissolved oxygen (Table 2). A large number of environmental variables were initially considered for use in developing fish habitat suitability models to eliminate the need for *a priori* specification of environmental conditions that may be important for describing abundance and distribution of forage fishes.

Static variables did not vary with time and were determined at locations where trawl tows were completed based on the midpoint of each tow. The distance to shoreline was estimated for each tow by calculating the shortest distance to the shoreline, even if the closest shoreline was an in-Bay island (Figure 3). A seabed grain size distribution was developed for the Chesapeake Bay and tributaries based on observed surface seabed grains size data (Moncure and Nichols 1968; Byrne et al. 1983; Kerhin et al. 1988; Velinsky 1994; Maryland Geological Survey 1996; Reid et al. 2005), as part of the development of the Chesapeake Bay hydrodynamic, wave, and sediment transport model. This baywide surface grain-size distribution map was used to estimate a seabed percent fine sediment at the location of each tow (Figure 3).

Environmental covariates were extracted from the Chesapeake Bay hydrodynamic model at the same time and location of the individual tows to allow us to couple fisheries observations with hindcasts of environmental covariates. For the Virginia survey, habitat covariates were extracted from the model at the midpoint of each tow. The Maryland survey used fixed trawl station locations, and we initially assumed that the tows occurred at those fixed sites. The recorded depth from each Maryland trawl tow, however, varied considerably from the depth at the fixed station location; we determined this by comparing the recorded depth with water depth from the numerical model and depth from a digital elevation model. We therefore used the recorded water depth observed during each tow to estimate the actual geographic location of each tow by identifying the closest location to the fixed station location that had a water depth similar to the observed water depth at the time of sampling.

The Chesapeake Bay model was used to hindcast environmental covariates at multiple temporal and spatial scales. The use of multiple scales to describe environmental conditions may provide more

accurate predictions of habitat suitability (Lecours et al. 2015). Dynamic variables were extracted from the Chesapeake Bay model for two time periods. The first period was the instantaneous value at the time of the trawl tow. The second period was the time-average of the 24.8 hours before the trawl tow, encompassing one tidal cycle. The instantaneous values represent the conditions at the time of the tow and the tidal-averaged values represent the conditions a fish at the location of the tow would have experienced over the preceding tidal cycle. Tidally-averaged, depth-averaged conditions were also obtained for salinity, temperature, and current speed (Table 2). We also considered covariates describing near-bed conditions (one m above the seabed), and maximum depth-averaged current speed over the tidal cycle or tidally-average current speed. Maximum depth-averaged current speed was examined because habitats exhibiting relatively high tidal-averaged current speeds may be used by fish as long as the maximum current speed does not exceed a threshold. A simple tidal-averaged covariate provides a measure of current speed during flood or ebb tide, which may be used by some species to aid in movements within the estuary (e.g., Brady and Targett 2013). Because vertical or horizontal gradients in current speed may act to aggregate food near complex currents or fronts, we also considered these covariates. The vertical gradient in the current speed was calculated as the difference between the current speed one meter above the bottom and one meter below the surface; the horizontal gradient in the current speed was calculated as the maximum difference in current speed between adjacent model grid cells.

Although habitat conditions near the seabed may be most relevant for understanding fish-habitat relationships for demersal species such as spotted hake, habitat use may respond more strongly to overall conditions in the water-column because even demersal fishes are not confined to near-bed habitats. For example, salinity stratification may influence the supply of food or DO to the near-bed region sampled by the trawl, and may be an indicator of forage fish occurrence. Thus, we used covariates describing salinity and temperature stratification; these covariates were calculated as the difference between instantaneous surface (top 1 m) and near-bed (1 m above seabed) conditions. In addition, because favorable habitats may be characterized by a range of conditions, we considered covariates based on the percent of time that near-bed conditions fell within a given range, for example, the percent of time that salinity exceeded 20 psu. The percent of time within a given salinity or temperature range was calculated over the same time interval used for tidal-averaging of the other covariates. We identified three salinity ranges (< 10 psu; 10 to 20 psu; > 20 psu), and three temperature ranges (< 10° C; 10 to 20° C; >20° C) consistent with observed patterns in fish communities in Chesapeake Bay (Tuckey and Fabrizio, *pers. obs.*).

***Validation of the Chesapeake Bay hydrodynamic model*** – Salinity and temperature from hindcasts of conditions for the 17-year period of study were validated using the CBIBS buoy data (n=7 stations), WQMP vertical profile data (n=13 stations), WQMP stations in the tributaries (n=38 stations), and temperature and salinity observations recorded at the time of fish sampling in Virginia (n=20,334 observations; Figure 4). The use of multiple data sources throughout the estuary provides a robust model validation and allowed us to validate the model across tidal, seasonal, and interannual time scales. The model was validated in the mainstem of the Chesapeake Bay and in the major tributaries using methods detailed in MacWilliams et al. (2015) and Irby et al. (2016); these methods used the means and correlations of the observed and estimated values, model skill (Willmott 1981), and target diagram statistics (Jolliff et al. 2009) to assess the accuracy of the model. For brevity, we present results based on target diagram statistics, as in Irby et al. (2016), who compared the accuracy of eight models of the Chesapeake Bay.

The target diagram statistics determine how the mean and variability of the model estimates related to those of the observed data (Jolliff et al. 2009; Hofmann et al. 2011). This approach uses the bias and the unbiased Root-Mean-Square Difference (ubRMSD) between the observations and estimates, which are normalized by the standard deviation ($bias_N$ and $ubRMSD_N$) to assess the accuracy of model-based estimates. The $ubRMSD_N$ was multiplied by the sign of the difference between the observed and modeled standard deviations to indicate overestimation (positive) or underestimation (negative) of the observed variability. On target diagrams, the Y axis is $bias_N$ and the X-axis is $ubRMSD_N$. The radial distance from the origin to each data point is the normalized total Root-Mean-Square Difference ($RMSD_N$, calculated as $RMSD_N = \sqrt{bias_N^2 + ubRMSD_N^2}$). The $RMSD_N$ is a dimensionless number, where values less than 1.0 indicate that model estimates were more accurate than simply estimating the mean of the observations. Thresholds established by MacWilliams et al. (2015) allow assessment of the accuracy of hydrodynamic model estimates such that an $RMSD_N$ less than 0.25 indicates very accurate estimates, 0.25 to 0.5 indicates accurate estimates, 0.5 to 1.0 indicates acceptable estimates, and greater than 1.0 indicates relatively poor agreement between model-based estimates and observations.

The validation of model-based estimates of salinity and temperature using the co-located fisheries survey observations from Virginia waters for 2000 through 2016 demonstrated that the salinity and temperature observed during fisheries sampling were accurately estimated; the $RMSD_N$ was less than 0.44 for salinity and less than 0.23 for temperature (Figure 5). The model was similarly accurate for both surface and bottom values. The model-based estimates of salinity were slightly biased high at high salinity values (> 15 psu), and model-based estimates of temperature were slightly biased low at low temperature (< 5°C).

Data from the WQMP were used to validate salinity and temperature for each year in the mainstem and the tributaries separately. For brevity, we report results for the mainstem during the two years evaluated in Irby et al. (2016). Modeled salinity and temperature were similar in accuracy in 2004 and 2005 to other models evaluated using similar methods for the same years (see Table 3 in Irby et al. 2016). Model validation for 2000 through 2016 indicated that the model-based estimates of bottom salinity in the mainstem were more accurate after 2005. Stratification was accurately estimated but was slightly lower in magnitude than the observed stratification. The depth-to-maximum stratification was also accurately estimated by the model, with depth-to-maximum stratification slightly closer to the water surface for the model-based estimates than what was observed.

The detailed validation of the model-based estimates of salinity and temperature indicated that the estimated values were similar in accuracy to the suite of Chesapeake Bay models evaluated by Irby et al. (2016). Our validation also demonstrated that the model accurately estimated temperature and salinity in the major tributaries and accurately estimated the salinity and temperature co-located with the fisheries catch data. We concluded that the Chesapeake Bay model is sufficiently accurate under a wide range of environmental conditions for hindcasting fish habitat conditions for multiple timescales and multiple spatial scales.

## Selection of influential habitat covariates

Boosted regression trees (BRTs) were used to select a subset of influential habitat covariates that explained variations in catch rates of fish (number of fish per 5-minute tow). Regression trees make minimal assumptions about the underlying distribution of the abundance of fish and are ideal for identifying covariates that are important drivers of abundance (Breiman et al. 1984). The regression tree algorithm uses recursive partitioning to explain variation in the response (catch rates), that is,

observations are repeatedly split into increasingly homogeneous groups based on threshold values of the predictors (habitat covariates; Breiman et al. 1984). Cross-validation was used to assess model fit and to ensure that the resultant trees were applicable to out-of-sample observations; cross-validation was achieved by fitting the tree to a subset of the data (the training set) and fit was assessed using the remaining data (the test set). Furthermore, the performance of regression tree algorithms may be improved with ensemble methods such as boosting, which aggregates multiple trees to enhance the stability of the resultant model (Knudby et al. 2010). All habitat covariates considered in BRT models were standardized to permit direct comparison of covariate importance (Schielzeth 2010). We used a Poisson response to model the number of fish captured per tow with the R package 'dismo' and the *gbm.step* procedure (R Core Team 2019; Elith et al. 2008; Elith and Leathwick 2017). Catches from the Virginia survey were expressed as numbers of fish per 5-minute tow and used without modification, but catches from the Maryland survey, which sampled fishes for 6 minutes, were expressed in 5-minute-tow equivalencies rounded to the nearest integer.

***Optimization of BRT models*** – Prior to fitting the BRTs to the fisheries observations from 2000 to 2016 and the associated habitat covariates, we optimized the model-fitting parameters of the BRTs by exploring the combination of parameter values that produced the lowest deviance for cross-validated data sets (Elith et al. 2008; Cameron et al. 2014). To determine optimal parameters values for the BRTs and because optimization is computationally intensive, we used a subset of observations (2010-2012; n=4,604 tows) that represented notably different environmental conditions (2011 was a wet year compared with 2010) as well as large differences in the relative abundance of forage species. BRT parameters were optimized separately for spotted hake, weakfish, spot, and bay anchovy using the *gbm.step* procedure in R (R Core Team 2019). Model fitting failed for bay anchovy, so we optimized the BRT parameters for winter (Dec-Jan-Feb) and summer (Jun-Jul-Aug) samples separately for this species.

Optimization focused on selection of the learning rate and tree complexity, which are model-fitting parameters assigned by the analyst. The learning rate determines how quickly the model approximates the observed data (Miller et al. 2016), and the tree complexity represents the level of interaction possible among the predictors. Another parameter selected by the analyst is the bag fraction, or the proportion of the data considered for training the model. Observations for the training subset are selected randomly without replacement for each model run and the remaining observations are used for cross-validation. Preliminary investigations suggested that a bag fraction of 0.75 was reasonable. Using this bag fraction, we fitted a series of trees to a range of learning rates (0.0005, 0.0050, 0.0075, 0.0100, 0.0200, 0.0300, 0.0400, 0.0500, 0.0750) and tree complexities (1, 3, 5, 10), similar to Cameron et al. (2014). For each species, we considered only those BRTs for which at least 1,000 trees were fit and selected parameters that reduced the deviance in the cross-validated data (Elith et al. 2008). We identified the optimal tree complexity for each species by graphically examining the change in cross-validated deviance across learning rates using SAS® software. Next, using the selected tree complexity, we identified the learning rate that produced the minimum cross-validated deviance.

To select influential habitat covariates, we fitted BRT models for each species for the period 2000 to 2016 ($N_{total}$=25,333 tows; $N_{Virginia}$=20,326 tows, $N_{Maryland}$ = 5,007 tows) using a bag fraction of 0.75 and values of the optimized species-specific learning rates and tree complexities determined by optimization. In addition, optimization runs indicated that the six covariates describing percent time were least informative, so these were not considered further. Therefore, a suite of 18 covariates (16 dynamic, 2 static; Table 2) were considered for the BRT models. These modeling results allowed us to identify and select a subset of important covariates for each species from the estimates of variable influence and scree plots produced by the *gbm.step* procedure (R Core Team 2019).

## Habitat suitability models

Habitat suitability models were used to assign habitat suitability scores and to quantify the extent of suitable habitat for forage fishes throughout the Chesapeake Bay and its tributaries from 2000 to 2016, across the four seasons. Habitat suitability models were estimated with the nonparametric histogram approach because this approach makes no assumption about the nature of the relationship between environmental features and fish abundance (Guan et al. 2016). Briefly, thresholds of environmental conditions that resulted in a gradient of suitability indices (SIs) from least suitable (0) to most suitable (1) were identified for each influential habitat covariate. The HSI was calculated as the mean of two or more environmental-condition-specific SIs, and also ranged between 0 and 1 to ease interpretation.

***Suitability index*** - We estimated suitability indices (SIs) for the range of observed values for each of the influential habitat covariates identified by the species-specific BRTs. We used the approach described in Tanaka & Chen (2015) to estimate SIs but applied a disjoint clustering method to identify 'natural clusters' of the habitat covariates for the histogram approach; we implemented this method with the *FastClus* procedure in SAS/STAT® software (version 9.4 of the SAS System for Microsoft). Tanaka & Chen (2015) fixed the number of individual bins to 10 for each habitat covariate, but we found that this resulted in bins with few observations (<5) or narrowly defined limits (e.g., bottom temperature between 16.2 and 16.3 °C). Thus, we allowed the number of bins to vary (but not exceed 10), and restricted cluster sizes to a minimum of 40 observations; in all cases, the smallest cluster included at least 54 observations, allowing a reasonable description of average abundance in each cluster. Once the clusters were defined, the SIs were estimated using

$$SI_{ij} = \frac{CPUE_{ij} - CPUE_{i,min}}{CPUE_{i,max} - CPUE_{i,min}}$$

where $SI_{ij}$ is the suitability index for cluster j of habitat covariate i, $CPUE_{ij}$ is the average catch (fish/km$^2$) observed in cluster j of habitat covariate i, and $CPUE_{i,min}$ and $CPUE_{i,max}$ are the minimum and maximum average catches observed across all clusters of habitat covariate i (Tian et al. 2009; Chang et al. 2012). In this manner, the SIs ranged between 0 and 1.0, with 1.0 indicating the most suitable (single factor) conditions and 0 the least. More explicitly, each covariate cluster, defined by a range of values, was associated with an SI score.

***Habitat suitability index*** - HSIs were calculated for each sample site and each grid cell in the hydrodynamic model by expressing the HSI as an average of the SIs across multiple habitat covariates. Thus, HSIs were calculated from the individual suitability indices, but we restricted the number of covariates in the HSI to those that were most influential as determined by the BRTs (Table 3). Care was taken to consider only those covariates that did not exhibit high correlations with other influential covariates, that is, only those covariates with $r^2 < 0.8$ were considered in the calculation of the HSIs. In this manner, we avoided overweighting of the HSI for a particular habitat condition.

As an average across multiple SIs, the HSI at a given site can be expressed as an arithmetic mean or a geometric mean (e.g., Brown et al. 2000; Tanaka and Chen 2015). A single approach to estimation of the HSI may not be appropriate for all species (e.g., Yu et al. 2019), so we explored the two models of the mean. The arithmetic mean model for the HSI is given by

$$HSI_{am} = \frac{SI_1 + SI_2 + SI_3 + \cdots + SI_p}{p}$$

where $SI_1$ is the suitability index for habitat covariate 1, $SI_2$ is the suitability index for covariate 2, and so forth; and *p* is the number of covariates considered (e.g., Hess and Bay 2000). The geometric mean model for the HSI is

$$HSI_{gm} = \sqrt[p]{SI_1 \times SI_2 \times SI_3 \times \ldots \times SI_p}$$

(e.g., Layher and Maughan 1985; Lauver et al. 2002; Tian et al. 2009). The geometric mean index applies the concept of a 'limiting factor' whereby a low SI for a single covariate results in a low $HSI_{gm}$ (Zajac et al. 2015). Regardless of formulation, HSIs ranged between 0 and 1. All HSI calculations were performed in SAS® or Matlab (MathWorks Inc.).

The area of suitable habitat throughout the Chesapeake Bay and its tributaries for each species and season was estimated using a GIS by summing the areas of individual hydrodynamic model grid cells where HSI exceeded a given threshold of suitability. For example, we calculated the area within the Chesapeake Bay region where HSI exceeded 0.4, 0.5, 0.6, 0.7, and 0.8 for each species and season. Areas of suitable habitat for each forage species were visualized through time using a GIS; this capability facilitated estimates of forage habitat in areas of Chesapeake Bay that are not routinely sampled by fisheries surveys. That is, we estimated the extent and location of suitable habitats for forage fishes using hindcasts of habitat covariates from the two environmental models and estimates of static features (distance to shore and sediment composition).

***Calibration of HSIs*** – We calibrated the HSI models by graphical examination of the relationship between the HSI and the average relative abundance for each species-season combination (e.g., Tanaka and Chen 2015); these graphs were produced using SAS® software. We used trimmed means as a measure of the average because these means are insensitive to the occasional extreme (outlier) catches observed for some species; data were trimmed by 5%. For a properly calibrated HSI, the mean relative abundance of forage fish is expected to increase as habitat conditions approach optimal for the species, that as, as the HSI increases from 0 to 1.0.

***Verification of modeling approach*** – We verified the BRT approach for selection of covariates and evaluated the reliability of the two formulations of HSI for forage fishes using a cross-validation approach. In this approach, we randomly selected ~70% of the fisheries observations (N=18,121) to comprise the training data set, and the remaining ~30% (N=7,212) was used as the test (or verification) data set. The random selection of samples was without replacement and followed a stratified design to ensure representation across years, seasons, and geographic areas. Training and test data sets were constructed separately for each species; note that for this analysis we omitted fall observations for bay anchovy because the BRT model failed to produce a boosted regression tree with at least 1000 trees (Elith et al. 2008). For each training data set, we fitted BRTs, selected influential covariates, and modeled the $HSI_{am}$ and $HSI_{gm}$. Note that the BRTs for each data set may have indicated a different number of influential covariates, as well as a different suite of influential covariates, than what was identified in the original run using all the samples (i.e., the original species-specific BRT fitted to observations from 25,333 tows). The resulting HSI models were applied to each of the test data sets to estimate the predicted HSIs. Due to computational intensity, 10 cross-validation data sets were generated (consistent with Pennino et al. 2020). The expected performance of the $HSI_{am}$ and $HSI_{gm}$ metrics for each species and season was evaluated with the root mean square error (RMSE), which was calculated as the standard deviation of the residuals (i.e., the difference between the predicted HSI and observed HSI for each location of fish sampling). RMSE is a commonly used metric to evaluate model performance; we used a *t*-test to assess differences in the mean RMSE. We retained the HSI formulation that exhibited the lower RMSE for further analyses. The training and test data sets were created using

the *SurveySelect* procedure in SAS/STAT® software, and the *gbm.step* function in R (R Core Team 2019) was used for identification of important covariates; the *FastClus* procedure in SAS/STAT® software was used to delineate habitat suitability bins, and the *SQL* procedure in SAS® was used for subsequent data management.

***Transferability of the HSI model*** – In addition to calibration of the species-specific HSIs and verification of our use of BRTs and HSIs to describe suitable habitats, we assessed the transferability of our modeling approach using a two-fold block validation (Wenger and Olden 2012; Fletcher et al. 2019). Transferability refers to the ability to predict the suitability of habitats in regions that were not used to develop the HSI models, in our case, Mobjack Bay. We followed the general guidance in Wenger and Olden (2012) and evaluated HSIs for fisheries observations from Mobjack Bay which were not used in the identification of influential covariates, or in the development of the HSI models. Thus, these data represent an independent quantitative assessment of our approach (Theuerkauf and Lipcius 2016). Relative abundances of juvenile fish in Mobjack Bay were expressed as the mean number of fish captured per $m^2$ in summer 2010, 2011, and 2012 (n=129 tows). We graphically examined the pattern of relative abundance per unit of suitable habitat (derived from the HSI model) for juvenile spot and juvenile weakfish in Mobjack Bay and compared these patterns to those observed in Chesapeake Bay. For this analysis, we adjusted the baywide indices of relative abundance for juvenile spot and bay anchovy in Chesapeake Bay by multiplying these values by a constant (10 for juvenile spot, 100 for bay anchovy); this allowed us to compare values in the same order of magnitude for Chesapeake Bay and Mobjack Bay. No adjustments were necessary for estimates of the indices of baywide relative abundance for weakfish in Chesapeake Bay. Because we had fisheries observations from only 3 years, we refrained from fitting predictive models to these data and instead used a graphical approach implemented with SAS® software. Similar patterns in relative abundance per unit of suitable habitat across years for Mobjack Bay and Chesapeake Bay indicate that the HSI models can be used to estimate suitable habitat extents in areas not sampled by surveys in Maryland and Virginia.

## Coupling of fisheries surveys and physical models

***Calculation of extent of suitable habitat*** – Results from the habitat-suitability models were used to define suitable habitat by identifying the range of conditions that support each of the forage fishes (e.g., salinities > 20 psu with current speeds < 1 m/s). To facilitate estimation of habitat suitability for each forage species, we considered environmental and physical conditions (e.g., distance to shore, percent fine sediment) at each hydrodynamic model grid cell. For each hydrodynamic model grid cell, we calculated the daily (24-hr) average value for environmental covariates because this temporal framework corresponded best with the tidal-averaged (24.8-hr average) period used for some of the environmental covariates. The median daily value was used to estimate a seasonal value for each environmental covariate. The habitat suitability model was then used to estimate the species- and season-specific HSI value for each hydrodynamic model grid cell; this allowed us to quantify suitable habitat area for each year, species, and season. Interannual variability in suitable habitat was examined using various thresholds to define 'suitable' habitat (e.g., HSI $\geq$ 0.5, HSI $\geq$ 0.6, etc.) and depicted graphically to illustrate changes in extent of suitable habitat between 2000 and 2016. Preliminary investigations revealed that annual patterns in suitable habitat extents were similar among the 0.5, 0.6, and 0.7 thresholds; the extent of habitats with HSIs $\geq$ 0.8 was too low to be useful. Therefore, we selected the 0.5 threshold for subsequent analyses; this threshold value was used by Theuerkauf and Lipcius (2016) to describe habitat suitability for the eastern oyster *Crassostrea virginica* in Chesapeake Bay and by Birkmanis et al. (2020) for pelagic sharks in Australia. We present habitat areas for the entire Chesapeake Bay and its tributaries, as well as estimates for Virginia waters, Maryland waters, the

Potomac River, and Mobjack Bay. Estimation of suitable habitat area over these select regions demonstrated how output from the Chesapeake Bay hydrodynamic and dissolved oxygen models can be combined with habitat suitability models to estimate seasonal and interannual variability in fish habitat over multiple spatial scales and to estimate the amount of suitable habitat outside the areas sampled by fisheries surveys (Potomac River, Mobjack Bay).

***Mapping of suitable habitat*** – To facilitate mapping of seasonal habitat conditions for each species, seasonal environmental covariates, SIs, and HSI were calculated for each hydrodynamic model grid cell. This approach resulted in maps that spanned the Chesapeake Bay and its tributaries and captured the interannual and seasonal variability in habitat conditions. In this manner, we mapped the species-specific seasonal HSIs at the spatial resolution of the hydrodynamic model because processes operating at small spatial scales may be masked when environmental conditions are averaged over large spatial scales (Windle et al. 2012).

### Relationship between suitable habitat extent and relative abundance of forage species

Annual and seasonal changes in the area of suitable habitat may affect the abundance of forage species in temperate ecosystems. We addressed this hypothesis by relating the annual time series of suitable habitat with annual estimates of baywide relative abundance for each of the forage species (objective 2). We limited the exploration of these relationships to the season during which each species was most vulnerable to the trawl gear: bay anchovy in summer and winter, juvenile spotted hake in spring, juvenile spot in summer and fall, and juvenile weakfish in summer and fall. Note that we were unable to fit a BRT model to habitat covariates for bay anchovy in fall, and thus, fall HSIs were not available for this species. For bay anchovy in winter, we used the relative abundance index from the Virginia survey because the Maryland survey did not sample in winter. We rank transformed the abundance indices because of the small number of observations (n=17 years except for bay anchovy in winter, n=16) and conducted nonparametric regression analyses to examine the relationship between rank abundance and extent of suitable habitat. For this analysis, we assumed the relationship was stationary, that is, the effect of suitable habitat extent on the abundance of forage fish was constant through time (e.g., Zeng et al. 2018). In addition, we tested the null hypothesis that the extent of suitable habitat in Chesapeake Bay did not vary between 2000 and 2016; as before, we rank-transformed the extent of suitable habitat (defined as areas with HSI$\geq$0.5). Computations for nonparametric regression analyses were performed with the *rank* and *glm* procedures in SAS®. Inspection of the residuals against the predicted values (ranked relative abundance or ranked extent of suitable habitat, obtained with the *sgplot* procedure in SAS®) suggested that the residuals were uniformly distributed in all cases (rank regression assumes that the residuals are continuous; Kloke and McCain 2012).

