Summer 8-2021

The Impacts of Climate Change on the Gulf of Maine Northern Shrimp (Pandalus borealis) Distribution, Reproduction, and Life

Hsiao-Yun Chang
hsiaoyun.chang@maine.edu

THE IMPACTS OF CLIMATE CHANGE ON THE GULF OF MAINE NORTHERN SHRIMP (*PANDALUS BOREALIS*) DISTRIBUTION, REPRODUCTION, AND LIFE CYCLE

By

Hsiao-Yun Chang

B.S. National Taiwan Ocean University, 2008

M.S. National Taiwan University, 2011

A DISSERTATION

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

(in Ecology and Environmental Sciences)

The Graduate School

The University of Maine

August 2021

Advisory Committee:

Yong Chen, Professor of Fisheries Population Dynamics, Advisor

Anne Richards, NOAA Northeast Fisheries Science Center scientist (retired)

Gayle Zydlewski, Professor of Marine Biology

Walter Golet, Assistant Professor of Marine Biology

Kate Beard-Tisdale, Professor of Spatial Information Science

THE IMPACTS OF CLIMATE CHANGE ON THE GULF OF MAINE NORTHERN SHRIMP (PANDALUS BOREALIS) DISTRIBUTION, REPRODUCTION, AND LIFE CYCLE

By Hsiao-Yun Chang

Dissertation Advisor: Dr. Yong Chen

An Abstract of the Dissertation Presented
in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy
(in Ecology and Environmental Sciences)
August 2021

The Gulf of Maine northern shrimp (*Pandalus borealis*) once supported a significant winter fishery for the Gulf of Maine (GOM). Although the shrimp fishery is not comparable to the lobster business, it provided fishermen and many coastal communities jobs and incomes in winters after lobster seasons. However, a moratorium has been put on the shrimp fishery since 2014 due to record low population abundance and perceived recruitment failures. The recruitment failures have been correlated with warming water temperatures over the past decade. The GOM has been recognized as experiencing rapid warming as a result of global climate change. Uncertain impacts of the changing environment on the life cycle and the fishery of northern shrimp could hamper the efforts to rebuild a resilient and sustainable shrimp fishery. Consequently, there is a pressing need to understand the impacts of climate change on the life cycle and population dynamics of the GOM northern shrimp. The objectives of this research project are to 1) develop a cost-effective sampling protocol for building a comprehensive fecundity database including maternal body size, and number and size of eggs; 2) examine the impacts of climate-induced environmental variabilities on egg mortality; 3) develop a complete

size-fecundity relationship in order to derive a robust estimate of reproductive potential of the population; 4) illuminate the effect of water temperature on spatial structure for each life stage; and 5) investigate the relationship between environmental variabilities and habitat suitability in northern shrimp spawners' distribution. The findings reveal that the GOM bottom temperatures might have changed considerably over the past fifty years; however, the correlation between water temperature and parasitic infection eggs was not significant and the changes in reproductive potential might be related to population density rather than bottom temperature. The results also showed that the distributions of mature groups were getting patchier and shifting northward, which were correlated with declined population abundance and warming temperature, respectively. Furthermore, the quality of habitat has declined significantly for mature groups in summer and fall, especially in the 2010s, which could result from warming temperature and subsequently lead to declined spawning stock biomass.

DEDICATION

This dissertation is dedicated to all the northern shrimp in the Gulf of Maine and the fishery communities, who were impacted by the collapse of the northern shrimp fishery.

ACKNOWLEDGMENTS

First and foremost, I would like to express my deepest gratitude to my advisor Yong Chen for his support and guidance throughout the process of completing this dissertation. I would also like to thank my committee members Anne Richards, Gayle Zydlewski, Kate Beard-Tisdale, and Walter Golet for their guidance and encouragement. I would also like to extend my sincere gratitude to Anne Richards and David Townsend for their meticulous and insightful review of the manuscripts. And special thanks to Rachael Klose and many other undergraduate assistants for the help they have provided for lab work and data collection. Their input and contribution to this project is greatly appreciated.

Furthermore, the biological samples and data provided by Northeast Fisheries Sciences Center (NEFSC) and Maine Department of Marine Resources (ME DMR) have been essential in the completion of this project; therefore, I am grateful for their generous support. And special thanks to Margaret Hunter and Katherine Thompson at ME DMR and Catherine Fillo and Peter Chase at NEFSC for providing technical help with this project's initial setup.

Additionally, I would like to thank all the lab mates for their support and feedback on my work, especially to Kisei Tanaka, Bai Li, and Mackenzie Mazur for their enthusiasm and help throughout this project.

This project would not have been accomplished without the funding support from the NOAA Salltonsall-Kennedy grant for providing funding support to this project.

I would like to thank my mentor, Chi-Lu Sun, for his encouragement and advice throughout my academic path. Finally, I would like to thank all my friends and family for their understanding and unconditional love and support during the portion of my life journey.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. AN INTRODUCTION TO NORTHERN SHRIMP (*PANDALUS BOREALIS*)

1.1 Biology, ecology, and life cycle

There are several populations of northern shrimp (*Pandalus borealis*) in the north Atlantic coastal shelf (Jorde et al. 2015), including Canada, eastern and western Greenland, Norway, Iceland, Alaska, and the Gulf of Maine (GOM). The GOM northern shrimp is considered a genetically distinct unit stock at the southernmost of their distribution, which makes them perceived to be very susceptible to environmental variabilities (Richards et al. 2012).

Northern shrimp is a protandric hermaphrodite (Shumway et al. 1985). They start their lives as males in inshore area in the Gulf of Maine (GOM). After 6-7 molts of larval stages, they grow into juveniles at around 5-9 months (Table 1.1). The molting rates depend on the water temperature (Stickney and Perkins 1977). After the age of 1 year old, they start to migrate to offshore areas and grow into mature adult males (Apollonio et al. 1986). They undergo the transitional stage at the age of 3 years before they turn into mature females (Shumway et al. 1985). Mating and spawning of mature males and females takes place in late summer and fall in offshore areas in the GOM (Shumway et al. 1985). After spawning, females attach the fertilized egg to their pleopods, becoming ovigerous females (Haynes and Wigley 1969; Apollonio et al. 1986). The ovigerous females carry the eggs until they migrate to inshore areas for shedding the embryos in winter and early spring (Shumway et al. 1985; Apollonio et al. 1986). Some females are able to complete the second reproduction cycle before they die at the age of 6; however, Haynes and Wigley (1969) indicated that there could be a high mortality rate for females after the first reproduction cycle.

Table 1.1. Life cycle of Gulf of Maine northern shrimp. Sp: spring, Su: summer, Fa: fall, Wi: winter, F1: Female I, OviF: Ovigerous Female, F2: Female II.

Year	0				1				2				3				4				5			
Season	Sp	Su	Fa	Wi	Sp	Su	Fa	Wi	Sp	Su	Fa	Wi	Sp	Su	Fa	Wi	Sp	Su	Fa	Wi	Sp	Su	Fa	Wi
Life Stage	Immature Male										Mature Male				Transition		F1	OviF			F2		OviF	

Apart from seasonal inshore-offshore migration, diurnal vertical migration was also observed for non-ovigerous shrimp at night when they arose off bottom for feeding, while ovigerous females remained on the bottom due to limited mobility (Apollonio and Dunton 1969; Apollonio et al. 1986). Northern shrimp is an opportunistic omnivore, feeding on polychaetes, molluscs, crustaceans, mud, etc., their feeding patterns vary with seasons and geographic locations (Apollonio and Dunton 1969; Shumway et al. 1985).

1.2 Fishery and stock status

Northern shrimp supported a significant winter fishery for the GOM, targeting ovigerous females when they move to the shore to shed their eggs in winter (ASMFC NSTC 2019). However, the GOM northern shrimp fishery experienced several boom-and-bust cycles since the fishery was established in 1938 (Fig. 1.1). The landings in the 1940s were reported to be low; with landings declining in the late 1940s (ASMFC NSTC 2019). The fishery stopped from 1954 through 1957 due to low abundance which was associated with high temperature at a 4-year lag

(Dow 1977). The fishery resumed in 1958, and the landings grew exponentially and reached above 10000 mt in the early 1970s (Fig. 1.1).

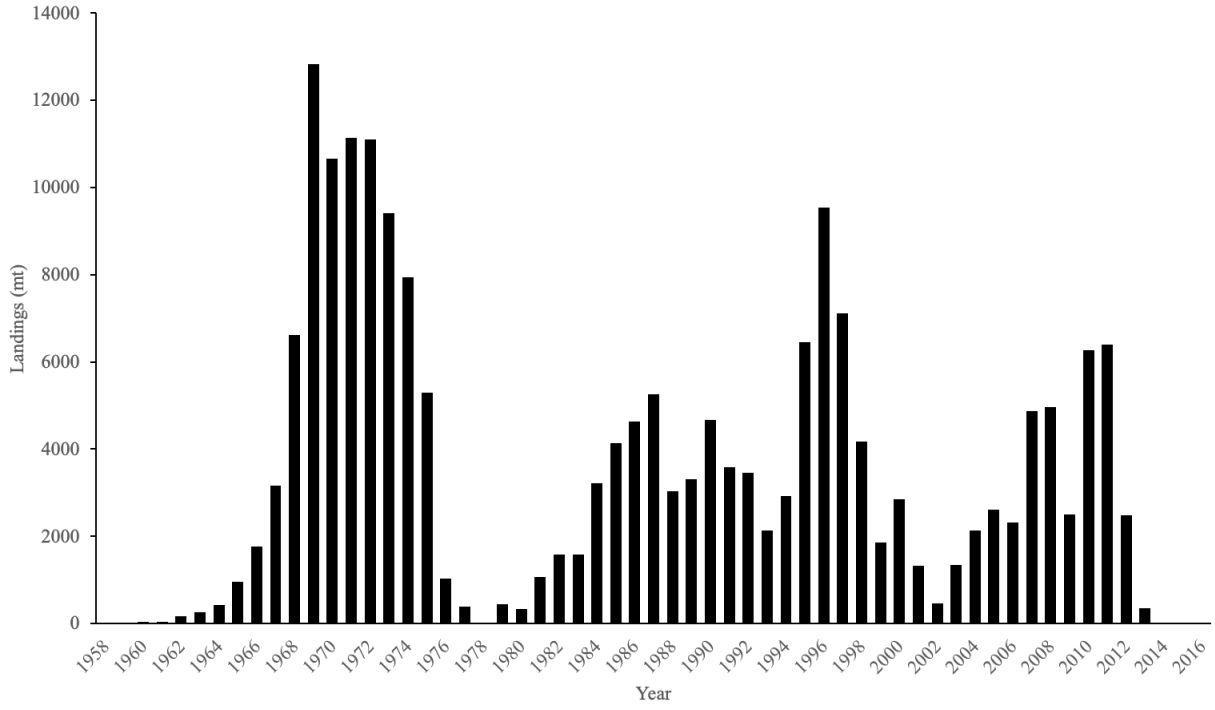


Figure 1.1. Annual commercial landings of the Gulf of Maine northern shrimp (data source: Atlantic States Marine Fisheries Commission Northern Shrimp Technical Committee, 2013).

The summer fisheries developed in the early 1970s which impacted northern shrimp at all ages including young males (Clark et al. 1999; ASMFC NSTC 2019). Clark et al. (1999) indicated that the summer fisheries was a major contributor to the increased fishing mortality and likely lead to subsequent recruitment failures. The fisheries declined very quickly and collapsed in 1977 with less than 500 mt landed (Clark et al. 1999; ASMFC NSTC 2019). The fishery was closed again in 1978. Since this second collapse of the shrimp fishery, the fishing months have been restricted to December through May (Clark et al. 1999; ASMFC NSTC 2019).

The third collapse happened in the 2010s. The landings declined from 2485 mt in 2012 to 345.5 mt in 2013 (Fig. 1.1). A moratorium has been imposed on the fishery since 2014, however the stock has not recovered yet. The third collapse of the shrimp fishery, also the most recent collapse, was due to low levels of abundance and recruitment failures in consecutive years (ASMFC NSTC 2019). The recruitment failures have been perceived to be correlated with warming water temperatures in the past decade (ASMFC NSTC 2019) as a result of global warming (Kavanaugh et al. 2017).

1.3 Overview

Temperature has been considered an important factor in shifting the seasonal migration patterns of geographic distribution (Apollonio et al. 1986) and in influencing phenology (Richards 2012) of the GOM northern shrimp. Various hypotheses that have been postulated for the recent collapse of the population were also temperature-related, including the temperature effects on parasite incidence, reproduction, habitat quality, and spatial structure. Therefore, this dissertation aims to examine these hypotheses and propose possible explanations for the collapse of the fishery.

Chapter 2 uses simulation of a resampling approach to evaluate two sampling strategies, developing a cost-effective protocol for collecting fecundity data from biological samples. Chapter 3 evaluates the effects of biotic and abiotic factors on parasite-infected egg mortality as well as proportion of infected females. Chapter 4 evaluates the effects of environmental variables on reproductive potential of northern shrimp population. Chapter 5 uses survey data to explore the relationships between spatial structure and water temperature and population abundance. Chapter 6 examines the temperature effects on habitat quality and the relation to spawning stock

biomass index. This dissertation aims to enhance our understanding of northern shrimp response to the environmental changes, providing important information for fishery management and conservation.

CHAPTER 2. EVALUATING SAMPLING STRATEGIES FOR COLLECTING SIZE-BASED FISH FECUNDITY DATA: AN EXAMPLE OF GULF OF MAINE NORTHERN SHRIMP *PANDALUS BOREALIS*

2.1 Abstract

Fecundity information is critical in determining reproductive potential of a population. Collecting fecundity data, however, can be cost prohibitive or ineffective if a sampling protocol is not well designed. Inappropriate sampling can lead to biased estimates of fecundity, which may result in biased estimate of reproductive potential. Processing egg samples tends to be time-consuming and labor-intensive. For many fish and crustacean species, fecundity is dependent on female sizes. Nevertheless, at extreme size classes, fecundity may decrease or level off due to senescence. In order to account for this maternal effect, female samples of a wide size range need to be collected for developing a complete relationship between fecundity and body sizes. Using the Gulf of Maine northern shrimp, *Pandalus borealis*, as an example, we evaluated two sampling strategies, simple random sampling and size-based stratified random sampling, with a different number of sampling locations and different number of animals sampled per sampling location or length interval. The study shows that both the sampling strategies, simple random sampling and size-based stratified random sampling, can generate representative samples. However, the simulation analysis suggests that when the population size distribution is skewed with a lack of large and/or small individuals, size-based stratified random sampling is preferred due to lower variation in differences of means and medians between samples and the population. This study provides a simulation framework for identifying a cost-effective sampling protocol that can improve the estimate of fecundity, leading to an improved estimate of fish population reproductive potential.

2.2 Introduction

For many crustaceans and fish species, reproductive output of a female individual tends to increase with body size as larger females have higher capacity to accommodate more eggs or offspring (Hannah et al. 1995; Hixon et al. 2014). However, the relationship between reproductive output and female body size is usually not linear. Instead, reproductive output tends to increase approximately exponentially with body size (Hixon et al. 2014; Barneche et al. 2018). At extreme size classes, however, reproductive output of a female may decrease or level off due to senescence (Shelton et al. 2012). In order to account for this maternal effect, a wide range of sizes of females should be collected for developing a comprehensive relationship between reproductive output and female body sizes in order to have a robust estimate of reproductive potential of a population (Marshall et al. 2006).

Sample sizes and locations may also influence the quality of fecundity estimates because of large variability in space and among individuals (Parsons and Tucker, 1986; Hannah et al. 1995). An insufficient number of samples may lead to underestimated or overestimated fecundity for a given size of fish. A large number of samples is usually encouraged for estimating biological traits of a population. However, collecting biological data such as fecundity can be very time-consuming and labor-intensive laboratory processes (Rogers et al. 2019). Excessive samples are not only a waste of resources, but also a source of unnecessary pressure on the population especially when the stock is in an unhealthy status. Therefore, to reach a balance between deriving robust estimates of life history traits and efficient use of available resources, an appropriate sampling design is important for collecting biological samples from a population.

Based on availability of resources and samples, two sampling designs are often used to collect biological data like fecundity: simple random sampling (Collins et al. 1998; Pennington and Helle, 2011) and stratified random sampling (Hannah et al. 1995). Simple random sampling is to randomly select samples from a population. Stratified random sampling is to divide the population into more than one group (e.g. length-intervals), and to randomly select samples from each group. In general, size-based stratified random sampling is theoretically more appropriate for collecting fecundity data, as it is more likely to include samples from each classification (length intervals), thus able to establishing a more complete biological database and fecundity-body size relationship over a full size range. However, it might not be feasible for some species whose gravid individuals are encountered by chance. In addition, it takes extra effort to classify each individual before randomly sampling from each stratum. In this case, simple random sampling is usually used as a default sampling strategy. Nevertheless, whether the samples collected by these 2 sampling schemes can be representative of the population is rarely discussed.

The Gulf of Maine (GOM) northern shrimp used to support a significant winter fishery for the New England states (ASMFC NSTC 2018), however the shrimp fishery has been on moratorium since 2014 due to presumed recruitment failures which were perceived to be a consequence of warming water temperature in the GOM in the past several years (Richards et al. 2012; ASMFC NSTC 2018). Recruitment is usually related to reproductive potential of a population which can be evaluated with fecundity. However, the relationship between shrimp body sizes and fecundity was estimated more than thirty years ago using 47 ovigerous females selected for size and wholeness of the egg mass (Haynes and Wigley 1969). These data were fitted with a parabola for estimating fecundity for females larger than 22-mm (Richards et al.

2012; ASMFC NSTC 2018): fecundity = -0.198 l2+128.81l-17821, where l is carapace length (0.1-mm). The body size-fecundity relationship estimated with the parabola was likely biased as small spawners were not included in their study and the estimated parabola equation generated negative values for fecundity when female carapace length was below 20-mm. Therefore, there is a pressing need to develop an updated fecundity database to provide more robust estimates for northern shrimp reproductive potential, which makes northern shrimp an appropriate case study.

The aim is to compare different sampling strategies for estimating fecundity for species such as northern shrimp that have maternal effects on fecundity and the number of ovigerous individuals were unevenly collected in sampling locations. The study can identify a cost-effective sampling design for collecting fecundity data, leading to improved fecundity estimation.

2.3 Materials and Methods

This study uses simulation of resampling approach to simulate different sampling strategy scenarios based on collected survey data.

2.3.1 NEFSC fall bottom trawl survey data

The GOM northern shrimp spawning season takes place in late summer and fall, and most females become ovigerous in fall. Therefore, the ovigerous females used for the fecundity study were sampled in the Northeast Fisheries Science Center (NEFSC) fall bottom trawl surveys which were designed for multispecies surveys in the northeast coastal areas. As the surveys are not specifically designed for northern shrimp, in the sampling location with presence of ovigerous females, the number of shrimp varied from one to several hundred among tows.

Given the limited resources, it is unrealistic to process all collected shrimp. Thus, there is a need to optimize the number of sampling locations in a year and number of shrimp collected in a sampling location. Moreover, as many other species are collected in the survey, which face similar needs, the methodology developed in this survey are applicable to other species.

The northern shrimp data and tow information were collected by NEFSC fall bottom trawl surveys (Smith 2002) from 2012 to 2016, including dorsal carapace length (DCL), life stage, date of catch, and longitude and latitude of sampling location. The DCLs of shrimp were measured to the nearest 0.1-mm, from the posterior limit of eye socket to the posterior limit of dorsal carapace (Haynes and Wigley 1969). Only ovigerous female data were used for simulation as the ultimate goal was to collect fecundity data based on maternal body sizes.

2.3.2 Simulation of resampling study

Data from 2012 to 2016 were resampled separately with two sampling strategies of simple random sampling and size-based stratified random sampling. Sampling locations were randomly resampled without replacement from each year's sampling locations for each scenario. Sampling intensity was determined by the number of shrimp of interest from a sampling location and the percentage of sampling locations in each year.

2.3.2.1 Simple Random Sampling

The sampling scenarios were considered with the percentage of sampling locations and number of shrimp sampled from each sampling location. Two potential sample sizes (i.e., 10 and 20) were considered for a sampling location in the simulation. For sampling locations with less than the required number of shrimp (i.e., 10 or 20), all shrimp in that location were used. For

sampling locations with more than the specified shrimp, 10 or 20 shrimp were randomly sampled without replacement (Fig. 2.1).

