

ABSTRACT

Title of Dissertation: FINE-SCALE SPATIAL ECOLOGY
OF AN ESTUARINE FORAGE FISH

Katelynn Marie Lankowicz
Doctor of Philosophy, 2022

Dissertation Directed by: Associate Professor Hongsheng Bi
Marine Estuarine Environmental Sciences

The spatial distribution of a species is shaped by the interactions of biotic and abiotic processes, which may cause unique patterns at any one spatial scale. This makes sampling at multiple spatial scales and across many habitat types essential to accurately determine and characterize the drivers of species abundance and distribution, though this has historically been difficult. In this dissertation, I examined the fine-scale spatial ecology of juvenile Atlantic menhaden (*Brevoortia tyrannus*) in a representative tidal estuarine tributary. Data were collected in the Patuxent River and its connected creeks using Adaptive Resolution Imaging Sonar (ARIS) cameras. The data were analyzed for trends in menhaden distribution at multiple spatial scales, and at aggregative extents from individual menhaden to clusters of menhaden schools. In Chapter 2, I used a generalized additive model (GAM) to characterize the density of individual menhaden across the lower Patuxent River. The highest density of menhaden was found within shallow-water tidal creeks, rather than in the Patuxent River channel. Water depth was a significant predictor of menhaden density. In Chapter 3, I developed techniques to identify the discrete schools into which

individual menhaden were aggregated. Schools typically consisted of fewer than 100 individuals. GAMs were used to characterize the density of menhaden within schools and the distance of schools from their nearest neighbors. Water depth, habitat, sampling month, and tidal stage were all significant predictors of menhaden school characteristics. In Chapter 4, I used spatial point pattern analysis to examine the spatial pattern of menhaden schools. Few identifiable trends were found. However, the results indicated that the relative impact of environmental and behavioral mechanisms on menhaden school distribution likely varies with spatial scale; environmental mechanisms have a greater impact at larger spatial scales, while behavioral mechanisms are more impactful at smaller spatial scales. Chapter 5 used the results of all empirical data analysis to inform the construction of two dynamic, spatially explicit individual-based models (IBMs) of menhaden movement. One IBM represented menhaden movement as only dependent on environmental stimuli, and the other represented menhaden movement as only dependent on the movements of conspecifics. Neither model was sufficient on its own to replicate the patterns of menhaden distribution observed in the empirical data. The spatial patterns that occur at any one space-time are likely the emergent result of the interactions between the two types of forces and individual variation in the perception of and response to local stimuli. The results are an important first step towards quantifying and resolving the relative impact of schooling behavior and environmental conditions on menhaden distribution. The research in this dissertation represents a small piece of estuarine forage fish spatial ecology. However, our understanding of any species' distribution in space and time is at its most comprehensive when it is seen as the emergent property of all its smaller components. The novel combination of fine-scale empirical data, statistical modeling, and individual-based models contribute to our understanding of the spatial ecology of forage fish in estuarine environments.

FINE-SCALE SPATIAL ECOLOGY
OF AN ESTUARINE FORAGE FISH

by

Katelynn Marie Lankowicz

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2022

Advisory Committee:

Dr. Hongsheng Bi, Chair/Advisor
Dr. Kenneth Rose
Dr. Ming Li
Dr. Edward Houde
Dr. Dong Liang

Chapter 2 © *Fisheries Research*, Elsevier B.V.
doi.org/10.1016/j.fishres.2020.105520

© Other materials copyright by
Katelynn Marie Lankowicz
2022

Acknowledgments

The end of a Ph.D. is a strange thing. It sits out in front of you without seeming to get any closer for so long, and then suddenly it's upon you. I've enjoyed the ride, and I want to thank the people who have made it possible. First, I'd like to thank my sources of funding: the Chesapeake Biological Laboratory Graduate Education Committee, the University of Maryland Ann G. Wylie Dissertation Fellowship, the Maryland Department of Natural Resources, and the National Science Foundation. Additional support was provided by the University of Maryland Jacob K. Goldhaber Award so I could attend conferences and present my research.

I'd also like to thank numerous people from Morgan State's Patuxent Environmental and Aquatic Research Laboratory, chiefly among them Dr. Chunlei Fan and Nikelene McLean, who helped collect sonar imaging data in the hot Maryland summers. I'd also like to thank Rob Nielsen for his boat-handling expertise and fabrication skills. I've received further help and had many interesting cultural conversations with Dr. Linlin Wang, Dr. Libin Zhang, Junting Song, Dr. Gang Lin, and Dr. Yuanzi Huo.

To my advisor, Dr. Hongsheng Bi: I don't think I can fully articulate how grateful I am to you. I can see your impact in every way I've developed as a scientist. Of course, you have taught me the traditional things you learn in a postgraduate degree like how to ask questions and where to find information. But more than that, you've been the understanding and compassionate voice I needed along the way. I appreciate you for your brilliance, your patience, and your kindness.

You always knew when I needed encouragement and when I needed a kick in the pants. I hope that I can do that for my own students one day.

To my committee members, Dr. Kenny Rose, Dr. Ed Houde, Dr. Dong Liang, and Dr. Ming Li: You have all given me so much help and guidance during this degree, and I am truly thankful. Your feedback, questions, and suggestions have served to point me in the right direction time and time again.

To Dr. Dave Secor: Thank you for your financial support and mentorship. Being able to participate in your lab activities, interact with you and your students, and contribute to the Nanticoke Sturgeon project gave me so much room to grow. I appreciate your taste in fieldwork music and your thoughtful advice.

To my labmates and friends, almost too numerous to count. Dr. Suzan Shahrestani, Kevin Kahover, Dr. Katie Martin, Alex Fireman, Samara Nehemiah, Nina Santos, Matt Guilherme, Maddy Lahm, Bri Groves, and Nicholas Coleman: I couldn't have made it this far without help from a small village worth of friends, colleagues, and roommates. Most of you pull triple duty in that regard. I want to thank you all for the myriad ways you've helped me get through this degree, whether that was lending a hand on a project, watching a terrible movie with me and a bunch of beer, loaning me some food, or just giving general encouragement.

To my best friend of 10+ years (!), Kate O'Brien: You really know how to keep me humble, and I do need that! Thank you for being the extroverted yin to my introverted yang since you flicked that dental rubber band into my hair in Mr. Hamman's class. I appreciate your sympathetic ear and all your expertly-curated TikToks.

To my Whale Hello pals, Nick Riddoch, Dr. Sarah Marley, and Dr. Phil Bouchet: Our Friday night Zoom hangouts are consistently the thing I look forward to every week. It has

been so wonderful to get to know you all. Thanks for accepting me into the fold and for all the invaluable advice.

To my dog, Jif Peanut Butter: you can't read, but I'll make sure to tell you in person how looking at your cute puppy eyes and floppy ears kept me from going off the deep end more than once. We've both gotten a lot grayer since 2016. You're a very good boy and co-pawthor, though I'm not sure you fit the criteria for me to list you as such in any journals.

To my grandparents, Ed & Maryann McErlean and Fran & Tony Lankowicz: So much of who I am came from the love and care you've always shown me. Pop pop, thank you for all the lessons I learned at "Professor Edward R. McErlean School." Maybe kindergarten was a little early to start multiplication, but I appreciate your belief in me. Grandma Maryann, I'll always remember the Golden Rule and learning to write my name by tracing over your letters. Papa Tony and Grandma Fran, thank you for teaching me checkers and chess, taking me to the library, and letting me pick up dozens of leaves, seeds, and pinecones to marvel over on our walks around your neighborhood.

To my sisters, Corinne and Allie Lankowicz: I can't tell you how many times I still think about spitting up grapes or the unhappy camper and absolutely lose it. It's strange how we're all so different, but still so very similar. I appreciate the change in perspective you both help me see, and the unwavering support I always feel from you.

To my dad and mom, John and Dr. Genevieve Lankowicz: I could fill pages with all the ways you've encouraged, inspired, and supported me. I think the first time I really started to wonder "What could be out there?" was that trip we took to Maine in 2001. I remember watching the lobster boats, finding the dead seal, and dropping pop pop's walkie-talkie in the drink. It all just filled me with a burning curiosity about the way aquatic life worked. I know you noticed

this curiosity, and you've helped me to chase it down ever since. I want to thank you for that, for all the books, classes, and trips, and for your endless patience as you (to this day) listen to me yammer on about the things I find interesting.

Finally, to my love, DOCTOR Leila Fouda: Well, now we've both done it. You have been my rock, and I don't think I could have made it here without you. You celebrated every victory with me and got me back on track when I ran out of steam. I am so lucky to have a partner that understands what I do and cares about my little fish. Though I will not forget the time you told me that you loved me but you didn't want to talk about R at 1 AM! I've appreciated every moment we've had together, whether that was walking around London, enjoying the southern Maryland sunshine, or just chatting on WhatsApp. I'm really looking forward to whatever comes next for us. There's so much left for us to learn, and facing the unknown isn't quite so scary when we can do it together. To quote our favorite, Ben Platt, "I don't know how this river runs, but I'd like the company through every twist and turn."

Table of Contents

Acknowledgements	ii
Table of Contents	vi
List of Tables	x
List of Figures	xi
Chapter 1: Introduction	1
1.1 Fisheries spatial ecology	2
1.2 The importance of forage fish	4
1.3 Atlantic menhaden	5
1.4 Menhaden spatial ecology and management	8
1.5 Bridging the gap: fine-scale spatial surveys and analyses	10
1.6 Individual-based models of fish movement	13
1.7 Dissertation framework	15
Chapter 2: Sonar imaging surveys fill data gaps in forage fish populations in shallow estuarine tributaries	18
2.1 Abstract	18
2.2 Introduction	19
2.3 Methods	23
2.3.1 Study area	23
2.3.2 Sonar imaging	24
2.3.3 Data processing	26
2.3.4 Statistical analysis	30
2.4 Results	35
2.4.1 Distribution between habitats	35
2.4.2 Distribution between depth strata	41
2.4.3 Seasonal variation	42
2.5 Discussion	44
2.5.1 Conclusions	48
2.6 Funding	49
2.7 Acknowledgements	50
Chapter 3: Characterizing juvenile Atlantic menhaden (<i>Brevoortia tyrannus</i>) distribu- tion and schooling in shallow estuarine waters	51

3.1	Abstract	51
3.2	Introduction	52
3.3	Methods	56
3.3.1	Study area	56
3.3.2	Data collection	57
3.3.3	Data processing	59
3.3.4	Statistical Methods	64
3.3.4.1	Generalized Additive Models	64
3.3.4.2	Cluster Coefficient	67
3.4	Results	67
3.4.1	Descriptive statistics	67
3.4.2	Generalized Additive Models	69
3.4.3	Clustering pattern	75
3.5	Discussion	76
3.5.1	Qualitative assessment of school characteristics	76
3.5.2	Behavioral drivers of school pattern and morphology	79
3.5.3	Environmental drivers of school pattern and morphology	81
3.5.4	Spatial clustering of schools	81
3.5.5	Potential limitations	82
3.5.6	Conclusions	84
3.6	Funding	84
3.7	Acknowledgements	85
Chapter 4:	Spatial point pattern analysis of juvenile Atlantic menhaden schools in shallow estuarine environments	86
4.1	Abstract	86
4.2	Introduction	87
4.3	Methods	91
4.3.1	Study area	91
4.3.2	Data collection	92
4.3.3	Data processing	93
4.3.4	Point pattern analysis	94
4.3.4.1	Linear networks	95
4.3.4.2	Pair correlation functions	97
4.4	Results	102
4.4.1	Spatial point pattern analysis	102
4.4.2	Comparative statistics	112
4.4.3	Varied spatial extents	115
4.5	Discussion	117
4.5.1	First-order structure	119
4.5.2	Second-order structure	121
4.5.3	Variation in spatial extent	122
4.5.4	Relative contribution of environmental and behavioral drivers	123
4.5.5	Study limitations	124
4.5.6	Conclusions	125

Chapter 5:	Spatially explicit individual-based models of juvenile Atlantic menhaden distribution and schooling	127
5.1	Abstract	127
5.2	Introduction	128
5.3	Methods	132
5.3.1	Empirical data	132
5.3.2	Simulating menhaden	133
5.3.3	Environment-responsive model	135
5.3.3.1	Model landscape	135
5.3.3.2	Scheduling and sub-models	136
5.3.3.2.1	<i>Sensing</i>	138
5.3.3.2.2	<i>Moving</i>	139
5.3.3.2.3	<i>Update individual characteristics</i>	140
5.3.4	Schooling-responsive model	140
5.3.4.1	Model landscape	140
5.3.4.2	Scheduling and sub-models	141
5.3.4.2.1	<i>Sensing</i>	141
5.3.4.2.2	<i>Moving</i>	142
5.3.4.2.3	<i>Update individual characteristics</i>	143
5.3.5	Identification of aggregations	144
5.3.6	Sensitivity to number of individuals and sensory distance	145
5.4	Results	146
5.4.1	Empirical data	146
5.4.2	Environment-responsive model	147
5.4.2.1	Habitat Suitability Index	147
5.4.2.2	Movement	148
5.4.2.3	Aggregations	148
5.4.3	Schooling-responsive model	151
5.4.3.1	Movement	151
5.4.3.2	Aggregations	152
5.4.4	Sensitivity tests	154
5.4.4.1	Number of individuals	154
5.4.4.2	Sensory distance	157
5.4.5	Comparison to empirical data	159
5.4.5.1	Environment-responsive model	159
5.4.5.2	Schooling-responsive model	160
5.5	Discussion	160
5.5.1	Sensitivity tests	163
5.5.2	Comparison to empirical data	165
5.5.3	Study limitations	166
5.5.4	Conclusions	166
Chapter 6:	Summary and conclusions	170
6.1	Novel technology applications	171
6.2	Habitat preferences	172

6.3	Characteristics of menhaden schools	173
6.4	Spatial pattern of menhaden schools	174
6.5	Impact of environmental and behavioral mechanisms on spatial patterns	175
6.6	Ecological implications	176
References		179

List of Tables

2.1	Monthly ARIS sampling output for 2016	33
2.2	Monthly ARIS sampling output for 2017	34
3.1	Terms for within-school density GAM	72
3.2	Terms for distance between NONS GAM	76
3.3	C_{dt} values for 2016 transects	77
3.4	C_{dt} values for 2017 transects	78
4.1	Point process models and point patterns for menhaden schools in 2016 transects .	103
4.2	Point process models and point patterns for menhaden schools in 2017 transects .	104
4.3	Point process models and point patterns for menhaden schools in combined and shortened extents	117
5.1	Summary statistics for sensitivity tests	158

List of Figures

2.1	ARIS sample image	28
2.2	Density of menhaden in the study area	37
2.3	Frequency of forage fish by habitat, transect, and depth bin	38
2.4	GAM smoothers	40
3.1	Map of Patuxent River transects	57
3.2	ARIS image of a school of menhaden	61
3.3	Illustration of overlapping ARIS images	63
3.4	Monthly comparisons of menhaden school characteristics	70
3.5	Within-school density GAM output	71
3.6	Between-school distance GAM output	74
4.1	Transect with no first- or second-order structure	105
4.2	Transect with second-order structure but no first-order structure	106
4.3	Transect with first-order structure but no second-order structure	107
4.4	Transect with first- and second-order structure	108
4.5	$\hat{\rho}$ plots of IPPP-modeled transects with longitude as a spatial covariate	110
4.6	$\hat{\rho}$ plots of IPPP-modeled transects with latitude as a spatial covariate	111
4.7	$\hat{\rho}$ plots of IPPP-modeled transects with depth as a spatial covariate	113
4.8	Values of pointwise distances r from all non-CSR point patterns	114
4.9	Values of pointwise distances r from all non-CSR point patterns in combined- extent transects	116
4.10	Values of pointwise distances r from all non-CSR point patterns in shorter-extent transects	118
5.1	Illustration of sensory-distance mediated sensory neighborhood	134
5.2	Suitability indices for environmental variables	137
5.3	Habitat suitability index values for the modeled grid at time step 1	138
5.4	Zones of influence around an individual in a schooling-responsive model	143
5.5	12-hour averaged HSI for 30-day modeled period	147
5.6	Positions of 1000 individuals at two randomly-selected time steps	148
5.7	Movement paths of 5 randomly-selected individuals over 3 modeled days in the environment-responsive model	149
5.8	Example of the alignment of individuals within aggregations from both models	150
5.9	Number of aggregations, number of individuals per aggregation, and within- aggregation density averaged over 12-hour moving windows for 1000 individuals in the environment-responsive model.	151

5.10	Positions of 100 individuals at two randomly-selected time steps	152
5.11	Movement paths of 5 randomly-selected individuals over 1 modeled day in the schooling-responsive model	153
5.12	Number of schools, number of individuals per school, and within-school density averaged over 12-hour moving windows for 100 individuals in the schooling- responsive model.	154
5.13	Environment-responsive model results with 1000, 2000, and 4000 individuals . .	156
5.14	Schooling-responsive model results with 100, 200, and 400 individuals	168
5.15	Environment-sensing model with variations in sensory distance	169

Chapter 1: Introduction

The distribution of a species is described by the density and pattern of individuals through space. Individuals are not naturally distributed at random or in a uniform fashion, but rather in identifiable spatial patterns such as patches and gradients ([Legendre and Fortin, 1989](#); [MacArthur, 1972](#)). The formation of these patterns implies the existence of an influential underlying process or processes that should be both measurable and relatively predictable ([MacArthur, 1972](#)). These are the first principles on which spatial ecology is built, and the tenets on which spatial ecologists seek to understand the formation and maintenance of a species' spatial distribution.

Accurately determining a species' spatial distribution and inferring the identities of the underlying processes is a challenging task. Distribution patterns can be highly complex, as they are influenced by an array of biotic and abiotic processes ([Wisz et al., 2013](#)). The interactions of these processes may be nonlinear, non-additive, and have emergent properties that can only be identified at higher organizational levels ([Ciannelli et al., 2008](#); [Newman et al., 2019](#)). Reaching a full and integrated understanding of spatial distribution depends on proper identification and knowledge of the drivers of pattern and density at all the scales on which they operate, as the patterns at any scale can have unique causes ([Levin, 1992](#)). This challenging task is broadly understood in ecology and is commonly referred to as the problem of pattern and scale ([Chave, 2013](#); [Dray et al., 2012](#); [Fortin et al., 2002](#); [Levin, 1992](#); [Rahbek, 2005](#); [Turner et al., 1989a](#);

Wiens, 1989).

There is no single natural scale at which all ecological processes occur, so spatial ecologists must choose the appropriate scale or scales to measure each pattern and process of interest. As the scale of a study shifts, the relevant controls on spatial patterns and processes change (Dray et al., 2012). Spatial scale is determined by its two components: extent, which is the scope of the study area, and grain, which is the minimum resolution (Turner et al., 1989b). Studies conducted at inappropriately small extents may fail to recognize patterns or influential driving processes operating out of sampling range. If an extent is too large for a study, it can lead to inefficiencies in sampling and processing time. Small grain size (e.g., one individual in a population) can also lead to sampling inefficiencies, or produce unrepresentative and unpredictable data. Inappropriately large grain size can incorporate aggregation error and mask spatial patterns and processes. The natural world is both fine-grained and of a large extent. Ecologists must therefore appropriately balance grain and extent when selecting all necessary scales to address their objectives.

1.1 Fisheries spatial ecology

Though ecologists have understood for decades that organisms have heterogeneous distributions through a range of spatial scales, sampling and modeling programs still often fail to consider multi-scale species-environment interactions. This is especially true in fisheries ecology (Ciannelli et al., 2008; Mason and Brandt, 1999). To this point, the biggest advances in fisheries science have been in determining the causes of fluctuations in abundance and modeling these population dynamics at a species level (Beverton and Holt, 1957; Cushing, 1990; Hjort, 1914; Ricker, 1975). In this approach, demographic and environmental variables are averaged over

large extents and at relatively large grain sizes to test their correlation. There is often a failure to consider how population demographics and environmental variables may vary with scale, leading to a potential misrepresentation of the processes and relationships that regulate fish survival, abundance, and distribution ([Ciannelli et al., 2008](#)). Incorporating a multitude of spatial scales in studies of fish distribution would not only be interesting for its own sake but also would elucidate some of these underlying processes and aid in understanding how species abundance changes over time and under variable environmental conditions.

The complexity of organism spatial distribution is partially driven by the heterogeneity and variability of the environment. Estuaries, as the confluence of fresh and marine waters, are some of the most spatially diverse and dynamic ecosystems on the planet ([Vinagre et al., 2017](#)). The sheer complexity and unpredictability of estuaries necessitates increased sampling effort needed to fully describe environmental conditions and organism populations ([Caeiro et al., 2003](#); [Elliott and McLusky, 2002](#); [Samedy et al., 2015](#)). Local and landscape-scale processes in estuaries are known to contribute to the creation of striking spatial patterns such as banded mussel beds, diatom-generated elevated mudflats, and fractal-like creeks which drive vegetation distribution ([van de Koppel et al., 2012](#)). The spatial heterogeneity inherent to estuaries also contributes to their high productivity; estuaries are highly biodiverse, and many organisms are estuarine-dependent for at least a portion of their life histories. The Chesapeake Bay, as the largest estuary in the United States and one of the largest estuaries in the world, is a prime representation of the spatial complexity, biodiversity, and productivity of most estuaries.

1.2 The importance of forage fish

From a human perspective, some of the most important organisms in estuaries and coastal marine ecosystems are forage fish. Forage fish is a collective term applied to a group of relatively small and highly abundant fish species which occupy intermediate trophic levels and therefore facilitate energy transfer from lower to higher trophic level species ([Bakun et al., 2010](#)). Although juvenile stages of many fish species could theoretically fit these requirements and thus be categorized as forage fish, this dissertation will use the term to refer to planktivorous, pelagic fish species which remain critical prey resources throughout their entire life histories ([Rountos, 2016](#)).

Depletion of forage fish and other similar-trophic level prey items has a significant negative impact on upper-trophic level predators, indicating some incidence of bottom-up control ([Cury et al., 2011](#); [Smith et al., 2011](#)). As such, forage fish are integral to the diets of marine mammals, fish-eating seabirds, and several life stages of piscivorous gamefish ([Hartman and Brandt, 1995](#); [Luo and Brandt, 1993](#); [Pikitch et al., 2012, 2014](#); [Walter III and Austin, 2001](#); [Weise and Harvey, 2008](#); [Wood and Houde, 2003](#)). Forage fish are relatively short-lived, have high fecundity, and respond rapidly to environmental factors, leading to large fluctuations in inter-annual abundance ([Pikitch et al., 2012](#)). This variability can have wide-reaching effects on the growth, reproductive rate, distribution, and abundance of their predators ([Gislason, 2003](#)).

In addition to their ecological value, forage fish also have high economic value to coastal regions. In 2006, global forage fish landings were worth approximately \$5.6 billion and supported other fisheries worth an additional \$11.3 billion ([Pikitch et al., 2012](#)). Several taxa are of particular importance; clupeids (herrings, sardines, and menhadens) are known for their high

productivity and valuable reduction, bait, and seafood fisheries. In many estuaries of the East Coast of the United States, including the Chesapeake Bay, clupeoid forage fish such as anchovies (*Anchoa* spp.) and menhaden (*Brevoortia* spp.) often comprise the majority of fish biomass (Gallagher et al., 2018; Jung and Houde, 2003; Taylor and Rand, 2003; Wood and Houde, 2003). Atlantic and Gulf menhaden (*Brevoortia tyrannus* and *Brevoortia patronus*, respectively) support two of the most valuable and largest fisheries by volume in North America. In 2019, total US menhaden landings approached 685,000 metric tons worth over \$156 million (National Marine Fisheries Service, 2021). Atlantic menhaden are both common to estuarine ecosystems of the eastern United States and highly valuable in an economic and ecological sense; this dissertation focuses on their spatial ecology.

1.3 Atlantic menhaden

Out of the many species of forage fish and seven extant species of menhaden, Atlantic menhaden has sometimes been titled the “most important fish in the sea” (Franklin, 2007). While this may be a North Atlantic-biased viewpoint, Atlantic menhaden has indeed supported North American communities as crop fertilizer since pre-colonial times (Franklin, 2007). Native Americans and early colonists mainly harvested menhaden at small scales for direct consumption, fertilizer, or agricultural feed (Clark, 2020). As the whaling industry began to fail in the 1840s with declining whale populations, fisher-farmers of New England noted the potential for menhaden oil to replace highly-valued whale oil (Clark, 2020; Goode, 1880). Whereas whaling required a significant input of resources to hunt increasingly few, large creatures in the open ocean, menhaden were both plentiful and much closer to shore. After the destruction of many boats special-

ized for long whale-chasing voyages during the American Civil War, commercialization of the much-more accessible menhaden fishery began in earnest ([Clark, 2020](#); [York, 2017](#)). Even in the early stages of industrialization, there was widespread recognition by fishers that menhaden contributed to coastal ecosystems as prey items for larger, commercially-valuable piscivorous fish ([Bolster, 2012](#); [Clark, 2020](#)). This often led to conflict between menhaden purse-seine fishers and hook-and-line fishers from other fisheries ([Clark, 2020](#); [Frye, 1978](#)). Though the fishery has gone through many changes in the last 180 years, including major expansion and contraction, Atlantic menhaden is still harvested for bait, reduction to oils, and rendering to agricultural or aquaculture feed today.

The life history of Atlantic menhaden contributed to the development into a major fishery and also contributes to their ecological value. Menhaden is estuarine-dependent and frequently found in near-shore coastal areas. Spawning hotspots occur on the coastal shelf off the Carolinas from December through February ([ASMFC, 2017](#)). Spawning can also occur in coastal bays and sounds ([ASMFC, 2020](#)). Adults will then undertake an extensive, size-structured northward migration; larger individuals will migrate as far as the Gulf of Maine. Older and more fecund individuals may be able to spawn along the migratory routes and at the northern feeding grounds ([Ahrenholz, 1991](#)). Southward migration of adults commences in the fall and the process is repeated. Recent studies have suggested that a portion of the menhaden population may reside year-round in the northern area ([Liljestrand et al., 2019](#)). Menhaden can live 10-12 years, though individuals more than six years old have been rare in fishery-dependent data since the 1960s ([ASMFC, 2020](#)).

After spawning and hatching events, menhaden eggs and larvae are transported northward and inshore by currents, winds, and tides ([Ahrenholz, 1991](#); [Hare et al., 1999](#)). The swimming

ability of larvae is limited, and thus they are still dependent on ocean circulation for transport to estuaries (Hare et al., 1999; Quinlan et al., 1999). They are further dependent on processes like wind forcing, residual bottom inflow, and selective tidal transport to enter estuaries and reach upstream nursery areas (Hare et al., 2005). Once inside estuaries, larvae metamorphose into juveniles.

Whereas larval menhaden are known to prey on zooplankton, post-metamorphic juveniles are filter-feeders and consume phytoplankton after development of gill-rakers (June and Carlson, 1971). Juveniles are thought to prefer sandy- or muddy-bottom habitats (SEDAR, 2020), and abundance is positively correlated with phytoplankton biomass (Friedland et al., 1989, 1996). Decreasing temperatures and declining food availability in the fall trigger emigration of the year-class out of the estuaries and into the coastal oceans, where they join the adult stock (Friedland et al., 1996; Friedland and Haas, 1988). Sexual maturity is linked to length, but most individuals reach sexual maturity at age one (ASMFC, 2020).

This dissertation does not attempt to resolve the fine-scale spatial ecology of all age classes of menhaden across their large spatial extent but instead focuses on juvenile menhaden in a tidal sub-estuary of the Chesapeake Bay. The Chesapeake Bay has long been considered the primary nursery habitat for Atlantic menhaden (Anstead et al., 2017; Simpson et al., 2016), with some studies estimating it provides up to 68.8% of recruitment to the spawning stock (Ahrenholz et al., 1989). Menhaden are among the most highly abundant fishes in the Chesapeake (Jung and Houde, 2003; Wood and Houde, 2003). Similar to other estuarine juvenile populations, the abundance of juvenile menhaden in the Chesapeake is positively correlated with primary production (Friedland et al., 2011; Houde et al., 2016).

1.4 Menhaden spatial ecology and management

In recent years, policymakers have pushed to change the management of many forage fish from single-species to ecosystem-based management methods ([Pikitch et al., 2004](#)). This stems from the recognition that though a forage species may not be overfished and overfishing may not be occurring, a depleted abundance of these fish could have negative consequences across their associated ecosystems. It is therefore important to have accurate and unbiased forage fish indices of abundance to develop ecosystem-based management methods.

Atlantic menhaden, as a species with high economic importance, high ecological importance, and broad distribution across the East Coast of the United States, has emerged as a high-profile case study in shifting to an ecosystem approach to fisheries management ([Buchheister et al., 2017](#); [Franklin, 2007](#)). Currently, Atlantic menhaden is managed as a single coastwide stock by the Atlantic States Marine Fisheries Commission ([ASMFC, 2020](#)). As discussed earlier, accurate representations of species distribution and abundance will require sampling across a range of spatial scales. However, gear types frequently used in menhaden and other forage fish surveys impose limitations on the spatial scales that can be analyzed. Many studies have noted flaws in both grain and extent of net-based survey methods like trawls and seines ([Buchheister et al., 2016](#); [Jung and Houde, 2003](#); [Liang et al., 2020](#); [Lucca and Warren, 2018, 2019](#); [Miller et al., 2008](#)). Seining surveys are conducted in extreme nearshore shallows (<2m), leading to a problem in extent by limitations in sampling volume and site selection. Trawl surveys can also have issues in extent driven by site selection; it is difficult to conduct trawls in shallow (<4m) waters due to vessel limitations. Trawls can also have flaws in grain size, as the highest resolution of spatial data they can provide is the length of the tow.

Despite these flaws, most menhaden indices of abundance included in stock assessments and the ASMFC's menhaden Interstate Fishery Management Plan (IFMP) are generated from seine or trawl survey methods ([ASMFC, 2017](#); [SEDAR, 2015](#)). There is an argument to be made that when combined, trawl and seine surveys have adequate grain and extent to create indices of relative abundance that can answer most questions surrounding menhaden stock dynamics. However, even when combined, these two methods under-sample or completely omit much of the extent of shallow waters between 2 and 4 meters deep, which is potentially a habitat of importance for menhaden and other forage fish.

There is yet another complication when examining forage fish spatial distribution. Many forage fish species, including menhaden, exhibit aggregative behaviors ([Pikitch et al., 2012](#)). Schooling, or aggregation with synchronized or polarized swimming behavior, is common ([Pitcher, 1986](#)) and results in forage fish rarely being seen alone ([Fréon et al., 2005](#)). Aggregative behaviors are therefore important to consider when describing and predicting menhaden spatial distribution. These aggregative behaviors can be described by five levels of increasing complexity: 1) individuals, 2) schools of individuals, 3) clusters of schools, 4) concentrations of clusters, and 5) the entire stock ([Fréon et al., 2005](#)). The influence of various biotic and abiotic factors operating at multiple spatial scales is known to create “patchy” patterns of forage fish schools, in which individuals group into schools in favorable areas of heterogeneous environments ([Gutiérrez et al., 2012](#); [Haugland and Misund, 2004](#); [Liang et al., 2020](#)). However, the dynamics of these aggregative levels are highly variable and poorly understood, particularly at the intermediate levels ([Beare et al., 2002](#)).

It is known that population inferences are unlikely to be robust if sampling methods are based on untested or poor assumptions of a species' distribution ([Anderson, 2003](#)). Relative

abundance estimates collected only from a few areas thought to be representative of the whole system cannot sufficiently account for spatial variation and may not lead to robust inferences of fish abundance and distribution. It is therefore important to understand how fish are distributed within all their habitats and to consider the influence of aggregative behaviors on their spatial patterns. The distribution data needed to address this data gap is on the scale of centimeters to tens of meters, which can be challenging to collect or parse out of existing datasets.

1.5 Bridging the gap: fine-scale spatial surveys and analyses

Hydroacoustic surveys have shown much promise in overcoming the extent and grain flaws inherent to net-based survey methods, as well as the ability to bridge the data gap present in shallow-water habitats ([Becker et al., 2017](#); [Boswell et al., 2010](#); [Lucca and Warren, 2018](#); [Samedy et al., 2015](#)). Hydroacoustic survey and data processing methods have proliferated in the fisheries ecology and fish management communities and are now seen as a viable replacement or supplement for traditional net-based methods ([Kubečka et al., 2012](#); [Liang et al., 2020](#); [Samedy et al., 2015](#); [Simmonds and MacLennan, 2005](#)). Several commercially important species have long-established hydroacoustic surveys as a major component of their abundance and distribution quantification. These surveys use echosounders to record high-resolution spatial data along selected transects within broad spatial scales, typically with a low temporal resolution (annual or biannual surveys). Alaska Pollock (*Gadus chalcogramma*) has been surveyed using a combination of acoustic and bottom trawl methods across much of the Bering Sea for the last 40 years ([Honkalehto et al., 2011](#)). Atlantic herring have been surveyed similarly in the Baltic, Barents, and North Seas since the late 1970s ([ICES, 2014](#); [Simmonds, 2009](#); [Toresen et al.,](#)

1998). Other groups with data from hydroacoustic surveys included in their abundance surveys are sprat, capelin, anchovies, sardines, and euphausiids (Gjøsæter et al., 1998; Hagström et al., 1989; Trenkel et al., 2011).

Echosounders are well-known, well-established methods to quantify the biomass of small pelagic fishes, even in shallow waters (Boswell et al., 2010, 2007; Lucca and Warren, 2018). However, echosounders are unable to identify individual fish. They rely on acoustic backscattering from fish bodies, especially the air-filled swim bladders of many species, to produce a signal which can then be attributed to fish biomass (Misund, 1997). Sonar imaging is a newer active acoustic survey method and uses recent technological advances to create near-video quality digital images by sending out high-frequency sound pulses and interpreting the returning echoes. This increase in image quality and resolution makes it possible to observe, enumerate, and measure individual fish. Sonar imaging cameras are also small enough that they can be deployed from small boats in shallow areas, expanding the extent that can be effectively surveyed. These qualities make sonar imaging cameras an excellent tool for identifying juvenile menhaden spatial distribution patterns at small spatial scales usually overlooked in net-based survey methods (Able et al., 2013, 2014).

Once empirical data gaps have been addressed with this new technology, spatial patterns which were previously obscured due to sampling issues in grain and extent can be observed. Many spatial fisheries ecology studies have focused on meso- or macro-scale spatial dynamics, with less attention paid to fine-scale spatial information (Barnett et al., 2019). With the availability of increasingly smaller-grain data, questions surrounding fine-scale spatial dynamics and aggregative behaviors can be addressed. Habitat usage patterns and correlations between juvenile menhaden abundance and previously un- or under-considered physical environment characteris-

tics, for example bathymetry will be of specific interest. Distribution patterns collected via sonar imaging with water quality information may also indicate if physical gradients such as temperature, salinity, dissolved oxygen, and turbidity are important spatial drivers at multiple scales. Several levels of aggregative behaviors can be concurrently observed and described, and their relationships to environmental covariates can be explored.

Although processing sonar imaging data is currently labor-intensive, the robust results and fine-scale detail on the spatial structure of menhaden are worth the effort. Further, advances in computing power and machine learning technology like artificial neural networks are rapidly decreasing processing time ([Boswell et al., 2008](#); [Shahrestani et al., 2017](#)). Semi-automated image processing procedures have been successfully developed for optical zooplankton identification and counting ([Bi et al., 2015](#); [Grosjean et al., 2004](#)) and stationary sonar imaging applications ([Shahrestani et al., 2017](#); [Yin et al., 2020](#)). Some progress has been made in developing image processing procedures to detect and track human divers using a mobile sonar imaging camera, but these developments have been very limited and recent ([Kvasić et al., 2019](#)). Though it has not been completed yet, Dr. Hongsheng Bi's laboratory at the University of Maryland Center for Environmental science is developing a similar method to automate the processing of mobile sonar imaging data and identify multiple species of interest.

Fine-grained data from mobile sonar imaging studies and other hydroacoustic studies make it possible to aggregate and assess juvenile menhaden distributions over a wide range of spatial scales, including increasingly smaller scales where few assessments have previously been made. Assessing species distribution over multiple spatial scales is extremely valuable, as it can aid in detecting explanatory variables not previously considered or processes that drive distribution patterns ([Dray et al., 2012](#); [McIntire and Fajardo, 2009](#)). However, wrangling fine-scale spatial

data can be complex. The volume of data necessitates significant computing power. Often, spatial data will need to be aligned and matched to analyze changes in distribution through time (Finley et al., 2014). Assumptions will need to be made about the mathematical function describing distribution across multiple spatial scales, and special consideration will need to be paid to the complex, nested spatiotemporal dependencies in the data. Several statistical methods attempt to analyze spatial distribution data with multiple explanatory variables across multiple spatial scales and aggregative levels, including generalized additive models (GAMs) and spatial point pattern analysis.

1.6 Individual-based models of fish movement

The level of detail observed in high-resolution empirical data can be represented to a relatively high degree using a spatially-explicit individual-based model (IBM; sometimes called agent-based model, or ABM) (van der Vaart et al., 2016). IBMs simulate the actions of individuals as they interact with each other and respond to the environment (DeAngelis and Mooij, 2005; Grimm and Railsback, 2005), and so have often been identified as extremely valuable tools for fisheries scientists and managers to make predictions of fish population dynamics in a changing environment (Huse and Ellingsen, 2008; Huse and Giske, 1998). IBMs have successfully been used to model bay anchovy (*Anchoa mitchilli*) across all life stages and multiple generations within the Chesapeake Bay (Rose et al., 1999). In contrast to bay anchovy, which often resides in the Chesapeake Bay over its entire life, Atlantic menhaden is estuarine-dependent and migrates between estuarine and coastal ocean ecosystems. The survival and transport of oceanic-phase menhaden larvae have been modeled using a coupled IBM- hydrodynamic model system (Quin-

[lan et al., 1999](#)), but the movement of young-of-the-year (YOY) menhaden within estuaries has not yet been modeled. Modeling movement in dynamic environments is expected to be challenging both due to a lack of knowledge of a species' perception of its environment and the inherent challenges of changing movement cues between time steps ([Watkins and Rose, 2017](#)). However, an accurate characterization of menhaden spatial distribution within their estuarine habitats from empirical fine-scale spatial distribution data can help overcome these problems. The addition of a spatially explicit IBM for juvenile menhaden to the body of work on Atlantic menhaden management could help answer questions surrounding their survival in estuaries and how their population dynamics may shift in a changing environment.

Individual-based models are also invaluable for identifying emergent properties at a population scale. Dozens of studies have explored the formation of flocks, schools, or herds of animals based on individual behaviors, beginning with Reynolds ([Reynolds, 1987](#)). The emergent property of fish schooling has been explored in simulated populations ([Huth and Wissel, 1992](#); [Parrish et al., 2002](#); [Reuter and Breckling, 1994](#)) or with limited laboratory observations ([Viscido et al., 2004](#)). An examination of how schooling behavior develops from collective individual behaviors and maintains itself in a spatially explicit, heterogeneous, and dynamic environment based on extensive and fine-scale field observations is a challenging but worthwhile goal, given its potential to elucidate some of the behavioral mechanisms and environmental characteristics driving fish movement and aggregation.

1.7 Dissertation framework

This dissertation will contribute to the body of work on the spatial ecology of Atlantic menhaden, an important forage fish. Spatial analysis of both empirical data and simulation modelings will be used to accomplish this goal. Chapters 2 through 4 are structured as stand-alone publications that explore the spatial characteristics of menhaden as seen in field data collected for this dissertation. Chapter 5 uses the knowledge gained from these efforts to create a spatially-explicit individual-based model of menhaden schooling and movement. Chapter 6 is a general conclusion based on all the data chapters and discusses both the greater impacts of this dissertation and possible avenues for future research.

Chapter 2 has been published in the peer-reviewed journal *Fisheries Research* as *Sonar imaging surveys fill data gaps in forage fish populations in shallow estuarine tributaries* (Lankowicz et al., 2020). Though sonar imaging methods are gaining traction in fisheries ecology, they are still rarely used in the quantitative analysis of fish distribution. One of the objectives of this Chapter was to test the efficacy of mobile sonar imaging to observe menhaden distribution. The resulting data showed not only the capabilities of mobile sonar imaging but also offered an entirely new perspective on menhaden distribution, one of an extremely fine (<1m) spatial resolution. Generalized Additive Models were used to characterize the density of individuals across the lower Patuxent River. Individuals were present more frequently and at significantly higher densities in shallow-water (2-4m) creek habitats than in deeper-water (> 4m) areas of the river channel. The results from this initial Chapter were used to inform the spatial models of menhaden distribution in all following chapters.

Chapter 3 has been revised and resubmitted to *Fisheries Research* as *Characterizing ju-*

venile Atlantic menhaden (Brevoortia tyrannus) distribution and schooling in shallow estuarine waters (2022). The focus in this Chapter shifts in aggregative extent from individuals to schools of individuals. Schools were identified, and their morphological characteristics (length, number of individuals per school, density of individuals in the school) and spatial patterns relative to each other were recorded. Generalized Additive Models were used to test the relationships between school spatial density and patterns and a set of environmental characteristics. Water depth, sampling month, and tidal stage were found to be significant drivers of menhaden school distribution. A simple clustering coefficient was used to determine the spatial pattern of menhaden schools, and schools were found to cluster at the spatial scale of the data collection transects. These findings illustrate how menhaden school morphological characteristics, spatial density, and spatial patterns can change with fluctuations in environmental conditions, as well as between spatially close but distinct habitat types.

Chapter 4 again increases in aggregative extent, shifting from schools of individuals to clusters of schools. This Chapter also explores the variation in menhaden school spatial patterns through multiple scales. This is accomplished through the use of spatial point pattern analysis adapted for line transects. The first- and second-order structure of menhaden school distribution was determined along each line transect. First-order structure refers to the uniformity of the probability of points occurring across a sampled space. Second-order structure refers to the dependence of points on each other.

Chapter 5 shifts from empirical data to simulations. The analyses of the previous chapters have provided a wealth of information on the distribution of menhaden individuals, schools, and clusters of schools. Schooling and higher-order aggregative levels have been identified as important drivers of overall menhaden spatial distribution. Some environmental characteristics,

like water depth and tidal condition, have also been identified as important. The results of Chapter 4 generated an interesting question: would a simulation of menhaden movement based on either environmental characteristics or behavioral mechanisms look more like the empirical data? To address this question, two separate spatially-explicit individual-based models were created; one in which simulated fish respond only to changing environmental conditions, and one in which simulated fish respond only to each other. Though we are not yet able to quantify the relative impact of each mechanism and process, these results are a first step in making an enduring and challenging ecological problem more tractable.

Chapter 2: Sonar imaging surveys fill data gaps in forage fish populations in shallow estuarine tributaries

2.1 Abstract

Forage fish are economically and ecologically important to many coastal marine ecosystems, but most forage fish species are not targeted in species-specific abundance surveys. In the Chesapeake Bay, Atlantic menhaden (*Brevoortia tyrannus*), an important forage species for many upper-trophic level consumers, are sampled with beach seines and midwater trawls. There is evidence that these traditional fish sampling gears are insufficient for observing menhaden abundance and distribution, which may lead to biased and inaccurate indices of abundance. We suggest this is due to omitting or under-sampling shallow estuarine creeks and nearshore areas. We tested this hypothesis in the Patuxent River, Maryland, and its surrounding shallow-water creeks, which are representative of many estuarine tributaries with varied bathymetry. Sonar imaging surveys were conducted using Adaptive Resolution Imaging Sonars (ARIS 1800 and 3000) and a generalized additive model was fit to the resulting data to model forage fish density. Results indicated that menhaden and other forage fish were present at significantly higher densities in shallow-water creeks than in deeper-water areas. Further, many forage fish were observed in creek habitats frequently omitted by both seines and trawls, suggesting that these survey de-

signs may result in biased forage fish indices of abundance. Therefore, we recommend current forage fish surveys in the Chesapeake Bay, especially those with the intent of developing juvenile menhaden indices of abundance, be supplemented with hydroacoustic surveys in shallow waters to improve accuracy and reduce bias.

2.2 Introduction

Forage fish are small, highly abundant, and planktivorous fish species that occupy intermediate trophic levels in estuarine food webs and therefore facilitate energy transfer from lower to higher trophic level species ([Bakun et al., 2010](#)). Depletion of forage fish and other similar-trophic level prey items has significant negative impacts on upper-trophic level predators, indicating some incidence of bottom-up control ([Cury et al., 2011](#); [Smith et al., 2011](#)). As such, forage fish are integral to the diets of marine mammals, fish-eating seabirds, and several life stages of piscivorous sportfish ([Hartman and Brandt, 1995](#); [Luo and Brandt, 1993](#); [Pikitch et al., 2012, 2014](#); [Walter III and Austin, 2001](#); [Weise and Harvey, 2008](#); [Wood and Houde, 2003](#)). Forage fish are short-lived, have high fecundity, and respond rapidly to environmental factors, leading to large fluctuations in inter-annual abundance ([Pikitch et al., 2012](#)). This variability can have wide impacts on the growth, reproductive rate, distribution, and abundance of their predators ([Gislason, 2003](#)).

Forage fish also have high economic value to coastal regions. Annually, global forage fish landings are worth approximately \$5.6 billion and support other fisheries worth an additional \$11.3 billion ([Pikitch et al., 2012](#)). Several species are of particular importance: clupeids (herrings, sardines, and menhadens) are known for their valuable reduction, bait, and seafood fish-

eries. Atlantic and Gulf menhaden (*Brevoortia tyrannus* and *Brevoortia patronus*, respectively; Clupeidae) support two of the oldest and largest fisheries by volume in North America. In 2017, total US menhaden landings approached 635,000 metric tons worth over \$125 million ([National Marine Fisheries Service, 2018](#)).

In recent years, policymakers have pushed to change the management of many forage fish from single-species to ecosystem-based management methods ([Buchheister et al., 2017](#); [Pikitch et al., 2004](#)). This stems from the recognition that though a forage species may not be overfished and overfishing may not be occurring, a depleted abundance of these fish could have negative consequences across their associated ecosystems. It is therefore important to have accurate and unbiased forage fish indices of abundance to develop ecosystem-based management methods. However, many forage species do not have targeted abundance surveys and are instead sampled concurrently with larger piscivorous species. Sites selected and gear types used in these surveys are not optimized for forage fish and may lead to issues in gear selection, evasion, and overall sampling bias.

Annual abundance surveys for many fish species in the Maryland portion of the Chesapeake Bay are conducted by the Maryland Department of Natural Resources (MD DNR) in a long-term survey designed to examine juvenile striped bass abundance relative to the total time-series of collected data. This survey also generates indices of relative abundance for Atlantic menhaden, Atlantic silverside (*Menidia menidia*, Atherinopsidae), bay anchovy (*Anchoa mitchilli*, Engraulidae), Spot (*Leiostomus xanthurus*, Sciaenidae), and many other forage species. In the MD DNR survey, a beach seine 1.24 m tall and 6.4 mm mesh size is swept by hand in nearshore shallows at 22 fixed sites during daylight hours every month ([Durell and Weedon, 2017](#)). Deployment is limited to areas shallower than the height of the net, which therefore restricts where surveys can

be completed. This method has a relatively low sample volume and the potential for high avoidance. Larger, faster fish may be able to outpace the net and avoid capture [Bayley and Herendeen \(2000\)](#). Some forage species, like the bay anchovy, are small enough as young-of-year that they may not fully recruit to mesh sizes used in the MD DNR survey until late August ([Wang and Houde, 1995](#)). Other environmental factors such as turbidity and light intensity may affect seine avoidance. Several studies on seine capture efficiency have found that high-light conditions (day-light hours) may lead to lower availability and detectability of sample species ([Allen et al., 1992](#); [Blackwell and Brown, 2005](#)).

Midwater trawls are a less-common method of studying forage fish in the Maryland portion of the Chesapeake Bay. The few and inconsistent forage fish surveys completed using this method use a net with a 3.8–7.6 mm mesh body and a 3 mm mesh codend to survey only areas 4m or deeper due to limitations set by survey and vessel draft ([Jung and Houde, 2003](#); [Miller et al., 2008](#)). Studies employing these methods have specifically stated they are not as effective at observing menhaden abundance as they are for other important fish species in a similar size range ([Jung and Houde, 2003](#)). This could indicate an issue in sampling site selection. Like beach seines, trawls may also have variability in capture efficiency driven by light intensity and turbidity, as well as further issues due to vessel avoidance ([Glass and Wardle, 1989](#)).

Combining seine and midwater trawl methods to observe forage fish abundance and distribution may not solve the issues of gear and site selection, as the two methods have non-overlapping depth ranges. All areas with a maximum depth in the 1.24–4 m range, as well as this depth range of the water column, are either not sampled or not sampled effectively for forage fish. This under-sampled area includes many shallow-water creeks. In the Maryland portion of the Chesapeake Bay, few attempts have been made to sample forage fish abundance in areas

not on a main river channel or main stem of the Chesapeake Bay. It is unclear what effect the omission of creeks and other shallow regions has on forage fish indices of abundance. However, it is known that population inferences are unlikely to be robust if sampling methods are based on untested or poor assumptions of a species' distribution through habitat types ([Anderson, 2003](#)). Relative abundance estimates collected only from a few areas thought to be representative of the whole system cannot account for spatial variation and may not lead to robust inferences of fish abundance and distribution. It is therefore important to understand how fish are distributed within their habitats, as this knowledge will help in designing effective and less-biased sampling procedures.

The principal objective of this study was to quantify the distribution of forage fish across multiple depth ranges within an estuarine region. We hypothesized that forage fish are less abundant in deeper waters and more abundant in shallower habitats such as estuarine creeks, which are under-studied and are potentially a critical data gap. To address this hypothesis, we used sonar imaging methods to survey forage fish in the Patuxent River, Maryland, USA. Sonar imaging technologies are ideal for surveys in estuaries, as they are non-invasive, not limited by water turbidity, and can be deployed from small boats in both shallow and deep waters. Data collected in our surveys were modeled using spatially explicit techniques and analyzed to find patterns of distribution across habitat types, water depth, and other explanatory parameters. With this detailed understanding of spatiotemporal dynamics of forage fish distribution, we are able to provide recommendations to improve future survey methods.