## Results

### Influential habitat covariates

Optimization of BRTs yielded values for the model-fitting parameters (learning rate, lr; tree complexity, tc) that varied among species: bay anchovy summer: lr=0.02, tc=3; bay anchovy winter lr=0.02, tc=3; juvenile spot lr=0.02, tc=10; juvenile spotted hake lr=0.01, tc=10; and juvenile weakfish lr=0.005, tc=10. These species-specific parameters were used to fit the BRTs to the 2000-2016 observations.

Environmental conditions and habitat features that comprised suitable habitats varied among species. Furthermore, the number of influential covariates identified by BRTs varied among species and ranged between four and seven (Table 3). Conditions at sites sampled in Maryland waters differed from those in Virginia waters: in Maryland, sampled habitats tended to be shallower, closer to shore, warmer in summer, and cooler in fall than habitats sampled in Virginia. Most notably, Maryland sites exhibited lower bottom DO concentrations in summer than Virginia sites (Figure 6). In addition, relative to sites in Virginia, Maryland sites exhibited lower surface salinities, less stratification in terms of salinity and temperature, lower current speeds and less stratification in current speeds (Figure 6). Water depth and one of the current speed metrics were consistently identified as influential covariates for all species; one of the temperature covariates was influential in describing suitable habitats for forage fishes in spring, summer, and fall, but was not selected for describing suitable habitats in winter (note, however, that the correlation between temperature and dissolved oxygen, which was important in winter, suggests that temperature may play a role in delineating suitable habitats in winter). Salinity defined suitable habitats for bay anchovy and juvenile spotted hake, and distance to shore explained suitable habitats for bay anchovy, juvenile spot, and juvenile weakfish. Bottom DO conditions delineated suitable habitats for bay anchovy in summer and winter and for all seasons for juvenile spot.

## Description of suitable habitats for forage fishes

The relationships between environmental covariates and their corresponding suitability index varied by species and season. Suitability indices reflected seasonal patterns in habitat use of forage species in a manner consistent with observations based on trawl survey catches. In Chesapeake Bay, for example, juvenile spot are more commonly observed away from the shoreline in summer, whereas in fall, juvenile spot may also be found close to shore. In contrast, juvenile weakfish use habitats near the shoreline in summer, and away from the shoreline in fall. Specific descriptions of suitable habitats for each species and season are presented below using the 0.5 threshold to define suitable habitat (i.e., conditions for which $SI \geq 0.5$).

For juvenile spot in summer, suitable habitats were more than 11,818 m from shore, shallower than 5.1 m or deeper than 10.2 m, and were characterized by tidal-averaged current speeds less than 0.000389 m/s and tidal-averaged temperature stratifications greater than 2.17°C (Figure 7). In addition, juvenile spot occupied habitats with low dissolved oxygen (< 4.8 $mgO_2$/l) in summer; maximum suitability occurred in habitats characterized by DO concentrations between 2.2 and 3.2 $mgO_2$/l (Figure 7). In fall, suitable habitats for juvenile spot were between 1,226 and 3,397 m from shore and deeper than 13.0 m; these habitats were characterized by tidal-averaged current speeds between 0.000064 and 0.000227 m/s, and tidal-averaged temperature stratifications greater than 2.78°C. Habitats with DO less than 6.2 $mgO_2$/l were occupied by juvenile spot in fall and maximum suitability occurred in habitats characterized by DO concentrations between 4.0 and 5.3 $mgO_2$/l.

Suitable habitats for juvenile weakfish in summer were characterized by tidal-averaged bottom temperatures greater than 25.9°C, depths greater than 7.7 m, and tidal-averaged current speed stratification between 0.066 and 0.293 m/s; these habitats were located less than 3,760 m from shore. Suitable habitats for juvenile weakfish in fall were characterized by tidal-averaged bottom temperatures greater than 24.5°C, depths greater than 10.4 m, tidal-averaged current speed stratification greater than 0.1 m/s, and were located generally more than 5,104 m from shore.

Suitable habitats for juvenile spotted hake in spring were deeper than 12.7 m, and characterized by tidal-averaged bottom temperatures between 5.3 and 14.2°C, tidal-averaged salinity stratification greater than 4.9 psu, and maximum depth-averaged current speeds that exceeded 0.5 m/s.

Suitable habitats for bay anchovy in summer occurred where the tidal-averaged salinity stratification ranged between 4.3 and 11.1 psu, substrate composition ranged between 1.2 and 14.4% fine sediment, distance to shore exceeded 3,784 m, tidal-averaged bottom temperatures ranged between 23.7 and 27.0°C, tidal-averaged surface salinity ranged between 17.1 and 26.0 psu, depth ranged between 5.1 and 16.1 m, and the horizontal gradient in tidal-averaged current speed was less than 0.000179 m/s. In winter, suitable habitats for bay anchovy were characterized by bottom DO concentrations between 6.6 and 10.4 $mgO_2/l$, distance to shore greater than 5,220 m, tidal-averaged surface salinity greater than 23.7 psu, substrate composition that ranged between 6.9 and 46.3% fine sediments, depths between 9.1 and 13.2 m, and horizontal gradients in tidal-averaged current speed less than 0.000084 m/s (Figure 8).

## Habitat suitability indices

***Verification of modeling approach*** – Bootstrap analyses verified that BRTs were a reliable means to select influential covariates; we found that, in general, the same or similar covariates were consistently identified as most influential among the 10 bootstrap realizations. For juvenile spotted hake, the HSI formulation based on the geometric mean ($HSI_{gm}$) provided the best approximation to the original HSI estimated for each sample as indicated by the significantly lower RMSE (Figure 9; $t = 4.56$, $P < 0.05$). Unlike the results for juvenile spotted hake, the HSI based on the arithmetic mean ($HSI_{am}$) performed better for bay anchovy (Figure 9; $t = -5.27$, $P < 0.05$); although we found no evidence for a difference in the mean RMSEs for the $HSI_{am}$ and the $HSI_{gm}$ for juvenile weakfish and juvenile spot (Figure 9; $t_{weakfish} = -1.65$, $P = 0.12$; $t_{spot} = -1.13$, $P = 0.27$), we used the $HSI_{am}$ for these species because the mean RMSE of the $HSI_{am}$ was consistently less than the mean RMSE of the $HSI_{gm}$.

Average relative abundance (estimated by the trimmed mean) for each of the four species increased as the HSI approached 1.0 (Figure 10), indicating proper calibration of the habitat suitability models. The ranges of observed HSI values across years were 0 to 0.92 for bay anchovy in summer; 0.04 to 0.98 for bay anchovy in winter; 0 to 0.95 for juvenile spotted hake in spring; 0.12 to 0.98 for juvenile spot in summer; 0.04 to 0.86 for juvenile spot in fall; 0.09 to 0.99 for juvenile weakfish in summer; and 0 to 0.89 for juvenile weakfish in fall (Figure 11).

***External validation of HSI models with fisheries catch data from Mobjack Bay*** – The extent of suitable summer habitat area for juvenile spot and juvenile weakfish in Mobjack Bay was about 1% of that observed in Chesapeake Bay, yet annual changes observed in Chesapeake Bay were reflected in Mobjack Bay (Figure 13). This congruent pattern suggested that annual fluctuations in environmental conditions in summer are coherent among areas. Annual changes in the index of relative abundance for juvenile spot per unit of suitable habitat were similar for Mobjack Bay and Chesapeake Bay in summer (Figure 13), indicating a consistent relationship between the extent of suitable habitat and juvenile spot abundance among the two areas. Similarly, annual changes in the index of relative abundance for juvenile weakfish and bay anchovy per unit of suitable habitat in Mobjack Bay displayed the same pattern observed for fish in Chesapeake Bay (Figure 13). The annual pattern for juvenile weakfish was less clear than that for juvenile spot due to the imprecision of the annual estimates of the index of relative abundance for juvenile weakfish (Figure 13).

## Suitable habitat extent for forage fishes

For a given species, habitat suitability and the extent of suitable habitat varied annually (e.g., Figure 14) and seasonally (e.g., Figure 15). The HSI maps were useful tools for visualizing the dynamic nature of habitats for forage species and for identifying times and locations that may potentially support high abundances of forage fishes.

***Juvenile spotted hake*** – We found a strong seasonal pattern in the extent of suitable habitat for juvenile spotted hake in Chesapeake Bay, such that little to no suitable habitat was available in summer and fall; suitable habitat area increased in winter, and was greatest in spring (Figure 15). This seasonal signal in the suitable habitat extent appeared to be driven by seasonal changes in water temperature in the Bay. Winter habitat extent increased during years when waters began to warm earlier (2012) than when waters warmed later in winter (2011). Similar seasonal patterns in the amount of suitable habitat were predicted for most of the subregions evaluated. During the 17 years examined, we estimated a lack of suitable habitat for spotted hake in Mobjack Bay. Except for 2007, suitable habitat in winter was sparse in the Potomac River until 2012, indicating that the Potomac River may now be warming earlier in the year or not cooling as much in winter. These results for the Potomac River also demonstrated that the pattern in the annual extent of suitable habitat in the tributaries does not always coincide with that observed in the bay proper. Suitable habitat areas for spotted hake in spring were relatively deep, away from the shoreline and marshes, and exhibited pronounced salinity stratification.

***Juvenile spot*** - The extent of suitable habitat for juvenile spot displayed a strong seasonal pattern, with relatively little suitable habitat in fall and winter and relatively more suitable habitat area in spring and summer (Figure 16). The seasonal signal in the extent of suitable habitats for juvenile spot appeared to be driven by the combination of the environmental and physical covariates and was not well described by a single covariate. The seasonal pattern in the amount of suitable habitat predicted for the entire Chesapeake Bay was also predicted for the Virginia, Maryland, and Potomac River subregions. In tributaries and embayments smaller than the Potomac River that are relatively close to the shoreline (e.g., Mobjack Bay), the extent of suitable habitat for juvenile spot in spring was generally greater than that in summer. For juvenile spot, suitable habitats were primarily found in shallow areas near the shoreline in spring, but in the deeper portions of the Chesapeake Bay and its tributaries in summer (Figure 16).

***Juvenile weakfish*** – The amount of suitable habitat for juvenile weakfish exhibited a strong seasonal pattern: little suitable habitat in winter and a greater extent of suitable habitat in summer; the extent of suitable habitats in fall was generally greater than in spring (Figure 17). In Mobjack Bay, the greatest extent of suitable habitat occurred in summer, and no suitable habitat occurred in fall and winter. For this species, we observed similar extents of suitable habitats in Virginia waters in summer and fall, but markedly lower extents of suitable habitats in fall compared with summer in Maryland waters. This result is due to the observation that HSI values in the mainstem of the bay in fall were greater in waters south of the Rappahannock River than in waters north of the Rappahannock River (Figure 17). In summer, suitable habitats for juvenile weakfish were found close to the shoreline of the Chesapeake Bay and in the tributaries; in fall, suitable habitats were located at the mouth of the Potomac River and in the lower Chesapeake Bay.

***Bay anchovy*** – Like other forage fishes, the estimated extent of suitable habitat for bay anchovy exhibited a strong seasonal pattern, with the greatest extent of suitable habitat area in spring, and generally greater extent of suitable habitat area in summer than in winter (Figure 18). This seasonal pattern was true for the Chesapeake Bay, and for Virginia waters and the Potomac River. The relative difference in suitable habitat area between spring and the other seasons was larger for the Potomac River than for the larger regions that include the mainstem portion of the Bay. In Maryland waters, a greater extent of suitable habitat area was often observed in winter compared with summer. The seasonal and interannual variability in suitable habitat extent for the smallest subregion, Mobjack Bay, was notably more complex than for the other regions. The HSI maps suggested that the distribution of bay anchovy throughout the Chesapeake Bay varied seasonally, with habitat conditions more suitable

for bay anchovy in the tributaries in spring than in summer and winter. Suitable habitat conditions occurred in areas that reflected the interplay of multiple environmental and physical covariates, and we were unable to succinctly describe suitable habitat areas for this species. Perhaps this is why we encountered difficulty in development of habitat suitability models for this species and why different covariates were used for each season.

## Suitable habitat and relative baywide abundance

Seasonal indices of relative abundance for forage fishes in Chesapeake Bay were variable across years, and interannual patterns in the Virginia survey were generally similar to those in the Maryland survey, particularly for juvenile spotted hake in spring, juvenile spot in summer and fall, and juvenile weakfish in summer (Figure 19). Inconsistent patterns were observed for weakfish in fall and bay anchovy in summer, suggesting that seasonal processes affecting abundance and habitat use of these species varied across regions (Maryland, Virginia) of the Chesapeake Bay. We note that the mean index of relative abundance for weakfish in fall was several orders of magnitude lower in Maryland waters than in Virginia, and as such, the Maryland survey index may not reflect the overall pattern of abundance for this species in the Chesapeake Bay. Similarly, the mean index of relative abundance for bay anchovy in summer was an order of magnitude lower in Maryland waters than in Virginia.

Two contrasting relationships were detected between the ranked baywide relative abundance index and the extent of suitable habitat for forage fishes, where the extent of suitable habitat was determined using an HSI threshold of 0.5. We observed a significant positive relationship between seasonal ranked baywide relative abundance and extent of suitable habitat for juvenile spot in summer ($F=4.57$, $P=0.05$; Table 4; Figures 20 and 21). The baywide relative abundance index for juvenile spot in summer was highly variable with contrasting periods of low abundance (2001, 2002, 2003, 2009) followed by a single year of high abundance (2010) and several years of low abundance (2014, 2015, 2016; Figure 21). Note that the extent of suitable summer habitat for juvenile spot exhibited no significant linear pattern across time ($F=0.01$, $P=0.93$; Table 4). We observed a similar relationship for bay anchovy in winter: the extent of suitable habitat in winter was a significant determinant of the ranked relative abundance of bay anchovy as determined by the VA survey ($F=19.98$, $P<0.01$; Table 4; Figures 20 and 21). Suitable winter habitat for bay anchovy also exhibited no systematic change through time ($F=0.17$, $P=0.69$; Table 4).

More commonly, we were unable to detect a significant relationship between the area of suitable habitat and the rank-transformed estimate of baywide relative abundance of juvenile spotted hake in spring, juvenile spot in fall, juvenile weakfish in summer and fall, and bay anchovy in summer (Table 4). For juvenile spotted hake in spring, we found no indication that the extent of suitable habitat was limiting, except perhaps in 2002 when the area of suitable habitat ($HSI \geq 0.5$) declined below 1,600 $km^2$ and the ranked abundance index was among the lowest observed in the time series. This, however, may be coincidental. The extent of suitable spring habitat for spotted hake varied without trend since 2000 ($F=0.21$, $P=0.65$; Table 4). The index of ranked abundance for juvenile spot in fall was highly variable, exhibiting periods of low abundance before and after 2010. The extent of suitable habitat for spot in fall exhibited no trend through time ($F=1.11$, $P=0.31$; Table 4) and was markedly less than in summer. We found no evidence of an effect of the extent of suitable fall habitat on the ranked relative abundance of juvenile spot in fall ($F=0.01$, $P=0.93$; Table 4). The baywide relative abundance of juvenile weakfish was variable in summer and fall, with little contrast across years (Figure 19). Relative baywide abundance of juvenile weakfish in summer 2001 was high, but otherwise, abundance remained somewhat stable. In contrast, the extent of suitable habitat for juvenile weakfish increased significantly in summer ($F=10.39$, $P<0.01$; Table 4; Figure 22A) and fall ($F=8.68$, $P=0.01$; Table 4; Figure 22B). The

relative abundance of juvenile weakfish, however, exhibited no detectable response to increases in the extent of suitable habitats in either summer or fall ($F_{summer}$=2.75, P=0.12; $F_{fall}$=0.12, P=0.73; Table 4). The relative abundance of bay anchovy in summer was highly variable and annual estimates were imprecise. Although the extent of suitable habitat for bay anchovy in summer increased significantly since 2000 (F=24.37, P<0.01; Table 4; Figure 22C), we were unable to detect a response in the ranked relative abundance of bay anchovy to changes in the extent of suitable habitat in summer (F=0.06, P=0.80; Table 4).

## Discussion

Our modeling framework combined the power of machine learning to identify influential habitat covariates with the flexibility of nonparametric approaches to characterize habitat suitability and the capabilities of GIS to quantify and depict suitable (and unsuitable) habitats for forage fishes in Chesapeake Bay from 2000 to 2016. We coupled information from fishery surveys with static features of the environment and modeled values of dynamic conditions to identify locations in the Chesapeake Bay that serve as suitable habitats for forage fishes. In an ecosystem-based approach, these locations may be targeted for protection (e.g., by limiting fishing activities that may incidentally capture or injure forage fishes) or restoration (e.g., by improving water quality conditions), thereby ensuring production of sufficient forage for predators. In addition, the consequences of aquatic habitat alterations, whether due to climate change or physical disturbances can be investigated using projections of environmental conditions and habitat suitability in the region (e.g., Brown et al. 2013). Importantly, our modeling approach for building forage-fish habitat suitability models was verified and validated for the Chesapeake Bay, thereby allowing estimation of habitat suitability for regions that are not routinely sampled by fishery surveys (e.g., Mobjack Bay, Potomac River). Furthermore, our results allow resource managers to focus protection measures on areas with critical habitats (e.g., locations that persistently support suitable habitats) in Chesapeake Bay and to identify environmental conditions that affect the suitability of habitats for forage fishes. We found annual patterns in suitable habitat extent that mirrored those of baywide relative abundance for two forage species; as such, estimates of the minimum habitat area required to produce a desired abundance (or biomass) of forage fish can be used to establish quantitative habitat targets (Kritzer et al. 2016) or spatial thresholds that may serve as spatial reference points for management (Reuchlin-Hugenholtz et al. 2016). In our study, two species exhibited a relationship between relative abundance and extent of suitable habitats – juvenile spot in summer and bay anchovy in winter. Quantitative habitat targets and spatial reference points for these species warrant further consideration.

Suitable seasonal habitat extents for forage species exhibited annual changes reflecting temporal heterogeneity in habitat conditions in Chesapeake Bay, with a strong seasonal signal. Current speed, water depth, and either temperature or dissolved oxygen were identified as important covariates for the four forage species we examined, and distance to shore was important for three of the four species; thus, suitable habitat conditions resulted from a complex interplay between water quality and the physical properties of the habitat. Variation in seasonal extents were more pronounced (e.g., juvenile spotted hake) than annual variations in suitable habitat extent indicating that the Chesapeake Bay serves as a nursery area for juvenile fishes but the nursery function is temporally restricted. For some species, the extent of suitable seasonal habitat increased since 2000 (juvenile weakfish in summer

and fall, and bay anchovy in summer), whereas for other species, extents varied annually with no clear trend. None of the species examined were at the southern limit of their geographic range, and as waters of the Chesapeake Bay continue to warm (Hinson et al. *in review*), we expect that suitable habitat extent may increase for species with broad thermal tolerances such as spot, bay anchovy, and weakfish. Salinity may, however, mediate the ability of species to tolerate higher temperatures and could serve to limit suitable habitats in the future. It is unclear how the interaction of salinity, temperature, and other covariates will affect habitats in Chesapeake Bay under expected climate-change scenarios. Laboratory-based investigations of the interactive effects of salinity on the thermal tolerances for these species could be informative.

The relationship between the extent of suitable habitat and relative abundance of forage species was species dependent and, when present, varied seasonally. Such relationships have not been widely explored for aquatic species; to our knowledge, only two other studies (Le Pape et al. 2003; Yu et al. 2019) attempted to relate extent of suitable habitat and relative abundance. In the Yu et al. (2019) study, a graphical assessment was used to note the consistency between declines in the relative abundance of neon flying squid *Ommastrephes bartramii* and the spatial shrinkage of suitable habitats in the northwest Pacific Ocean (Yu et al. 2019). In contrast to Yu et al. (2019), we used nonparametric regression to statistically evaluate the strength of such relationships for forage fishes in Chesapeake Bay and found a positive relationship between suitable habitat extent and baywide relative abundance of bay anchovy in winter and juvenile spot in summer, suggesting that environmental conditions may affect the carrying capacity of the Chesapeake Bay for these two forage species during a portion of the year. Seasonal variation in the location of suitable habitats has been demonstrated for several estuarine species, and linked to variation in freshwater input (e.g., Rubec et al. 2019). In Chesapeake Bay, freshwater input influences salinity and salinity stratification, however, we identified additional hydrodynamic covariates such as temperature stratification and current speed that contributed to variation in suitable habitats. For example, for bay anchovy, the suitability of winter habitats was partly determined by dissolved oxygen and the horizontal gradient in the tidally averaged current speed. Although water temperature was not considered in the HSI model for bay anchovy in winter due to the correlation with dissolved oxygen, we cannot rule out temperature as an important covariate describing habitat conditions for bay anchovy in winter. Interestingly, the HSI model for juvenile spot also included bottom dissolved oxygen; suitable habitats exhibited DO levels less than 4.0 mg $O_2$/l in summer and less than 5.3 mg $O_2$/l in fall, suggesting that juvenile spot use habitats that may be considered marginal for DO. Low DO conditions are associated with warmer waters and indeed, juvenile spot were more likely to be observed in habitats where the tidally-averaged temperature stratification exceeded 2.2°C in summer and 2.7°C in fall. Habitat conditions in the Chesapeake Bay region were not limiting for juvenile spot in fall, suggesting that suitable habitat extent in fall exceeds that necessary to support the population of spot that remains in the system by the end of summer. One possible explanation for the decoupling of habitat extent and relative abundance for juvenile spot in fall concerns temperature. Mean water temperatures in Chesapeake Bay are greatest in late August-early September and the increased metabolic rates and energy demands of predators during this time may increase predation mortality on juvenile spot. Rising water temperatures in early fall may also affect the phenology of spot emigration resulting in earlier emigration and fewer juvenile spot remaining in the Chesapeake Bay during fall in warmer years. Although we were unable to detect a decline in the extent of suitable fall habitats for juvenile spot, warming water temperatures in fall associated with directional climate change may deteriorate habitat conditions for this species. Continued monitoring of fall abundances and a better understanding of the cues that trigger spot emigration in fall will be necessary to address this hypothesis.

A decoupling of favorable habitat extent and relative abundance was also observed for juvenile spotted hake in spring and juvenile weakfish in summer and fall. Factors other than the extent of suitable habitat likely affected relative seasonal abundances of these species-season combinations as well as juvenile spot in fall and bay anchovy in summer. That is, even in years when suitable habitat extent was relatively low, the availability and extent of seasonal habitats were sufficient to support these forage species and therefore, other factors such as predation (Minello et al. 1989) or food availability (Tableau et al. 2016) may have contributed to changes in relative abundances. In particular, our results for juvenile weakfish are consistent with the observed increase in natural mortality rates for this species in the 2000s (ASMFC 2019; Krause et al. 2020). Although suitable habitat extent for juvenile weakfish in summer and fall increased significantly since 2000, abundance of juvenile weakfish was affected by factors other than (or in addition to) the extent of suitable habitat. For weakfish, the sources of increased natural mortality remain unclear, but increased levels of predation and interspecific competition are believed to have played a role (ASMFC 2019). Abundances of other forage fishes may be limited by a potential lack of prey, or by declines in suitable habitats used by earlier life stages such as eggs and larvae. Another possible reason for the lack of a relationship between suitable habitat area and baywide relative abundance is that the seasonal relative abundance indices for these species were imprecise and hence were statistically invariable across time (based on 95% credible intervals of the baywide hierarchical index). Thus, a 'good' year with relatively high mean relative abundance index was not statistically discernible from a 'poor' year with relatively low index; this lack of contrast may have hampered our ability to detect a relationship between relative abundance indices and suitable habitat extents. Such results suggest that if population abundance is changing, the sampling intensity (temporal, spatial) of current fisheries surveys is insufficient to detect such changes. Alternatively, abundance may be fairly stable (but variable) across years.

In this study, we included information from two fishery-independent surveys, one of which (the Maryland survey) yielded fewer annual observations but sampled shallow-water habitats across a large portion of the estuarine salinity gradient. Overall, our fisheries observations were collected at a relatively fine spatial resolution (> 100 sites sampled/month) and high temporal intensity (monthly), thereby minimizing biases due to seasonal or short-term habitat use (e.g., by sampling only one or two months each year). Our initial concern that inclusion of observations from a temporally less intense survey could weaken our ability to detect relationships between the extent of suitable habitat and relative abundance of forage fishes was unfounded. For example, we were able to observe an effect of DO on the relative baywide abundance of juvenile spot in summer because low DO conditions are more prevalent in summer in Maryland waters of the Chesapeake Bay. The Chester River, Eastern Bay, Little Choptank River and Patuxent River in particular exhibited DO levels in summer that were lower than what was typically observed in Virginia waters in summer. In addition, sites sampled in Maryland waters also provided observations from shallow habitats close to shore, and these conditions were not well sampled in Virginia waters, thus, the two surveys together provided observations from a greater range of environmental conditions commonly encountered in the Chesapeake Bay region.