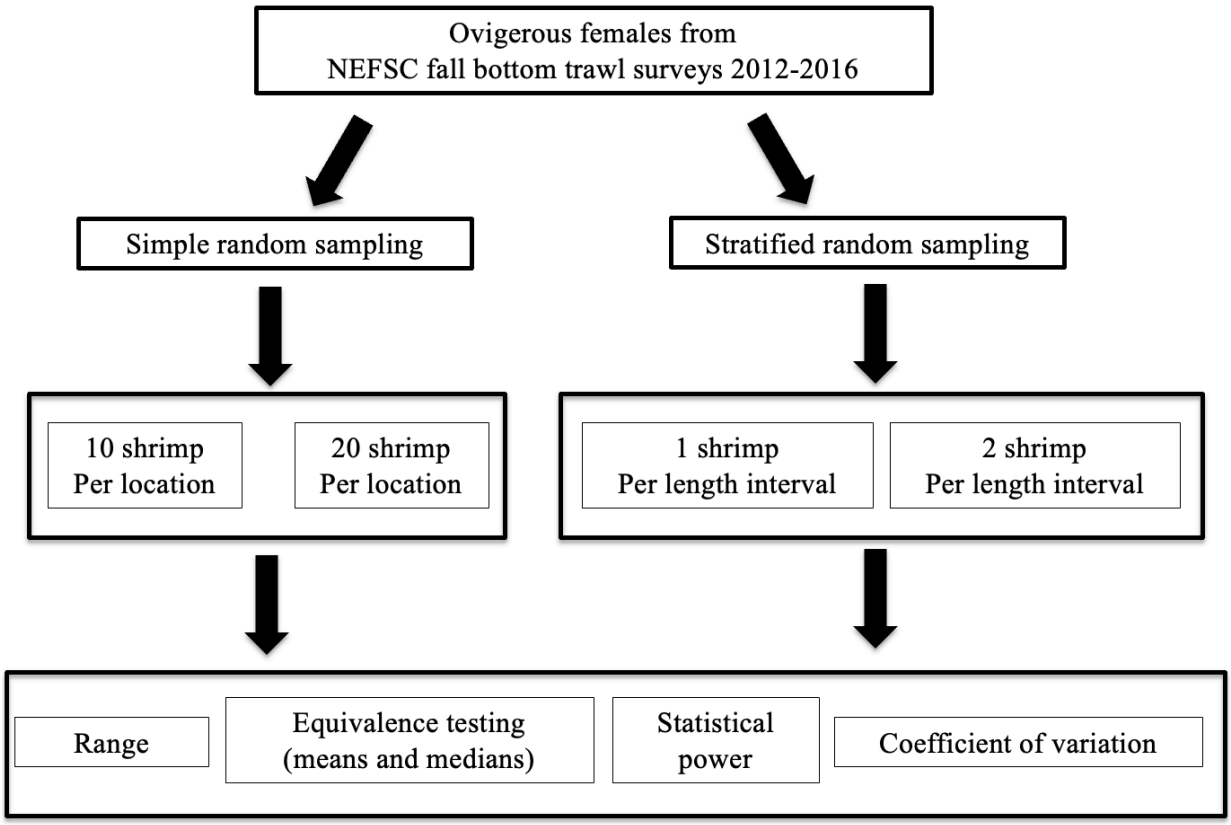


Figure 2.1. A flowchart illustrating the procedure of the simulation analysis. Scenarios of 10 and 20 shrimp per sampling location were considered for simple random sampling at different sampling intensity (percentage of sampling locations). Scenarios of 1 and 2 shrimp per 1.5-mm length interval were considered for stratified random sampling at different sampling intensity. Range of simulated dorsal carapace lengths, equivalence testing of means and medians, statistical power, and coefficient of variation were used for examining the simulated samples in each scenario.

2.3.2.2 Stratified Random Sampling

For stratified random sampling, minimal and maximal lengths were determined to be the minimum and maximum DCLs of samples collected in a year with a length interval of 1.5-mm. A given number (1 or 2) of shrimp was sampled from each length interval until no more shrimp were available in that length interval. The sampling scenarios were developed with a different sampling intensity and number of shrimp sampled from each length interval. For sampling locations which had less than 10 shrimp collected, all shrimp in that location were used for 1-shrimp scenarios (20 shrimp for 2-shrimp scenarios, Fig. 2.1).

2.3.2.3 Equivalence testing

Null hypothesis significance testing framework is commonly used in ecology to examine the differences between the two groups (Martinez-Abrain 2008; Beninger et al. 2012). However, it is criticized in some ecological studies for the following reasons: (1) a lack of significance ($P>\alpha$) simply means there is not sufficient evidence to reject the null hypothesis, but it does not mean the null hypothesis is true (Brosi and Biber 2009; Beninger et al. 2012; Lakens 2017); and (2) the statistical power needed to detect a difference is low. Alternatively, two one-sided equivalence tests within a frequentist framework can be used to ascertain effect quality by specifying meaningful effect size based on biological or ecological understanding (Parkhurst 2001; Lakens 2017). Moreover, the lower and upper bounds constructed with a priori specified effect size allow the researchers to evaluate significant differences with reduced type II error defined in traditional hypothesis testing (Parkhurst 2001; Brosi and Biber 2009). Therefore, instead of using traditional null hypothesis testing, we use two one-sided equivalence testing for the simulated data in each scenario.

Before we performed equivalence testing, a difference of 1.5-mm (Δ) was determined as the minimum effect size that we would like to detect. Effect size was defined as the magnitude of the observed difference (Beninger et al. 2012). Our data suggested that mean DCL of ovigerous females was around 25-mm, which is equivalent to an age of 3.5 years based on age-DCL growth curve (ASMFC NSTC 2018) with age 3 being estimated at 23.5-mm and age 4 at 26.5-mm. We thus determined the effect size interval at 1.5-mm, as shrimp in DCLs smaller or larger than 1.5-mm are likely to be at a different age of years. The lower and upper bounds of equivalence intervals for each sample were constructed as (Nakagawa and Cuthill 2007; Lakens 2017):

$$(m_s - m_y) \pm t_{\alpha,df} s_{pooled} \sqrt{\frac{1}{n_s} + \frac{1}{n_y}} \quad (1)$$

$$s_{pooled} = \sqrt{\frac{{s_s}^2(n_s-1)+{s_y}^2(n_y-1)}{n_s+n_y-2}} \quad (2)$$

where m_s = mean (or median) DCL of samples from a given scenario in year y; m_y = mean (or median) DCL of all samples collected in year y; $t_{\alpha,df}$ = t statistic at a significance level of α at degree of freedom at df; α = 0.05, df = $n_s + n_y - 2$; n_s = number of samples of a given scenario; n_y = number of samples collected in year y; s_s = standard deviation of samples from a scenario in year y; and s_y = standard deviation of all the samples collected in year y.

Two one-sided tests were performed to means and medians of samples simulated from each scenario in each year. The null hypothesis is $ei_l \leq \Delta$ and $ei_u \geq \Delta$, and the alternative hypothesis is $-\Delta < equivalence\ interval < \Delta$, where ei_l = lower bound of equivalence interval, ei_u = upper bound of equivalence interval. Both components in the stated null

hypothesis must be false to reject the null hypothesis. Thus, if the equivalence interval falls within the equivalence interval, the difference between the means or medians is smaller than the magnitude of effect size we specified.

Statistical power of detecting the specified effect size ($\Delta = 1.5$-mm) was estimated with the number of samples simulated in each scenario at the significance level of 0.05. Statistical power of 0.95 was set as a reference instead of traditional 0.8, as we assume the cost of committing a type II error was the same as that of committing a type I error (Peterman 1990; Di Stefano 2003). Coefficient of variation (CV) was also calculated for evaluating the dispersion of samples for each simulation scenario. All analyses were performed in R 3.5.1 (R Core Team 2018).

2.4 Results

2.4.1 Number of sampling locations in each year

The total yearly number of sampling locations and total number of ovigerous females collected in each year from 2012 to 2016 were shown in Table 1. Our data showed that the mean DCL of ovigerous females varied between 24.08 and 25.86 from 2012 to 2016 (Fig. 2.2). In addition, samples collected in 2014 deviated from normal distribution with a mean at 25.43-mm-DCL and a median of 26.5-mm-DCL, and with an unusual wide standard deviation (SD) of 2.89-mm (SD varied from 1.52 to 1.66 in the other four years).

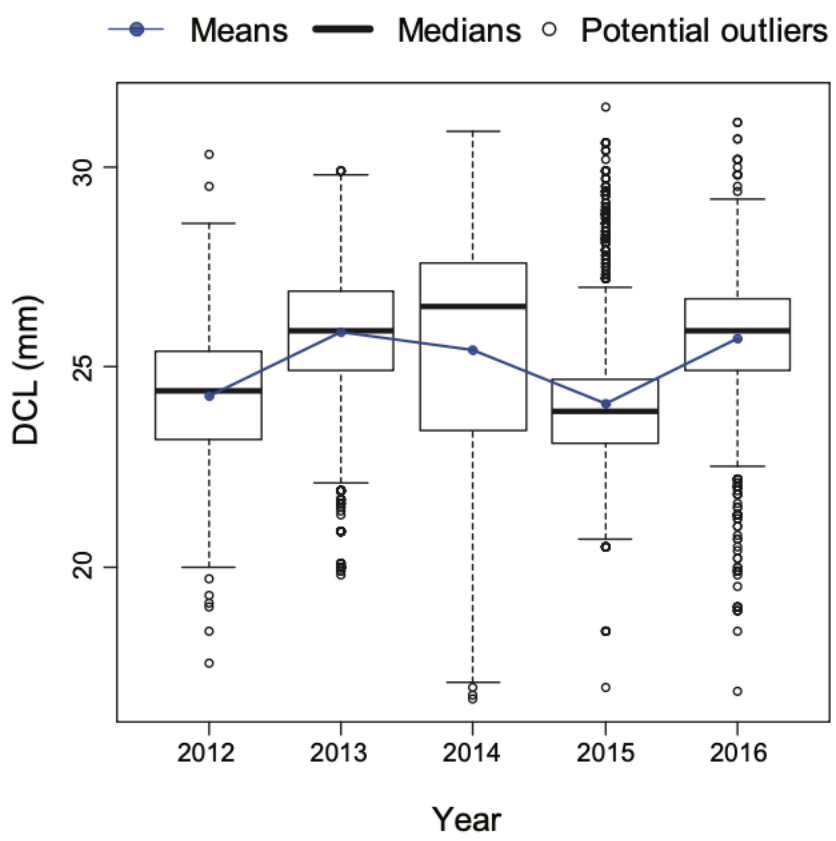


Figure 2.2. Boxplots of dorsal carapace length (DCL) of ovigerous female northern shrimp *Pandalus borealis*, collected from Northeast Fisheries Science Center (NEFSC) fall bottom trawl surveys from 2012 to 2016. The blue symbols are means and the horizontal bars in the boxes are medians. The lower and upper limits of the boxes are the first (Q1) and third (Q3) quartiles (25th and 75 percentiles). The difference between Q1 and Q3 is interquartile range (IQR). Potential outliers are defined as observation points falling outside the range of Q1-1.5*IQR and Q3+1.5*IQR. If potential outliers are presented, the whiskers extend to 1.5 times the IQR from Q1 or Q3. If no outliers are presented, the whiskers extend to the minima and maxima of the distributions.

2.4.2 Equivalence tests

The equivalence tests of means for all the scenarios showed that most equivalence intervals of means fell within the specified effect size interval when at least 20% of the sampling locations were sampled except for 2014 (Fig. 2.3). Similar results could be found in tests for the difference in medians (Fig. 2.4). The equivalence interval of medians barely fell within the effect size interval for simulated samples in 2014 even if all stations were sampled.

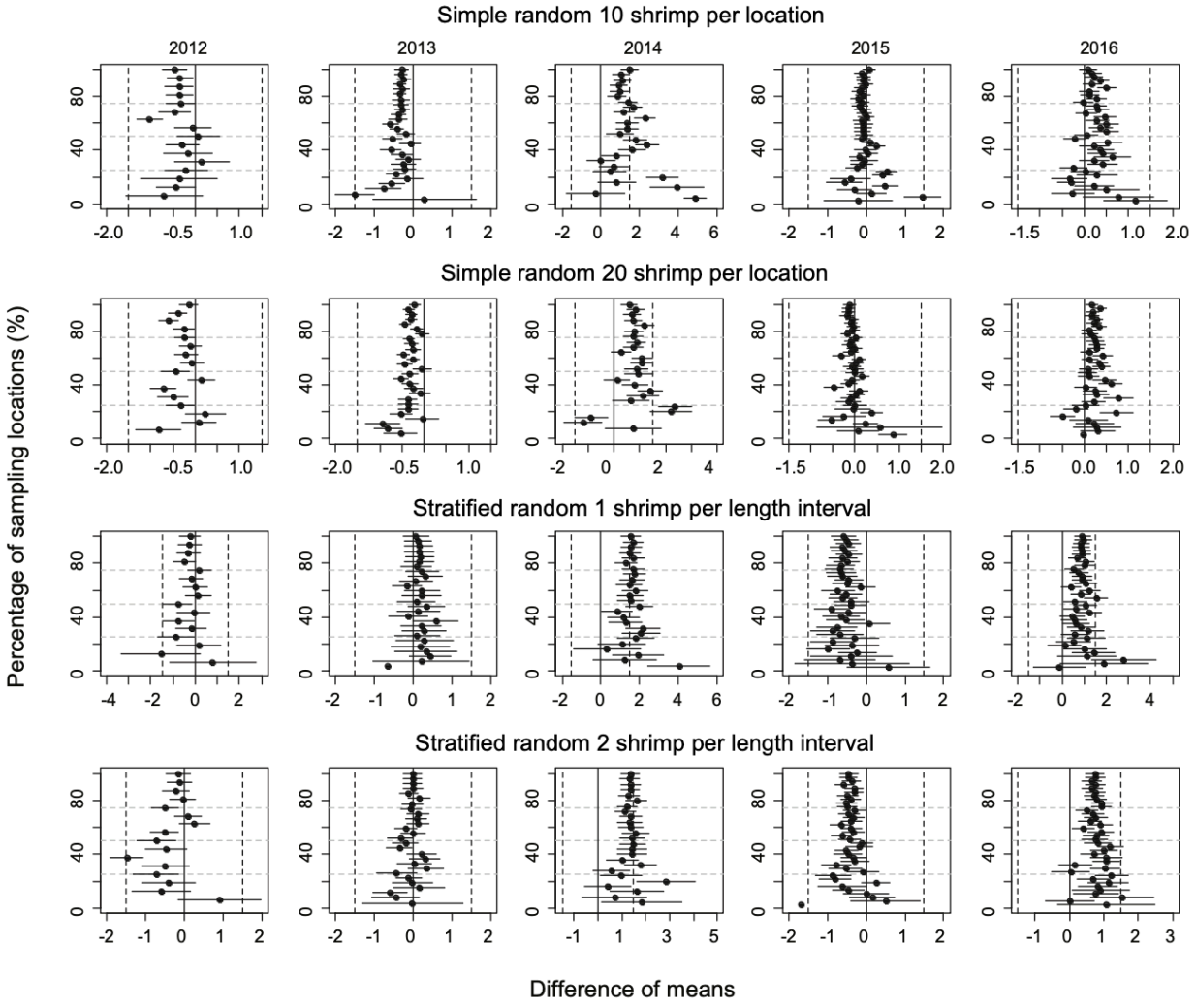


Figure 2.3. Differences between means of samples in each scenario and the population (all shrimp collected in a given year) and 90% confidence intervals (dashed lines) with equivalence bounds (-1.5 and 1.5) for each scenario at percentage of sampling locations for each year.

Vertical solid lines denote mean differences at zero. Gray dashed lines are y-axis grid lines, denoting 25, 50, and 75% of sampling locations.

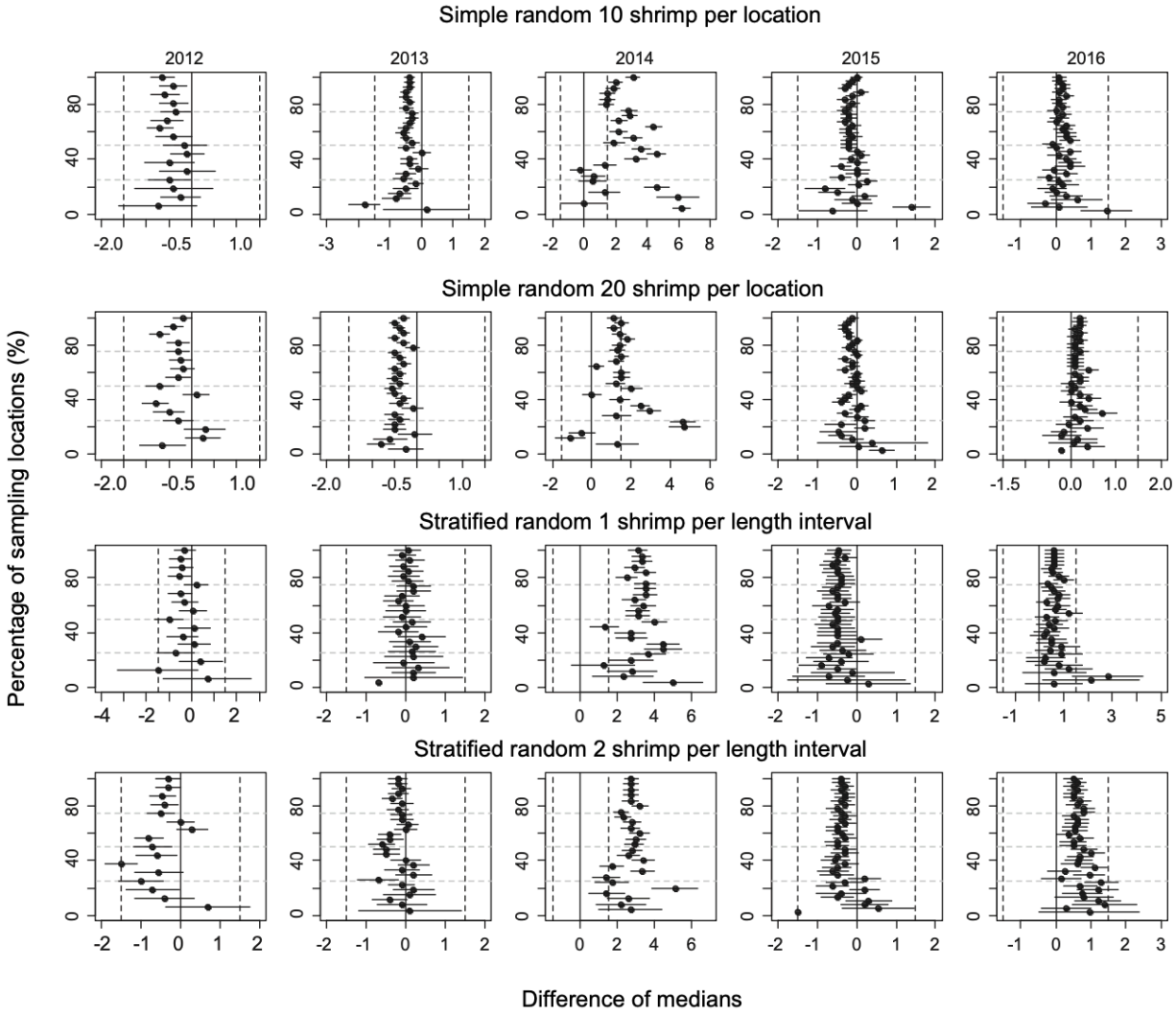


Figure 2.4. Differences between medians of samples in each scenario and the population (all shrimp collected in a given year) and 90% confidence intervals (dashed lines) with equivalence bounds (-1.5 and 1.5) for each scenario at percentage of sampling locations for each year. Vertical solid lines denote mean differences at zero. Gray dashed lines are y-axis grid lines, denoting 25, 50, and 75% of sampling locations.

For means of 20-shrimp scenarios in 2014, the equivalence intervals started to fall within the specified effect size interval when more than 50% of the sampling locations were sampled. When less than 50% of the locations were sampled in 2014, both sampling strategies failed to reject the null hypothesis. However, the differences in means of simple random sampling had a wider variation than those of stratified random sampling scenarios (Fig. 2.3).

As for the equivalence tests of medians for 2014 samples, almost all scenarios failed to reject the null hypothesis (Fig. 2.4). Similar to the equivalence tests of means, when less than 50% of the locations were sampled, the median differences for random sampling method tended to have larger variations than those of stratified random sampling.

2.4.3 Statistical power

The statistical power of detecting the minimal effect size (Δ =1.5-mm) increased with sampling intensity, when more than 20% of sampling locations were sampled, all scenarios could reach the statistical power of 0.95 except for scenarios of 2014 (Fig. 2.5). Simulated samples of 2014 could reach the statistical power of 0.95 when at least 30% of the locations were sampled. There was a trade-off between the number of shrimp per location (or length interval) and percentage of sampling locations. Given a sampling strategy, more numbers of shrimp per sampling location (or length interval) could reach the statistical power of 0.95 with a lower percentage of sampling locations. The coefficients of variation were mostly below 0.1 for each scenario except scenarios in 2014 due to large standard deviation of DCL collected in 2014 (Fig. 2.5).

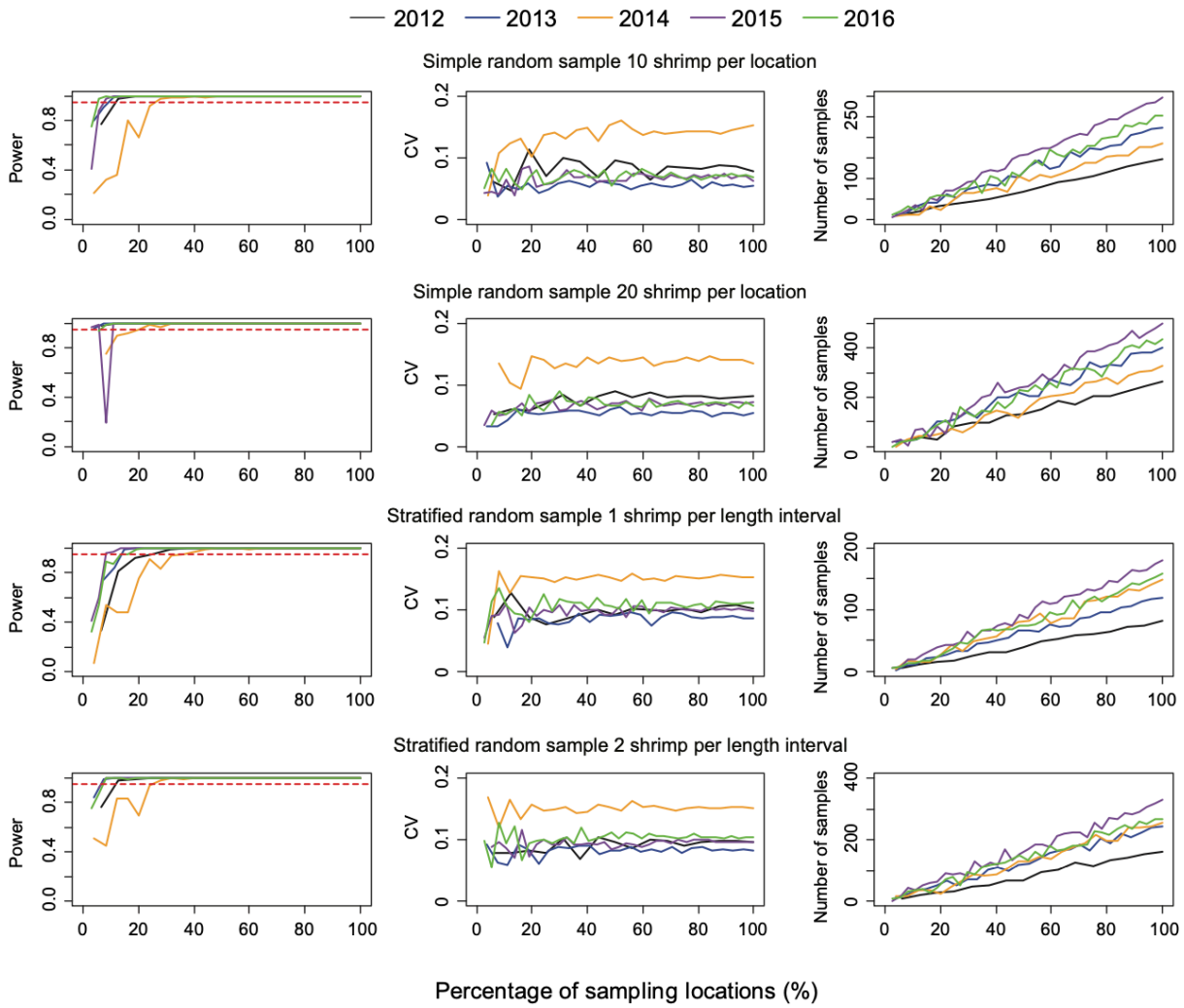


Figure 2.5. Relationships between statistical power (Power), coefficient of variation (CV), number of samples, and percentage of sampling locations for each scenario. Left column: statistical power; middle column: coefficient of variation; right column: number of samples. Top 2 rows: 10 and 20 shrimp per sampling location for simple random sampling; bottom 2 rows: 1 and 2 shrimp per length interval for stratified random sampling.