2.3 Methods

2.3.1 Study area

The Patuxent River is a western-shore tributary of the Chesapeake Bay estuary and was used in this study as an example of varying habitats within a shallow-water system. The Patuxent has a total length of about 80 km and varies from limnetic conditions at its headwaters (salinity approx. 0 ppt) to mesohaline conditions at its mouth (salinity 12–18 ppt). We focused on surveying the mesohaline and tidally influenced area of the river between the Gov. Thomas Johnson Bridge and St. Leonard Creek (river km 5–12). We completed 8 surveys in the summer of 2016 (May – August) and 9 surveys in the summer of 2017 (June – August) which resulted in over 32 h of sonar imaging data and 1.4 million individual images. Two cross-river transects were sampled during each survey. The North Patuxent transect (sampling range: 1.0–21.9 m; mean depth sampled: 7.2 m) samples from the mouth of St. Leonard Creek to a fishing dock in Greenwell State Park on the opposite shore, and the South Patuxent transect (sampling range: 1.1–33.8 m; mean depth sampled: 10.7 m) samples from the mouth of Cuckold Creek to a series of breakwaters on the Calvert County shore. The similarities in the physical characteristics of the North and South Patuxent transects led us to group them into a single habitat type, which we called the river channel habitat. In the results section, we will describe how this grouping is further justified by similarities in observations of forage fish. Three major creeks within the survey area were also sampled: St. Leonard Creek (sampling range: 1.1–6.1 m; mean depth sampled: 3.9 m), Hellen Creek (sampling range: 1.1–4.1 m; mean depth sampled: 2.8 m), and Cuckold Creek (sampling range: 1.1–6.0 m; mean depth sampled: 3.8 m). These creeks are relatively narrow and shallow

environments. Like the two river channel transects, the three creek transects were similar to each other and therefore grouped and labeled as creek habitat. Together, the five transects serve to observe forage fish distribution between the river and creek environments and across a wide range of water depths.

2.3.2 Sonar imaging

Active acoustic survey methods such as side-scan sonar and echosounders have been successfully used to quantify the biomass of small pelagic fishes in many habitats, including shallow-water estuarine environments ([Boswell et al., 2010](#); [Lucca and Warren, 2018](#)). Sonar imaging is a newer active acoustic survey method, but has also been successfully used to assess fish abundance with high accuracy ([Able et al., 2013, 2014](#); [Boswell et al., 2008](#); [Rieucan et al., 2015](#)). Sonar imaging cameras create digital images by sending out sound pulses and interpreting the returning echoes and are therefore ideal tools in turbid environments like estuaries. Though these cameras are becoming increasingly precise and accessible to researchers, there are still few studies using sonar imaging to analyze patterns of aquatic species distribution across different depth ranges and habitat types. It is important to note that most published sonar imaging studies have been static point deployments, which cannot identify continuous patterns of spatial distribution. However, a few recently published studies have proven the feasibility of using mobile sonar imaging surveys to observe individual estuarine fish, which led us to believe the same methods would be suitable to quantify the spatial distribution of forage fish ([Able et al., 2013, 2014](#); [Becker et al., 2017](#)).

We used the Adaptive Resolution Imaging Sonar (ARIS) 1800 and 3000 sonar imaging cameras (Sound Metrics Corp, Bellevue, WA) to complete our surveys. These cameras can be

towed to increase spatial coverage and are capable of producing near-video quality images of the water column, which makes enumerating individual fish across a continuous transect possible. ARIS cameras also collect distance, target object size, and camera depth information, which allows us to later calculate the volume of water sampled and the depth of the water column in the camera field of view. The cameras are relatively small and can be powered via a portable generator or car battery, which allows for deployment from small boats (< 26 ft).

The ARIS 1800 was deployed in 2016 and the ARIS 3000 in 2017. The ARIS 1800 uses an identification frequency of 1.8 MHz, in which 96 beams $0.3^{\circ}(\text{H}) \times 14^{\circ}(\text{V})$ make up a $28^{\circ}(\text{H}) \times 14^{\circ}(\text{V})$ field of view capable of sampling objects up to 15 m from the camera lens. The ARIS 1800 can also use a detection frequency of 1.1 MHz, in which 96 beams $0.5^{\circ}(\text{H}) \times 14^{\circ}(\text{V})$ make up a $28^{\circ}(\text{H}) \times 14^{\circ}(\text{V})$ field of view capable of sampling objects 15–35 m from the camera lens. The ARIS 3000 uses an identification frequency of 3.0 MHz, in which 128 beams $0.2^{\circ}(\text{H}) \times 14^{\circ}(\text{V})$ make up a $30^{\circ}(\text{H}) \times 14^{\circ}(\text{V})$ field of view capable of sampling up to 5 m from the camera lens. For depths sampled 5–15 m from the camera lens, the ARIS 3000 maintains the same number and dimensions of beams but uses a detection frequency of 1.8 MHz. As with most imaging technology, image resolution decreases with increasing imaging range. We chose to use the ARIS 3000 in 2017 over the ARIS 1800 because it produces images with higher resolution, though it sacrifices sampling range. Surveys completed using the ARIS 1800 primarily used its identification frequency (sampling range < 15 m) and had an average resolution of 5.8 mm per pixel. Surveys completed with the ARIS 3000 primarily used its detection frequency and had an average resolution of 2.9 mm per pixel. We do not expect there to be a significant change in detection efficiency due to differences in camera model and image resolution. All recorded images of forage fish had an image resolution sufficient for a human expert to count the individual

fish present. ARIS surveys were initiated in 2016 with two transects across the lower Patuxent River (labeled North and South) and one transect in St. Leonard Creek (Table 2.1). In 2017, we continued sampling the same three transects and expanded our spatial extent to also include Hellen Creek and Cuckold Creek (Table 2.2). We limited the camera range to 15 m for the majority of images collected in 2016 to maximize the tradeoff between image quality and the volume of water sampled. Most forage fish were observed to be in the upper part of the water column, and so in 2017 we further refined this maximum imaging distance to 6 m. The ARIS was mounted to a pole fixed to the side of a small research vessel. All research vessels used in this study had a draft shallower than 0.5 m. When sampling, the ARIS was held approximately 0.25 m below and at a 35-50° angle from the surface for an increased field of view compared to a vertical deployment. On average, we collected 12–15 images per second. While collecting data, the research vessel moved at 0.5–1 m s⁻¹ to avoid compromising image quality. A handheld GPS unit (Garmin GPSMAP 64 s, Garmin Ltd, Olathe, KS) recorded location every 10-15 s. A CTD (RBR Ltd, Ottawa, Canada) continuously recorded surface water temperature and salinity. In model selection stages, temperature and salinity were dropped as explanatory variables. This was due to both a lack of significance and incomplete data, as the CTD was deployed on only 11 of the 17 cruises completed in the study period.

2.3.3 Data processing

After collection, ARIS data were downloaded and split into individual files based on date and transect. Each file, which consisted of thousands of individual images, was then analyzed in ARISFish software. Images were sharpened and false-colored to increase the contrast between

objects and the background for best definition and visibility, then manually inspected for forage fish (Figure 2.1). Schooling behavior, school size, and samples of individual fish sizes were all used to identify appropriate forage fish species and exclude all other fish species we observed in our surveys. Images were filtered to remove objects with less than 3 cm² visible surface area. This size filter is large enough to remove noise, yet small enough to retain most images of juvenile bay anchovy, the smallest forage fish of interest. Age 1+ anchovy in the Chesapeake Bay are expected to have an average body length near 57 mm in June. A substantial portion of bay anchovy young-of-the-year (YOY) are hatched May-June, and on average YOY individuals hatched by mid-July will reach approximately 19 mm body length and recruit to the survey gear in August (Wang and Houde, 1995). Though individual YOY bay anchovy would be difficult to see in filtered images, it is easy to identify the dense schools they move in, at which point appropriate image magnifications can be made to count individual fish. The number of forage fish in every frame was recorded. ARIS files were then read into MATLAB (MathWorks, Natick, MA), where header information including camera depth, start range (closest distance from the camera lens in the field of view), end range (furthest distance from the camera lens in the field of view), pitch (angle of deployment), and pixel resolution were extracted for each frame using code developed and provided by Sound Metrics. GPS data were downloaded and interpolated to a temporal resolution of 1 s.

River distance, or the shortest navigable distance between each point of a transect and a set source point, was calculated in ArcMap 10.1 (ESRI, Redlands, CA). Source points for creeks were set near the head of each creek. Source points for the two river transects were set at the mouth of the nearest creek. We expected forage fish abundance to be concentrated in the shallowest portions of the creeks (which corresponded to further distances from the creek mouth)

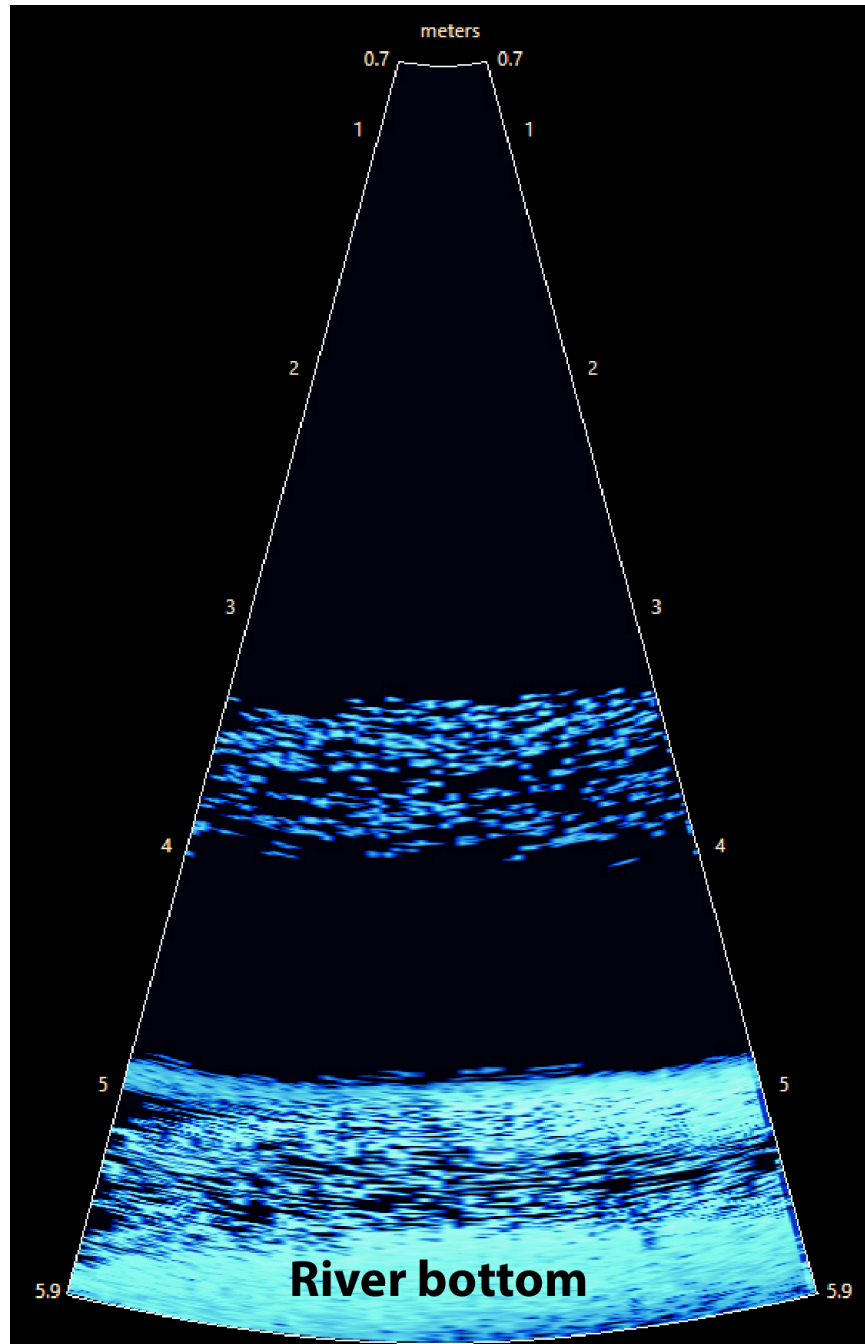


Figure 2.1: Sample output of ARIS imaging data. Data are sharpened, filtered, and false-colored using the same procedure as developed during image processing. Numbers down the edges of the image indicate distance away from the camera lens in meters. The wide, bright strip at the bottom of the image is the river bottom. The objects 3.5–4.0 m from the camera lens are forage fish between 7 and 10 cm long.

with relatively fewer fish dispersing to and remaining in deeper river channel habitat.

Echograms, a continuous image of the camera field of view over a transect where one-pixel width represents one frame, were extracted for each file using ARISFish software. It is important to note that echograms are line-of-sight and are therefore angled at the pitch of the sonar camera for each frame. They do not show the vertical depth of the water column but are used in the process of calculating this parameter. Echograms were adjusted in Adobe Photoshop (Adobe, San Jose, CA) so the river bottom, if visible, was represented by fully white pixels and the water column was fully black pixels. Adjusted echograms were read into MATLAB. Image height in pixels (IH) was recorded and the height of the river bottom in pixels (BH) was calculated per frame. From here, the visible range (VR) of the field of view was calculated based on IH, BH, range end (RE), and range start (RS, Eq. 1). The vertical depth of the water column (WD) was then calculated based on VR, sonar pitch, and the camera's depth when deployed (Eq. 2). Note that ARIS cameras record pitch as relative to gravity and negative in the "nose-down" direction so that a camera pointed straight down records a pitch of -90°.

$$VR = \frac{(IH - BH) * (RE - RS)}{IH} + RS$$

$$WD = [\cos(90^\circ - \text{absolute value}(\text{sonar pitch})) * VR] + \text{camera depth}$$

For images in which the river bottom was deeper than the maximum range of the camera or technological issues prevented sonar pitch and camera depth from recording, vertical water depth was estimated from an interpolated map of soundings created from NOAA hydrographic survey 6877.

Volume per frame was calculated as a truncated pyramid with a rectangular base to account for the angled field of view. This calculation was based on the visible range, the minimum

range, the distance the camera traveled in 10 s, and the known angles of the three-dimensional field of view. More detail on calculating volume can be found in Shahrestani et al. [Shahrestani et al. \(2017\)](#). Once image processing was complete and all parameters were calculated, resulting frame-by-frame datasets were merged in R (version 3.5.3). Parameters were averaged over 1 s, then aggregated into bins with a width equal to the average distance traveled in 10 s. This was chosen as the bin size because it is approximately the amount of time it takes for the research vessel, traveling at $0.5\text{--}1\text{ m s}^{-1}$, to pass completely through and out of the area seen in the initial field of view. This binning process also helped to reduce artificial population inflation caused by counting the same fish in sequential frames. In all ARIS observations, less than 2% of forage fish schools were in view for more than 10 s ($n = 51$). The density of fish in each bin was then calculated as the mean count of fish per 1 m^3 sampled.

2.3.4 Statistical analysis

After data processing, relationships between forage fish density and explanatory variables were noted to be non-linear and have nested and complex spatio-temporal dependence structures. The high spatial resolution and complex spatio-temporal dependencies of the dataset cannot appropriately be analyzed using classical design-based inference methods ([Thompson, 2012](#)). Data were heavily skewed and heteroskedastic due to the patchy distribution of schooling forage fish and could not be transformed to meet normality and homoscedasticity assumptions. Consecutive 10 s aggregated bins within the same transect were spatially autocorrelated and unlikely to represent true replicate draws of the forage fish population, resulting in a violation of independence assumptions. In the classical sampling design sense, each transect should be treated as a sample

of size one from its corresponding habitat stratum. However, this requires the unrealistic assumption that temporal pseudoreplication is not a concern and confounding of sampling area and time does not occur. For these reasons, classical design-based statistics are not appropriate to analyze these data.

We reasoned that a model-based approach would be more suitable for disentangling the complex and nested spatio-temporal dependence structures of the data. A generalized additive model (GAM) was chosen as the modeling framework due to its ability to represent the relationship between a response variable and multiple non-linear and unknown smooth functions of predictor variables. An appropriately fitted GAM can provide quantitative analysis for highly skewed, zero-inflated data with complex spatial and temporal dependence structures. Therefore, a GAM was fitted to the data to summarize the interaction of sampling date, sampling month, water depth, and river distance as predictor variables. The abundance of forage fish was chosen as the response variable (Eq. 3). The fitted GAM used a negative binomial distribution family due to the zero-inflated and overdispersed nature of the forage fish count data. Restricted maximum likelihood (REML) methods were used to find the smoothness of each predictor variable and penalize overfitting, as this method is less prone to being influenced by local minima. The model selected was:

$$\text{Fish abundance} \sim s(\text{sampling event}) + s(\text{Month}) + s(\text{River distance}) + s(\text{depth}) + \text{offset}(\log(\text{volume}))$$

In this model, $s(\dots)$ is a cubic spline smoother function. Sampling dates were labeled 1–17 in chronological order, creating the sampling event variable. Recall that water depth does not refer to the depth at which individual fish were observed, but rather the averaged depth of the water column over an aggregated 10 s bin. The model was stratified by habitat type, as data

exploration at this stage showed a divergence in patterns of fish density as related to predictor variables between the two habitats. An offset of log-transformed volume was included to account for variability in sampling effort between 10 s bins, and effectively normalized our abundance response variable by volume of water sampled, resulting in a measure of forage fish density.

An artificial dataset was created to begin using model-based inferences to test the statistical significance of the observations. Simulated draws of forage fish density from the fitted GAM represent the dependence structures present in the observed data. Moreover, it is possible to filter out confounding factors and focus on identifying the influence of a single factor at a time. Model-based comparisons of simulated forage fish density across spatial groups can be conditioned to remove temporal effects and will follow more realistic assumptions than their design-based equivalents. Temporal effects can also be considered with model-based comparisons of fish density across sampling events. The artificial dataset included frame-by-frame latitude, longitude, river distance, water depth, and sampled water volume values from a hypothetical sampling day of all five transects. River distance values were calculated using the same methods as the observed dataset. The water depth of each frame was interpolated from NOAA hydrography survey 6877. The artificial dataset, like the observed dataset, was aggregated into 10 s bins and average river distance and water depth were calculated per bin. The water volume sampled for each artificial aggregate bin was then calculated as a truncated rectangular pyramid. Water volume sampled was not allowed to exceed the maximum volume sampled per bin in the actual observations. This would correspond to when the river bottom was deeper than the camera's maximum viewing distance (approximately 15 m), and therefore sampling volume would be at its maximum possible value.

To test for differences in forage fish density in spatial groups without confounding tempo-

ral effects, the month and sampling event of the artificial dataset points were set to values where GAM output showed their respective smoothers were equal to zero. This corresponds to values of sampling event and month where forage fish density of the study area was equal to the average density across all temporal replicates, which effectively removes the temporal variation and nested effect of sampling event within sampling month. We then tested for differences in forage fish density across all five transects, the two habitat types, and three depth strata. Depth strata were set to represent the functional ranges of beach seines and midwater trawls. The shallowest depth stratum (< 1.24 m) corresponds to nearshore shallows that are best sampled with beach seines. The deepest depth stratum (> 4.00 m) are best sampled with a midwater trawl. The remainder is the area with a maximum depth between 1.24 and 4.00 m and is not well sampled by either method, but can be observed using the sonar imaging camera.

Table 2.1: Monthly ARIS sampling output for 2016 as described by sampling events, spatial bins observed, median forage fish density, mean forage fish density, variance, volume sampled, and habitat type. Transects sampled were St. Leonard Creek (SLC), the North Patuxent (NPX), and South Patuxent (SPX). Hellen Creek (HLC) and Cuckold Creek (CUC) were not sampled until 2017.

2016								
Month	Transect	Sampling events	Spatial bins	Median dens. (fish m^{-3})	Mean dens. (fish m^{-3})	Var.	Vol. (m^3)	Habitat
June	SLC	2	504	0.000	0.022	0.009	28038	Creek
	NPX	2	192	0.000	0.001	0.000	45505	Channel
	SPX	1	100	0.000	0.002	0.000	20381	Channel
July	SLC	3	1016	0.000	0.012	0.004	79833	Creek
	NPX	3	587	0.000	0.004	0.001	82276	Channel
	SPX	3	480	0.000	0.006	0.002	39520	Channel
Aug.	SLC	2	660	0.000	0.018	0.00	56650	Creek
	NPX	2	426	0.000	0.002	0.00	39177	Channel
	SPX	2	328	0.000	0.000	0.00	32688	Channel

The predicted fish density of the artificial dataset was simulated 999 times from the fitted GAM. Results were then split by transect, and the average density of forage fish in each transect was found for all 999 simulations. A model-based algorithm was then used to test the difference

Table 2.2: Monthly ARIS sampling output for 2017 as described by sampling events, spatial bins observed, median forage fish density, mean forage fish density, variance, volume sampled, and habitat type. Transects sampled were Hellen Creek (HLC), Cuckold Creek (CUC), St. Leonard Creek (SLC), the North Patuxent (NPX), and South Patuxent (SPX).

2017								
Month	Transect	Sampling events	Spatial bins	Median dens. (fish m ⁻³)	Mean dens. (fish m ⁻³)	Var.	Vol. (m ³)	Habitat
June	HLC	3	403	0.000	0.027	0.030	11583	Creek
	CUC	3	406	0.000	0.024	0.015	18537	Creek
	SLC	3	804	0.000	0.011	0.006	43695	Creek
	NPX	3	633	0.000	0.001	0.000	41275	Channel
	SPX	3	511	0.000	0.001	0.000	27571	Channel
July	HLC	3	376	0.014	0.101	0.083	12864	Creek
	CUC	3	394	0.019	0.076	0.019	25433	Creek
	SLC	3	743	0.000	0.043	0.012	49023	Creek
	NPX	3	395	0.000	0.003	0.000	29543	Channel
	SPX	3	468	0.000	0.008	0.001	35560	Channel
Aug.	HLC	2	261	0.136	0.247	0.095	5761	Creek
	CUC	2	292	0.10	0.134	0.176	13495	Creek
	SLC	3	785	0.000	0.114	0.127	36804	Creek
	NPX	2	454	0.000	0.008	0.001	31985	Channel
	SPX	2	385	0.000	0.012	0.014	28810	Channel

in forage fish density between transects. First, all transects were combined and the 100% confidence limits for the overall mean were computed. Next, the probability of all transect means being within the overall confidence limits was computed to estimate the probability of all transect means being equal (p). Finally, pairwise comparisons of transect means were applied using the same simulation method with Bonferroni adjustment for multiple testing. The same procedure was applied to test for differences between river channel and creek habitat and between the three depth strata.

The effects of nested temporal variation were also tested using model-based simulations of forage fish distribution. The artificial dataset was replicated 17 times to represent the 17 sampling days analyzed and included in the GAM. Each dataset replicate was given a unique sampling day identification number, and replicates were nested in sampling month to duplicate

how the sampling was carried out. For example, artificial sampling day 1 was nested within May and artificial sampling days 4, 5, and 6 were nested within July. The predicted fish density of each artificial sampling day was simulated 999 times from the fitted GAM. Results were split by sampling day, then each sampling day was split into the river channel and creek habitats. This was done after results from the tests of spatial effects showed significant differences between the two habitat types which could not be removed from the test of temporal effects. The average density of forage fish in each group was found for all 999 simulations. Pairwise differences in forage fish density for each habitat and sampling day combination were then tested using the same procedure as was used to test for differences in spatial groups.

2.4 Results

2.4.1 Distribution between habitats

Through two years of ARIS deployments, we surveyed over 870,000 m³ of water and counted 3151 schools of forage fish. Though the ARIS can spatially resolve images enough to identify individual fish, it cannot provide enough detail while being towed to identify down to a species level. Instead, visual observations of fish at the surface while towing, previous stationary studies ([Shahrestani et al., 2017](#)), schooling behavior ([Zhang et al., 2017](#)), and seasonal length information allowed us to distinguish between larger piscivorous fish, juvenile menhaden, and smaller forage fish. Smaller forage fish are expected to be mostly bay anchovy and Atlantic silversides, which are highly abundant in the Bay ([Gallagher et al., 2018](#); [Jung and Houde, 2003](#); [Wood and Houde, 2003](#)). Of all our observations of fish, juvenile menhaden were the most frequently observed, which is unsurprising given that they are also considered to be among the most

highly abundant species in the Chesapeake Bay (Jung and Houde, 2003; Wood and Houde, 2003).

The majority of spatial bins with forage fish present ($n = 2148$) were in the three creek transects (Figure 2.2), (Figure 2.3). Analyzing forage fish density in the whole of the study area showed large spatial and temporal variation in the distribution of forage fish between the creek and river channel habitats (Table 2.1 and Table 2.2). Forage fish density ranged from 0.0–6.3 individuals m^{-3} with a mean density of 0.05 ± 0.20 individuals m^{-3} in the creek transects. In the river channel transects, density ranged from 0.0–1.8 individuals m^{-3} with a mean density of $< 0.001 \pm 0.04$ individuals m^{-3} .

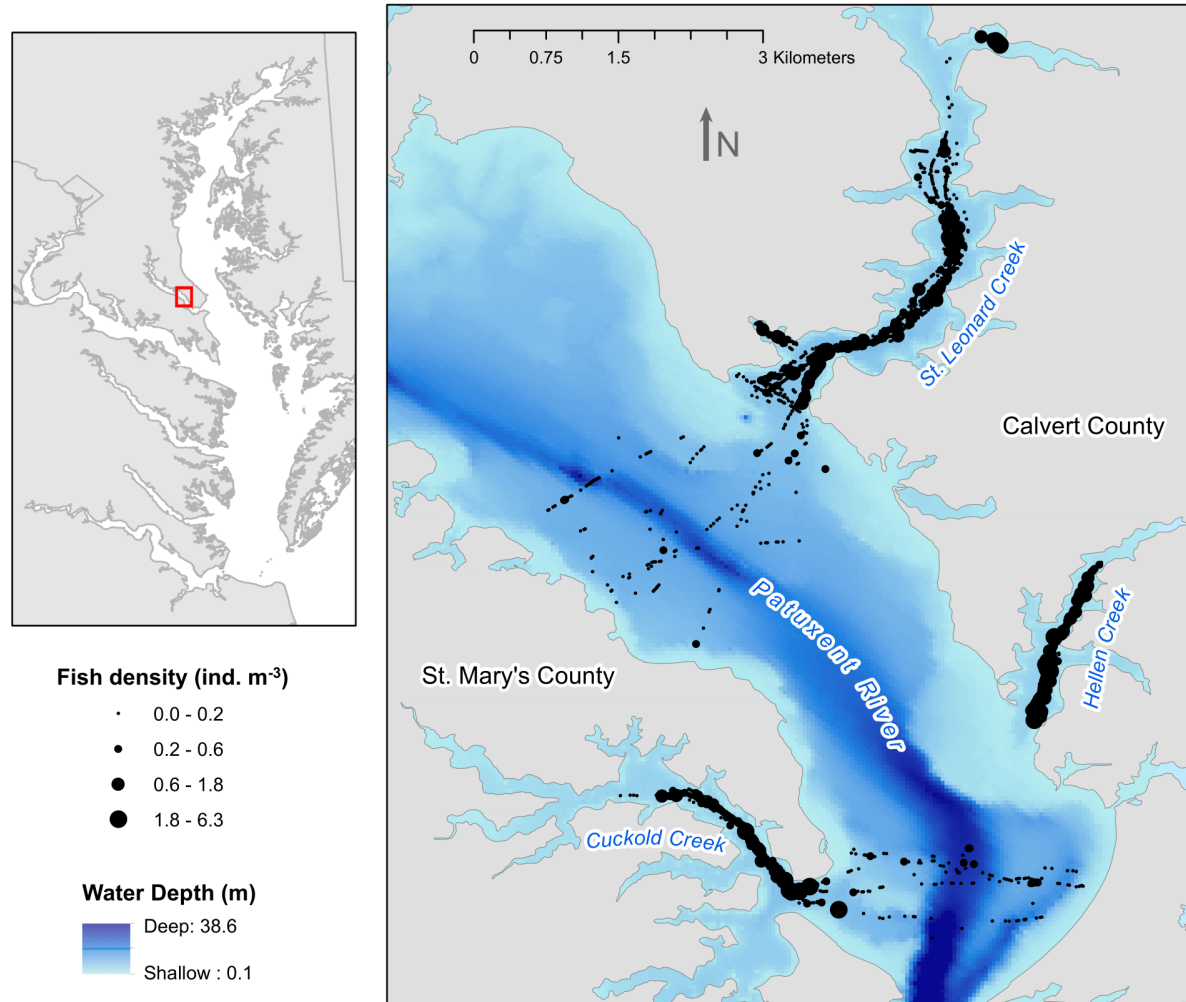


Figure 2.2: (a) Map of the Chesapeake Bay. Red bounding box indicates the study site in the lower portion of the Patuxent River. (b) Composite forage fish distribution through the Patuxent River for the study period. Black dots represent an observation of > 0 individual fish m^{-3} and dots are sized on a graduated scale where larger dots indicate a higher density of fish. Water depth is reported as mean low water interpolated from NOAA hydrographic survey 6877.

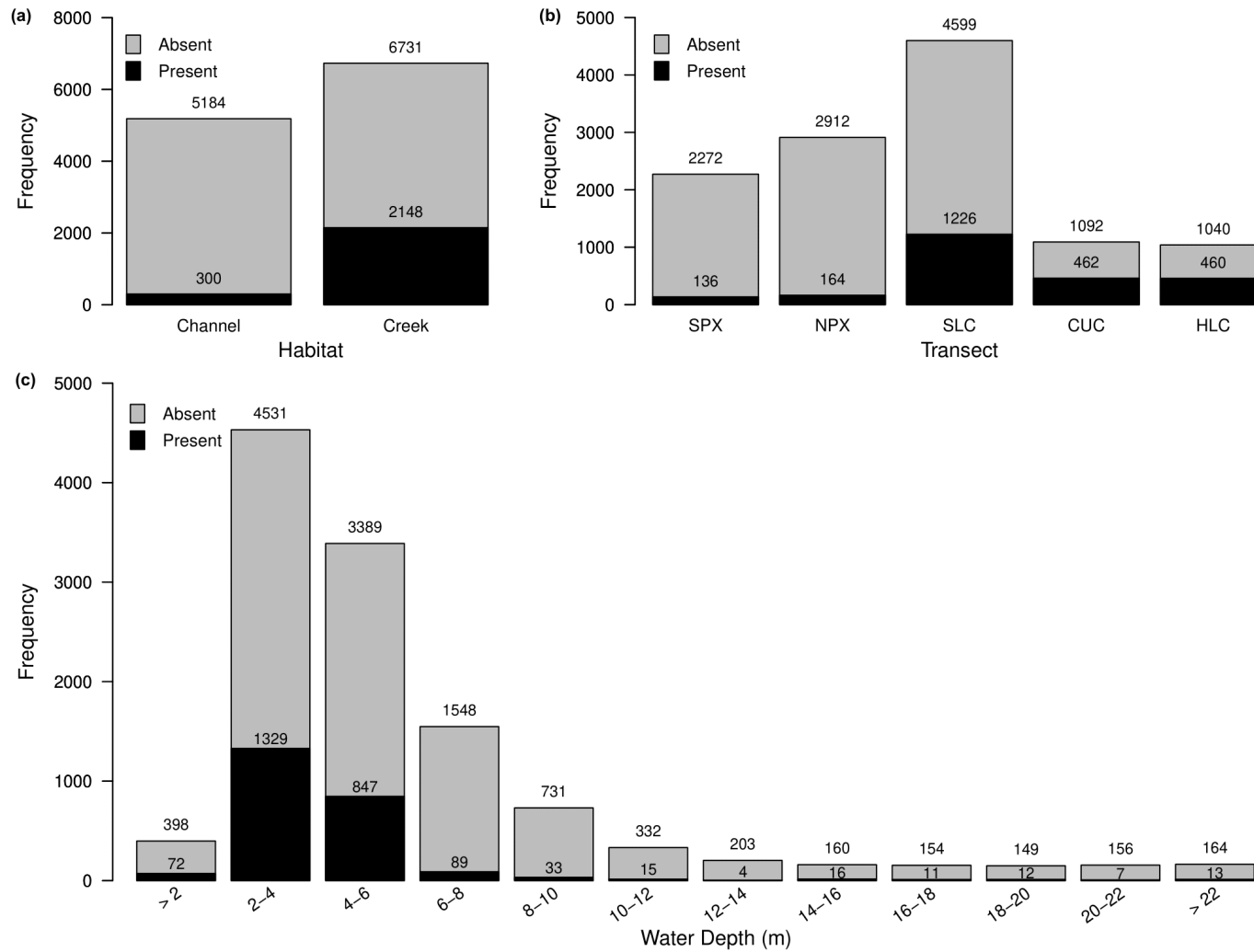


Figure 2.3: (a) Frequency of forage fish presence in 10 s bins of the creek and channel habitat. (b) Frequency of forage fish presence in 10 s bins of each transect. Transects included are the South Patuxent (SPX), North Patuxent (NPX), St. Leonard Creek (SLC), Cuckold Creek (CUC), and Hellen Creek (HLC). (c) Frequency of forage fish presence in 10 s bins across the vertical water column.

The generalized additive model allowed us to characterize forage fish distribution across the spatial and temporal range of the study. The output of GAM smooth functions is centered on the mean response variable taken over the set of predictor variables. The mean is then set to zero with variation in the response variable seen as positive or negative departures from this centered mean. Explanatory variables remaining in the final model explained 38.5% of deviance, and each explanatory variable had a smoother term significantly different from zero ($p < 0.001$), therefore contributing to the overall model fit. As stated earlier, the interactions between forage fish density and predictor variables differed between river channel and creek habitats.

The effects of river distance on forage fish density varied widely between the river channel and creek transects. Observations were taken 0.01–2.61 km from source points in the river channel habitats and 0.11–3.90 km from source points in the creek habitats. In all points of the river distance range in the river channel habitat, forage fish density was less than the mean of the whole study area (Figure 2.4). Creek results were more mixed; areas within 0.75 km of the head of the creek had higher forage fish density than the mean, as did areas 2.79–3.73 km away from the creek head (Figure 2.4). This indicates a bimodal distribution, with more fish being found in the head and mouth areas of the creeks and fewer in the middle.

Simulated forage fish density was found to differ significantly between the creek and channel habitats in the artificial dataset ($p < 2e^{-16}$). In all 999 simulations, the density of forage fish was always higher in the creeks than in the river channel. We then looked at pairwise comparisons of fish density across all five transects and found the same pattern. There is a 0% probability of all means being equal. Transects grouped into the same habitat type did not differ significantly from each other but often differed significantly from the transects in the other habitat. The only pairwise comparison between a creek transect and river channel transect to not have a significant

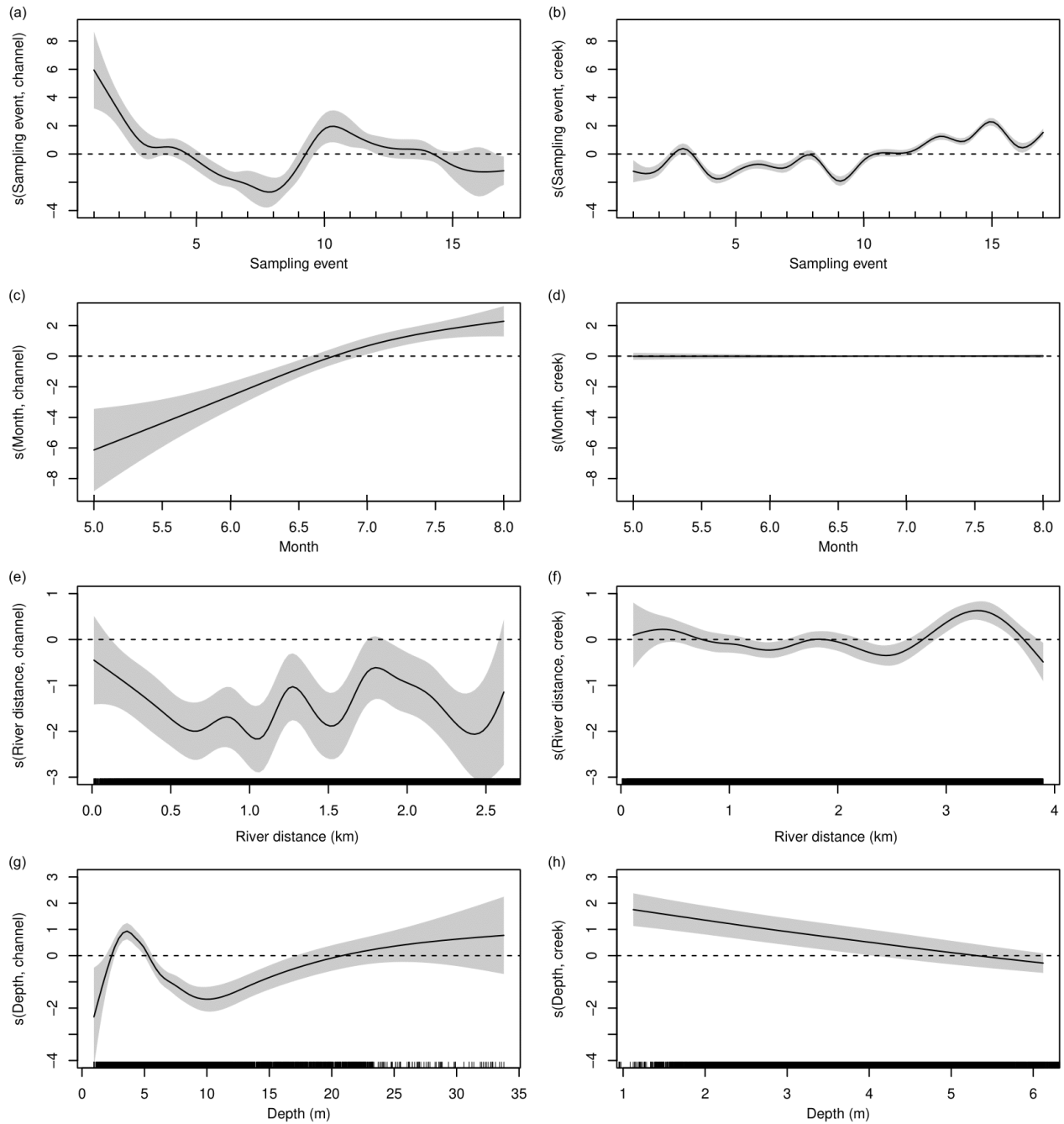


Figure 2.4: Sampling event, month, river distance, and depth smoothers for the generalized additive model. Channel habitat smoothers are in the left column and creek habitat smoothers are in the right column.

difference is that between Hellen Creek and the South Patuxent. St. Leonard Creek and Cuck-old Creek transects had strong significant differences from both the North and South Patuxent

transects (all $p < 0.04$).

2.4.2 Distribution between depth strata

Sampled water depth in the entire study area varied between 1 and 34 m with a mean sampling depth of 6 ± 4 m. Most observations occurred in waters 2–4 m deep ($n = 4531$, 38.1% of total observations) (Figure 2.3). Most positive observations both in total number and proportion also occurred in waters of this depth ($n = 1329$, 29.3% of positive observations) (Figure 2.3). The distribution of observation points was not equal across water depths and relatively few observations were taken in waters deeper than 10 m, but proportionally fewer fish were seen in these deeper areas.

Water depth at the point of observation also had a significant effect on the density of forage fish. Sampled areas in the river channel transects were 1–34 m deep with a mean sampling depth of 9 ± 5 m. Here, forage fish density was highest in waters 2–5 m deep (Figure 2.4). There is another peak in density at depths greater than 21 m, but this effect is likely exaggerated due to the small sample size, especially in waters deeper than 23 m. This assumption is supported when looking at the 95% confidence interval for deeper waters; it extends both above the highest density seen in the habitat and below the mean density for the entire study area. Sampled areas in the creek transects were 1–6 m deep, with a mean sampling depth of 4 ± 1 m. The effect of water depth on fish density in the creeks was much more straightforward, with a positive effect on density seen for all areas less than 5 m deep (Figure 2.4).

Pairwise differences in fish density in the artificial dataset across depth strata representative of different sampling methods were then tested. Because habitat type was an important explana-

tory factor for forage fish density, we split the simulated data into six groups— shallow (0.1–1.24 m), medium (1.25–4.00 m), and deep (> 4.00 m) — within both the creek and river channel habitats. Pairwise comparisons of these groups showed that the shallow channel group differed significantly from all other groups; this is likely because only 2 of 999 simulations produced a non-zero result. The medium channel group differed significantly from only the shallow channel group, and the deep channel group differed significantly from all groups except the medium channel group. The shallow creek group had the highest variance and group mean of all groupings and was significantly different from all groups except the medium channel. The medium and deep creek groups were significantly different from all groups except the medium channel group and each other.

2.4.3 Seasonal variation

Forage fish density throughout the whole study area varied monthly. Mean fish density in May was 0.001 ± 0.006 individuals m^{-3} . Density increased to 0.005 ± 0.040 individuals m^{-3} in June, 0.012 ± 0.050 individuals m^{-3} in July, and 0.027 ± 0.106 individuals m^{-3} in August. Using sampling event as a predictor variable allowed us to roughly examine differences in forage fish density across the two sampling years. In 2016, 8 sampling events were completed. In 2017, 9 sampling events were completed. The sampling event smoother function included in the GAM was given 17 knots to match the 17 events. In the river channel, forage fish density was higher than the mean density across the study period for May through early June sampling events in 2016 (sampling events 1–4) and mid-June through late July sampling events in 2017 (sampling events 9–14). Note that in 2017, no samples were taken in May. The first two cruises in 2016,

(sampling events 1 and 2) had the highest forage fish density over the mean. Creek habitat showed a different pattern of fish density across sampling events. In 2016, only sampling event 3 showed a higher forage fish density than the mean. In 2017, all sampling events after June 8 (sampling event 9) had higher fish density than the mean. Recall that only one creek transect was sampled in 2016, while in 2017 sampling was expanded to include three total creeks. When data included in the GAM is limited to only the three transects taken in both sampling years, the pattern of results does not change.

The month smoother function was included in the GAM to examine the variation of forage fish density across each sampling month without considering annual variation. It was given 4 knots, therefore allowing each sampling month to influence the overall pattern of the smoother function. In the river channel habitat, forage fish density was lower than the mean density across the study period until late June (Figure 2.4). After approximately June 20th, fish density reached the mean across the study period and continued to climb at a gradual rate. Forage fish density in the creek habitats is approximately equal to the mean fish density for all four months but is slightly less than the mean until early July (Figure 2.4). After this point, fish density increases slightly over the mean density. When these results are checked against a GAM with data input limited to the three transects taken in both sampling years, the pattern again does not change.

Comparing pairwise differences in fish density through the nested sampling event and month variables in the artificial dataset is more difficult due to the sheer number of groups that must be compared (34 total groups, 17 sampling events per habitat). However, it is immediately clear from Figure 2.4 that fish density in the creek habitats was higher in the latter half of the 2017 sampling season compared to the other sampling events. In river channel habitats, fish density was highest in the early sampling events in 2016, with a smaller density peak in the early

sampling events in 2017. The artificial dataset showed the 10 highest group densities to all be within the creek habitat. The highest of all group means is seen in the creek habitat on the 15th sampling event, which corresponds to mid-August 2017. Pairwise comparisons of all sampling events for both habitat types can be found in the supplementary materials (Tables SM1, SM2, SM3).

2.5 Discussion

The distribution and density of forage fish across the Patuxent River waterscape were found to vary significantly with temporal (monthly and annually) and spatial (water depth, habitat, and distance from “source points”) factors. The highest fish density and the majority of non-zero density patches were observed in areas with a maximum depth between 1 and 6 m, most of which are within the creeks. Creeks are rarely sampled for forage fish in the Maryland portion of the Chesapeake Bay, but our surveys provide compelling evidence that they are important habitats. Regardless of depth, the creek habitat had a higher density and variance of the forage fish population than the river channel habitat. Though the creek and river channel habitats are adjacent and highly connected on a hydrological level, our results indicate fine-scale differences in habitat utilization. These differences should be considered when designing sampling surveys to observe forage fish abundance and distribution.

Through sonar imaging surveys and generalized additive modeling, we have found that forage fish density varies considerably across habitat type, month, river distance, and water depth. This variation is not likely to be well accounted for in sampling methods with highly restrictive deployment requirements and few sampling locations, such as beach seines. Midwater trawls are

also likely to report dampened variance of forage fish density compared to our findings due to their omission of shallow, narrow, and densely populated creek environments. Dampened variability reported in indices of abundance could lead to the forage fish populations being perceived as relatively stable and consistent when in truth they fluctuate widely on multiple spatial and temporal scales.

Though indices of abundance relative to fishing effort created across a whole ecosystem may provide a robust enough estimate of population abundance for some management efforts, there is great utility in developing an understanding of the fine-scale spatial variation in population distribution and abundance. Bias in trends of abundance could be due to spatial heterogeneity of a species' distribution, and this bias could be reduced with the inclusion of spatially explicit and high-resolution survey data ([Babcock et al., 2005](#)). This kind of information would be highly valuable in ongoing attempts to create ecosystem-based fisheries management plans for menhaden, as it is difficult to balance the yield of various fisheries with the population of menhaden left in the water for ecosystem services without an accurate picture of population size and distribution.

Variation in forage fish density among the three creek habitats could be driven by variation in hydrographic features. Slight differences in the physical properties of each creek, such as the ratio of creek mouth depth to width, surrounding land use, freshwater input, and bottom substrate may lead to variation in flow and turbidity. These factors lead to heterogeneous spatial structure in the distribution of phytoplankton and, in turn, drive spatial variation in the distribution of forage fish. For example, forage fish density is often highest in Hellen Creek. This creek is the shortest and shallowest transect sampled and is surrounded by mainly agricultural and forested land, leading to allochthonous resource input and retention. More sampling and hydrographic

modeling will need to be conducted to determine the strength of this relationship between creek hydrographic features and the density of forage fish.

Forage species such as Atlantic menhaden have long been known to have high annual variability in population recruitment. Annual variability in the strength of forage fish year classes could be affected by several climate and hydrographic factors, especially those that alter larval transport into estuarine nursery grounds and the conditions of food availability in those nurseries. Though a longer time series is needed to draw robust inferences on patterns of annual variation, our results indicate increased forage fish density across the whole study area in 2017 as compared to 2016. The difference in forage fish densities across both sampling years can be seen especially when comparing the results of the river channel month and sampling event GAM smoothers. Forage fish were present in approximately 14% of spatial bin observations in 2016 and increased to 25% in 2017. Most observations of fish in 2017 occurred in either July or August. This shift in the spatio-temporal distribution of forage fish could lead to the patterns in smoother terms seen in the GAM output.

Forage fish density was higher in all creek transects than both river channel transects, regardless of the month. Forage fish are seen in relatively low densities in all habitats early in the growing season (May and June) and increase in density through the end of the sampling period (late August). This increase in density through the season could be indicative of juvenile forage fish recruiting to the ARIS survey. Bay anchovy spawning begins in mid-May and peaks in July. Surviving larvae can grow anywhere from 0.30 to 0.85 mm day⁻¹ depending on environmental conditions and would reach the juvenile stage beginning in mid-June ([Jung and Houde, 2004](#); [Zastrow et al., 1991](#)). Anchovy spawned at the peak in mid-July would reach the juvenile stage in mid-August. Increasing forage fish density later in the summer could also be indicative of ju-

venile menhaden migrating downstream to the study sites from their oligohaline larval and early juvenile nursery sites.

As stated earlier, we expect hydrographic differences in spatially explicit areas of our surveyed area are what drive differences in forage fish abundance and distribution. A major factor defining these hydrographic differences would be water depth. Creek transects tended to be much shallower than the river channel transects and have a higher abundance of forage fish. In these creeks, forage fish density was highest in areas shallower than 5 m. Sampled areas in the creeks deeper than 5 m still had forage fish density relatively close to the mean across the whole study area. Approximately 96% of the sampled area in the creeks was shallower than 5 m, with the maximum creek depth being 6.3 m in St. Leonard Creek.

River channel habitat had very different patterns for forage fish density. Generalized additive modeling results expected areas shallower than 2 m to have a much lower density than the rest of the sampled area, albeit with high uncertainty due to low sampling frequency. Most of the river channel area is well beyond 2 m deep, with only narrow nearshore bands being this depth or shallower. This result suggests that beach seine survey methods may under-represent the abundance of some forage fish species, especially in river channel habitat. This result could be further refined with more sampling of these shallow regions, especially as there may be some difference in species composition in these areas. For example, mummichogs (*Fundulus heteroclitus*) and Atlantic silversides are likely to be most common in the shallowest habitats where seines can be effective, but menhaden and bay anchovy may not be present as frequently there. Forage fish are seen in higher densities in the deepest river channel waters (> 21 m), though again low sampling frequency leads to high uncertainty. Water this deep occurs in a thin strip with a steep slope in the center of the South Patuxent transect, as well as in a few patches in the center of the North

Patuxent transect.