Hydrodynamic models and other numerical models of environmental conditions can provide information on dynamic habitat features that are not measured at the time of sampling and represent a significant step towards refining spatial relationships between fish and their environment (e.g., Crear et al. 2020). Consideration of such information may yield habitat models with greater predictive accuracy (Scales et al. 2017). In our study, we used daily and tidal-averaged conditions to develop habitat suitability models that reasonably reflected the relationship between (daily) environmental conditions and relative abundance (at the tow level) of forage fishes in Chesapeake Bay. This fine-scale approach is

preferable to one that uses seasonal averages of habitat conditions to build habitat suitability models (Scales et al. 2017). Indeed, the temporal resolution of the environmental covariates used to build the suitability model affects the scale of inference. For example, Woodland et al. (2020) recently described annual patterns in the distribution and abundance of forage fishes and invertebrates relative to patterns of predation and environmental conditions in the Chesapeake Bay and its major tributaries. Seasonal changes, however, could not be addressed in that study because several habitat conditions were represented by annual means (e.g., average discharge from tributaries, and the Atlantic Multidecadal Oscillation [AMO] index). Large-scale climatic changes as indexed by the AMO affect mean abundances of bay anchovy and juvenile spot (Woodland et al. 2020), and our findings for bay anchovy in winter and juvenile spot in summer are consistent with results presented in Woodland et al. (2020). Specifically, our results suggested that environmental conditions contribute to the observed variation in relative abundance of these forage species. Although we did not consider large-scale climate indices *per se*, we did examine small-scale environmental indicators of climate change (temperature, salinity) and demonstrated how these changes affect habitat suitability and relative baywide indices of abundance for bay anchovy in winter and juvenile spot in summer. Our study did not, however, address possible lags in the response of forage species to environmental warming, nor did we consider the effect of warming rates; rather than examining multiple drivers of forage abundance, we focused on the characterization and role of suitable habitat. Unlike Woodland et al. (2020) who found greater relative abundance of bay anchovy in the upper bay than in the lower bay, our indices of relative abundance for bay anchovy in summer were an order of magnitude lower in Maryland than those observed in Virginia waters, perhaps because we lacked samples from deep sites (>2.7 m) in Maryland (compared with maximum depths in Virginia of 13.2 m). Similar to Woodland et al. (2020), we found that bay anchovy and juvenile spot exhibited higher relative abundances in southern tributaries than in northern tributaries of the Chesapeake Bay, but our spatial depiction of suitable habitat conditions throughout the system allowed us to examine the fine-scale spatial distribution of suitable habitats. Such depictions are helpful for identifying the geographic focus of management efforts to protect or restore habitats.

We used BRTs to identify influential covariates from a large suite of possible covariates, similar to Georgian et al. (2019) who used random forests to select 13 covariates from a set of 30 possible covariates describing habitat conditions of deep-sea environments. Satellite imagery, ocean observing systems, and hydrodynamic models yield a multitude of environmental descriptors of habitat and these data are commonly used to study habitat ecology of fishes. As the number of habitat descriptors available to researchers increases, variable reduction techniques are critical to selection of influential covariates, that is, covariates that are useful to explain the variation in observed abundance and distribution of aquatic organisms. We applied BRTs to identify a subset of covariates useful in describing habitat conditions that affect the relative abundances of forage fishes in Chesapeake Bay. Based on our observations with BRTs, tree complexity can play a large role in improving the outcome of cross-validation and model fitting, and should therefore be optimized based on the data under consideration. Many researchers either fail to optimize regression trees or when optimization is implemented, only a single parameter is optimized (typically learning rate; e.g., Georgian et al. 2019, Yu et al. 2020) after using the default bag fraction (0.75) and selecting an arbitrary value for tree complexity (typically between 2 and 5; e.g., Georgian et al. 2019, Pennino et al. 2020, Yu et al. 2020). Due to the lack of consistency among published studies, approaches and guidelines for optimization of regression trees for ecological data warrant further research, particularly as BRTs appear useful as a variable selection technique.

Across the four species we examined, the geometric mean formulation of the HSI was best for juvenile spotted hake, but otherwise, the arithmetic mean formulation was preferred. The $HSI_{gm}$ is widely used but may penalize the index too harshly for mobile species that can tolerate broad variations in environmental conditions, including sub-optimal conditions for limited periods of time. For instance, in areas where DO is less than 2.0 mg/L, the individual suitability index for DO is likely to be 0; in this case, the value of the $HSI_{gm}$ is also 0, but other environmental conditions in these areas may be suitable, even optimal, and thus, the overall habitat suitability may not be well indexed by an $HSI_{gm}$ value of 0. For instance, the hypoxia tolerance of juvenile spot is insensitive to a broad range of temperatures commonly observed in Chesapeake Bay (10 – 30° C; Marcek et al. 2019), so an area characterized by DO < 2.0 mg/L and bottom temperatures around 25° C may be more suitable than indicated by the $HSI_{gm}$ index. Temperature, and other factors including salinity, may affect hypoxia tolerance in fishes, but the nature of such interactions is quite variable among fishes (Rogers et al. 2016). As such, seasonal models may be better able to reflect changes in habitat suitability for species that exhibit temperature-mediated environmental tolerances such as hypoxia tolerance in striped bass (Lapointe et al. 2014) and summer flounder (Capossela et al. 2012). Selection of the $HSI_{gm}$ or $HSI_{am}$ formulation clearly depends on species, and thus, rather than arbitrary selection, we recommend consideration of multiple formulations of the HSI and use of data-driven analyses and assessment of model performance to inform selection (Chang et al. 2012; Tanaka and Chen 2015; Yu et al. 2019; this study).

## Future applications

We examined habitat suitability for four forage fishes that are important components of the food web structure in Chesapeake Bay as indicated by contemporary trophic analyses of fish predators in the region. As the system continues to warm, forage species such as penaeid shrimp will likely increase in abundance (Tuckey et al. *in press*); we expect penaeid shrimp will become a significant component of the diet of many predators in Chesapeake Bay as they have in other estuarine systems (Minello and Zimmerman 1983; Minello et al. 1989; Fujiwara et al. 2016). Indeed, we hypothesize that penaeid shrimp may become as important as Mysids currently are for fishes of Chesapeake Bay (Buchheister and Latour 2015). Habitat suitability models for penaeid shrimp could be developed from existing data from trawl surveys in Virginia and Maryland, but supplementation of such surveys with additional shallow-water surveys, particularly in Virginia, will be required.

The accuracy of projections of habitat suitability models to areas not sampled by fishery-independent surveys or to years not included in the models is overestimated by commonly used measures of model performance (Wenger and Olden 2012), and thus, model transferability must be assessed. Transferability refers to model generality and how well a model can project into new geographic regions or times (Elith and Leathwick 2009). Elith et al. (2010) and Wenger and Olden (2012) demonstrate this with presence/absence data, but there is a need to extend this approach to abundance data and to develop guidelines for efficient estimation. Transferability assessments may not be necessary if "projections do not extend beyond the conditions represented by the data used to fit the model," that is, as long as projections avoid extrapolation beyond the range of predictors used in the training data set or extension to areas where novel combinations of predictors occur (Conn et al. 2015). Our fisheries observations came from two surveys that sampled across a broad geographic area in the largest estuary in the US; these observations represent 17 years of monthly sampling, and as such, reflect the breadth of habitat conditions that fishes are likely to encounter in Chesapeake Bay. The Virginia survey employed a stratified random design and the Maryland survey used a targeted (non-random) design; when observations are spatially extensive, the integration and use of information from surveys based on different sampling designs can provide models with good predictive performance

(Soranno et al. 2020). Furthermore, when the number of observations used to fit the model is sufficiently large (we used observations from 25,333 tows), then projections for unsampled areas within the same time frame are considered interpolations (Elith and Leathwick 2009; Soranno et al. 2020). We consider our projections of HSIs in Mobjack Bay, Potomac River, and other subestuaries that were not routinely sampled by fishery-independent surveys to be interpolations and valid for assessment of habitat conditions in non-sampled areas during the timeframe of the study (2000-2016).

Projections of habitat suitability into the future and under novel climate conditions, however, can be more problematic (Elith et al. 2010). Indeed, models that are useful descriptors of current conditions and distributions of species may lead to poor predictions in other times (Elith et al. 2010). For example, climate change is expected to exert effects on the structure of fish communities in river-dominated estuaries through changes in temperature and salinity (Feyrer et al. 2015), and these changes may affect how fishes use estuarine habitats. The habitat suitability models that we developed may be used to assess the suitability of Chesapeake Bay habitats under various climate-change scenarios, but such applications will require additional research and model building. In particular, additional transferability assessments are recommended. More specifically, our modeling approach should be cross-validated by non-random assignment of observations to temporally distinct groups to ensure that the relationships between fish abundance and habitat conditions (e.g., temperature) are not extrapolated beyond the range of the data used to fit the model (Wenger and Olden 2012). Globally, the last five years have been four of the warmest years on record (Arguez et al. 2020) and as such, may provide a reasonable indication of the relationship between forage fish abundance and habitat conditions in Chesapeake Bay in the future. We considered habitat conditions and fish habitat relationships up to and including 2016 and our current model would require updating with more contemporary information. Habitat conditions in the last 5 years, however, may not foretell conditions in the future. Hinson et al. (*in review*) report that the Chesapeake Bay warmed by about 0.7° C since 1985, or about 0.24°C per decade; this rate of warming is similar to the predicted rate of warming of sea surface temperature in the North Atlantic by 2050 under a 'reduced emissions' scenario (Hinson et al. *in review*). The Chesapeake Bay is experiencing this rate today. Moreover, surface- and bottom-water warming in Chesapeake Bay is primarily driven by warming air temperatures in the region. Thus, it is reasonable to assume that increasing air temperatures associated with global warming will continue to affect warming of the Chesapeake Bay. On a global scale, future years have a strong likelihood of remaining near record levels (i.e., setting new records; Arguez et al. 2020), and we therefore recommend that habitat suitability projections for Chesapeake Bay be constrained within a relatively short time frame (5-10 years) to ensure projections do not extrapolate beyond the maximum observed temperature conditions. The reliability of those projections can be assessed using metrics such as the generalized independent variable hull (gIVH, Conn et al. 2015) or multivariate environmental similarity surfaces (MESS, Elith et al. 2010). The gIVH extends Cook's concept of influence from a linear regression framework to complex models that incorporate multiple covariates to describe abundance in space and time. MESS provides a means to index the similarity of an observation to the reference data used to build the model (Elith et al. 2010). Formalized approaches such as gIVH and MESS are necessary to evaluate whether projections are extrapolations or interpolations (Elith et al. 2010; Conn et al. 2015), and thus, to guide inferences.

Finally, we note that the uncertainties associated with habitat suitability modeling and resulting projections are rarely assessed (Elith and Leathwick 2009). We are aware of only one study that explored uncertainty of HSI model projections (e.g., Zajac et al. 2015), but such uncertainties are useful for understanding the limitations of model-based results for conservation and fisheries management.

Uncertainties may arise from model specification (e.g., the type of model used, or covariates omitted from the model) and from the observations used to fit the model (e.g., samples may not represent the population of interest, or sample size may be inadequate). Ensemble approaches have been used to partially mitigate model uncertainty, but ensemble models do not fully overcome the limitations of the individual component models (Elith et al. 2010). Model and observational uncertainties may affect habitat suitability projections in different ways, and the analysis of uncertainty for HSI models warrants further research (Zajac et al. 2015) and engagement with environmental statisticians. In particular, a flexible hierarchical modeling framework recently described by Hefley and Hooten (2016) and which incorporates point processes to model presence/absence and count data appears to be a promising unifying approach for modeling the distribution of species across time and space.

## Acknowledgments

We thank Jian Shen (Virginia Institute of Marine Science, VIMS) and Ben Marcek (Ohio State University) for extending the numerical model that allowed us to consider bottom DO conditions. We are grateful to Jack Buchanan (VIMS) and Vaskar Nepal (VIMS) for application of the DO model and its validation. Rachel Dixon (VIMS) and Shannon Smith (VIMS) assisted in the verification of our HSI modeling approach. We thank Chris Walstrum (MD DNR) for providing fisheries data from the Maryland trawl survey. We appreciate the steadfast sampling in Virginia tidal waters by Captains Hank Brooks, Voight 'Bubba' Hogge, and Wendy Lowery, as well as the scientific crew of the VIMS Juvenile Fish Trawl Survey that served between 2000 and 2016 – Reid Broadwell, Aimee Comer, Christopher Davis, Julia Ellis, Deane Estes (deceased), Courtney Ford, Brian Gallagher, Pat Geer, Paul Gerdes, Jennifer Greaney, Chris Hager, Rebecca Hailey, James Harrison, Amanda Hewitt, Emily Loose, Leonard Machut, Marcel Montane, Taylor Moore, Katherine Nickerson, Ryan Norris, Ashleigh Rhea, Jillian Swinford, Steve Thornton, and Anya Voloshin. Graduate students in the Fabrizio lab – Mark Henderson, Karen (Capossela) Horodysky, Ben Marcek, Vaskar Nepal, Lauren Nys, Olivia Phillips, Cemil Saglam (Ege University, Izmir, Turkey), Ryan Schloesser, Branson Williams, and Justine Woodward – also participated in these surveys. Finally, we acknowledge funding for this research from the NOAA Chesapeake Bay Office.

# Literature Cited


Arbeider, M., C. Sharpe, C. Carr-Harris, and J. W. Moore. 2019. Integrating prey dynamics, diet, and biophysical factors across an estuary seascape for four fish species. Marine Ecology Progress Series 613: 151–169.

Arguez, A., S. Hurley, A. Inamdar, L. Mahoney, A. Sanchez-Lugo, and L. Yang. 2020. Should we expect each year in the next decade (2019–28) to be ranked among the top 10 warmest years globally? Bulletin of the American Meteorological Society 101: E655–E663 https://doi.org/10.1175/BAMS-D-19-0215.1.

Atlantic States Marine Fisheries Commission (ASMFC). 2019. Weakfish Stock Assessment Update Report. http://www.asmfc.org/uploads/file/5de7fc7c2019WeakfishAssessmentUpdate.pdf.

Bever, A. J., and M. L. MacWilliams. 2013. Simulating sediment transport processes in San Pablo Bay using coupled hydrodynamic, wave, and sediment transport models. Marine Geology 345: 235–253.

Bever, A. J., M. L. MacWilliams, B. Herbold, L. R. Brown, and F. V. Feyrer. 2016. Linking hydrodynamic complexity to Delta smelt (*Hypomesus transpacificus*) distribution in San Francisco estuary, USA. San Francisco Estuary & Watershed Science 14(1). http://escholarship.org/uc/item/2x91q0fr.

Birkmanis, C. A., J. J. Freer, L. W. Simmons, J. C. Partridge, and A. M. M. Sequeira. 2020. Future distribution of suitable habitat for pelagic sharks in Australia under climate change models. Frontiers in Marine Science 7: 570. doi: 10.3389/fmars.2020.00570.

Blaber, S. J. M., and T. J. Blaber. 1980. Factors affecting the distribution of juvenile estuarine and inshore fish. Journal of Fish Biology 17: 143-162.

Brady, D. C., and T. E. Targett. 2013. Movement of juvenile weakfish *Cynoscion regalis* and spot *Leiostomus xanthurus* in relation to diel-cycling hypoxia in an estuarine tidal tributary. Marine Ecology Progress Series 491: 199-219.

Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone. 1984. Classification and regression trees. Chapman and Hall, NY.

Brown, L. R., W. A. Bennett, R. W. Wagner, T. Morgan-King, N. Knowles, F. Feyrer, D. H. Schoellhamer, M. T. Stacey, and M. Dettinger. 2013. Implications for future survival of Delta smelt from four climate change scenarios for the Sacramento-San Joaquin Delta, California. Estuaries and Coasts 36: 754-774.

Brown, S. K., K. R. Buja, S. H. Jury, and M. E. Monaco. 2000. Habitat suitability index models for eight fish and invertebrate species in Casco and Sheepscot Bays, Maine. North American Journal of Fisheries Management 20: 408-435.

Buchheister, A., and R. J. Latour. 2011. Trophic ecology of Summer Flounder in lower Chesapeake Bay inferred from stomach content and stable isotope analyses. Transactions of the American Fisheries Society 140: 1240-1254.

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.

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.

Byrne, R. J., C. H. Hobbs, III, and M. J. Carron. 1983. Baseline sediment studies to determine distribution, physical properties, sedimentation budgets, and rates in the Virginia portion of the Chesapeake Bay. Final Report to the US EPA Chesapeake Bay Program. Gloucester Point, Virginia: Virginia Institute of Marine Science.

Camaclang, A. E., M. Maron, T. G. Martin, and H. P. Possingham. 2015. Current practices in the identification of critical habitat for threatened species. Conservation Biology 29: 482-492.

Cameron, M. J., V. Lucieer, N. S. Barrett, C. R. Johnson, and G. J. Edgar. 2014. Understanding community-habitat associations of temperate reef fishes using fine-resolution bathymetric measures of physical structure. Marine Ecology Progress Series 506: 213-229.

Capossela, K. M., R. W. Brill, M. C. Fabrizio, and P. G. Bushnell. 2012. Metabolic and cardiorespiratory responses of summer flounder *Paralichthys dentatus* to hypoxia at two temperatures. Journal of Fish Biology 81: 1043–1058.

Casulli, V. and P. Zanolli. 2002. Semi-implicit numerical modelling of non-hydrostatic free-surface flows for environmental problems. Mathematical and Computer Modelling 36: 1131-1149.

Casulli, V. and P. Zanolli. 2005. High-resolution methods for multidimensional advection-diffusion problems in free-surface hydrodynamics. Ocean Modelling 10 (1-2): 137-151.

CBP (Chesapeake Bay Program). 2015. Forage Fish Outcome Management Strategy 2015-2025, V1. https://www.chesapeakebay.net/documents/22031/1d_forage_ms_6-24-15_ff_formatted.pdf.

CBP (Chesapeake Bay Program). 2018. Forage Fish Outcome Management Strategy 2015-2025, V2. https://www.chesapeakebay.net/documents/22031/2018-2019_forage_management_strategy.pdf.

Champion, C., A. J. Hobday, G. T. Pecl, and S. R. Tracey. 2019. Oceanographic habitat suitability is positively correlated with the body condition of a coastal-pelagic fish. Fisheries Oceanography 29: 100-110.

Chang, Y.-J., C.-L. Sun, Y. Chen, S.-Z. Yeh, and G. Dinardo. 2012. Habitat suitability analysis and identification of potential fishing grounds for swordfish, *Xiphias gladius*, in the South Atlantic Ocean. International Journal of Remote Sensing 33: 7523-7541.

Cheng, R. T., and V. Casulli. 2002. Evaluation of the UnTRIM model for 3-D Tidal Circulation. Proceedings of the 7th International Conference on Estuarine and Coastal Modeling, St. Petersburg, Florida. November 2001; pp. 628–642.

Conn, P. B. 2010. Hierarchical analysis of multiple noisy abundance indices. Canadian Journal of Fisheries and Aquatic Sciences 67: 108-120.

Conn, P. B., D. S. Johnson, and P. L. Boveng. 2015. On extrapolating past the range of observed data when making statistical predictions in ecology. PLoS ONE 10(10): e0141416. doi:10.1371/journal.pone.0141416.

Craig, J. K. and L. B. Crowder. 2005. Hypoxia-induced habitat shifts and energetic consequences in Atlantic croaker and brown shrimp on the Gulf of Mexico shelf. Marine Ecology Progress Series 294: 79 – 94.

Crear, D. P., R. J. Latour, M. A. M. Friedrichs, P. St.-Laurent, and K. C. Weng. 2020. Sensitivity of a shark nursery habitat to a changing climate. Marine Ecology Progress Series 652: 123–136.

Day, Jr., J. W., C. A. S. Hall, W. M. Kemp, and A. Yáñez-Arancibia. 1989. Estuarine ecology. John Wiley & Sons, New York.

Du, J. and J. Shen. 2014. Decoupling the influence of biological and physical processes on the dissolved oxygen in the Chesapeake Bay. Journal of Geophysical Research Oceans 120: 78–93.

Efron, B. and R. Tibshirani. 1986. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. Statistical Science 1(1): 54-77.

Elith, J. and J. R. Leathwick 2009. Species distribution models: ecological explanation and prediction across space and time. Annual Review of Ecology, Evolution, and Systematics 40: 677-697

Elith, J. and J. Leathwick. 2017. Boosted regression trees for ecological modeling. R vignette: https://cran.biodisk.org/web/packages/dismo/vignettes/brt.pdf

Elith, J., M. Kearney, and S. Phillips. 2010. The art of modelling range-shifting species. Methods in Ecology and Evolution 1: 330‑342.

Fabrizio, M. C., J. P. Manderson, and J. P. Pessutti. 2013. Habitat associations and dispersal of black sea bass from a mid-Atlantic Bight reef. Marine Ecology Progress Series 482: 241-253.

Feyrer, F., J. E. Cloern, L. R. Brown, M. A. Fish, K. A. Hieb, and R. D. Baxter. 2015. Estuarine fish communities respond to climate variability over both river and ocean basins. Global Change Biology 21: 3608-3619.

Fletcher, Jr., R. J., T. J. Hefley, E. P. Robertson, B. Zuckerberg, R. A. McCleery, and R. M. Dorazio. 2019. A practical guide for combining data to model species distributions. Ecology 100(6):e02710.

Forte, M. F., J. L. Hanson, L. Stillwell, M. Blanchard-Montgomery, B. Blanton, R. Leuttich, H. Roberts, J. Atkinson, and J. Miller. 2011. Coastal Storm Surge Analysis System Digital Elevation Model. United States Army Corps of Engineers, Engineering and Research Development Center. ERDC/CHL TR-11-1. March 2011.

França, S., M. J. Costa, and H. N. Cabral. 2009. Assessing habitat specific fish assemblages in estuaries along the Portuguese coast. Estuarine, Coastal and Shelf Science 83: 1–12.

Fujiwara, M., C. Zhou, C. Acres, F. Martinez-Andrade. 2016. Interaction between penaeid shrimp and fish populations in the Gulf of Mexico: Importance of shrimp as forage species. PLoS ONE 11(11): e0166479. doi:10.1371/journal.pone.0166479

Fulford, R. S., M. S. Peterson, W. Wu, and P. O. Grammer. 2014. An ecological model of the habitat mosaic in estuarine nursery areas: Part II – Projecting effects of sea level rise on fish production. Ecological Modelling 273: 96-108.

Georgian, S. E., O. F. Anderson, and A. A. Rowden. 2019. Ensemble habitat suitability modeling of vulnerable marine ecosystem indicator taxa to inform deep-sea fisheries management in the South Pacific Ocean. Fisheries Research 211: 256-274.

Gillson, J. 2011. Freshwater flow and fisheries production in estuarine and coastal systems: where a drop of rain is not lost. Reviews in Fisheries Science 19: 168-186.

Glaspie, C. N., M. Clouse, K. Huebert, S. A. Ludsin, D. M. Mason, J. J. Pierson, M. R. Roman, and S. B. Brandt. 2019. Fish diet shifts associated with the northern Gulf of Mexico hypoxic zone. Estuaries and Coasts 42: 2170–2183.

Guan, L., Y. Chen, and J. A. Wilson. 2016. Evaluating spatio-temporal variability in the habitat quality of Atlantic cod (*Gadus morhua*) in the Gulf of Maine. Fisheries Oceanography 26: 83-96.

Harford, W. J., S. G. Smith, J. S. Ault, and E. A. Babcock. 2016. Cross-shelf habitat occupancy probabilities for juvenile groupers in the Florida Keys coral reef ecosystem. Marine and Coastal Fisheries 8: 147-159.

Hatin, D., J. Munro, F. Caron, and R. D. Simons. 2007. Movements, home range size, and habitat use and selection of early juvenile Atlantic sturgeon in the St. Lawrence estuarine transition zone. In: J. Munro, ed., Anadromous sturgeons: habitats, threats, and management. American Fisheries Society, Bethesda, MD.

Hefley, T. J., and M. B. Hooten. 2016. Hierarchical species distribution models. Current Landscape Ecology Reports 1: 87–97.