2.4.4 Sample size

The numbers of shrimp simulated in each scenario increased with sampling intensity, and simple random sampling strategy tended to generate larger sample sizes than stratified random

sampling strategy at a given sampling intensity (Figs. 2.5 and 2.6). When 20% of sampling locations were sampled, the total numbers of shrimp in the simulation for five years ranged from 129 to 349 for different strategies with different intensity (Fig. 2.6). When 30% of the locations were sampled, the total numbers of shrimp increased to 215-612 (Fig. 2.6).

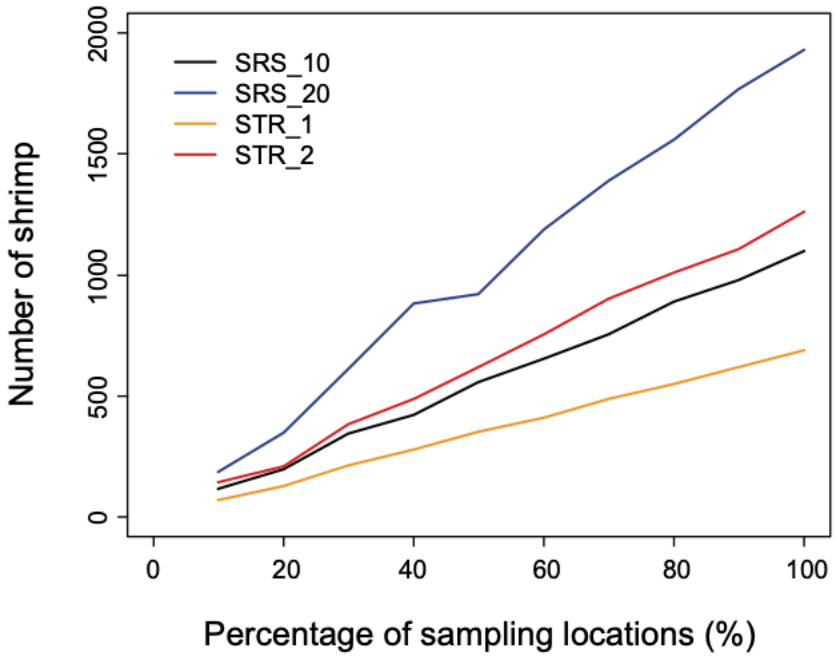


Figure 2.6. Relationships between total number of samples and percentage of sampling locations. SRS_10 and SRS_20 are scenarios of 10 and 20 shrimp per sampling location for simple random sampling. STR_1 and STR_2 are scenarios of 1 and 2 shrimp per length interval for stratified random sampling.

The means, medians, and ranges of samples simulated in each scenario were compared with the assumed populations (samples collected from the surveys) in each year (Fig. 2.7). When more than 20% of the locations were sampled, the simulated samples could include the central 95% of DCL of the assumed population for both sampling strategies. When less than 50% of the

location were sampled, the stratified random sampling, as expected, was more likely to include the minimum and maximum of DCLs of the assumed population than the simple random sampling.

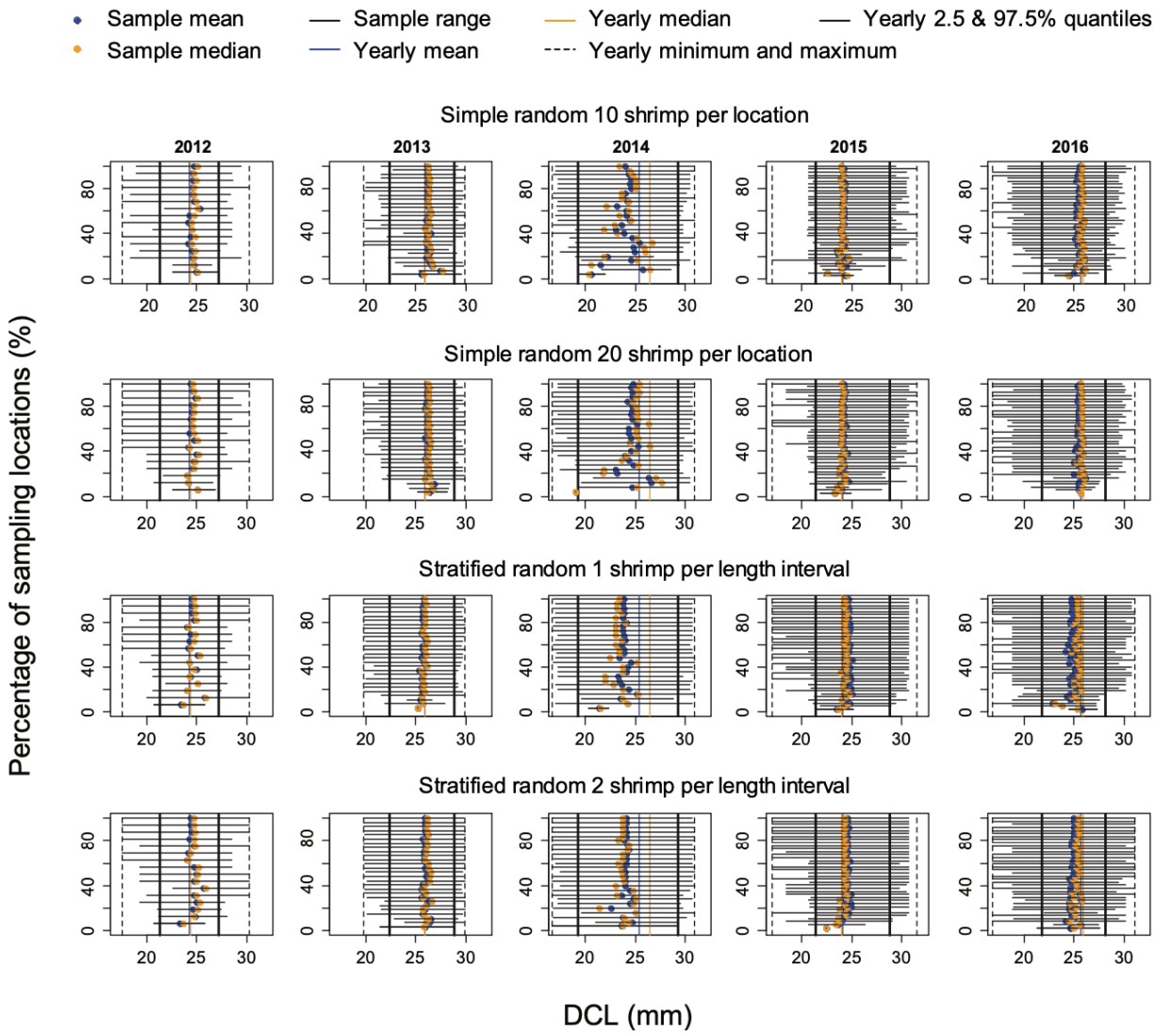


Figure 2.7. A summary of the ranges, means, medians, and the central 95% intervals of dorsal carapace lengths (DCLs) for the assumed populations (all samples collected from the surveys in a year) and samples simulated in each scenario.

2.5 Discussion

The results of equivalence testing showed that there were no large differences between samples simulated with simple random sampling and stratified random sampling strategies when the population distribution is approximately normal. Both sampling strategies can collect samples that are representative of the population (i.e., including the central 95% of the distribution) and the means and medians did not significantly differ from the specified effect size when more than 20% sampling locations were sampled. However, if we conducted traditional null hypothesis significance testing, many of the simulated samples would suggest statistical significance as the confidence interval of error did not include zero, which might not be biologically significant. The results suggested the merits of equivalence testing over traditional null hypothesis significance testing with the ability to detect a biologically meaningful or ecologically important effect size (Parkhurst 2001; Brosi and Biber 2009).

The number of shrimp simulated for each scenario with different strategies, in general, linearly increased with the number of sampling locations. However, as the surveys were not specifically designed for northern shrimp, number of shrimp collected at a station could be only a few. Therefore, the ultimate sampling intensity (number of shrimp simulated for a scenario) was not exactly proportional to the number of locations sampled. An extreme example was the 20-shrimp scenario with three sampling stations with simple random sampling strategy, which had only four DCLs simulated in that scenario. The statistical power was hence low (Fig. 2.5). Our simulation reflects the discrepancy between samples collected in multispecies surveys and ideal sampling for fecundity data. Care should be taken to adjust sampling strategy in such circumstances.

Increasing sampling intensity by either raising the number of shrimp per location, length interval, or the number of sampling locations can reduce sampling error and increase statistical power. However, the cost of increasing sampling intensity may not be effective as the magnitude of precision that can be improved is trivial when sampling intensity is above a certain level (Pennington et al. 2002). Although both the sampling strategies we adopted in this study suggested that the equivalence interval can fall within the effect size interval when at least 20% of the locations were sampled (except for 2014), we determined stratified random sampling may be a more effective sampling strategy for collecting fecundity data as it pointed to a lower sample size compared to the simple random sampling.

With stratified random sampling at a fixed overall sampling size (number of shrimp simulated for all five years), based on the trade-off between the number of shrimp per length interval and the percentage of the locations, a desired statistical power can be achieved at a lower percentage of sampling locations for 2-shrimp per length interval scenarios. However, the stratified random sampling strategy with one shrimp per length interval is preferred in this case, as a higher percentage of sampling locations allows a broader spatial coverage of the study area. Therefore, the optimal sample size for collecting fecundity data was estimated at 215 shrimp for five years (30% of the locations) with size-based stratified random sampling.

Both sampling strategies generated unrepresentative samples which were significantly different from the specified effect size when less than 50% of the locations were sampled for 2014 due to the skewed distribution of DCLs in 2014. Generally, it is not possible to know the length distribution of the population which is usually assumed to be approximately normally or log-normally distributed. Caution is warranted when many small spawners are observed in the population, which could be a sign of early sexual maturity resulting from fishing pressure,

environmental changes and consequent food availability to females (O'Brien 1999; Koeller et al. 2007). Spawners at small sizes make less contribution per individual to reproductive potential of a population, as small spawners tend to produce fewer offspring per individual with lower survival rates (Shelton et al. 2012; Barneche et al. 2018).

Aanes and Volstad (2015) used a simulation approach to evaluate subsampling strategies for collecting age data for Northeast Arctic cod (*Gadus morhua*), suggesting that length-stratified sampling is more effective than simple random sampling because length-stratified sampling can ensure a better coverage of the age composition when age data were collected from a small subsample of measured lengths of fish. Our findings agree with Aanes and Volstad (2015). For the purpose of collecting fecundity data, stratified random sampling strategy is preferred over simple random sampling when the size distribution of ovigerous females is actually skewed with many small spawners (deviating from the assumed normally distributed population). Because it is often not possible to have enough resources for a high sampling intensity, and simple random sampling is more likely to generate a biased sample in a low sampling intensity (Figs. 2.3, 2.4, and 2.7). Conversely, although stratified random sampling also generates biased samples, the variation of means and medians of samples are relatively stable when sampling intensity is low. Furthermore, laboratory process for collecting fecundity data can be very time-consuming and labor-intensive. The time needed for processing a shrimp to collect fecundity data is generally 3-4 hours. Given a sampling intensity of 20% of the sampling location, the 10-shrimp simple random sampling scenario generates a larger number of sample size than the 1-shrimp per length interval stratified random sampling scenario by 69 shrimp. Thus, the simple random sampling may take 207 additional hours (69 shrimp × 3 hours), which would cost additional $4140 (i.e.,

207 hours × $20 per hour per person) for laboratory process alone. Our analyses suggest that length-stratified random sampling is a more cost-effective strategy for collecting fecundity data.

The shrimp samples Haynes and Wigley (1969) used for collecting fecundity data ranged from 22 to 31-mm-DCL. Except for 2014, the central 95% of ovigerous females collected from the survey ranged from a similar interval of 22-28-mm-DCL in this study. However, it appeared that if shrimp outside the central 95% length interval were excluded from the regression of length and fecundity, the regressed relationship may not be able to provide reliable estimates of fecundity for the population as the fecundity-DCL relationship developed with 47 female shrimp by Haynes and Wigley (1969) generates negative numbers for shrimp at DCLs<20-mm. It suggested that, when estimating size-based fecundity for a population, (1) a complete range of size data is necessary for developing a fecundity-body size relationship; (2) several years of samples may be needed for building a complete fecundity database; and (3) parabola equation should be used with caution as it may generate biologically meaningless estimates of fecundity (negative values). Estimating the magnitude of the bias in reproductive potential of a population is beyond the scope of this study. Consequently, before we take a further step into investigation of the misestimates of fecundity, there is a pressing need to develop a new fecundity-DCL relationship with proper sampling design for collecting fecundity data.

This study proposes a simulation framework that can be used to develop a cost-effective sampling strategy for estimating fecundity data for many marine fish and crustacean species which share the characteristics of (1) a strong maternal effect on fecundity (i.e., number of offspring increase with female body sizes; Haynes and Wiley, 1969); (2) number of individuals collected varied among sampling locations and number of sampling locations varied by year; and (3) extensive length frequency data have been collected for multiple years which can be used for

sampling design. Collecting fecundity data can be very time-consuming and labor-intensive. Insufficient samples may result in biased estimates; however, excess samples can be a waste of resources. Therefore, an appropriate sampling design for optimizing effective sample size is needed for building a complete fecundity database. We advocate the use of equivalence testing and power analysis before collecting samples in order to determine biologically meaningful effect size instead of statistical significance in traditional null hypothesis significance testing.

CHAPTER 3. POSSIBLE CLIMATE-INDUCED ENVIRONMENTAL IMPACTS ON PARASITE-INFECTION RATES OF NORTHERN SHRIMP *PANDALUS BOREALIS* EGGS IN THE GULF OF MAINE

3.1 Abstract

The Gulf of Maine northern shrimp *Pandalus borealis* population once supported a significant commercial winter fishery for the New England states. However, the fishery has been on moratorium since 2014 due to consecutive recruitment failures. The issue of parasite-infected eggs, so-called 'white eggs,' has long been identified for the Gulf of Maine northern shrimp, which makes shrimp eggs nonviable and subsequently hampers the recruitment potential. Furthermore, the proportion of infected females was observed to increase with water temperature. As Gulf of Maine temperatures have been increasing for decades, it is important to re-visit issues related to white eggs to evaluate possible impacts of climate-induced environmental changes on the white egg infection rates. We used biological samples collected by the Northeast Fisheries Science Center in 2012−2016 to evaluate the probability that a female shrimp was infected (P_{inf}) and the proportion of white eggs in an infected female shrimp (p_{we}). Although P_{inf} was high, with an average of 73.81% over the Gulf of Maine, p_{we} was mostly <5%. The variation in both P_{inf} and p_{we} examined in this study was not well explained by environmental factors or female body size. However, the average rates of both P_{inf} and p_{we} observed in this study were higher than those observed in the 1960s when the bottom temperatures were cooler. The results can be used to account for egg mortality and provide information on potential impacts of possible climate-induced variability on shrimp population dynamics.

3.2 Introduction

The Gulf of Maine (GOM) has been experiencing intense warming over the past 15 year at a higher rate than 99% of the global oceans (Pershing et al. 2015). The sea surface temperature has increased at an average rate of 0.26°C $year^{-1}$ since 2004 (Mills et al. 2013). Such rapid changes in temperature have impacts on suitable thermal habitats (Kleisner et al. 2017), life cycles (Richards et al. 2012), and fisheries (Pershing et al. 2015) of various species.

GOM northern shrimp *Pandalus borealis* once supported a significant winter fishery for the New England states (Clark et al. 2000). However, a moratorium has been in place since 2014 due to recruitment failures for several consecutive years (ASMFC NSTC 2018). The recruitment failures have been correlated with warming water temperatures in the GOM for the past decade (Richards et al. 2012). GOM northern shrimp are considered sensitive and vulnerable to climate-induced environmental variabilities as they are distributed at the southern end of their distribution in the North Atlantic Ocean (Richards et al. 2012; ASMFC NSTC 2018). However, the mechanism of how recruitment in this species is affected by the environment has not been identified.

Fecundity is a key characteristic that can be used for estimating the reproductive potential of a population (Hannah et al. 1995). Fecundity is affected by many biotic and abiotic factors, including maternal body size, food availability to females, population density, environmental conditions, and diseases (Haynes and Wigley 1969; Apollonio et al. 1986; Hannah et al. 1995).

Non-viable eggs (‘white eggs’) have been a known issue for GOM northern shrimp since the 1960s (Haynes and Wigley 1969; Elliot 1970; Stickney 1981a; Apollonio et al. 1986; Parsons and Tucker 1986). Stickney (1978) examined white eggs and inferred that they are infected by a group of dinoflagellates which eventually cause egg death. The parasitic organisms enter the

eggs and develop into plasmodia inside the eggs (Shields 1994). The yolk materials are depleted and replaced by the plasmodia of the parasite after a week from the start of infection (Stickney 1978). The egg ruptures by the end of the second week, and the flagellated cells are released at 8°C (Stickney 1978). Meyers et al. (1994) suggested that the dinoflagellate-like parasite that occurred in shrimp eggs was a *Peridinium* sp.; however, the exact species of parasite infecting northern shrimp eggs has yet to be conclusively identified.

Parasite incidences were also perceived to be correlated with water temperature (Stickney 1981b; Apollonio et al. 1986). The proportion of infected female shrimp was observed to increase with water temperature (Apollonio et al. 1986). Stickney (1981b) suggested that the annual variation in fecundity could be related to environmental variability and/or egg disease, indicating that the influence of temperature on egg development may have an effect on recruitment. However, the effect of parasitism on egg mortality has not been quantified, and the relationship between infection and temperature remains unclear.

Northern shrimp are assumed to have a high natural mortality at early life stages (ASMFC NSTC 2018), but egg mortality is rarely addressed. We hypothesized that an increased white egg infection rate, potentially a result of climate-induced changes in thermal habitat, can lead to increased egg mortality and subsequent recruitment failure, which has been identified as a main cause for the collapse of the GOM northern shrimp fishery. Given large changes in the GOM thermal habitat (Pershing et al. 2015), this hypothesis needs to be carefully considered and tested to identify key drivers resulting in the collapse of the GOM northern shrimp fishery. This study aimed to test the above hypothesis by examining the effects of biotic and abiotic factors on parasitic infection incidence. Such an investigation can enhance our understanding of egg mortality and provide important information about the impact of parasitism on reproductive

potential and subsequently better estimate recruitment of the northern shrimp population in the GOM.

3.3 Materials and Methods

Biological samples were collected by the National Oceanic and Atmospheric Administration (NOAA) ship 'Albatross IV' during the Northeast Fisheries Science Center (NEFSC) fall bottom trawl surveys in the GOM in 2012−2016 (Fig. 3.1). The sampling locations were randomly determined from each stratum, and the number of stations within a stratum was generally proportional to the area of the stratum and the overall variation of multispecies distribution among strata (Smith 2002). Details of the survey sampling design and vessel configuration were documented by Stauffer (2004) and Politis et al. (2014). The spawning season of GOM northern shrimp starts in late summer (August−September). Biological samples were collected in October and November, when most females have extruded and attached eggs to their pleopods. Fall is also the time of year when most infected eggs were observed (Haynes and Wigley 1969; Elliot 1970; Apollonio et al. 1986). Only egg-bearing females were used in this study.