This study emphasizes the importance of quantifying forage fish spatial distribution, then creating long-term survey designs from this knowledge. A baseline understanding of forage fish habitat usage and distribution in estuaries is vital as climate change and other anthropogenic forces rapidly alter these shallow, often highly urbanized regions. Nutrient loading, time-shifted temporal cues, altered precipitation patterns, and increased usage of hardened shorelines may change habitat quality even more quickly in small, shallow creeks. Climatic regime shifts such as the North Atlantic Oscillation (NAO) and Atlantic Multi-Decadal Oscillation (AMO) may also affect hydrodynamic transport processes that deliver coastal-spawning Atlantic menhaden larvae to estuarine nursery grounds ([Austin, 2002](#); [Simpson et al., 2017](#); [Wood and Austin, 2009](#)). These environmental factors may have complex, synergistic relationships that make forecasting shifts in forage fish population and habitat usage very difficult. Currently, there is some evidence of low recruitment success of juvenile menhaden stocks in the Chesapeake Bay as compared to the historic average, but there are no satisfactory explanations of the cause of this decrease ([ASMFC, 2017](#); [Houde, 2011](#); [SEDAR, 2015](#)). It is further unclear if the lower recruitment success is a matter of decreased larval transport into the Bay or decreased survival of larvae to juvenile stages once in nursery grounds. Acoustic surveys of juvenile menhaden abundance through their growing season and across their utilized habitat may shed some light on this issue.

2.5.1 Conclusions

Sonar imaging surveys give a far more descriptive and comprehensive explanation of forage fish habitat usage and density across the Patuxent River than traditional trawling and seining

methods. Though we recognize the importance of preserving the consistency of longstanding time-series data, it is clear that traditional methods omit or under-sample areas with high forage fish density, and instead likely incorporate inconsistent bias into annual relative indices of abundance. We recognize that sonar imaging surveys have limitations as well; data processing alone is a huge challenge due to the sheer number of images that must be processed. Effectively sampling extreme nearshore shallow regions is also a difficulty due to the survey boat draft and the camera's minimum focus distance requirements. Vessel avoidance rates are relatively unknown and may be spatiotemporally variable. However, advances in computer science fields such as machine learning are quickly improving the performance of automated image processing procedures, and thus it will soon be possible for sonar image data to be processed on a similar time scale to net-generated catch data. Other advances in deployment technologies like autonomous surface vehicles will likely lead to increased access to shallow areas, and concurrent video-equipped drone deployments may allow us to quantify the avoidance of fish in surface waters. Therefore, we recommend current forage fish surveys in the Chesapeake Bay, especially those with the intent of developing juvenile menhaden indices of abundance, be supplemented with sonar imaging surveys in shallow waters to improve accuracy and reduce bias.

2.6 Funding

Lankowicz and Bi were supported by the National Science Foundation [grant number 1602488].

2.7 Acknowledgements

The authors thank Edward Houde and two anonymous reviewers for their feedback and suggestions on improving this manuscript. The authors also thank Suzan Shahrestani, Linlin Wang, Nathan Hirtle, Nikelene McLean, and Gang Lin for their assistance in fieldwork and image processing.

Chapter 3: Characterizing juvenile Atlantic menhaden (*Brevoortia tyrannus*) distribution and schooling in shallow estuarine waters

3.1 Abstract

Forage fish are a major food source for a diverse array of economically valuable and culturally important predators. The spatial distribution of forage species is driven by numerous and often interacting biological and environmental factors, which leads to a patchy distribution of forage fish arranged into discrete schools. There is a lack of fine-scale *in situ* fish distribution and schooling characteristics data, particularly in shallow waters that are not sampled well by net-based gear. To address this knowledge gap, an Adaptive Resolution Imaging Sonar (ARIS) was used to observe juvenile Atlantic menhaden (*Brevoortia tyrannus*) within the Patuxent River and three connected shallow creeks in the summers of 2016 and 2017. The resulting data were used to identify schools of menhaden and determine school morphological characteristics and spatial pattern. An increasing abundance of individual menhaden and an increasing number of discrete menhaden schools were observed over the summer season. Differences in the number of menhaden individuals and schools between the creek and river channel habitats were identified; more menhaden individuals and more menhaden schools were found in the creeks. Two Generalized Additive Models (GAMs) highlighted water depth and month as important environmental

drivers of menhaden school spatial density and patterns. Finally, a simple clustering coefficient was used to determine the spatial pattern of menhaden schools within the study area, and schools were found to be clustered together rather than randomly or uniformly dispersed. This study highlights how juvenile menhaden school morphological characteristics, spatial density, and spatial patterns can change with fluctuations in population size, as well as between spatially close but distinct habitat types. These findings on menhaden could be useful in understanding seasonal and habitat-dependent distributions of similar estuarine forage fish species.

3.2 Introduction

The distribution of a species is described by the density and pattern of organisms through space. Density is a measurement of the concentration of organisms in a defined area, and spatial pattern describes whether organisms or groups of organisms are positioned in predictable arrangements. Ecological studies have found that organisms are not naturally arranged either at random or in a uniform fashion, but rather in identifiable spatial patterns such as patches and gradients ([Legendre and Fortin, 1989](#); [MacArthur, 1972](#)) The formation of these patterns implies the existence of one or more underlying repetitive processes which should be both measurable and relatively predictable ([MacArthur, 1972](#)). It is well-established that both environmental variability and biotic interactions affect species distribution by producing highly complex patterns at multiple spatial and temporal scales ([Wisz et al., 2013](#)).

Estuaries, as the confluence of fresh and marine waters, are among the most spatially diverse and dynamic ecosystems on the planet ([Vinagre et al., 2017](#)). From both an ecological and economic perspective, forage fish may be the most important organisms in estuaries and coastal

marine ecosystems. Forage fish is a general term for several small- to medium-sized, highly abundant, and planktivorous fish species that occupy intermediate trophic levels and therefore facilitate energy transfer from lower to higher trophic level species ([Bakun et al., 2010](#); [Pikitch et al., 2012](#)). They are integral to the diets of a diverse array of species including pinnipeds, cetaceans, seabirds, and several life stages of piscivorous fishes ([Hartman and Brandt, 1995](#); [Luo and Brandt, 1993](#); [Pikitch et al., 2012, 2014](#); [Walter III and Austin, 2001](#); [Weise and Harvey, 2008](#); [Wood and Houde, 2003](#)). Forage fish landings are among the most voluminous and valuable in the world; in 2006, they were estimated to be worth \$5.6 billion annually and support other fisheries worth an additional \$11.3 billion ([Pikitch et al., 2012](#)).

Schooling, or aggregation with synchronized and polarized swimming behavior ([Pitcher, 1986](#)), is common among the numerous species of forage fish ([Pikitch et al., 2012](#)). Forage fish are rarely seen alone ([Fréon et al., 2005](#)). Schooling is, therefore, an important consideration in defining the spatial patterns of forage fish. The spatial pattern and density of schools are affected by local and regional abiotic drivers such as thermal structure ([Fréon et al., 2005](#); [Kimmerer et al., 2009](#)), turbidity ([Borner et al., 2015](#); [Kimbrell and Morrell, 2015](#)), tidal condition and current speed ([Zamon, 2003](#)), habitat and habitat complexity ([Rodriguez-Pinto et al., 2020](#); [Szedlmayer and Able, 1996](#)), oxygen concentration ([Domenici et al., 2017](#)), and water depth ([Blaber and Blaber, 1980](#)). Behavioral drivers such as predation pressure and intraspecific competition for food resources also play a role in structuring fish distribution across multiple spatial scales ([Hixon et al., 2012](#)). The combination of these drivers results in “patchy” patterns of forage fish schools, with concentrations of schools found within favorable areas of heterogeneous environments.

Forage fish aggregation can be described by five levels of increasing complexity: 1) individuals, 2) schools of individuals, 3) clusters of schools, 4) concentrations of clusters and 5) the

entire stock ([Fréon et al., 2005](#)). The dynamics of these aggregative levels are highly variable and the links between them are poorly understood, in part because spatial fisheries ecology has focused on examining large shifts in abundance by averaging demographic and distribution data over large spatial scales ([Ciannelli et al., 2008](#); [Mason and Brandt, 1999](#)). Modeling efforts and laboratory experiments have provided information at the individual level ([Parrish and Turchin, 1997](#); [Viscido et al., 2004](#)), but there is limited *in situ* information at the school level ([Beare et al., 2002](#)). This is particularly true in shallow water environments, where water depth limits data collection by longstanding net-based methods ([Lucca and Warren, 2018](#)). Fine-scale distribution data needed to address this data gap is on the scale of centimeters to tens of meters, which is difficult to parse out of more aggregative data collection methods and can be challenging to collect on its own without disrupting the target species. This lack of data could cause a potential misrepresentation of the mechanisms and relationships that regulate fish survival, abundance, and distribution ([Ciannelli et al., 2008](#)), which could further impact the assessment and management of fish populations ([Beare et al., 2002](#); [Lucca and Warren, 2018](#); [Pitcher et al., 1996](#)).

Incorporating fine-scale, school-level information to what is already known about the spatial pattern, density, and extent of forage fish species could contribute to accurate assessments of population size and distribution. This information could potentially elucidate some of the underlying local and regional mechanisms that act upon spatial pattern ([Bartolino et al., 2017](#)). In turn, more detailed information about drivers of forage fish spatial pattern could assist in designing more efficient surveys and creating better predictive models of distribution shifts in a changing climate ([Bartolino et al., 2017](#)). Understanding fish spatial pattern at all relevant scales and aggregation levels also has implications for observing and understanding trophic interactions, feeding energetics, and reproduction ([Bartolino et al., 2017](#); [Curtis, 2004](#)).

The principal objective of this study was to quantify fine-scale (meters to tens of meters) changes in spatial density and pattern characteristics of juvenile Atlantic menhaden schools (*Brevoortia tyrannus*, hereafter: menhaden) with regard to both local and regional environmental drivers and behavioral drivers. Juvenile menhaden were selected as the species of interest for this research because they are among the most abundant species in the Chesapeake Bay ([Jung and Houde, 2003](#); [Wood and Houde, 2003](#)) and because they were the forage species most commonly seen at the water surface while collecting data and most commonly identified within the dataset. Discrete schools of menhaden were identified and the number of individuals per school, length per school in meters, and distance between nearest observed neighboring schools (NONS) in meters were quantified in two habitats and three sampling months.

Generalized additive models (GAMs) were constructed to identify the effects of the environmental and behavioral drivers of these fine-scale menhaden schooling characteristics. Behavioral drivers included in statistical modeling efforts were the distance between NONS ([Overholtz, 2002](#)) and the density of fish within a school (individuals per unit area) ([Haugland and Misund, 2004](#)). Local and regional environmental drivers included in statistical modeling efforts were tidal stage ([Zamon, 2003](#)), water depth ([Blaber and Blaber, 1980](#)), habitat ([Szedlmayer and Able, 1996](#)), and month ([Szedlmayer and Able, 1996](#)). Previous research conducted in the study area concluded that the density of menhaden was spatially variable between habitats and temporally variable among sampling months ([Lankowicz et al., 2020](#)).

3.3 Methods

3.3.1 Study area

This study took place in the river channel and associated creeks of the Patuxent River, which is a western-shore tributary of the Chesapeake Bay ([Figure 3.1](#)). The Patuxent River is approximately 80 km long and increases in salinity from freshwater conditions at its headwaters to mesohaline conditions at its mouth (salinity 12-18 ppt). Data collection focused on river km 5-12, which is a mesohaline and tidally-influenced area with multiple connecting shallow creeks. The Patuxent River channel was characterized by the North Patuxent and South Patuxent cross-channel transects, which sampled relatively deep waters (North Patuxent mean sampling depth: 7.2 m, South Patuxent mean sampling depth: 10.7m). The transects were similar in length. Both transects end at a creek mouth; the North Patuxent connects to St. Leonard Creek at its eastern endpoint and the South Patuxent connects to Cuckold Creek at its western endpoint. The physical characteristics of these two transects are similar enough to group them as a single habitat type, referred to here as the river channel habitat.

The three largest creeks connected to this area of the Patuxent River were also sampled: St. Leonard Creek (mean sampling depth 4.2m), Hellen Creek (mean sampling depth 2.8m), and Cuckold Creek (mean sampling depth 3.8m). The St. Leonard Creek transect was the longest in all sampling efforts. Hellen Creek was often the shortest, closely followed by Cuckold Creek. All three creeks are much narrower and shallower environments than the Patuxent River channel. Because the creeks were similar in their physical characteristics, they were grouped and referred to as the creek habitat.

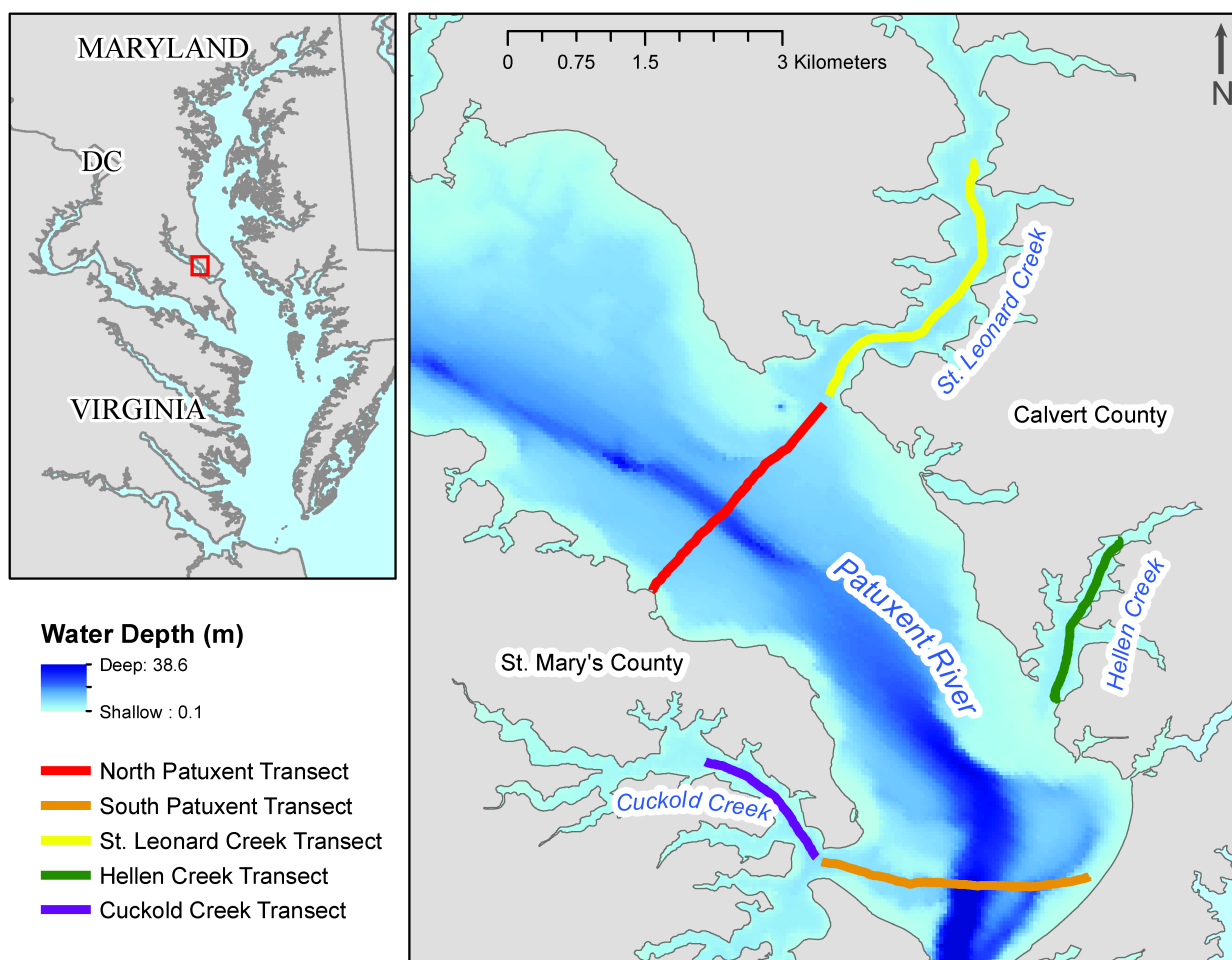


Figure 3.1: Position of the study area in the Chesapeake Bay region (A) and location of the five transects within the study area (B). The North and South Patuxent transects characterized the river channel habitat, while St. Leonard, Hellen, and Cuckold Creek transects characterized the creek habitat.

3.3.2 Data collection

Two ARIS systems were utilized during the two-year study period. The ARIS 1800 was deployed across the North and South Patuxent transects and the St. Leonard Creek transect from June to August of 2016. Transects were sampled approximately twice every month. Most data were collected using a 1.8 MHz frequency with an average downrange resolution of 5.8 mm and a maximum range of 15 m. The ARIS 3000 replaced the ARIS 1800 in the 2017 sampling

season due to its increased image resolution. Data were collected at least twice monthly from June to August of 2017 in the three original transects, plus Cuckold and Hellen Creeks. The ARIS 3000 uses an identification frequency of 3.0 MHz at ranges of less than 5m and a detection frequency of 1.8 MHz at 5-15 m. Both frequencies can resolve images in the near field to 2.9 mm. Data collected using the ARIS 3000 averaged a downrange resolution of 2.9 mm and a maximum range of 6 m. For both sampling years and camera types, 12-15 images were collected per second. More information on the specifics of each ARIS camera model and beam array can be found in Lankowicz et al. [Lankowicz et al. \(2020\)](#).

It is unlikely that the change in the camera model or maximum sampling range would affect sampling efficiency. Juvenile menhaden are expected to be 50-80 mm long (total length) in June ([Annis et al., 2011](#)), which is much larger than the average resolution per pixel (5.8 mm) of both cameras. The resolution of all recorded images is sufficient for an expert to count the individual menhaden, especially given their large size relative to resolution per pixel. The maximum camera range was reduced in 2017 because an exploratory analysis of the 2016 data found very few schools of fish deeper than 6 m. Therefore, the tradeoff in image quality for decreased sampling range was preferable, and the likelihood of falsely observing no menhaden in an area where they were present was minimal.

Both ARIS cameras were mounted on a polearm fixed to the side of a small research vessel and held approximately 0.25m deep and at a 35-50° angle relative to the surface. This angle allowed for increased sampling volume as compared to a fully vertical deployment, and unlike a fully horizontal deployment allowed for a cross-section examination of fish distribution through the vertical water column. The research vessel moved at approximately 1 m s⁻¹ while recording to avoid compromising image quality. A handheld GPS unit (Garmin GPSMAP 64s, Garmin Ltd.,

Ottawa Canada) recorded the location every 10-15 s. Once recorded, GPS data were interpolated to provide geolocation information for every ARIS image.

After data collection, tidal information for the study area was downloaded from the National Oceanographic and Atmospheric Administration's Center for Operational Oceanographic Products and Services. Water level is recorded every six minutes from the Chesapeake Biological Laboratory research pier, which is near the southern edge of the study area. Water level was interpolated to the second and used to assign every ARIS image a tidal stage identifier; either rising or falling tide.

3.3.3 Data processing

ARIS data collection results in hundreds of thousands of still images. ARISFish software (Sound Metrics Corp., Bellevue, WA) was used to view the target objects in these images and identify possible menhaden. A 3cm filter was applied to the images to remove background noise and highlight target objects that could be menhaden. One reader manually identified and counted the target images in each filtered ARIS image. Juvenile menhaden are expected to be 50-80 mm long (total length) in June and grow to 100-180 mm (total length) by the end of August ([Annis et al., 2011](#)). Objects outside of this seasonal growth curve were not considered to be menhaden and thus not counted. Further, juvenile menhaden are schooling fish and are rarely seen alone ([Fréon et al., 2005](#)). If a target object of appropriate size was tracked across sequential images but remained alone, it was not considered to be a menhaden. Only ARIS images with more than two target objects of an appropriate size that were tightly clustered and oriented in the same direction were considered to have possible menhaden. The menhaden seen in every ARIS image were

counted and recorded.

The spatial organization of individual menhaden into schools is clear in the ARIS data (Figure 3.2). Schools would first present in an ARIS image as a few densely-packed target objects oriented in the same direction. In the following several images, the number of target objects would quickly grow to be upwards of 100 individuals of similar size. Schools would pass through the ARIS field of view and eventually, all individuals would move beyond the edges of the camera range. As such, schools were defined as an aggregation of target objects of an appropriate size in a continuous series of more than two images.

After potential schools were identified, discrete schools had to be separated. A discrete school was defined as a potential school with 15 or more images devoid of menhaden immediately preceding and following it. If fewer than 15 images separated potential schools, it was assumed all imaged fish were within a single partially-sampled school, and the full shape of the school did not match the ARIS field of view. The 15-image separation results in a temporal separation of at least one second and spatial separation of at least one meter between discrete schools. This distance was chosen because the maximum voluntary sustained swim speed of juvenile menhaden has been estimated at 2.5 body lengths per second (Durbin and Durbin, 1975; Durbin et al., 1981; Friedland et al., 1984). Menhaden have been observed to swim faster, but only in laboratory flow-through experiments (Hartwell and Otto, 1978). It is unlikely that menhaden in natural environments would sustain high swim speeds along the transect paths; if they are not disturbed by the sampling method they would not likely deviate from natural behavior, and if they are disturbed by the sampling method they would move away from the perceived threat. Thus, since the largest menhaden observed were approximately 17 cm, at a swim speed of 2.5 body lengths per second the maximum displacement of an individual is approximately 42.5 cm per second. As

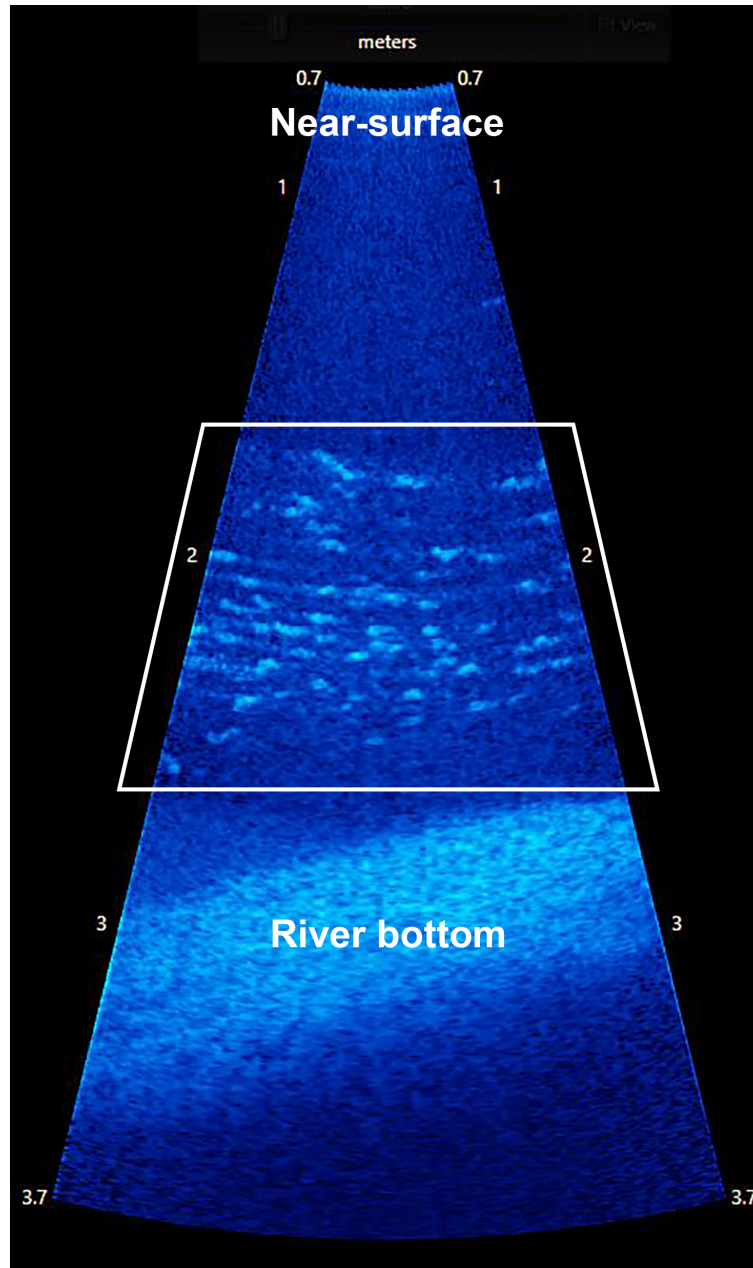


Figure 3.2: ARIS image of a partial school of juvenile Atlantic menhaden. The aggregation, bounded by a rectangle, consists of approximately 60 individuals averaging 7cm in length. Near-surface and river bottom are indicated on the image. Numbers along the edges of the ARIS field of view indicate distance from the camera lens.

the research vessel was moving approximately 1 meter per second and the ARIS was collecting 12-15 images per second, a separation of 15 images is enough to avoid labeling one school of fish as multiple schools. It also follows that as expected menhaden swim speed is slower than

the speed of the research vessel during data collection, so it is unlikely that the same school of menhaden is encountered more than once along a transect.

Accurately determining the number of individuals in each identified school requires a mechanism to correct for the artificial population inflation caused by the high temporal frequency of data collection ($12\text{-}15 \text{ images s}^{-1}$) outpacing the movement speed of the menhaden, resulting in the same individual being observed in multiple sequential images. The method chosen was a simple calculation of spatial overlap between sequential images (Figure 3.3). GPS data were interpolated to determine the distance traveled in every image. Each ARIS image is approximately trapezoidal. The area sampled per image, the minimum range of camera field of view, distance to the river bottom or maximum range of camera field of view (whichever is closer), the orientation of the camera relative to the surface and research vessel, and the internal angles of the beam array are all known. Therefore, the (x, y) coordinates of the vertices of each image can be calculated as if they are in a Cartesian plot and the overlapping space between sequential images can be determined. It is important to note that in 2016, the ARIS was oriented perpendicular to the direction of travel along a transect. This means the overlap between sequential images would occur in the wider (28°) plane of the trapezoidal prism created by the beam array. In 2017, the camera orientation was changed to be in the same direction as travel along a transect for increased image quality, causing spatial overlap in sequential images to occur in the narrower (14°) plane of the beam array. MATLAB (MathWorks, Natick MA) was used to calculate the spatial position of image vertices given the field of view dimensions, boat speed, and geographic position. Polygons representing each ARIS image were plotted and the area of overlap between sequential images was calculated. Finally, the ratio of overlap between sequential images was calculated as the area of overlapping space between two sequential images divided by the total area of the first image.

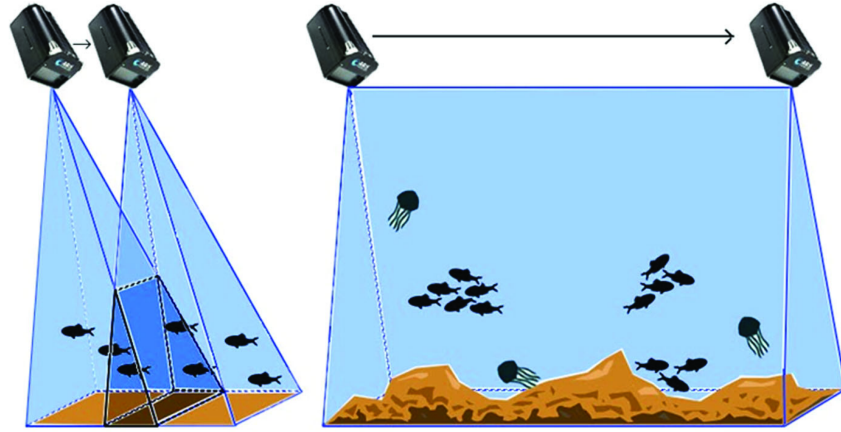


Figure 3.3: Illustration of overlapping sequential ARIS frames when imaging forage fish (A). Fish in the overlapping (darker) portion are counted in both frames, and thus a correction must be applied to remove the replication. (B) shows the outcome of this correction, in which multiple sequential frames are smoothed with areal overlap corrections to create a more accurate count of fish observed.

Calculating the number of fish in each observed school was also a multi-step process. The first, middle, and last images taken of each school were identified. It was also necessary to identify the image within each school with the highest number of individual fish, which was labeled the “maximum instantaneous abundance.” This image was used as the starting point to calculate school abundance because the total number of fish observed in a school cannot be lower than the maximum instantaneous abundance. From this starting image, “new” fish seen in sequential images were calculated by moving through the images in each school sequentially, multiplying the number of fish observed in each image by $(1 - \text{overlap ratio})$. Because the image with maximum instantaneous abundance was often in the middle of a school, this process needed to be done both forward and backward through the sequence of images that makes up a school. “New” fish seen in each image per school were summed and added to the maximum instantaneous abundance, then rounded to the nearest integer. This procedure resulted in the approximation of the number of fish observed in each school and removed over-counting attributable to a fish being

observed in multiple sequential images.

Length per school was calculated as the distance between the geolocations of the first and last images per school interpolated from GPS data. Distance between NONS was calculated in R ([R Core Team, 2021](#)) using functions in the geosphere package ([Hijmans, 2019](#)). Potential neighboring schools were limited to schools within the same transect area and sampling date. This limitation was imposed to avoid artificially inflating the distance between NONS when the closest identified school was in a different transect area, which could be separated from the target school's transect area by unsampled areas or land. If the calculated length of a school or calculated distance between NONS was equal to 0, it was considered a measurement error due to either the accuracy of the handheld GPS (reported accuracy to 3 m) or an issue with the frequency of recording data. These schools were removed from the results. Water depth at the midpoint of observation of each school was interpolated from NOAA hydrographic survey 6877 ([National Oceanographic Service, 2012](#)).

3.3.4 Statistical Methods

3.3.4.1 Generalized Additive Models

With only two years of sampling data, it is difficult to achieve a robust quantification of inter-annual variability in menhaden abundance and schooling metrics. Therefore, data from both years were pooled for analysis. Data collection and processing resulted in information on the number of fish in each school, the length of each school, and the distance between NONS. These drivers of spatial pattern are at least partially affected by fish behavioral mechanisms related to schooling, and so they were labeled behavioral drivers. Information was also available on the

habitat type, month, water depth, and tidal stage in which each school was observed, which were considered environmental drivers. The number of fish per school variable and length of school variable were found to be highly correlated (Spearman correlation = 0.68). Including both terms in statistical modeling would result in structural collinearity. Thus, for modeling purposes, a composite school density variable was created by dividing the number of fish per school by the length of the school.

Generalized Additive Models (GAMs) are flexible and commonly-used tools for modeling fish population and spatial dynamics ([Bartolino et al., 2017](#); [Brady and Targett, 2013](#); [Buchheister et al., 2016](#); [Ciannelli et al., 2008](#)). Two GAMs were created to model menhaden school spatial patterns and characteristics. The first GAM relates a set of environmental and behavioral drivers to the within-school density of menhaden. The second GAM relates these drivers to the distances between NONS. Given the wide variation in spatial scale between the local (within-school density) and regional (distance between NONS) response variables, creating two separate GAMs was more appropriate than creating a single multivariate GAM. GAMs were fit using the `mgcv` package in R according to standards set out in [Wood \(2017\)](#).

Autocorrelation, especially spatial autocorrelation, is a concern for data collected in continuous transects like the ARIS images in this study. To account for spatial autocorrelation and avoid the violation of the independence of residuals assumption, a smooth of the spatial locations was included in both GAMs. Spline correlograms with 1000 bootstrapped resamples were used to test for spatial autocorrelation as a continuous function of distance. Mantel tests were used in place of spline correlograms when there were too few schools in a transect to fit a spline correlogram.

The final variable included in the GAMs was a random effect to account for the stochasticity

of schooling. Random effects are used to model the subject- or group-specific effects without estimating an intercept for each subject or group. It was assumed each school was observed only once and therefore a variable was created to group similar schools. The number of individuals per school was chosen as this grouping variable. There is evidence that the number of individuals in a school affects phenomena as varied as feeding success (Foster et al., 2001) and the likelihood for that school to stay together or break apart (Niwa, 2004). Only 8% of all schools observed had a unique number of individuals, and the remainder of schools had up to 46 other schools with the same number of individuals. Schools with the same number of fish were grouped and assigned the same level of a categorical school identifier variable.

Models were initialized with all possible explanatory variables and interaction terms. Residuals of initial models indicated that the raw data had issues with homogeneity of variance and normality, so all numeric variables were power-transformed. Backward selection was used to remove non-significant terms. Candidates for removal were identified by examining the p-values of the GAM outputs, determining the extent to which the confidence band for a smoother term includes 0, and testing the significance of all terms. AIC and explained deviance were used to compare all models. The model with the lowest AIC, highest explained deviance, and no remaining non-significant terms was selected as the final model. Residuals of models were plotted and examined for violations of independence, normality, or equality of variance assumptions. The final models were:

Within-school density $\sim s(lon, lat) + s(water\ depth) + s(distance\ between\ NONS) + interaction(habitat, tidal\ stage) + interaction(month, tidal\ stage) + random\ effect(school\ identifier)$

Distance between NONS $\sim s(lon, lat) + s(water\ depth) + s(within-school\ density) + interaction(habitat, tidal\ stage) + interaction(month, tidal\ stage)$

where $s(\dots)$ represents a cubic spline smoother term. As stated earlier, within-school density, the distance between NONS, and water depth were power-transformed before modeling.

3.3.4.2 Cluster Coefficient

A clustering coefficient was used to characterize the spatial pattern of menhaden schools when more than 2 schools were observed in a transect on a single sampling day (Lucca and Warren, 2019). The spatial pattern of menhaden schools could not be described across the entire sampling area because the transects were not connected. Nor could spatial pattern be described at a coarser temporal grain, as it would be inappropriate to treat schools observed on different days as spatially proximate. The clustering coefficient was calculated as:

$$C_{dt} = \frac{MeanDDN}{MeanASD}$$

where C_{dt} is the clustering coefficient, mean DNN is the mean distance between NONS of all nearest-neighbor pairs within a transect on a single day, and mean ASD is the mean distance between all schools within a transect on a single day. Distances between pairs of schools were calculated in R. C_{dt} values range from 0 to 1. Values of less than 0.5 indicate a clustering pattern, while values greater than 0.5 indicate uniform distribution.

3.4 Results

3.4.1 Descriptive statistics

We collected and examined 1,478,854 ARIS images over two years, which represents approximately 32.5 hours of real-time data collection. Menhaden schools were identified in 97,496

of these images, representing approximately 7% of the collected images. After accounting for the artificial population inflation of seeing the same individual in sequential frames, a total of 327,314 unique individuals were counted in the study area. These were assigned to 2,335 discrete schools. Schools were comprised of 4 to 5,388 unique individuals, with a median of 58 individuals per school. School length ranged from 0.4 to 144.0 m, with a median length of 2.8 m. This resulted in a density of 3.1 to 599.9 individuals per meter, with a median density of 20.6 individuals per meter. Distance between NONS ranged from 1.4 m to 153.7 m, with a median of 10.0 m between neighboring schools.

Differences in these school metrics were noticed between the two habitat types and the three sampling months. More schools were observed in the creek habitat as compared to the river channel in every sampling month (Figure 3.4A). The number of total observed schools increased in July and August as compared to June. Similarly, the abundance of individuals was higher in the creek habitat as compared to the river channel in every sampling month (Figure 3.4B). Abundance increased every month from June to August. The power-transformed number of individuals per school was higher in the creek habitat as compared to the river channel habitat for every month and was relatively consistent between months (Figure 3.4C). Power-transformed length of schools was relatively consistent between sampling habitats and months (Figure 3.4D). Power-transformed within-school density was higher in the creek as compared to the river channel every month, and within-school density peaked in July for both habitats (Figure 3.4E). Power-transformed distance between NONS was shorter in the creek as compared to the river channel. In the creek habitat, the distance between NONS decreased every month. In the river channel habitat, the distance between NONS was relatively consistent between months (Figure 3.4F).

3.4.2 Generalized Additive Models

Model selection for the within-school density GAM indicated that water depth was a significant environmental predictor variable. Two environmental interactions were also significant: the interaction of habitat and tidal stage and the interaction of month and tidal stage. Distance to NONS and the random effect based on the number of individuals per school, which could both be linked to behavioral mechanisms associated with schooling, were also retained as significant predictors of within-school density. The spatial location smoother term (interaction of latitude and longitude) was also retained as significant. The explanatory variables do not have issues of collinearity. Significance for all terms in this model can be found in [Table 3.1](#). Spline correlograms, Mantel tests, and residuals from the final within-school GAM gave little evidence that the data were spatially autocorrelated, which would have violated the assumption of independence. Of the 48 sampled transects with more than two schools of fish, only 4 had residuals that indicated spatial autocorrelation. The relationship of within-school density to the explanatory variables was plotted ([Figure 3.5](#)).

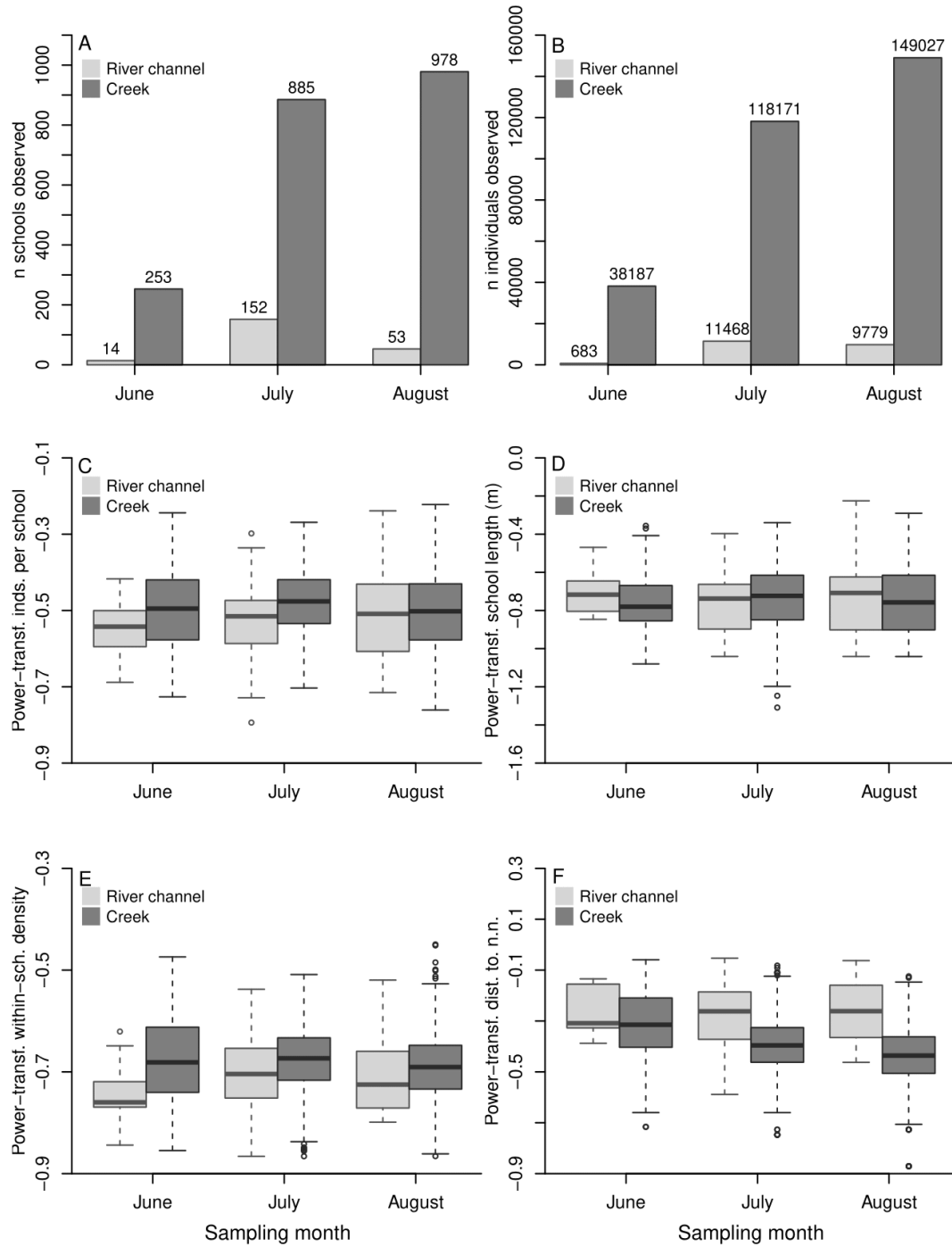


Figure 3.4: Monthly comparisons of juvenile menhaden school morphological characteristics by habitat. A. Schools observed per month and habitat. B. Unique individuals observed per month and habitat. C. Power-transformed number of individuals per school observed per month and habitat. D. Power-transformed length of schools observed per month and habitat. E. Power-transformed within-school density observed per month and habitat. F. Power-transformed distance between nearest observed neighboring school per month and habitat.

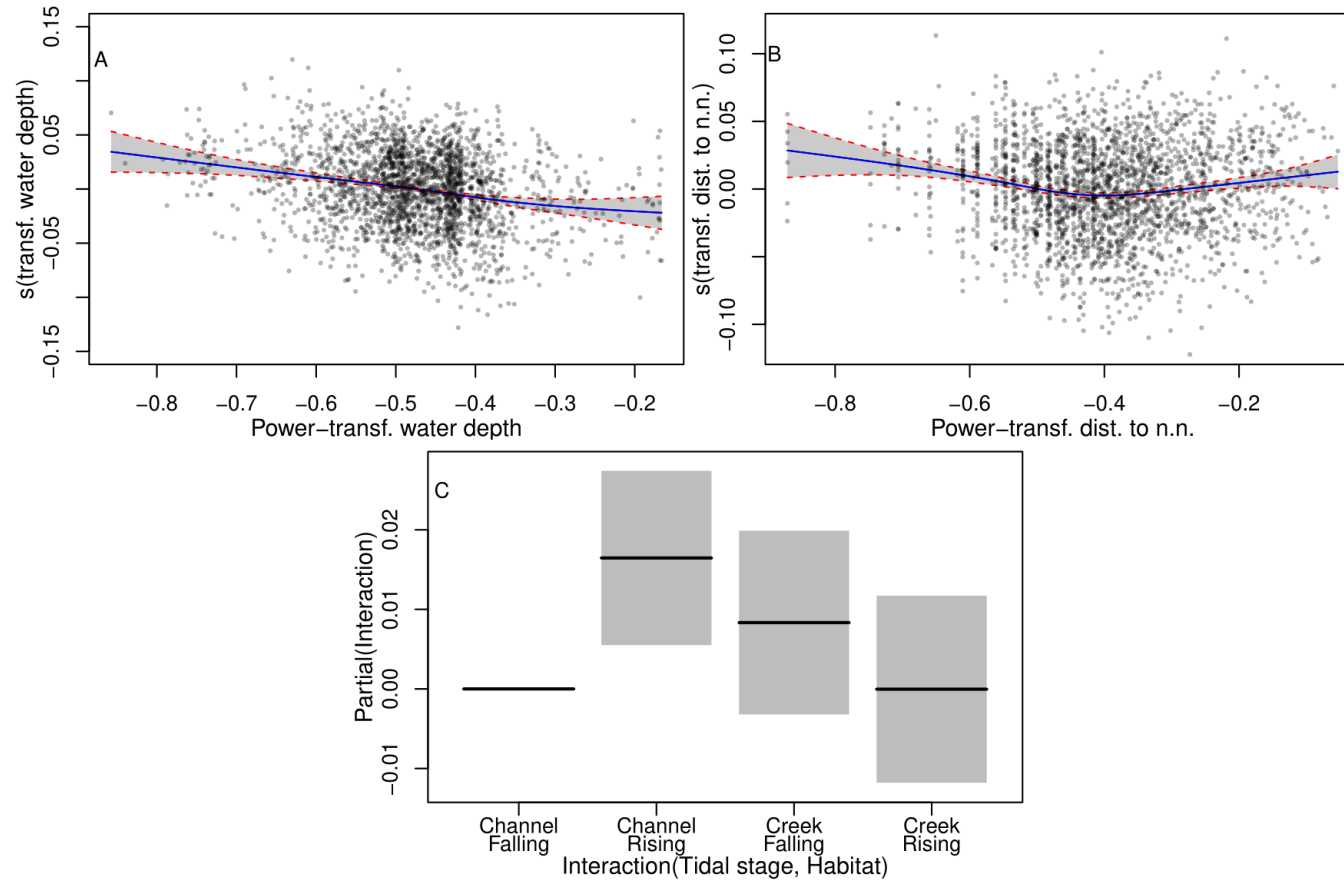


Figure 3.5: Smooth terms and parametric coefficient partials for the explanatory variables of the within-school density general additive model. A. Parametric coefficient partials for categorical abundance. B. Smooth term for power-transformed distance to the nearest observed neighboring school. C. Parametric coefficient partials for the tidal stage. D. Smooth term for power-transformed water depth.

Table 3.1: Terms for within-school density GAM

A. Parametric coefficients	Estimate	Std. Error	t-value	p-value	Sig.
(Intercept)	-0.70951	0.00600	-118.188	< 0.0001	***
Interaction(status, hab) rising.channel	0.01646	0.00559	2.946	0.0033	**
Interaction(status, hab) falling.creek	0.00834	0.00589	1.415	0.1572	
Interaction(status, hab) rising.creek	-0.00003	0.00600	-0.006	0.9955	
B. Smooth terms	edf	Ref.df	F-value	p-value	Sig.
s(lon, lat)	7.738	10.34	4.269	<0.0001	***
s(transf. water depth)	2.321	9.00	11.009	< 0.0001	***
s(transf. dist. to n.n.)	3.628	9.00	14.626	< 0.0001	***
s(random school effect)	323.819	444.00	6.735	< 0.0001	***

Menhaden schools in shallower waters had significantly higher within-school densities than the mean density across the whole study area, while schools in deeper waters had lower within-school densities (Figure 3.5A). Back-transforming the water depth variable shows that the shift from above the mean to below the mean occurred at approximately 4.0 m. The smoother term for the power-transformed distance between NONS indicated it had a non-linear relationship with within-school density. If the distance between two NONS was either near the minimum or near the maximum, they had higher within-school density than the mean. Schools with nearest observed neighbor distances near the median had lower within-school density than the mean (Figure 3.5B). The back-transformation of this variable indicates that schools closer than 5.8 m or further than 23.0 m from their NONS had higher within-school density.

The interaction of tidal stage and habitat used falling tide in the river channel habitat as the reference level. Menhaden schools at rising tide in the channel habitat had significantly higher within-school densities than the reference level, while schools in both rising and falling tides in the creeks did not have within-school densities significantly different from the reference level. Finally, the interaction between tidal stage and month used falling tide in June as the reference level. Schools had significantly lower within-school densities on rising tides in July and August

when compared to the reference level.

Model selection for the between-school distance GAM indicated that water depth, the interaction between tidal stage and habitat, and the interaction between tidal stage and month were all significant environmental predictor variables. Within-school density was also retained as a significant predictor and is likely to represent behavioral mechanisms related to schooling. The spatial location smoother term (interaction of latitude and longitude) was also retained as significant. The explanatory variables do not have issues of collinearity. Significance for all terms in this model can be found in [Table 3.2](#). Though the residuals of the final model did not indicate a violation of the assumption of independence, spline correlograms and Mantel tests indicated the presence of spatial autocorrelation. Of the 48 sampled transects with more than two schools of fish, 21 had issues of autocorrelation that frequently presented as positive spatial autocorrelation at short distances ($< 100\text{m}$). The relationship of within-school density to the explanatory variables was plotted ([Figure 3.6](#)).

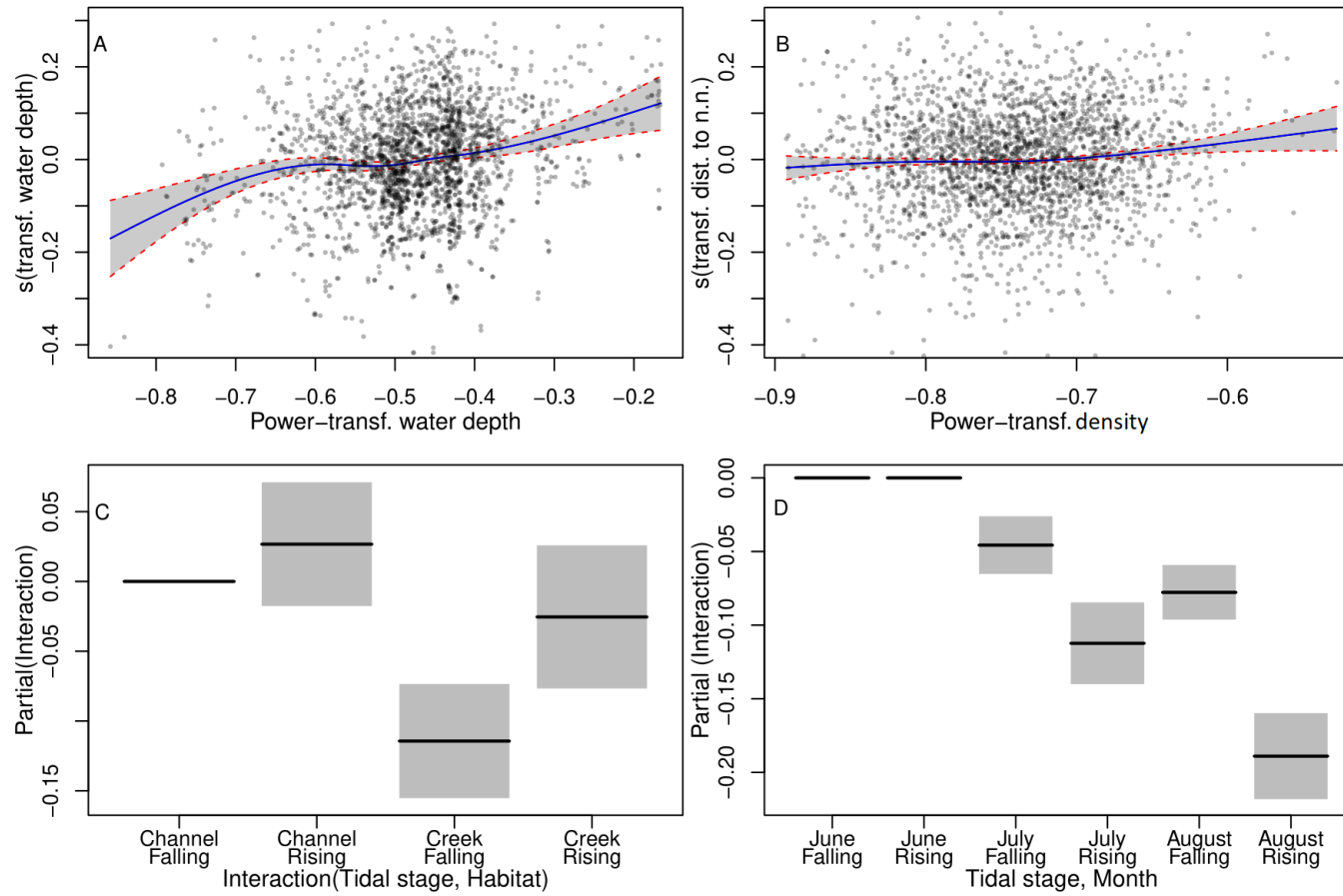


Figure 3.6: Smooth terms and parametric coefficients partials for the explanatory variables of the between-school distance generalized additive model. A. Smooth term for power-transformed water depth. B. Smooth term for power-transformed within-school density. C. Parametric coefficient for tidal stage- habitat interaction. D. Parametric coefficient for month-tidal stage interaction.