Hess, G. R., and J. M. Bay. 2000. A regional assessment of windbreak habitat suitability. Environmental Monitoring and Assessment 61: 237–254.

Hinson, K. E., M. A. M. Friedrichs, P. St-Laurent, F. Da, and R. G. Najjar. *In Review*. A thirty-year analysis of Chesapeake Bay warming. Journal of American Water Resources Association.

Hofmann, E. E., B. Cahill, K. Fennel, M. A. M. Friedrichs, K. Hyde, C. Lee, A. Mannino, R. G. Najjar, J. E. O'Reilly, J. Wilkin, and J. Xue. 2011. Modeling the dynamics of continental shelf carbon. Annual Review of Marine Science 3: 93–122.

Ihde, T.F., E.D. Houde, C.F. Bonzek, and E. Franke. 2015. Assessing the Chesapeake Bay forage base: existing data and research priorities. STAC Publication Number 15-005. http://www.chesapeake.org/pubs/346_Ihde2015.pdf

Irby, I. D., M. A. M. Friedrichs, C. T. Friedrichs, A. J. Bever, R. R. Hood, L. W. J. Lanerolle, M. Li, L. Linker, M. E. Scully, K. Sellner, J. Shen, J. Testa, H. Wang, P. Wang, and M. Xin. 2016. Challenges associated with modeling low-oxygen waters in Chesapeake Bay: a multiple model comparison. Biogeosciences 13: 2011–2028.

Jacobs, J. M., M. R. Rhodes, A. Baya, R. Reimschuessel, H. Townsend, and R. M. Harrell. 2009. Influence of nutritional state on the progression and severity of mycobacteriosis in striped bass *Morone saxatilis.* Diseases of Aquatic Organisms 87: 183–197.

Jenkins, G. P., D. Spooner, S. Conron, and J. R. Morrongiello. 2015. Differing importance of salinity stratification and freshwater flow for the recruitment of apex species of estuarine fish. Marine Ecology Progress Series 523: 125-144.

Jolliff, J. K., J. C. Kindle, I. Shulman, B. Penta, M. A. M. Friedrichs, R. Helber, and R. A. Arnone. 2009. Summary diagrams for coupled hydrodynamic-ecosystem model skill assessment. Journal of Marine Systems 76: 64–82.

Kerhin, R. T., J. P. Halka, D. V. Wells, E. L. Hennessee, P. J. Blakeslee, N. Zoltan, and R. H. Cuthbertson. 1988. Chesapeake Bay Earth Science Study (CBESS): Physical Properties of Surficial Sediments, Chesapeake Bay, Maryland (Tabular Data). Baltimore, Maryland: Maryland Geological Survey. http://www.mgs.md.gov/coastal_geology/baysedata.html.

Kloke, J. D., and J. W. McKean. 2012. Rfit: Rank-based estimation for linear models. The R Journal 4: 57-64.

Knudby, A., A. Brenning, and E. LeDrew. 2010. New approaches to modelling fish-habitat relationships. Ecological Modelling 221: 503-511.

Kostecki, C., F. Le Loc'h, J.-M. Roussel, N. Desroy, D. Huteau, P. Riera, H. Le Bris, and O. Le Pape. 2010. Dynamics of an estuarine nursery ground: the spatio-temporal relationship between the river flow and the food web of the juvenile common sole as revealed by stable isotope analysis. Journal of Sea Research 64: 54-60.

Krause, J. R., J. E. Hightower, J. A. Buckel, J. T. Turnure, T. M. Grothues, J. P. Manderson, J. E. Rosendale, and J. P. Pessutti. 2020. Using acoustic telemetry to estimate weakfish survival rates along the U.S. east coast. Marine and Coastal Fisheries 12: 241-257.

Kritzer, J. P., M. Delucia, E. Greene, C. Shumway, M. F. Topolski, J. Thomas-Blate, L. A. Chiarella, K. B. Davy, and K. Smith. 2016. The importance of benthic habitats for coastal fisheries. BioScience 66: 274-284.

Kupschus, S. 2003. Development and evaluation of statistical habitat suitability models: an example based on juvenile spotted seatrout *Cynoscion nebulosus*. Marine Ecology Progress Series 265: 197-212.

Lankowicz, K. M., H. Bi, D. Liang, and C. Fan. 2020. Sonar imaging surveys fill data gaps in forage fish populations in shallow estuarine tributaries. Fisheries Research 226: 105520. https://doi.org/10.1016/j.fishres.2020.105520.

Lapointe, D., W. K. Vogelbein, M. C. Fabrizio, D. T. Gauthier, and R. W. Brill. 2014. Temperature, hypoxia, and mycobacteriosis: effects on adult striped bass *Morone saxatilis* metabolic performance. Diseases of Aquatic Organisms 108: 113‑127.

Large, W., and S. Pond. 1981. Open ocean momentum flux measurements in moderate to strong winds. Journal of Physical Oceanography 11: 324–336.

Latour, R. J., J. Gartland, C. F. Bonzek, and R. Johnson. 2008. The trophic dynamics of summer flounder (*Paralichthys dentatus*) in Chesapeake Bay. Fishery Bulletin 106: 47–57.

Latour, R. J., J. Gartland, and C. F. Bonzek. 2017. Spatiotemporal trends and drivers of fish condition in Chesapeake Bay. Marine Ecology Progress Series 579: 1-17.

Lauver, C. L., W. H. Busby, and J. L. Whistler. 2002. Testing a GIS model of habitat suitability for a declining grassland bird. Environmental Management 30: 88–97.

Layher, W. G., and O. E. Maughan. 1985. Spotted bass habitat evaluation using an unweighted geometric mean to determine HSI values. Proceedings of the Oklahoma Academy of Science 65: 11–17.

Lecours, V., R. Devillers, D. C. Schneider, V. L. Lucieer, C. J. Brown, and E. N. Edinger. 2015. Spatial scale and geographic context in benthic habitat mapping: review and future directions. Marine Ecology Progress Series 535: 259-284.

Le Pape, O., F. Chauvet, S. Mahevas, P. Lazure, D. Guerault, and Y. Desaunay. 2003. Quantitative description of habitat suitability for the juvenile common sole in the Bay of Biscay (France) and the contribution of different habitats to the adult population. Journal of Sea Research 50: 139-149.

Little, A. G., I. Loughland, and F. Seebacher. 2020. What do warming waters mean for fish physiology and fisheries? Journal of Fish Biology 97: 328–340.

MacWilliams, M. L., A. J. Bever, and E. Foresman. 2016. 3-D simulations of the San Francisco estuary with subgrid bathymetry to explore long-term trends in salinity distribution and fish abundance. San Francisco Estuary & Watershed Science 14(2). http://escholarship.org/uc/item/5qj0k0m6.

MacWilliams, M. L., A. J. Bever, E. S. Gross, G. A. Ketefian, and W. J. Kimmerer. 2015. Three-dimensional modeling of hydrodynamics and salinity in the San Francisco Estuary: an evaluation of model accuracy, X2, and the low salinity zone. San Francisco Estuary and Watershed Science 13(1). http://dx.doi.org/10.15447/sfews.2015v13iss1art2.

Manderson, J. P., J. Pessutti, J. G. Hilbert, and F. Juanes. 2004. Shallow water predation risk for a juvenile flatfish (winter flounder; *Pseudopleuronectes americanus*, Walbaum) in a northwest Atlantic estuary. Journal of Experimental Marine Biology and Ecology 304: 137-157.

Manderson, J. P., L. Palamara, J. Kohut, and M. J. Oliver. 2011. Ocean observatory data are useful for regional habitat modeling of species with different vertical habitat preferences. Marine Ecology Progress Series 438: 1-17.

Marcek, B. J., R. W. Brill, and M. C. Fabrizio. 2019. Metabolic scope and hypoxia tolerance of Atlantic croaker (*Micropogonias undulatus* Linnaeus, 1766) and spot (*Leiostomus xanthurus* Lacepède, 1802), with insights into the effects of acute temperature change. Journal of Experimental Marine Biology and Ecology 516: 150–158.

Maryland Geological Survey. 1996. Baltimore Harbor Surficial Sediments: Texture and Chemistry Baltimore, Maryland (Tabular Data). Maryland Department of Natural Resources/Maryland Geological Survey. http://www.mgs.md.gov/publications/data_pages/baltoharbdata.html.

Moncure, R., and M. Nichols. 1968. Characteristics of sediments in the James River Estuary, Virginia. Special Scientific Report No. 53. Virginia Institute of Marine Science. www.vims.edu/GreyLit/VIMS/ssr053.pdf.

Mather, M. E., J. T. Finn, K. H. Ferry, L. A. Deegan, and G. A. Nelson. 2009. Use of non-natal estuaries by migratory striped bass (*Morone saxatilis*) in summer. Fishery Bulletin 107:329–338.

Minello, T. J., and R. J. Zimmerman. 1983. Fish predation on juvenile brown shrimp, *Penaeus aztecus* Ives: the effect of simulated *Spartina* structure on predation rates. Journal of Experimental Marine Biology and Ecology 72: 211-231.

Minello, T. J., R. J. Zimmerman, and E. X. Martinez. 1989. Mortality of young brown shrimp *Penaeus aztecus* in estuarine nurseries. Transactions of the American Fisheries Society 118: 693-708.

Murdy, E. O., R. S. Birdsong and J. A. Musick. 1997. Fishes of Chesapeake Bay. Smithsonian Institution Press.

Parsons, D. M., C. Middleton, K. T. Spong, G. Mackay, M. D. Smith, and D. Buckthought. 2015. Mechanisms explaining nursery habitat association: how do juvenile snapper benefit from their nursery habitat? PLoS ONE AR e0122137.

Peebles, E. B., S. E. Burghart, and D. J. Hollander. 2007. Causes of interestuarine variability in bay anchovy (*Anchoa mitchilli*) salinity at capture. Estuaries and Coasts 30: 1060-1074.

Pennino, M. G., M. Coll, M. Albo-Puigserver, E. Fernández-Corredor, J. Steenbeek, A. Giráldez, M. González, A. Esteban, and J. M. Bellido. 2020. Current and future influence of environmental factors on small pelagic fish distributions in the northwestern Mediterranean Sea. Frontiers in Marine Science 7: article 622. doi: 10.3389/fmars.2020.00622.

Powers, S. P., and J. N. Kittinger. 2002. Hydrodynamic mediation of predator-prey interactions: differential patterns of prey susceptibility and predator success explained by variation in water flow. Journal of Experimental Marine Biology and Ecology 273: 171-187.

R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Reid, J. M., J. A. Reid, C. J. Jenkins, M. E. Hastings, S. J. Williams, and L. J. Poppe. 2005. usSEABED: Atlantic coast offshore surficial sediment data release: U.S. Geological Survey Data Series 118, version 1.0. http://pubs.usgs.gov/ds/2005/118/.

Reuchlin-Hugenholtz, E., N. L. Shackell, and J. A. Hutchings. 2016. Spatial reference points for groundfish. ICES Journal of Marine Science 73: 2468-2478.

Rubec, P. J., C. Santi, Y. Ghile, and X. Chen. 2019. Modeling and mapping to assess spatial distributions and population numbers of fish and invertebrate species in the lower Peace River and Charlotte Harbor, Florida. Marine and Coastal Fisheries 11: 328–350.

Rogers, N. J., M. A. Urbina, E. E. Reardon, D. J. McKenzie, and R. W. Wilson. 2016. A new analysis of hypoxia tolerance in fishes using a database of critical oxygen level ($P_{crit}$). Conservation Physiology 4 (1). https://doi.org/10.1093/conphys/cow012.

Ruiz, G. M., A. H. Hines, and M. H. Posey. 1993. Shallow water as a refuge habitat for fish and crustaceans in non-vegetated estuaries: an example from Chesapeake Bay. Marine Ecology Progress Series 99: 1-16.

Scales, K. L., E. L. Hazen, M. G. Jacox, C. A. Edwards, A. M. Boustany, M. J. Oliver, and S. J. Bograd. 2017. Scale of inference: on the sensitivity of habitat models for wide ranging marine predators to the resolution of environmental data. Ecography 40: 210–220.

Schielzeth, H. 2010. Simple means to improve the interpretability of regression coefficients. Methods in Ecology and Evolution 1: 103–113.

Schlenger, A. J., E. W. North, Z. Schlag, Y. Li, D. H. Secor, K. A. Smith, and E. J. Niklitschek. 2013. Modeling the influence of hypoxia on the potential habitat of Atlantic sturgeon: a comparison of two methods. Marine Ecology Progress Series 483: 257-272.

Scully, M. 2013. Physical controls on hypoxia in Chesapeake Bay: A numerical modeling study. Journal of Geophysical Research: Oceans 118: 1239–1256.

Shen, J., H. Wang, M. Sisson, and W. Gong. 2006. Storm tide simulation in the Chesapeake Bay using an unstructured grid model. Estuarine, Coastal and Shelf Science 68: 1–16.

Sisson, M., H. Wang, Y. Li, J. Shen, A. Kuo, W. Gong, M. Brush, and K. Moore. 2010. Development of Hydrodynamic and Water Quality Models for the Lynnhaven River System. Final Report to the U.S. Army Corps of Engineers and City of Virginia Beach. Special Report No. 408. November 2010. http://www.vims.edu/GreyLit/VIMS/sramsoe408.pdf.

Soranno, P. A., K. S. Cheruvelil, B. Liu, Q. Wang, P.-N. Tan, J. Zhou, K. B. S. King, I. M. McCullough, J. Stachelek, M. Bartley, C. T. Filstrup, E. M. Hanks, J.-F. Lapierre, N. R. Lottig, E. M. Schliep, T. Wagner, and K. E. Webster. 2020. Ecological prediction at macroscales using big data: Does sampling design matter? Ecological Applications 30(6): e02123. Doi: 10.1002/eap.2123.

Tableau, A., A. Brind'Amour, M. Woillez, and H. Le Bris. 2016. Influence of food availability on the spatial distribution of juvenile fish within soft sediment nursery habitats. Journal of Sea Research 111: 76-87.

Tanaka, K., and Y. Chen. 2015. Spatiotemporal variability of suitable habitat for American lobster in Long Island Sound. Journal of Shellfish Research 34: 531-543.

Theuerkauf, S. J., and R. N. Lipcius. 2016. Quantitative validation of a habitat suitability index for oyster restoration. Frontiers in Marine Science 3: 64. doi: 10.3389/fmars.2016.00064.

Tian, S., X. Chen, Y. Chen, L. Xu, and X. Dai. 2009. Evaluating habitat suitability indices derived from CPUE and fishing effort data for *Ommatrephes bratramii* in the northwestern Pacific Ocean. Fisheries Research 95: 181-188.

Tuckey, T. D. and M. C. Fabrizio. 2013. Influence of survey design on fish assemblages: implications from a study in Chesapeake Bay tributaries. Transactions of the American Fisheries Society 142:957-973.

Tuckey, T. D. and M. C. Fabrizio. 2020. Estimating relative juvenile abundance of ecologically important finfish in the Virginia portion of Chesapeake Bay. Final report submitted to Virginia Marine Resources Commission.

Tuckey, T. D., J. Swinford, M. C. Fabrizio, H. Small, and J. Shields. *In press.* Penaeid shrimp in Chesapeake Bay: population growth and black gill disease syndrome. Marine and Coastal Fisheries.

USACE (United States Army Corps of Engineers). 2016. Federal Navigation Channels: Controlling Depths Reports. http://www.nab.usace.army.mil/Missions/Civil-Works/Nav-Maps/.

Velinsky, D., T. L. Wade, C. E. Schlekat, B. L. McGee, and B. J. Presley. 1994. Tidal river sediments in the Washington DC area I. Distribution and sources of trace metals. Estuaries 17: 305–320.

Wang, H. V., J. D. Loftis, D. Forrest, W. Smith, and B. Stamey. 2015. Modeling storm surge and inundation in Washington, DC, during hurricane Isabel and the 1936 Potomac River great flood. Journal of Marine Science and Engineering 3: 607–629.

Wenger, S. J., and J. D. Olden. 2012. Assessing transferability of ecological models: an underappreciated aspect of statistical validation. Methods in Ecology and Evolution 3: 260–267.

Willmott, C. 1981. On the validation of models. Physical Geography 2: 184–194.

Windle, M. J. S., G. A. Rose, R. Devillers, and M.-J. Fortin. 2012. Spatio-temporal variations in invertebrate–cod–environment relationships on the Newfoundland–Labrador Shelf, 1995–2009. Marine Ecology Progress Series 469: 263–278.

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

Xu, J., W. Long, J. D. Wiggert, L. W. J. Lanerolle, C. W. Brown, R. Murtugudde, and R. R. Hood. 2012. Climate forcing and salinity variability in the Chesapeake Bay, USA. Estuaries and Coasts 35: 237–261.

Yu, H., A. R. Cooper, and D. M. Infante. 2020. Improving species distribution model predictive accuracy using species abundance: Application with boosted regression trees. Ecological Modelling 432: 109202. https://doi.org/10.1016/j.ecolmodel.2020.109202.

Zajac, Z., B. Stith, A. C. Bowling, C. A. Lantimm, and E. D. Swain. 2015. Evaluation of habitat suitability index models by global sensitivity and uncertainty analyses: a case study for submerged aquatic vegetation. Ecology and Evolution 5: 2503–2517.

Zeng, X., K. R. Tanaka, Y. Chen, K. Wang, and S. Zhang. 2018. Gillnet data enhance performance of rockfishes habitat suitability index model derived from bottom-trawl survey data: A case study with *Sebasticus marmoratus*. Fisheries Research 204: 189-196.

Zhang, H., S. A. Ludsin, D. M. Mason, A. T. Adamack, S. B. Brandt, X. Zhang, D. G. Kimmel, M. R. Roman, W. C. Boicourt. 2009. Hypoxia-driven changes in the behavior and spatial distribution of pelagic fish and mesozooplankton in the northern Gulf of Mexico. Journal of Experimental Marine Biology and Ecology 381 (Supplement): S80–S91.

Table 1. Number of sites sampled monthly by fishery surveys in Virginia and Maryland waters of Chesapeake Bay, March 2000 to November 2016, and total number of sites for each year. We omitted observations from January 2000, February 2000, and December 2016 in Virginia from consideration because these represented incomplete sampling in the three-month seasonal period (Dec-Jan-Feb). Blank values in the Maryland portion of the table indicate that no sampling was completed in those months and years.

| Year | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Virginia Survey** | | | | | | | | | | | | | |
| 2000 | | | 66 | 101 | 103 | 111 | 109 | 111 | 114 | 110 | 111 | 98 | 936 |
| 2001 | 30 | 30 | 30 | 66 | 110 | 111 | 111 | 111 | 114 | 111 | 110 | 94 | 1032 |
| 2002 | 66 | 90 | 66 | 90 | 96 | 96 | 96 | 97 | 95 | 96 | 96 | 96 | 1078 |
| 2003 | 66 | 95 | 66 | 96 | 96 | 111 | 111 | 111 | 78 | 144 | 111 | 105 | 1181 |
| 2004 | 57 | 114 | 66 | 105 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 105 | 1224 |
| 2005 | 66 | 105 | 66 | 105 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 90 | 1224 |
| 2006 | 66 | 105 | 66 | 105 | 110 | 111 | 111 | 111 | 111 | 111 | 78 | 105 | 1175 |
| 2007 | 66 | 105 | 66 | 105 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 105 | 1224 |
| 2008 | 66 | 105 | 66 | 105 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 105 | 1224 |
| 2009 | 66 | 105 | 66 | 105 | 111 | 111 | 111 | 111 | 111 | 111 | 110 | 105 | 1223 |
| 2010 | 66 | 105 | 66 | 105 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 105 | 1224 |
| 2011 | 66 | 105 | 66 | 105 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 105 | 1224 |
| 2012 | 66 | 105 | 66 | 105 | 111 | 111 | 111 | 111 | 111 | 111 | 92 | 105 | 1205 |
| 2013 | 66 | 105 | 66 | 105 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 105 | 1224 |
| 2014 | 66 | 97 | 74 | 105 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 105 | 1224 |
| 2015 | 66 | 77 | 94 | 105 | 111 | 111 | 110 | 110 | 111 | 111 | 111 | 105 | 1222 |
| 2016 | 66 | 103 | 66 | 105 | 111 | 111 | 111 | 110 | 111 | 111 | 111 | | 1221 |
| **Maryland Survey** | | | | | | | | | | | | | |
| 2000 | | | | | 37 | 36 | 37 | 37 | 36 | 37 | | | 220 |
| 2001 | | | | | 37 | 37 | 37 | 37 | 36 | 37 | | | 221 |
| 2002 | | | | | 36 | 37 | 37 | 37 | 37 | 37 | | | 221 |
| 2003 | | | | | 53 | 50 | 53 | 53 | 53 | 52 | | | 314 |
| 2004 | | | | | 53 | 53 | 53 | 53 | 53 | 53 | | | 318 |
| 2005 | | | | | 53 | 53 | 53 | 53 | 53 | 53 | | | 318 |
| 2006 | | | | | | 53 | 53 | 53 | 53 | 53 | | | 256 |
| 2007 | | | | | 45 | 53 | 53 | 53 | 53 | 53 | | | 310 |
| 2008 | | | | | 37 | 53 | 53 | 53 | 53 | 53 | | | 302 |
| 2009 | | | | | 37 | 53 | 53 | 52 | 53 | 53 | | | 301 |
| 2010 | | | | | 53 | 53 | 53 | 53 | 53 | 53 | | | 318 |
| 2011 | | | | | 53 | 53 | 53 | 53 | 50 | 53 | | | 315 |
| 2012 | | | | | 53 | 53 | 53 | 53 | 53 | 53 | | | 318 |
| 2013 | | | | | 53 | 53 | 53 | 53 | 53 | 53 | | | 318 |
| 2014 | | | | | 53 | 53 | 53 | 53 | 53 | 53 | | | 318 |
| 2015 | | | | | 53 | 53 | 53 | 53 | 53 | 53 | | | 318 |
| 2016 | | | | | 52 | 52 | 52 | 52 | 52 | 52 | | | 312 |

Table 2. List of the 24 static and dynamic habitat features considered for optimization of boosted regression trees (BRTs) for forage fishes in Chesapeake Bay, 2010 - 2010. With the exception of the six 'percent time' covariates, the same covariates were used to fit the BRTs to the 2000-2016 observations, allowing us to identify influential covariates for use in habitat suitability models.

| Type | Habitat covariate | Units |
|---|---|---|
| Static | Sediment composition (percent fine sediment) | % |
| Static | Distance to shore | m |
| Dynamic | Water depth | m |
| Dynamic | Bottom dissolved oxygen | mg $O_2$/l |
| Dynamic | Tidal-averaged depth-averaged salinity | psu |
| Dynamic | Tidal averaged surface salinity | psu |
| Dynamic | Tidal averaged bottom salinity | psu |
| Dynamic | Tidal averaged salinity stratification | psu |
| Dynamic | Tidal-averaged depth-averaged temperature | ° C |
| Dynamic | Tidal averaged bottom temperature | ° C |
| Dynamic | Tidal averaged surface temperature | ° C |
| Dynamic | Tidal averaged temperature stratification | ° C |
| Dynamic | Tidal averaged depth-averaged current speed | m/s |
| Dynamic | Maximum depth-averaged current speed | m/s |
| Dynamic | Tidal averaged surface current 1 m below surface | m/s |
| Dynamic | Tidal averaged bottom current, 1 m above bottom | m/s |
| Dynamic | Tidal averaged vertical stratification in current speed | m/s |
| Dynamic | Tidal averaged horizontal gradient in current speed | m/s |
| Dynamic | Percent time bottom waters < 10 C | % |
| Dynamic | Percent time bottom waters between 10 and 20 C | % |
| Dynamic | Percent time bottom waters > 20 C | % |
| Dynamic | Percent time bottom waters < 10 psu | % |
| Dynamic | Percent time bottom waters between 10 and 20 psu | % |
| Dynamic | Percent time bottom waters > 20 psu | % |

Table 3. Influential covariates (checked and shaded boxes) identified by boosted regression trees for bay anchovy, juvenile spot, juvenile spotted hake, and juvenile weakfish from Chesapeake Bay, 2000-2016. The number of influential habitat covariates is denoted in parentheses after the species' name.

| Habitat covariate | Juvenile spotted hake (4) | Juvenile weakfish (4) | Juvenile spot (5) | Bay anchovy-winter (6) | Bay anchovy-summer (7) |
|---|---|---|---|---|---|
| *Physical covariates* | | | | | |
| Distance to shore | | ✓ | ✓ | ✓ | ✓ |
| Percent fine sediment | | | | ✓ | ✓ |
| Water depth | ✓ | ✓ | ✓ | ✓ | ✓ |
| *Temperature covariates* | | | | | |
| Tidal average bottom temperature | ✓ | ✓ | | | ✓ |
| Tidal average temperature stratification | | | ✓ | | |
| *Salinity covariates* | | | | | |
| Tidal average surface salinity | | | | ✓ | ✓ |
| Tidal average salinity vertical stratification | ✓ | | | | ✓ |
| *Dissolved oxygen* | | | | | |
| Bottom dissolved oxygen | | | ✓ | ✓ | |
| *Current speed covariates* | | | | | |
| Tidal average current speed stratification | | ✓ | | | |
| Tidal average current speed horizontal gradient | | | ✓ | ✓ | ✓ |
| Maximum depth-averaged current speed | ✓ | | | | |

Table 4. Results from the nonparametric regression analyses for juvenile spotted hake, juvenile spot, juvenile weakfish, and bay anchovy from Chesapeake Bay, 2000- 2016. The model fit to the data was $Y_i = \beta_0 + \beta_1 X_{1i} + \varepsilon_i$ where $Y_i$ is the rank-transformed response, n is the sample size, $\beta_0$ is the overall average response (intercept), $\beta_1$ is the regression coefficient (slope), $X_{1i}$ is the value of the predictor for observation *i*, and $\varepsilon_i$ is the unexplained random error. The F-statistic is used to test the significance of the model. Extent of suitable habitat was calculated for each season as the sum of the areas throughout Chesapeake Bay with HSI $\geq$ 0.5.

| Rank-transformed response (Y) | Predictor ($X_1$) | n | F | P |
|---|---|---|---|---|
| *Juvenile spotted hake in Spring* | | | | |
| Relative abundance | Extent of suitable habitat | 17 | 0.01 | 0.91 |
| Extent of suitable habitat | Year | 17 | 0.21 | 0.65 |
| *Juvenile spot in Summer* | | | | |
| Relative abundance | Extent of suitable habitat | 17 | 4.57 | 0.05 |
| Extent of suitable habitat | Year | 17 | 0.01 | 0.93 |
| *Juvenile spot in Fall* | | | | |
| Relative abundance | Extent of suitable habitat | 17 | 0.01 | 0.93 |
| Extent of suitable habitat | Year | 17 | 1.11 | 0.31 |
| *Juvenile weakfish in Summer* | | | | |
| Relative abundance | Extent of suitable habitat | 17 | 2.75 | 0.12 |
| Extent of suitable habitat | Year | 17 | 10.39 | <0.01 |
| *Juvenile weakfish in Fall* | | | | |
| Relative abundance | Extent of suitable habitat | 17 | 0.12 | 0.73 |
| Extent of suitable habitat | Year | 17 | 8.68 | 0.01 |
| *Bay anchovy in Summer* | | | | |
| Relative abundance | Extent of suitable habitat | 17 | 0.06 | 0.80 |
| Extent of suitable habitat | Year | 17 | 24.37 | <0.01 |
| *Bay anchovy in Winter* | | | | |
| Relative abundance | Extent of suitable habitat | 16 | 19.98 | <0.01 |
| Extent of suitable habitat | Year | 16 | 0.17 | 0.69 |

Figure 1. Sample sites (filled circles) of two surveys used to assess relative abundance of forage fishes in Chesapeake Bay, 2000-2016. (A) Sites sampled by the VIMS Juvenile Fish Trawl Survey in a representative month; site selection is based on a random stratified survey design, thus, sites sampled in any given month vary randomly, but the number of sites sampled remains constant, with only minor changes due to weather or vessel complications. During 2000 to 2016, this survey sampled 1,224 sites/year. (B) Sites sampled by the Maryland Blue Crab Summer Trawl Survey; fixed sites are sampled monthly between May and October. The star symbols denote multiple sample sites in the region (for clarity, individual sites are not shown).