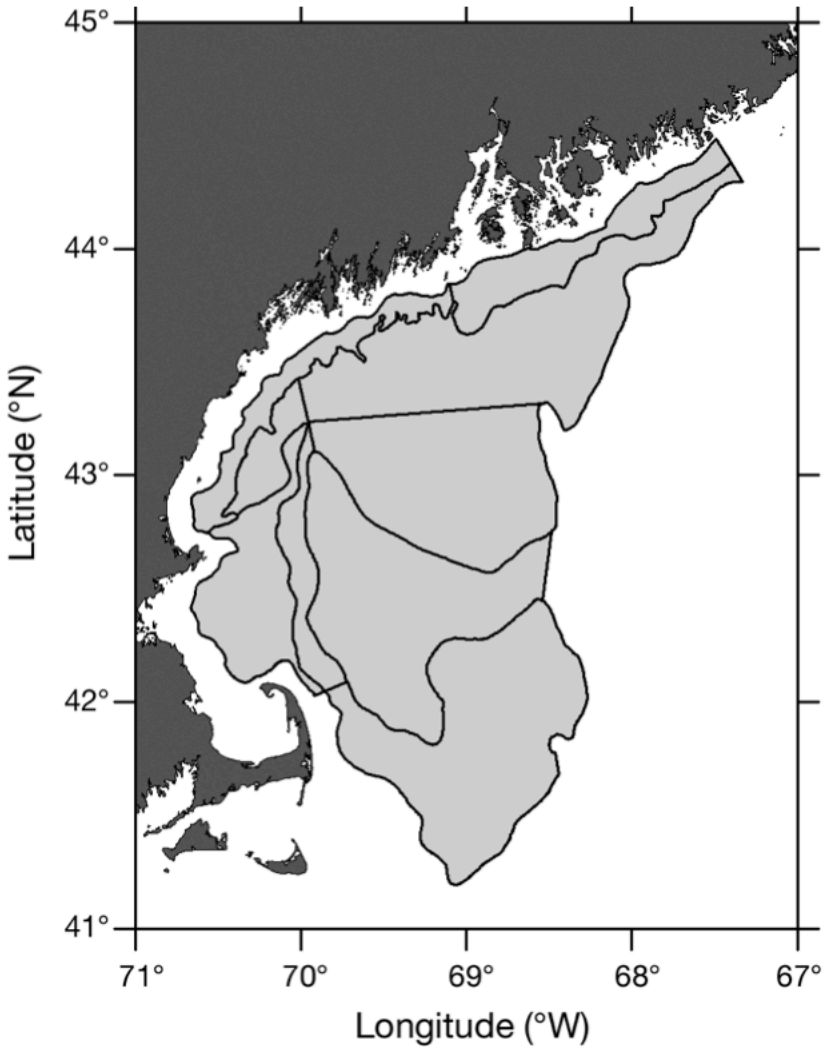


Figure 3.1. Sampling area of the Northeast Fisheries Science Center bottom trawl surveys in fall (black outlines indicate different strata) in the Gulf of Maine during 2012−2016

Abiotic data including bottom temperature, bottom salinity, and latitude and longitude of the sampling location were collected by the NEFSC fall bottom trawl surveys for each tow at each sampling location. Bottom temperature and salinity data were measured with electronic profiling conductivity/temperature/ depth instruments (Smith 2002).

Biological data including dorsal carapace length (DCL, mm; Table 1) and numbers of both normal and white eggs were collected. Ten ovigerous females of different sizes were collected from each sampling location with a stratified random sampling design. Length intervals with the range from minimum to maximum DCL of a bag (samples from a sampling location) were divided into 10 length intervals. If there were more than 10 shrimp collected in a sampling location, 1 shrimp was randomly chosen from each size class available at that location. Although a pre-planned sampling design was followed, more than half of the biological samples of eggs rotted due to improper preservation. Therefore, biological samples were taken from all available sampling locations and rotten samples were excluded from the analysis.

Table 3.1. List of abbreviation used in this chapter.

	Abbr.	Definition
Data	P_{we}	Proportion of white eggs of an egg mass
	P_{inf}	Possibility of a female is infected
	DCL	Dorsal carapace length
	BOTTEMP	Bottom temperature
	BOTSAL	Bottom salinity
Statistics	ε_y	Random effect of year
	ε_l	Random effect of sampling location
	Marginal R^2_{GLMM}	Proportion of variance explained by the fixed effects in the model
	Conditional R^2_{GLMM}	Proportion of variance explained by both fixed and random effects
	ΔAIC	Difference in AICs between the best model and the competing model
	LOOCV	Leave-one-out cross-validation
	RMSE	Root mean squared error
	VIF	Variance inflation factor

Eggs were removed from the body of each sampled shrimp. As white eggs can be easily distinguished from normal eggs by eye (Fig. 3.2), each egg mass was examined for the presence of white eggs. An individual was considered infected if it carried at least 1 white egg. All

removed eggs were preserved in 10% neutral formalin (Parsons and Tucker 1986), and then later imaged under a trinocular dissecting microscope (OMAX V434B- L54P-C140U) mounted with a digital camera (OMAX A35140U3). White eggs are easily recognized due to their irregular swollen shape and opaque white coloration (Fig. 3.2; Haynes and Wigley 1969). The preservation of formalin did not lead to misidentification of egg infection status. Numbers of normal eggs (noninfected) and white eggs (infected) were counted with assistance from ObjectJ (Schneider et al. 2012). The proportion of white eggs for each individual was the number of white eggs divided by the total number of normal and white eggs.

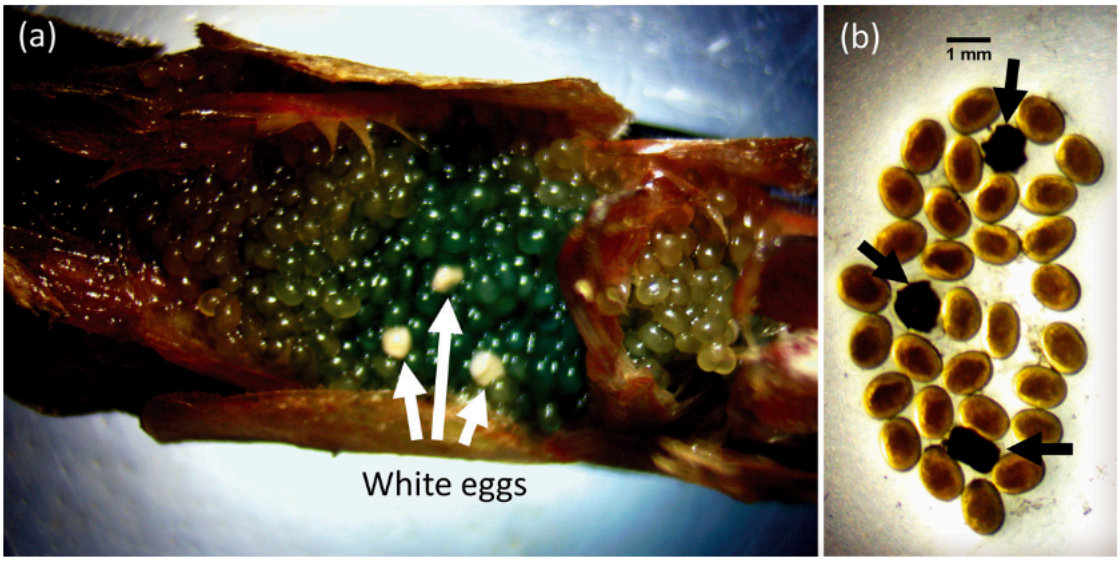


Figure 3.2. Eggs of northern shrimp. (a) Fresh eggs attached to a female's abdomen. (b) Eggs preserved in 10% neutral formalin after 1 wk. The arrows indicate parasite-infected eggs (so-called ‘white eggs’)

Before fitting a model to the data, a variance inflation factor (VIF) analysis was conducted to identify multicollinearity. A threshold VIF of 3 was set for evaluating possible collinearity between predictors in the data set (Schmiing et al. 2013; Brosset et al. 2019). All predictors were centered at means in order to have biologically meaningful intercepts when fitting models to the data.

Binomial generalized linear models (Bolker 2008; Zuur et al. 2009) were used to model the probability of an individual being infected (P_{inf}), and beta generalized linear models were used to model the proportion of white eggs (p_{we}) through a logit link function (Bolker 2008):

$$logit(y) \sim \alpha + \beta_1 x_{1,i} + \beta_2 x_{2,i} + \cdots + \beta_k x_{k,i} + \varepsilon_y + \varepsilon_l + \varepsilon_i$$

where y is P_{inf} or P_{we}, x_k is the k^{th} explanatory variable, α and β's are parameters, ε_y and ε_l are random effects of year and sampling location, and ε_i is residual errors. Year and sampling location are included in the model as random effects because multiple samples were collected from the same location or the same year which might cause pseudoreplications (Hurlbert 1984; Weltz et al. 2013). Including these nested random effects can account for variations among years and sampling locations (Bolker et al. 2009; Thorsen and Minto 2015). Only 3 out of 73 locations were resampled in different years. Therefore, crossed random effects were not considered in the models.

Models with different combinations of predictors were built and compared. Predictors were excluded from a model if the effect sizes (estimated regression coefficients, i.e. the magnitude of an effect) were not significantly different from zero, and then the data were refitted. The p-values from the Wald tests for significance testing tend to be anti-conservative (Luke 2017). Therefore, when the p-values from the Wald tests were <0.05, parametric bootstrap

methods were used to obtain p-values from likelihood ratio test statistics (Luke 2017) to evaluate the significance.

Model selection was based on graphical inspection, Akaike's information criterion (AIC, Akaike 1973), and marginal and conditional pseudo R^2 for generalized linear mixed models (R^2_{GLMM}) developed by Nakagawa et al. (2017) for evaluating the goodness of fit of each model. The AIC provides a measure of model fitting with model complexity accounted for (Akaike 1973). The marginal R^2_{GLMM} represents the proportion of variance explained by the fixed effects in the model, and the conditional R^2_{GLMM} represents the proportion of variance explained by both fixed and random effects (Nakagawa et al. 2017; Bartoń 2019). Predictive performance of models was evaluated by root mean squared error (RMSE) estimated with leave-one-out cross-validation (Zuur et al. 2009; Arlot and Celisse 2010).

The fixed effects were first assessed when all random effects were included in the models (Zuur et al. 2009; Barr et al. 2013). The model with only random effects was therefore the null model. After the variable selection was determined, random effect was dropped from the model if the variance was estimated to be zero (Pasch et al. 2013). The model with the lowest AIC and RMSE and highest R^2 was selected as the optimal model (Zuur et al. 2009; Mac Nally et al. 2018).

Residual diagnosis was graphically examined through histograms and plots of Pearson residual errors from the best model against fitted values for evaluating the distribution and deviation of residual errors (Zuur et al. 2009). Semivariance of Pearson residual errors from the best model was examined for the presence of spatial autocorrelation (Cressie 1993; Pebesma 2004; Gräler et al. 2016).

Three potential outliers were removed (see Section 3.4) from the dataset. Models were refitted to the new dataset with outliers removed. The procedures for the model selection described above were followed. All analyses were performed with R version 3.5.1 (R Core Team 2018).

The bottom temperature observed in this study was 5.7−11.7°C with a mean of 8.1°C. However, the bottom temperatures observed in December 1967 to January 1968 and November to December 1968 by Apollonio et al. (1986) were 4.3−8.6°C with a mean of 6°C. The proportion of infected females (P_{inf}) and bottom temperatures collected by Apollonio et al. (1986) were then incorporated with data collected in this study. P_{inf} and bottom temperature data were aggregated by sampling locations in order to be incorporated with data of Apollonio et al. (1986). To examine the effect of bottom temperature on P_{inf}, we fitted a model to the combined data through a logit link function with an assumed binomial error distribution.

3.4 Results

3.4.1. Probability of an individual being infected

A total of 565 females collected from 73 locations in 2012−2016 were examined for infection. The overall infection rate was 73.81%, with over 50% of the sampling locations having infection rates higher than 90%. Thus, more than 90% of the sampled females had at least 1 white egg in more than half of the sampling locations.

All explanatory variables (bottom temperature, salinity, and DCL) had VIFs <3. Collinearity between variables was thus not considered an issue in the models. Including bottom temperature and DCL in the models did not help explain the variation in P_{inf} (Table 3.2), suggesting that bottom temperature and DCL had no significant impacts on the infection rate. P_{inf}

was estimated to be 83.2% at an average salinity and was negatively correlated with bottom salinity, with 1‰ increase in salinity resulting in a 27.93% decrease in P_{inf} (Fig. 3.3). Although the effect of bottom salinity was statistically significant (p = 0.045) based on the parametric bootstrapped likelihood ratio test, the proportion of variation that bottom salinity could explain was rather small, less than 6%.

Table 3.2. Model statistics for the null model and the competing models for factors associated with possibility of a female being infected (see Table 3.1 for abbreviations). *: p<0.05

Model (Binomial)	Marginal R^2_{GLMM}	Conditional R^2_{GLMM}	ΔAIC	LOOCV RMSE
$P_{inf} \sim$ BOTSAL* $+ \varepsilon_l$	0.048	0.548	0.000	0.330
$P_{inf} \sim$ BOTSAL* $+ \varepsilon_y + \varepsilon_l$	0.048	0.548	2.000	0.330
$P_{inf} \sim$ DCL + BOTSAL* $+ \varepsilon_y + \varepsilon_l$	0.052	0.551	2.793	0.331
$P_{inf} \sim + \varepsilon_l$	NA	0.564	3.064	0.329
$P_{inf} \sim$ BOTTEMP + BOTSAL* $+ \varepsilon_y + \varepsilon_l$	0.053	0.550	3.444	0.331
$P_{inf} \sim$ DCL + BOTTEMP + BOTSAL* $+ \varepsilon_y + \varepsilon_l$	0.056	0.553	4.283	0.331
$P_{inf} \sim + \varepsilon_y + \varepsilon_l$	NA	0.564	5.064	0.329
$P_{inf} \sim$ DCL $+ \varepsilon_y + \varepsilon_l$	0.002	0.568	6.233	0.329
$P_{inf} \sim$ BOTTEMP $+ \varepsilon_y + \varepsilon_l$	0.008	0.567	6.394	0.329
$P_{inf} \sim$ DCL + BOTTEMP $+ \varepsilon_y + \varepsilon_l$	0.010	0.571	7.601	0.329

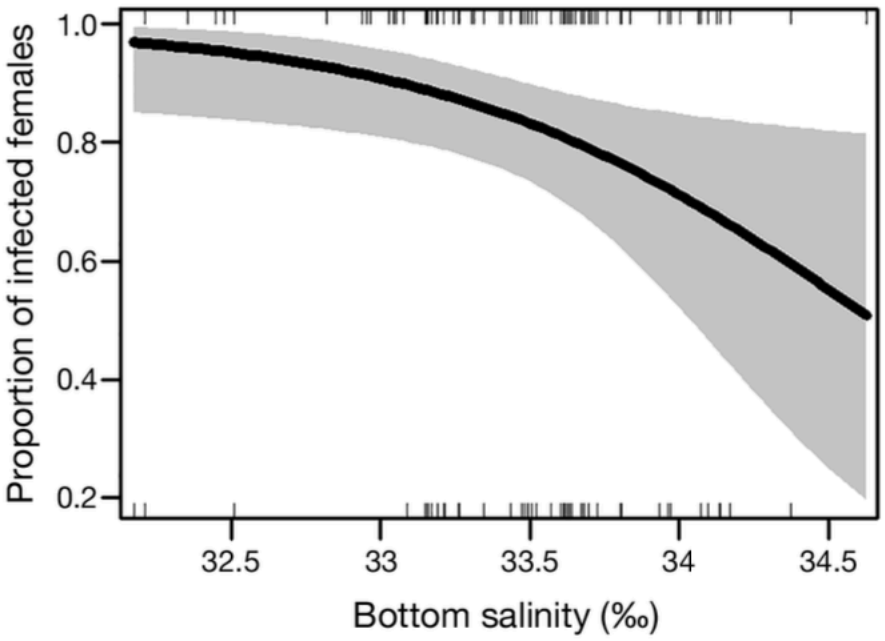


Figure 3.3. Relationship between the proportion of infected female northern shrimp and bottom salinity. Shaded area denotes the 95% confidence interval. Vertical bars at the top and bottom are positive and negative residuals

More than 50% of the variation can be explained by the random effect of sampling locations (Table 3.2). The differences between the intercept of each sampling location and the overall intercept of the model was approximately normally distributed. This implies that the P_{inf} varies among sampling locations, but none of the variables we used is the main driver. The variance of the random effect of year was estimated to be zero, and it was therefore excluded from the model. The semivariance of model residual errors did not suggest the presence of spatial autocorrelation (Fig. 3.4). All models had similar predictive performance, with RMSEs ranging from 0.33−0.34 (Table 3.2), suggesting the differences between predictions and estimations ranging from 0.33−0.34.

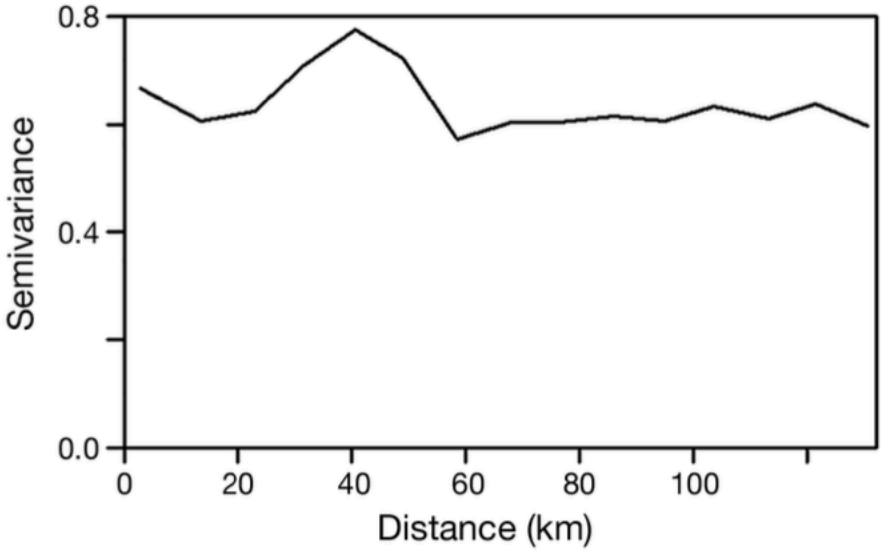


Figure 3.4. Variogram computed with residuals of the generalized linear mixed model with bottom salinity and the random effect of sampling location for the proportion of infected female northern shrimp

A total of 90 observations of bottom temperatures and P_{inf} were combined with data from Apollonio et al. (1986). Bottom temperature was statistically significant ($p < 0.001$) and explained 41% of the variation, indicating that bottom temperature had a significant effect on P_{inf}. Mean P_{inf} at mean bottom temperature (7.7°C) of the data set was 41.86% and increased to 62.56% when bottom temperature increased 1°C from the mean. The bottom temperatures observed by Apollonio et al. (1986) during 1967−1968 were below 9°C, and P_{inf} observed was lower than 50%. The bottom temperatures observed in our study were 5.7−11.7°C with a mean of 8.1°C, and more than 75% of the sampling locations had $P_{inf} > 90\%$ over the observed bottom temperatures with high variations (Fig. 3.5). The model results showed that when bottom temperature was higher than 8°C, at least 50% of females were infected.

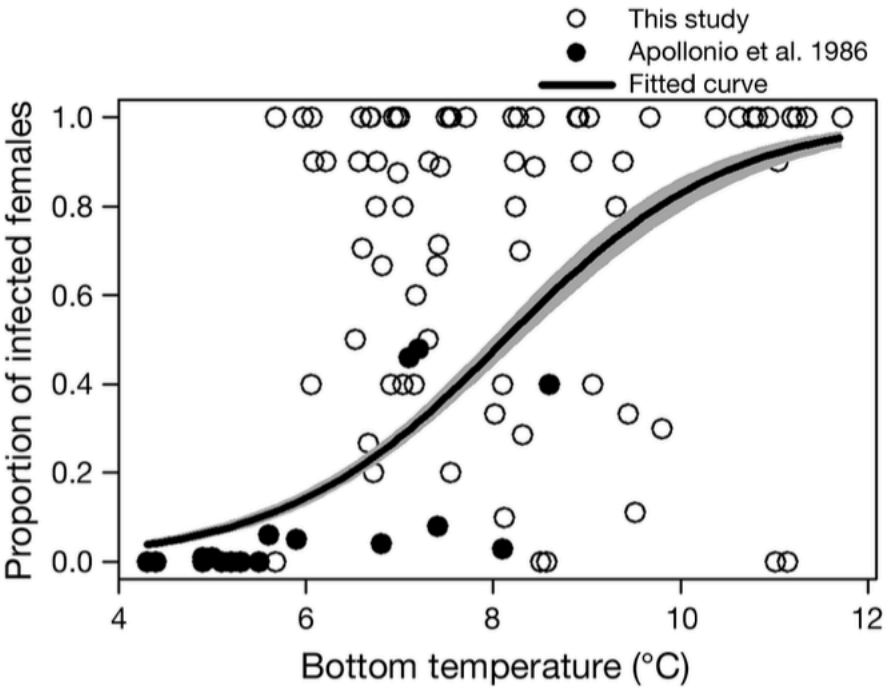


Figure 3.5. Relationship between proportion of infected female shrimp and bottom temperature fitted with combined data from this study and from Apollonio et al. (1986). Gray shading denotes the 95% confidence interval

3.4.2. Proportion of white eggs

A total of 82 infected females were randomly selected from 32 sampling locations during 2012− 2015 and examined for p_{we}. p_{we} ranged from <1 to 33%, with more than 90% of the examined individuals having less than 5% of white eggs. The overall average p_{we} was 2.16%.

The VIFs for all variables were <3. The best model selected by the lowest AIC had bottom salinity and DCL as fixed effects and sampling location as a random effect (Table 3.3). Both bottom salinity and DCL together explained 14.5% of the variation in p_{we}. Adding bottom temperature to the model improved the marginal R^2GLMM by less than 2%, but the model goodness-of-fit (AIC) decreased, indicating overfitting. The parametric bootstrapped likelihood ratio tests suggested statistical significance for the effect size of DCL ($p = 0.01$) but not for

bottom salinity (p = 0.06). The estimated p_{we} at mean DCL (25.22 mm) at mean bottom salinity (33.75‰) was 1.41%. Given the effect sizes of bottom salinity and DCL, p_{we} decreased by 0.83% on average with a 1‰ increase in bottom salinity and by 0.15% with a 1 mm increase in DCL (Fig. 3.6). Three potential outliers of observations with high p_{we} (>10%) were graphically identified (Fig. 3.6). These potential outliers were removed from the data and the models were refitted.