Menhaden schools in shallower water had shorter distances between NONS than the mean, while schools in deeper water had longer distances between NONS (Figure 3.6A). The back-transformation of the water depth variable indicates that this transition occurs around 4.0 m. The smoother term for power-transformed within-school density indicated that schools with a higher within-school density would have longer distances between NONS than the mean. Schools with lower within-school density would have shorter distances than the mean (Figure 3.6B). The back-transformation of this variable indicates that the transition occurs around 33 individuals per meter. Schools at falling tides in the river channel were used as the reference level for the interaction of tidal stage and habitat. Schools at falling tide in the creek had significantly shorter distances between NONS than the reference level ($p < 0.001$), while schools at rising tides in both the creek and channel habitats were not significantly different from the reference level. Schools at falling tides in June were used as the reference level for the interaction between the month and tidal stage. Schools in all other levels except for those at rising tides in June had significantly shorter distances between NONS than the reference level (Table 3.2). The output for the rising tide-June combination indicates there is not a big enough sample size to make a comparison to the reference level.

3.4.3 Clustering pattern

A total of 20 transects were sampled in 2016. Of the 20 sampled transects, 16 transects had more than 2 observed menhaden schools and could be analyzed for clustering patterns (Table 3.3). Of these, 15 had C_{dt} values less than 0.5, suggesting menhaden schools are spatially clustered. One of these transects had a C_{dt} value of 0.49, which is evidence of spatial randomness. The

Table 3.2: Terms for distance between NONS GAM

A. Parametric coefficients	Estimate	Std. Error	t-value	p-value	Sig.
(Intercept)	-0.238349	0.021195	-11.246	< 0.0001	***
Interaction(status, hab) rising.channel	0.026701	0.022623	1.180	0.238	
Interaction(status, hab) falling.creek	-0.114356	0.020870	-5.479	< 0.0001	***
Interaction(status, hab) rising.creek	-0.025422	0.026168	-0.971	0.331	
Interaction(status, month) rising.June	0.000000	0.000000	NaN	NaN	
Interaction(status, month) falling.July	-0.045673	0.009996	-4.569	< 0.0001	***
Interaction(status, month) rising.July	-0.112339	0.014156	-7.936	< 0.0001	***
Interaction(status, month) falling.August	-0.077754	0.009454	-8.225	< 0.0001	***
Interaction(status, month) rising.August	-0.189001	0.014896	-12.688	< 0.0001	***
B. Smooth terms	edf	Ref.df	F-value	p-value	Sig.
s(lon, lat)	14.171	18.45	3.468	< 0.0001	***
s(transf. water depth)	4.317	9.00	4.471	< 0.0001	***
s(transf. within-school dens.)	2.543	9.00	1.908	< 0.0001	***

C_{dt} values were less than 0.1 for 8 transects, providing strong evidence for clustering within the respective transects. Only one transect in 2016 had a $C_{dt} > 0.5$, which would indicate uniform dispersion.

In 2017, 40 transects were sampled. Of all transects, 32 had more than 2 menhaden schools and could be analyzed for clustering patterns (Table 3.4). Of these, 30 had C_{dt} values less than 0.5. One transect had a C_{dt} value of 0.49, which is evidence of spatial randomness. The C_{dt} values for 23 transects were < 0.1 , providing strong evidence for clustering along the transects. Only 2 sampled transects in 2017 had C_{dt} values which indicated uniform dispersion.

3.5 Discussion

3.5.1 Qualitative assessment of school characteristics

Understanding the relationship between the fine-scale spatial pattern, density, and extent of juvenile menhaden schools and the local and regional environmental and behavioral factors

Table 3.3: C_{dt} values for 2016 transects

Date	Transect	n School	C_{dt}
08/11	St. Leonard Creek	100	0.018
08/26	St. Leonard Creek	119	0.019
07/28	St. Leonard Creek	77	0.023
06/28	St. Leonard Creek	76	0.036
07/14	South Patuxent	20	0.051
07/14	St. Leonard Creek	22	0.064
07/21	St. Leonard Creek	43	0.069
06/15	St. Leonard Creek	33	0.081
07/21	North Patuxent	19	0.109
07/14	North Patuxent	8	0.122
06/15	South Patuxent	10	0.129
08/11	North Patuxent	4	0.268
08/26	North Patuxent	4	0.389
07/21	South Patuxent	6	0.397
07/28	North Patuxent	6	0.495
08/26	South Patuxent	3	0.723

that act upon spatial pattern could improve assessments of the species' population and spatial dynamics ([Bartolino et al., 2017](#)). The abundance of juvenile menhaden individuals in the study area peaked in August, mainly driven by an increase in abundance in the creek habitat. Peak abundance of individuals in the river channel occurred in July. The number of schools in the study area followed the same pattern; the highest number of discrete schools in the study area occurred in August, although the highest number of discrete schools in the river channel occurred in July. As similar transect lines were followed and similar amounts of data were collected during sampling events, this observed increase in abundance is likely evidence of menhaden juveniles moving downriver into the study area from the fresher upriver habitats in the Patuxent River they inhabit as larvae.

Finding a higher frequency of both juvenile menhaden individuals and juvenile menhaden schools in the creek habitat has implications for population survey methods. Previously, it was

Table 3.4: C_{dt} values for 2017 transects

Date	Transect	n School	C_{dt}
08/11	St. Leonard Creek	204	0.011
08/31	St. Leonard Creek	130	0.014
07/18	St. Leonard Creek	136	0.015
08/11	Cuckold Creek	126	0.015
07/24	Hellen Creek	115	0.016
07/07	Cuckold Creek	104	0.018
08/31	Hellen Creek	136	0.018
07/24	St. Leonard Creek	94	0.019
08/11	Hellen Creek	118	0.021
07/18	Cuckold Creek	79	0.027
08/31	North Patuxent	18	0.027
07/07	South Patuxent	39	0.032
07/18	Hellen Creek	74	0.032
06/27	Cuckold Creek	47	0.034
07/24	Cuckold Creek	68	0.034
07/07	St. Leonard Creek	40	0.04
08/31	Cuckold Creek	45	0.045
06/27	St. Leonard Creek	37	0.053
07/18	North Patuxent	10	0.054
07/18	South Patuxent	12	0.068
07/24	South Patuxent	17	0.074
08/11	North Patuxent	18	0.075
07/07	Hellen Creek	33	0.079
07/24	North Patuxent	13	0.108
06/16	St. Leonard Creek	14	0.127
06/16	Hellen Creek	10	0.13
06/08	Cuckold Creek	13	0.137
06/16	Cuckold Creek	10	0.299
06/27	Hellen Creek	5	0.305
06/08	St. Leonard Creek	5	0.497
08/11	South Patuxent	4	0.531
06/08	Hellen Creek	3	0.617

found that juvenile menhaden are present at significantly higher densities in the creek habitat compared to the river channel habitat (Lankowicz et al., 2020). The river channel has a larger volume of water compared to the creek habitat, but it also has fewer schools comprised of either approximately the same number of individuals or fewer individuals than schools in the creeks.

A greater proportion of the population may reside in the relatively small creek habitat. It is also possible that the increased relative sampled area in the narrow, shallow creeks as compared to the river channel affects the frequency of menhaden observations. Regardless, the results provide evidence that robust population estimates require surveys of shallow estuarine creek habitats.

The traditional understanding of menhaden school morphological characteristics through net-based fisheries-independent survey methods and fisheries-dependent harvests is that schools are vast in both geometric size and number of individuals. While this is true for adult and near-adult menhaden in deeper waters, the schooling characteristics of earlier life stages in shallow estuarine waters were not previously well described. Juvenile menhaden observed in the study area were organized into many smaller schools (tens to hundreds of individuals) as compared to a few extremely large schools (thousands of individuals).

3.5.2 Behavioral drivers of school pattern and morphology

The output from both the within-school density and distance between NONS generalized additive models indicates that school morphological characteristics and spatial pattern are affected by several local behavioral drivers. The within-school GAM indicated that grouping schools by size and using these groups to simulate the random effects of schooling had a significant effect on the model output. The number of individuals per school is expected to affect processes such as feeding success ([Foster et al., 2001](#)) and the likelihood a school will stay together or break apart ([Niwa, 2004](#)). The structure of this variable as a random effect is similar to previous approaches ([Anderson, 1981](#)) and indicates that schooling dynamics are dependent on both the relative size of observed schools and random effects distributed over the range of

those school sizes. The between-school GAM did not retain this variable as a significant factor, perhaps because the spatial pattern of schools is affected more by the distribution of favorable patches within a heterogeneous environment than the morphological characteristics of the school.

The within-school density GAM retained the power-transformed distance between the NONS variable as a significant term. A relationship between within-school density and the number of schools in a habitat area has been noted. Additionally, a relationship between the distance between NONS and the number of schools in a habitat area has been noted. The number of schools in an area may link these two behavioral mechanisms, as schools with short distances to NONS tended to have higher within-school density. The distance between NONS GAM retained the power-transformed within-school density variable as a significant term. It has been speculated that the size of a fish school affects the probability that it will break up or reform, which would in turn affect the distance between NONS. Further, in the three-spine stickleback (*Gasterosteus aculeatus*) group size has been found to affect the distance between nearest neighboring individuals within a school (Ward et al., 2017), and drivers that shape the position of fish relative to each other within a school could have emergent properties that result in the spatial structuring of schools relative to each other. Behavioral drivers and their interactions likely affect the spatial patterns of menhaden schools at both local and regional spatial scales. The use of within-school density and distance between NONS variables in models both as response and predictor variables (though not within the same model) is designed to both test the direct effects of one term on the other, and to implicitly include the effects of behavioral drivers for which explicit information is unavailable.

3.5.3 Environmental drivers of school pattern and morphology

The output of both GAMs indicated that water depth and interactions between tidal stage, habitat, and sampling month were significant environmental factors to model menhaden school spatial pattern and schooling characteristics. Water depth, and likely habitat type, are linked to a suite of other environmental variables, including temperature gradients, food availability, and currents, all of which are known to affect the spatial distribution of estuarine fish (Spares et al., 2012). The month variable likely encompasses temporal changes in several local and regional environmental variables that affect menhaden, like temperature (Ding and Elmore, 2015), dissolved oxygen concentration (Scully, 2016), and phytoplankton prey availability (Marshall and Affronti, 1992). Month would also have a positive correlation to menhaden body length, as the individuals grow through the season. Other studies relating fish schooling to the tidal condition have found forage fish schools have higher densities at high tide (Couperus et al., 2016), which would correspond to late flood periods through early ebb periods. However, lacking measurements for these environmental variables, these relationships cannot be analyzed.

3.5.4 Spatial clustering of schools

Menhaden, like other forage fish, are known to have patchy spatial distributions across heterogeneous environments (Gutiérrez et al., 2012; Haugland and Misund, 2004; Liang et al., 2020). Menhaden schools are likely to cluster, especially in the creek habitat. Of the 60 transects sampled over both years, 48 had more than 2 schools of juvenile menhaden. Of these, 43 had cluster coefficients consistent with moderate to very strong evidence of schools arranged in spatial clusters along the transects. Only 3 of the sampled transects had C_{dt} values indicative of

uniform dispersal. These uniform dispersal patterns tended to occur in the river channel habitat (2 of 3 occurrences) and with few observed schools. The remaining 2 transects had C_{dt} values close to 0.5, which would indicate spatial randomness along the transects. The tendency of menhaden schools to cluster likely indicates that menhaden, either individually or as schools, are detecting and responding to environmental and behavioral cues. Previous research indicates that phytoplankton distribution is a critical environmental driver of menhaden distribution (Friedland et al., 1989, 1996). Output from the GAMs indicates that population abundance, within-school density, and the distance between NONS, water depth, and tidal stage are among these cues.

3.5.5 Potential limitations

Fish schools are highly dynamic and likely to merge and break apart frequently in response to environmental and behavioral cues (Pitcher et al., 1996; Soria et al., 1998). It is difficult to observe and quantify *in-situ* schooling accurately using any instrument or sampling approach. As with all studies, there are limitations to the ARIS instrument and sampling approach. While some of the identified fish may be forage species other than menhaden, the combined evidence of visual identification of menhaden at the water surface during data collection, previous stationary ARIS studies of menhaden body shape (Shahrestani et al., 2017), previous stationary ARIS studies of menhaden schooling behavior (Zhang et al., 2017), and seasonal size information for juvenile menhaden give confidence that most of the fish identified in this study are juvenile menhaden. Though the ARIS itself is a non-invasive sampling tool, menhaden, and other organisms may attempt to avoid the boat motion and engine noise from the sampling vessel. The effects of vessel avoidance are unquantified, so the scope of this potential limitation is still unknown.

There is likely measurement error associated with the recording frequency and accuracy of the handheld GPS, which is accurate to 3m. Schools that had clear measurement errors were excluded from the analysis. However, the school length, density, and distance between NONS variables might still contain some measurement error. The dataset also likely has some spatial and temporal autocorrelation — menhaden schools close together in space and time could be influencing each other. This may increase the likelihood of type 1 error when comparing schooling metrics.

The transect-based deployment of the ARIS can be considered an Eulerian approach, in which individual fish must pass through a field of view to be considered. The distribution of schools outside the field of view is unknown. Further, the behavior of the individuals and schools before or after observation is unknown, and so it is uncertain if the size and morphology of the schools are representative of a stable, ordered collective state or a less-stable transitional state ([Calovi et al., 2015](#); [Tunstrøm et al., 2013](#)). Further, schools larger in area than the field of view will not be entirely captured. Schools that are equal to or smaller than the field of view must be centered within the field of view, or they also will not be entirely captured.

Despite its limitations, fine-scale sonar imaging is a useful approach to observing the distribution of forage fishes in shallow estuarine environments. The results suggest that juvenile menhaden schooling characteristics and spatial pattern are affected by a suite of local and regional environmental and behavioral drivers.

3.5.6 Conclusions

The sonar imaging techniques employed in this research allowed for a non-invasive and quantitative determination of menhaden schooling morphology and distribution at a fine spatial scale, and within an area where few previous assessments have been made. These data are both novel and valuable. The characteristics, pattern, and abundance of menhaden schools observed at fine scales and in heterogeneous environments can be aggregated and projected to predict distribution and abundance at larger scales. Further, the clear differences in menhaden distribution by habitat type and with water depth are evidence that current survey methods for menhaden in deeper waters should be augmented with surveys conducted in shallower habitats. Furthermore, these data collection and analysis methods are not exclusive to menhaden and have already proven useful to observe other species that inhabit waters that are difficult to sample with established methods ([Kerschbaumer et al., 2020](#)). The descriptive and comprehensive information on fish schooling characteristics and distribution provided by sonar imaging could provide valuable insights and predictive tools for management efforts.

3.6 Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

3.7 Acknowledgements

The authors thank Dong Liang for his intellectual contributions to the early iterations of this study. I also thank Chunlei Fan, Suzan Shahrestani, Linlin Wang, Nathan Hirtle, Nikelene McLean, and Gang Lin for their assistance in fieldwork and image processing.

Chapter 4: Spatial point pattern analysis of juvenile Atlantic menhaden schools in shallow estuarine environments

4.1 Abstract

Atlantic menhaden is an ecologically and economically important forage fish species and is known to aggregate during most of its life stages. The drivers of these aggregations and the resulting patterns are variable with spatial scale. Understanding the conditions under which aggregations of menhaden are formed at any scale will contribute to our overall understanding of its ecology and could aid in the effective management of the population. Spatial data on the distribution of menhaden schools in the lower Patuxent River and its connecting creeks, all shallow estuarine tributaries, were collected using non-invasive sonar imaging. The locations of the schools were treated as points and analyzed using spatial point pattern analysis. Previous explorations into the patterns of menhaden distribution in this area indicated that there are important distinctions in environmental variables and menhaden distribution between the Patuxent River channel and the various connecting creeks. Point pattern analysis was therefore conducted at three different extents to test the effects of extent on menhaden school spatial point patterns. At spatial scales of tens of meters to kilometers, variations in spatially-linked environmental factors create point patterns with first-order structure. At spatial scales of meters to tens of meters, in-

interactions between schools create point patterns with second-order structure. Characterizing the drivers behind menhaden school distributions at varying spatial scales has important implications for the understanding and management of the population.

4.2 Introduction

Environmental conditions and behavioral mechanisms act on multiple spatial scales and self-organizational levels of a species and serve to shape its overall spatial distribution. These underlying processes are inherently spatial and varied, and species detect and respond to these variations to create spatially-structured populations ([Grados et al., 2012](#)). This results in each range of spatial scales having a unique set of spatial patterns created by an equally unique set of environmental conditions, behavioral mechanisms, and their interactions ([Levin, 1992](#)). It is difficult to identify spatial patterns and the underlying processes that shape them, a problem articulated by [Levin \(1992\)](#) as “the problem of pattern and scale.” There is no single natural scale at which all ecological processes occur, thus scale should be used as a means to understand patterns in natural variation. It is therefore important to identify the most relevant scale to observe and characterize variation in species distribution while understanding that the perceived patterns and overall distribution may change with the scale of observation ([Dray et al., 2012](#); [Levin, 1992](#)). For species that self-organize into increasingly complex spatial aggregations, it is also important to consider variations in spatial pattern between the organizational levels, these being 1) the individual, 2) aggregations of individuals, 3) clusters of aggregations, 4) concentrations of clusters and 5) the entire population ([Fréon et al., 2005](#)).

Aggregations of organisms are widely observed across many ecosystems, spatial scales,

and taxa of interest, and are an important behavioral mechanism behind species' spatial distributions. Aggregative behaviors are common to many fishes. Approximately half of all extant fish species will form aggregations for at least a portion of their life cycles, and a quarter will aggregate for their entire life cycles ([Shaw, 1978](#)). These aggregations can be defined as either shoals or schools. Shoals are “loose” social groupings that are understood to benefit individuals within the group by reducing predation pressure on the individual ([Hamilton, 1971](#)). Schools are specialized social groupings, in which conspecifics of similar size coordinate their behavior to swim in the same direction at the same rate ([Fréon et al., 2005](#)). Schooling is therefore an emergent property that is only observable at a scale greater than just the individual.

Schooling behavior is common in numerous species often categorized as forage fish ([Alder et al., 2008](#); [Pikitch et al., 2012](#)). These small- to medium-sized pelagic fish facilitate energy transfer between trophic levels ([Bakun et al., 2010](#); [Pikitch et al., 2012](#)) and are integral to the diets of a diverse array of species including pinnipeds, cetaceans, seabirds, and several life stages of commercially and ecologically important piscivorous fishes ([Hartman and Brandt, 1995](#); [Luo and Brandt, 1993](#); [Pikitch et al., 2012, 2014](#); [Walter III and Austin, 2001](#); [Weise and Harvey, 2008](#); [Wood and Houde, 2003](#)). Clupeids (menhadens, herrings, and sardines) are particularly well-known for their economic value and ecological impacts. They typically form large, dense, and patchily distributed schools within areas of suitable habitat ([Ahrenholz, 1991](#); [Boyd et al., 2015](#); [Haugland and Misund, 2004](#); [Liang et al., 2020](#)). The spatial structure of the suitable areas is shaped at large scales by environmental factors such as predator and prey abundance, thermal structure, turbidity, and current speed and direction, which combine to create a heterogeneous environment ([Boyd et al., 2015](#)).

Schools are known to aggregate into clusters within patches of suitable habitat, further

adding to the patchy nature of forage fish distribution (Fréon and Misund, 1998; Le Fur and Simon, 2009; Petitgas and Levenez, 1996; Swartzman, 1997). The causes for these clusters, as well as their spatially-explicit formation and maintenance, have been studied with varying success *in situ*. Various reasons for cluster formation examined in the literature include temperature, depth, hydroclimatic conditions (Swartzman, 1997), population abundance (Petitgas et al., 2001), information transmission between schools (Haugland and Misund, 2004), and food availability (Bertrand et al., 2008).

Even at relatively small spatial scales within patches of suitable habitat, changes in the distribution of forage fish schools and clusters of schools can have significant ecological impacts. Shifts in fish aggregations can affect the distributions and behaviors of animal and human predators (Benoit-Bird et al., 2013; Langseth et al., 2014; Potier et al., 1997). Characterizing the features and spatial scales of both schools and clusters of schools is therefore important both to understand the ecology of forage fishes and to effectively manage their stocks (Benoit-Bird and Waluk, 2021).

It is difficult to collect data on the dynamics of each of the aggregative levels of forage fish species with net-based survey methods like trawls and seines, as these methods are limited in their spatial grain. Net-based methods can only provide information at the spatial grain of a single tow or net pull. The extent of areas covered by both trawls and seines is also limited. Trawls require waters deep enough for research vessels ($> 4\text{m}$) and seines are typically conducted in extreme nearshore shallows ($< 2\text{m}$). Finally, net-based methods are invasive and will likely affect the behavior and spatial patterns of forage fish schools in the area of operations (Trumpickas et al., 2020). Hydroacoustic surveys have been established as effective methods for avoiding the extent, grain, and invasive issues of net-based methods (Becker et al., 2017; Boswell et al.,

2010; Kubečka et al., 2012; Liang et al., 2020; Lucca and Warren, 2018; Samedy et al., 2015; Simmonds and MacLennan, 2005). Typically, multi-frequency echosounders are used to produce and identify acoustic backscatter from fish bodies, which is then quantified into fish biomass (Misund, 1997; Simmonds and MacLennan, 2005). This method works well to efficiently sample large volumes of water but cannot identify individual fish.

Most of the studies of fish clustering dynamics have occurred in large basins or continental shelf environments (Benoit-Bird and Waluk, 2021; Bertrand et al., 2008; Brochier et al., 2018; Le Fur and Simon, 2009; Lucca and Warren, 2019; Petitgas and Levenez, 1996; Swartzman, 1997). There are very few studies that attempt to parse the clustering dynamics of forage fish in shallow estuarine or extreme nearshore coastal environments [for example, see Lucca and Warren (2018)]. This is partially due to the depth requirements of net-based methods, and partially because spatial fisheries ecology has focused on examining large shifts in abundance by averaging demographic and distribution data over large spatial scales (Ciannelli et al., 2008; Mason and Brandt, 1999)).

Sonar imaging is a newer active acoustic survey method that is similar to echosounders in that it uses the return of high-frequency sound signals to identify fish but can interpret these return signals into digital images with up to a 3mm per pixel resolution. This allows for the identification and enumeration of individuals. Sonar imaging can be conducted from small research vessels at speeds up to 1 m s^{-1} (Lankowicz et al., 2020). Therefore, sonar imaging can cover a large spatial extent, including areas too shallow for net-based methods or larger research vessels. This high-resolution, minimally-invasive method capable of sampling an area of a relatively large extent is therefore an ideal tool to explore the multiscale spatial dynamics of forage fish schools and clusters in estuarine environments.

Atlantic menhaden (*Brevoortia tyrannus*, hereafter: menhaden) is an appropriate study species to observe schooling and clustering dynamics in estuarine environments. Menhaden aggregate into shoals or schools at every life stage after early larval development, have a spatially-varied life history, and inhabit many estuarine and coastal environments (Friedland et al., 1996). Menhaden are estuarine dependent; larvae are transported into estuaries where they metamorphose into juveniles and grow rapidly (Ahrenholz, 1991; ASMFC, 2020; Hare et al., 1999, 2005; SEDAR, 2020). Post-metamorphic juveniles are known to aggregate in areas with high primary production (Friedland et al., 1989, 1996, 2011; Houde et al., 2016). This aggregative behavior likely results in a series of emergent spatial patterns, where individuals aggregate into schools and schools aggregate into clusters.

This work examined the dynamics of juvenile menhaden schools and clusters of schools in a shallow-water estuarine tributary as observed in field data collected with a sonar imaging camera. Point pattern analysis was used to identify spatial structures that depart from complete spatial randomness and to identify spatial covariates that may lead to a variable intensity of school distribution within the study area. The effect of spatial extent on the outcomes was explored by repeating the analysis at both smaller and larger spatial extents.

4.3 Methods

4.3.1 Study area

Sonar imaging transects were conducted in the Patuxent River, a western-shore tributary of the Chesapeake Bay. The Chesapeake Bay and its tributaries are important nursery grounds for menhaden (Hildebrand and Schroeder, 1928) and juvenile menhaden are among the most highly

abundant fish in the Chesapeake Bay ([Jung and Houde, 2003](#); [Wood and Houde, 2003](#)). Sampling took place in daylight hours (0800-1400) in the mesohaline and tidally-influenced region between river kilometers 5 and 12. Two cross-channel transects, the North Patuxent and South Patuxent, characterized juvenile menhaden school distribution in the main river channel. The mean sampling depth in the North Patuxent was 7.2m, while the mean sampling depth in the South Patuxent was 10.7m. The transects were similar in length and orientation, and both connect to the mouth of a creek. The two transects were grouped into a single habitat type, labeled the river channel habitat, due to their physical similarities.

Juvenile menhaden school distribution was also characterized in the three largest creeks connected to this area of the Patuxent River. These were St. Leonard Creek (mean sampling depth 4.2m), Hellen Creek (mean sampling depth 2.8m), and Cuckold Creek (mean sampling depth 3.8m). All sampled creeks are much narrower and shallower than the Patuxent River channel. The St. Leonard Creek transect was typically the longest sampled transect in all sampling efforts, while Hellen and Cuckold Creeks were often the shortest. The three creek transects were similar in physical characteristics and were therefore grouped and labeled the creek habitat.

4.3.2 Data collection

Two Adaptive Resolution Imaging Sonar (ARIS) systems were utilized during the two-year study period. An ARIS 1800 was used in 2016 and an ARIS 3000 was used in 2017. Both ARIS cameras were mounted on a polearm fixed to the side of a small research vessel and held approximately 0.25m deep and at a 35-50° angle relative to the surface. A handheld GPS unit (Garmin GPSMAP 64s, Garmin Ltd., Ottawa Canada) concurrently recorded GPS information. The ARIS

1800 was deployed across the North and South Patuxent transects and the St. Leonard Creek transect from June to August of 2016. The maximum range of the camera was typically 10m and transects were sampled approximately twice every month. The ARIS 3000 replaced the ARIS 1800 in the 2017 sampling season due to its increased image resolution. Data were collected at least twice monthly from June to August of 2017 in the three original transects, plus Cuckold and Hellen Creeks. The maximum range of the camera was typically 6m. For both sampling years and camera types, 12-15 images were collected per second. It is unlikely that the change in the camera model or sampling range would affect the ability to observe menhaden schools. The resolution per pixel of both camera models is sufficient to identify individual menhaden, and an exploratory analysis of school depth found few schools deeper than 6m. More information on the specifics of each ARIS camera model and beam array can be found in [Lankowicz et al. \(2020\)](#).

4.3.3 Data processing

ARIS images were manually processed for the presence of juvenile menhaden using ARIS-Fish software (Sound Metrics Corp., Bellevue, WA). ARIS cameras cannot resolve images taken during mobile deployments enough to identify to a species level. Therefore, a set of standards was created to categorize target objects as either potential menhaden or not menhaden. For each image, target objects were considered to be menhaden if there were more than two objects and the objects were of an appropriate size compared to a menhaden seasonal growth curve ([Annis et al., 2011](#)), were tightly clustered, and were generally oriented in the same direction. Single target objects were not retained for further analysis, as they cannot be considered schools and it was considerably harder to identify target objects without observing their movement behavior in

groups. Three single target objects in 2016 and two single target objects in 2017 were discarded due to this threshold.

Schools of menhaden in ARIS data first appear as a few, densely packed target objects of similar size moving in the same direction. Over the next several consecutive images, the number of target objects in the field of view would expand to upwards of 100 individuals. The target objects would eventually all pass through the opposite edge of the field of view from where they entered. Thus, discrete schools were relatively easy to observe and were defined as a group of target objects meeting the criteria to be identified as forage fish in a continuous series of more than two images, separated from other discrete schools by more than 1m. More detail on the identification of discrete menhaden schools can be found in Chapter 3.

4.3.4 Point pattern analysis

Point pattern analysis is a commonly-used method in the field of spatial ecology to characterize the patterns and drivers that create the overall distribution of a set of events within a defined window of observation ([Wiegand and Moloney, 2013](#)). The main objective of point pattern analysis is to determine if the pattern is a result of some underlying point process, or if it is simply random. The pattern is characterized by its first- and second-order structure. First-order structure describes the spatial variation in the distribution of the points, referred to as intensity, caused by a spatial covariate (e.g. latitude or longitude). Second-order structure describes whether the location of a point in the pattern is dependent on the location of one or more of the other points. For this study, schools of menhaden were treated as points in a spatial point pattern within windows of observation characterized by the data collection transects.

4.3.4.1 Linear networks

Data collection was not continuous throughout the study area. The five transects were completed separately and did not often connect. There were occasions in which the North Patuxent and St. Leonard Creek transects were connected, but they were artificially separated and treated as unique transects because of both the difference in the number of schools and the characteristics of the differing habitat types. Overall, the study area is best described as a set of one-dimensional, curvilinear transects situated within the same two-dimensional plane. The transects are comprised of many short line segments joined together at nodes, and therefore each transect can be considered a separate, non-branching linear network. The distribution of menhaden schools in this network is a spatial point pattern that is driven by a spatial point process.

It is difficult to compute both first-order intensity and second-order correlative statistics for spatial point patterns along transects or linear networks. The geometry of the network is unique relative to each spatial location along the network, and therefore the two-dimensional plane containing the network cannot be considered a homogeneous space ([Baddeley et al., 2015](#)). Traditional statistical techniques for two-dimensional spaces cannot be used due to the one-dimensional nature of the line transects. The influence of spatial point processes and the distribution of spatial points is limited to along the network, and cannot be extrapolated to areas that are not along the network path. Therefore, the spatial point pattern of menhaden schools along the transects was investigated using the network pair correlation function (PCF) with Ang's geometric correction as recommended by [Ang et al. \(2012\)](#) and [Okabe and Yamada \(2001\)](#).

Though the ARIS does compress three-dimensional volume to a two-dimensional image with both height and width, the transects were treated as one-dimensional and the position of

fish schools across the “width” of the camera field of view was not considered. This decision was made for multiple reasons. First, the maximum width of the field of view was often several orders of magnitude shorter than the total transect length. This would mean the contribution of school positions in the “width” dimension to the overall spatial pattern would be negligible in comparison to the contribution of school positions along the length of the transects. Further, the methods of spatial analysis used in this study are based on Ripley’s K-function, and the distribution of objects is considered a spatial point process in which the area of the object is negligible in comparison to the size of the study area ([Ripley, 1981](#)). A school of menhaden often occupied the entirety of the width of the camera field of view, so this assumption would be violated in the “width” dimension. Finally, the distribution of schools outside of the transect paths was not interpolated. Schools of menhaden certainly navigate the Patuxent River and its creeks outside of the defined transects paths, but little data was collected to determine the shore-to-shore distribution of menhaden in the creeks or alongshore distribution of menhaden in the Patuxent River channel. It would be inappropriate to interpolate the spatial distribution of menhaden schools to areas where data was not collected.

The method for placing schools of menhaden along a transect was to use the spatial location of the median frame in the sequence of frames in which a distinct school was present. Distances between schools are therefore not edge-to-edge, but midpoint-to-midpoint. This approach is more consistent with the assumption that the area of the spatial objects is negligible in comparison to the size of the study area. Though this recorded spatial position is the true location of the central beam of the ARIS field of view during recording, the camera has a finite field of view and a portion of the school may be outside that field of view. Thus, it may not be the true location of the menhaden school’s geometric center or center of density. Regardless, this approach is the

most consistent way to record menhaden school positions and examine menhaden school spatial patterns within the collected data.

Analyses were completed in R ([R Core Team, 2021](#)) using the `spatstat` package ([Baddeley et al., 2015](#)). GPS data interpolated to 10-second increments were split into distinct transect paths for each sampling event. A 10-second increment is roughly equivalent to 10 meters of distance traveled. Transect paths were then converted into networks of line segments. Each recorded position in the GPS data was represented as a node and Euclidean distance between nodes became “edges,” or connecting line segments. The positions of observed menhaden schools were overlaid onto the linear network as events. Transects with fewer than three observed schools were not retained for analysis, as more than two points are needed to analyze a spatial point pattern. Transects could not be combined to increase the number of points as, again, the transects were not often connected. Therefore, five transects completed in 2016 were removed from the analysis, as were six transects completed in 2017. One transect completed in St. Leonard Creek in 2016 was also removed from the analysis as the data were collected across many short and disconnected segments, rather than the single longer transect that characterized all other data collection efforts in this area.

4.3.4.2 Pair correlation functions

Linear-corrected PCFs were used to determine if the distribution of menhaden schools was significantly different than what would be expected under the null hypothesis of complete spatial randomness (CSR) at all possible shortest-path distances between schools along the transects. PCFs are often considered the most informative spatial point pattern summary statistic, as they

can capture scale-dependent neighborhood properties of points (Illian et al., 2008; Wiegand et al., 2013). Though there are many spatial summary statistics appropriate for two-dimensional planar applications, only the Ripley's K and PCFs have been adapted for linear applications. Using Ripley's K in addition to the PCF would be redundant, and using it in place of the PCF would be less informative (Wiegand et al., 2013).

PCFs relate the position of spatial points in a point pattern using rings with radius r . The linear-network adapted PCF is similar to the standard PCF, but the rings are confined to the one-dimensional lines of the network. Both the standard and the linear network-adapted PCF return the expected number of points in a ring of radius r centered at an arbitrary point in the pattern divided by the overall intensity λ (Dale et al., 2002; Ripley, 1981; Wiegand et al., 2013). Spatial points closer to the chosen radius of the rings are weighted more heavily than those further from the radius. In cases of complete spatial randomness, the PCF for all values of $r \geq 0$ is equal to 1. The significance of deviations away from this value is determined through Monte Carlo envelopes generated by simulating a null model of the spatial point process hundreds of times. It is therefore important to choose an appropriate null model of the spatial point process. An un-modeled correlation between the intensity of the points and an underlying spatial covariate could incorrectly identify areas of high intensity as second-order structure when they are actually a reflection of first-order structure. If the intensity of the points, λ , is uniform through the area and not dependent on a spatial covariate, then a Homogeneous Poisson Point Process (HPPP) is an appropriate null model. HPPPs assume there is no first-order structure. If the point pattern is dependent on an underlying spatial process, then intensity λ will be spatially variable with the covariate and this variation will need to be included in the model. An inhomogeneous Poisson point process (IPPP) would be the best choice for a null model in this case.

The results from previously published work ([Lankowicz et al., 2020](#)) and other analyses (Lankowicz, in revision) indicate that the best choice for a null model varies with transect. Though the overall spatial pattern of menhaden does not likely depend upon the absolute space-time positioning of the schools, some trends were noted. Depth was identified as an important spatial covariate for menhaden distribution. In addition, menhaden density was found to be elevated at the heads and mouths of creeks and the shores of the Patuxent River. Depth, latitude, and longitude were therefore considered to be likely spatial covariates for the spatial point pattern of schools with first-order structure. The locations of the menhaden schools were projected using a NAD83 Maryland State Plane projection, and so coordinates were reported in easting and northing meters. It is understood that latitude and longitude represent the distance along the transects, as transects were unidirectional and generally straight. In the creek transects, these spatial coordinates would represent the distance from the mouth of the creek to its head. In river channel transects, they would represent the distance across the river channel from shore to shore. Water depth information came from NOAA hydrographic survey 6877 and was averaged over the 10-meter bins that made up the transect paths. HPPP models and IPPP models with latitude, longitude, and water depth spatial covariates were compared for each transect. The best model was selected via AIC.

Once the best null model was selected, an appropriate bandwidth for kernel density estimate was identified. PCFs for observed data are estimated with kernel smoothing methods, where the bias and variance of the kernel depend heavily on the chosen bandwidth of the kernel. The best bandwidth for each transect was determined by likelihood cross-validation methods and retained for use in the estimation of the PCF. Identifying the bandwidth in this way allows for the underlying structure of the data to be represented without under- or over-fitting. When the best

model was an IPPP, the nonparametric estimate of intensity as the function of the covariate, $\hat{\rho}$, was plotted given the estimated bandwidth value.

Finally, a linear-corrected PCF for each transect was calculated and plotted in R. Monte Carlo methods were used to generate 999 simulations of complete spatial randomness given the distribution of the null model. The simulations were used to create and plot a pointwise critical envelope, where the upper and lower limits at each possible pointwise distance r are equal to the largest and smallest value generated at that point for the 999 simulations. If the observed PCF for the point pattern remains within the envelope, there is no evidence of second-order structure at that pointwise distance r and the null hypothesis of CSR fails to be rejected. If the plotted PCF of the data passes above this envelope, it is evidence of clustering at that pointwise distance r . If the plotted function passes below the envelope, it is evidence of uniform dispersion at that pointwise distance r . Both types of deviations are evidence of second-order structure in the point pattern and indicate that the position of each point is affected by the position of at least some of the other points.

As this is a two-sided Monte Carlo test with 999 simulations, the significance level for any pointwise distance r is equal to $2/(1 + n \text{ simulations})$, or 0.002. It is important to note that these envelopes are not confidence intervals. The likelihood that the function would pass outside the envelope due to random chance at any pointwise distance r along the possible range is much higher than the calculated pointwise significance level ([Baddeley et al., 2015](#)). To attribute a p -value of 0.002 to the overall evidence of spatial non-randomness would be a problem of multiple comparisons. However, plotting the deviations of the PCF along the entire range of pointwise distances r allows for a view of the spatial pattern across multiple spatial scales.

Two-sided maximum absolute deviation (MAD) tests were run on the PCF results for each

of the transects that had evidence of deviations from CSR. This test is equivalent to creating a constant and global envelope of variance, where the width of the envelope is equal to the biggest deviation from the theoretical PCF generated at any pairwise distance r , and declaring a statistically significant outcome if the observed function passes out of the global envelope anywhere along the possible range of r (Baddeley et al., 2015). This test allows for a straightforward approach to determining if a spatial point pattern is from a CSR point process while avoiding the problem of multiple comparisons.

The result of spatial point pattern analysis is the determination of whether a point pattern comes from a homogeneous or inhomogeneous point process, and whether the point pattern is clustered, dispersed, or completely spatially random. Each of the sampled transects was labeled as such and the results were compared. A Chi-square test of independence was used to test the relationship between the identified spatial point pattern and habitat type for each transect. In cases where there were too few observed or expected values for a traditional Chi-square test of independence, p-values were simulated based on 2000 replicates. Kruskal-Wallis tests were used to test the relationship between the number of schools and the spatial point process model or point pattern. This non-parametric method of comparison was chosen because the number of schools per transect was highly right-skewed. Kruskal-Wallis tests were also used to test if there was a relationship between transect length and identified point process model or point pattern. The range of r values that had evidence of deviation from complete spatial randomness were qualitatively compared within each habitat type.

It was also of interest to test how the results of the analysis would change given different spatial extents. As mentioned earlier, there were some sampling events in which the North Patuxent and St. Leonard Creek transects were connected. In these instances, the analyses were re-run

on the large, connected line network that encompassed both transects. Analyses were also re-run for a small section of St. Leonard Creek from the creek mouth to approximately 500m upriver (approximately -76.493° N latitude). The resulting spatial point patterns for both the connected and shortened transects were compared with the outcomes from the initial, transect-defined analysis.

4.4 Results

Across 2 years of sampling, 60 transects were completed in the study area. Spatial point pattern analysis within linear networks requires an unbroken network and more than 3 points in the point pattern. Out of all 60 transects, 43 met these requirements. An additional 3 transects with sufficient numbers of schools were excluded from analysis because crossing boat traffic during data collection resulted in disrupted transect paths. The full results of point pattern analysis are represented in [Table 4.1](#) and [Table 4.2](#). Plots for PCFs of all transects can be found in Appendix A.

4.4.1 Spatial point pattern analysis

Spatial point patterns in 20 of the transects had no evidence of first-order structure. This indicates that the intensity of points is uniform within the transect line and is not structured by any spatial covariate. Eleven of the 20 point patterns with no first-order structure occurred in the river channel habitat. Most of the point patterns with no first-order structure also had no evidence of second-order structure (example: [Figure 4.01](#)). Point patterns without second-order structure can be called completely spatially random, as the location of any point is not dependent on the

Table 4.1: Output from model selection and PCFs for transects sampled in 2016. Reported values include month and habitat in which the transect occurred, the selected best point process model (H = Homogeneous Poisson, I = Inhomogenous Poisson), spatial covariate for inhomogenous Poisson point processes (X-coordinate, Y-coordinate, or depth), identified Point Pattern (PP; CSR = completely spatially random, DIV = divergence from CSR), the number of schools in the transect, and the length of the transect in meters.

Month	Habitat	Model	Covar	PP	n Schools	Length (m)
June	Channel	H	-	CSR	10	1812
June	Creek	I	Y	CSR	33	4091
June	Creek	I	Y	DIV	60	4040
July	Creek	I	Y	CSR	9	4130
July	Channel	H	-	CSR	8	2737
July	Channel	H	-	CSR	6	1691
July	Creek	I	Y	DIV	42	4041
July	Channel	I	Y	CSR	19	2708
July	Creek	I	Depth	DIV	60	4219
July	Channel	I	Y	CSR	6	2878
Aug.	Creek	H	-	DIV	79	4029
Aug.	Channel	H	-	CSR	4	2359
Aug.	Creek	I	Depth	DIV	100	4224
Aug.	Channel	H	-	CSR	4	3111

location of any other point in the pattern. It was noted that most of these CSR point patterns had fewer than 20 schools of menhaden (12 out of 15 CSR patterns). Only 5 point patterns without first-order structure had second-order structure: 3 of these were in creek habitat and 2 were in river channel habitat (example: [Figure 4.02](#)). All but one of these point patterns had more than 20 schools of menhaden.

Spatial point patterns in the other 23 transects did have evidence of first-order structure, and therefore had a spatially-variable intensity of points dependent on a spatial covariate. Most of these occurred in the creek habitat (19 of 23) and most had more than 20 points in the point pattern (16 of 23). Fifteen point patterns with first-order structure did not have second-order structure (example: [Figure 4.03](#)); 3 of these were in the river channel habitat and the other 12 were in the creek habitat. Only 8 point patterns had both first- and second-order structure (example: [Figure 4.4](#)), and all but 1 of these occurred in the creek habitat.

Table 4.2: Output from model selection and PCFs for transects sampled 2017. Reported values include month and habitat in which the transect occurred, the selected best point process model (H = Homogeneous Poisson, I = Inhomogenous Poisson), spatial covariate for inhomogenous Poisson point processes (X-coordinate, Y-coordinate, or depth), identified Point Pattern (PP; CSR = completely spatially random, DIV = divergence from CSR), the number of schools in the transect, and the length of the transect in meters.

Month	Habitat	Model	Covar	PP	n Schools	Length (m)
June	Creek	I	Y	CSR	5	1780
June	Creek	H	-	CSR	10	2967
June	Creek	H	-	CSR	10	1779
June	Creek	I	Y	CSR	14	2347
June	Creek	H	-	DIV	47	2560
June	Creek	H	-	CSR	5	1834
June	Creek	H	-	CSR	37	3725
July	Creek	I	X	CSR	104	1881
July	Creek	I	Y	CSR	33	2140
July	Channel	H	-	DIV	39	1777
July	Creek	I	Y	DIV	40	652
July	Creek	H	-	CSR	79	1525
July	Creek	I	Depth	CSR	74	3334
July	Channel	H	-	CSR	12	2492
July	Creek	I	Depth	DIV	136	1968
July	Channel	H	-	CSR	10	2656
July	Creek	I	Y	CSR	68	2267
July	Creek	H	-	CSR	115	2720
July	Channel	H	-	DIV	17	3578
July	Creek	I	Y	CSR	94	1838
July	Channel	H	-	CSR	13	2118
Aug.	Creek	I	X	CSR	126	2848
Aug.	Channel	H	-	CSR	4	2756
Aug.	Creek	H	-	DIV	204	2274
Aug.	Channel	I	X	CSR	18	1396
Aug.	Creek	I	Depth	DIV	45	2775
Aug.	Creek	I	Depth	CSR	136	1548
Aug.	Creek	I	Depth	CSR	130	3798
Aug.	Channel	I	X	DIV	18	2595

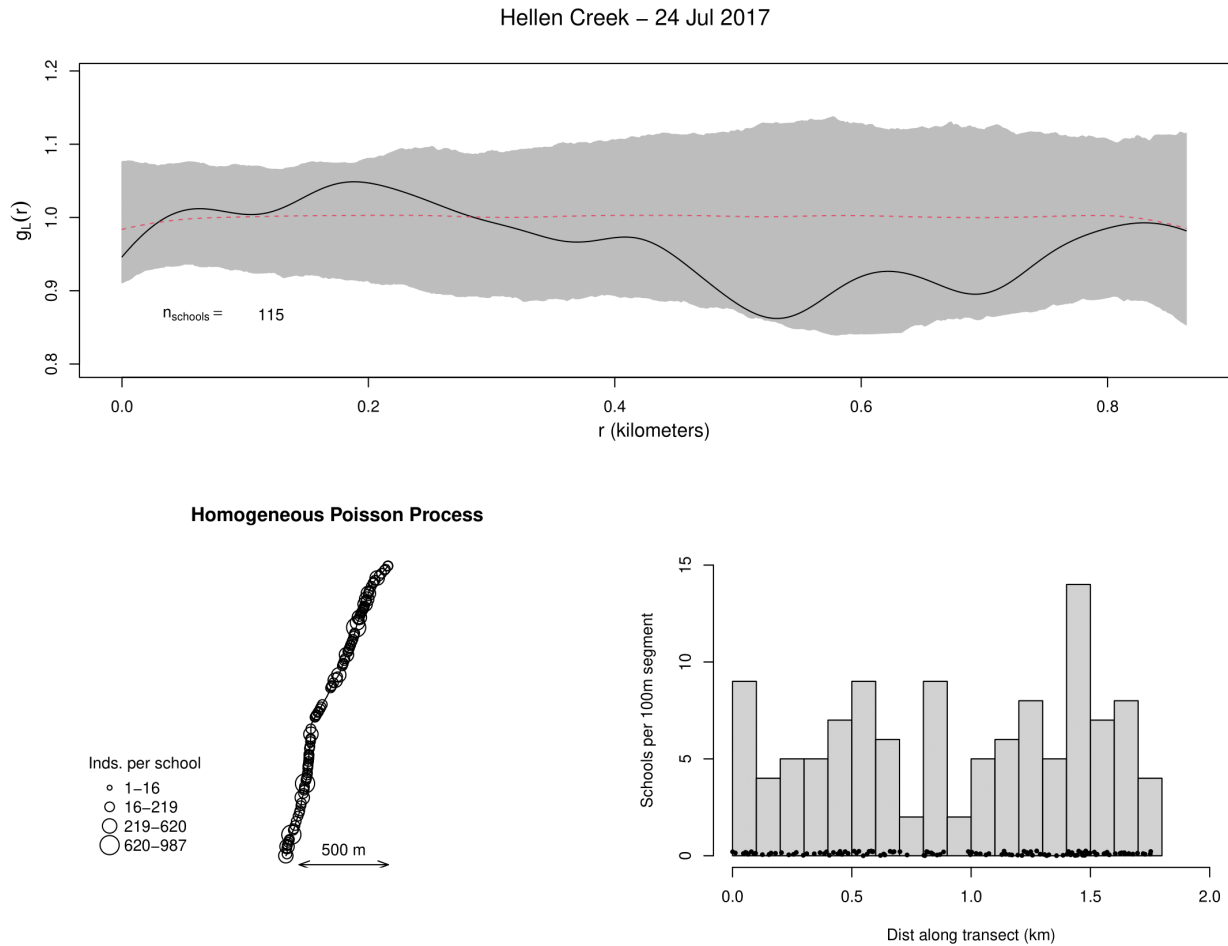


Figure 4.1: All plots are based on a transect with no first- or second-order structure. Counterclockwise from top: PCF output indicating no statistically significant second-order structure. Representation of the transect and the linear point pattern of menhaden schools positioned along the transect. Histogram of the number of schools per 100 m of transect length, with real distance along the transect represented as points.