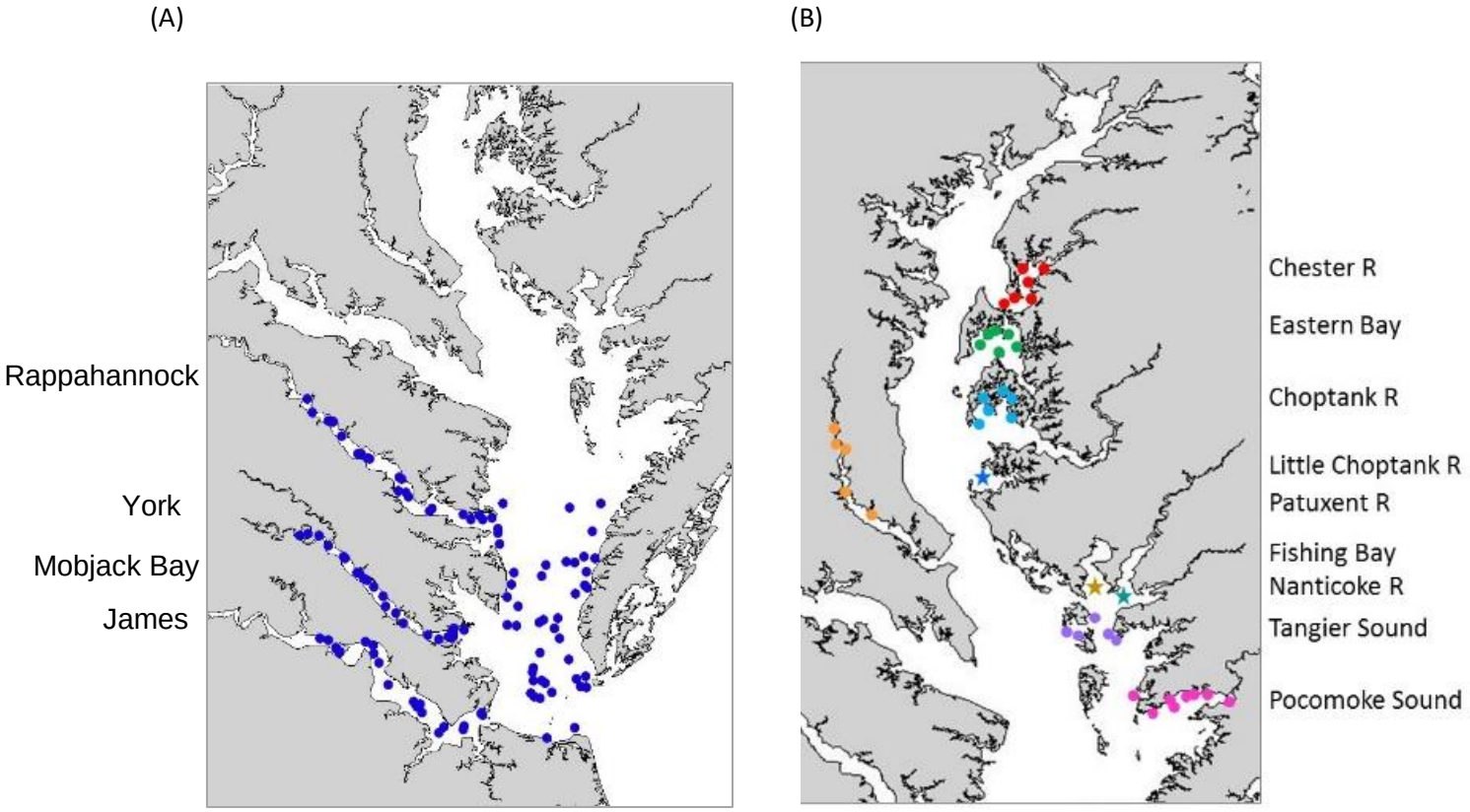


Figure 2. Model domain and boundary conditions for the 3-dimensional UnTRIM Chesapeake Bay model used to hindcast environmental conditions for forage fishes, 2000-2016.

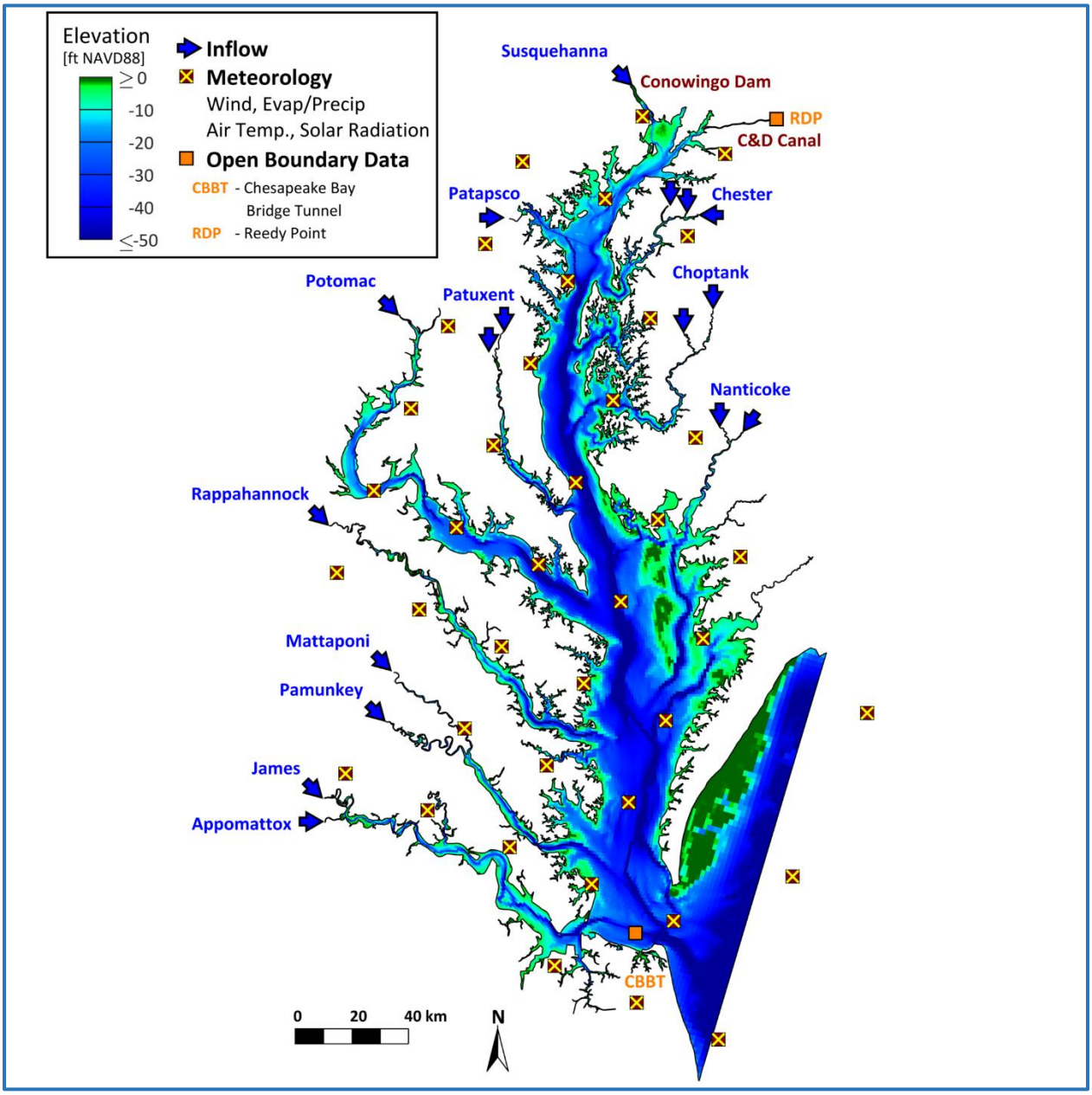


Figure 3. Distance to shoreline (km) and sediment composition (% fines) of the seabed in Chesapeake Bay.

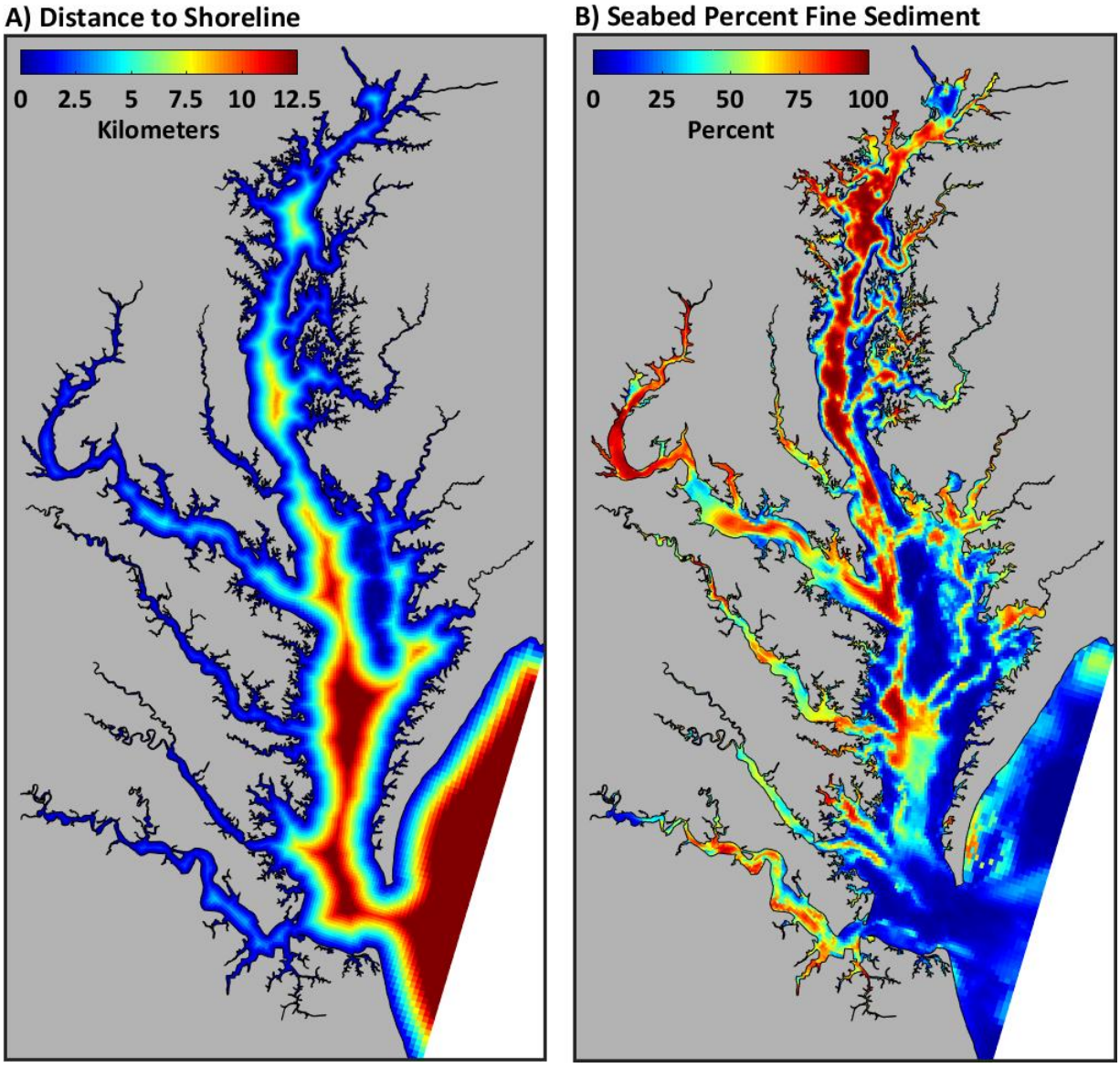


Figure 4. Locations used for validation of the 3-dimensional hydrodynamic model for Chesapeake Bay.

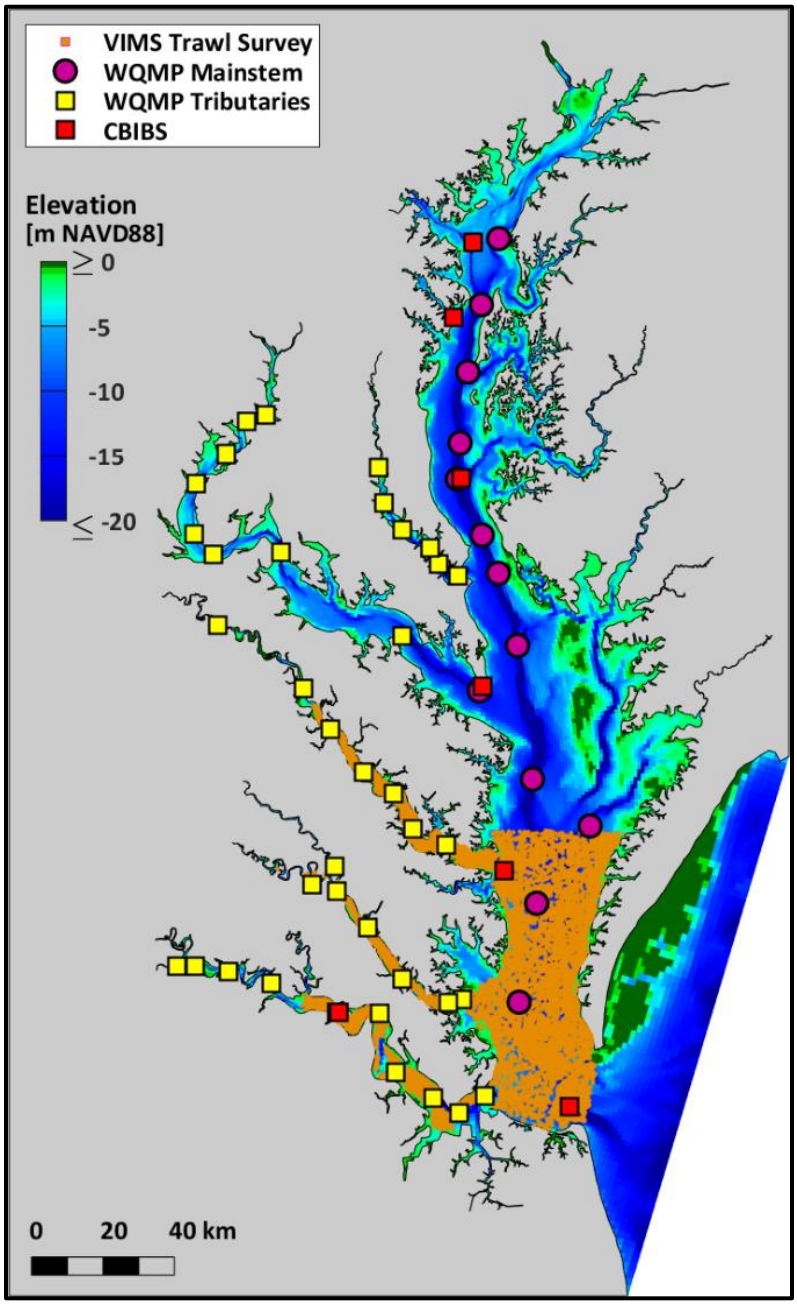


Figure 5. Validation of salinity and temperature hindcasts from the hydrodynamic model of Chesapeake Bay, 2000-2016. Scatter plots compare model estimates to observations recorded during the fisheries sampling for (A) bottom salinity and (B) bottom temperature during 2012; the black line is the 1 to 1 line. Target diagrams show each year (red) for (C) bottom salinity, (D) bottom temperature, (E) surface salinity, and (F) surface temperature. Text marks are the 2-digit year starting in 2000.

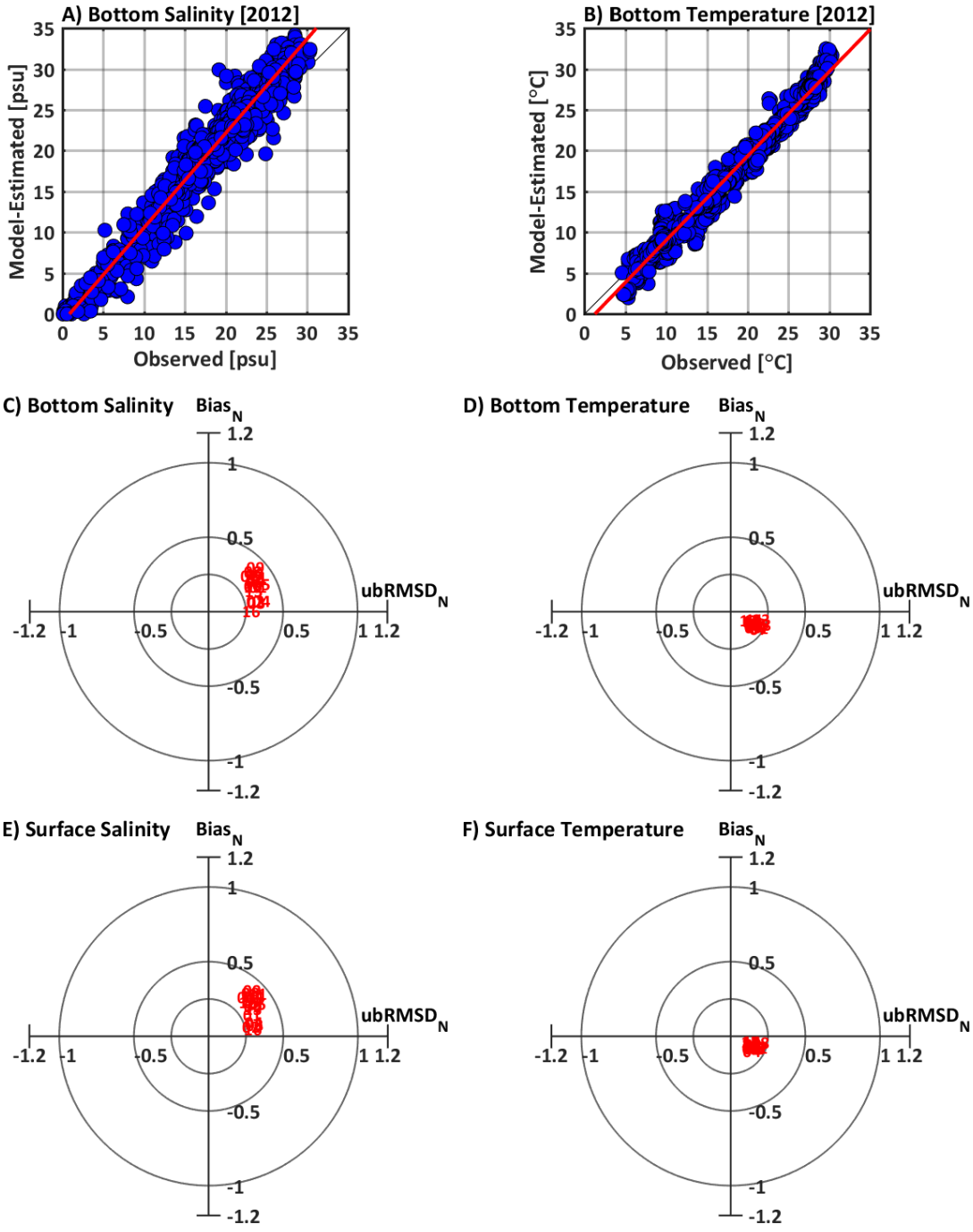
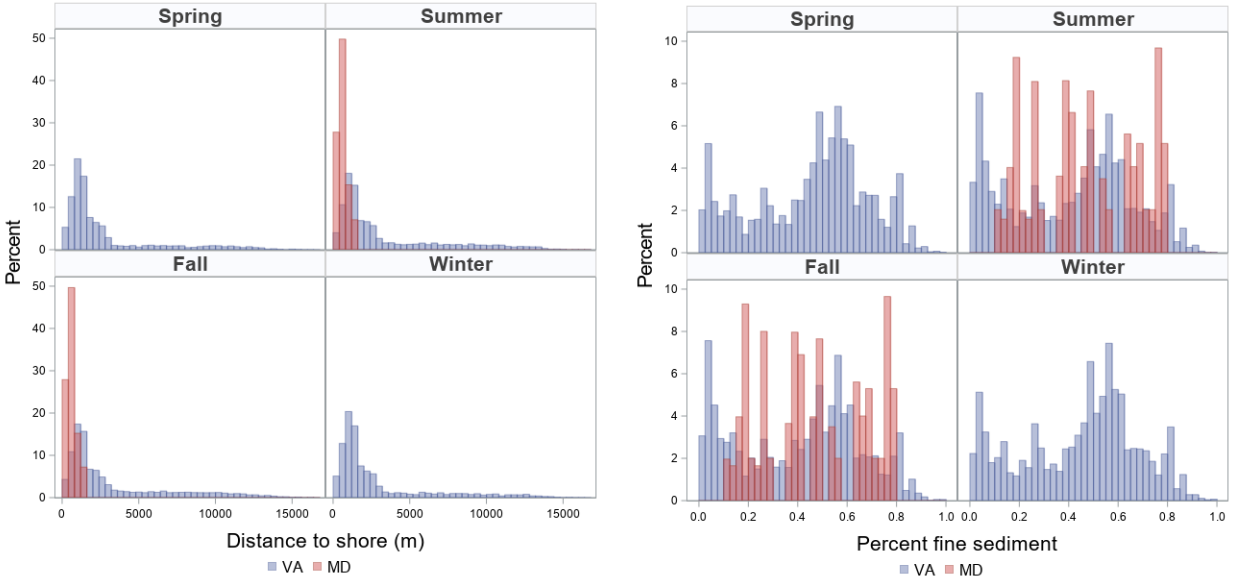
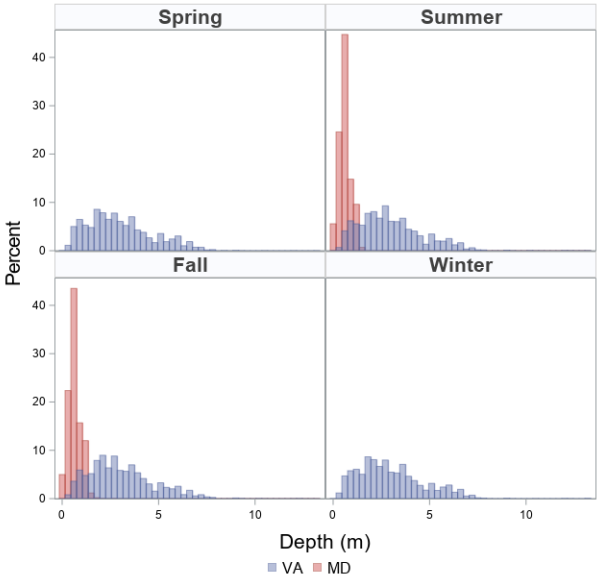
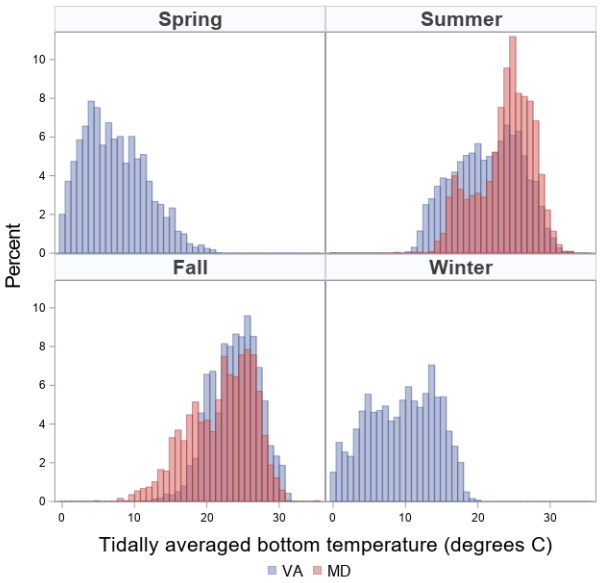