Table 3.3. Model statistics for the null model and the competing models for factors associated with proportion of white eggs of an egg mass (see Table 3.1 for abbreviations).*: p<0.05; **: p<0.01

Models (Beta)	Marginal R^2_{GLMM}	Conditional R^2_{GLMM}	ΔAIC	LOOCV RMSE
P_{we} ~DCL** +BOTSAL* $+ \varepsilon_l$	0.145	0.590	0	0.019
P_{we} ~DCL** +BOTSAL* $+\varepsilon_y + \varepsilon_l$	0.145	0.590	2.000	0.019
P_{we} ~DCL** +BOTTEMP +BOTSAL* $+\varepsilon_y + \varepsilon_l$	0.163	0.573	2.081	0.019
P_{we} ~DCL* $+\varepsilon_y + \varepsilon_l$	0.061	0.602	4.318	0.019
P_{we} ~DCL** +BOTTEMP $+\varepsilon_y + \varepsilon_l$	0.010	0.589	4.337	0.018
P_{we} ~ BOTSAL* $+\varepsilon_y + \varepsilon_l$	0.074	0.613	6.700	0.020
P_{we} ~ $+ \varepsilon_l$	NA	0.629	6.749	0.020
P_{we} ~ BOTTEMP +BOTSAL* $+\varepsilon_y + \varepsilon_l$	0.088	0.596	7.141	0.020
P_{we} ~ $+\varepsilon_y + \varepsilon_l$	NA	0.629	8.749	0.020
P_{we} ~ BOTTEMP $+\varepsilon_y + \varepsilon_l$	0.030	0.617	9.194	0.020

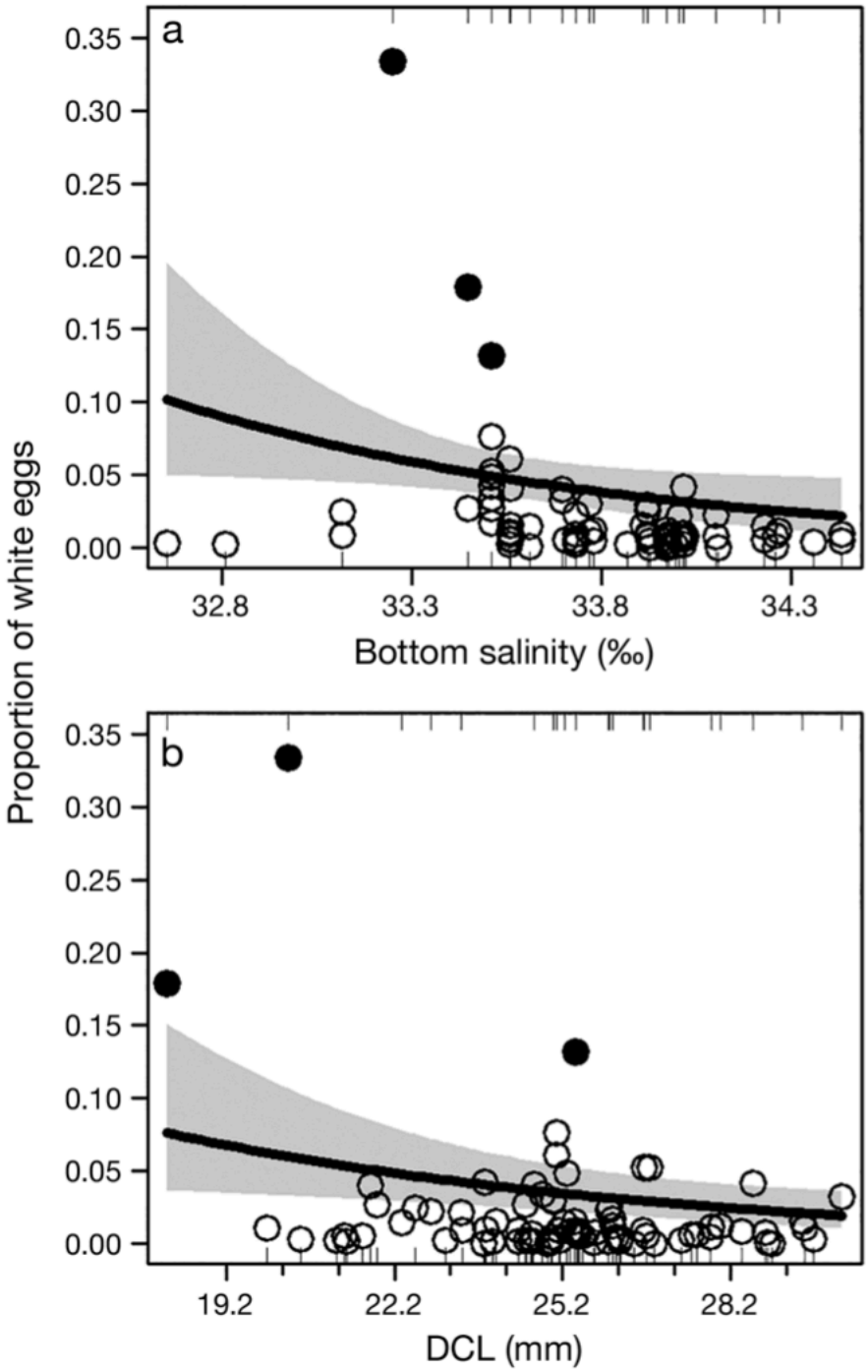


Figure 3.6. Relationships between proportion of parasite-infected shrimp eggs ('white eggs') and (a) bottom salinity and (b) dorsal carapace length (DCL). Circles: all observations; closed circles: potential outliers (but included in the regressions). Gray shading denotes the 95% confidence interval

After the 3 potential outliers were removed, all observations of p_{we} were below 10% for each individual. The new best model was with bottom temperature as a fixed effect and sampling location as a random effect (Table 3.4). p_{we} for an individual at mean bottom temperature (7.21°C) was estimated at 1.1%, and decreased by 0.84% with a 1°C increase in bot- tom temperature. The proportion of variation in p_{we} explained by bottom temperature was 6.5%.

Table 3.4. Model statistics for the null model and the competing models for factors associated with proportion of white eggs of an egg mass (see Table 3.1 for abbreviations). *: p<0.05; **: p<0.01

Models (Beta)				Marginal R^2_{GLMM}	Conditional R^2_{GLMM}	ΔAIC	LOOCV RMSE
P_{we} ~	BOTTEMP		$+\varepsilon_l$	0.065	0.387	0.000	0.011
P_{we} ~			$+\varepsilon_l$	NA	0.377	1.070	0.011
P_{we} ~	BOTTEMP		$+\varepsilon_y+\varepsilon_l$	0.065	0.387	2.000	0.011
P_{we} ~	BOTTEMP+BOTSAL		$+\varepsilon_y+\varepsilon_l$	0.087	0.379	2.872	0.011
P_{we} ~			$+\varepsilon_y+\varepsilon_l$	NA	0.377	3.070	0.011
P_{we} ~		BOTSAL	$+\varepsilon_y+\varepsilon_l$	0.023	0.367	3.671	0.011
P_{we} ~DCL	+BOTTEMP		$+\varepsilon_y+\varepsilon_l$	0.064	0.387	3.972	0.011
P_{we} ~DCL	+BOTTEMP+BOTSAL		$+\varepsilon_y+\varepsilon_l$	0.087	0.380	4.812	0.011
P_{we} ~DCL			$+\varepsilon_y+\varepsilon_l$	0.002	0.378	4.885	0.011
P_{we} ~DCL		+BOTSAL	$+\varepsilon_y+\varepsilon_l$	0.026	0.368	5.434	0.011

All explanatory variables included in the models were not statistically significant, suggesting that these variables had no significant effects on p_{we}. Furthermore, the proportion of variation in p_{we} explained by the fixed effects was less than 9%, even if all 3 variables were included in the model. The variance of the random effect of year was estimated to be zero and was thus dropped from the model. The random effect of sampling location was able to account for 37.7% of the variation. The predictive errors were not improved by inclusion of variables.

The difference in predictive error (RMSE) between the models with and without potential outliers was around 1%, suggesting that all models have similar predictive performance. The semivariance of model residuals did not suggest the presence of spatial autocorrelation.

The results show that p_{we} varied among sampling locations in the GOM. However, more than 95% of the observations of p_{we} were <5%; thus p_{we} did not appear to be associated with any of the explanatory variables measured in this study.

3.5 Discussion

Although bottom salinity showed significant impacts on the probability that a female was infected (P_{inf}), the magnitude of the salinity effect was small, and thus did not seem to be biologically relevant. The possibility of a female carrying at least 1 white egg was high; however, changes in bottom salinity did not appear to be the main driver.

The issue of white eggs has been observed and associated with temperature since the 1960s, and P_{inf} seemed to vary considerably among years in the GOM. P_{inf} was between 0 and 74% in the 1960s (Haynes and Wigley 1969; Apollonio et al. 1986) and between 55 and 92% in the 1970s (Stickney 1980). In our study, mean P_{inf} was 73.81% during 2012−2016 and was not correlated with bottom temperatures. However, based on the bottom temperatures observed in the 1960s by Apollonio et al. (1986) and in our study, more observations of higher temperatures were recorded during 2012−2016. Furthermore, the model with combined data suggested that P_{inf} increased with bottom temperature.

Apollonio et al. (1986) indicated that when temperature was lower than 5.5°C, P_{inf} was zero or low. Our observations corresponded to those of Apollonio et al. (1986), as P_{inf} was high at temperatures above 6°C. The findings and comparisons from these 2 studies imply that (1) the

environment in the GOM regarding bottom temperature might have changed considerably over the past 50 year, and (2) warming temperatures had a negative impact on P_{inf}.

Although P_{inf} seems to be high, p_{we} was mostly below 5%, with an average of 2.16% and 3 observations higher than 10%. p_{we} was around 1% in the 1960s (Haynes and Wigley 1969) and 1−5% in the 1970s (Stickney 1980). Occasional high p_{we} was also observed. Stickney (1981a) indicated that p_{we} was mostly 2−5%, but could be up to 30%. It is likely that the magnitude of the temperature effect was too small to be statistically detectable in our study, as most (more than 75%) of the observed p_{we} values were lower than 3% at warm temperatures compared to temperatures collected in the 1960s by Apollonio et al. (1986).

Haynes and Wigley (1969) reported 1% p_{we} on average for an individual, with 0−27 white eggs per female observed from 47 individuals (2.3 white eggs on average). In this study, we observed 0−173 white eggs per female from 143 examined individuals, with 13.7 white eggs on average for an individual. Assuming that bottom temperatures during Haynes and Wigley's (1969) surveys were similar to the temperature data collected by Apollonio et al. (1986), temperature might influence the number of white eggs, as the observed number of white eggs per female in our study was much higher than observed in the 1960s (Haynes and Wigley 1969). Furthermore, correlations between p_{we} or number of white eggs per female and female body size were not found in this study or in the study by Haynes and Wigley (1969). However, temperature data were not provided with p_{we} or number of white eggs per female observed in the 1960s and 1970s for further analyses and comparisons.

Although p_{we} seems to be low, and high p_{we} only happens occasionally, the possibility of a female carrying at least 1 white egg was generally high and appears to be negatively impacted by warming water temperatures. However, the extent of the white egg issue on reproductive

potential requires a study of integrated information on quality and quantity of eggs of individuals in the shrimp population, because reproductive potential is related to the fecundity− female body size relationship as well as size frequency of females in the population. The results derived from our study provide essential information that can be used to improve the estimation of reproductive potential of northern shrimp in the GOM.

The issue of white eggs has been reported for the GOM northern shrimp stock since the 1960s, while it seems to have not been an issue for other shrimp stocks. Temperature has been hypothesized as a contributing factor affecting variability in recruitment of several northern shrimp stocks in the North Atlantic in the last decade (Wieland and Siegstad 2012; Jónsdóttir et al. 2013; Brosset et al. 2019). Although the exact species of parasite infecting the shrimp eggs has not been conclusively identified, the results of this study suggest that the infection of white eggs is associated with bottom temperature, providing a potential underlying mechanism that affects egg mortality and subsequent recruitment of northern shrimps in the North Atlantic, as warming water temperatures have been widely observed in the North Atlantic (Mills et al. 2013). However, more studies are needed to evaluate the trend in white egg infection rates with the continuing increases in water temperature in the GOM. Biological samples may also need to be collected from other northern shrimp stocks (e.g. Atlantic Canada and Greenland) to identify if white eggs are present in these more northern stocks. With a continuing increase in water temperature throughout the distributional range of northern shrimp, white eggs may be observed in areas in which they have not been observed previously. A broader monitoring program across the whole distributional range of northern shrimp needs to be developed to improve our understanding of the mechanisms of white egg infections and possible consequences on the northern shrimp fisheries throughout the North Atlantic Ocean.

In addition to white eggs, other diseases such as black spot gill syndrome have also been reported for northern shrimp in the North Atlantic (Lee et al. 2019). Black spot gill syndrome has been an issue for northern shrimp since the 1960s (Rinaldo and Yevich 1974). Although the parasitic ciliate (*Synophrya* sp.) was recently identified (Lee et al. 2019), there is a lack of evidence linking black spot gill syndrome to the recruitment failures and the declines of the fishery in recent years. Alternatively, black spot gill syndrome could have an impact on the recruitment failures of GOM northern shrimp. Further investigation is needed to evaluate the impacts of diseases on the shrimp populations in the GOM as well as the North Atlantic.

CHAPTER 4. EFFECTS OF ENVIRONMENTAL FACTORS ON REPRODUCTIVE POTENTIAL OF THE GULF OF MAINE NORTHERN SHRIMP (*Pandalus borealis*)

4.1 Abstract

The northern shrimp (Pandalus borealis) once supported a significant winter fishery in the Gulf of Maine (GOM). However, the population collapsed in 2012 and a fishery moratorium has been in effect since 2014 due to record low levels of spawning stock biomass and persistent recruitment failure. An important parameter in determining population dynamics, fecundity, has not been evaluated for more than 30 years, during which time the GOM has warmed significantly. In this study, we quantified three reproductive characteristics related to fecundity: potential fecundity (PF, number of viable eggs), relative fecundity (RF, number of viable eggs per gram of body weight), and egg size (ES) for GOM northern shrimp. The results showed that PF was strongly related to body size, while RF was independent of body size. Egg size declined with increasing body size for larger females. Egg size is related to size at larval hatch, suggesting that although larger females produced more eggs, those eggs may produce smaller larvae. In contrast with previous studies, PF and RF were positively correlated with bottom temperature in the relatively warm years of our study. We hypothesize that the positive temperature-fecundity relationship we observed may not reflect a direct effect, but possibly a compensatory response relating to decreased population density during the time period of our study. In addition, the environmental effects we observed may to some extent reflect progression of the inshore migration of females. The information derived in this study can help us have a better understanding of environmental effects on reproductive potential for climatically vulnerable species such as the GOM northern shrimp.

4.2 Introduction

Northern shrimp once supported a significant winter fishery in the Gulf of Maine (GOM) (Clark et al., 2000). However, landings have fluctuated widely and the fishery has collapsed three times since its inception in the 1930s (Clark et al., 2000; ASMFC NSTC, 2018). The most recent collapse occurred in 2012, and a moratorium has been imposed on the shrimp fishery since 2014 due to low stock size and recruitment failure for several successive years (ASMFC NSTC, 2018). Recent recruitment failures have been associated with unfavorably warm water temperatures in the GOM (ASMFC NSTC, 2018), which has been warming faster than 99% of the global oceans (Pershing et al., 2015). The intense warming rates have shifted hatch timing of northern shrimp (Richards, 2012), driven distributional (Nye et al., 2010; Kleisner et al., 2017) and phenological shifts (Staples et al., 2018; Staudinger et al., 2019) of various marine species, and impacted fisheries in the GOM (Mills et al., 2013; Pershing et al., 2015; Mazur et al., 2018).

Northern shrimp are sequential hermaphrodites (Shumway et al., 1985). They hatch as males in near-shore areas in GOM during winter-spring. During their second year of life, they migrate to offshore areas where they mature as males and remain until they transform to mature females at presumed age 3. They then reproduce for 2 years as females. Following molting and mating, females extrude eggs that are fertilized externally and attached to their pleopods. The eggs are carried for several months before the females move to near-shore areas to release the offspring (Shumway et al., 1985; Richards, 2012). Unlike many fish species, most crustacean species' egg development is synchronous and fecundity is determinant (Parsons and Tucker, 1986). This means that all fertilized eggs develop at the same pace and no new eggs are extruded

during incubation, thus the number of eggs (reproductive potential) can be estimated at the onset of spawning.

A well-defined stock-recruitment relationship is highly beneficial to fisheries management because it facilitates the estimation of exploitation levels that ensure future recruitment and sustainable fisheries (Hannah et al., 1995). Stock can be represented by several measures, including parental stock biomass, mature female stock biomass, or egg production (Hilborn and Walters, 1992). Of these, egg production estimated from size-specific (or age-specific) fecundity is considered the best measure of spawning stock (Morgan et al., 2011). An advanced understanding of reproductive potential and its relationship with environmental variables could help elucidate the causes of recruitment failure of northern shrimp in the GOM.

Direct and indirect effects of climatic factors are important for determining species' vulnerability to climate change (de los Ríos, 2018). GOM northern shrimp are perceived to be particularly susceptible to environmental variability as they are at the southernmost limit of their distribution in the North Atlantic Ocean (Shumway et al., 1985). A few studies have estimated size-fecundity relationships for GOM northern shrimp (Apollonio and Dunton, 1969; Haynes and Wigley, 1969; Apollonio et al., 1986) and one of these examined possible temperature effects on fecundity, but found no correlation between spatial variation in bottom temperature and fecundity (Apollonio et al., 1986). However, these studies were conducted more than fifty years ago, before the steep warming trend began (Richards et al., 2012; Pershing et al., 2015). In the current study, we provide a contemporary estimate of the size-fecundity relationship for GOM northern shrimp and examine possible effects of environmental variability on potential fecundity (PF), relative fecundity (RF) and egg size (ES) in order to better understand environmental

effects on reproductive potential of northern shrimp, and their possible relations to recruitment failures.

4.3 Materials and Methods

4.3.1 Biological samples and environmental data

Samples of female shrimp were collected during bottom trawl surveys conducted in the GOM by the Northeast Fisheries Science Center (NEFSC) during October and November, 2012-2016 (Fig. 4.1 and Table 4.1). The NEFSC bottom trawl surveys were designed to capture a broad range of species across the northeast continental shelf of the United States. The survey uses a stratified random sampling design in which sampling locations are randomly selected within strata that are defined by bathymetry and latitude (Politis et al., 2014). The number of stations within a stratum is generally proportional to the area of the stratum and the overall variation of multispecies distribution among strata (Politis et al., 2014). At each sampling station, bottom temperature and bottom salinity data were measured with electronic profiling conductivity/temperature/depth sensors (Politis et al., 2014). Samples of the shrimp catch were frozen at sea, then thawed and processed at NEFSC for collecting size and life stage data, then refrozen for later laboratory processing for fecundity data collection. A total of 143 samples collected from 46 sampling locations in October and November 2012-2016 were used for estimating PF, RF, and ES (Table 4.1).

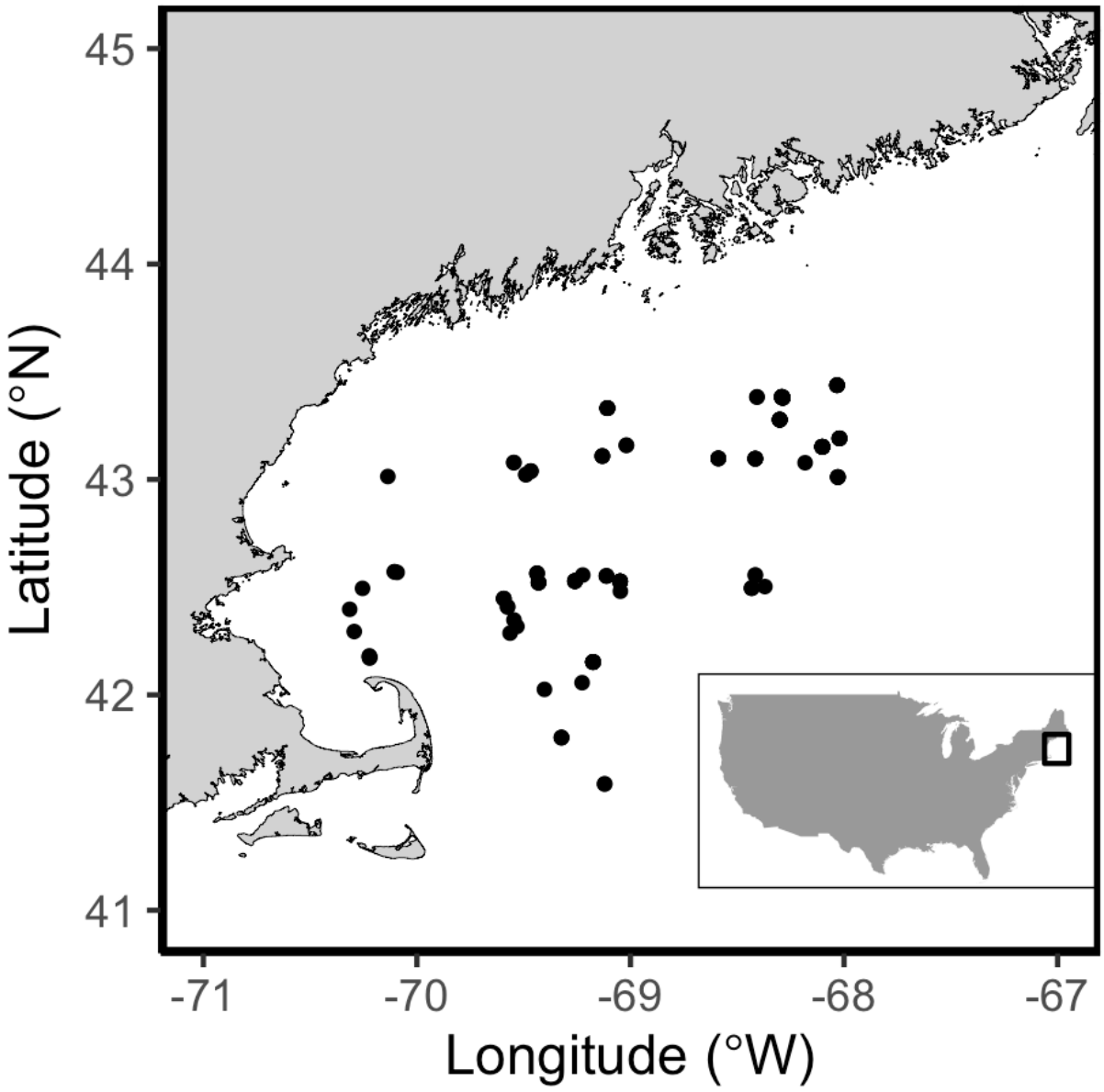


Figure 4.1. Sampling locations of the Northeast Fisheries Science Center fall bottom trawl surveys where northern shrimp (*Pandalus borealis*) were collected in the Gulf of Maine during 2012-2016

Table 4.1. Number of stations where northern shrimp (*Pandalus borealis*) were collected and number of northern shrimp successfully used for estimating fecundity and egg size

Year	Number of stations	Number of northern shrimp used	Sampling period
2012	7	14	Oct 3-Nov 10
2013	5	8	Nov 13-Nov 15
2014	12	40	Oct 31-Nov 10
2015	21	76	Oct 21-Nov 5
2016	1	5	Nov 7
Total	46	143	

4.3.2 Laboratory procedures

Only shrimp bearing early developmental stage (non-eyed) larvae (Haynes and Wigley, 1969) were used for estimating fecundity. These could be readily identified under a microscope because the eggs were transparent and eyes had not yet formed. Non-eyed individuals comprised 99.9% of the samples collected. Fecundity (number of viable eggs of an individual female) estimated from our samples taken in the fall were considered PF rather than realized fecundity (number of eggs successfully hatched per female) due to potential egg loss during the remaining 3-4 months of incubation. RF was estimated as the number of eggs per gram of female body weight (potential fecundity/female body weight without eggs), which is a measure of individual

reproductive investment (Pereira et al., 2017). Parasite-infected eggs (white eggs; Haynes and Wigley, 1969, Apollonio et al., 1986, Chang et al., 2020) were recognized under a microscope based on their appearance. The proportion of white eggs ranged from 0-33% with an average of 2.16% of all eggs for an individual (Chang et al., 2020). These white eggs were excluded from the analyses as they are non-viable (Chang et al., 2020).