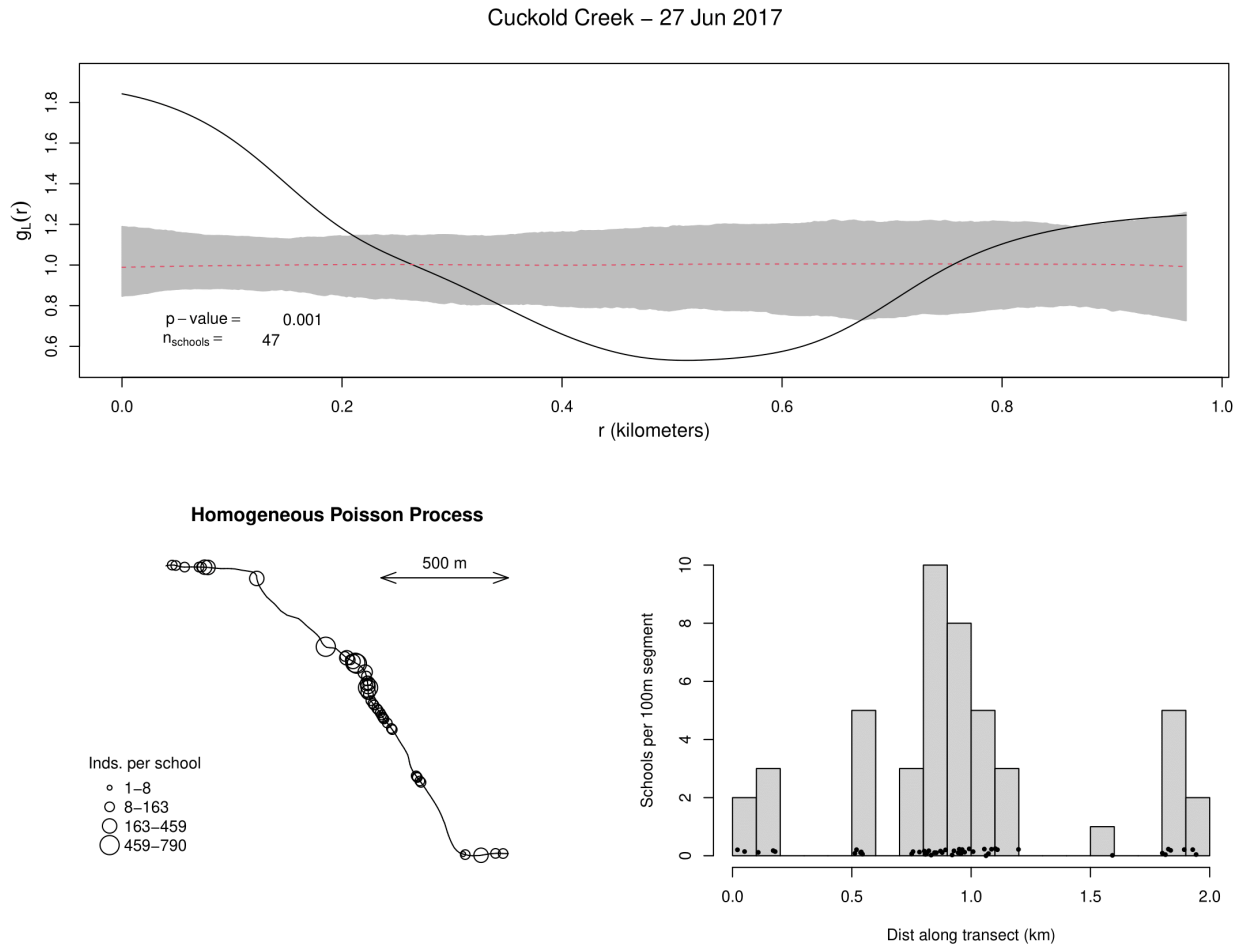


Figure 4.2: All plots are based on a transect with second-order structure, but no first-order structure. Counterclockwise from top: PCF output indicating statistically significant deviations from CSR, p-value reported in the bottom left of the plot. Representation of the transect and the linear point pattern of menhaden schools positioned along the transect. Histogram of the number of schools per 100 m of transect length, with real distance along the transect represented as points.

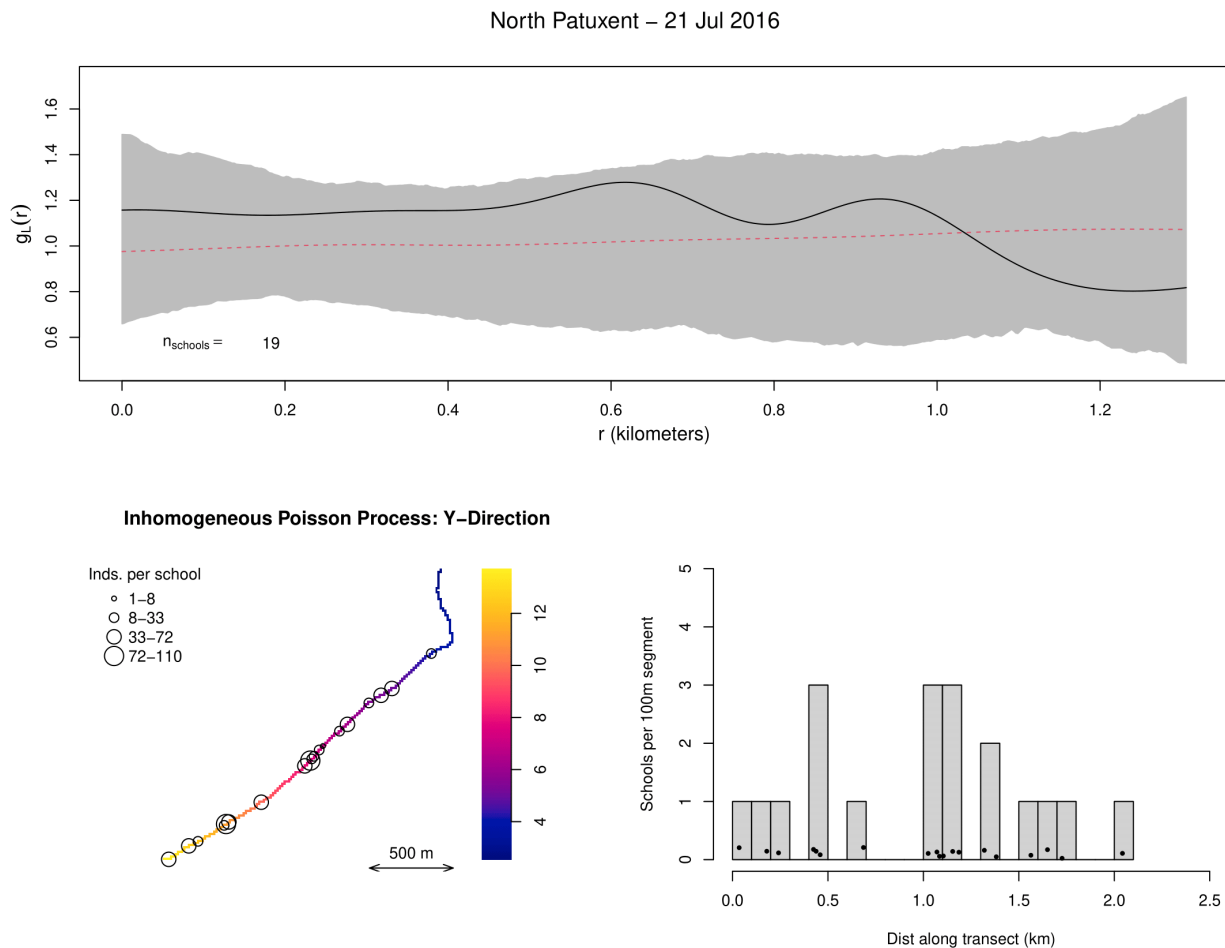


Figure 4.3: All plots are based on a transect with first-order structure, but no second-order structure. Counterclockwise from top: PCF output indicating no statistically significant second-order structure. Representation of the transect and the linear point pattern of menhaden schools positioned along the transect, with the color of the line representing the number of points expected given the variation in spatial intensity. Histogram of the number of schools per 100 m of transect length, with real distance along the transect represented as points.

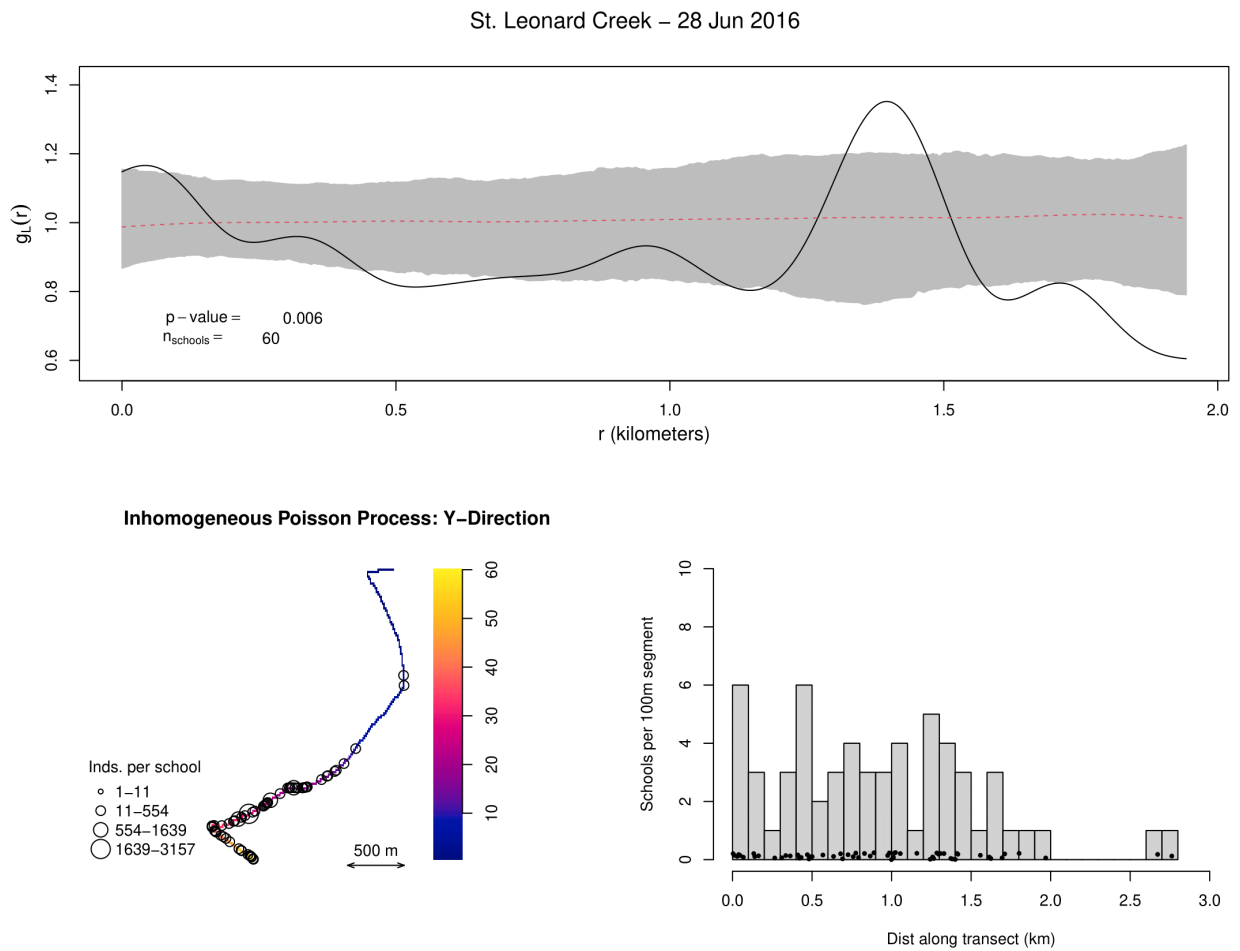


Figure 4.4: All plots are based on a transect with both first- and second-order structure. Counter-clockwise from top: PCF output indicating statistically significant deviations from CSR, p-value reported in the bottom left of the plot. Representation of the transect and the linear point pattern of menhaden schools positioned along the transect, with the color of the line representing the number of points expected given the variation in spatial intensity. Histogram of the number of schools per 100 m of transect length, with real distance along the transect represented as points.

Point patterns with first-order structure most often had variation in point intensity λ with a coordinate-based spatial covariate. Of the 23 total point patterns with first-order structure, longitude was identified as the best spatial covariate for 4 point patterns and latitude was identified as the best spatial covariate for 12 point patterns. The best spatial covariate for all point patterns in the river channel habitat with first-order structure was either latitude or longitude. Plots of $\hat{\rho}$ indicated that the highest intensity λ of points occurred at either of the extremes of the ranges of the spatial coordinate ([Figure 4.05](#) and [Figure 4.06](#)); this translates to a higher density of menhaden schools at the heads and mouths of creeks or nearer to the shores of the river channel than would be expected with constant intensity. The location of the highest spatial intensity along the possible range of latitudes or longitudes flipped often. For example, the area of a creek transect with the highest intensity λ of points would often change from head to mouth with consecutive sampling events.

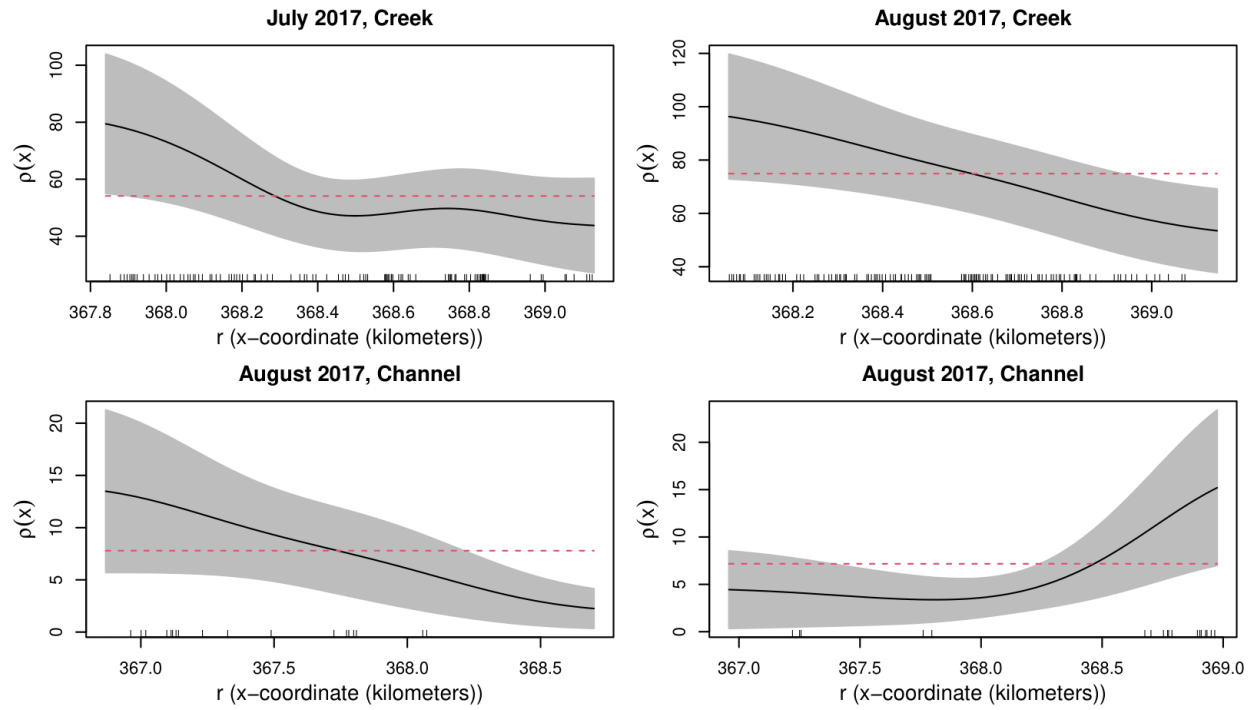


Figure 4.5: $\hat{p}$ plots of IPPP-modeled transects with longitude as a spatial covariate. Grey bands are the 95% CI, and red lines represent the average for the range of longitude.

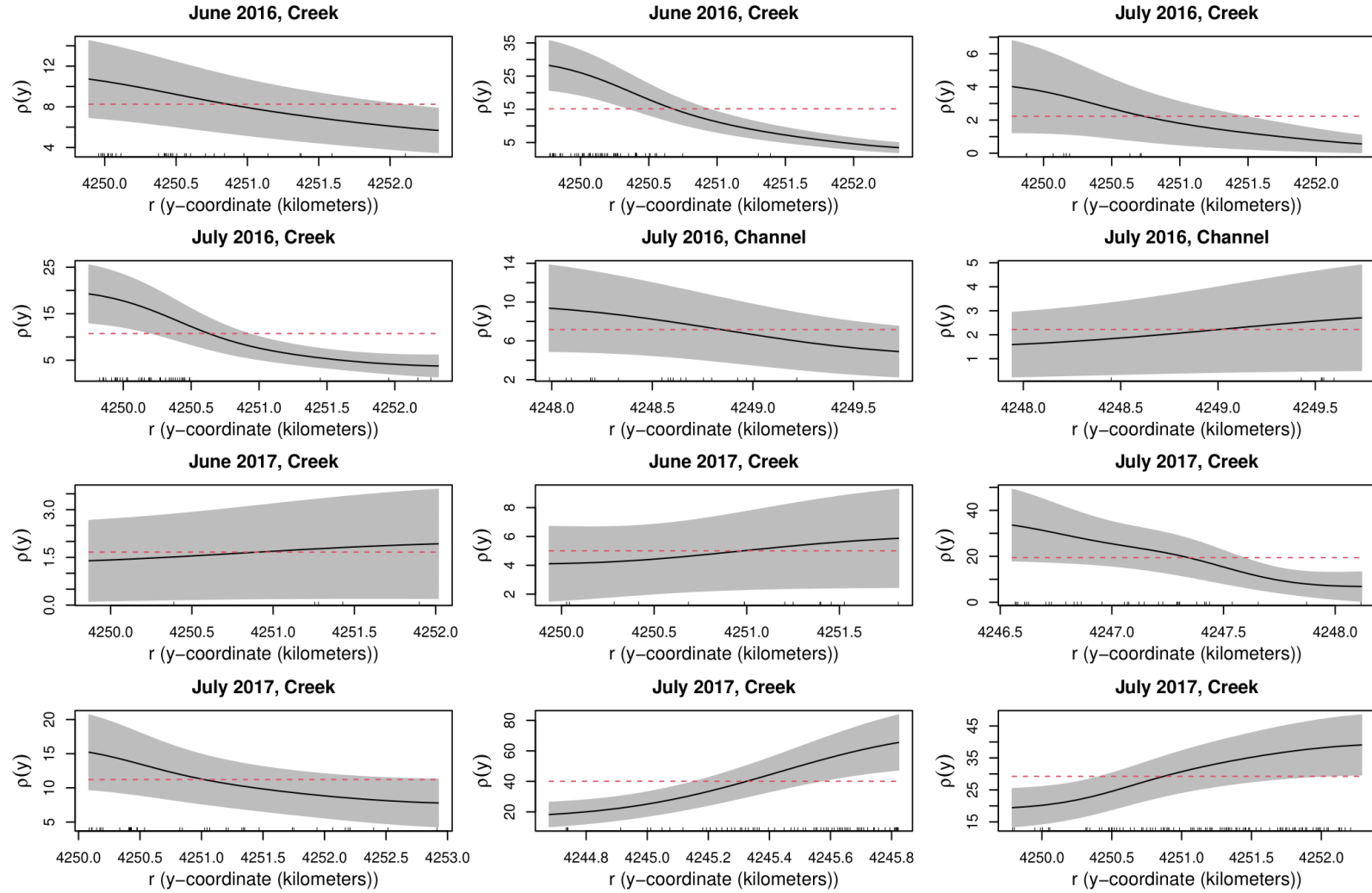


Figure 4.6: $\hat{p}$ plots of IPPP-modeled transects with latitude as spatial covariate. Grey bands are the 95% CI, and red lines represent the average for the range of latitude.

Depth was identified as the best spatial covariate for 7 point patterns, all of which were in creek habitat. The shape of the $\hat{\rho}$ curve generally indicated that menhaden schools were more likely to be found in areas of deeper waters (> 4 m) than would be expected if the intensity was constant (Figure 4.07). There were two instances in which this trend was flipped, and more schools were found in shallower areas (< 4 m) than would be expected with constant intensity.

Most point patterns that had second-order structure followed a pattern of clustering at small values of r and dispersion at larger values of r (Figure 4.08). Of the 13 point patterns that were significantly different from CSR, all but 1 had evidence of clustering at values of r less than 100 m. This is true regardless of the habitat type the point pattern came from or the spatial structure (or lack of spatial structure) of the intensity of the points. In short, this means that there are more schools within 100 m from any arbitrarily selected point along any of these transects than would be expected under CSR. Further, 9 of the 13 non-CSR point patterns had evidence of dispersion at values of r approaching the length of the transect path. This means there are fewer schools than would be expected under CSR at relatively long distances from any point along these transects.

4.4.2 Comparative statistics

The river channel habitat was previously found (Chapter 3) to have fewer observed menhaden schools than the creek habitat. Point patterns in the river channel habitat mostly have neither first- nor second-order structure and are therefore CSR point patterns. A Chi-square test of independence was used to test the relationship between habitat type and first-order structure. The result ($p= 0.024$) indicates that these categories are significantly negatively associated. There are significantly fewer point patterns in the river channel habitat with first-order structure than

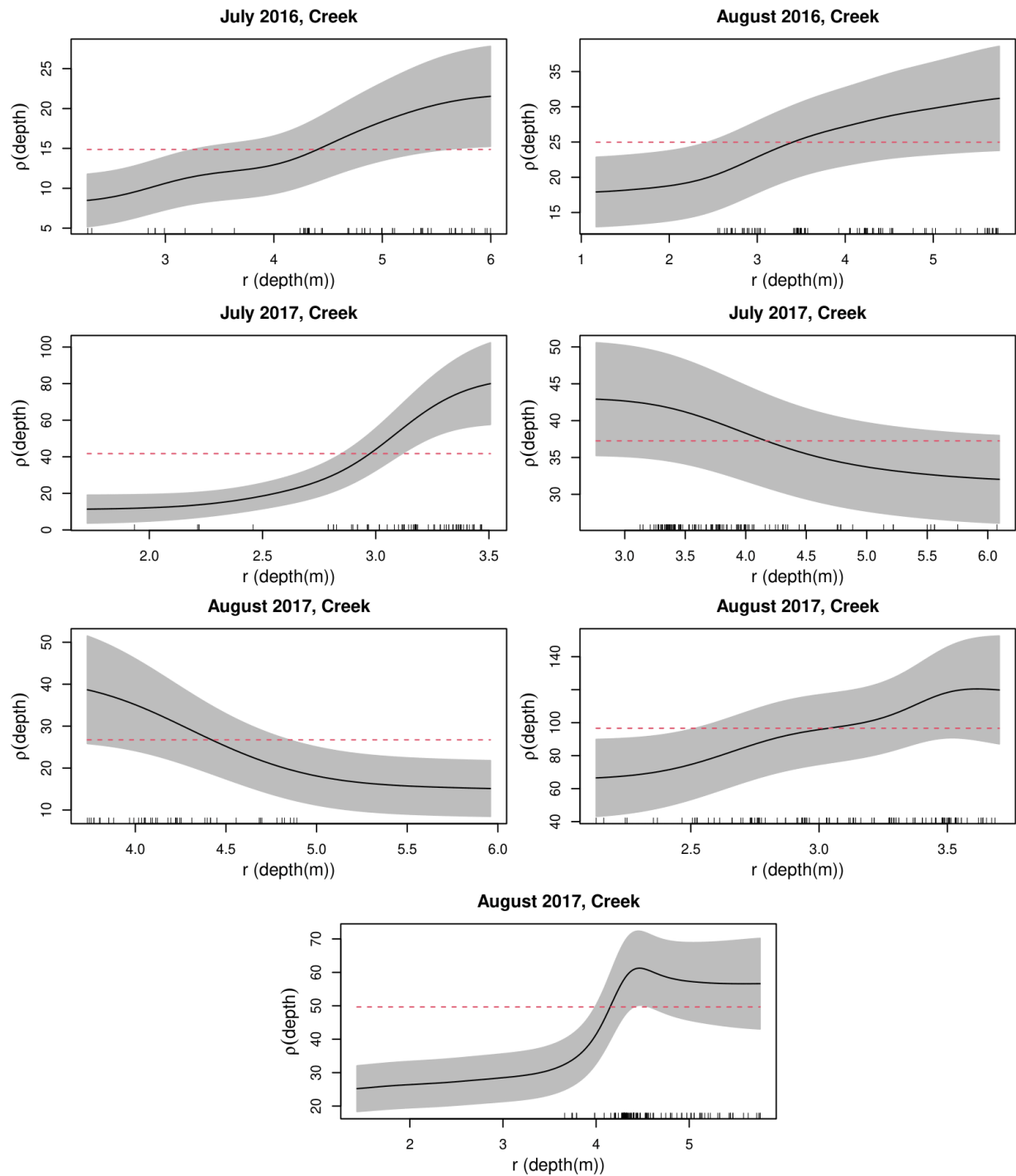


Figure 4.7: $\hat{\rho}$ plots of IPPP-modeled transects with water depth (m) as a spatial covariate. Grey bands are the 95% CI, and red lines represent the average for the range of water depth.

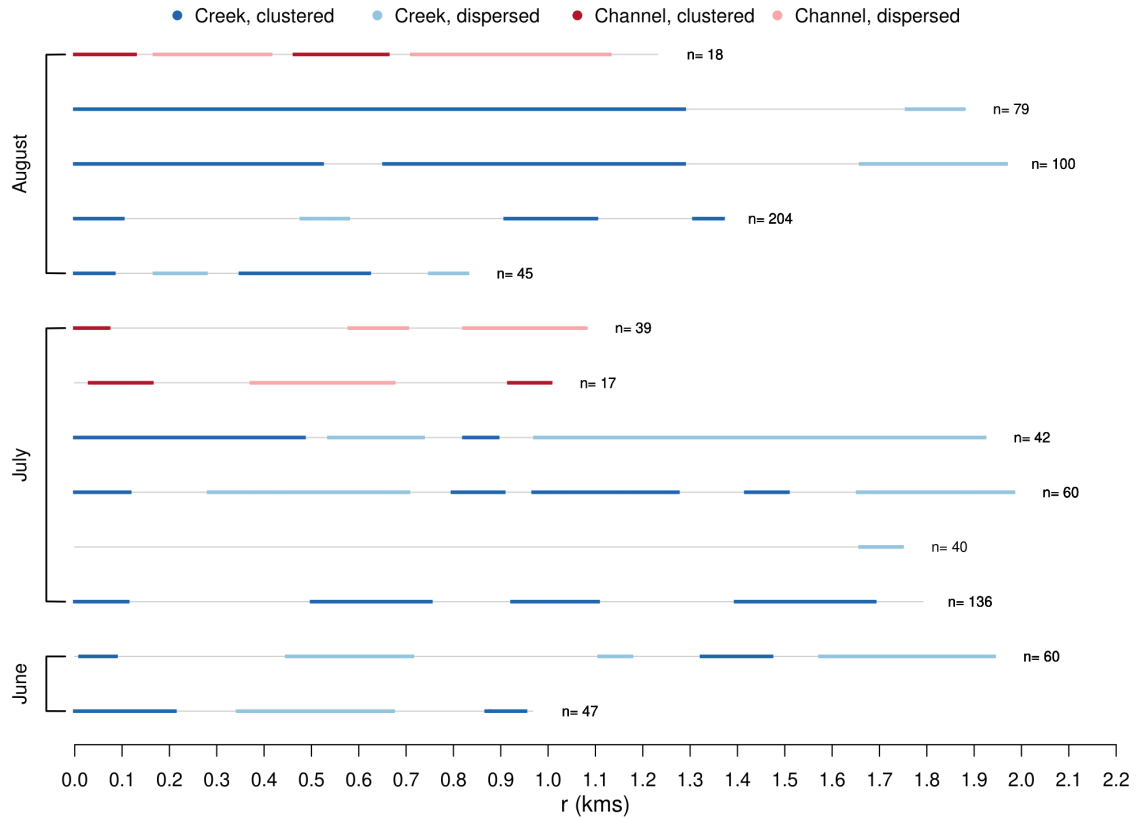


Figure 4.8: Values of pointwise distances r for all non-CSR point patterns. Each horizontal line represents one transect that had statistically significant results for second-order structure as indicated by MAD tests. Underlying light gray lines represent the length of the transect. The color of the patches indicates which habitat the transect is from—blue colors represent the creek habitat and red colors represent the river channel habitat. Darker colors indicate evidence of clustering at that pointwise distance r and lighter colors indicate evidence of dispersion at that pointwise distance r . n values indicate the number of menhaden schools in the transect.

expected if there were no association. A separate Chi-square test of independence provided no evidence of an association between habitat type and point pattern. A Kruskal-Wallis test indicated that point patterns with no first-order structure had significantly fewer schools than point patterns with first-order structure ($p = 0.015$). A second Kruskal-Wallis test indicated that point patterns with no second-order structure had significantly fewer schools than point patterns with second-order structure ($p = 0.015$). Further tests found no significant relationship between transect length and point process model ($p = 0.342$) or transect length and point pattern ($p = 0.146$).

4.4.3 Varied spatial extents

Point process modeling and PCFs were repeated for the transects with longer and shorter extents. The North Patuxent and St. Leonard Creek transects were connected for 10 of the 15 sampling events, which provided an opportunity to test the spatial point patterns of schools across a larger extent. A 500m subset of St. Leonard Creek served as means to test the spatial point patterns of schools across a smaller extent for the same sampling events. PCF plots for all combined-extent transects can be found in Appendix B, and PCF plots for all shorter-extent transects can be found in Appendix C.

The point patterns in all of the 10 combined-extent transects had first-order structure, with longitude identified as the best spatial covariate for 3 of these and latitude identified as the best spatial covariate for the remaining 7. All combined-extent point patterns had evidence of clustering at small values of pointwise distances r , typically less than 400m [Figure 4.09](#). Several also had evidence of dispersion at longer pointwise distances r closer to the maximum length of the transects.

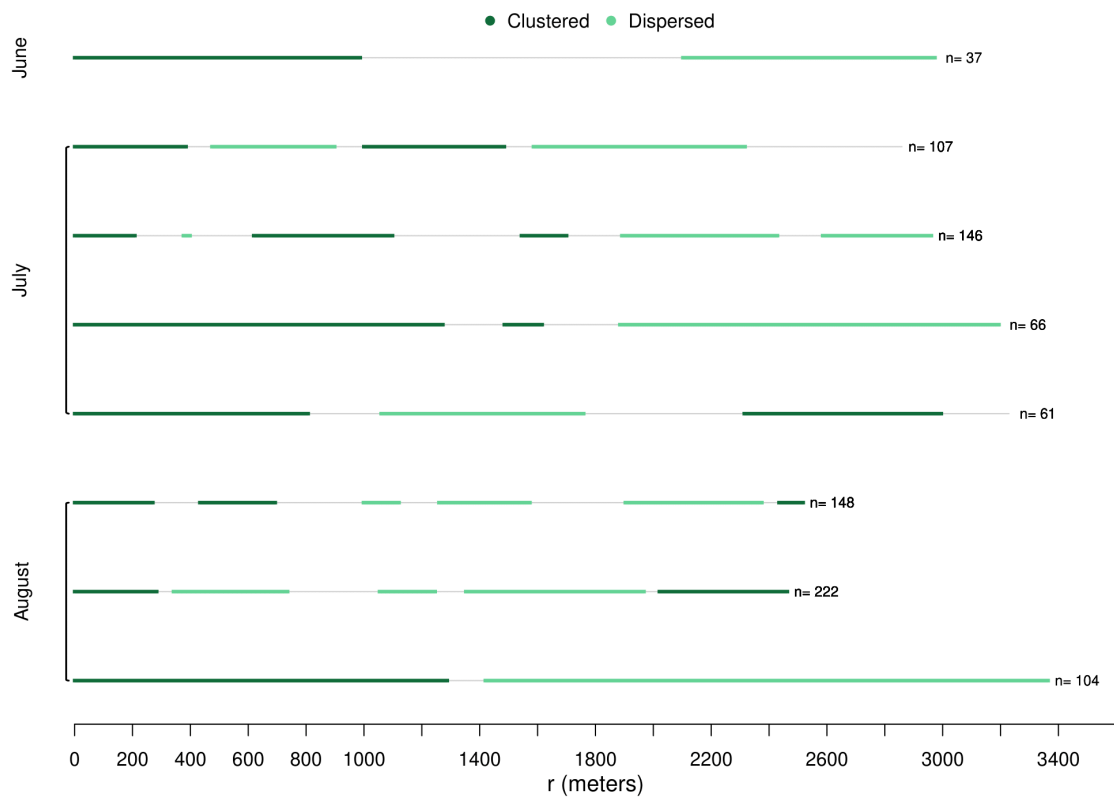


Figure 4.9: Values of pointwise distances r from all non-CSR point patterns in combined-extent transects. Each horizontal line represents one combined-extent transect that had statistically significant results for second-order structure as indicated by MAD tests. Underlying light gray lines represent the length of the transect. Darker colors indicate evidence of clustering at that pointwise distance r and lighter colors indicate evidence of dispersion at that pointwise distance r . n values indicate the number of menhaden schools in the transect.

Analysis of the point patterns within the shorter-extent transects did not identify the same first-order structure, spatial covariates, or second-order structure as those within the combined-extent transects (Table 4.3). Half of the 10 shorter-extent point patterns had no first-order structure. The other 5 did have first-order structure; depth was identified as the best spatial covariate for 4 of these, and longitude as the best spatial covariate for the 1 remaining shorter-extent point pattern. Four of the point patterns had evidence of second-order structure (Figure 4.10). Although all 4 presented evidence of dispersion at values of pointwise distance r approaching the maximum length of the transect, only 2 presented evidence of clustering at short pointwise distances r .

Table 4.3: Output from model selection and PCFs for combined (North Patuxent and St. Leonard Creek) extents and shortened (Southern ~500 m of St. Leonard Creek) extents. Reported values include the date in which the transect was sampled, the selected best point process model (H = Homogeneous Poisson, I = Inhomogenous Poisson), spatial covariate for inhomogenous Poisson point processes (X-coordinate, Y-coordinate, or depth), identified Point Pattern (PP; CSR = completely spatially random, DIV = divergence from CSR), and the number of schools in the extents.

Month	Combined Extent				Shortened Extent			
	Model	PP	Covar	n Schools	Model	PP	Covar	n Schools
June	I	CSR	X	14	H	CSR	-	4
June	I	DIV	Y	37	H	CSR	-	9
July	I	CSR	Y	17	H	CSR	-	6
July	I	DIV	Y	61	H	CSR	-	36
July	I	DIV	Y	66	I	DIV	Depth	36
July	I	DIV	Y	146	H	DIV	-	28
July	I	DIV	Y	107	I	CSR	Depth	12
Aug.	I	DIV	Y	104	I	DIV	X	25
Aug.	I	DIV	X	222	I	CSR	Depth	64
Aug.	I	DIV	X	148	I	DIV	Depth	50

4.5 Discussion

Techniques developed during previous research on this dataset (Lankowicz, in revision) allowed for the identification of discrete schools of menhaden in shallow estuarine waters. The analyses completed here extend our knowledge of menhaden spatial ecology by moving from the

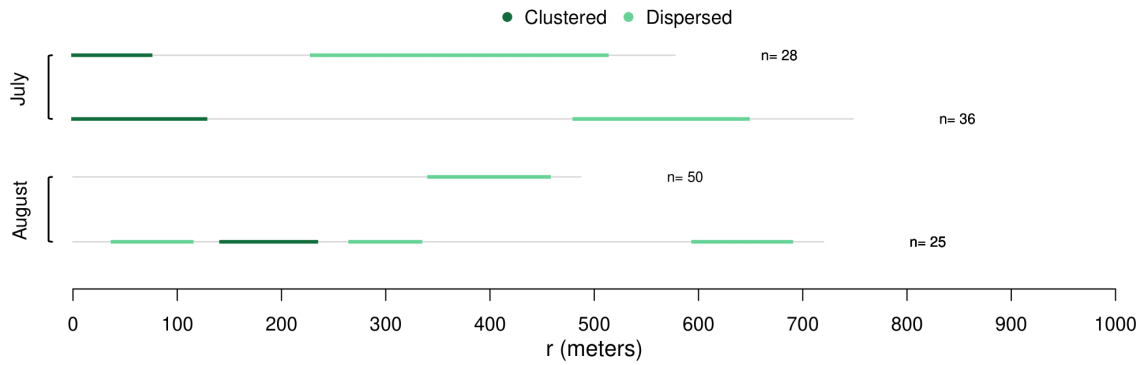


Figure 4.10: Values of pointwise distances r from all non-CSR point patterns in shorter-extent transects. Each horizontal line represents one shorter-extent transect that had statistically significant results for second-order structure as indicated by MAD tests. Underlying light gray lines represent the length of the transect. Darker colors indicate evidence of clustering at that pointwise distance r and lighter colors indicate evidence of dispersion at that pointwise distance r . n values indicate the number of menhaden schools in the transect.

school organizational level to the cluster of schools organizational level. The spatial locations of schools in a transect observed during a sampling event were treated as a spatial point pattern driven by an unknown underlying spatial point process. Through model selection and PCFs, the first-order structure, spatial covariates contributing to first-order structure, and second-order structure were identified for point patterns within 43 sampled transects conducted in two distinct habitats.

Though there is no consensus on the minimum number of spatial points or the size of the observation window needed to apply spatial point pattern analysis, it is notably difficult to accurately identify point patterns and underlying point processes with small sample sizes (Ben-Said, 2021). Point pattern analysis has been completed in plant ecology studies with between 4 and 50 points (Cordero et al., 2016; Wehenkel et al., 2015). There is a consensus within the field of spatial ecology that there is still value in conducting spatial point pattern analysis for point patterns with less than 50 points, particularly as exploratory studies in areas with limited data

availability ([Ben-Said, 2021](#); [Ben-Said et al., 2020](#)).

4.5.1 First-order structure

First-order structure describes the spatial variation in the density of points that comprise a point pattern. If a point pattern has first-order structure, the intensity of points will vary with a spatial covariate. For forage fish like menhaden, these spatial covariates are likely to be multiscale environmental forces and factors such as temperature, food availability, and oceanic currents ([Bertrand et al., 2008](#); [Friedland et al., 1989, 1996](#)). It can be assumed that first-order structure of the point patterns in this study, representing menhaden school distribution, are also mainly the result of spatially- and temporally-variable environmental factors.

The number of point patterns with and without first-order structure was relatively even. HPPP models were identified as the best point process models for 20 of the 43 point patterns, indicating that they did not have first-order structure. 11 of these point patterns were within the river channel habitat, and a chi-square test of independence indicated there was a negative association between the river channel habitat and first-order structure. The river channel does not likely have homogeneous environmental conditions, so this result was somewhat unexpected. A Kruskal-Wallis test indicated that there was a significant difference in the number of menhaden schools in point patterns with and without first-order structure. Point patterns without first-order structure tended to have fewer schools. The river channel habitat is known to have both a lower density of menhaden individuals ([Lankowicz et al., 2020](#)) and fewer schools of menhaden (Lankowicz, in revision). It seems likely that the negative association between point patterns in the river channel habitat and first-order structure is related to the low number of schools typically found in these

areas, and is perhaps not an accurate representation of menhaden school distribution. The exploration into why river channel transects do not have as many observed schools of menhaden as the creek transects cannot directly be answered by spatial point pattern analysis with the spatial covariates available. Previous research has modeled individual menhaden ([Lankowicz et al., 2020](#)) and menhaden school distribution (Lankowicz, in revision) as related to a suite of spatially- and temporally-variable environmental factors. These factors, for example depth, time of year, and tidal stage, likely also serve to structure the distribution of clusters of menhaden schools. Currently, spatio-temporal data of potentially important environmental factors like Chlorophyll-A and dissolved oxygen concentration are not collected in the study area at sufficient resolutions to include in spatial point pattern analysis as spatial covariates. This information would be invaluable to future studies of menhaden school distribution in estuaries.

More point patterns in the creek habitat had first-order structure than would be expected if there was no association between habitat type and first-order structure. Because menhaden school distribution is likely driven at the spatial scale of a transect by multiple spatially-variable environmental factors, this result was expected. What was more notable was the number of individuals per school identified in the creek point patterns with no first-order structure: 6 of the 9 had more than 20 schools of fish. These point patterns should theoretically have enough points to identify first-order structure if it exists. As multiple processes can create or explain the same spatial pattern (a concept often referred to as equifinality), it is relatively uninformative to speculate as to what caused the lack of first-order structure in these point patterns ([Perry et al., 2006](#)).

When spatial point patterns were found to have first-order structure, latitude or longitude was most often identified as the best spatial covariate to describe the variation in point intensity

λ . As noted earlier, these spatial coordinates represent the distance along a transect. There are many physical forces and environmental factors that are likely to vary across these habitats, but again, data to describe these forces and factors at an appropriate grain for this study do not yet exist. Food availability is likely to be an important environmental factor, as it is known to drive juvenile menhaden distribution in other estuaries ([Friedland et al., 1989, 1996](#)).

4.5.2 Second-order structure

Second-order structure indicates that the location of any point in a point pattern is dependent on at least some of the other points in the point pattern. In terms of fish, this nicely matches the mechanisms of schooling and clustering. Individual fish sense and respond to each other in simple ways to create schools ([Aoki, 1982](#); [Huth and Wissel, 1992](#)). This behavior may scale with increasing organizational levels so that schools also seem to sense and respond to each other as an emergent property of individual behaviors. It follows that the point patterns created from these behaviors should have second-order structure.

Though only 3 of the 13 point patterns with second-order structure came from the river channel habitat, Chi-square tests of independence did not find an association between second-order structure and habitat type. This could indicate that the behaviors that drive schooling and clustering do not vary between the two habitat types. There was a notable association between second-order structure and the number of schools. Spatial point patterns with second-order structure had significantly more schools than their counterparts without second-order structure. The detection of second-order structure in a point pattern with few points could be a limitation of the point pattern analysis approach. However, it could also indicate that clusters of schools are

simply more likely to form when more schools are present.

4.5.3 Variation in spatial extent

In this work, characterizing spatial point patterns within spatially close but distinct habitats (river channel and creek) was of interest. Extent, as a component of spatial scale, has an impact on observed spatial patterns ([Meentemeyer and Box, 1987](#); [Turner et al., 1989a,b](#)). This was demonstrated in my research when the spatial point pattern analyses were re-run for areas of a smaller extent than a transect and areas of a larger extent than a transect. When a subset of a creek transect was analyzed, much of the variation of the spatial patterns was lost, likely because of a reduction in the range of values of the underlying driver variables. When transects from the two habitats were combined into an area of larger extent, larger-scale processes masked the point patterns within each habitat type. As demonstrated by Kruskal-Wallis tests, transects with and without first-order structure did not have significantly different lengths, nor did transects with and without second-order structure. Therefore, the noted differences in the spatial pattern were probably not simply because of a variation in transect length. It is more likely that the extent better matched the spatial scale of an underlying process creating the point pattern at that chosen scale. The most appropriate choice in extent was to match it to the objectives of the study (i.e., single transects within each habitat type) so that the underlying mechanisms of menhaden spatial distribution within these areas were more likely to be accurately represented.

Chapter 2 related the locations of menhaden schools to each other using a simple daily clustering coefficient (C_{dt}) based on work by [Lucca and Warren \(2019\)](#). Using this approach, most transects had evidence of clustering and very few had evidence of random or dispersed

distribution. The results from this more traditional point pattern analysis seem contradictory in comparison. However, it is important to note that the daily clustering coefficient can only determine point patterns at the resolution of the entire transect. The linear network-adapted spatial point pattern analysis completed in this chapter is better equipped to deal with the unique geometry and constraints of transect-based sampling and is further able to resolve patterns at very fine spatial resolution. There is value in these approaches. The daily clustering coefficient is useful as a simple snapshot of fish school spatial patterns in larger areas. The more complex approach outlined here is more useful in accurately identifying the multiscale patterns of fish school distributions and the spatial drivers of these patterns.

4.5.4 Relative contribution of environmental and behavioral drivers

The variation in first- and second-order structure of point patterns in the two habitat types indicates a variable contribution of environmental and behavioral factors to menhaden school spatial patterns. First-order structure is likely caused by environmental forces which create a heterogeneous environment of suitable and unsuitable patches. Second-order structure indicates that the spatial location of organisms or aggregations of organisms is dependent on the location of other organisms in the studied area, which is representative of schooling behavior in menhaden. It is likely that the overall spatial point pattern of menhaden schools, as described by their first- and second-order structure, is impacted by environmental and behavioral factors, and the relative importance of these factors is spatially and temporally variable.

Previous studies on the relative contribution of environmental and behavioral factors to the distribution of a forage fish species indicated that behavioral drivers like schooling behavior were

more impactful at small spatial scales (10s of centimeters to 10s of meters) and environmental factors like food availability were more impactful at larger spatial scales (10s of meters to 1000s of kilometers) (Bertrand et al., 2008). The results of this study were similar. The influence of second-order structure is most consistent at small values of pointwise distance r . Most point patterns with second-order structure had evidence of clustering at distances less than 100m. There was also a high proportion of point patterns that had evidence of dispersion at values of pointwise distance r that approached the total length of the transects. As these distances were often greater than 1 km, this dispersion may be more reflective of unmodeled first-order structure than it is of true second-order structure.

4.5.5 Study limitations

There are several potential limitations to both the ARIS sampling method and the spatial point pattern analysis method. The ARIS has a finite field of view and objects will invariably exist outside the observable window. As such, there are many more schools of menhaden in the study areas than what was observed. Further, the locations identified as the center of each school may not be the true geometric center of the schools if a school extended beyond the ARIS field of view. Avoidance behavior is also a concern. As mobile sonar imaging is fairly novel, there is not yet a method to identify or quantify the impacts of possible avoidance behavior. Fish schools are highly dynamic and likely to merge and break apart in response to many external cues (Pitcher et al., 1996; Soria et al., 1998). Noise from the vessel platform or the ARIS itself could be a cue that impacts the distribution of the menhaden schools. It is important to note that limitations in observation windows and possible avoidance behaviors are a problem for all physical sampling

methods, including well-established instruments like trawls, seines, and echosounders. Despite its limitations, fine-scale sonar imaging has proven to be a useful approach that produces novel information on menhaden spatial dynamics in shallow estuarine environments.

Characterizing point patterns along line transects has been a long-established approach in plant ecology, but it is far less common for mobile species. Point pattern analysis is difficult for highly mobile organisms, particularly within restricted observation windows. However, this approach is still possible if the locations of the organisms within a limited window of time are treated as points within a point pattern (example: [O'Driscoll \(1998\)](#)). More difficulty arises when there are very few points within a point pattern, which may result in an inaccurate representation of the true first- and second-order structure ([Ben-Said, 2021](#)). The true spatial point pattern can be further obfuscated if the spatial extent of the study area is not carefully considered. Though it is not possible to state with certainty that the extent chosen in the study encapsulated all environmental and behavioral mechanisms that contribute to menhaden school distribution in the creek and river channel habitats, the comparison of spatial point patterns from larger and smaller extents indicates that most major processes were included.

4.5.6 Conclusions

This research exemplifies the utility of sonar imaging cameras to identify novel spatial dynamics of forage fish like menhaden in under-studied shallow estuarine waters. The sampling method is non-invasive and able to sample a relatively large extent given its very fine spatial resolution, which allowed for an examination of menhaden school and cluster spatial distribution.

Spatial point pattern analysis has not often been used as a tool to examine fish school

distribution in transect-based sampling approaches. Though this method has limitations when dealing with highly mobile species and few schools within a transect, point pattern analysis did reveal multiscale structural mechanisms behind menhaden school distribution. A concrete quantification of the contribution of environmental and behavioral drivers to menhaden school spatial patterns is still elusive given limitations in fine-scale environmental data availability. However, the relative impact of environmental and behavioral drivers found in this study matches that of previous works on clupeiforms ([Bertrand et al., 2008](#)). At large spatial scales, variations in environmental drivers are more impactful and create point patterns with first-order structure. At small spatial scales, schooling behaviors are more visible and create point patterns with second-order structure. Understanding the conditions under which aggregations of menhaden are formed at any spatial scale will contribute to our overall understanding of their ecology and could aid in the effective management of their population.

Chapter 5: Spatially explicit individual-based models of juvenile Atlantic menhaden distribution and schooling

5.1 Abstract

Individual-based models (IBMs) are valuable tools to explore how interactions between individuals, processes, and the environment create emergent behaviors. IBMs are particularly well-suited to track the movement of individuals through dynamic and spatially explicit environments. An examination of how fish schooling behavior develops from collective individual behaviors and how schools are distributed within a heterogeneous and dynamic environment could elucidate some of the mechanisms that drive fish aggregation and movement. With this in mind, IBMs were developed to model the behaviors of juvenile Atlantic menhaden, a species of forage fish that frequently schools, in a shallow estuarine environment. Two IBMs were created, one in which individuals only sense and respond to the environment and one in which individuals only sense and respond to each other. The results of the two models were compared to empirical data on menhaden schooling patterns collected in a representative estuarine creek using a sonar imaging camera. The environment-responsive model did not produce aggregations consistent with the empirical data. The schooling-responsive model produced schools with similar numbers of individuals and within-school densities as schools in the empirical data, but the movement pat-

terns and persistence of these schools were inconsistent with the empirical data. This indicates that menhaden behavior is shaped by both biotic and abiotic forces and that the influence and interactions between these forces may be scale-dependent. Further, the IBMs were found to be sensitive to the number of individuals and the distance to which individuals were able to sense stimuli. Therefore, the mechanisms that control the formation and maintenance of fish schools may be an emergent property of the interaction between biotic and abiotic processes. These IBMs are the first step towards quantitative assessments of the impact of schooling behaviors and environmental conditions on the fine-scale distributions of menhaden in estuarine environments.