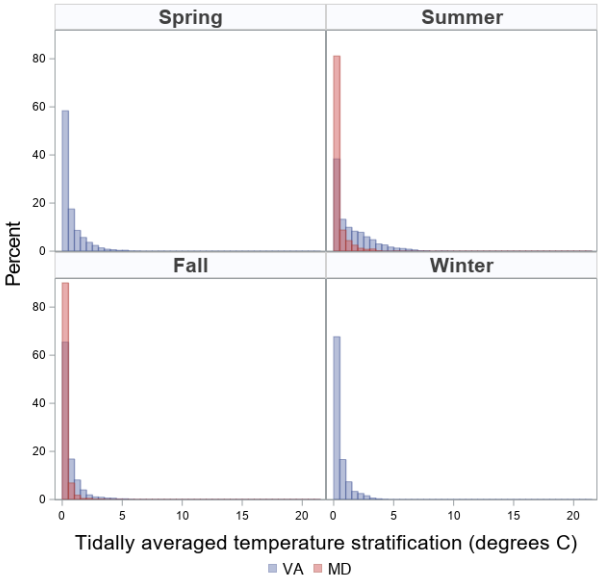
Figure 6. Frequency histograms of seasonal environmental and physical conditions at sites sampled by the Virginia and Maryland surveys in the Chesapeake Bay and its major tributaries, 2000 – 2016. The 11 covariates depicted are those identified as influential covariates for juvenile spotted hake, juvenile weakfish, juvenile spot, and bay anchovy.

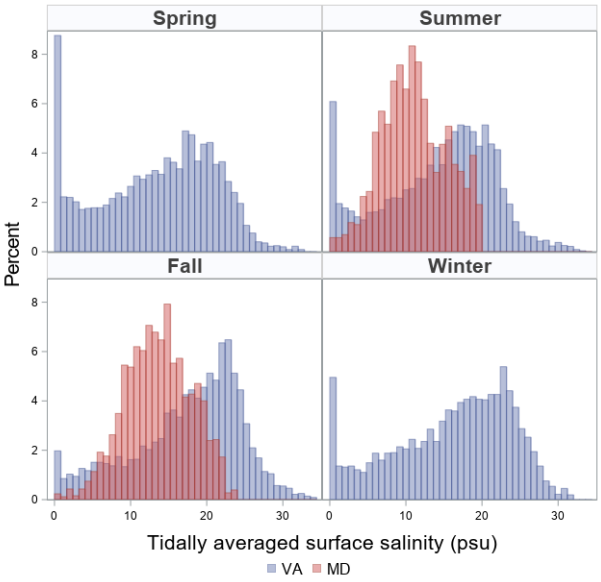


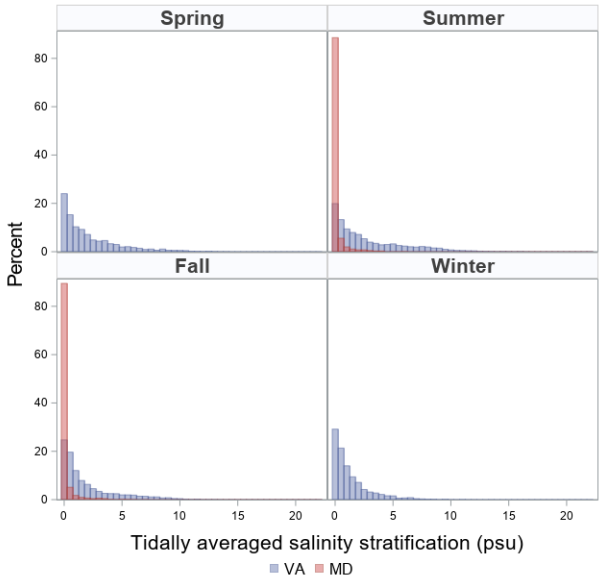



## Figure 6. continued

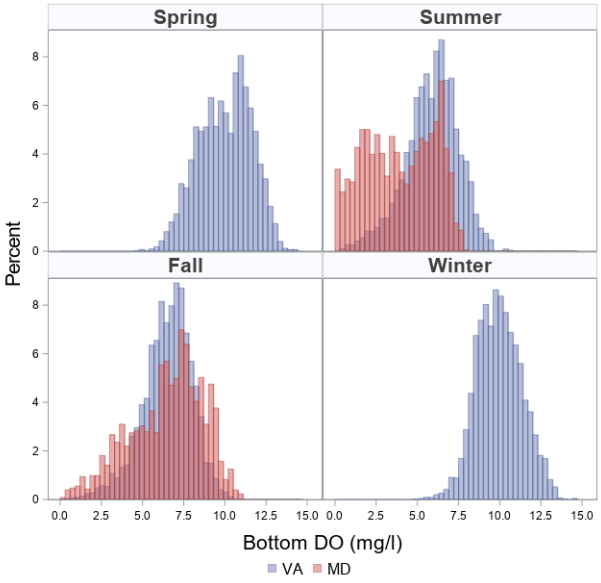


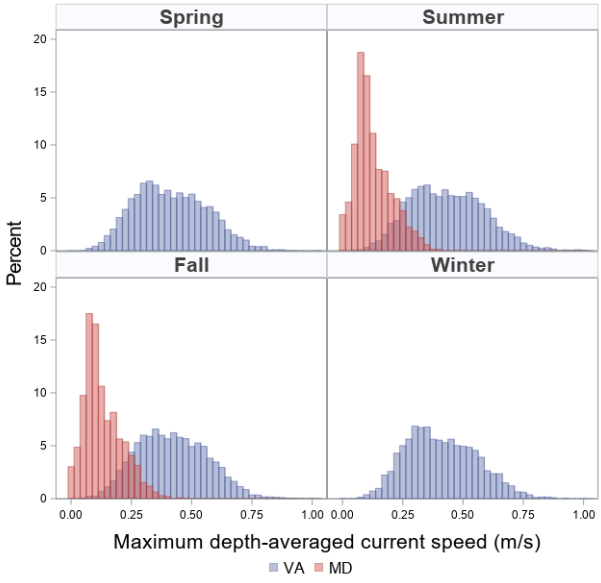



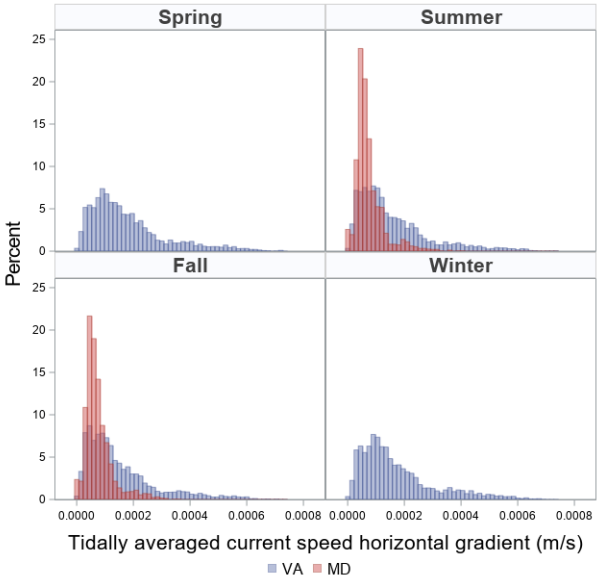



## Figure 6. continued



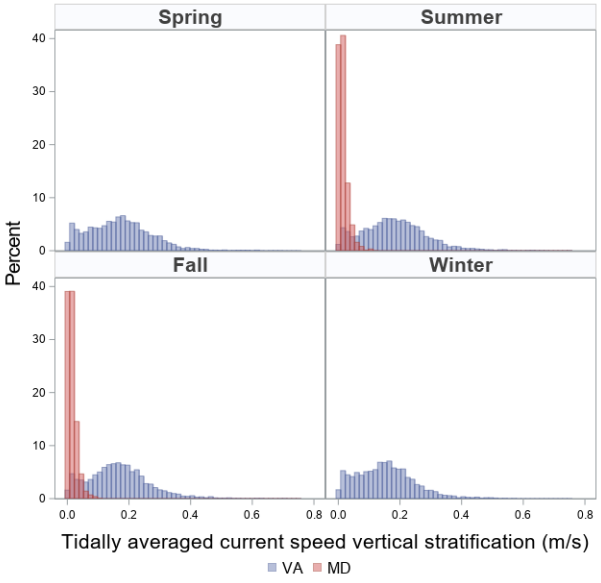



Figure 7. Suitability indices for 5 habitat covariates for juvenile spot in summer (May-July) for Chesapeake Bay, 2000-2016: bottom dissolved oxygen, distance to shore, water depth, horizontal gradient of tidal-averaged current speed, and tidal-averaged temperature stratification.

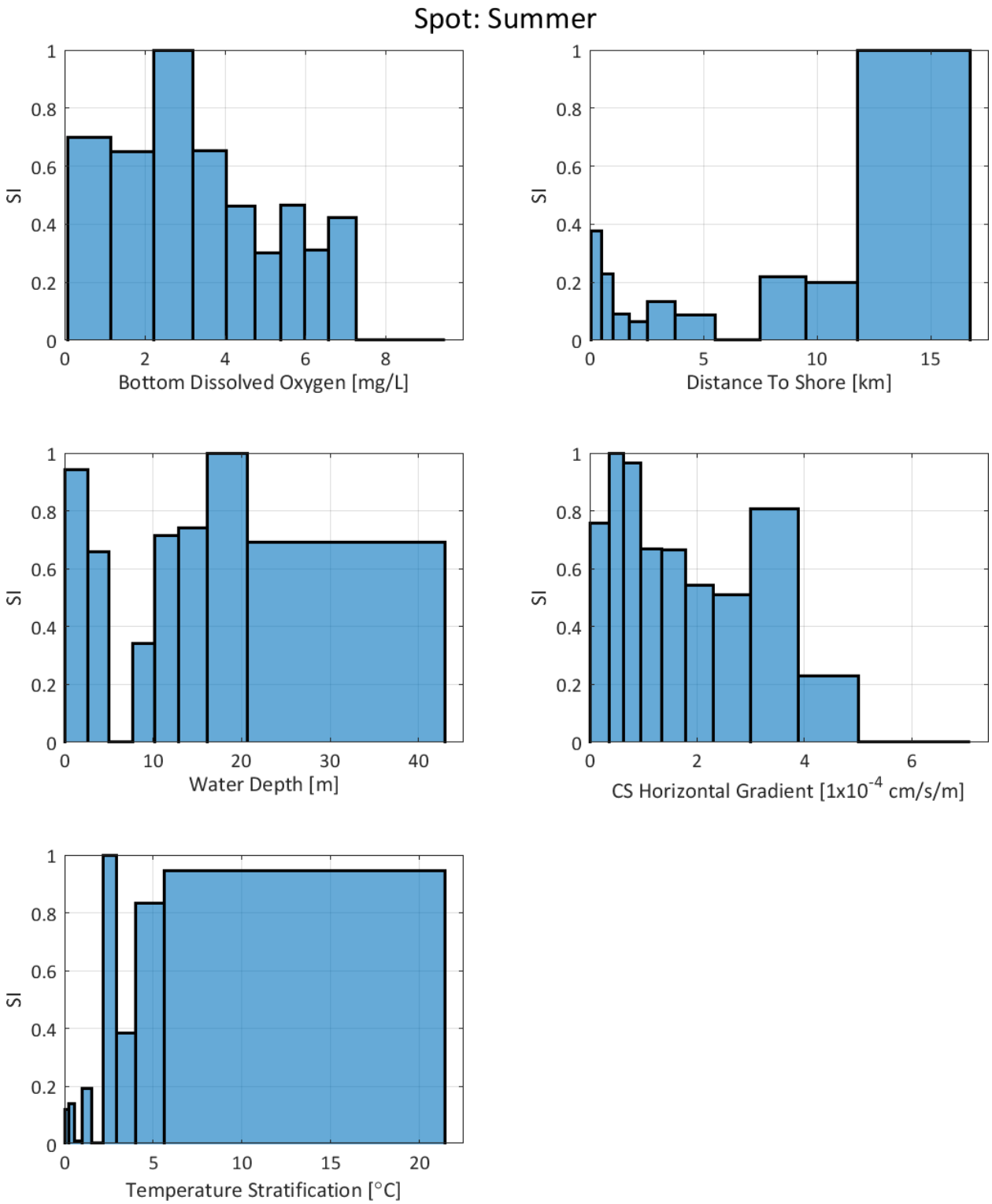


Figure 8. Suitability indices for 6 habitat covariates for bay anchovy in winter (December-February) for Chesapeake Bay, 2000-2016: bottom dissolved oxygen, distance to shore, tidal-averaged surface salinity, percent fine sediment, water depth, and horizontal gradient of tidal-averaged current speed.

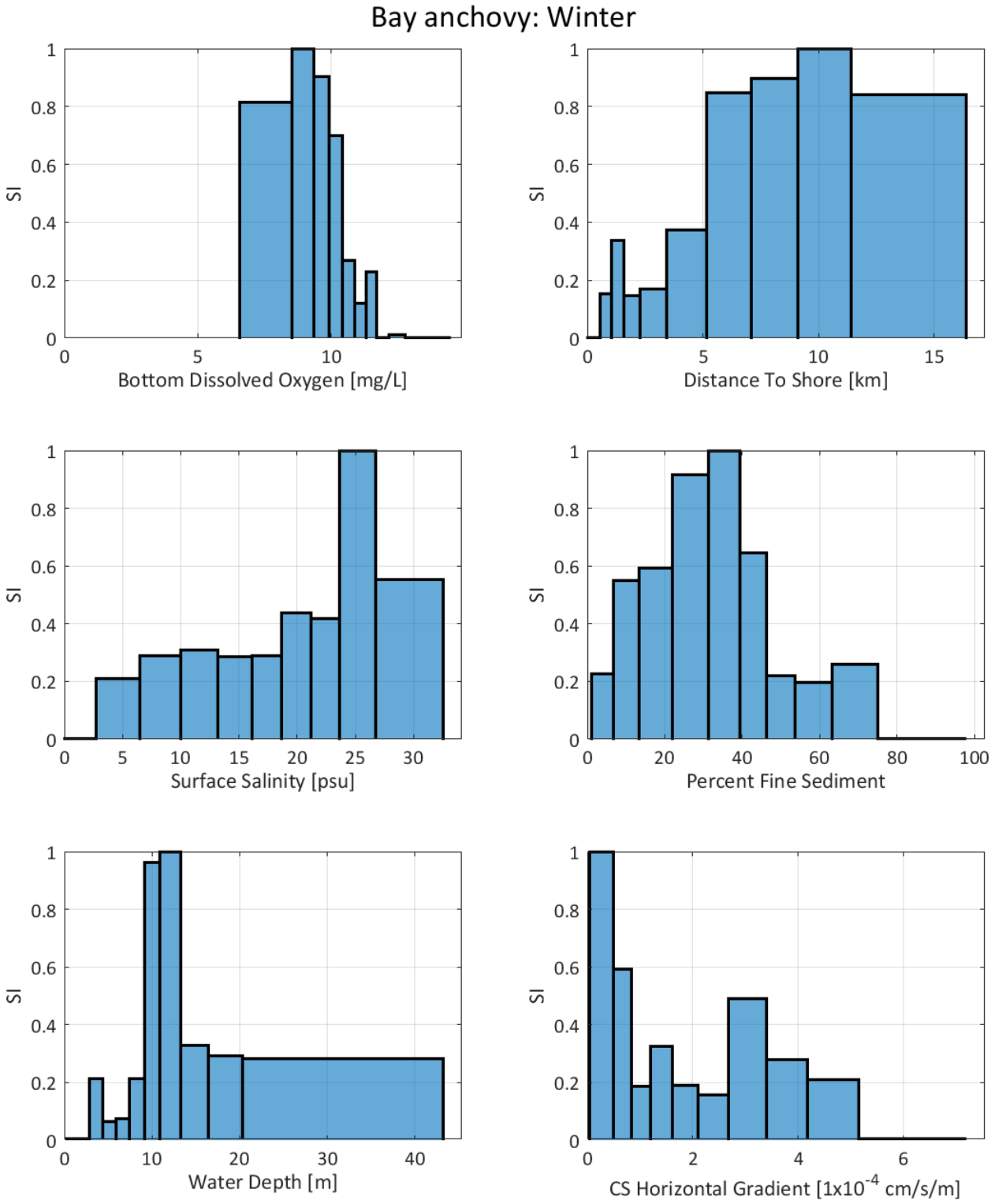


Figure 9. Box-and-whisker plots of the root mean-square error (RMSE) for two formulations of the habitat suitability index (HSI) estimated from 10 bootstrap replicates of 25,333 observations for forage species in Chesapeake Bay from 2000 to 2016. Bootstrap replicates were partitioned into training (~70%) and test (~30%) sets: the training set was used to identify influential covariates and to estimate the habitat suitability index (HSI); the test set was used to predict the HSI, which was then compared with the HSI from the training set and used to calculate RMSEs. The RMSEs were estimated across seasons for all species, except for bay anchovy, which lacked observations in fall. The diamond symbol in the plot displays the mean, the horizontal blue line is the median, the top and bottom of the boxes are the quartiles, the whiskers represent 1.5 times the interquartile range, and the open circles denote outliers. Significant differences in the mean RMSEs were detected for bay anchovy and hake; we did not detect a difference in mean RMSEs for spot or weakfish. Lower values of the mean RMSE indicate better model performance.

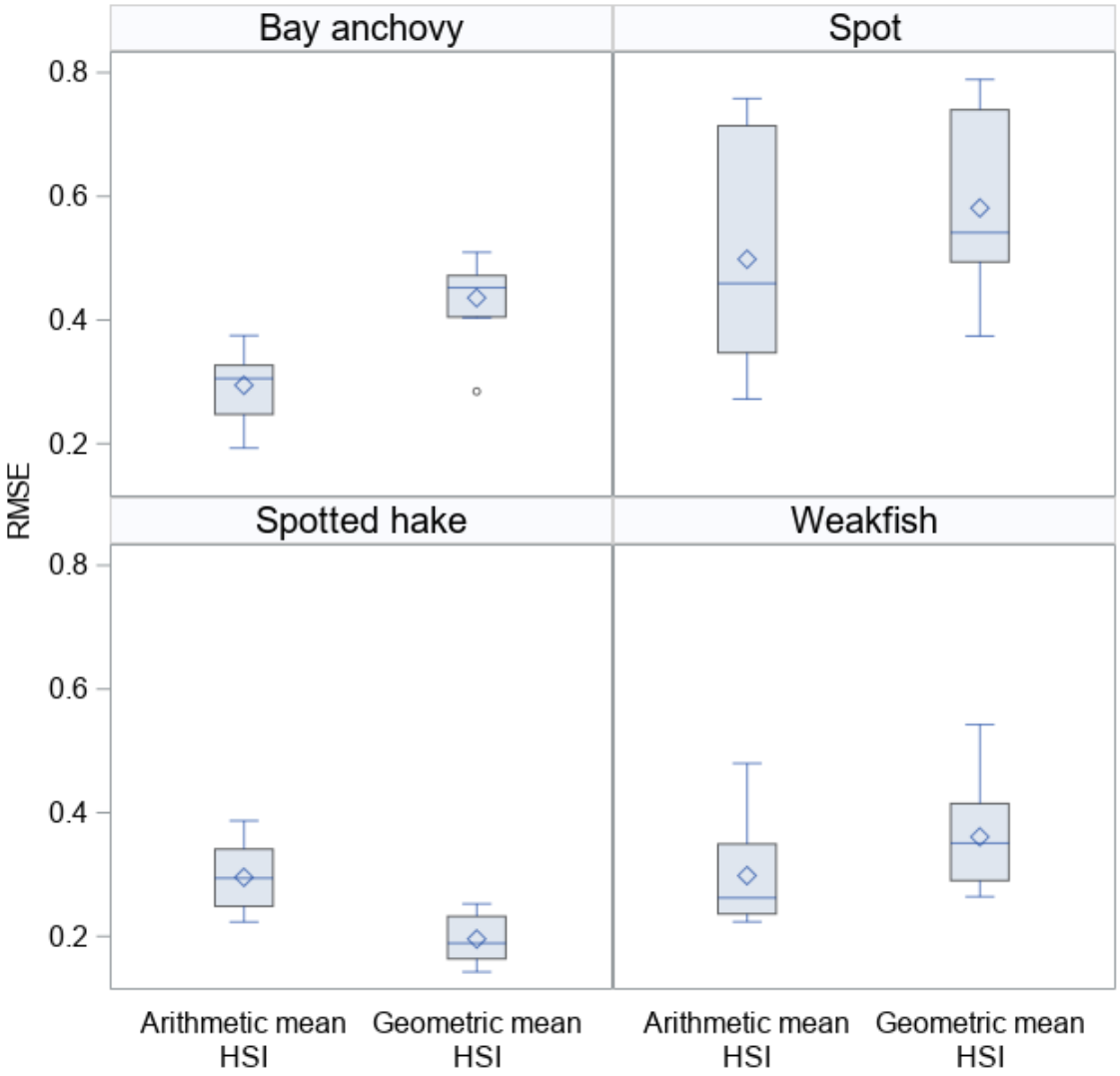


Figure 10. Relationship between HSI and trimmed mean catches of forage fishes in Chesapeake Bay, 2000-2016: juvenile spotted hake in spring, juvenile spot in summer, juvenile spot in fall, juvenile weakfish in summer, juvenile weakfish in fall, bay anchovy in summer, and bay anchovy in winter. For juvenile spotted hake, the $HSI_{gm}$ is shown, whereas the $HSI_{am}$ is shown for other species (see text). Note the different y-axes.