An analysis of the optimal number of samples needed for estimating fecundity indicated that length-stratified random sampling is more cost-effective than simple random sampling assuming the fecundity of northern shrimp is size-dependent (Chang and Chen, 2020). Therefore, we used length-stratified sampling which divided the range of shrimp sizes (dorsal carapace length, DCL, mm) into 10 length intervals (approximately 1 mm per length interval) and sampled 1 shrimp from each length interval. All available samples were processed, adhering as closely as possible to the protocols developed in Chang and Chen (2020).

Shrimp were processed for fecundity estimation by thawing the specimens at room temperature and using forceps to gently tease the eggs masses off the pleopods. Biological data including DCL (mm), lateral carapace length [LCL, mm; for comparing with Apollonio et al., (1986)], wet weight of each egg mass (g), and female body weight [air-thawed wet weight (g)] without eggs were measured. Egg samples were then preserved in 10% neutral formalin (Parsons and Tucker, 1986) for later processing. Eggs were removed from formalin, stained with toluidine blue and imaged under a trinocular dissecting microscope (OMAX V434BL54P-C140U) mounted with a digital camera (OMAX A35140U3). Egg masses were gently teased apart using forceps. ObjectJ (Schneider et al., 2012) was used to count all eggs (except white eggs) in each egg mass and measure their longest diameter (μm). ES of an individual female was defined as the average diameter (μm) of viable eggs in her egg mass.

4.3.3 Statistical analysis

Generalized additive mixed models (GAMMs; Wood, 2017) with negative binomial, Gaussian, and lognormal distributions were used to examine the relationships between predictors and PF, RF, and ES, respectively. The distributions chosen were based on the distributions of the dependent variables (Zuur et al., 2009). A variance inflation factor (VIF) analysis was conducted to identify multicollinearity before fitting models to the data. Predictors with VIFs exceeding 3 were considered collinear with other variables, a threshold VIF of 3 was thus set for evaluating possible collinearity between predictors in the data set (Schmiing et al., 2013; Brosset et al., 2019). The form of the GAMMs was:

$$g(y) \sim \alpha + \sum_{i=1}^{k} f_i(x_i) + \varepsilon_y + \varepsilon_l + \varepsilon$$

where g() is the link function, y is PF, RF or ES, f_i is the i^{th} smooth function based on thin plate regression splines, x_i is the i^{th} explanatory variable, ε_y and ε_l are random effects of year and sampling location, and ε is residual error. Year and sampling location were included in the models as nested random effects to account for possible spatial or temporal pseudoreplication and for variation among years and sampling locations (Hurlbert, 1984; Zuur et al., 2009; Weltz et al., 2013, Thorsen and Minto, 2015).

The VIF analysis indicated that multicollinearity was not an issue as all explanatory variables (DCL, bottom temperature, and bottom salinity) had VIFs < 3. Therefore, models of all combinations of the three explanatory variables were built (Fisher et al., 2018). Model selection was based on full-subsets information theoretic approaches (Anderson and Burnham, 2002; Fisher et al., 2018), using Akaike's information criterion corrected for small sample sizes (AICc;

Akaike, 1973; Hurvich and Tsai, 1989), Bayesian information criterion (BIC; Wit e al., 2012), AICc weights (ωAICc) and BIC weights (ωBIC, Fisher et al., 2018), deviance explained by the model, and graphical inspection. Root mean squared error (RMSE) estimated with leave-one-out cross-validation (LOOCV) was used to evaluate predictive performance of models (Zuur et al., 2009; Arlot and Celisse, 2010). The random structure was first assessed by AICc with all the smooth terms included in the models (Zuur et al., 2009) using restricted maximum likelihood (REML) methods. The random effects included in the model that had the lowest AICc were considered the optimal random structure.

After the random components were determined, models with different combinations of predictors were developed and compared using maximum likelihood methods. The base model for estimating PF included DCL based on previous studies (Haynes and Wigley, 1969; Apollonio et al., 1986). Little is known about the effect of DCL on RF and ES for northern shrimp, so the effect of DCL was not included in the null model for RF and ES.

Models with differences in AICc (ΔAICc) <2 were considered candidate models that were well supported by the data (Anderson and Burnham, 2002). Models with the lowest AICc scores were refit with REML (Zuur et al., 2009; Wood, 2017) and effect plots of these models were presented. Biological relevance and the relationships between response variables and predictors were graphically examined. Relative importance of predictors was quantified as the sum of AICc weights of models in which a predictor is present (Anderson and Burnham, 2002; Fisher et al., 2018).

The residuals of the optimal model were graphically evaluated with QQ-plots and plots of Pearson residual errors against fitted values to inspect any patterns in the residual errors (Zuur et al., 2009). Semivariance of Pearson residual errors from the most optimal model was examined for the presence of spatial autocorrelation (Cressie, 1993; Pebesma, 2004; Gräler et al., 2016). Simple linear regression models were fit to bottom temperature and bottom salinity data with day of year for exploring potential effects of spawning migration on the environmental relationships with reproductive characteristics.

4.3.4 Population fecundity

To fully compare with Richards et al. (2012), a GAM model was built for individual PF using only DCL as the explanatory factor in the model.

$$\mathrm{PF}_i = f(DCL_i) + \varepsilon_i$$

The size-fecundity relationship was then applied to the stratified mean population size composition estimated in Richards et al (2012) for estimating population fecundity:

$$N_{i,y} = \sum_{s=1}^{k} \frac{N_{s,i,y} A_{s,y}}{\sum A_{s,y}}$$

$$Population\ fecundity_y = \sum_{i=1}^{m} PF_{i,y} N_{i,y}$$

where s=stratum, i=length, y=year, $N_{s,i,y}$ = mean number of ovigerous female at length i in stratum s in year y, A_{sy} = area of stratum s in year y.

The same data of $N_{i,y}$ were used to estimate population fecundity based on the model developed in Richards et al. (2012):

$$Population\ fecundity_y = \sum_{i=1}^{m} \left(-0.198{L_{i,y}}^2 + 128.81L_{i,y} - 17821\right)N_{l,y}$$

where $L_{i,y}$ is the i^{th} DCL in 0.1 mm in year y.

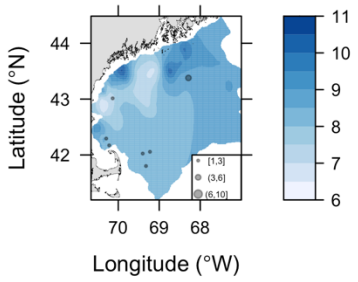
DCLs of the full range were used for the PF model in this study, and only ovigerous females ≥ 20 mm DCL were used for Richards et al.'s (2012) model.

Variability in northern shrimp PF was compared with previous studies (Haynes and Wigley, 1969; Apollonio et al., 1986). RMSEs were estimated using the data and corresponding models in each study. The DCLs of Apollonio et al. (1986) were converted from lateral dorsal carapace length (LCL) based on the equation: LCL = 1.464+1.086DCL (n=653, r^2=0.92) which was developed using data collected in this study (Fig. A1).

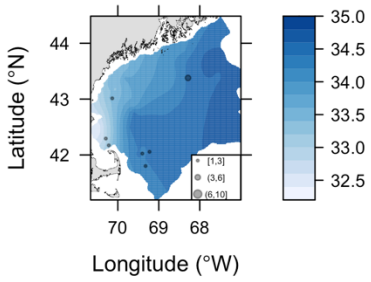
4.4 Results

Maps of the sampling locations, number of samples, and contours of interpolated bottom temperature and salinity data collected by NEFSC fall bottom trawl surveys during October-November 2012-2106 are shown in Fig. 4.2.

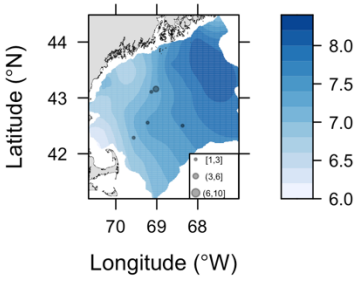
BOTTEMP 2012



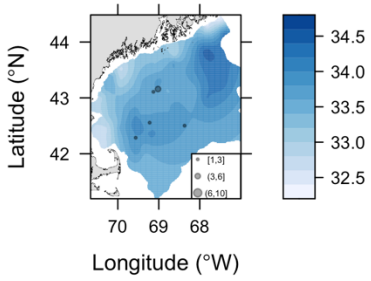
BOTSAL 2012



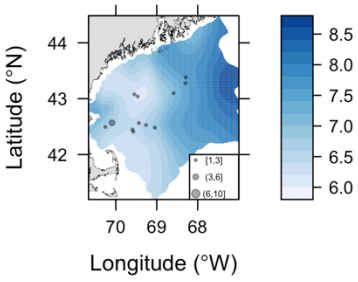
BOTTEMP 2013



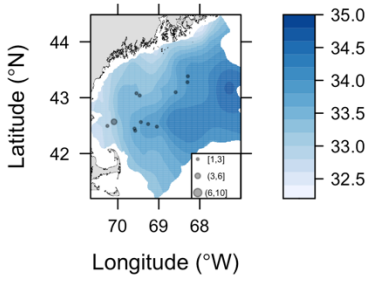
BOTSAL 2013



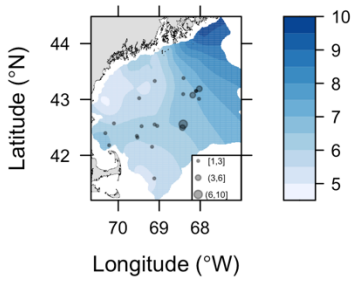
BOTTEMP 2014



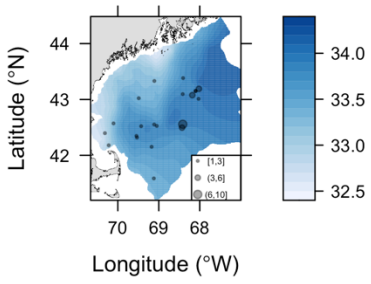
BOTSAL 2014



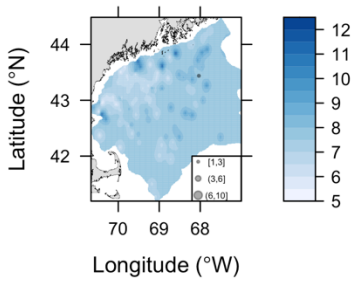
BOTTEMP 2015



BOTSAL 2015



BOTTEMP 2016



BOTSAL 2016

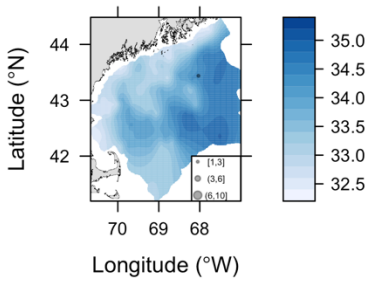


Figure 4.2. Sampling locations of the Northeast Fisheries Science Center fall bottom trawl surveys where northern shrimp (*Pandalus borealis*) were collected in the Gulf of Maine during 2012-2016. Different sizes of dots denote number of shrimp used from each location. Maps of contours of interpolated bottom temperature and bottom salinity were overlaid on the bottom.

4.4.1. Potential fecundity (PF)

The observed PF ranged from 124-3557 eggs per female, and the observed DCL ranged 18.15-30.84 mm. The average PF was 1442 eggs (standard deviation, SD, 620.73) for an average sized female (25 mm DCL) (Table 4.2). The spatial distribution of DCL in each year is shown in Figure A.3. Spatial autocorrelation was not suggested by semivariance over the sampling area (Figure A.4).

Table 4.2. Table 2. Observed ranges, means, and standard deviations of response and explanatory variables in this study, Richards et al. (2012), and Apollonio et al. (1986). DCL = dorsal carapace length; BOTTEMP = bottom temperature; BOTSAL = bottom salinity; PF = potential fecundity; RF = relative fecundity; ES = egg size; SD = standard deviation; n = number of samples; RMSE = root mean squared error. The DCLs of Apollonio et al. (1986) were converted from lateral dorsal carapace length (LCL) based on the equation: LCL = 1.464+1.086DCL (n=653, r^2=0.92) which was developed using data collected in this study.

Study	DCL (mm)			BOTTEMP (°C)			BOTSAL (‰)			PF (number of eggs)			RF (number of eggs/g)			ES (µm)			n	RMSE
	range	mean	SD	range	mean	SD	range	mean	SD	range	mean	SD	range	mean	SD	range	mean	SD		
This study	18.15-30.84	25	2.56	5.67-11.01	7.6	1.19	32.42-34.63	33.83	0.43	124-3557	1442.2	620.7	15.0-289.3	161.9	52.3	891.8-1335.9	1091.6	69.2	143	417
Richards et al. (2012)	22.24-30.82	27	2.4	NA	NA	NA	NA	NA	NA	882-3396	2425.9	628.2	NA	NA	NA	NA	NA	NA	47	288.8
Apollonio et al. (1986)	21.93-30.78	25.65	1.66	NA	NA	NA	NA	NA	NA	734-2775	1616.4	344.2	NA	NA	NA	NA	NA	NA	202	221.6

The optimal model for PF included the random effect of year, and the smoothers of bottom temperature and bottom salinity (DCL was included in the base model, Table 4.3). This model had strong support from the data with an AICc weight of 0.98. The model was able to explain 48.4% of the deviance, and the predictive error (i.e. the average bias between prediction and observation) was 476.5 eggs. Both bottom temperature and bottom salinity were significant factors explaining the variation in PF, as the best model outcompeted other models (Fig. 4.3).

Table 3. Model statistics of the best models within 2 ΔAICc for potential fecundity (PF), relative fecundity (RF), and egg size (ES) of northern shrimp (*Pandalus borealis*). DE=deviance explained, EDF=estimated degrees of freedom, LOOCV RMSE=leave-one-out cross validation root mean squared error, ωAICc=AICc weights, ωBIC=BIC weights, DCL=dorsal carapace length, TEMP=bottom temperature, SAL=bottom salinity. σ_y and σ_l =random effects of year and sampling location. The DCL in the PF model is in parentheses as DCL was included in the null model.

Models								DE	EDF	LOOCV RMSE	ΔAICc	ΔBIC	ωAICc	ωBIC
PF	~	(DCL) +	TEMP +	SAL	+ σ_y			0.484	7.155	476.494	0	1.236	0.977	0.334
RF	~		TEMP +	SAL	+ σ_y	+ σ_l		0.569	28.031	52.075	0	0	0.456	0.782
RF	~			SAL	+ σ_y	+ σ_l		0.582	29.792	51.433	1.821	4.569	0.183	0.08
ES	~	DCL	+	SAL		+ σ_l		0.59	30.81	65.376	0	3.356	0.388	0.133
ES	~			SAL		+ σ_l		0.564	29.203	67.431	0.387	0	0.32	0.713
ES	~		TEMP			+ σ_l		0.573	30.624	69.667	1.841	3.414	0.154	0.129

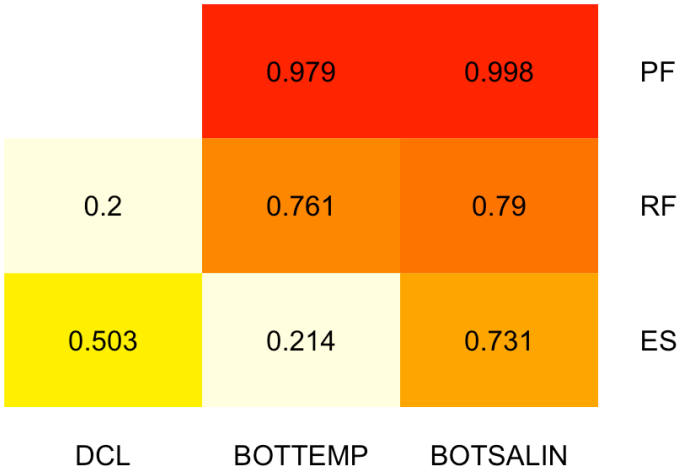


Figure 4.3. A heatmap of importance scores of each variable estimated from full-subsets generalized additive mixed models with information theoretic approaches for potential fecundity (PF), relative fecundity (RF), and egg size (ES). DCL=dorsal carapace length (mm); BOTTEMP = bottom temperature; BOTSALIN = bottom salinity.

The partial effects of predictors are shown in Fig. 4.4. Both DCL and bottom temperature had estimated degrees of freedom of 1, which means the relationships between PF and these two variables are linear via the link function. PF increased with DCL, showing that larger females are able to produce more eggs (Fig. 4.4a). Biological senescence (i.e. an asymptotic relationship) was not observed for PF. PF was also positively correlated with bottom temperature (Fig. 4.4b). The relationship between PF and bottom salinity was nonlinear (Fig. 4.4c). PF increased with bottom salinity when the bottom salinity was lower than 33.9‰, and decreased slightly when bottom salinity was higher than 33.9‰. The semivariance of the Pearson residual errors did not suggest a pattern of spatial autocorrelation (Fig A.5a). No clear patterns were found in the QQ plot and Pearson residuals against fitted values (Fig A.6a-b).

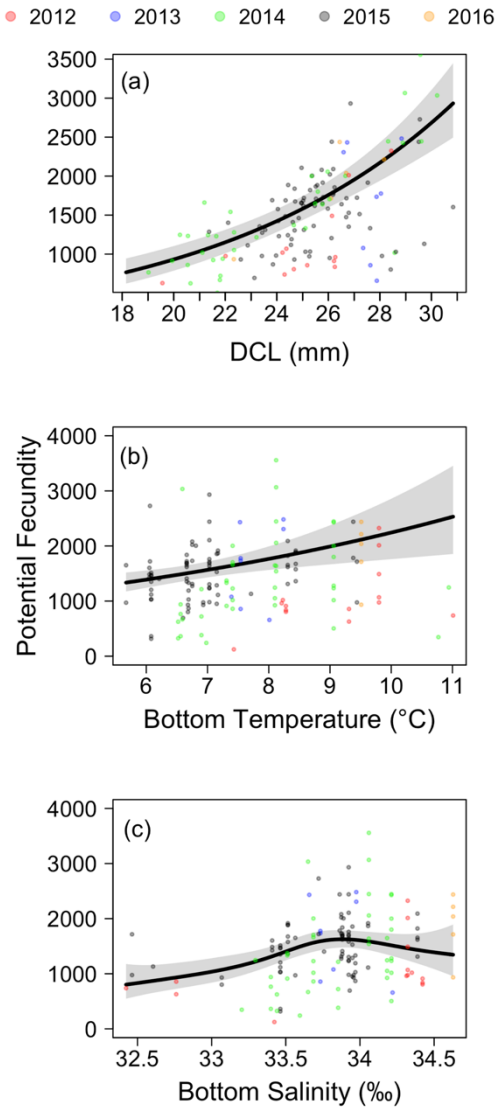


Figure 4.4. Partial effects of each variable included in the best model with lowest AICc on potential fecundity (PF). The black solid lines denote the modeled relationship, the gray bands denote the 95% confidence intervals about the estimated relationship, and colored dots are observations of each year. (red=2012, blue=2013, green=2014, black=2015, and orange=2016; DCL=dorsal carapace length)

The relationship between bottom temperature and bottom salinity is shown in Fig. 4.5 (slope=0.076, $p<0.05$, $r^2=0.04$). The observed bottom temperature where shrimp were collected increased with day of year with a slope of 0.028 (Fig. 4.6a, $p<0.05$, $r^2=0.03$). The observed bottom salinity where shrimp were collected decreased with day of year with a slope of -0.013 (Fig. 4.6b, $p<0.05$, $r^2=0.06$).