5.2 Introduction

The ecology of any one species, community, or ecosystem is inherently spatial ([Fletcher and Fortin, 2018](#)). The interactions between and among individuals, processes, and the environment are profoundly linked to spatial location. Therefore, knowing the distribution and movement patterns of a species is critical for effective management both at a single-species level and an ecosystem level. The incorporation of spatially-explicit data into population modeling is difficult, as the large extent and fine spatial grain of natural processes are at odds with the need for parsimony in most modeling structures ([DeAngelis and Mooij, 2005](#)). Further, the modeling of variation among individuals, particularly in the perception of and response to spatially-varied processes, is a complex task. Classical population modeling approaches assume populations are monoliths and the response of all individuals will be identical, and so do not address spatial or individual variability ([DeAngelis and Gross, 1992](#)). Modeling species dynamics, especially spatially-linked dynamics, without adequate consideration of individual variability and spatial

dependence is not likely to reproduce the reality of ecological systems ([DeAngelis and Grimm, 2014](#); [DeAngelis and Gross, 1992](#); [Łomnicki, 1988](#)).

Individual-based models (IBMs) can represent complex, highly-detailed real-world patterns and processes by simulating the actions of individuals as they interact with each other and their environment ([DeAngelis and Mooij, 2005](#); [Grimm and Railsback, 2005](#); [van der Vaart et al., 2016](#)). IBMs offer an approach to population and community modeling that can be individually variable, spatially explicit, and dynamic ([Boyd et al., 2018](#); [McLane et al., 2011](#); [Tyler and Rose, 1994](#); [Watkins and Rose, 2013, 2014](#); [Werner et al., 2001](#)). Since the 1980s, the rise in popularity of IBMs as modeling and predictive tools for species behavior and distribution has matched an increase in both the availability of fine-resolution spatial data and computing power ([Huston et al., 1988](#); [Watkins and Rose, 2013](#)). IBMs can be valuable tools for fisheries scientists and managers to model and predict fish population dynamics in changing environments ([Huse and Ellingsen, 2008](#); [Huse and Giske, 1998](#)). The ability to track the state and position of individuals throughout the model run makes IBMs well-suited to examine the movement and habitat use of mobile species ([Grimm and Railsback, 2005](#); [Heinänen et al., 2018](#); [Scutt Phillips et al., 2018](#)). As such, IBMs can produce detailed distribution maps for important species. These distribution maps are considered critical to the development and implementation of ecosystem-based fisheries management (EBFM) ([MSFCMA, 2007](#)), and so it stands to reason that IBMs could contribute to EBFM efforts.

Movement through a heterogeneous space determines an individual's exposure to several influential biotic and abiotic processes, which in turn affect the individual's growth, reproduction, and mortality ([Tyler and Rose, 1994](#); [Watkins and Rose, 2014, 2017](#)). Movement modeled in IBMs is structured as a set of rules which determine the timing of departure and intended

destination of individuals. These rules are frequently conditioned on the ability of an individual to perceive and respond to the local environment (Railsback et al., 1999; Watkins and Rose, 2017). Movement may also be conditioned on ecological interactions with conspecifics. These interactions between individual fish create complex aggregative structures like schools or shoals (Hemelrijk and Kunz, 2004; Huth and Wissel, 1992; Parrish, 1989; Vabø and Nøttestad, 1997). One of the most important components of an IBM is an accurate representation of the movement of individuals, though this is often one of the most difficult aspects of individual-based modeling. IBMs require good knowledge of the ecological and physiological processes and factors that produce the emergent behaviors of modeled individuals (Grimm, 1999; Grimm and Railsback, 2005). This kind of large-scale, highly-detailed knowledge is not easily gathered in aquatic environments, though technological developments in animal tagging and remote sensing have rapidly increased the availability, extent, and resolution of both animal movement data and correlated environmental conditions (Rose et al., 2010, 2001; Scutt Phillips et al., 2018). Despite these improvements, it is still difficult to identify a species' perception of its environment and how it will respond to temporally-variable movement cues (Demšar et al., 2015; Nathan et al., 2008; Watkins and Rose, 2017).

An examination of how schooling behavior develops from collective individual behaviors and how schools are distributed within a spatially explicit, heterogeneous, and dynamic environment could elucidate some of the behavioral mechanisms and environmental characteristics that drive fish movement and aggregation. Several forage fish species would be suitable candidates for this type of study. Forage fish is a collective term for many small and highly abundant fish species which occupy intermediate trophic levels and therefore facilitate energy transfer from lower to higher trophic level species (Bakun et al., 2010). Most species of forage fish aggregate

and move in schools (Fréon et al., 2005; Pikitch et al., 2012). Spatially explicit IBMs have been used to explore the population dynamics of several forage fishes (Brochier et al., 2008; Catalán et al., 2013; Mullon et al., 2002), but as of yet, no IBMs have been developed to explore how schooling and movement rules shape the overall distribution of a forage fish species.

This study developed IBMs to examine the formation of juvenile Atlantic menhaden (*Brevoortia tyrannus*, hereafter: menhaden) schools and the spatial patterns of those schools in a simulated estuarine environment. Menhaden is an economically and ecologically important forage species to many Northwestern Atlantic ecosystems and is known to school (SEDAR, 2020). The simulated environment will be based on a tributary of the Chesapeake Bay. Juvenile menhaden utilize the Chesapeake Bay as a nursery ground and are among the most abundant fishes in the Bay (Jung and Houde, 2003; Wood and Houde, 2003). IBMs are not commonly used to examine spatial or population dynamics of menhaden (exception: Quinlan et al. (1999)), though they have been used in a multitude of studies on bay anchovy (*Anchoa mitchilli*), another important Chesapeake Bay forage species (Adamack et al., 2012; Cowan, 1999; Rose et al., 1999; Wang et al., 1997). These bay anchovy IBMs were made possible by abundant information on the ecology of bay anchovy (Houde and Zastrow, 1991). With the knowledge of juvenile menhaden spatial distribution patterns and school characteristics gained from the previous three chapters, it was possible to create IBMs of menhaden spatial dynamics.

Schooling patterns and distributions were modeled using two separate IBMs—one in which the simulated fish sense and respond to environmental conditions and a second in which simulated fish sense and respond to each other. Though both environmental conditions and the movements of conspecifics are expected to affect the movement of an individual, the sensory models were not used in the same IBM. The relative contribution of both sensory mechanisms to the spatial

pattern of schooling fish is likely scale-dependent ([Bertrand et al., 2008](#)). The quantification of how to combine these sensory mechanisms to simulate realistic aggregations is a complex challenge. Separating the sensory models and comparing the characteristics and spatial patterns of any resulting aggregations to each other and fine-scale empirical data is a good first step towards making this challenge more tractable.

5.3 Methods

In broad terms, both the environment-responsive model and the schooling-responsive model consist of a spatially explicit landscape, a set of non-uniform individuals, and a set of sub-models that control the movement and characteristics of the individuals. In both IBMs, the modeled landscapes were designed to have edges that connected at the x-minimum and x-maximum to reduce the likelihood of edge effects. Edges at the y-minimum and y-maximum were not connected, and individuals with movement paths that would cross the y-minimum or -maximum were instead reflected towards the center of the grid. A brief description of both models will be provided in this section; a full description and code to reproduce both models are provided in Appendix D.

5.3.1 Empirical data

An Adaptive Resolution Sonar Imaging (ARIS) camera was used to identify and enumerate menhaden in the lower Patuxent River, which is a mesohaline and tidally-influenced estuarine tributary. Sampling took place across the river channel and within the three largest creeks connected to the sampled portion of the river. The distribution of these individuals into discrete schools and the characteristics of the schools were calculated. Details on the specific methods

used and transect locations have been provided in the previous three chapters. The empirical data used in this chapter as a comparison to model outputs were taken specifically from St. Leonard Creek in June. St. Leonard Creek is the largest of the three creeks connecting to the lower Patuxent River, and the area sampled was approximately 2500m long by 500m wide. This area consistently harbored thousands of menhaden. It was sampled twice in June 2016 and three times in June 2017.

5.3.2 Simulating menhaden

Both the environment-responsive and schooling-responsive models are reliant on the simulation of realistic menhaden-like individuals with non-uniform characteristics. The exact value of any characteristic for each individual was generated from a folded normal distribution, where the edges of the folds are the possible ranges of the characteristic as taken from the literature. These included body length in mm ([Annis et al., 2011](#)), growth potential for one model time step (mm per s seconds) ([Annis et al., 2011](#); [Humphrey et al., 2014](#); [Luo et al., 2001](#)), swimming speed in body lengths per second ([Durbin and Durbin, 1975](#); [Luo et al., 2001](#)), and sensory distance in meters.

A sensing mechanism must be created for the simulated individuals to sense any stimuli. The chosen mechanism for this modeling effort is a modified neighborhood search. Simulated individuals can sense stimuli within a circular neighborhood centered on their location, which is labeled the sensory neighborhood. The radius of this circle is the sensory distance. For both the environment-responsive model and the schooling-responsive model, the sensory neighborhood was limited to an arc 135° to the left and right of the bearing of the individual ([Figure 5.01](#)).

This prevents individuals from sensing the area directly behind themselves. This method is common to models of fish aggregation (Aoki, 1982; Gautrais et al., 2008; Huth and Wissel, 1992, 1994; Oboshi et al., 2002). Though post-metamorphic juvenile menhaden are obligate filter feeders (Friedland et al., 2005) and perhaps do not rely on visual cues to respond to environmental conditions, this method of defining sensory neighborhoods was retained for the environment-responsive model. This was done both to keep the models as similar as possible apart from the sensory stimulus and because the first-order correlated random walk style of movement used would prevent unrealistic and abrupt “about-faces.”

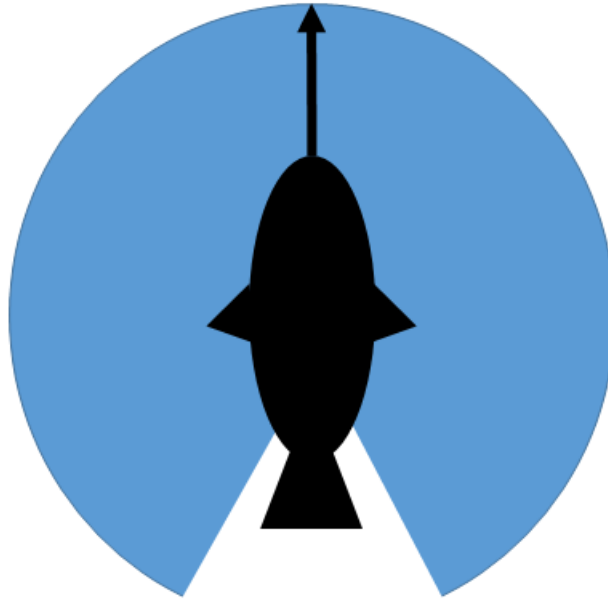


Figure 5.1: Illustration of sensory-distance mediated sensory neighborhood. The sensory distance of an individual creates a circle of radius r around the center of the individual. Bearing, represented by the black arrow, is the direction the individual currently faces. The final representation of the sensory neighborhood of an individual is the partial circle created by an arc that begins 135 degrees to the left of the bearing and ends 135 degrees to the right of the bearing. This sensory neighborhood is represented in blue in this figure.

The initial position and bearing of all individuals were randomized within the modeled landscapes. The value of an individual’s bearing is defined by the unit circle, where $0/360^\circ$ is

represented by the positive side of the x-axis.

5.3.3 Environment-responsive model

5.3.3.1 Model landscape

Simulation of a realistic environment similar to shallow estuarine areas was important for the environment-responsive model. Field data to simulate these conditions do not exist at the fine spatio-temporal scale needed for this IBM. Output from a Semi-implicit Cross-scale Hydro-science Integrated System Model – Integrated Compartment Model (SCHISM-ICM) was used instead; this modeled space is based on the Corsica River, which is a part of the Chester River system and a sub-tributary on the eastern shore of the Chesapeake Bay. SCHISM-ICM was created by Richard Tian of the Chesapeake Bay Program. Model results were subset spatially and temporally to represent the conditions that would likely occur in shallow estuarine creeks in early summer. Environmental conditions for the surface layer of a portion of the model were selected and interpolated to create a 2500m long by 500m wide grid, which approximate the dimensions of St. Leonard Creek (see Chapters 2, 3, and 4). Grid cells were 10m by 10m.

SCHISM-ICM output was subset and interpolated to create hourly 2-dimensional grids of temperature (°C), salinity (PSU), dissolved oxygen (mg L^{-1}), chlorophyll (ug L^{-1}), and tidally-influenced water depth (m). To relate environmental conditions to menhaden movement and behavior, suitability indices were created for the range of each environmental variable, where 0 represents poor suitability for juvenile menhaden and 1 represents superior suitability (Figure 5.02). These suitability indices were generated from either values in the literature or other a priori knowledge of menhaden distribution (Christmas et al., 1982; Dalyander and Cerco, 2010;

Lankowicz et al., 2020; Lynch et al., 2010; SEDAR, 2020). The overall habitat suitability index (HSI) of each grid cell was calculated as the geometric mean of all the suitability index inputs:

$$HSI_{gm} = \sqrt[p]{SI_{EV1} \times SI_{EV2} \times \dots \times SI_{EVp}}$$

where SI_{EV1} is the suitability for environmental variable 1, etc., and p is the total number of environmental variables considered. This method of calculation is highly sensitive to low values; therefore, a single low suitability index value will result in a low HSI value within the grid cell. HSI values closer to 0 indicate poor habitat, while values closer to 1 indicate superior habitat (Figure 5.03). The simulated individuals sense and respond to this value, preferentially moving towards more suitable habitats.

Qualitative explorations into the correlation between the inputs to HSI were completed. This was done because retaining highly correlated suitability indices would essentially “double count” the contribution of one process to overall HSI. Removing either temperature or salinity variables did not impact the overall HSI of the modeled space, which indicates that they are highly correlated in space and time. Neither temperature nor salinity alone has been found to affect menhaden distribution (Friedland et al., 1996; Wilkens and Lewis, 1971), but temperature is known to play an important role in initiating menhaden seasonal migrations (Friedland and Haas, 1988). Therefore, the temperature suitability index was retained and the correlation was removed by excluding the salinity suitability index from the final HSI models.

5.3.3.2 Scheduling and sub-models

The model landscape (HSI grid) updated every 60 minutes, and environmental conditions were static between updates. Sub-models are used to determine the direction and magnitude of

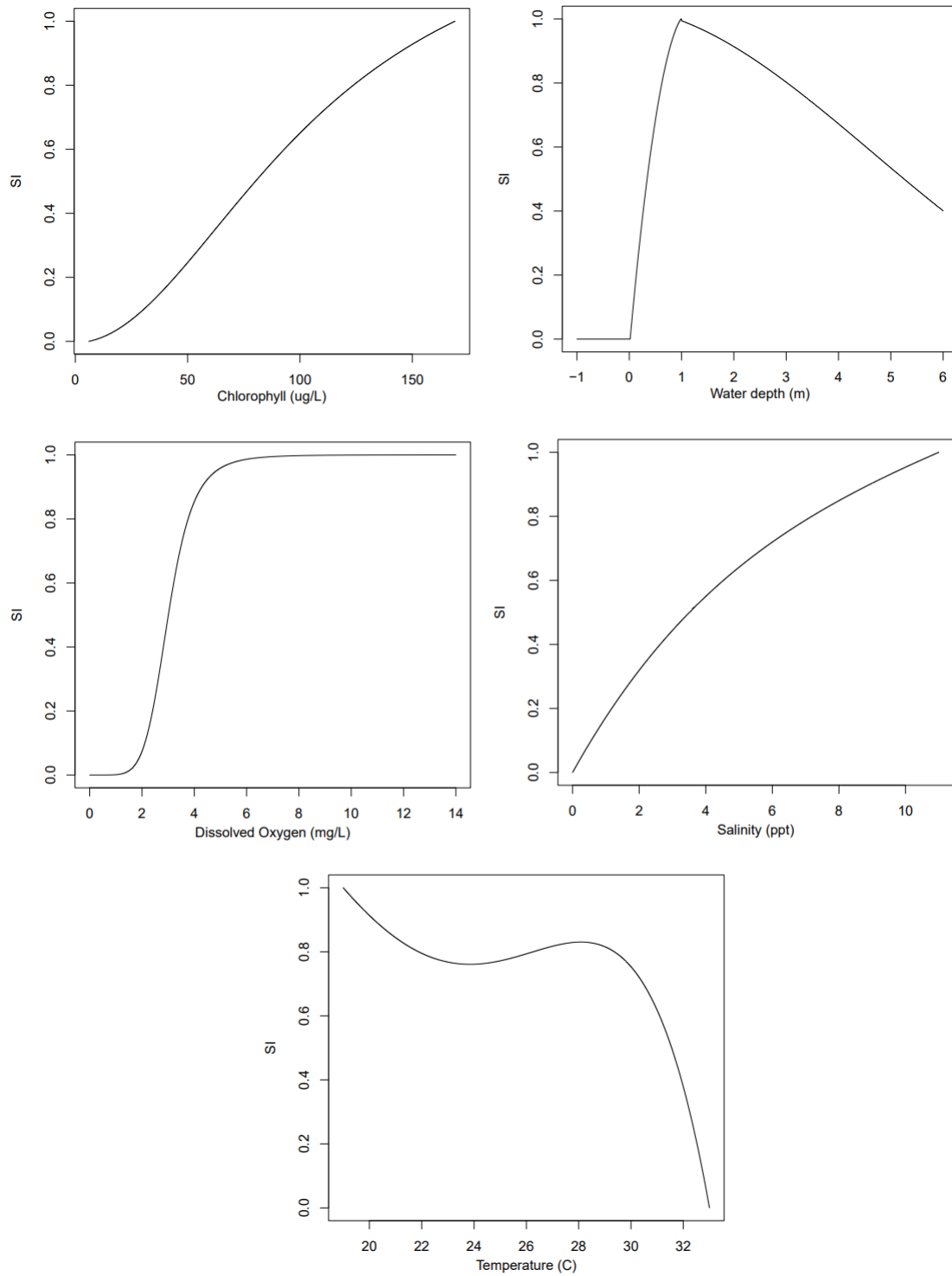


Figure 5.2: Suitability indices for environmental variables used in the environment-responsive IBM.

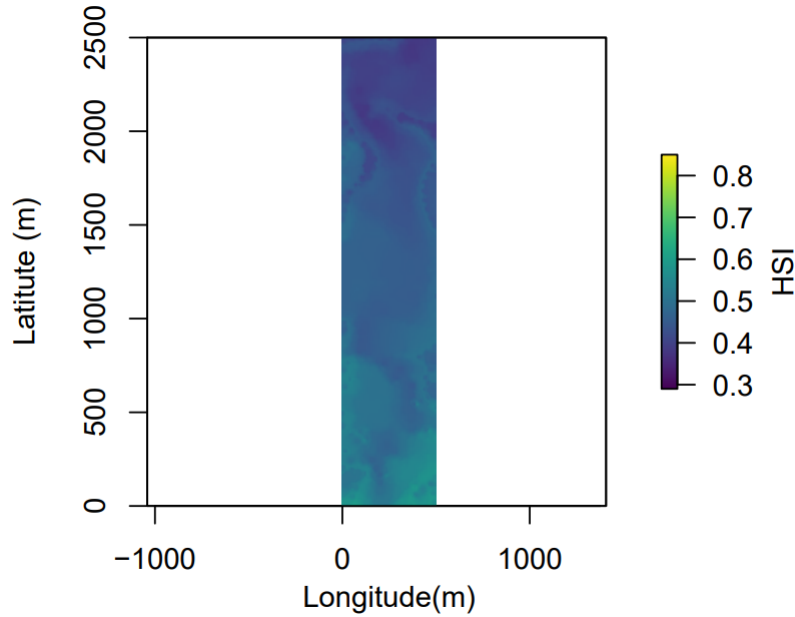


Figure 5.3: Habitat suitability index values for the modeled grid at time step 1. Darker colors indicate poor-quality habitat and lighter colors indicate higher-quality habitat.

a vector that characterizes the movement of an individual within the model landscape. The sub-models are scheduled to update the position of all individuals at a given time step, and the value of this time step is the temporal scale of the model. The temporal scale and the spatial scale of the sub-models must match the process that is being modeled to produce realistic outcomes. For the environment-responsive model, sensing and moving sub-models were scheduled to update the positions of all individuals every 30 seconds, which offers sufficient resolution to create realistic movements. The model was run to represent 30 days of real-world time for 1000 individuals. Individuals first sensed the environment in their respective sensory neighborhoods and then moved accordingly.

5.3.3.2.1 Sensing The sensory sub-model returned the HSI values of all grid cells within the sensory neighborhood of an individual. If an individual was positioned near the x-minimum or

x-maximum, the sensory neighborhood included sensing around the connected edge of the grid. Because HSI is partially conditioned on water depth, the connection between HSI values at the x-minimum and x-maximum for any position along the y-axis is relatively smooth. The same cannot be said for HSI values at the y-minimum and y-maximum for any position along the x-axis—the modeled water depth is much deeper for lower y-position values (the “bottom” edge of the grid) than for higher y-position values (the “top” edge of the grid), which affects the HSI values at these positions.

Each individual then compared the HSI values of all grid cells within their sensory neighborhoods. Menhaden are not expected to have a perfect perception of the environment and perfect decision-making skills. Grid cells must have a difference of more than 0.01 HSI units to be considered different from the surrounding cells. Each individual evaluated all cells within its sensory neighborhood and determined if one or more of these grid cells had a noticeably higher HSI than its current cell. The individual then was assigned a new bearing that brought it towards this desired cell. If the current cell was identified as having the highest HSI or if all cells were perceived to have equivalent HSI, then the individual was assigned a new bearing at random. The choice of random bearing was a first-order correlated random walk, which limited the selection of a new bearing to be within the bounds of the sensory neighborhood. This prevents unnatural “about-faces.”

5.3.3.2.2 Moving The movement sub-model was run after the sensory sub-model. It returned the new bearing of each individual chosen by the sensory sub-model. It then returned the body length and swim speed of each individual and calculated the distance traveled for the length of the time step. From these values, the new x- and y-coordinates for each individual’s position

were calculated. If the movement of an individual carried it out of the modeled landscape, the individual was returned to the sensory sub-model and a new bearing was calculated. This process was repeated until a bearing that kept the individual within bounds was selected.

5.3.3.2.3 Update individual characteristics Immediately after the two sub-models were run, the characteristics of the individuals were updated. The growth potential, sensory distance, and swimming speed of the individuals were redrawn from the same folded normal distributions of the characteristics initialized in the model setup. The body length of the individual was adjusted; body length for any individual i at any time step ts greater than 1 was dependent on the spatial location of the fish xy and was calculated as

$$Bodylength_{i,ts} = Bodylength_{i,ts-1} + Growthpotential_{i,ts} * HSI_{i,ts,xy}$$

After all sub-models and updates were completed, the resulting array of individuals and their characteristics was saved and the model advanced a time step.

5.3.4 Schooling-responsive model

5.3.4.1 Model landscape

Simulation of the environment was not necessary for the schooling-response model, as the individuals were not coded to sense the environment. However, a model landscape was still needed to position the simulated individuals in space. Schooling models are highly computationally expensive, as model structure dictates that the position of all individuals must be compared at every time step to determine which, if any, neighboring individuals will affect the movement of an individual under consideration. Therefore, the modeled grid was scaled down to a 500m

long by 100m space, and only 100 individuals were simulated. As in the environment-responsive model, grid cells were 10m by 10m.

5.3.4.2 Scheduling and sub-models

The schooling-responsive model required a much higher temporal resolution to create realistic outcomes that approximate aggregative behavior (K. Rose, personal communication). If the time step is too coarse, individuals will not be able to “catch up” to their neighbors and create aggregations. However, if the time step is too fine, the model may become intractable due to high computational effort. After some exploratory analysis, 2 seconds was determined to be a sufficiently fine time step to create realistic movement without overburdening computational effort. Modeled time was set to 7 days.

5.3.4.2.1 *Sensing* The sensory sub-model created a point-distance matrix of the Euclidean distance between all pairs of individuals. The distance value takes into consideration the connection between the x-minimum and x-maximum edges of the modeled space. The sub-model then iteratively looped through all individuals. In this loop, point distances from a considered individual to all other individuals were drawn from the point-distance matrix. Then, all individuals further away than the sensory distance of the considered individual were removed from the subset. The next step considered the shape of the sensory neighborhood of the individual. Neighbors within the “blind spot” of the sensory neighborhood were excluded and did not influence the behavior of the individual under consideration. The identity of, bearing of, and distance to the nearest neighbor within the sensory neighborhood as identified by Euclidean distance were recorded.

The influence of the nearest neighbor on a considered individual’s movement was depen-

dent on the distance between the pair (Figure 5.04). These “schooling rules” that define the interactions between individuals are well-established in the efforts to model fish aggregations and are known to create emergent schooling behaviors (Aoki, 1982; Huth and Wissel, 1992, 1994). If the distance between the pair was less than 0.01m, then the considered individual was repulsed from its nearest neighbor. It was assigned a random new bearing to carry it away from the neighbor. If the distance between the pair was greater than 0.01m but less than 1m, then the considered individual aligned its bearing to match its nearest neighbor. If the distance to the nearest neighbor was greater than 1m, then the considered individual was attracted to its neighbor and moved towards it. If there were no neighbors within the sensory neighborhood of an individual, then the individual was assigned a new bearing at random. Just as in the environment-responsive model, the choice of random bearing is a first-order correlated random walk and the selected bearing must be within the sensory neighborhood. As there is little information available on the schooling structure of forage fish like menhaden, the choice of 0.01m and 1m as the repulsion and attraction zone limits, respectively, was arbitrary.

5.3.4.2.2 Moving The movement sub-model was run after the sensory sub-model. It completed the same function as the movement sub-model from the environment-responsive IBM. It returned the new bearing of each individual chosen by the sensory sub-model, and the body length and swim speed of each individual. It then calculated the distance traveled for the length of the time step and the new x- and y-coordinates for each individual’s position. If the movement of an individual carried it out of the modeled landscape, the individual was returned to the sensory sub-model and a new bearing was calculated. This process was repeated until the individual is randomly assigned a bearing that did not carry it out of bounds.

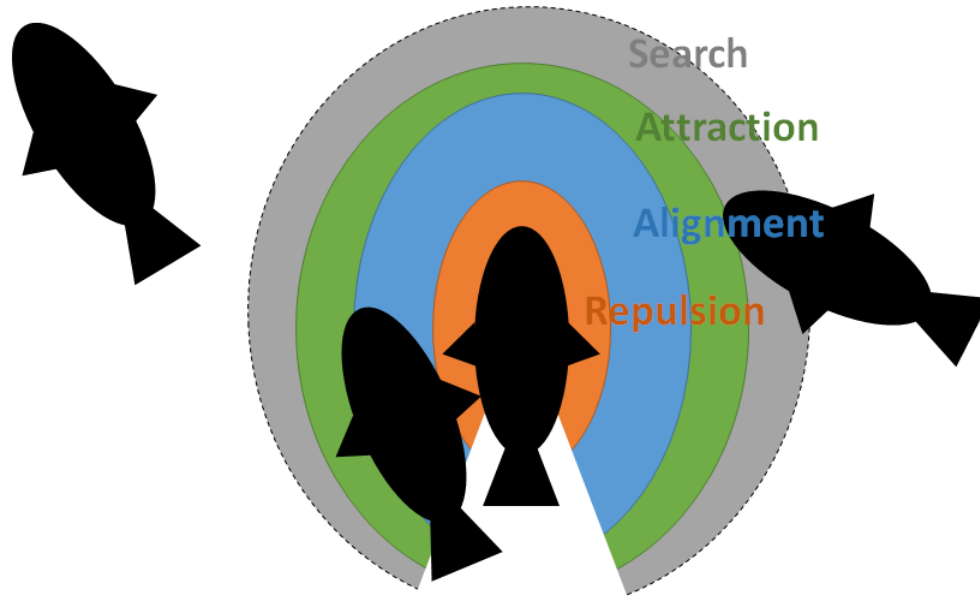


Figure 5.4: Adapted from [Aoki \(1982\)](#). Zone of influence around an individual in a schooling-responsive model. Interactions between individuals are structured based on these distance-based zones. The overall size and shape of the combined zones of influence are equivalent to an individual's sensory neighborhood.

5.3.4.2.3 Update individual characteristics The characteristics of individuals in the schooling-responsive model were updated in the same way as in environment-responsive model. Growth potential, sensory distance, and swimming speed for each individual were redrawn from the folded normal distributions of the variables. Body length was calculated given the body length at the previous time step and the growth potential. After all sub-models and updates were completed, the resulting array of individuals and their characteristics was saved and the model advanced a time step.

5.3.5 Identification of aggregations

Once both models complete the prescribed number of time steps, the positions of individuals at every time step were analyzed for the presence of aggregations. The characteristics used to identify potential aggregations were taken from the results of empirical data analysis presented in Chapters 2, 3, and 4. It is important to note that the non-specific word aggregation is used here so that groups of individuals from both model types can be referred to when describing the steps in the analysis, as it was found that aggregations formed in the environment-responsive model did not have the characteristics of a school. Aggregations from the schooling-responsive model were likely to consist of individuals swimming in a coordinated and polarized manner and were therefore be referred to as schools ([Fréon et al., 2005](#)). Aggregations from the environment-responsive model were not polarized and did not seem to coordinate movement, and therefore did not fit the definition of either schools or shoals ([Hamilton, 1971](#)).

Each time step in a model's outcome was recorded as a matrix with n rows representing the number of individuals and k columns representing the tracked characteristics of those individuals (e.g. sensory distance, body length, etc.). Aggregations were identified for all time steps. First, a Euclidean distance matrix was created to find the distance between all pairs of individuals. An agglomerative hierarchical model was used to identify individuals within 1 meter of each other as belonging to the same aggregation. An aggregation must consist of 4 or more individuals. Not all individuals within an aggregation need to be within the same 1-meter distance; for example, if 4 individuals were positioned in a straight line where each individual was 99 cm away from the next closest individual, this would still be considered an aggregation. The number of aggregations per time step was recorded, as was the number of individuals in each aggregation and the position

of the geometric center of each aggregation. The empirical data covered in Chapter 3 provided measures of the number of aggregations and the number of individuals per aggregation to which the modeled results were compared.

A concave-hull polygon was drawn around all the points of an aggregation. The area of this polygon was calculated, as was the longest chord within the polygon. The longest chord distance represented the length of the school in meters. The density of individuals within an aggregation was calculated as individuals per meter length. Because the longest chord always served as the representation of aggregation length, the individual per meter length density measure was highly conservative. The empirical data covered in Chapter 3 provided measures of individuals per meter length to which the modeled results were compared.

Point pattern analysis of the resulting aggregations was intended but was not possible. The original objective was to determine if the aggregations of menhaden were positioned relative to each other in any identifiable patterns. However, point pattern analysis requires at least four points, and frequently the IBMs produced fewer than this required number of aggregations.

5.3.6 Sensitivity to number of individuals and sensory distance

The effect of the number of individuals and sensory distance ranges of individuals within a model on the number and characteristics of resulting aggregations was tested. The environment-responsive and schooling-responsive models were re-run with variations in the number of individuals. The environment-responsive model was also re-run with variations in the sensory distance of the individuals.

The environmental-responsive IBM was completed with 1000 individuals. Subsequent runs

were completed with 2000 and 4000 individuals. The schooling-responsive IBM was completed with 100 individuals. Subsequent runs were completed with 200 and 400 individuals. For the initial runs of the environment-responsive model, the range of sensory distances of individuals was set as a folded normal distribution between 15 and 30 meters. In subsequent runs, sensory distances were either doubled or halved. All runs of both IBMs with varied numbers of individuals or sensory distances were analyzed for aggregations in the same way as the original runs.

5.4 Results

5.4.1 Empirical data

The model landscape of both IBMs was based on St. Leonard Creek, which is the longest estuarine creek that connects to the lower Patuxent River. The environment-responsive model was set up to match the dimensions of this area (2500m long by 500m wide), while the schooling-responsive model was a scaled-down version (500m long by 100m wide) with the same ratio of length to width. Across the 5 sampling events, between 5 and 76 schools were observed in the Creek (median: 33 schools). The average number of individuals per school varied between 31.5 individuals and 125 individuals (median: 44 individuals). The total number of individuals observed in each transect varied from 324 to 11432 individuals, with a median of 1957 individuals observed. Average density within the schools, reported as individuals per meter of school length, varied from 12.3 to 44.3 individuals per meter (median: 15.3 individuals per meter).

5.4.2 Environment-responsive model

5.4.2.1 Habitat Suitability Index

The habitat suitability index of the model landscape was highly dynamic (Figure 5.05). It was uncommon that all grid cells would have a similar HSI value, as chlorophyll and dissolved oxygen suitability values shifted rapidly in both space and time. Areas of high suitability were typically located near the left and right edges of the grid, which represented the shallower regions near the shoreline (Figure 5.03). When average HSI was calculated over a 1-hour moving window for the entire modeled period, a clear diel pattern was noted. HSI was highest in daylight hours and lowest at night. Average HSI remained fairly consistent, but dropped dramatically at the end of the modeled period due to decreased suitability of both chlorophyll and dissolved oxygen levels.

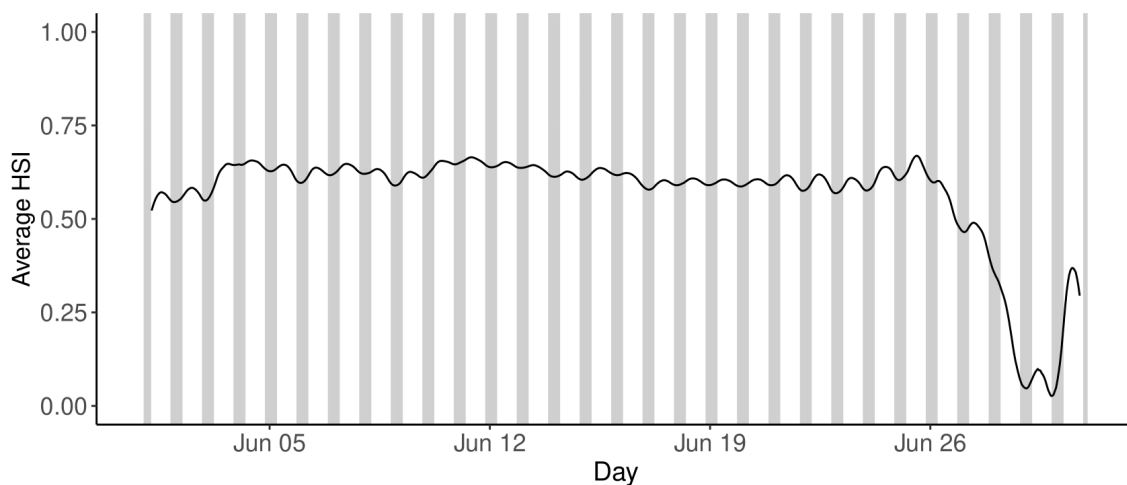


Figure 5.5: HSI averaged with 12-hour moving windows through the 30-day modeled period. Daylight hours have light background and nighttime hours have darker background.

5.4.2.2 Movement

The modeled individuals tracked the shifting location of regions with high HSI. Individuals sensed and moved towards areas of high suitability within their sensory neighborhoods (Figure 5.06). Individuals frequently “wrapped around” the grid, meaning they traversed the continuous connection between the x-minimum and x-maximum (Figure 5.07).

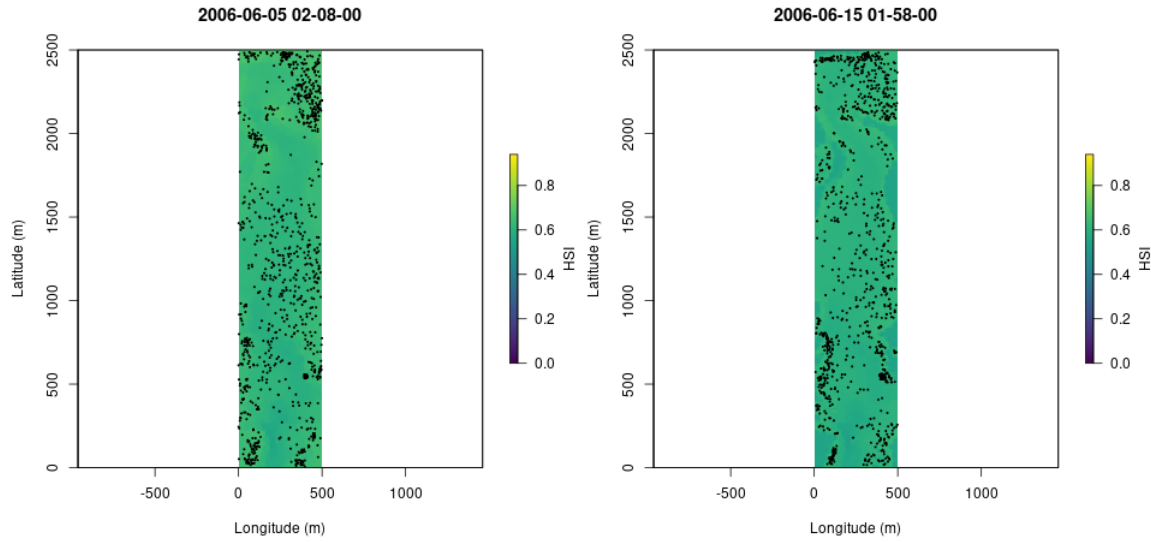


Figure 5.6: Positions of 1000 individuals at two randomly-selected time steps. The background color represents HSI; darker values indicate poor habitat, while lighter colors indicate superior habitat. Individuals are noted to avoid poor habitat.

5.4.2.3 Aggregations

The environment-responsive IBM produced aggregations that cannot be described as schools. Coordinated and polarized swimming is required for an aggregation to be considered a school. The movement rules of this model produced aggregations where individuals were not coordinated, and instead, the bearings of all individuals in an aggregation were uniformly dispersed

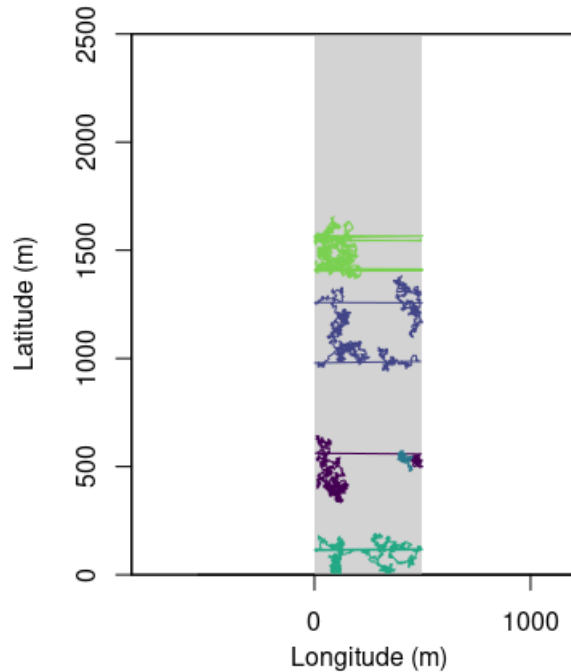


Figure 5.7: Movement paths of 5 randomly-selected individuals over 3 modeled days in the environment-responsive model. Note that long, straight horizontal lines that traverse the entire x-range of the modeled space are not representations of individuals moving 500 meters in a time step; these lines represent individuals “wrapping” around the grid from the x-minimum to the x-maximum in a time step.

around the unit circle ([Figure 5.08](#)).

Very few aggregations were identified at each time step. Because the model was initialized with randomly-placed individuals, this was expected for the “primer” period at the beginning of the model. This period lasted approximately 7 days and was removed when calculating summary statistics. The number of aggregations remained low through the entire 30-day modeled period ([Figure 5.09](#)). When averaged over 12-hour moving windows, between 4 and 8 aggregations were seen in the modeled space. The number of aggregations declined sharply at the end of the modeled period, which coincided with the sharp decline in average HSI. The number of individuals per school also remained low throughout the modeled period, with between 6.0 and

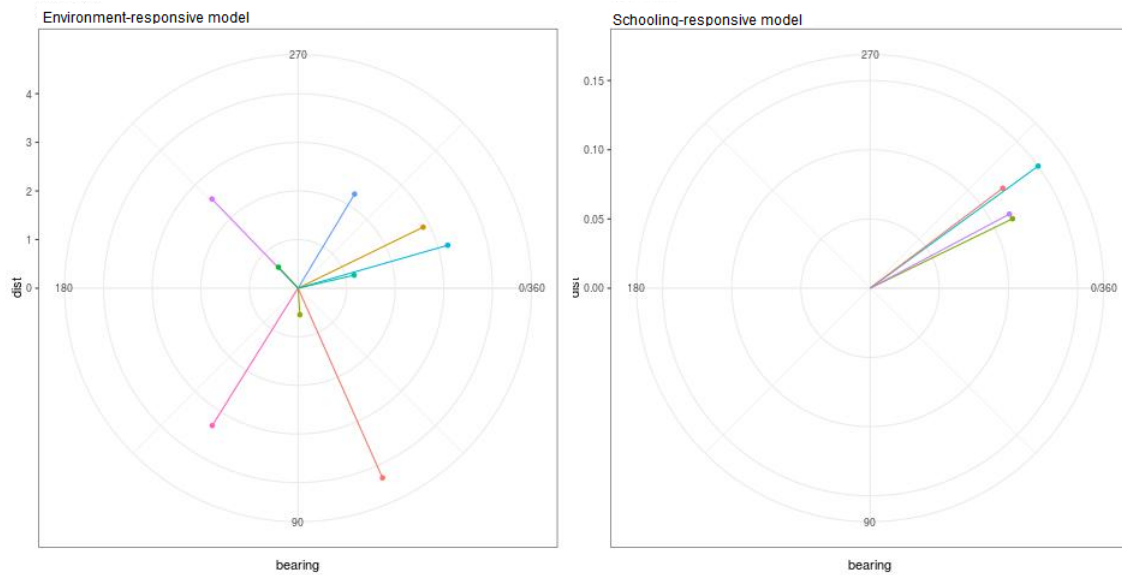


Figure 5.8: Example of the alignment of individuals within aggregations from both models. The left panel is output from the environment-responsive model and the right panel is output from the schooling-responsive model. The movement of individuals is represented as vectors. The direction of the vector is equal to the bearing of the individual and the magnitude of the vector is equal to the distance the individual traveled in a time step.

8.3 individuals per school, and declined sharply at the end of the modeled period with both the number of aggregations and HSI (Figure 5.09). The rapid decline in the suitability of chlorophyll and dissolved oxygen at the end of the modeled period caused the drop in the average HSI of the modeled space. These two environmental variables were highly dynamic in space, whereas temperature was fairly constant across the modeled space and tidal depth had slight and consistent semi-diurnal fluctuations. Therefore, when chlorophyll and dissolved oxygen levels declined, the modeled space had consistently low HSI values in all cells. With little variation between cells and overall low values, there were no suitable patches for the modeled individuals to sense and move toward, and thus few aggregations were formed.

The density of individuals within an aggregation was also low, which is another indication that the aggregations cannot be considered schools. Averaged over a 12-hour moving window,

density within an aggregation ranged from 0.2 to 3.1 individuals per meter (Figure 5.09). Similar to both the number of aggregations and the number of individuals per aggregation, this metric also declined sharply at the end of the modeled period with a decline in HSI.

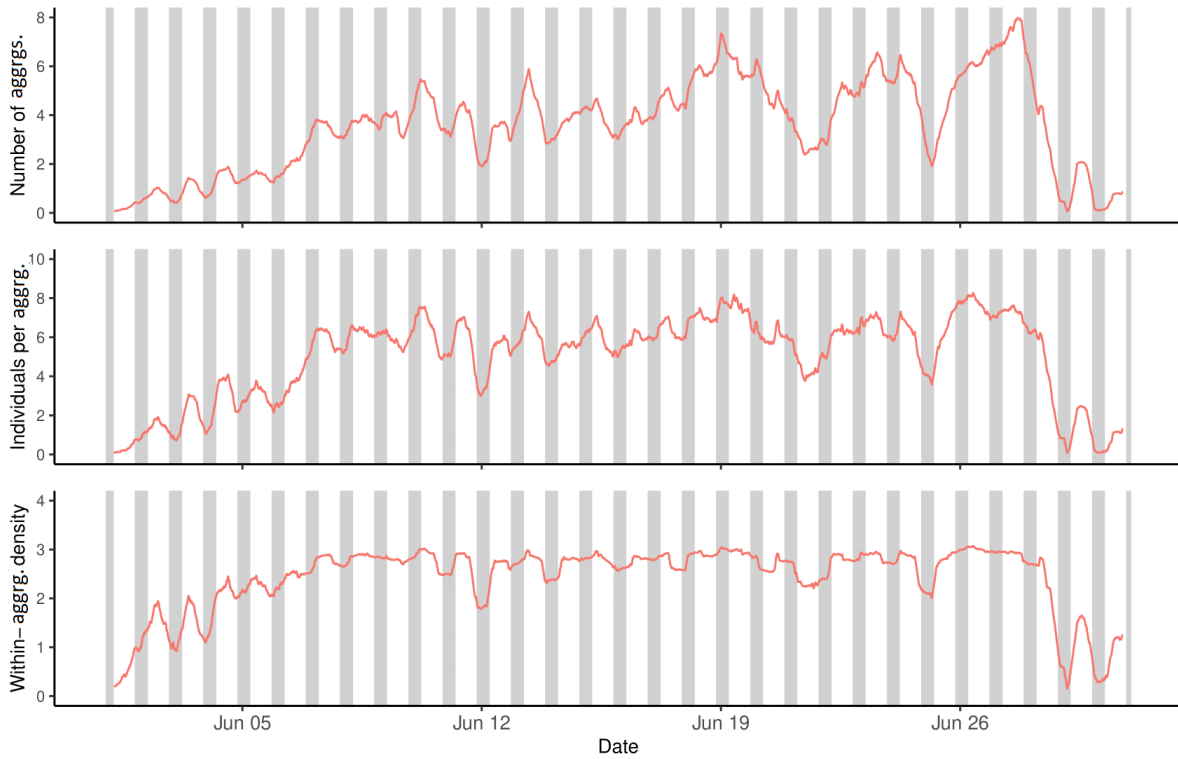


Figure 5.9: Number of aggregations, number of individuals per aggregation, and within-aggregation density (individuals per meter length) averaged over 12-hour moving windows for 1000 individuals in the environment-responsive model. Background color indicates daylight condition. Light-colored background bars indicate daytime, dark-colored background bars indicate nighttime.

5.4.3 Schooling-responsive model

5.4.3.1 Movement

The modeled individuals sensed and responded to each other. Individuals typically remained within the schools that were formed (Figure 5.10). Similar to the environment-responsive

model, individuals frequently “wrapped around” the grid (Figure 5.11). This behavior was in part because schools were not disrupted from moving in the x-direction once they stabilized, though they would be reflected off the y-minimum and -maximum of the grid space if they approached it.

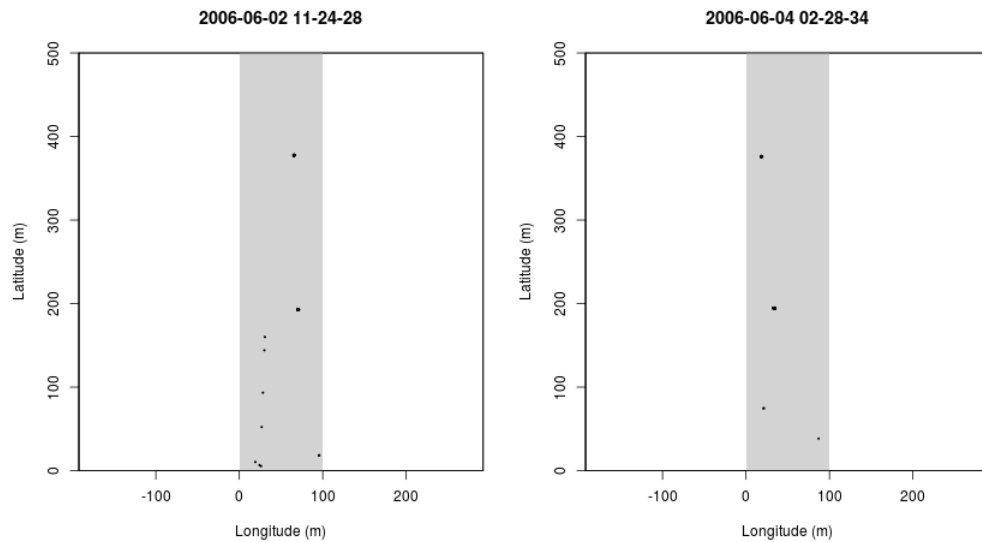


Figure 5.10: Positions of 100 individuals at two randomly-selected time steps. In both sub-figures, the individuals have aggregated into several tight schools.

5.4.3.2 Aggregations

The schooling-responsive IBM produced aggregations in which individuals moved in a coordinated and polarized manner (Figure 5.08). Therefore, these aggregations can be described as schools. Similar to the environment-responsive model, the school-responsive model needed a “primer” period to stabilize. The model needed approximately 3 days to stabilize, and this “primer” period was removed when calculating summary statistics. Very few schools were formed at every time step. However, unlike the environment-responsive model, this was not

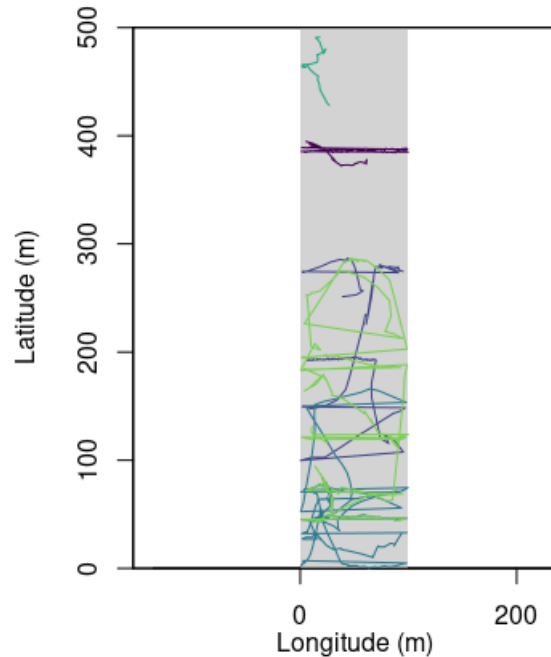


Figure 5.11: Movement paths of 5 randomly-selected individuals over 1 modeled day in the schooling-responsive model. Note that long, straight horizontal lines that traverse the entire x-range of the modeled space are not representations of individuals moving 100 meters in a time step; these lines represent individuals “wrapping” around the grid from the x-minimum to the x-maximum in a time step.

because schools were uncommon. The scarcity occurred because the 100 individuals settled into 2 large, distinct schools, and individuals did not leave the schools they had joined (Figure 5.12). When averaged over 12-hour moving windows, there were between 46 and 50 individuals in each school. The density of each school fluctuated between 24.5 and 27.5 individuals per meter (Figure 5.12).