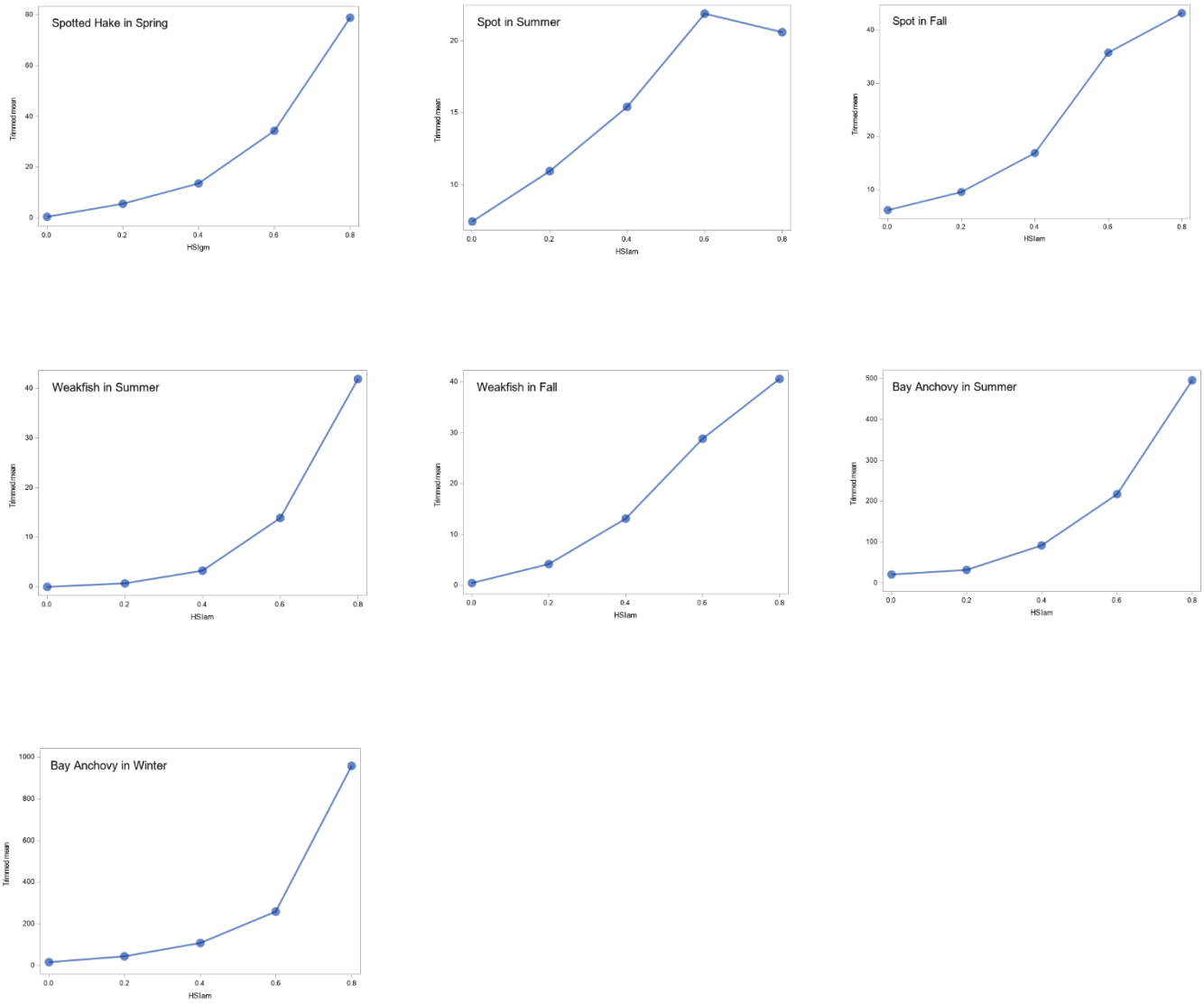


Figure 11. Frequency histograms of the estimated habitat suitability indices (HSI) for juvenile spotted hake in spring ($HSI_{gm}$), juvenile spot in summer and fall ($HSI_{am}$); juvenile weakfish in summer and fall ($HSI_{am}$); and bay anchovy in summer and winter ($HSI_{am}$) pooled for 2000 – 2016. $HSI_{gm}$ is the geometric mean HSI and $HSI_{am}$ is the arithmetic mean HSI.

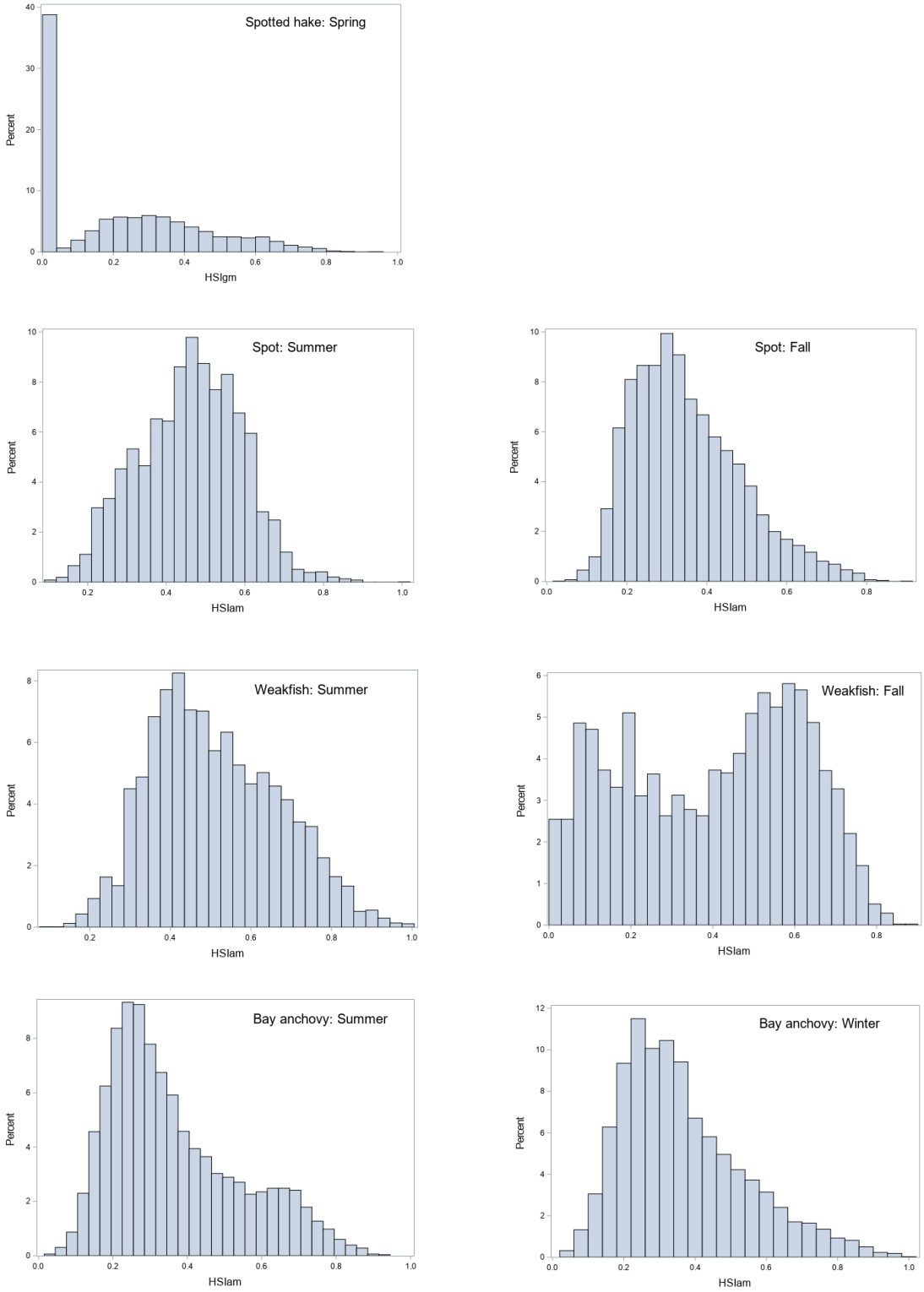


Figure 12. Extent of suitable habitats ($km^2$) for juvenile spot, juvenile weakfish, and bay anchovy in Chesapeake Bay and Mobjack Bay in summer 2010, 2011, and 2012. Suitable habitat was defined as habitats with $HSI_{am}$ values $\geq$ 0.5. The Mobjack Bay fisheries observations were not used to develop the habitat suitability model and thus, these independent data were used for external validation of the modeling approach.

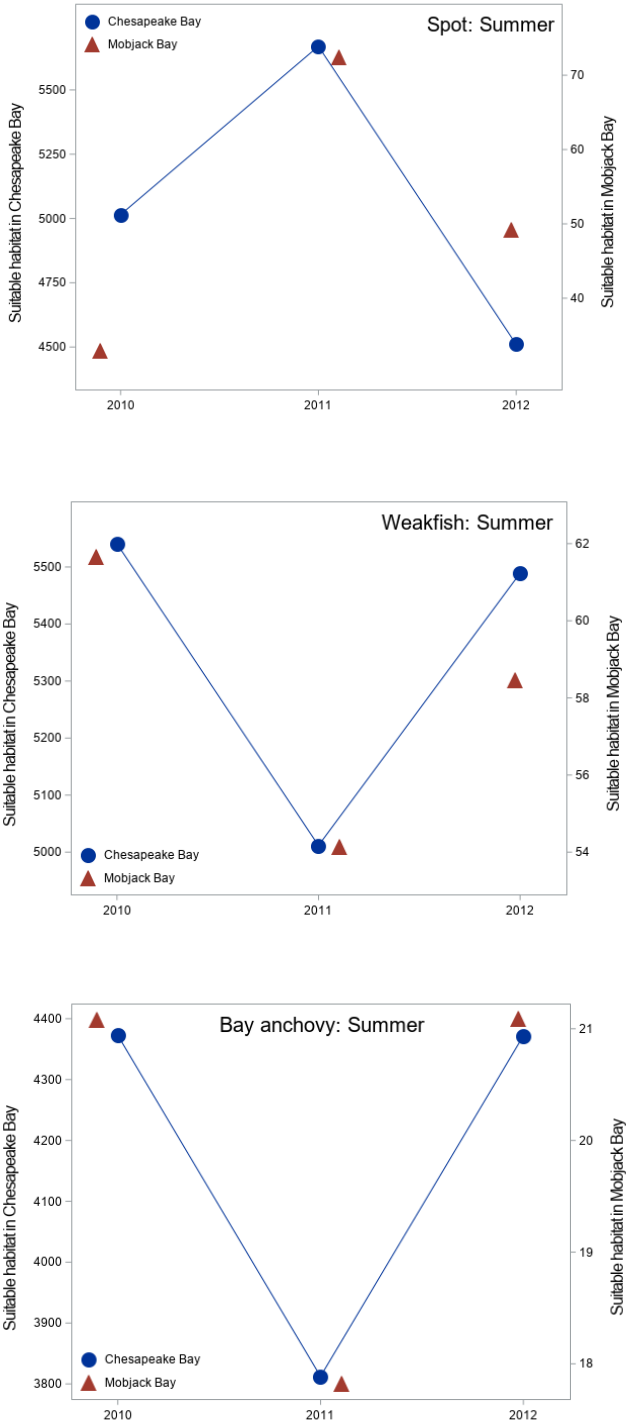


Figure 13. Patterns in the relative abundance per unit ($km^2$) of suitable summer habitat for juvenile spot, juvenile weakfish, and bay anchovy in Chesapeake Bay and Mobjack Bay in 2010, 2011, and 2012. Suitable habitat was defined as areas with $HSI_{am}$ values $\geq$ 0.5. The Mobjack Bay fisheries data were not used to develop the habitat suitability model and thus, these independent data were used for external validation of the modeling approach. The points were jittered to improve clarity.

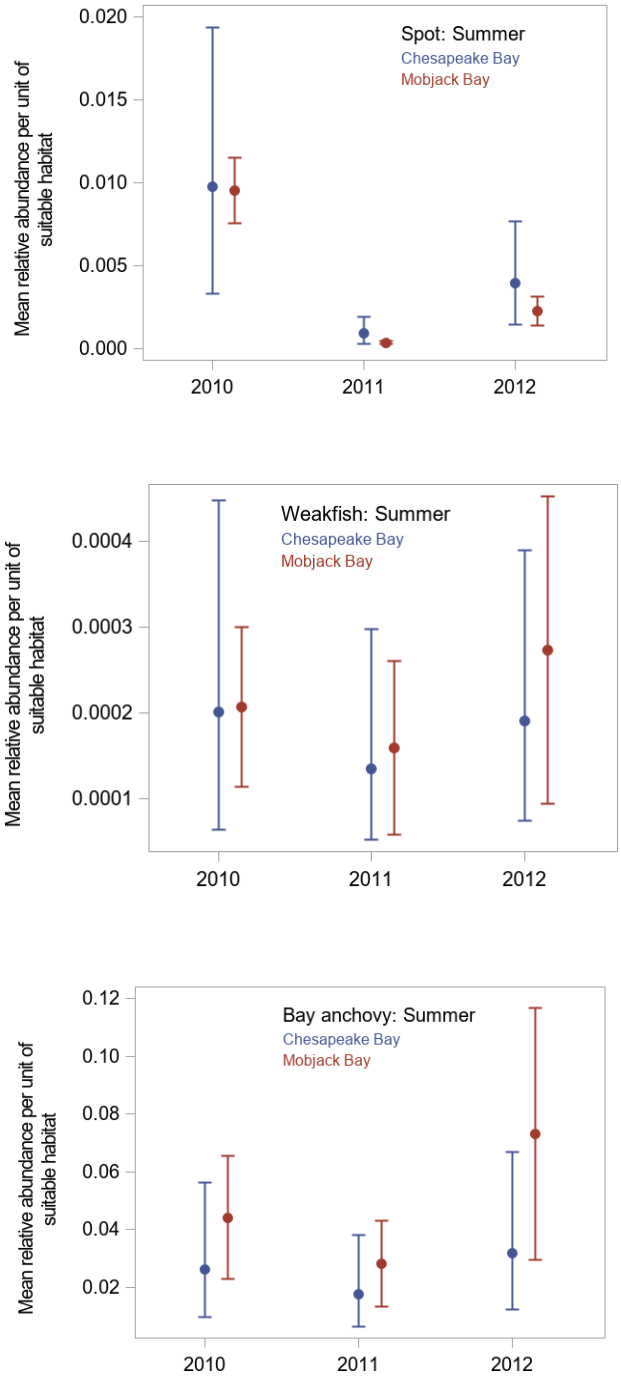


Figure 14. Habitat suitability for juvenile spot in 2011 (A) and 2014 (B) in Chesapeake Bay showing the annual variation in the extent of suitable habitats in summer for this species. The habitat suitability index ranges from 0 (red) indicating poor habitat to 1 (dark blue), with any shade of blue indicating suitable habitat.

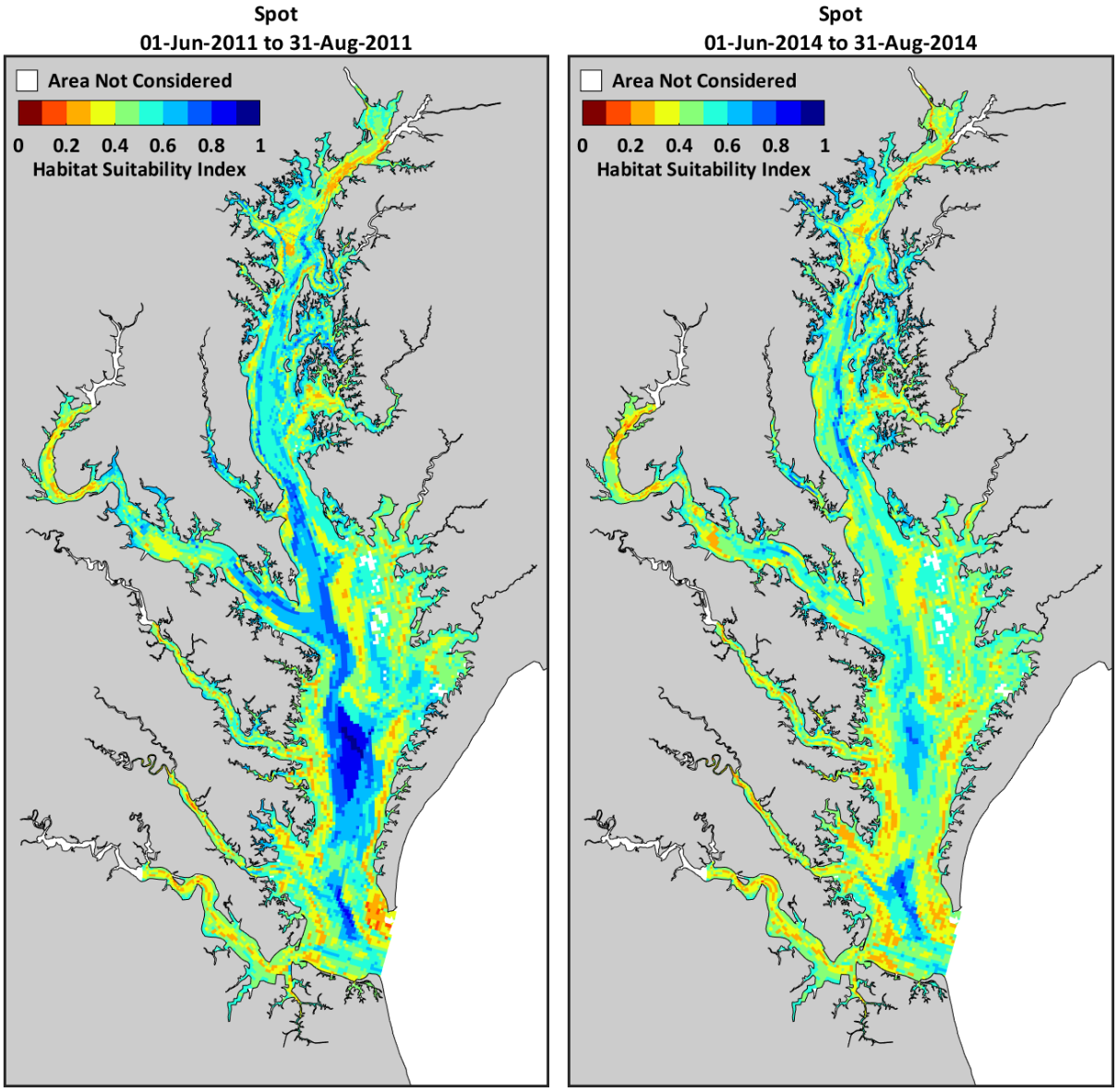


Figure 15. Habitat suitability for juvenile spotted hake in 2012 in Chesapeake Bay showing seasonal variation in the extent of suitable habitats for this species. The habitat suitability index ranges from 0 (red) indicating poor habitat to 1 (dark blue), with any shade of blue indicating suitable habitat.

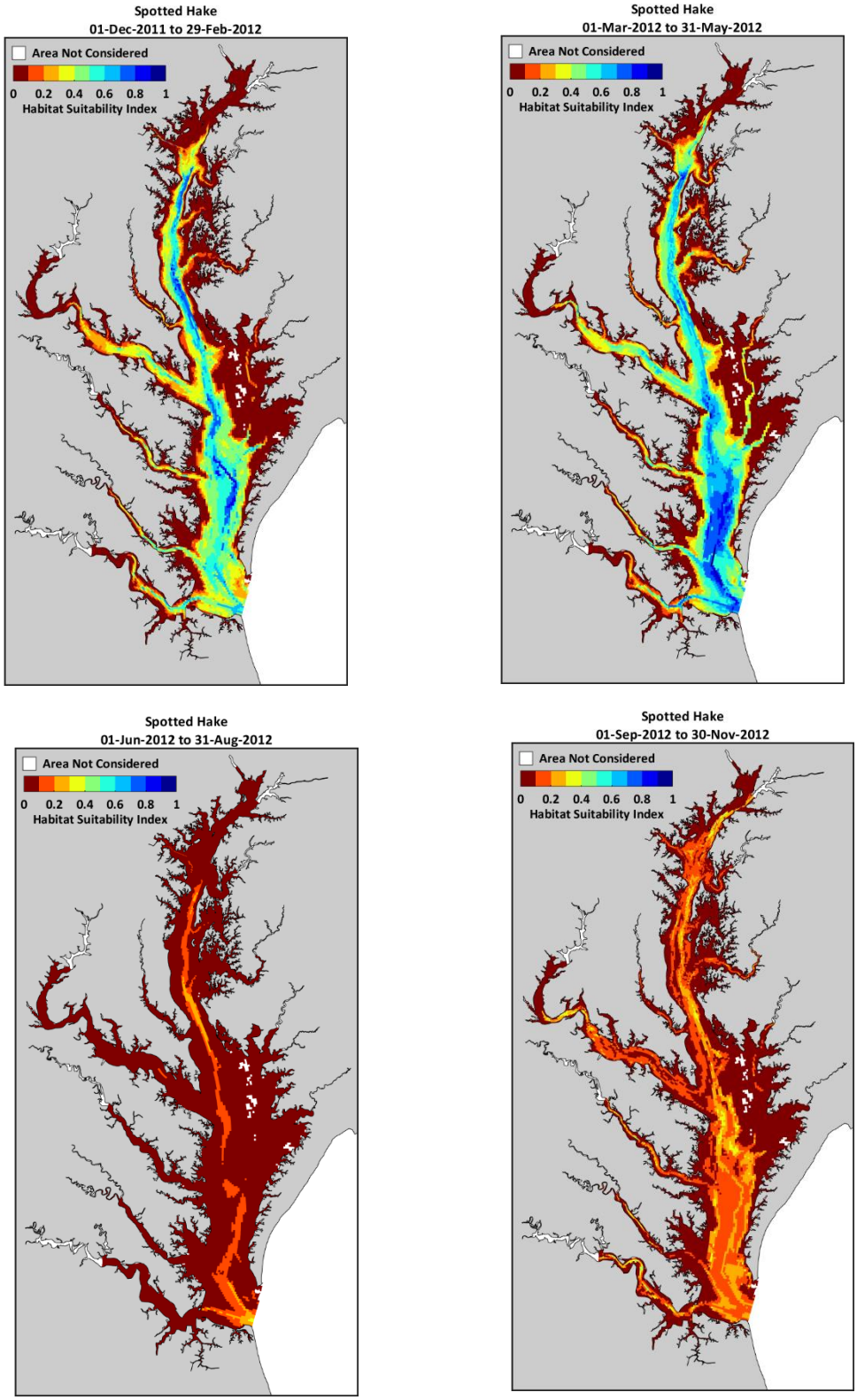


Figure 16. Habitat suitability for juvenile spot in 2011 in Chesapeake Bay showing seasonal variation in the extent of suitable habitats for this species. The habitat suitability index ranges from 0 (red) indicating poor habitat to 1 (dark blue), with any shade of blue indicating suitable habitat.

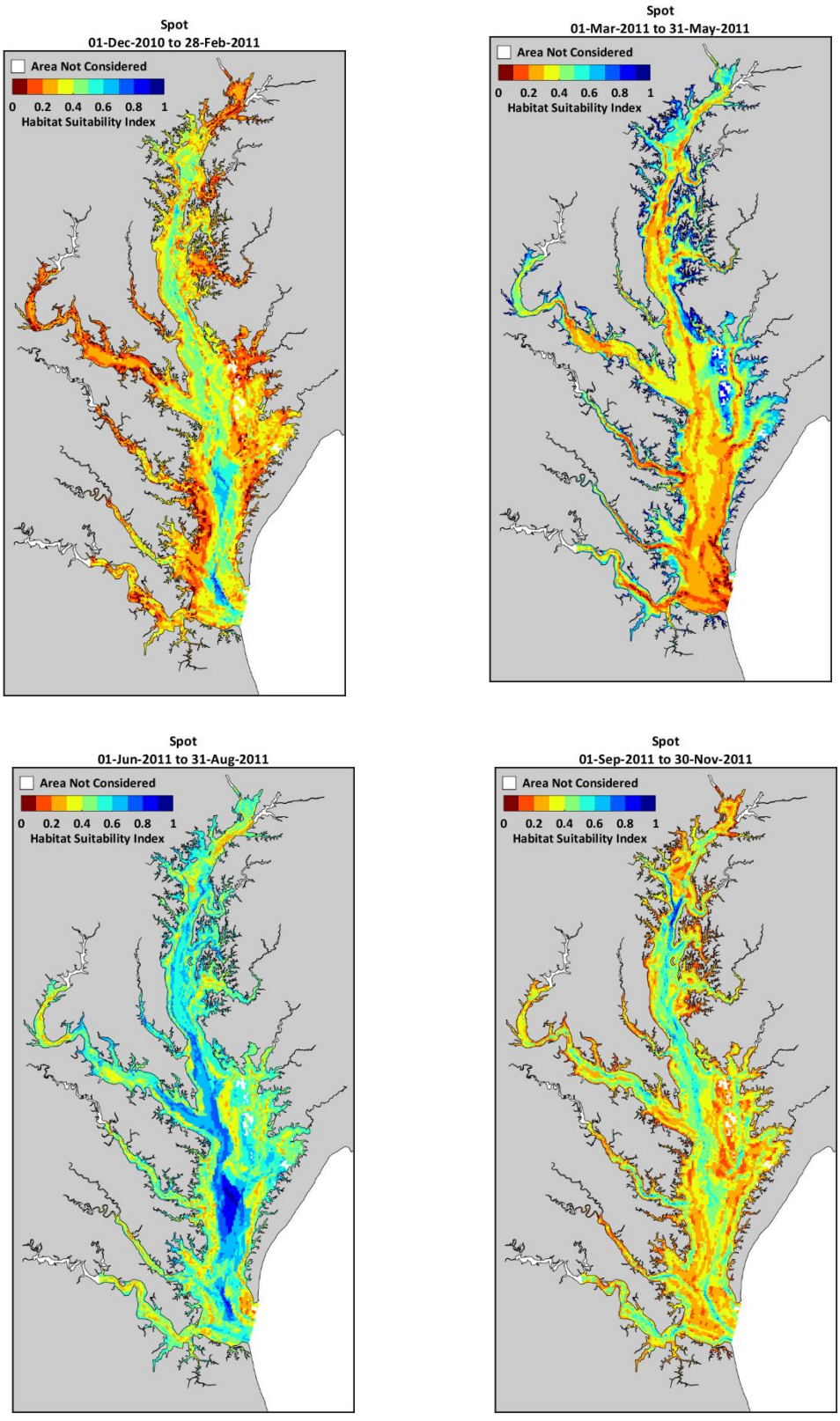


Figure 17. Habitat suitability for juvenile weakfish 2011 in Chesapeake Bay showing seasonal and spatial variation in the extent of suitable habitats for this species. The habitat suitability index ranges from 0 (red) indicating poor habitat to 1 (dark blue), with any shade of blue indicating suitable habitat.