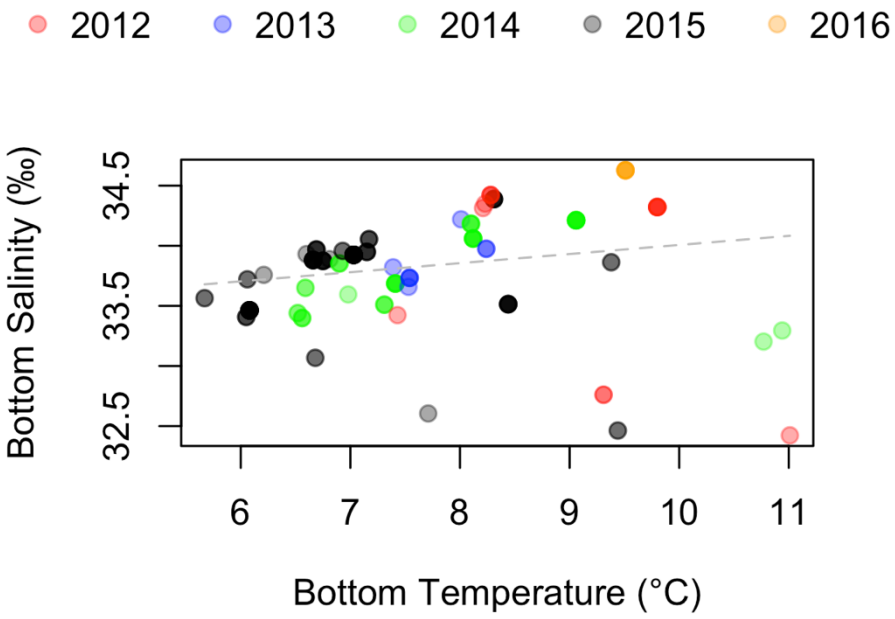


Figure 4.5. Relationships between bottom temperature and bottom salinity. The gray dashed lines are linear regression lines, and colored dots are observations of each year. (red=2012, blue=2013, green=2014, black=2015, and orange=2016)

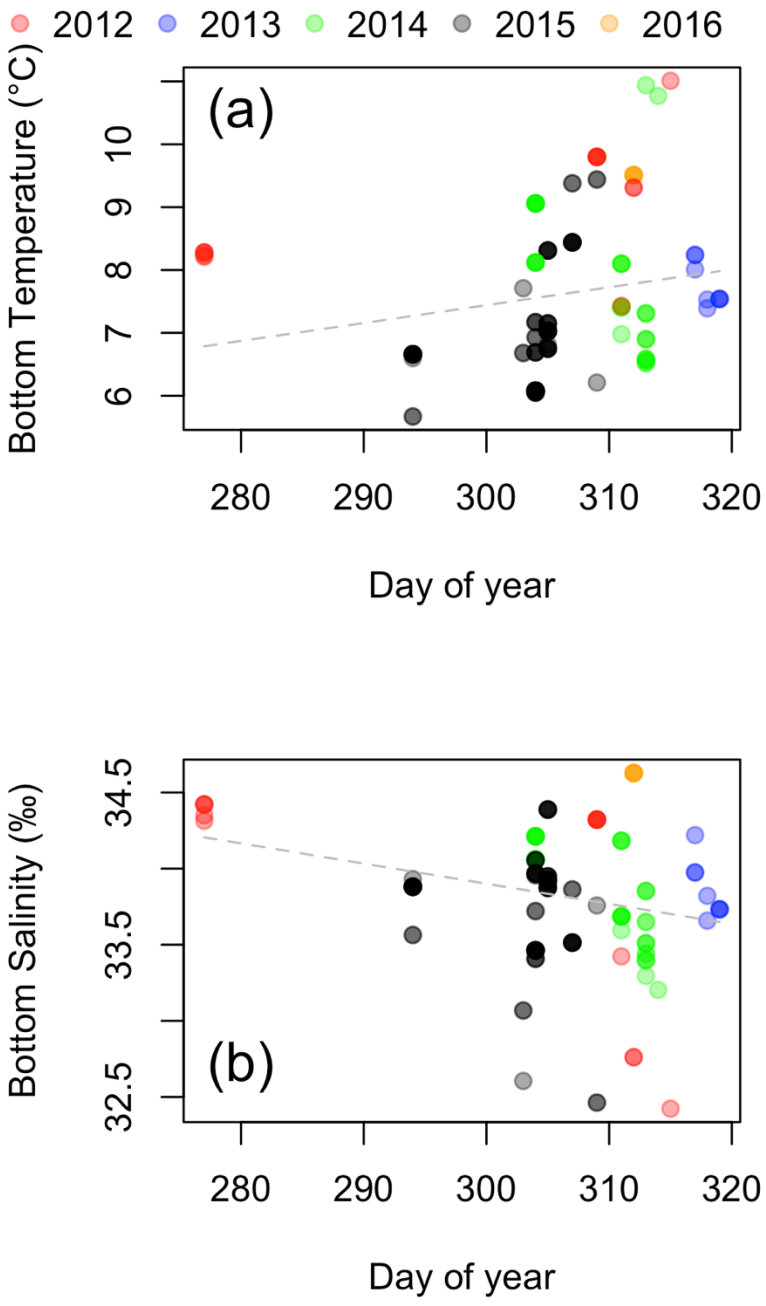


Figure 4.6. Relationships between (a) bottom temperature and (b) bottom salinity with day of year. The gray dashed lines are linear regression lines, and colored dots are observations of each year. (red=2012, blue=2013, green=2014, black=2015, and orange=2016)

4.4.2 Relative fecundity (RF)

The observed RF ranged 15.0-289.3 eggs/g of body weight with a mean of 161.9 (SD=52.3), and the observed body weight (without eggs) ranged 3.33-15.8 g (SD=2.4) (Table 4.2). The optimal random structure for RF included sampling location and year in the model selected by AICc. The best model selected by both AICc and BIC included bottom temperature and bottom salinity with ωAICc=0.46 and ωBIC=0.78 (Table 4.3). This model explained 57% of the deviance. Both bottom salinity and bottom temperature were identified as important factors across the top models (relative importance scores of 0.79 and 0.76 respectively). RF was not driven by DCL as DCL was not found in the top models and had a relatively low importance score (Fig. 4.3).

RF increased linearly with bottom temperature (Fig. 4.7a), indicating that females with higher reproductive investment were found at higher bottom temperature. Similar to PF, RF increased with bottom salinity when bottom salinity was lower than 33.9‰, and it leveled off when bottom salinity was higher than 33.9‰ (Fig. 4.7b). The semivariance did not suggest a presence of spatial autocorrelation (Fig A.5b). No patterns were found in the QQ-plot and Pearson residuals against fitted values (Fig A.6c-d).

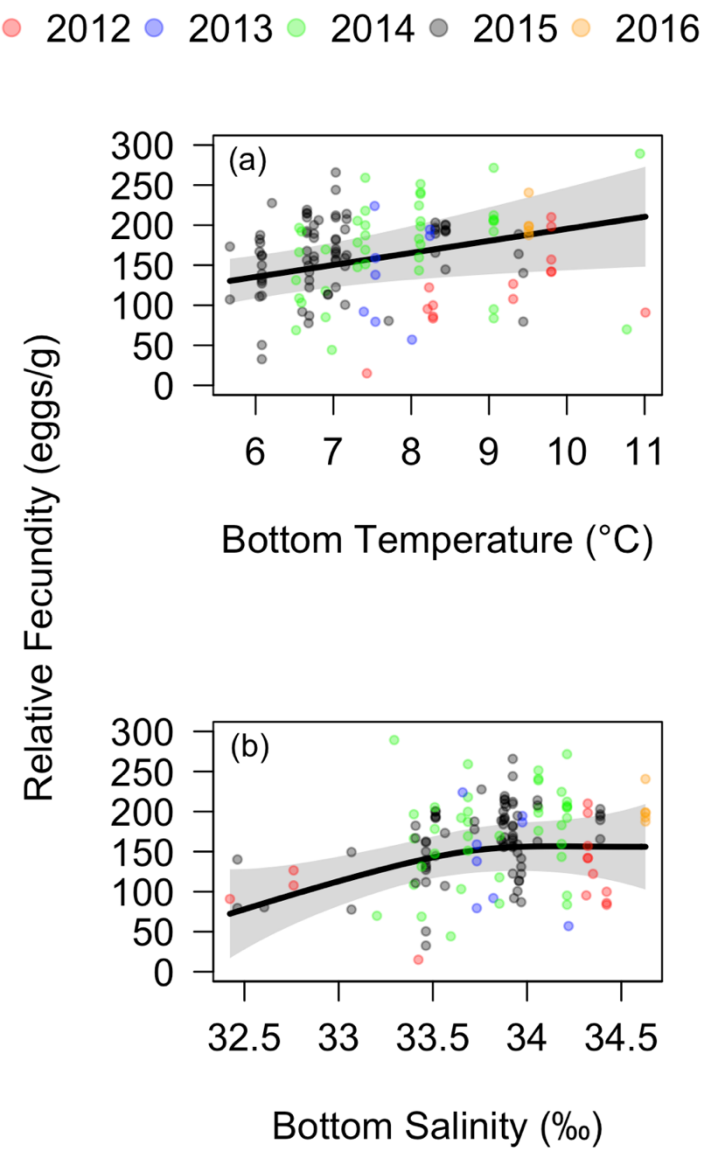


Figure 4.7. Partial effects of each variable included in the best model with lowest AICc on relative fecundity (RF). The black solid lines denote the modeled relationship, the gray bands denote the 95% confidence intervals about the estimated relationship, and colored dots are observations of each year. (red=2012, blue=2013, green=2014, black=2015, and orange=2016)

4.4.3 Egg size (ES)

The optimal random effect structure selected by AIC included sampling location. The model with the lowest AICc included DCL and bottom salinity with an ωAICc of 0.39 (Table 4.3). The second best model which included a single factor of bottom salinity was preferred by BIC with a ωBIC of 0.71 (ωAICc=0.32). The third ranking model included only bottom temperature, with ωAICc=0.15, showing relatively lower importance of bottom temperature for explaining the variation in ES. Bottom salinity (importance score=0.73) and DCL (importance score=0.50) had higher importance for explaining ES than did bottom temperature (importance score=0.21) (Fig. 4.3). The top three models within 2 ΔAICc had similar percentages of deviance explained and predictive errors, showing a level of model uncertainty.

The average ES increased with DCL when females were smaller than 25 mm DCL (Fig. 4.8a), then started decreasing with DCL when females were larger than 25 mm DCL. For the relationship between ES and bottom salinity, the average ES linearly decreased with bottom salinity, with larger eggs found in lower bottom salinity (Fig. 4.8b). The presence of spatial autocorrelation was not suggested by the semivariance of the Pearson residuals (Fig A.5c). No patterns were found in the plot of Pearson residuals against fitted values (Fig A.6f), although the QQ-plot right tail deviated from a normal distribution due to a few large values of ES observations (Fig A.6e).

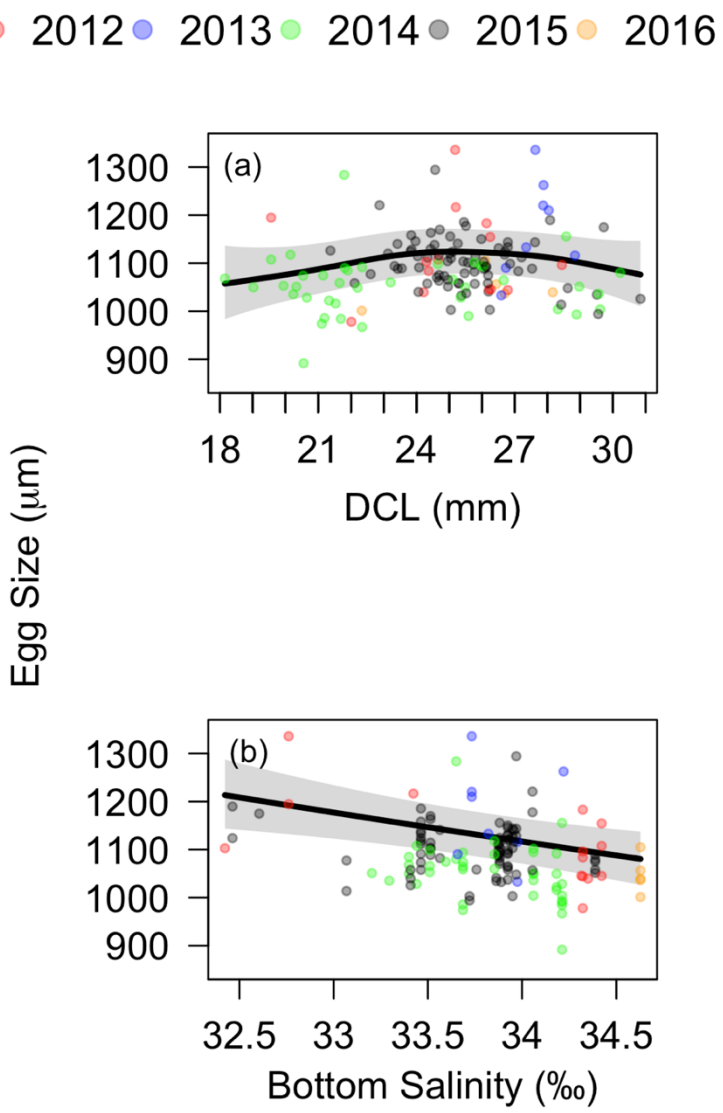


Figure 4.8. Partial effects of each variable included in the best model with lowest AICc on egg size (ES). The black solid lines denote the modeled relationship, the gray bands denote the 95% confidence intervals about the estimated relationship, and colored dots are observations of each year. (red=2012, blue=2013, green=2014, black=2015, and orange=2016; DCL=dorsal carapace length)

4.4.4 Population fecundity comparison

The females used in Haynes and Wigley (1969) for estimating a size-fecundity relationship ranged 22.2-30.8 mm, and the observed PF ranged 881.7-3396.2 eggs. The stock assessment (ASMFC NSTC 2018) and Richards et al. (2012) used a parabola to fit Haynes and Wigley's (1969) size-fecundity data (Fig. 4.9). Their estimated PF was generally higher than that estimated in this study at a given DCL. For example, the estimated PF for a 25 mm female in this study was 1394.0 eggs and was 44% higher in Richards et al (2012) (2007 eggs). Estimated PF became negative for shrimp with DCL<20 mm when a parabola was used [as in Richards et al. (2012)]; however, most shrimp < 20 mm DCL are males. The average absolute difference between fecundity of females > 22 mm DCL estimated by Richards et al. (2012) and this study was 34.0%.

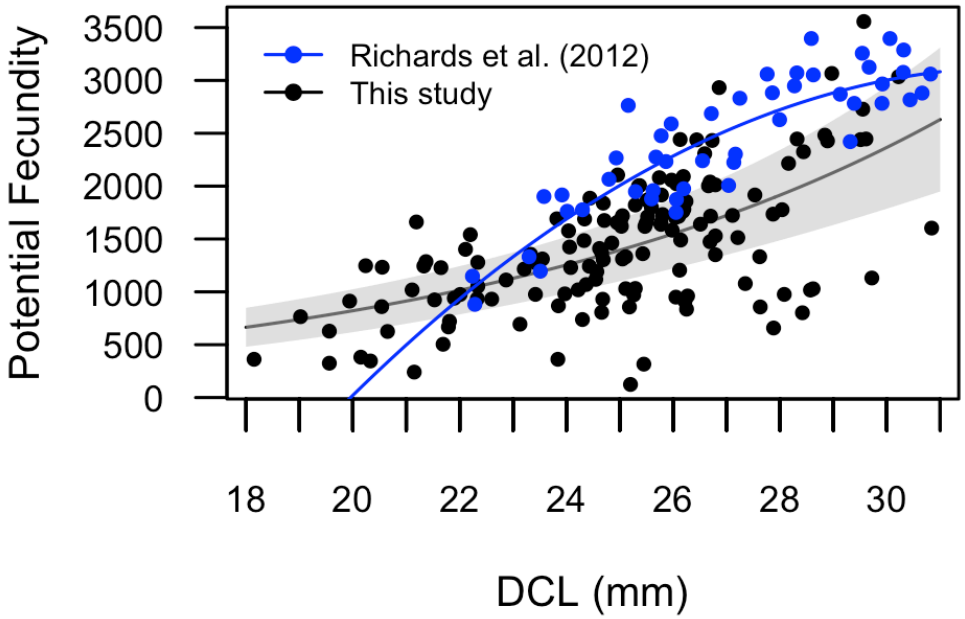


Figure 4.9. Relationships between potential fecundity and dorsal carapace length (DCL) estimated by Richards et al. (2012) and this study for northern shrimp (*Pandalus borealis*). Black

dots denote observations in this study, and blue dots denote observations in Richards et al. (2012). The gray shadowed area is the 95% confidence interval for the estimated size-fecundity relationship in this study.

Population PF estimated in this study using Richards et al.'s (2012) model is shown in Fig. 4.10. The population PF estimated by these two models had similar trends, although Richards et al.'s (2012) population PF was almost always higher than the fecundity estimated in this study over the time series except for 2014. The average absolute difference between population PF was 0.173 million eggs.

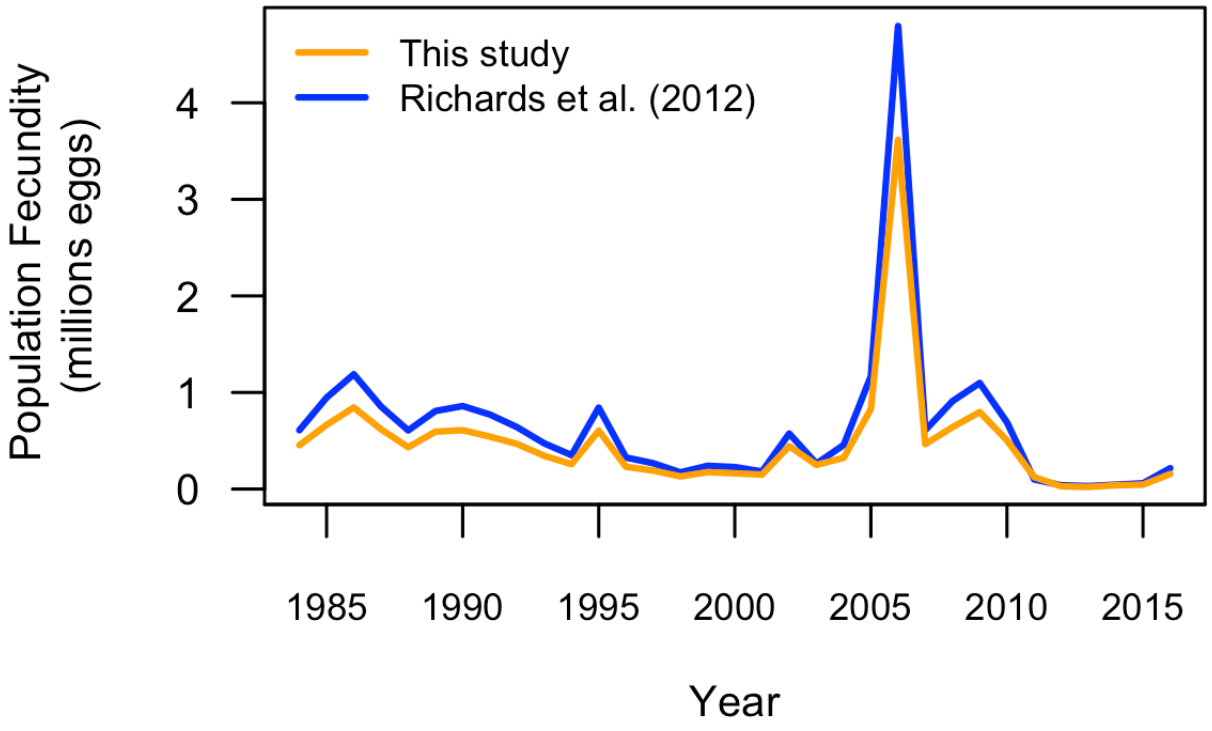


Figure 4.10. Annual population fecundity of northern shrimp estimated by PF (potential fecundity) models developed in this study and Richards et al. (2012) using Atlantic States

Marine Fisheries Commission (ASMFC) summer shrimp survey data during 1984-2016. The orange line denotes population fecundity estimated from the PF model in this study, the blue line denotes Richards et al. (2012).

4.5 Discussion

The reproductive characteristics investigated in this study describe different aspects of reproductive investment. PF describes the potential number of eggs a female can produce, and the size-fecundity relationship is typically used to estimate egg production of populations (Hilborn and Walters, 1992). RF puts egg production into the context of energetics by taking into account the size (energy content) of individual females (Dautov et al., 2004; Pereira et al., 2017). RF can also be easily applied to the population size composition for estimating reproductive potential. ES can be an indicator of egg quality, with larger eggs generally thought to be of higher quality as larger eggs contain higher amounts of yolk (Dautov et al., 2004; Zimmermann et al., 2015). This in turn may be related to larval survival (Ramirez-Llodra et al., 2000; Dautov et al., 2004; Zimmermann et al., 2015). All these reproductive characteristics and consequently the reproductive potential of the population may be affected by environmental factors such as water temperature and salinity (Apollonio and Dunton, 1969).

4.5.1 Potential fecundity and relative fecundity

Relatively higher variability in northern shrimp PF was observed in this study compared to previous studies (Haynes and Wigley, 1969; Apollonio et al., 1986; Table 4.2). The higher RMSE could be due to sample sizes, ranges and variation in DCL, differences in spatiotemporal scales, sampling seasons, and higher natural variability. High variability in PF between years,

seasons, or areas had been observed for Pandalid shrimp by several previous studies (Parsons and Tucker, 1986; Hannah et al., 1995; Jónsdóttir, 2018). Although the PF varied substantially among locations, spatial autocorrelation was not suggested by the spatial variogram for PF in this study.

Several studies have investigated the effects of water temperature on fecundity of northern shrimp (Nunes, 1984; Apollonio et al., 1986; Parsons and Tucker, 1986). Temperature effects were not consistently detected, but when they were, the relationships were negative, i.e. fecundity was lower at higher temperatures. Apollonio et al. (1986) noted that differences in individual fecundity of northern shrimp collected from eight sampling locations in the GOM in August and September 1968 were not significant and could not be correlated with the ambient bottom temperatures. Parsons and Tucker (1986) also found no clear relationship between individual fecundity and ambient temperatures using samples collected intermittently from several locations in Canadian waters (northwest Atlantic) during 1971-1982. However, Nunes (1984) concluded that cold (3°C) to moderate (6°C) temperatures were more suitable than high (9°C) temperature in laboratory studies of northern shrimp egg production in Alaska, as fecundity-at-size was higher for northern shrimp in 3 and 6°C. Furthermore, Apollonio et al. (1986) reported a negative relationship between predicted mean fecundity of a 25 mm DCL female and annual bottom temperature (April-July) during 1968-1982 (non-consecutive years). In contrast with these previous studies, we found that potential fecundity was positively correlated with temperature. However, bottom temperatures in the GOM have changed considerably since those earlier studies, and northern shrimp population size is much smaller (ASMFC NSTC 2019). The bottom temperature observed in Apollonio et al.'s (1986) study in October-December 1967-1968 ranged ~2-6°C, and the bottom temperature observed in Oct-Dec

in this study ranged 5.7-11°C with a mean of 7.6°C. Bottom temperatures in NEFSC bottom trawl surveys in October and November 1967-1968 (temporal and spatial coverage comparable to our study) ranged 3.6-10.1 °C (mean 6.5).