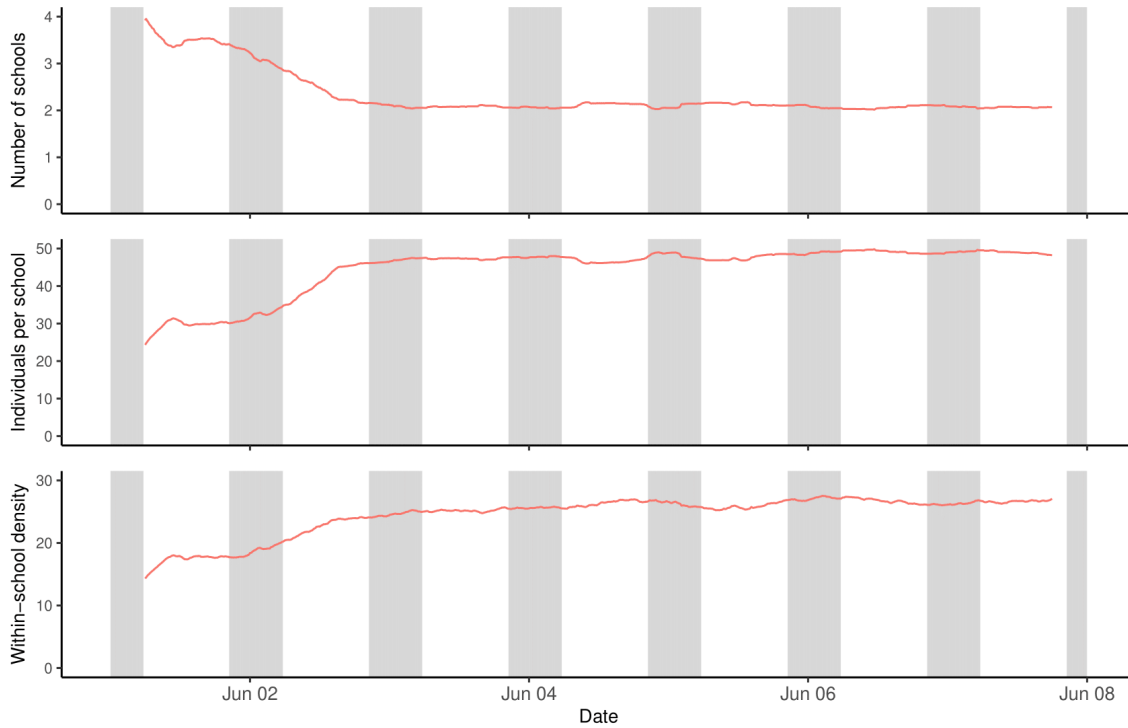


Figure 5.12: Number of schools and number of individuals per school averaged over 12-hour moving windows for 100 individuals in the schooling-responsive model. Though environmental conditions were not considered, background colors for day and night periods were still included so it is easier to visualize the changing of the metrics over the modeled period. Light-colored background bars indicate daytime, dark-colored background bars indicate nighttime.

5.4.4 Sensitivity tests

5.4.4.1 Number of individuals

The environment-responsive model was run with 1000, 2000, or 4000 simulated individuals. All other variables, including model space size, were held at their original values when adjusting the number of individuals (Table 5.01). Therefore, increasing the number of simulated individuals also increased the density of individuals within the model space. When the number of simulated individuals increased, the number of aggregations increased (Figure 5.13A). Though the decline in HSI was noted to cause a decrease in the number of aggregations in the original

1000-individual model, this decrease was not as severe in the 2000-individual model. The increased density of individuals within the model space increases the likelihood that individuals would be positioned close enough together to be identified as an aggregation. It is apparent that the 4000-individual model was best able to produce aggregations in periods of low habitat suitability, as peak number of aggregations was reached during times of low HSI. At this elevated density of individuals in the model space, enough individuals were in proximity to the smaller and fewer patches of relatively high-suitability habitat to form aggregations.

Increasing the number of simulated individuals in the model had a similar effect on the number of individuals per aggregation. Models with more individuals typically had more individuals per aggregation (Figure 5.13B). This is likely also a density-related effect, as described above.

Increasing the number of individuals in the model did not have a clear effect on the within-aggregation density (Figure 5.13C). The 1000- and 2000-individual models had approximately the same diel trends, and within-aggregation density declined sharply with the decline in HSI. The 4000-individual model also showed a slight decline in within-aggregation density with the decline in HSI, though it was not as evident as the other two models. However, within-aggregation density of the 4000-individual model did not show the same diel trends.

The schooling-responsive model was run with either 100, 200, or 400 individuals (Table 5.01). All other variables were held equal when adjusting the number of individuals. Whereas the number of schools in the 100-individual model was very stable at 2 schools over the length of the modeled time, the number of schools fluctuated in the 200- and 400-individual models (Figure 5.14A). The 200-individual model always had more schools than the 100- and 400-individual models. The instability of schools in the higher-density models indicates a non-linear connection

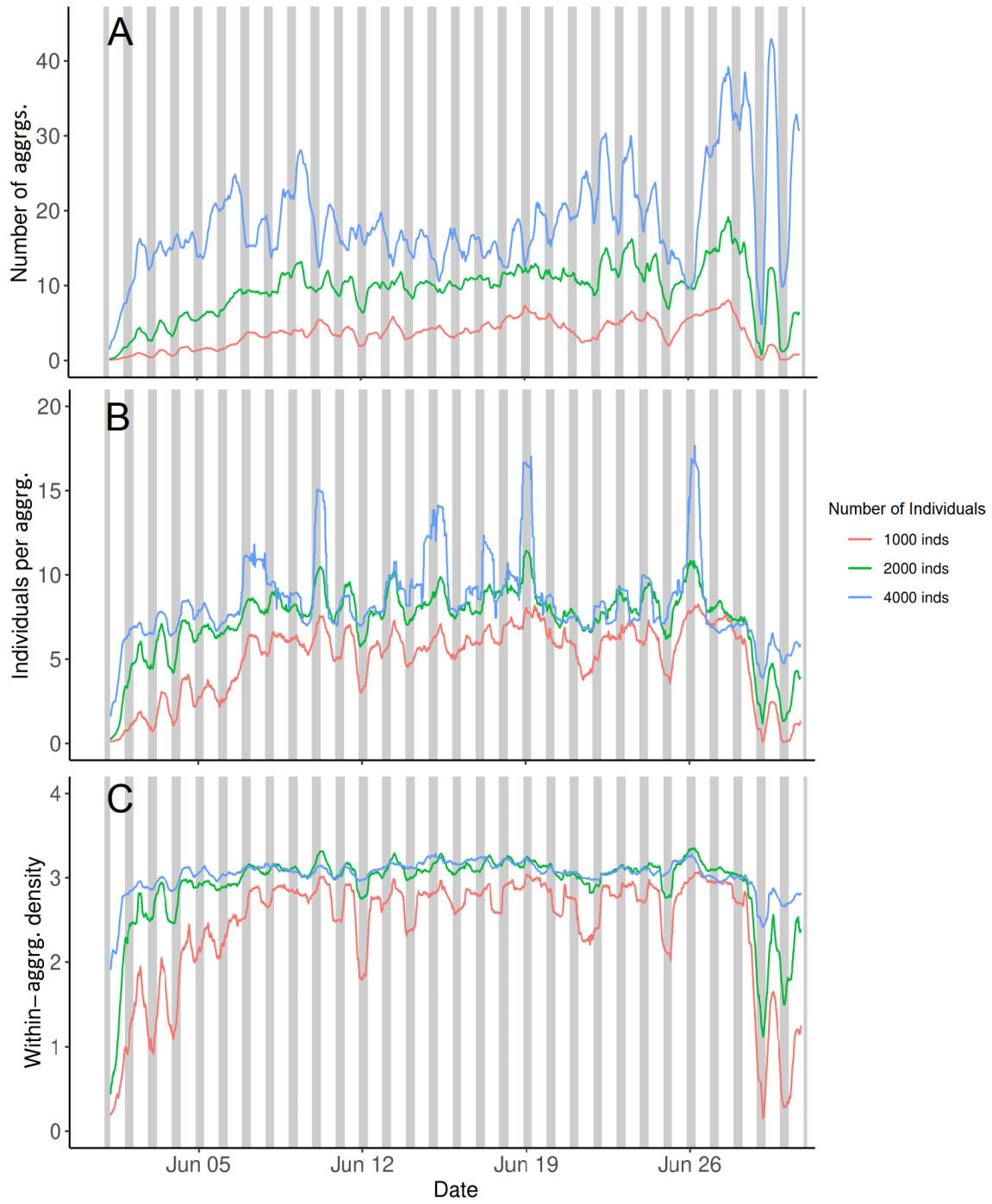


Figure 5.13: Number of aggregations, individuals per aggregation, and within-aggregation density (individuals per meter length) averaged over 12-hour moving windows in the environment-responsive model with 1000, 2000, and 4000 individuals. Light-colored background bars indicate daytime, dark-colored background bars indicate nighttime.

between the density of individuals in the model space and the persistence of schools. This relationship should be tested further with more model runs at different numbers of individuals and different model space sizes.

The number of individuals per school reached a stable state in all three models, but was much higher in the 100-individual model than in the 200- and 400-individual models (Figure 5.14B). The same was true for the within-school density; the number of fish per meter school length reached a stable state in all three models but was much higher in the 100-individual model (Figure 5.14C).

5.4.4.2 Sensory distance

The original environment-responsive model was run with 1000 individuals with a sensory range of 15-30 meters. Variations in the sensory range were tested while holding the number of individuals at 1000 (Table 5.1). The sensory distance was either halved (7.5-15m) or doubled (15-30m) from the original values. The number of aggregations, number of individuals per aggregation, and within-aggregation density did not change in predictable ways when the sensory distance was changed (Figure 5.15). Doubling the sensory distance resulted in very few aggregations, few individuals per aggregation, and low within-aggregation density when HSI was stable. However, when HSI was low or variable, the model with doubled sensory distance produced more aggregations with more individuals and higher within-aggregation densities than the default or halved sensory distances. When HSI was low or variable, the halved-sensory distance model predictably produced fewer aggregations with fewer individuals and lower within-aggregation densities. During periods of stable HSI, the halved-sensory distance model had similar trends

in number of aggregations, individuals per aggregation, and within-aggregation density as the original sensory distance model ([Target5.15](#)).

Table 5.1: Summary statistics for sensitivity tests. Variations of the environment-responsive model were run with either 1000, 2000, or 4000 individuals. Further variations of the environment-responsive model were run with 1000 individuals and either half the original sensory distance or twice the original sensory distance. Variations of the schooling-responsive model were run with either 100, 200, or 400 individuals. Metrics reported for each model run include the number of schools, the number of individuals per school, and the density within the school (individuals per meter length).

Model	# Inds.	Sensory distance	Metric	Minimum	Median	Mean	Maximum
Environment	1000	15-30m	# schools	0.06	4.09	4.17	8.11
			# individuals	0.08	6.07	5.69	8.27
			density	0.15	2.79	2.58	3.07
Environment	1000	7.5-15m	# schools	0.01	2.2	2.18	4.59
			# individuals	0	4.62	4.44	8.88
			density	0.02	2.51	2.27	3.16
Environment	1000	30-60m	# schools	0.02	0.9	1.65	13.87
			# individuals	0	1.24	1.44	5.67
			density	0.05	1.24	1.21	2.77
Environment	2000	15-30m	# schools	0.74	10.51	10.62	19.19
			# individuals	1.18	8	7.79	11.44
			density	1.12	3.09	2.99	3.35
Environment	4000	15-30m	# schools	4.81	17.54	19.49	42.94
			# individuals	3.85	8.45	8.83	17.68
			density	2.42	3.09	3.07	3.29
Schooling	100	15-30m	# schools	2.02	2.08	2.09	2.18
			# individuals	46.01	47.98	48.05	49.8
			density	24.48	26.2	26.12	27.53
Schooling	200	15-30m	# schools	6.5	7.18	7.19	7.82
			# individuals	6.05	6.63	6.68	7.43
			density	6.29	6.62	6.64	7.06
Schooling	400	15-30m	# schools	2.58	3.57	3.58	4.44
			# individuals	2.86	4.1	4.14	5.29
			density	5.01	5.4	5.41	6.3

5.4.5 Comparison to empirical data

5.4.5.1 Environment-responsive model

Menhaden observed in the empirical data were noted to coordinate their swimming movements to form schools. The environment-responsive model did not produce aggregations with this synchronized behavior, and therefore the aggregations could not be considered schools. The environment-responsive model did form some non-school aggregations in areas of relatively high habitat suitability, though most individuals were not affiliated with any aggregation. On average, fewer than 8 aggregations were present in the modeled landscape at any time, and these aggregations typically had fewer than 8 individuals. The frequency of lone individuals is extremely unusual as compared to the empirical data, where singular menhaden-like fish were observed fewer than 5 times across dozens of hours of sampling.

Though the number of aggregations observed in the environment-responsive model outcomes is within the range of the number of schools observed in the empirical data (5-76 schools), the number of individuals per aggregation is much lower. The environment-responsive model had an average of 5.7 individuals per aggregation, while the empirical data consistently showed 40-45 individuals per school. Further, the results of the environment-responsive IBM indicated that the number of individuals per aggregation fluctuates with the number of individuals present in the model landscape, whereas in the empirical data the number of individuals per school was similar despite fluctuations in the number of individuals observed within the study area. Finally, the density of individuals per meter of aggregation length is well below the reported value from the empirical data (average for environment-responsive IBM 2.6 inds m^{-1} , average for empiri-

cal data 15.3 inds m^{-1}), which strengthens the assertion that these aggregations should not be considered schools.

5.4.5.2 Schooling-responsive model

The schooling-responsive model was designed so that polarized aggregation behavior was imposed on the individuals. Therefore, the aggregations that formed did have close alignment between individuals and could be considered schools. Far fewer individuals were modeled in the schooling-responsive models (100 - 400 individuals) than were seen in the empirical data (324 to 11432 individuals). With this critical difference in mind, it is best to compare the number of individuals per school to the empirical data rather than the number of schools formed. The schooling-responsive model produced schools with an average of 48 individuals and a density of 24-28 individuals per meter. These values are comparable to the empirical data. However, increasing the number of individuals in the schooling-responsive IBM did result in a dramatic decrease in the number of individuals per school and the density of those schools ([Table 5.1](#)), which was not consistent with the patterns seen in the empirical data. The sensitivity of model behavior to number of simulated individuals indicates a missing mechanism that serves to structure the organization of schools relative to each other.

5.5 Discussion

The best-available data from decades of published science on menhaden habitat preferences and spatial dynamics, as well as empirical data collected and analyzed in the previous three chapters, were used to inform two separate IBMs. The first IBM was a representation of

individual behavior and movement if it only occurred in response to environmental stimuli. The second IBM represented individual behavior if it only occurred in response to conspecifics. Neither model was able to completely represent the behaviors and spatial patterns observed in the empirical data, though the schooling-responsive model was able to approximate the number of individuals per school and within-school density seen in the empirical data. This lack of similarity was to be expected, as fish behavior and spatial patterns are likely the results of both biotic and abiotic processes and their interactions, not wholly one or the other.

Environmental and behavioral factors are known to operate on different spatio-temporal scales and thus impact fish at different scales ([Beare et al., 2002](#); [Bertrand et al., 2008](#); [Heinänen et al., 2018](#); [Petitgas and Levenez, 1996](#)). In these IBMs, the variation in temporal scale was demonstrated by the difference in time step length between the environment-responsive model and the schooling-responsive model. While a time step of 30 seconds was sufficient for the environment-responsive IBM to create the movement patterns and spatial distributions that were expected, a much smaller time step was needed for the schooling model. The mechanisms for schooling require individuals to respond quickly to each other, or they will not be able to coordinate their swimming ([Parrish et al., 2002](#)). The same quick responses are not needed for fish in the environment-responsive model. Estuaries are known for rapid changes in environmental conditions ([Vinagre et al., 2017](#)), but the conditions between two spatially proximate points should not change dramatically within 30 seconds.

The extremely stable schools created by the 100-individual schooling-responsive model are an illustration of the importance of disturbance mechanisms. Real-world fish schools are highly dynamic and are likely to merge or break apart frequently in response to stimuli ([Pitcher et al., 1996](#); [Soria et al., 1998](#)). Diel patterns in aggregations and spatial patterns have been observed for

clupeids (Skaret, 2003). Diel patterns in menhaden aggregations have been noted in the Patuxent River, with schools forming during the day and breaking apart at night (Lankowicz, unpublished data). However, when only presented with the behavioral mechanism for schooling, the 100-individual schooling-responsive model resulted in schools that persisted for days. Though it is not within the scope of this work, it would be interesting to add mechanisms to adjust the behavior of the simulated individuals in future versions of these models to better represent these diel patterns.

In contrast, the 200- and 400-individual schooling-responsive models did not produce stable schools. The highly stochastic nature of these schools indicated a non-linear relationship between school formation and the density of individuals within the modeled space. This relationship is not representative of the empirical data, in which schools were noted to have a similar number of individuals and within-school density regardless of population size. The empirical data indicated that an emergent property of schooling is the ability to interact as a school with another school. The mechanisms that underly this ability were clearly not captured in the model, as the individuals would frequently abandon their school in an often-unsuccessful attempt to join a more distant school.

The environment-responsive IBM indicated that environmental conditions do play a role in structuring the number and positioning of menhaden aggregations. HSI had a strong, diel cyclical pattern in the modeled landscape. The diel pattern of the modeled space HSI was probably caused by the effects of light conditions and tidal fluctuations on chlorophyll and dissolved oxygen concentration, which were often the limiting factors in a grid cell's HSI value. Diel patterns in menhaden aggregations were not as evident, though this should be expected due to the stochasticity that comes from having non-uniform individuals and no schooling mechanism. When values

of HSI were relatively stable, fluctuations in the number of aggregations, number of individuals per aggregation, and density within the aggregations were still evident. These fluctuations could be the result of the non-uniform ability of the fish to perceive their environments, as well as non-uniform swimming speeds and body lengths. When HSI values decreased dramatically at the end of the modeled period due to declines in the suitability of chlorophyll and dissolved oxygen conditions, few aggregations were formed. This indicates that the effects of poor habitat will outweigh the effects of individual variation.

5.5.1 Sensitivity tests

The sensitivity tests indicated that both the number of individuals and the range of sensory distance greatly affect the movement and spatial patterns of simulated menhaden. Adjusting the number of individuals in the model affected the number of aggregations and the number of individuals per aggregation in both the environment-responsive and the schooling-responsive IBMs. This result is consistent with other empirical studies, which have found that the number of individuals in an area affects the formation and maintenance of schools ([Niwa, 2004](#)). When the number of individuals in the environment-responsive model was increased to 4000, the impact of low HSI on the formation of aggregations was less evident, likely due to an overall increased density of individuals within the modeled space. In both models, more individuals in the model resulted in more aggregations with fewer individuals, rather than increasing the number of individuals within a stable number of aggregations. This result does not match what was found in the empirical data, where an increasing number of individuals within the study area correlated to an increasing number of schools but a relatively stable number of individuals per school. The

IBMs are either missing a mechanism that serves to structure the number of individuals within an aggregation, or the structure of aggregation is an emergent property of combining environmental and behavioral mechanisms. Identifying and including the controlling mechanism or set of controlling mechanisms would improve the realism of the model outcomes.

Sensory distance also had a considerable impact on the structure of aggregations within the IBMs. All other things held equal, halving the range of sensory distance of individuals within the environment-responsive model resulted in a scaled response in the number of aggregations. However, the number of individuals within the aggregations remained roughly the same. This is likely caused by the spatial autocorrelation in HSI between spatially proximate points. When sensory distance is halved the individuals cannot follow the spatially-mediated gradient in HSI because this gradient is only evident at larger spatial scales. Therefore, the likelihood of modeled individuals “finding their way out” of poor habitat is decreased. Individual fish are not fully capable of exploring the modeled space and finding higher-quality areas within which to aggregate. When individuals do find higher-quality habitat, they will structure their movements to remain within it.

When the sensory distance was doubled, the number of aggregations decreased. This was likely caused by a mismatch in perception against the best spatial scale to perceive the spatially-mediated gradient in HSI. Individuals with doubled sensory distances were constantly perceiving a better habitat to move towards, rather than settling in an area of relatively high HSI. Because the modeled landscape was updated frequently (every hour, or 120 time steps), aggregations were unlikely to occur. However, when HSI declined sharply at the end of the model, the fish with increased sensory distance were better able to perceive the fewer and smaller patches of relatively high-quality habitat. Therefore, they formed into many more aggregations than were

seen in the original and half-sensory-distance models.

5.5.2 Comparison to empirical data

Neither IBM was able to completely replicate the patterns and characteristics of menhaden schools as seen in the empirical data. However, the schooling-responsive model had promising results. It produced schools with similar characteristics to the empirical data, though the spatial organization and movement of these schools were not realistic. It is possible that with further refinement, the schooling-responsive model could produce simulated schools with characteristics and spatial patterns comparable to those seen in the field. The results indicate that menhaden spatial patterns, both in how individuals aggregate into schools and how schools are distributed through a landscape, are created by a combination of interactions with both the environment and conspecifics. This corroborates the theory that fish spatial distribution patterns result from the interaction of multiple processes at several scales. Particularly for schooling forage species like menhaden, this means schooling may have a higher impact at small spatial scales and environmental forces may have a higher impact at larger spatial scales (eg [Bertrand et al. \(2008\)](#)).

The modeled number of individuals within a defined area was found to have an impact on the resulting behaviors and spatial patterns. It was notable that changing the number of individuals highly affected the number of individuals per school in both IBMs, as the number of individuals per school stayed relatively consistent in the empirical data despite a large variation in the number of individuals in the study area. This result indicates that both models are missing a mechanism or set of mechanisms that control the number of individuals per school.

5.5.3 Study limitations

There are numerous challenges when implementing IBMs. Though the use of IBMs is increasing in many ecological sub-fields, they are still highly complex and it is difficult to communicate the results. It is important to strike a balance between representing all pertinent processes and over-complicating (or over-parameterizing) the inputs when setting up an IBM ([van der Vaart et al., 2016](#)). Though the intention behind over-parameterization is often a desire to be as detailed and realistic as possible, the outcome will be reduced realism and predictions that are only relevant to the modeled space ([van der Vaart et al., 2016](#)). If model predictions are to be emergent rather than imposed, they should be produced through an accurate mechanical representation of the relationships between biotic and abiotic processes rather than attempting to pattern-match an outcome to empirical data ([Grimm and Railsback, 2005](#)). Though the best available information was used to parameterize the two IBMs developed in this work, questions remain regarding the realism behind the assumed relationships. For example, there is no published information on the possible range of sensory distances of juvenile menhaden. The chosen sensory distances and “schooling rules zones” (i.e. repulsion, attraction, and alignment) were arbitrary. Sensory distance was found to have a significant impact on the formation and size of aggregations. Collecting this information in a laboratory setting and incorporating the results into the IBMs would greatly improve confidence in the ability of the models to represent realistic processes.

5.5.4 Conclusions

Although the IBMs did not completely represent menhaden school characteristics and spatial patterns seen in the empirical data, there were similarities. The schooling-responsive model

was able to produce schools of a similar size in number of individuals and within-school density as was typically seen in the representative estuarine tributary. With refinement, the schooling-responsive model could produce menhaden schools that behave in the same way as schools seen in the empirical data. The sensitivity of the schooling-responsive model to the size of the simulated population was not consistent with the relationship between schooling and population size in the empirical data, indicating a missing mechanism or mechanisms that structure the spatial pattern of schools. The mechanisms represented in the schooling-responsive model are at fine spatial scales, and the missing mechanism or mechanisms that would create stable schools regardless of population size are likely at a coarser spatial scale. Future research needs to be conducted to quantify the contribution of fine- and coarse-scale processes to menhaden school characteristics and spatial structure. This research, built from the understanding of menhaden spatial ecology as it exists in published literature and from the analysis of empirical data in the previous three chapters, is a substantial first step in addressing this ecological challenge.

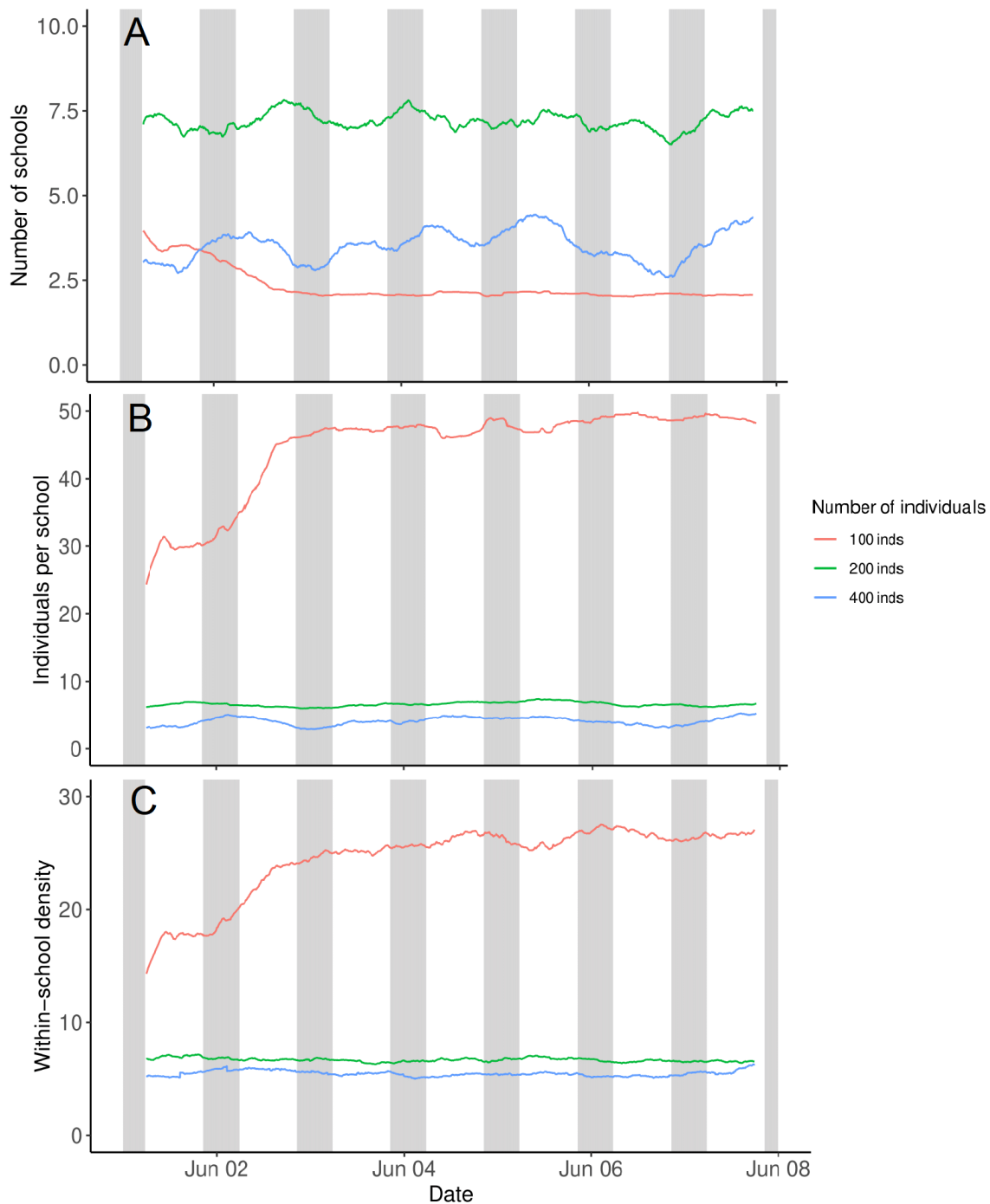


Figure 5.14: Number of schools, individuals per school, and within-school density averaged over 12-hour moving windows in the schooling-responsive model with 100, 200, and 400 individuals. Though environmental conditions were not considered, background colors for day and night periods were still included so it is easier to visualize the changing of the metrics over the modeled period. Light-colored background bars indicate daytime, dark-colored background bars indicate nighttime.

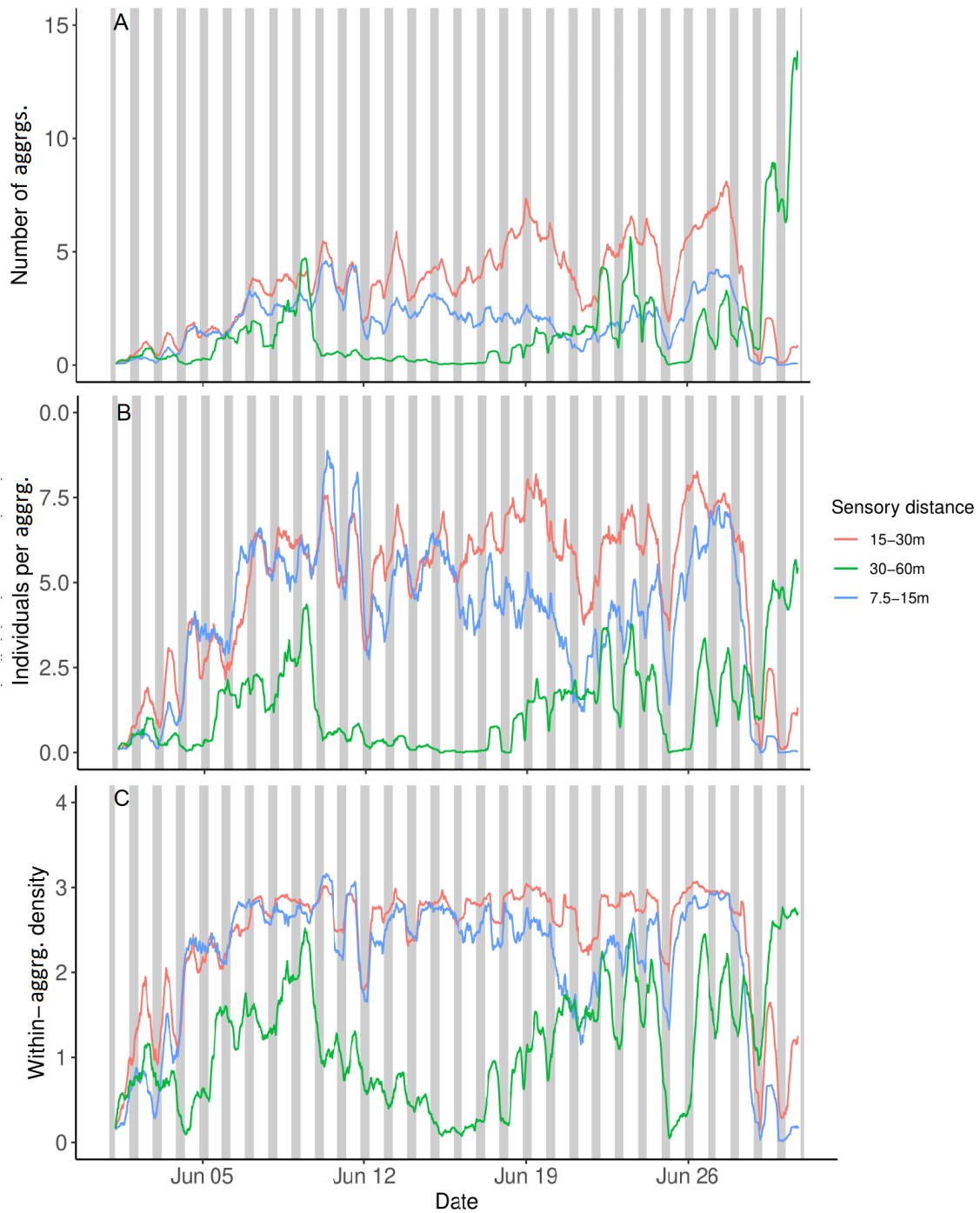


Figure 5.15: Number of aggregations, individuals per aggregation, and within-aggregation density (individuals per meter length) averaged over 12-hour moving windows in the environment-responsive model with sensory set to 15-30m (default), 7.5-15m, or 30-60m. Light-colored background bars indicate daytime, dark-colored background bars indicate nighttime.

Chapter 6: Summary and conclusions

The varied effects of space on organism distribution, movement, and survival have sparked ecological questions for hundreds of years. The questions addressed in this dissertation were in that same vein, and so fall under the wide umbrella of spatial ecology. Atlantic menhaden and other forage fish are good candidates for spatial ecology studies, as their spatial patterns are highly dynamic and linked to many biotic and abiotic processes that operate at multiple spatial scales ([Bertrand et al., 2008](#); [Gutiérrez et al., 2012](#); [Liang et al., 2020](#)). This makes sampling at multiple spatial scales and across many habitat types essential to accurately determine and characterize the drivers of species abundance and distribution, but information on most fish species' spatial dynamics has previously focused only on large spatial scales ([Ciannelli et al., 2008](#)). With the use of sonar imaging and through a series of complementary studies in a representative tidal estuarine tributary, I have attempted to resolve fine-scale patterns in juvenile menhaden habitat use; quantify the effects of environmental factors like temperature, salinity, dissolved oxygen, tidal stage, and chlorophyll concentration on menhaden distribution; and examine the role of individual behaviors in creating emergent schooling patterns. The combination of fine-scale empirical data, statistical modeling, and individual-based models used in these studies contributes novel information to our understanding of juvenile menhaden spatial dynamics.

6.1 Novel technology applications

The adaptation of sonar imaging technology for mobile, transect-based surveys to observe the abundance and spatial distributions of menhaden was a significant achievement of this research (Chapter 2). The combined evidence of visual identification of menhaden at the water surface during data collection, previous stationary ARIS studies of menhaden body shape ([Shahrestani et al., 2017](#)), previous stationary ARIS studies of menhaden schooling behavior ([Zhang et al., 2017](#)), and seasonal size information for juvenile menhaden indicate that vast majority of fish identified in this study are juvenile menhaden. Recent shiftings from single-species management emphasis to ecosystem-based management have highlighted the need to collect detailed data on the interactions between fish species and their ecosystems and the variability of these interactions ([Anstead et al., 2021](#); [Pikitch et al., 2004](#)). Sonar imaging cameras have proven to be a useful tool in collecting abundance and distribution data on fishes at extremely fine scales, in areas that are not sampled well by traditional net-based methods, and with minimal disturbance to the fish. Unlike most active acoustic survey methods, mobile ARIS allows for the identification and enumeration of individuals. Mobile ARIS is also a relatively efficient data collection method—it is not labor-intensive for the researchers to use and continuously collects large volumes of data.

There are still challenges remaining for the optimization of sonar imaging techniques for forage fish; the avoidance response to either vessel noise or the sound signals of the sonar camera itself is unknown. In my use of mobile ARIS surveys across multiple Chesapeake Bay tributaries, I have noted variable responses to the approaching vessel and sonar imaging camera. Avoidance behaviors could be individually-variable, or they could be linked to species and habitat type.

Further studies would be needed to clarify the response.

The massive quantities of data produced by sonar imaging also pose a challenge. Other established active acoustic survey methods have image processing software and algorithms that speed up the analysis of raw data. The methods for automated or semi-automated processing of sonar imaging data are typically limited to stationary applications ([Shahrestani et al., 2017](#)). These methods are not suitable for mobile sonar imaging data, as the images produced are highly dynamic and therefore difficult for machine learning algorithms to process. I, and others who use mobile sonar imaging, must process the data manually in the proprietary software created by the manufacturer of the sonar imaging camera ([Becker et al., 2017](#); [Kerschbaumer et al., 2020](#)). This is time-consuming and lends itself to reader bias and uncertainty. Machine learning algorithms to semi-automate the processing of mobile sonar imaging data are under development but are not yet ready for applications in fisheries research.

6.2 Habitat preferences

With the use of GAMs, I found that menhaden density varies considerably throughout the summer within the lower Patuxent River (Chapter 2). Habitat type, water depth, and month were all statistically significant predictors of the distribution and density of menhaden. The creek habitats were noted to have significantly higher densities of menhaden than the Patuxent River channel habitat. This result was partially mediated by water depth—menhaden were also present at significantly higher densities in shallower habitats, and the creeks are all much shallower (average sampling depth approximately 4 m) than the river channel (average sampling depth approximately 8 m).

6.3 Characteristics of menhaden schools

Chapters 3 and 4 used the same empirical dataset as Chapter 2, but the organization of menhaden into discrete schools was taken into account. Chapter 2 focused on individuals, but Chapter 3 emphasized aggregative extent to examine the density of schools across the study area. I developed and implemented a procedure to identify discrete schools of menhaden observed in the sonar imaging data. In Chapter 3, I found that the menhaden identified in the lower Patuxent River were typically organized into many smaller schools (tens to hundreds of individuals) as compared to a few large schools (thousands of individuals). This is a sharp contrast to the schools of millions of individuals found in the coastal oceans and harvested with purse seines.

A GAM indicated that the density within a school (number of individuals per meter length) was affected by water depth, distance to the nearest neighboring school, a random effect based on the number of individuals per school, and interactions between tidal stage, habitat, and sampling month. A separate GAM found that the distance between nearest neighboring schools was affected by water depth, within-school density, spatial location (x- and y-coordinates), and interactions between tidal stage, habitat, and sampling month. It is expected that ontogeny and size would play a role in structuring menhaden aggregations as well, and though these were not explicitly included as explanatory variables in GAMs, they could be implicitly included under the sampling month variable. The GAMs indicate that both environmental and behavioral mechanisms structure the size and spatial pattern of menhaden schools.

6.4 Spatial pattern of menhaden schools

In Chapter 3, a daily clustering coefficient was used to explore the spatial pattern of menhaden schools with respect to each other. The results indicated that these schools were clustered at the spatial scale of the sampled transects. In Chapter 4, spatial point pattern analysis adapted for use along one-dimensional linear transects was used to explore the spatial patterns of menhaden schools at a higher spatial resolution. The spatial locations of schools within a transect were treated as a point pattern, and these point patterns were analyzed for spatial variation linked to environmental spatial covariates (first-order structure) and spatial variation linked to the location of other points in the pattern (second-order structure). The results were mixed, and very few obvious trends were noted. A negative association between the river channel habitat type and the presence of first-order structure was noted. Point patterns with no first-order structure or no second-order structure had fewer schools than point patterns that did have first- or second-order structure. There is not a recognized number of points or size of the observation window needed to draw robust conclusions from spatial point pattern analysis, and some studies on rare plants use as few as 4 points ([Ben-Said, 2021](#); [Cordero et al., 2016](#); [Wehenkel et al., 2015](#)). However, it is known that it may be difficult to accurately identify point patterns and underlying point processes with fewer than 50 points. The river channel transects often did not have more than 20 identified schools of menhaden, so the negative association between the river channel and first-order structure may be due to the low sample size. There is still value in these point pattern analyses as exploratory studies in areas with limited availability of high-resolution data suitable for point pattern analysis ([Ben-Said, 2021](#); [Ben-Said et al., 2020](#)).

6.5 Impact of environmental and behavioral mechanisms on spatial patterns

Point pattern analysis highlighted the variation in the relative contribution of environmental and behavioral drivers of menhaden spatial pattern across multiple scales. First-order structure is likely caused by environmental forces which create a heterogeneous environment of suitable and unsuitable patches. First-order structure of menhaden schools was driven by spatial covariates like y-coordinate (representative of distance along a transect) and depth. The gradients of these spatial covariates are best seen at spatial scales of tens of meters to hundreds of meters. Second-order structure indicates that the spatial location of organisms or aggregations of organisms is dependent on the location of other organisms in the studied area, which is representative of schooling behavior in menhaden. Second-order structure was most consistent at small spatial scales, between meters and tens of meters. It is likely that the true spatial point pattern of menhaden schools within the lower Patuxent River at any time is shaped by both environmental and behavioral mechanisms, as well as the stochasticity of variation in individual behaviors.

Finally, in Chapter 5, two dynamic and spatially explicit individual-based models of menhaden schooling and movement were built based on findings from all analyses of the empirical data. IBMs are valuable tools to explore how interactions between individuals, processes, and the environment create emergent behaviors. They are well-suited to track the movement of individuals through dynamic and spatially explicit environments, which also makes them highly capable of quantifying aggregative behaviors. The first IBM was a representation of individual behavior and movement if it only occurred in response to environmental stimuli. The second IBM represented individual behavior if it only occurred in response to conspecifics. Neither model was able to accurately represent the behaviors and spatial patterns observed in the empirical data.

The environment-responsive model successfully replicated the movement of individuals within a heterogeneous environment consisting of suitable and unsuitable areas, but did not create the schools of individuals seen in the empirical data. The schooling-responsive model did create realistic schools of an appropriate size, but the individuals were not as close together within schools as they were in the empirical data. Further, the schooling-responsive model created schools that persisted for an unnaturally long time. The lack of realism in the model outputs, as compared to the empirical data, was expected given our understanding of the spatially-variable impact of both environmental and behavioral forces on menhaden schooling.

6.6 Ecological implications

The results of empirical data analysis indicated that the creek habitat was important for menhaden. At all scales of aggregation (individual, school, and clusters of schools), significant differences in density and spatial structure of menhaden were observed between the creek and river channel habitats. More schools of menhaden were identified in the creeks than in the river channel, though the number of individuals per school in the two habitats was not significantly different. Environmental factors like phytoplankton prey availability, dissolved oxygen concentration, and tides and water flow are expected to contribute to the differences in menhaden populations between the two habitats. Predation of menhaden was unquantified in this research, but variability in predator distribution could also be a factor in structuring menhaden distribution.

There was also variation in the density of menhaden among the three creeks that were sampled. The creeks have differences in features like surrounding land use, freshwater input, and depth, all of which would affect the physical circulation of water within the creeks and therefore

the distribution of phytoplankton. Atlantic menhaden is known to form spatial patterns based on the distribution of phytoplankton prey ([Friedland et al., 1989, 1996](#)), these features could be a good indication of forage fish presence. More research is required needed to determine the strength of this possible relationship.

If global climate change or other environmental challenges alter the suitability of shallow creeks for menhaden or their phytoplankton prey items, it would likely impact the health of the overall estuarine population of menhaden. Negative impacts to the menhaden population could ripple out and negatively affect the populations of a wide array of predatory species. The clear recognition of the critical role of Atlantic menhaden in the ecosystem has spurred scientific and management efforts to develop ecosystem approaches to its management ([Anstead et al., 2021](#); [SEDAR, 2020](#)).

The individual-based modeling efforts outlined in Chapter 5 represent a first step in addressing an enduring ecological objective, which is quantifying the relative impact of various environmental and behavioral forces on the abundances and distributions of organisms at multiple spatial scales. The modeling results, though not useful on their own to predict menhaden abundances or distributions, are an indication that schooling is more important at small spatial and temporal scales, and environmental conditions are more important at larger spatial and temporal scales. The spatial patterns that occur at any location and time are likely the emergent result of the interactions between the two types of forces and the stochasticity of individual variation in the perception of and response to local stimuli.

The parts of this dissertation, even combined, are only a small piece of menhaden spatial ecology. However, our understanding of any species' spatial ecology is at its best and most comprehensive when it is seen as the emergent property of all of its smaller components. Looking

forward, it would be interesting to expand the spatial extent of this work and compare the trends in menhaden spatial distribution and abundance in other environments. The menhaden observed in the lower Patuxent River eventually emigrate to the main stem of the Chesapeake Bay, and then to the coastal Atlantic Ocean. There are clear differences in the environmental conditions and intra- and inter-species interactions that fish will face in these three areas of increasing spatial extent. Learning how those differences create spatial patterns and affect abundance at all applicable spatial scales could help managers make robust decisions, and could further improve our ability to predict shifts in menhaden abundance and distribution with a changing climate.

Bibliography

- Able, K. W., Grothues, T. M., and Kemp, I. M. (2013). Fine-scale distribution of pelagic fishes relative to a large urban pier. *Marine Ecology Progress Series*, 476:185–198.
- Able, K. W., Grothues, T. M., Rackovan, J. L., and Buderman, F. E. (2014). Application of mobile dual-frequency identification sonar (DIDSON) to fish in estuarine habitats. *Northeastern Naturalist*, 21(2):192–209.
- Adamack, A. T., Rose, K. A., Breitburg, D. L., Nice, A. J., and Lung, W. S. (2012). Simulating the effect of hypoxia on bay anchovy egg and larval mortality using coupled watershed, water quality, and individual-based predation models. *Marine Ecology Progress Series*, 445:141–160.
- Ahrenholz, D. W. (1991). Population biology and life history of the North American Menhadens, *Brevoortia* spp. *Marine Fisheries Review*, 53(4):3–19.
- Ahrenholz, D. W., Guthrie, J. F., Krouse, C. W., Lenarz, W. H., and Gunderson, D. R. (1989). Results of Abundance Surveys of Juvenile Atlantic and Gulf Menhaden, *brevoortia tyrannus* and *B. patronus*. Technical report, NOAA NMFS.
- Alder, J., Campbell, B., Karpouzi, V., Kaschner, K., and Pauly, D. (2008). Forage fish: From ecosystems to markets. *Annual Review of Environment and Resources*, 33(1):153–166.
- Allen, D. M., Service, S. K., and Ogburn-Matthews, M. V. (1992). Factors influencing the collection efficiency of estuarine fishes. *Transactions of the American Fisheries Society*, 121:234–244.
- Anderson, D. R. (2003). Response to Engeman: Index values rarely constitute reliable information. *Wildlife Society Bulletin (1973-2006)*, 31(1):288–291.
- Anderson, J. J. (1981). A stochastic model for the size of fish schools. *Fishery Bulletin*, 79(2):315–323.
- Ang, Q. W., Baddeley, A., and Nair, G. (2012). Geometrically corrected second order analysis of events on a linear network, with applications to ecology and criminology. *Scandinavian Journal of Statistics*, 39(4):591–617.
- Annis, E. R., Houde, E. D., Harding, L. W., Mallonee, M. E., and Wilberg, M. J. (2011). Calibration of a bioenergetics model linking primary production to Atlantic menhaden *Brevoortia Tyrannus* growth in Chesapeake Bay. *Marine Ecology Progress Series*, 437:253–267.

- Anstead, K. A., Drew, K., Chagaris, D., Schueller, A. M., McNamee, J. E., Buchheister, A., Nesslage, G., Uphoff Jr., J. H., Wilberg, M. J., Sharov, A., Dean, M. J., Brust, J., Celestino, M., Madsen, S., Murray, S., Appelman, M., Ballenger, J. C., Brito, J., Cosby, E., Craig, C., Flora, C., Gottschall, K., Latour, R. J., Leonard, E., Mroch, R., Newhard, J., Orner, D., Swanson, C., Tinsman, J., Houde, E. D., Miller, T. J., and Townsend, H. (2021). The Path to an Ecosystem Approach for Forage Fish Management: A Case Study of Atlantic Menhaden. *Frontiers in Marine Science*, 8:607657.
- Anstead, K. A., Schaffler, J. J., and Jones, C. M. (2017). Contribution of nursery areas to the adult population of Atlantic Menhaden. *Transactions of the American Fisheries Society*, 146:36–46.
- Aoki, I. (1982). A simulation study on the schooling mechanism in fish. *Bulletin of the Japanese Society of Scientific Fisheries*, 48(8):1081–1088.
- ASMFC (2017). Atlantic Menhaden stock assessment update. Technical report, ASMFC, Arlington, VA.
- ASMFC (2020). Atlantic Menhaden assessments overview. Technical report, ASMFC, Arlington, VA.
- Austin, H. M. (2002). Decadal oscillations and regime shifts, a characterization of the Chesapeake Bay marine climate. In McGinn, N. A., editor, *Fisheries in a Changing Climate. American Fisheries Society Symposium 32*, pages 155–170. American Fisheries Society, Bethesda, MD.
- Babcock, E. A., Pikitch, E. K., McAllister, M. K., Apostolaki, P., and Santora, C. (2005). A perspective on the use of spatialized indicators of ecosystem-based fishery management through spatial zoning. *ICES Journal of Marine Science*, 62:469–476.
- Baddeley, A., Rubak, E., and Turner, R. (2015). *Spatial Point Patterns: Methodology and Applications with R*. Chapman and Hall/ CRC, Boca Raton, FL.
- Bakun, A., Babcock, E. A., Lluch-Cota, S. E., Santora, C., and Salvadeo, C. J. (2010). Issues of ecosystem-based management of forage fisheries in "open" non-stationary ecosystems: The example of the sardine fishery in the Gulf of California. *Reviews in Fish Biology and Fisheries*, 20(1):9–29.
- Barnett, L. A. K., Ward, E. J., Jannot, J. E., and Shelton, A. O. (2019). Dynamic spatial heterogeneity reveals interdependence of marine faunal density and fishery removals. *Ecological Indicators*, 107:105585.
- Bartolino, V., Tian, H., Bergström, U., Jounela, P., Aro, E., Dieterich, C., Markus Meier, H. E., Cardinale, M., Bland, B., and Casini, M. (2017). Spatio-temporal dynamics of a fish predator: Density-dependent and hydrographic effects on baltic sea cod population. *PLoS ONE*, 12(2).
- Bayley, P. B. and Herendeen, R. A. (2000). The efficiency of a seine net. *Transactions of the American Fisheries Society*, 129:901–923.