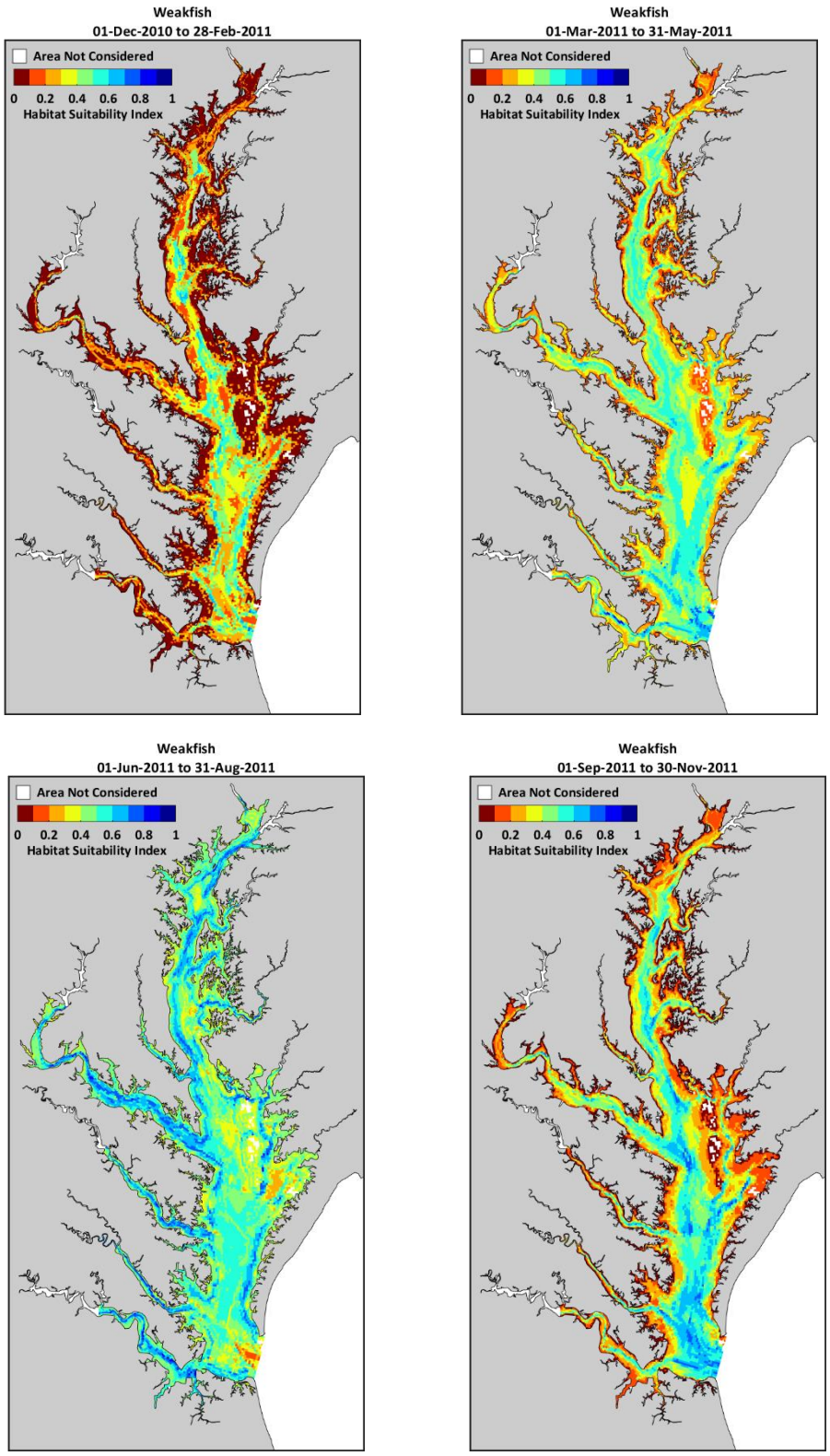


Figure 18. Habitat suitability for bay anchovy in 2011 in Chesapeake Bay showing seasonal and spatial variation in the extent of suitable habitats for this species. The habitat suitability index ranges from 0 (red) indicating poor habitat to 1 (dark blue), with any shade of blue indicating suitable habitat. Note that a fall habitat suitability model was not available for this species.

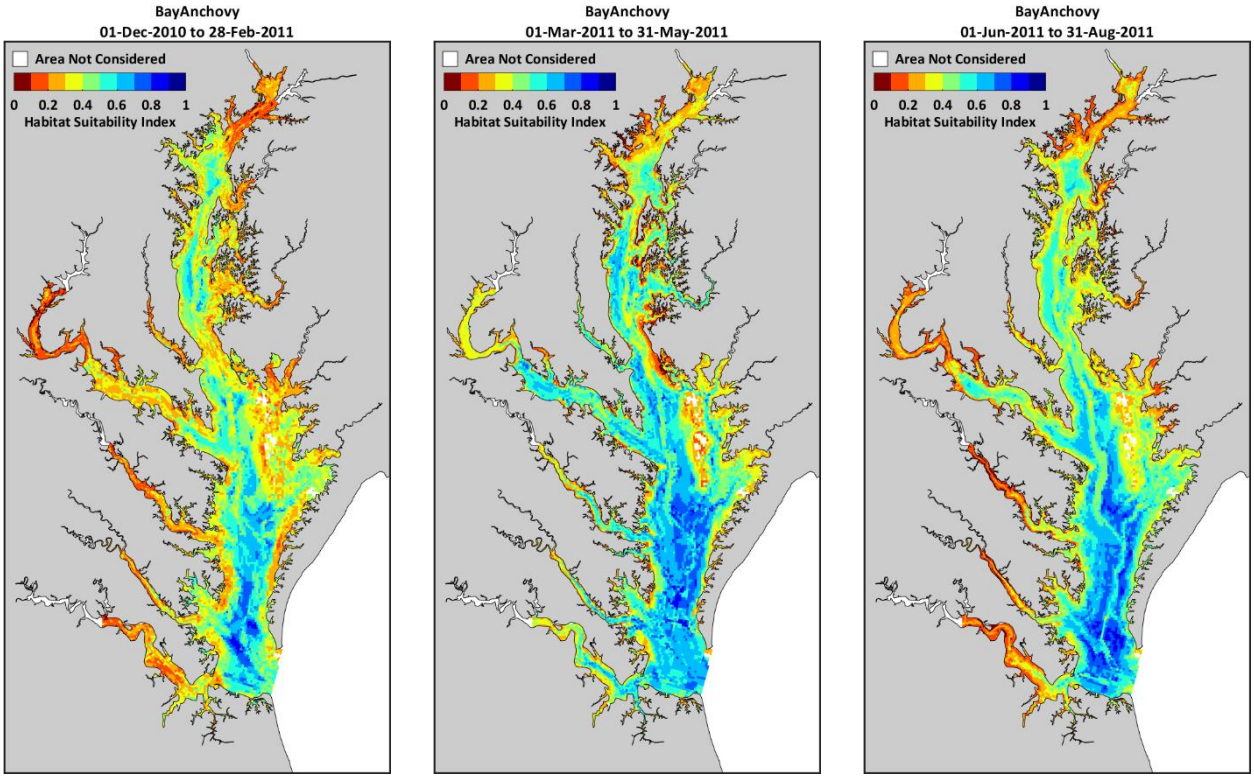
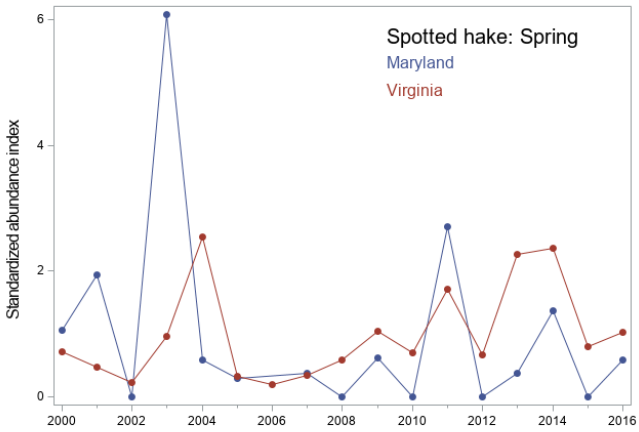
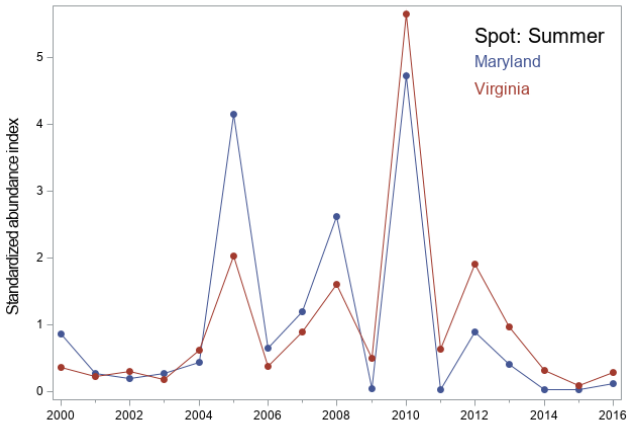
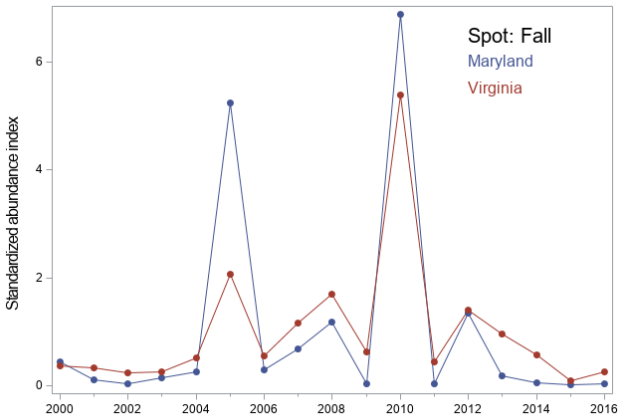


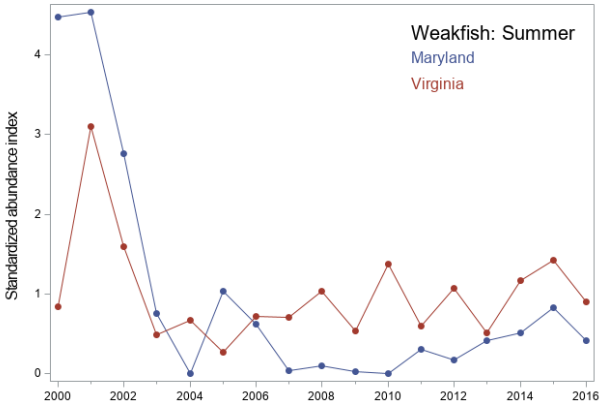
Figure 19. Seasonal standardized abundance indices for forage fishes (juvenile spotted hake, juvenile spot, juvenile weakfish, and bay anchovy) in Maryland (blue) and Virginia (red) waters of Chesapeake Bay, 2000-2016. Seasons were Feb-Apr for spring, May-Jul for summer, Aug-Oct for fall, and Nov-Jan for winter. The seasonal relative abundance index was estimated as the mean catch per unit effort (CPUE), calculated as the number of fish captured divided by the area swept by the trawl (number/km$^2$). Seasonal CPUEs were standardized to a mean of 1.0 across the 17 years, thus, patterns of abundance can be readily compared across states, but these standardized abundance indices do not reflect absolute differences in estimated mean CPUEs within a given year. For example, mean relative abundances (CPUEs) of spotted hake in spring were at least one order of magnitude greater in Virginia than in Maryland. Note that bay anchovy were not sampled in Maryland in winter, and thus, only the standardized index for Virginia is depicted.

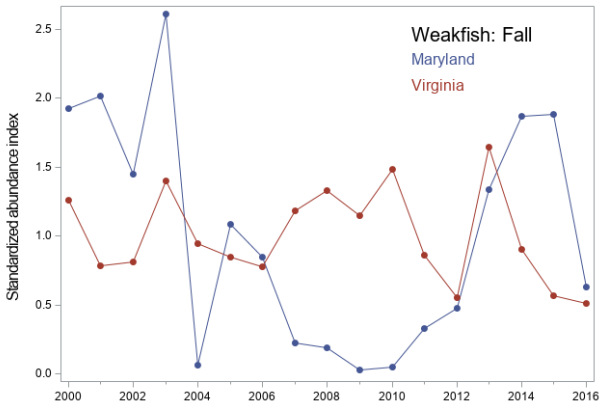


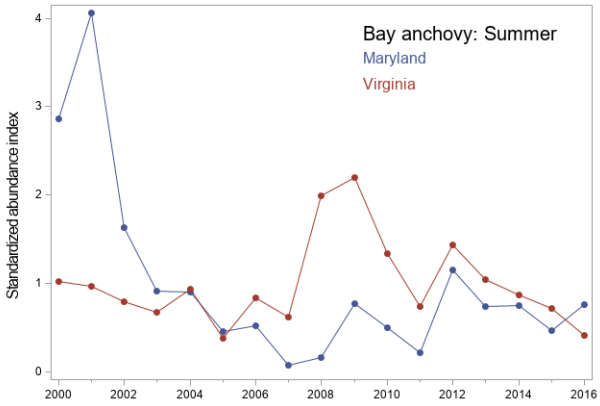



## Figure 19. continued



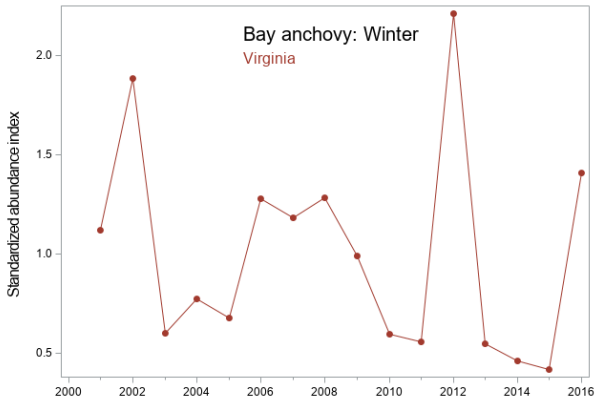
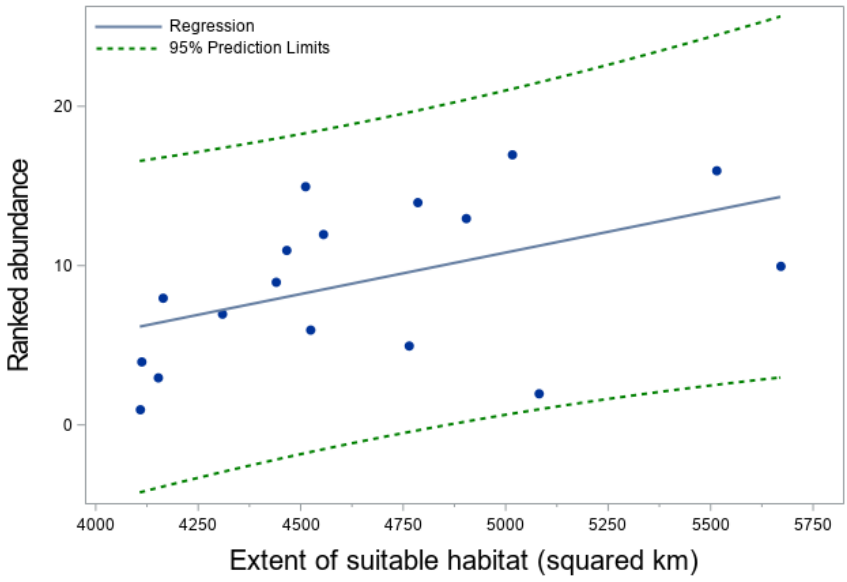



Figure 20. Nonparametric relationship between rank abundance and extent of suitable habitat ($km^2$) for (A) juvenile spot in summer and (B) bay anchovy in winter in Chesapeake Bay, 2000 – 2016. Observations are depicted by blue circles; the solid line is the nonparametric regression fit to the observations, and the dashed line is the 95% prediction limit. Values of $HSI_{am} \geq 0.5$ were considered suitable habitat.

(A)




(B)

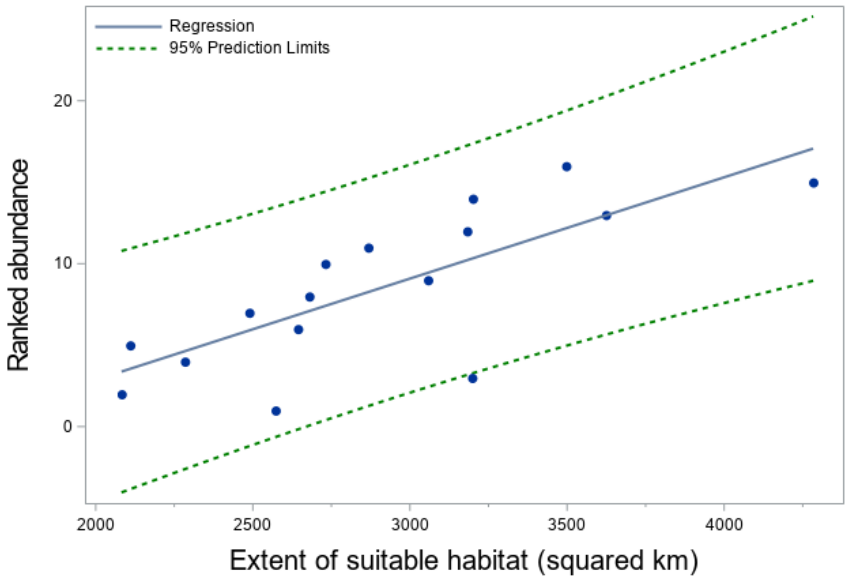
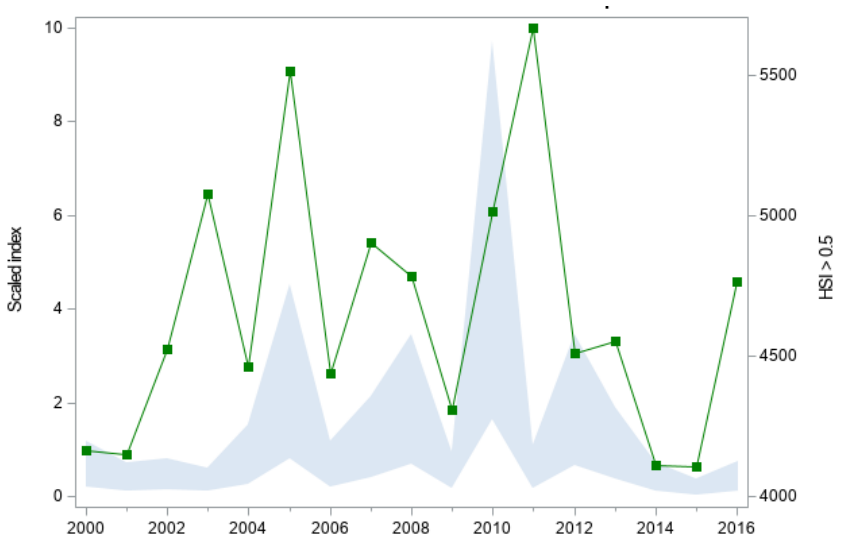


Figure 21. Relative abundance (scaled index) and extent of suitable habitat ($km^2$) for (A) juvenile spot in summer and (B) bay anchovy in winter in Chesapeake Bay, 2000 – 2016. Relative abundance (blue polygon) is depicted with a 95% credible interval; area of suitable habitat is denoted with green squares. Values of $HSI_{am} \geq 0.5$ were considered suitable habitat.

(A)




(B)

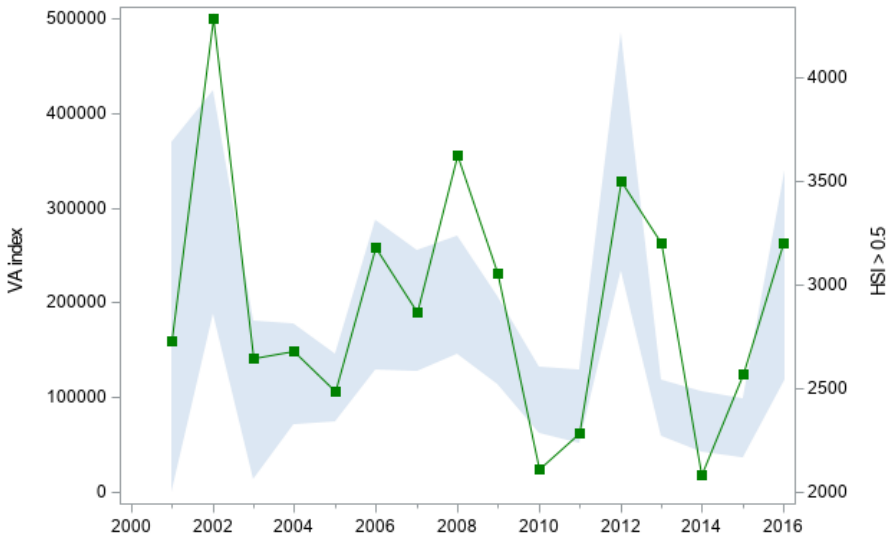
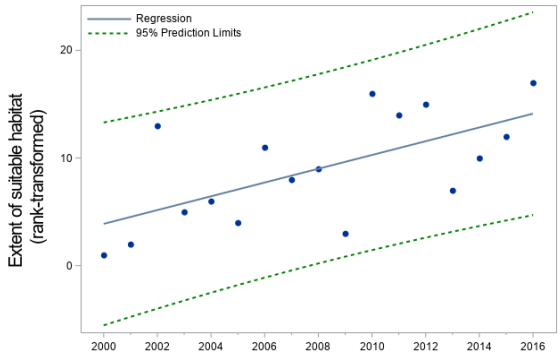
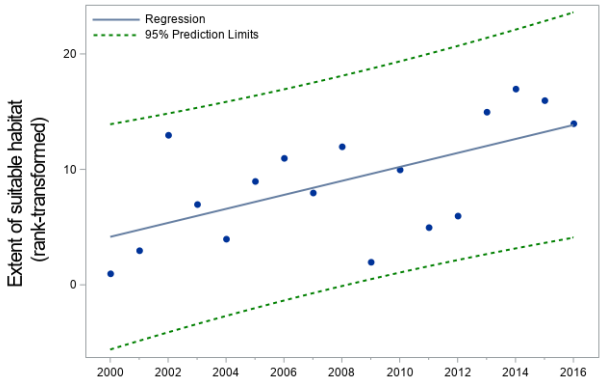


Figure 22. Pattern of change in the extent of suitable habitat for (A) juvenile weakfish in summer, (B) juvenile weakfish in fall, and (C) bay anchovy in summer in Chesapeake Bay, 2000 – 2016. The extent of suitable habitat was rank-transformed due to the low sample size (n=17), and the regression line estimates the nonparametric fit to the data. All regression slopes were positive and significantly different from 0 (Table 4). Values of $HSI_{am} \geq 0.5$ were considered suitable habitat.

(A)




(B)




(C)