- Beare, D. J., Reid, D. G., and Petitgas, P. (2002). Spatio-temporal patterns in herring (*Clupea Harengus* L.) school abundance and size in the northwest North Sea: Modelling space-time dependencies to allow examination of the impact of local school abundance on school size. *ICES Journal of Marine Science*, 59(3):469–479.
- Becker, A., Whitfield, A. K., Cowley, P. D., and Cole, V. J. (2017). Does water depth influence size composition of estuary-associated fish? Distributions revealed using mobile acoustic-camera transects along the channel of a small shallow estuary. *Marine and Freshwater Research*, 68:2163–2169.
- Ben-Said, M. (2021). Spatial point-pattern analysis as a powerful tool in identifying pattern-process relationships in plant ecology: An updated review. *Ecological Processes*, 10:56.
- Ben-Said, M., Ghallab, A., Lamrhari, H., Carrier, J., Linares, J. C., and Taiqui, L. (2020). Characterizing spatial structure of *Abies Marocana* forest through point pattern analysis. *Forest Systems*, 29(1):e104.
- Benoit-Bird, K. J., Battaile, B. C., Heppell, S. A., Hoover, B., Irons, D., Jones, N., Kuletz, K. J., Nordstrom, C. A., Paredes, R., Suryan, R. M., Waluk, C. M., and Trites, A. W. (2013). Prey patch patterns predict habitat use by top marine predators with diverse foraging strategies. *PLoS ONE*, 8(1):e53348.
- Benoit-Bird, K. J. and Waluk, C. M. (2021). Remote acoustic detection and characterization of fish schooling behavior. *The Journal of the Acoustical Society of America*, 150:4329.
- Bertrand, A., Gerlotto, F., Bertrand, S., Gutiérrez, M., Alza, L., Chipollini, A., Díaz, E., Espinoza, P., Ledesma, J., Quesquén, R., Peraltilla, S., and Chavez, F. P. (2008). Schooling behaviour and environmental forcing in relation to anchoveta distribution: An analysis across multiple spatial scales. *Progress in Oceanography*, 79:264–277.
- Beverton, R. J. H. and Holt, S. J. (1957). *On the Dynamics of Exploited Fish Populations*. Springer, Dordrecht.
- Bi, H., Guo, Z., Benfield, M. C., Fan, C., Ford, M., Shahrestani, S., and Sieracki, J. M. (2015). A semi-automated image analysis procedure for *in situ* plankton imaging systems. *PLoS ONE*, 10(5):1–17.
- Blaber, S. J. M. and Blaber, T. G. (1980). Factors affecting the distribution of juvenile estuarine and inshore fish. *Journal of Fish Biology*, 17(2):143–162.
- Blackwell, B. G. and Brown, M. I. (2005). Comparison of day and night shoreline seine catches in two South Dakota glacial lakes. *Journal of Freshwater Ecology*, 20(1):79–83.
- Bolster, J. (2012). *The Mortal Sea: Fishing the Atlantic in the Age of Sail*. Harvard University Press, Cambridge, MA.
- Borner, K. K., Krause, S., Mehner, T., Uusi-Heikkilä, S., Ramnarine, I. W., and Krause, J. (2015). Turbidity affects social dynamics in Trinidadian guppies. *Behavioral Ecology and Sociobiology*, 69(4):645–651.

- Boswell, K. M., Wilson, M. P., and Cowan, J. H. J. (2008). A semiautomated approach to estimating fish size, abundance, and behavior from dual-frequency identification sonar (DIDSON) data. *North American Journal of Fisheries Management*, 28:799–807.
- Boswell, K. M., Wilson, M. P., MacRae, P. S. D., Wilson, C. A., and Cowan, J. H. J. (2010). Seasonal estimates of fish biomass and length distributions using acoustics and traditional nets to identify estuarine habitat preferences in Barataria Bay, Louisiana. *Marine and Coastal Fisheries: Dynamics, Management, and Ecosystem Science*, 2:83–97.
- Boswell, K. M., Wilson, M. P., and Wilson, C. A. (2007). Hydroacoustics as a tool for assessing fish biomass and size distribution associated with discrete shallow water estuarine habitats in Louisiana. *Estuaries and Coasts*, 30(4):1–11.
- Boyd, C., Woillez, M., Bertrand, S., Castillo, R., Bertrand, A., and Punt, A. E. (2015). Bayesian posterior prediction of the patchy spatial distributions of small pelagic fish in regions of suitable habitat. *Canadian Journal of Fisheries and Aquatic Sciences*, 72(2):290–303.
- Boyd, R., Roy, S., Sibly, R., Thorpe, R., and Hyder, K. (2018). A general approach to incorporating spatial and temporal variation in individualbased models of fish populations with application to Atlantic mackerel. *Ecological Modelling*, 382:9–17.
- Brady, D. C. and Targett, T. E. (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.
- Brochier, T., Auger, P.-A., Pecquerie, L., Machu, E., Capet, X., Thiaw, M., Mbaye, B. C., Braham, C.-B., Ettahiri, O., Charouki, N., Sène, O. N., Werner, F. E., and Brehmer, P. (2018). Complex small pelagic fish population patterns arising from individual behavioral responses to their environment. *Progress in Oceanography*, 164:12–27.
- Brochier, T., Lett, C., Tam, J., Fréon, P., Colas, F., and Ayón, P. (2008). An individual-based model study of anchovy early life history in the northern Humboldt Current system. *Progress in Oceanography*, 79(2-4):313–325.
- Buchheister, A., Miller, T. J., and Houde, E. D. (2017). Evaluating ecosystem-based reference points for Atlantic menhaden. *Marine and Coastal Fisheries: Dynamics, Management, and Ecosystem Science*, 9:457–478.
- Buchheister, A., Miller, T. J., Houde, E. D., Secor, D. H., and Latour, R. J. (2016). Spatial and temporal dynamics of Atlantic menhaden (*Brevoortia Tyrannus*) recruitment in the Northwest Atlantic Ocean. *ICES Journal of Marine Science*, 73(4):1147–1159.
- Caeiro, S., Painho, M., Goovaerts, P., Costa, H., and Sousa, S. (2003). Spatial sampling design for sediment quality assessment in estuaries. *Environmental Modelling & Software*, 18:853–859.
- Calovi, D. S., Lopez, U., Schuhmacher, P., Chaté, H., Sire, C., and Theraulaz, G. (2015). Collective response to perturbations in a data-driven fish school model. *Journal of the Royal Society Interface*, 12(104).

- Catalán, I. A., Macías, D., Solé, J., Ospina-Álvarez, A., and Ruiz, J. (2013). Stay off the motorway: Resolving the pre-recruitment life history dynamics of the European anchovy in the SW Mediterranean through a spatially-explicit individual-based model (SEIBM). *Progress in Oceanography*, 111:140–153.
- Chave, J. (2013). The problem of pattern and scale in ecology: What have we learned in 20 years? *Ecology Letters*, 16:4–16.
- Christmas, J., McBee, J., Waller, R., and Sutter III, F. (1982). Habitat suitability index models: Gulf menhaden. Technical Report FWS/OBS-82/10.23, U.S. Department of the Interior, Fish and Wildlife Service.
- Ciannelli, L., Fauchald, P., Chan, K. S., Agostini, V. N., and Dingsør, G. E. (2008). Spatial fisheries ecology: Recent progress and future prospects. *Journal of Marine Systems*, 71:223–236.
- Clark, T. P. (2020). Mining the Sea: A within-case comparative analysis of the Atlantic menhaden fishery in the age of capital. *Sociology of Development*, 6(2):222–249.
- Cordero, I., Jimenez, M., Delgado, J., Villegas, L., and Balaguer, L. (2016). Spatial and demographic structure of tara stands (*Caesalpinia Spinosa*) in Peru: Influence of present and past forest management. *Forest Ecology and Management*, 377:71–82.
- Couperus, B., Gastauer, S., Fässler, S. M., Tulp, I., van der Veer, H. W., and Poos, J. J. (2016). Abundance and tidal behaviour of pelagic fish in the gateway to the Wadden Sea. *Journal of Sea Research*, 109:42–51.
- Cowan, J. (1999). Modeling effects of increased larval mortality on bay anchovy population dynamics in the mesohaline Chesapeake Bay: Evidence for compensatory reserve. *Marine Ecology Progress Series*, 185:133–146.
- Curtis, K. A. (2004). Fine scale spatial pattern of Pacific sardine (*Sardinops Sagax*) and northern anchovy (*Engraulis Mordax*) eggs. *Fisheries Oceanography*, 13(4):239–354.
- Cury, P. M., Boyd, I. L., Bonhommeau, S., Anker-Nilssen, T., Crawford, R. J. M., Furness, R. W., Mills, J. A., Murphy, E. J., Osterblom, H., Paleczny, M., Piatt, J. F., Roux, J.-P., Shannon, L. J., and Sydeman, W. J. (2011). Global seabird response to forage fish depletion - One-third for the birds. *Science*, 334:1703–1706.
- Cushing, D. H. (1990). *Plankton Production and Year-Class Strength in Fish Populations: An Update of the Match/Mismatch Hypothesis*, volume 26. Academic Press.
- Dale, M. R. T., Dixon, P., Fortin, M.-J., Legendre, P., Myers, D. E., and Rosenberg, M. S. (2002). Conceptual and mathematical relationships among methods for spatial analysis. *Ecography*, 25(5):558–577.
- Dalyander, P. S. and Cerco, C. F. (2010). Integration of a fish bioenergetics model into a spatially explicit water quality model: Application to menhaden in Chesapeake Bay. *Ecological Modelling*, 221(16):1922–1933.

- DeAngelis, D. L. and Grimm, V. (2014). Individual-based models in ecology after four decades. *F1000 Prime Reports*, 6(39).
- DeAngelis, D. L. and Gross, L. J., editors (1992). *Individual-Based Models and Approaches in Ecology: Populations, Communities and Ecosystems*. Chapman and Hall, New York, NY.
- DeAngelis, D. L. and Mooij, W. M. (2005). Individual-Based Modeling of Ecological and Evolutionary Processes. *Annual Review of Ecological and Evolutionary Systems*, 36:147–68.
- Demšar, U., Buchin, K., Cagnacci, F., Safi, K., Speckmann, B., Van de Weghe, N., Weiskopf, D., and Weibel, R. (2015). Analysis and visualisation of movement: An interdisciplinary review. *Movement Ecology*, 3(1):5.
- Ding, H. and Elmore, A. J. (2015). Spatio-temporal patterns in water surface temperature from Landsat time series data in the Chesapeake Bay, U.S.A. *Remote Sensing of Environment*, 168:335–348.
- Domenici, P., Steffensen, J. F., and Marras, S. (2017). The effect of hypoxia on fish schooling. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1727).
- Dray, S., Péliissier, R., Couteron, P., Fortin, M.-J., Legendre, P., Peres-Neto, P. R., Bellier, E., Bivand, R., Blanchet, F. G., De Caceres, M., Dufour, A. B., Heegaard, E., Jombart, T., Munoz, F., Oksanen, J., Thioulouse, J., and Wagner, H. H. (2012). Community ecology in the age of multivariate multiscale spatial analysis. *Ecological Monographs*, 82(3):257–275.
- Durbin, A. G. and Durbin, E. G. (1975). Grazing rates of the Atlantic menhaden *Brevoortia Tyrannus* as a function of particle size and concentration. *Marine Biology*, 33(3):265–277.
- Durbin, A. G., Durbin, E. G., Verity, P. G., and Smayda, T. J. (1981). Voluntary swimming speeds and respiration rates of a filter-feeding planktivore, the Atlantic menhaden, *Brevoortia Tyrannus* (Pisces: Clupeidae). *Fishery Bulletin*, 78(4):877–886.
- Durell, E. Q. and Weedon, C. (2017). Striped Bass Seine Survey Juvenile Index Web Page. dnr.maryland.gov/fisheries/Pages/juvenile-index.aspx.
- Elliott, M. and McLusky, D. S. (2002). The need for definitions in understanding estuaries. *Estuarine, Coastal and Shelf Science*, 55:815–827.
- Finley, A. O., Banerjee, S., and Cook, B. D. (2014). Bayesian hierarchical models for spatially misaligned data in R. *Methods in Ecology and Evolution*, 5(6):514–523.
- Fletcher, R. and Fortin, M.-J. (2018). *Spatial Ecology and Conservation Modeling: Applications with R*. Springer International Publishing.
- Fortin, M.-J., Dale, M. R. T., and ver Hoef, J. (2002). Spatial Analysis in Ecology. In El-Shaarawi, A. H. and Piegorsch, W. W., editors, *Encyclopedia of Environmetrics*, volume 4, pages 2051–2058. John Wiley & Sons, Chichester, UK.

- Foster, E. G., Ritz, D. A., Osborn, J. E., and Swadling, K. M. (2001). Schooling affects the feeding success of Australian salmon *Arripis Trutta* when preying on mysid swarms *Parame-sopodopsis Rufa*. *Journal of Experimental Marine Biology and Ecology*, page 14.
- Franklin, H. B. (2007). *The Most Important Fish in the Sea: Menhaden and America*. Island Press.
- Fréon, P., Cury, P. M., Shannon, L. J., and Roy, C. (2005). Sustainable exploitation of small pelagic fish stocks challenged by environmental and ecosystem changes: A review. *Bulletin of Marine Science*, 76(2):385–462.
- Fréon, P. and Misund, O. A. (1998). *Dynamics of Pelagic Fish Distribution and Behaviour: Effects on Fisheries and Stock Assessment*. "Fishing News" Books. Wiley-Blackwell.
- Friedland, K. D., Ahrenholz, D. W., and Guthrie, J. F. (1989). Influence of plankton on distribution patterns of the filter-feeder *Brevoortia Tyrannus* (Pisces: Clupeidae). *Marine Ecology Progress Series*, 54(1):1–11.
- Friedland, K. D., Ahrenholz, D. W., and Guthrie, J. F. (1996). Formation and seasonal evolution of Atlantic menhaden juvenile nurseries in coastal estuaries. *Estuaries*, 19(1):105–114.
- Friedland, K. D., Ahrenholz, D. W., and Haas, L. W. (2005). Viable gut passage of cyanobacteria through the filter-feeding fish Atlantic menhaden, *Brevoortia Tyrannus*. *Journal of Plankton Research*, 27(7):715–718.
- Friedland, K. D. and Haas, L. W. (1988). Emigration of juvenile Atlantic menhaden, *Brevoortia Tyrannus* (Pisces: Clupeidae), from the York River estuary. *Estuaries*, 11(1):45–50.
- Friedland, K. D., Haas, L. W., and Merriner, J. V. (1984). Filtering rates of the juvenile Atlantic menhaden *Brevoortia Tyrannus* (Pisces: Clupeidae), with consideration of the effects of detritus and swimming speed. *Marine Biology*, 84(2):109–117.
- Friedland, K. D., Lynch, P. D., and Gobler, C. J. (2011). Time series mesoscale response of Atlantic menhaden *Brevoortia Tyrannus* to variation in plankton abundances. *Journal of Coastal Research*, 27(6):1148–1158.
- Frye, J. (1978). *The Men All Singing: The Story of Menhaden Fishing*. Donning, Norfolk, VA.
- Gallagher, B. K., Fabrizio, M. C., and Tuckey, T. D. (2018). Estimation of juvenile Striped Bass relative abundance in the Virginia portion of Chesapeake Bay. 2017 annual report to the Virginia Marine Resources Commission. Technical report, Virginia Institute of Marine Science, College of William and Mary.
- Gautrais, J., Jost, C., and Theraulaz, G. (2008). Key behavioural factors in a self-organised fish school model. *Annales Zoologici Fennici*, 45(5):415–428.
- Gislason, H. (2003). The effects of fishing on non-target species and ecosystem structure and function. In Sinclair, M. and Valdimarsson, G., editors, *Responsible Fisheries in the Marine Ecosystem*, page 255. FAO, Cambridge, MA.

- Gjøsæter, H., Dommasnes, A., and Røttingen, B. (1998). The Barents Sea Capelin stock 1972-1997. A synthesis of results from acoustic surveys. *Sarsia*, 83(6):497–510.
- Glass, C. W. and Wardle, C. S. (1989). Comparison of the reactions of fish to a trawl gear, at high and low light intensities. *Fisheries Research*, 7:249–266.
- Goode, G. B. (1880). *A History of the Menhaden*. Orange Judd, New York.
- Grados, D., Fablet, R., Ballón, M., Bez, N., Castillo, R., Lezama-Ochoa, A., and Bertrand, A. (2012). Multiscale characterization of spatial relationships among oxycline depth, macrozooplankton, and forage fish off Peru using geostatistics, principal coordinates of neighbour matrices (PCNMs), and wavelets. *Canadian Journal of Fisheries and Aquatic Sciences*, 69:740–754.
- Grimm, V. (1999). Ten years of individual-based modelling in ecology: What have we learned and what could we learn in the future? *Ecological Modelling*, 115(2-3):129–148.
- Grimm, V. and Railsback, S. F. (2005). *Individual-Based Modeling and Ecology*. Princeton University Press, Princeton, NJ.
- Grosjean, P., Picheral, M., Warembourg, C., and Gorsky, G. (2004). Enumeration, measurement, and identification of net zooplankton samples using the ZOOSCAN digital imaging system. *ICES Journal of Marine Science*, 61:518–525.
- Gutiérrez, M., Castillo, R., Segura, M., Peraltilla, S., and Flores, M. (2012). Trends in spatio-temporal distribution of Peruvian anchovy and other small pelagic fish biomass from 1966-2009. *Latin American Journal of Aquatic Research*, 40(3 SPL. ISS.):633–648.
- Hagström, O., Palmén, L.-E., Häkansson, N., Kästner, D., Rothbart, H., Grygiel, W., and Wyszynski, M. (1989). Acoustic estimates of the herring and the sprat stocks in the Baltic proper October 1989. Technical report, Committee Meeting of the International Council for the Exploration of the Sea.
- Hamilton, W. (1971). Geometry for the selfish herd. *Journal of Theoretical Biology*, 31:295–311.
- Hare, J. A., Quinlan, J. A., Werner, F. E., Blanton, B. O., Govoni, J. J., Forward, R. B., Settle, L. R., and Hoss, D. E. (1999). Larval transport during winter in the SABRE study area: Results of a coupled vertical larval behaviour- three- dimensional circulation model. *Fisheries Oceanography*, 8(2):57–76.
- Hare, J. A., Thorrold, S., Walsh, H. J., Reiss, C., Valle-Levinson, A., and Jones, C. M. (2005). Biophysical mechanisms of larval fish ingress into Chesapeake Bay. *Marine Ecology Progress Series*, 303:295–310.
- Hartman, K. J. and Brandt, S. B. (1995). Trophic resource partitioning, diets, and growth of sympatric estuarine predators. *Transactions of the American Fisheries Society*, 124(4):520–537.
- Hartwell, S. I. and Otto, R. G. (1978). Swimming performance of juvenile menhaden (*Brevoortia Tyrannus*). *Transactions of the American Fisheries Society*, 107(6):793–798.

- Haugland, E. K. and Misund, O. A. (2004). Evidence for a clustered spatial distribution of clupeid fish schools in the Norwegian Sea and off the coast of southwest Africa. *ICES Journal of Marine Science*, 61(7):1088–1092.
- Heinänen, S., Chudzinska, M. E., Brandi Mortensen, J., Teo, T. Z. E., Rong Utne, K., Dok-sæter Sivle, L., and Thomsen, F. (2018). Integrated modelling of Atlantic mackerel distribution patterns and movements: A template for dynamic impact assessments. *Ecological Modelling*, 387:118–133.
- Hemelrijk, C. K. and Kunz, H. (2004). Density distribution and size sorting in fish schools: An individual-based model. *Behavioral Ecology*, 16(1):178–187.
- Hijmans, R. J. (2019). *Geosphere: Spherical Trigonometry*. R package version 1.5-10.
- Hildebrand, S. F. and Schroeder, W. C. (1928). Fishes of Chesapeake Bay. *US Fisheries Bulletin*, 43(1):1–366.
- Hixon, M. A., Anderson, T. W., Buch, K. L., Johnson, D. W., Mcleod, J. B., and Stallings, C. D. (2012). Density dependence and population regulation in marine fish: A large-scale, long-term field manipulation. *Ecological Monographs*, 82(4):467–489.
- Hjort, J. (1914). Fluctuations in the great fisheries of Northern Europe. *Rapports et Procès-Verbaux*, 20:1–228.
- Honkalehto, T., Ressler, P. H., Towler, R. H., and Wilson, C. D. (2011). Using acoustic data from fishing vessels to estimate walleye pollock (*Theragra Chalcogramma*) abundance in the eastern Bering Sea. *Canadian Journal of Fisheries and Aquatic Sciences*, 68:1231–1242.
- Houde, E. D. (2011). *Managing the Chesapeake’s Fisheries: A Work in Progress*. Maryland Sea Grant, College Park, MD.
- Houde, E. D., Annis, E. R., Harding, L. W. J., Mallonee, M. E., and Wilberg, M. J. (2016). Factors affecting the abundance of age-0 Atlantic menhaden (*Brevoortia Tyrannus*) in Chesapeake Bay. *ICES Journal of Marine Science*, 73(9):2238–2251.
- Houde, E. D. and Zastrow, C. E. (1991). Bay anchovy, *Anchoa Mitchilli*. In Funderburk, S., Muhursky, J., Jordan, S., and Riley, D., editors, *Habitat Requirements for Chesapeake Bay Living Resources*. Living Resources Subcommittee, Chesapeake Bay Program, Annapolis, MD, second edition.
- Humphrey, J., Wilberg, M. J., Houde, E. D., and Fabrizio, M. C. (2014). Effects of temperature on age-0 atlantic menhaden growth in Chesapeake Bay. *Transactions of the American Fisheries Society*, 143(5):1255–1265.
- Huse, G. and Ellingsen, I. (2008). Capelin migrations and climate change: A modelling analysis. *Climatic Change*, 87:177–197.
- Huse, G. and Giske, J. (1998). Ecology in Mare Pentium: An individual-based spatio-temporal model for fish with adapted behaviour. *Fisheries Research*, 37:163–178.

- Huston, M., DeAngelis, D. L., and Post, W. (1988). New computer models unify ecological theory. *BioScience*, 38(10):682–691.
- Huth, A. and Wissel, C. (1992). The simulation of the movement of fish schools. *Journal of Theoretical Biology*, 156(3):365–385.
- Huth, A. and Wissel, C. (1994). The simulation of fish schools in comparison with experimental data. *Ecological Modelling*, 75–76(C):135–146.
- ICES (2014). Manual of International Baltic Acoustic Surveys (IBAS). Technical report, International Council for the Exploration of the Sea, Copenhagen, Denmark.
- Illian, J., Penttinen, A., Stoyan, H., and Stoyan, D. (2008). *Statistical Analysis and Modelling of Spatial Point Patterns*. John Wiley & Sons, Ltd, West Sussex.
- June, F. C. and Carlson, F. T. (1971). Food of young Atlantic mehnaden, *Brevoortia Tyrannus*, in relation to metamorphosis. *Fishery Bulletin*, 68(3):493–512.
- Jung, S. and Houde, E. D. (2003). Spatial and temporal variabilities of pelagic fish community structure and distribution in Chesapeake Bay, USA. *Estuarine, Coastal and Shelf Science*, 58:335–351.
- Jung, S. and Houde, E. D. (2004). Production of bay anchovy *Anchoa Mitchilli* in Chesapeake Bay: Application of size-based theory. *Marine Ecology Progress Series*, 281:217–232.
- Kerschbaumer, P., Tritthart, M., and Keckeis, H. (2020). Abundance, distribution, and habitat use of fishes in a large river (Danube, Austria): Mobile, horizontal hydroacoustic surveys vs. a standard fishing method. *ICES Journal of Marine Science*, 77(5):1966–1978.
- Kimbell, H. S. and Morrell, L. J. (2015). Turbidity influences individual and group level responses to predation in guppies, *Poecilia Reticulata*. *Animal Behaviour*, 103:179–185.
- Kimmerer, W. J., Gross, E. S., and Macwilliams, M. L. (2009). Is the response of estuarine nekton to freshwater flow in the San Francisco estuary explained by variation in habitat volume? *Estuaries and Coasts*, 32:375–389.
- Kubečka, J., Godø, O. R., Hickley, P., Prchalová, M., Říha, M., Rudstam, L., and Welcomme, R. (2012). Fish sampling with active methods. *Fisheries Research*, 123–124:1–3.
- Kvasić, I., Mišković, N., and Vukić, Z. (2019). Convolutional neural network architectures for sonar-based diver detection and tracking. In *OCEANS 2019 - Marseille*, pages 1–6, Marseille, France. IEEE.
- Langseth, B. J., Purcell, K. M., Craig, J. K., Schueller, A. M., Smith, J. W., Shertzer, K. W., Creekmore, S., Rose, K. A., and Fennel, K. (2014). Effect of changes in dissolved oxygen concentrations on the spatial dynamics of the Gulf menhaden fishery in the Northern Gulf of Mexico. *Marine and Coastal Fisheries: Dynamics, Management, and Ecosystem Science*, 6:223–234.

- Lankowicz, K. M., Bi, H., Liang, D., and Fan, C. (2020). Sonar imaging surveys fill data gaps in forage fish populations in shallow estuarine tributaries. *Fisheries Research*, 226.
- Le Fur, J. and Simon, P. (2009). A new hypothesis concerning the nature of small pelagic fish clusters: An individual-based modelling study of *Sardinella Aurita* dynamics off West Africa. *Ecological Modelling*, 220:1291–1304.
- Legendre, P. and Fortin, M.-J. (1989). Spatial pattern and ecological analysis. *Vegetatio*, 80:107–138.
- Levin, S. A. (1992). The problem of pattern and scale in ecology. *Ecology*, 73(6):1943–1967.
- Liang, D., Nesslage, G. M., and Wilberg, M. J. (2020). A spatial simulation approach to hydroacoustic survey design: A case study for Atlantic menhaden. *Fisheries Research*, 222:105402.
- Liljestrand, E. M., Wilberg, M. J., and Schueller, A. M. (2019). Estimation of movement and mortality of Atlantic menhaden during 1966–1969 using a Bayesian multi-state mark-recovery model. *Fisheries Research*, 210:204–213.
- Łomnicki, A. (1988). *Population Ecology of Individuals*. Princeton University Press, Princeton, NJ.
- Lucca, B. M. and Warren, J. D. (2018). Acoustically Measured Distribution and Abundance of Atlantic Menhaden (*Brevoortia Tyrannus*) in a Shallow Estuary in Long Island, NY. *Estuaries and Coasts*, 41:1436–1447.
- Lucca, B. M. and Warren, J. D. (2019). Fishery-independent observations of Atlantic menhaden abundance in the coastal waters south of New York. *Fisheries Research*, 218:229–236.
- Luo, J. and Brandt, S. B. (1993). Bay anchovy *Anchoa Mitchilli* production and consumption in mid-Chesapeake Bay based on a bioenergetics model and acoustic measures of fish abundance. *Marine Ecology Progress Series*, 98:223–236.
- Luo, J., Hartman, K. J., Brandt, S. B., Cerco, C. F., and Rippeto, T. H. (2001). A spatially-explicit approach for estimating carrying capacity: An application for the Atlantic Menhaden (*Brevoortia Tyrannus*) in Chesapeake Bay. *Estuaries*, 24(4):545–556.
- Lynch, P. D., Brush, M. J., Condon, E. D., and Latour, R. J. (2010). Net removal of nitrogen through ingestion of phytoplankton by Atlantic menhaden *Brevoortia Tyrannus* in Chesapeake Bay. *Marine Ecology Progress Series*, 401:195–209.
- MacArthur, R. H. (1972). *Geographical Ecology: Patterns in the Distribution of Species*. Harper and Row, New York, NY.
- Marshall, H. G. and Affronti, L. F. (1992). Seasonal phytoplankton development within three rivers in the lower Chesapeake Bay region. *Virginia Journal of Science*, 43(1A):15–23.
- Mason, D. M. and Brandt, S. B. (1999). Space, time, and scale: New perspectives in fish ecology and management. *Canadian Journal of Fisheries and Aquatic Sciences*, 56(S1):1–3.

- McIntire, E. J. B. and Fajardo, A. (2009). Beyond description: The active and effective way to infer processes from spatial patterns. *Ecology*, 90(1):46–56.
- McLane, A. J., Semeniuk, C., McDermid, G. J., and Marceau, D. J. (2011). The role of agent-based models in wildlife ecology and management. *Ecological Modelling*, 222(8):1544–1556.
- Meentemeyer, V. and Box, E. (1987). Scale effects in landscape studies. In Turner, M. G., editor, *Landscape Heterogeneity and Disturbance*, pages 15–34. Springer-Verlag, New York, NY.
- Miller, T. J., Nye, J. A., and Loewensteiner, D. L. (2008). Development and implementation of the Chesapeake Bay fishery-independent multispecies survey (CHESFIMS). Technical report, Chesapeake Biological Laboratory, University of Maryland Center for Environmental Science, Solomons, MD.
- Misund, O. A. (1997). Underwater acoustics in marine fisheries and fisheries research. *Reviews in Fish Biology and Fisheries*, 7:1–34.
- MSFCMA (2007). *Magnuson-Stevens Fishery Conservation and Management Act*. U.S. Department of Commerce: NOAA, National Marine Fisheries Service.
- Mullon, C., Cury, P., and Penven, P. (2002). Evolutionary individual-based model for the recruitment of anchovy (*Engraulis capensis*) in the southern Benguela. *Canadian Journal of Fisheries and Aquatic Sciences*, 59(5):910–922.
- Nathan, R., Getz, W. M., Revilla, E., Holyoak, M., Kadmon, R., Saltz, D., and Smouse, P. E. (2008). A movement ecology paradigm for unifying organismal movement research. *PNAS*, 105(49):19052–19059.
- National Marine Fisheries Service (2018). Fisheries of the United States, 2017. Technical report, US Department of Commerce, NOAA Current Fisheries Statistics No. 2017, Silver Spring, MD.
- National Marine Fisheries Service (2021). Fisheries of the United States, 2019. Technical report, US Department of Commerce, NOAA Current Fisheries Statistics No. 2019, Silver Spring, MD.
- National Oceanographic Service (2012). NOS Hydrographic Surveys Collection [h06877].
- Newman, E. A., Kennedy, M. C., Falk, D. A., and McKenzie, D. (2019). Scaling and complexity in landscape ecology. *Frontiers in Ecology and Evolution*, 7:293.
- Niwa, H.-S. (2004). Space-irrelevant scaling law for fish school sizes. *Journal of Theoretical Biology*, 228(3):347–357.
- Oboshi, T., Kato, S., Mutoh, A., and Itoh, H. (2002). Collective or Scattering: Evolving Schooling Behaviors to Escape from Predator. *Artificial Life - ALIFE*.
- O’Driscoll, R. L. (1998). Description of spatial pattern in seabird distributions along line transects using neighbour K statistics. *Marine Ecology Progress Series*, 165:81–94.

- Okabe, A. and Yamada, I. (2001). The K-function method on a network and its computational implementation. *Geographical Analysis*, 33(3):271–290.
- Overholtz, W. J. (2002). The Gulf of Maine-Georges Bank Atlantic herring (*Clupea Harengus*): Spatial pattern analysis of the collapse and recovery of a large marine fish complex. *Fisheries Research*, 57(3):237–254.
- Parrish, J. K. (1989). Re-examining the selfish herd: Are central fish safer? *Animal Behaviour*, 38(6):1048–1053.
- Parrish, J. K. and Turchin, P. (1997). Individual decisions, traffic rules, and emergent pattern in schooling fish. In Parrish, J. K. and Hamner, W. M., editors, *Animal Groups in Three Dimensions: How Species Aggregate*, pages 126–142. Cambridge University Press, New York, NY.
- Parrish, J. K., Viscido, S. V., and Grünbaum, D. (2002). Self-organized fish schools: An examination of emergent properties. *Biological Bulletin*, 202(3):296–305.
- Perry, G. L. W., Miller, B. P., and Enright, N. J. (2006). A comparison of methods for the statistical analysis of spatial point patterns in plant ecology. *Plant Ecology*, 187(1):59–82.
- Petitgas, P. and Levenez, J. (1996). Spatial organization of pelagic fish: Echogram structure, spatio-temporal condition, and biomass in Senegalese waters. *ICES Journal of Marine Science*, 53:147–153.
- Petitgas, P., Reid, D., Carrera, P., Iglesias, M., Georgakarakos, S., Liorzou, B., and Massé, J. (2001). On the relation between schools, clusters of schools, and abundance in pelagic fish stocks. *ICES Journal of Marine Science*, 58(6):1150–1160.
- Pikitch, E. K., Boersma, P. D., Boyd, I. L., Conover, D. O., Cury, P. M., Essington, T. E., Heppell, S. S., Houde, E. D., Mangel, M., Pauly, D., Plagányi, É., Sainsbury, K. J., and Steneck, R. S. (2012). Little fish, big impact: Managing a crucial link in ocean food webs. Technical report, Lenfest Ocean Program, Washington, DC.
- Pikitch, E. K., Rountos, K. J., Essington, T. E., Santora, C., Pauly, D., Watson, R., Sumaila, U. R., Boersma, P. D., Boyd, I. L., Conover, D. O., Cury, P. M., Heppell, S. S., Houde, E. D., Mangel, M., Plagányi, É., Sainsbury, K. J., Steneck, R. S., Geers, T. M., Gownaris, N., and Munch, S. B. (2014). The global contribution of forage fish to marine fisheries and ecosystems. *Fish and Fisheries*, 15(1):43–64.
- Pikitch, E. K., Santora, C., Babcock, E. A., Bakun, A., Bonfil, R., Conover, D. O., Dayton, P., Doukakis, P., Fluharty, D., Heneman, B., Houde, E. D., Link, J., Livingston, P. A., Mangel, M., McAllister, M. K., Pope, J., and Sainsbury, K. J. (2004). Ecosystem-based fishery management. *Science*, 305(5682):346–347.
- Pitcher, T. J. (1986). Functions of Shoaling Behaviour in Teleosts. In Pitcher, T. J., editor, *The Behaviour of Teleost Fishes*, pages 294–337. Springer, Boston, MA, first edition.

- Pitcher, T. J., Misund, O. A., Fernö, A., Totland, B., and Melle, V. (1996). Adaptive behaviour of herring schools in the Norwegian Sea as revealed by high-resolution sonar. *ICES Journal of Marine Science*, 53(2):449–452.
- Potier, M., Petitgas, P., and Petit, D. (1997). Interaction between fish and fishing vessels in the Javanese purse seine fishery. *Aquatic Living Resources*, 10(3):149–156.
- Quinlan, J. A., Blanton, B. O., Miller, T. J., and Werner, F. E. (1999). From spawning grounds to the estuary: Using linked individual-based and hydrodynamic models to interpret patterns and processes in the oceanic phase of Atlantic menhaden *Brevoortia Tyrannus* life history. *Fisheries Oceanography*, 8(SUPPL. 2):224–246.
- R Core Team (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rahbek, C. (2005). The role of spatial scale and the perception of large-scale species-richness patterns. *Ecology Letters*, 8:224–239.
- Railsback, S. F., Lamberson, R. H., Harvey, B. C., and Duffy, W. E. (1999). Movement rules for individual-based models of stream fish. *Ecological Modelling*, 123(2-3):73–89.
- Reuter, H. and Breckling, B. (1994). Selforganization of fish schools: An object-oriented model. *Ecological Modelling*, 75–76:147–159.
- Reynolds, C. W. (1987). Flocks, herds, and schools: A distributed behavioral model. *Computer Graphics*, 21(4):25–34.
- Ricker, W. E. (1975). Computation and interpretation of biological statistics of fish populations. *Bulletin of the Fisheries Research Board of Canada*, 191:1–382.
- Rieucou, G., Boswell, K. M., Kimball, M. E., Diaz, G., and Allen, D. M. (2015). Tidal and diel variations in abundance and schooling behavior of estuarine fish within an intertidal salt marsh pool. *Hydrobiologia*, 753:149–162.
- Ripley, B. D. (1981). *Spatial Statistics*. Wiley, Hoboken, NJ.
- Rodriguez-Pinto, I. I., Rieucou, G., Handegard, N. O., and Boswell, K. M. (2020). Environmental context elicits behavioural modification of collective state in schooling fish. *Animal Behaviour*, 165:107–116.
- Rose, K. A., Allen, J. I., Artioli, Y., Barange, M., Blackford, J., Carlotti, F., Cropp, R., Daewel, U., Edwards, K., Flynn, K., Hill, S. L., HilleRisLambers, R., Huse, G., Mackinson, S., Megrey, B., Moll, A., Rivkin, R., Salihoglu, B., Schrum, C., Shannon, L. J., Shin, Y.-J., Smith, S. L., Smith, C., Solidoro, C., St. John, M., and Zhou, M. (2010). End-to-end models for the analysis of marine ecosystems: Challenges, issues, and next steps. *Marine and Coastal Fisheries: Dynamics, Management, and Ecosystem Science*, 2(1):115–130.
- Rose, K. A., Cowan, J. H., Winemiller, K. O., Myers, R. A., and Hilborn, R. (2001). Compensatory density dependence in fish populations: Importance, controversy, understanding and prognosis: Compensation in fish populations. *Fish and Fisheries*, 2(4):293–327.

- Rose, K. A., Cowan, J. H. J., Clark, M. E., Houde, E. D., and Wang, S.-B. (1999). An individual-based model of bay anchovy population dynamics in the mesohaline region of Chesapeake Bay. *Marine Ecology Progress Series*, 185:113–132.
- Rountos, K. J. (2016). Defining forage species to prevent a management dilemma. *Fisheries*, 41(1).
- Samedy, V., Wach, M., Lobry, J., Selleslagh, J., Pierre, M., Josse, E., and Boët, P. (2015). Hydroacoustics as a relevant tool to monitor fish dynamics in large estuaries. *Fisheries Research*, 172:225–233.
- Scully, M. E. (2016). Mixing of dissolved oxygen in Chesapeake Bay driven by the interaction between wind-driven circulation and estuarine bathymetry. *Journal of Geophysical Research: Oceans*, 121(8):5639–5654.
- Scutt Phillips, J., Gupta, A. S., Senina, I., van Sebille, E., Lange, M., Lehodey, P., Hampton, J., and Nicol, S. (2018). An individual-based model of skipjack tuna (*Katsuwonus Pelamis*) movement in the tropical Pacific ocean. *Progress in Oceanography*, 164:63–74.
- SEDAR (2015). SEDAR 40 - Atlantic Menhaden stock assessment report. Technical report, SEDAR, North Charleston, SC.
- SEDAR (2020). SEDAR 69 - Atlantic Menhaden benchmark stock assessment report. Technical report, ASMFC, North Charleston, SC.
- Shahrestani, S., Bi, H., Lyubchich, V., and Boswell, K. M. (2017). Detecting a nearshore fish parade using the adaptive resolution imaging sonar (ARIS): An automated procedure for data analysis. *Fisheries Research*, 191:190–199.
- Shaw, E. (1978). Schooling fishes: The school, a truly egalitarian form of organization in which all members of the group are alike in influence, offers substantial benefits to its participants. *American Scientist*, 66(2):166–175.
- Simmonds, E. J. (2009). Evaluation of the quality of the North Sea herring assessment. *ICES Journal of Marine Science*, 66(8):1814–1822.
- Simmonds, E. J. and MacLennan, D. N. (2005). *Fisheries Acoustics: Theory and Practice*. Blackwell Publishing, Oxford, second edition.
- Simpson, C. A., Bi, H., Liang, D., Wilberg, M. J., Schueller, A. M., Nessler, G. M., and Walsh, H. J. (2017). Spawning locations and larval dispersal of Atlantic menhaden during 1977–2013. *ICES Journal of Marine Science*, 74(6):1574–1586.
- Simpson, C. A., Wilberg, M. J., Bi, H., Schueller, A. M., Nessler, G. M., and Walsh, H. J. (2016). Trends in relative abundance and early life survival of Atlantic Menhaden during 1977–2013 from long-term ichthyoplankton programs. *Transactions of the American Fisheries Society*, 145(5):1139–1151.

- Skaret, G. (2003). Spawning of herring: Day or night, today or tomorrow? *Aquatic Living Resources*, 16(3):299–306.
- Smith, A. D. M., Brown, C. J., Bulman, C. M., Fulton, E. A., Johnson, P., Kaplan, I. C., Lozano-Montes, H., Mackinson, S., Marzloff, M., Shannon, L. J., Shin, Y.-J., and Tam, J. (2011). Impacts of fishing low-trophic sevel species on marine ecosystems. *Science*, 333(6046):1147–1150.
- Soria, M., Petitgas, P., and Bahri, T. (1998). On the size of fish schools and clusters: A spatial analysis of multibeam sonar images in the Mediterranean Sea. *ICES CM 1998/J:8*.
- Spares, A. D., Stokesbury, M. J. W., O’Dor, R. K., and Dick, T. A. (2012). Temperature, salinity and prey availability shape the marine migration of Arctic char, *Salvelinus Alpinus*, in a macrotidal estuary. *Marine Biology*, 159(8):1633–1646.
- Swartzman, G. (1997). Analysis of the summer distribution of fish schools in the Pacific Eastern Boundary Current. *ICES Journal of Marine Science*, 54:105–116.
- Szedlmayer, S. T. and Able, K. W. (1996). Patterns of seasonal availability and habitat use by fishes and decapod crustaceans in a southern New Jersey estuary. *Estuaries*, 19(3):697–709.
- Taylor, J. C. and Rand, P. S. (2003). Spatial overlap and distribution of anchovies (*Anchoa Spp.*) and copepods in a shallow stratified estuary. *Aquatic Living Resources*, 16(3):191–196.
- Thompson, S. K. (2012). *Sampling*. Wiley.
- Toresen, R., Gjøsæter, H., and de Barros, P. (1998). The acoustic method as used in the abundance estimation of capelin (*Mallotus Villosus* Müller) and herring (*Clupea Harengus* Linné) in the Barents Sea. *Fisheries Research*, 34:27–37.
- Trenkel, V. M., Ressler, P. H., Jech, J. M., Guannoulaki, M., and Taylor, C. (2011). Underwater acoustics for ecosystem-based management: State of the science and proposals for ecosystem indicators. *Marine Ecology Progress Series*, 442:285–301.
- Trumpickas, J., Pinder, M., and Dunlop, E. S. (2020). Effects of vessel size and trawling on estiamtes of pelagic fish backscatter in Lake Huron. *Fisheries Research*, 224:105430.
- Tunstrøm, K., Katz, Y., Ioannou, C. C., Huepe, C., Lutz, M. J., and Couzin, I. D. (2013). Collective states, multistability and transitional behavior in schooling fish. *PLoS Computational Biology*, 9(2).
- Turner, M. G., Dale, V. H., and Gardner, R. H. (1989a). Predicting across scales: Theory development and desting. *Landscape Ecology*, 3(3/4):245–252.
- Turner, M. G., O’Neill, R. V., Gardner, R. H., and Milne, B. T. (1989b). Effects of changing spatial scale on the analysis of landscape pattern. *Landscape Ecology*, 3(3/4):153–162.
- Tyler, J. A. and Rose, K. A. (1994). Individual variability and spatial heterogeneity in fish population models. *Reviews in Fish Biology and Fisheries*, 4(1):91–123.

- Vabø, R. and Nøttestad, L. (1997). An individual based model of fish school reactions: Predicting antipredator behaviour as observed in nature. *Fisheries Oceanography*, 6(3):155–171.
- van de Koppel, J., Bouma, T. J., and Herman, P. M. J. (2012). The influence of local- and landscape-scale processes on spatial self-organization in estuarine ecosystems. *The Journal of Experimental Biology*, 215:962–967.
- van der Vaart, E., Johnston, A. S. A., and Sibly, R. M. (2016). Predicting how many animals will be where: How to build, calibrate and evaluate individual-based models. *Ecological Modelling*, 326:113–123.
- Vinagre, C., Costa, M. J., and Dunne, J. A. (2017). Effect of spatial scale on the network properties of estuarine food webs. *Ecological Complexity*, 29:87–92.
- Viscido, S. V., Parrish, J. K., and Grünbaum, D. (2004). Individual behavior and emergent properties of fish schools: A comparison of observation and theory. *Marine Ecology Progress Series*, 273:239–249.
- Walter III, J. F. and Austin, H. M. (2001). Diet composition of large Striped Bass (*Morone Saxatilis*) in Chesapeake Bay. *Fish Bulletin*, 101(2):414–423.
- Wang, S. B., Cowan Jr, J. H., Rose, K. A., and Houde, E. D. (1997). Individual-based modelling of recruitment variability and biomass production of bay anchovy in mid-Chesapeake Bay. *Journal of Fish Biology*, 51(sA):101–120.
- Wang, S. B. and Houde, E. D. (1995). Distribution, relative abundance, biomass and production of bay anchovy *Anchoa Mitchilli* in the Chesapeake Bay. *Marine Ecology Progress Series*, 121:27–38.
- Ward, A. J. W., Schaerf, T. M., Herbert-Read, J. E., Morrell, L., Sumpter, D. J. T., and Webster, M. M. (2017). Local interactions and global properties of wild, free-ranging stickleback shoals. *Royal Society Open Science*, 4(7):170043.
- Watkins, K. S. and Rose, K. A. (2013). Evaluating the performance of individual-based animal movement models in novel environments. *Ecological Modelling*, 250:214–234.
- Watkins, K. S. and Rose, K. A. (2014). The effects of spatial and temporal resolution in simulating fish movement in Individual-Based Models. *Transactions of the American Fisheries Society*, 143(5):1143–1160.
- Watkins, K. S. and Rose, K. A. (2017). Simulating individual-based movement in dynamic environments. *Ecological Modelling*, 356:59–72.
- Wehenkel, C., Brazão-Protázio, J. M., Carrillo-Parra, A., Martínez-Guerrero, J. H., and Crecente-Campo, F. (2015). Spatial distribution patterns in the very rare and species-rich *Picea Chihuahuana* tree community (Mexico). *PLoS ONE*, 10(10):e0140442.
- Weise, M. J. and Harvey, J. T. (2008). Temporal variability in ocean climate and California sea lion diet and biomass consumption: Implications for fisheries management. *Marine Ecology Progress Series*, 373:157–172.

- Werner, F. E., Quinlan, J. A., Lough, R. G., and Lynch, D. R. (2001). Spatially-explicit individual based modeling of marine populations: A review of the advances in the 1990s. *Sarsia*, 86(6):411–421.
- Wiegand, T., He, F., and Hubbell, S. P. (2013). A systematic comparison of summary characteristics for quantifying point patterns in ecology. *Ecography*, 36(1):92–103.
- Wiegand, T. and Moloney, K. A., editors (2013). *Handbook of Spatial Point-Pattern Analysis in Ecology*. CRC Press, New York, NY.
- Wiens, J. (1989). Spatial scaling in ecology. *Functional Ecology*, 3:385–397.
- Wilkens, F. and Lewis, R. (1971). Abundance and distribution of young Atlantic menhaden, *Brevoortia Tyrannus*, in the White Oak Estuary, North Carolina. *Fishery Bulletin, U.S.*, 69:783–789.
- Wisz, M. S., Pottier, J., Kissling, W. D., Pellissier, L., Lenoir, J., Damgaard, C. F., Dormann, C. F., Forchhammer, M. C., Grytnes, J. A., Guisan, A., Heikkinen, R. K., Høye, T. T., Kühn, I., Luoto, M., Maiorano, L., Nilsson, M. C., Normand, S., Öckinger, E., Schmidt, N. M., Termansen, M., Timmermann, A., Wardle, D. A., Aastrup, P., and Svenning, J. C. (2013). The role of biotic interactions in shaping distributions and realised assemblages of species: Implications for species distribution modelling. *Biological Reviews*, 88(1):15–30.
- Wood, R. J. and Austin, H. M. (2009). Synchronous multidecadal fish recruitment patterns in Chesapeake Bay, USA. *Canadian Journal of Fisheries and Aquatic Sciences*, 66:496–508.
- Wood, R. J. and Houde, E. D. (2003). Variability in the dynamics of forage fish abundances in Chesapeake Bay: Retrospective analysis, models and synthesis. Technical report, NOAA Chesapeake Bay Office, Chesapeake Bay Fisheries Research Program, Annapolis, MD.
- Wood, S. N. (2017). *Generalized Additive Models: An Introduction with R*. Chapman and Hall/CRC, Boca Raton, FL.
- Yin, T., Zang, X., Hou, Z., Jacobson, P. T., Mueller, R. P., and Deng, Z. (2020). Bridging the gap between laboratory and field experiments in american eel detection using transfer learning and convolutional neural network. In *Proceedings of the 53rd Hawaii International Conference on System Sciences*, pages 932–939.
- York, R. (2017). Why Petroleum Did Not Save the Whales. *Socius*, 3:1–13.
- Zamon, J. E. (2003). Mixed species aggregations feeding upon herring and sandlance schools in a nearshore archipelago depend on flooding tidal currents. *Marine Ecology Progress Series*, 261:243–255.
- Zastrow, C. E., Houde, E. D., and Morin, L. G. (1991). Spawning, fecundity, hatch-date frequency, and young-of-the-year growth of bay anchovy *Anchoa Mitchilli* in mid-Chesapeake Bay. *Marine Ecology Progress Series*, 73(2,3):161–171.

Zhang, L., Yang, Y., Bi, H., Shahrestani, S., and Fan, C. (2017). Self-organized fish schooling: From individual variability to synchrony. In *Association for the Sciences of Limnology and Oceanography 2017 Ocean Sciences Meeting.*, Honolulu, Hawaii.