In addition to temperature effects (see Discussion for ES), northern shrimp fecundity could also be affected by factors such as density-dependence (Shumway et al., 1985; Moraes-Valenti et al., 2010; Jónsdóttir, 2018). It has been suggested that ovigerous female GOM northern shrimp may gravitate towards the coldest water available to them when they encounter thermal gradients (Stickney and Perkins, 1977, Shumway et al., 1985). Furthermore, Chang et al. (in prep) found that the variability around the center of gravity of ovigerous northern shrimp distribution was positively correlated with population abundance and negatively correlated with bottom temperature. Therefore, we hypothesized that population density of ovigerous females may be negatively associated with bottom temperature in the GOM, and fecundity is assumed to be negatively correlated with population density due to competition for food and cannibalism of eggs of neighbor females (Elliot, 1970). More data and further analysis are needed to test this hypothesis.

The effect of bottom salinity has been rarely investigated. In our study, around 53% of the females were found in the range of 33.5-34.2‰, and bottom salinity had a very high importance score (0.998) for explaining PF. In the GOM, females undertake an extensive onshore migration to hatch their brood in relatively shallow, less saline water (Fig 4.2, Haynes and Wigley 1969, Apollonio et al. 1986). In our study, bottom salinity was negatively correlated with day of year, which likely reflects the progress of this onshore migration. Thus the high importance score of salinity may reflect changes in the ambient environment during females ' inshore migration, rather than a direct association between PF and salinity. Females that have

reached lower salinities presumably have been carrying their egg masses longer, so may have had greater egg loss by the time of sampling.

All three of the explanatory factors we investigated (size, bottom temperature and salinity) had high importance scores for explaining variation in fecundity; however, the correlations between PF and variables suggested by the models may not imply causal effects. Other habitat variables associated with bottom temperature or bottom salinity may be important, such as dissolved oxygen or food availability. Brillon et al. (2005) observed that females provide parental care to embryos by beating the pleopods, presumably to oxygenate the egg mass. These habitat variables can also directly or indirectly influence females 'fitness, and consequently influence reproductive potential.

RF can be used as a measure of reproductive investment of an individual female. Unlike PF which increased with body size, RF was independent of DCL. Pereira et al. (2017) also found no association between reproductive investment and female body size, suggesting that individuals at all sizes make similar reproductive investment per unit of body. RF was dependent on bottom temperature and salinity and both variables had high importance scores for explaining relative fecundity. The effects of bottom temperature and salinity on RF are similar to that on PF, confirming that higher temperatures and higher salinities were correlated with higher PF and RF during incubation. Bottom temperature and salinity in our study reflected the offshore environment in October and November 2012-2016. As female shrimp migrate to inshore areas for hatching, their exposure to the environment at different stages may change as shown in the correlation between bottom temperature/salinity and day of year. These environmental effects during incubation on the final quantity and quality of eggs need further investigation.

4.5.2 Size-fecundity relationships

The size-fecundity relationship can be an important input in stock assessment (ASMFC NSTC, 2018). Size-fecundity relationships have been estimated for Pandalid shrimp using parabolic (Richards et al., 2012), power (Parsons and Tucker, 1986; Hannah et al., 1995), and linear regression models (Apollonio et al., 1986). In all of these, the PF was strongly correlated with female body size; however, a high level of variation may exist within sizes. Hannah et al. (1995) indicated that the curvature of the size-fecundity relationship might not be easily detected due to high variation in fecundity at a given length. Therefore, this study used a GAMM to have more flexibility to account for curvilinearity and possible biological senescence, making a biological meaningful size-fecundity relationship covering females at a wide range of sizes.

Comparison between fecundity studies is complicated by several factors in addition to different model structures. These include size range of females sampled; methods for counting number of eggs; and spatial and temporal variation. A lack of small females may result in incomplete size-fecundity relationship and lead to biased estimates of fecundity for small females (e.g., Haynes and Wigley, 1969). Zero to five percent (mean=1%) of the ovigerous females in our study were smaller than 20 mm, and 0-16% (mean=5%) were smaller than 22mm during 1991-2018. There could be biases in methods for counting number of eggs if eggs were subsampled and fecundity estimated by expansion. Instead of using estimated number of eggs by weight, we counted all the eggs of an egg mass for each female in this study. Our estimates of PF are thus likely more accurate than fecundity estimated from expanded subsamples. In terms of spatial and temporal variation, fecundity may differ between years and areas due to changes in environmental conditions or diseases (Parsons and Tucker, 1986; Chang et al., 2020). Assuming Haynes and Wigley's (1969) samples (collected during 1953-1966, mostly 1963-1965) were

collected from similar environmental conditions as Apollonio et al. (1986) in the 1960s, the changes in the environment could also be a reason for the differences in the size-fecundity relationships between Richards et al. (2012) and this study.

In this study, spatial patterning or significant differences between years were not observed, which is consistent with the observations of Apollonio et al. (1986). Given the fact that the GOM has been experiencing rapid warming in the past two decades (Pershing et al., 2015; Kavanaugh et al., 2017), it seems possible that the size-fecundity relationship could change due to changes in the environmental conditions. However, the sampling years in this study were mostly warm years and the time series too short to identify a pattern in temporal changes.

4.5.3 Egg size

ES can be a measure of egg quality as larger eggs contain higher amounts of yolk and produce larger offspring which may be better prepared for competition and predator avoidance (Ramirez-Llodra et al., 2000; Dautov et al., 2004; Zimmermann et al., 2015). Therefore, ES could be a key factor in determining the success of larval survival. Wieland (2004) observed reduced size-at-sex change of Greenlandic northern shrimp in the late 1990s and inferred that it could have resulted from smaller eggs due to shorter incubation periods under warmer water temperatures (Wieland, 2005). This implies that temperature could have lagged effects on egg and larval sizes and consequently population productivity. In our study, DCL had a moderate importance score for explaining the variation in ES and the relationship between ES and DCL was nonlinear, which might not be detected by linear regression models (Nunes, 1984; Ahamed and Ohtomi, 2011). Larger females (DCL>27 mm) produced more eggs but the average ES was smaller, possibly reflecting biological senescence or a tradeoff in quantity versus quality. If egg

quality and offspring survival rates are positively associated with ES, then larger females do not necessarily produce offspring with higher survival rates than females at average sizes (around 25 mm DCL). However, our results could also be explained by size-dependent spawning times. If large females spawn later than small ones, larger females would have smaller eggs at an earlier developmental stage at a given sampling time. Apollonio et al. (1986) indicated that females of an older age group spawn later than the younger age group although this has not been well-documented, and the delay in spawning time between age groups was not estimated.

Although Clarke et al. (1991) suggested that factors determining reproductive output and egg size are different, in our study bottom salinity was identified as an important factor for explaining PF, RF, and ES with high importance scores across models. However, PF and RF were positively correlated with bottom salinity, while ES was negatively correlated with bottom salinity. This likely reflects progress of the inshore migration because eggs develop and ES increases over time while bottom salinity decreases as the females migrate towards the shore.

4.5.4 Other factors

The models we built in this study were able to explain 48-59% of the variation in PF, RF, and ES. Other factors that might explain variation in reproductive potential and fitness of female shrimp include food availability (Hannah et al., 1995), population density (Apollonio et al., 1986), disease (Chang et al., 2020), and egg loss during incubation (Elliot, 1970; Skuladottir et al., 1978; Stickney, 1981; Nunes, 1984; Apollonio et al., 1986; Brillon et al., 2005). Skuladottir et al. (1978) estimated 30-54% egg loss during seasonal migration in Iceland, and Brillon et al. (2005) observed high variation in individual egg loss (1-99%) during incubation in laboratory experiments. The ovigerous females sampled in this study were collected from an early phase of

incubation (October-November), thus the fecundity estimates would not reflect total egg loss during the entire incubation period (5-6 months). Further investigations are needed for evaluating egg loss during incubation for GOM northern shrimp.

Several explanations have been observed or postulated for egg loss, including diseases (Apollonio et al., 1986), incomplete fertilization (Parsons and Tucker, 1986), cannibalism by neighboring shrimp, parental behavior (Elliot, 1970), and temperature (Brillon et al., 2005). Brillon et al. (2005) observed higher egg loss when ambient temperature of females increased, but Nunes (1984) reported highest egg loss at low temperature. Brillon et al. (2005) suggested that more intense movements of pleopods may be needed in warmer temperatures for ovigerous females to supply the higher oxygen demand of developing embryos, which might increase egg loss during incubation.

In addition to egg loss, ambient temperature during incubation could also affect size of larvae at hatching (Nunes, 1984; Brillon et al., 2005). Although the temperature effect was less important than DCL and bottom salinity on ES in this study, in laboratory studies larger larvae were observed to hatch from eggs incubated at lower temperatures with higher survival rates and growth rates in Alaskan waters (Nunes, 1984) and in the St. Lawrence estuary (Brillon et al., 2005).

A number of hypotheses have been postulated or tested for explaining the most recent collapse of the GOM northern shrimp population. Chang et al. (2020) examined the effects of bottom temperature and salinity on the incidence of parasitized eggs (white eggs) of the GOM northern shrimp, but correlations between white eggs and environmental factors were not significant. Richards (2012) indicated that the hatching timing and duration of the GOM northern shrimp hatch period has shifted due to warming water temperatures; however, the match-

mismatch theory (Hjort 1914, Cushing 1990) was not able to explain variation in shrimp survival at early life stages (Richards et al. 2016). Richards and Hunter (2021) presented evidence that the sudden shrimp population collapse might have been due to a spike in predation by longfin squid (*Doryteuthis pealeii*) after a distributional shift during the 2012 heatwave. Charleson (2020) observed decreases in size-at-transition of northern shrimp at higher water temperatures, which could result in smaller average female size and consequent skewed population size structures, leading to decreased reproductive potential. Chang et al. (in prep) examined habitat suitability for adult GOM northern shrimp during summer and fall and found that increasing proportions of low quality habitat were correlated with a declining spawning stock biomass index at a 2-year lag. Consequently, in addition to changes in reproductive potential of the GOM northern shrimp, high predator pressure and loss of suitable habitat might have significant impacts on population size.

4.6 Conclusions

This study examined three different measures of reproductive output of GOM northern shrimp and their relationships with biotic and abiotic factors. PF quantified reproductive potential of an individual female, and the size-fecundity relationship derived from this study can be used for estimating egg production of the shrimp population. PF was most strongly correlated with female body size, yet our results showed that the ambient environmental conditions (bottom temperature and salinity) were also important for explaining the variation in PF. RF was independent of female body size, suggesting that relative reproductive output is constant across female size. RF was affected by environmental conditions similarly to PF. ES declined in large females (> 27 mm DCL), which could indicate biological senescence or may reflect differences

in spawning times of large and small females. In contrast with previous studies, PF and RF were positively correlated with bottom temperature in the relatively warm years of our study. We hypothesize that the positive temperature-fecundity relationship we observed may not reflect a direct effect, but possibly a compensatory response relating to decreased population density during the time period of our study. In addition, the environmental effects we observed may to some extent reflect progression of the inshore migration of females.

GOM northern shrimp are considered vulnerable and sensitive to environmental variability as they are at the southernmost of their distribution. Various hypotheses for the most recent collapse of the GOM northern shrimp population examined or tested in this and previous studies suggest there may be multiple factors undermining population growth and sustainability, primarily through indirect effects. Accordingly, reproductive potential should be monitored and considered in evaluating population dynamics for such vulnerable species, because it may play an important role in the process of population recovery.

CHAPTER 5. TEMPERATURE AND ABUNDANCE EFFECTS ON SPATIAL STRUCTURES OF NORTHERN SHRIMP (*PANDALUS BOREALIS*) AT DIFFERENT LIFE STAGES IN THE OCEANOGRAPHICALLY VARIABLE GULF OF MAINE

5.1 Abstract

The Gulf of Maine (GOM) northern shrimp, *Pandalus borealis*, once supported a significant winter fishery, but a moratorium has been placed on the fishery since 2014 due to recent recruitment failures that have been attributed to unfavorably warm water temperatures. The GOM is at the southernmost end of the northern shrimp's range, suggesting its population dynamics and distribution may be vulnerable to warming water temperatures. While spatial distributions may provide important insights on temporal changes in abundance, the spatial structures of the GOM northern shrimp have not been thoroughly explored. In this study we used survey data to estimate density-based spatial indicators for GOM northern shrimp at various life stages, and explore temporal changes in the spatial structures of northern shrimp. We also examined the relationships between abundance and bottom temperature with the season- and life stage-specific spatial distribution indicators. We observed patchier distributions over time with a distributional shift toward the shore for adult shrimp, which were associated with declining population abundance and warming bottom water temperatures, respectively. These season- and life stage-specific density-dependent spatial distribution indicators depict spatial structures for the GOM northern shrimp, providing important information for a better understanding of shrimp population dynamics under climate change and improved management.

5.2 Introduction

The northern shrimp, *Pandalus borealis*, a caridean decapod, is an ecologically important and commercially-harvested demersal species in the Gulf of Maine (GOM) that once supported a significant winter fishery targeted on egg-bearing females. Commercial shrimp landings in the GOM over the last five decades have fluctuated widely, from a low of less than 100 mt to more than 12,000 mt, during which the fishery has experienced three major collapses (ASMFC NSTC [Atlantic States Marine Fisheries Commission Northern Shrimp Technical Committee] 2018). The fishery has been under a moratorium since 2014 as a result of repeated recruitment failures (ASMFC NSTC 2018), which have been attributed to unfavorably warm water temperatures in recent years (Richards et al. 2012; ASMFC NSTC 2018).

Northern shrimp are protandrous hermaphrodites and have a complex life history. In the Gulf of Maine, they mature at age 2 as males, before transitioning to females at presumed age 3 (Table 1.1). These age 3 females spawn with mature males in the deeper waters offshore in the western GOM in late summer and early fall. During late fall and early winter, adult females (presumed age 4 to 5) carry their developing eggs externally and begin their inshore migration to shallower (~50 m) inshore waters in the northern GOM (Haynes and Wigley 1969; Incze unpublished). The females brood the developing eggs externally as they migrate inshore where, arriving in winter, the larvae are released. Following a period of planktonic development that lasts several weeks, the larvae settle to the benthos as juveniles in late spring and early summer, remaining in inshore waters as males for one to three years before returning to deeper offshore waters. Spawning and the transition of males to females at age 3, all occur in deep, offshore waters of the western GOM (reviewed in Shumway et al. 1985; Apollonio et al. 1986; Clark et

al. 2000; Richards et al. 2012, and summarized in Table 1.1). It is during the winter months in inshore GOM waters when the egg-bearing females are targeted by the commercial fishery.

The temperature dependence of shrimp population dynamics is not straightforward. The inshore-offshore migrations by maturing males and egg-bearing females are thought to be triggered by seasonal water temperature cues (Apollonio and Dunton 1969; Shumway et al. 1985; Apollonio et al. 1986), and Apollonio et al. (1986) suggested that the GOM northern shrimp are intolerant of abnormally warm water temperature during certain stages in their life cycle. More recent analyses (Clark et al. 2000) support the contention of Apollonio and Dunton (1969), Haynes and Wigley (1969), and Apollonio et al. (1986) that shrimp reside in the deeper offshore waters of the western GOM in summer where bottom water temperatures generally remain the coldest (4 to 6 ºC), with the 6 ºC bottom isotherm delineating the boundary of shrimp's preferred temperatures (only lower abundances of shrimp were found outside that boundary in bottom waters warmer than 6 ºC).

In addition to water temperatures, other factors – e.g., salinity, depth, and bottom sediment grain sizes – have been reported to influence the distributions of northern shrimp (Shumway et al. 1985; Apollonio et al. 1986; Wieland 2005). In their study, Worm and Myers (2003) ruled out predation by cod as affecting significantly shrimp population abundance in the GOM. However, Richards and Hunter (2021) reported that the most recent collapse of the northern shrimp population could be due to the invasion of longfin squid (*Doryteuthis pealeii*) into the GOM during the 2012 heat wave, resulting in high predator pressure on the northern shrimp population. As for bottom-up influences, variability in food limitation accompanying temperatures changes should perhaps be considered. Shumway et al. (1985) characterized shrimp as opportunistic omnivores, functioning both as predators and scavengers. Haynes and

Wigley (1969) reported a direct relationship between shrimp abundances and high organic content of fine sediments in the GOM, and cite various earlier workers who observed the same apparent preference of shrimp for organic-rich, fine-grained bottom sediments which was correlated with depth. Apollonio et al. (1986) reported that shrimp feed on various benthic invertebrates and detrital phytoplankton and other organic detrital remains and that their examination of stomach contents revealed mostly mud and unrecognizable debris. There has been little if any follow up to these early studies on the food and feeding of shrimp, however. The general consensus is that water temperature exerts an overriding control on the dynamics of their abundance (Dow 1977; Worm and Myers 2003) and recruitment (Richards et al. 2012), and the fishery moratorium in place since 2014 was a reaction to record low levels of recruitment and spawning stock biomass as well as presumed unfavorable environmental conditions for the GOM northern shrimp (ASMFC NSTC 2019).

Based on satellite sea surface temperature (SST) analyses, surface temperatures in the GOM have been argued to be warming rapidly over the past two decades (e.g., Mills et al. 2013; Pershing et al. 2015). Such analyses of high spatial and temporal resolution satellite surface temperature data stand in stark contrast to available data on actual measurements of bottom temperatures, which, though far less extensive in their spatial and temporal coverage, would seem to be more relevant to bottom-dwelling (demersal) shrimp populations. Although the bottom temperature is relatively more stable than the SST in the GOM, it was shown that the bottom temperature has also been warming over the past three decades (Kavanaugh et al. 2017).

Water temperatures in the GOM are driven by a complex and highly variable interaction between both positive and negative seasonal surface heat fluxes, as well as advective processes that exchange GOM water masses with very different waters (different temperatures and

salinities) farther offshore, outside the GOM, and beyond the 200 isobath (Townsend et al. 2006; Townsend et al. 2015). Deep and bottom waters of the GOM are the result of influxes of relatively warm and salty Slope Waters from beyond the edge of the continental shelf; this Slope Water layer may extend upward from the bottom to depths less than 75 m (Bigelow 1927). Additional inflow to the GOM occurs as a surface layer of cold and relatively fresh Scotian Shelf Water that enters from the east and around Cape Sable, Nova Scotia, as a continuation of the Nova Scotia Current (Smith 1983; 1989). Sandwiched between these two layers in the interior GOM resides seasonally a cold intermediate water layer: convective sinking and mixing of cold surface waters produced in winter creates the relatively cold water temperatures at depth, which subsequent seasonal warming at the surface isolates as a cold intermediate water layer that slowly erodes during the remainder of the year (Hopkins and Garfield 1979). The bottom water temperatures that can be expected to affect shrimp populations in the western GOM will therefore reflect those water mass intrusions into the eastern GOM and their subsequent mixing as they spread to the western GOM, and the intensity of winter convection in the western GOM (Brown and Beardsley 1978).

The changes in bottom temperature would be expected to have impacts on spatial distribution and abundance of northern shrimp in the GOM. Based on density-dependent habitat selection (DDHS), the relationship between abundance and spatial distribution is positive—individuals would occupy the most optimal habitat when population abundance is low; as abundance increases, the population density in the most optimal habitat increases, resulting in decreasing habitat quality due to competition for food and other resources (Reuchlin-Hugenholtz et al. 2015; Thorson et al. 2016). As a result, as abundance increases, individuals start to spread into suboptimal habitats (Reuchlin-Hugenholtz et al. 2015; Thorson et al. 2016). The spatial

distribution indices that incorporate spatial structure information may therefore provide useful information to conservation and fisheries management. Nevertheless, the spatial structure of northern shrimp has not been explored in detail.

In this study we used survey data to estimate density-based spatial distribution indicators and examine the temporal changes of these spatial distribution indicators. We explored the relationship between abundance and bottom temperature with the spatial distribution indicators for northern shrimp at various life stages, and examined if these indicators provide important information and may be useful proxies of abundance. Specifically, we used survey data to estimate evenness and proportions of low/high density areas of the population to see how these spatial indicators are correlated with abundance and bottom temperature. Furthermore, we used center of gravity of each life stage to evaluate possible distributional shifts and develop hypothesis for the observed distributional shift.

5.3 Materials and Methods

5.3.1 Data

Shrimp samples and bottom water temperatures were collected by NOAA's Northeast Fisheries Science Center (NEFSC) during their fall bottom trawl surveys from 1991 to 2018, and by the ASMFC during their summer shrimp surveys from 1984 to 2019 (ASMFC NSTC 2018). Data collected by NEFSC spring bottom trawl surveys were not used due to incomplete time series. During the NEFSC fall surveys since 1991 and during the ASMFC shrimp surveys since 1984, bottom temperatures were measured with expendable bathythermographs (XBTs) or with a conductivity, temperature and depth profiler (CTD).

The ASMFC summer surveys were designed specifically for northern shrimp in the GOM for monitoring relative abundance, providing data required for northern shrimp annual stock assessment (Cao et al. 2017; ASMFC NSTC 2018). The NEFSC fall bottom trawl surveys were designed for multispecies such as demersal fish species and invertebrate species including northern shrimp, providing relative abundance for each species in the GOM (Politis et al. 2014).

Both surveys were conducted with a stratified random sampling design, following consistent sampling protocols. The sampling locations were allocated throughout the western GOM which was divided into several strata based on depth and geographical location (Fig. 5.1). For the summer northern shrimp survey, historical fishing pattern was also considered for stratification (Clark 1989; Cao et al. 2017). The sampling locations were randomly selected from each stratum, and the number of stations within a stratum was generally proportional to the area of the strata and the overall variation of multispecies distribution among strata for fall surveys (Politis et al. 2014); the importance of the stratum to the northern shrimp assessment was considered for summer survey (Cao et al. 2017). Details of the survey sampling design and vessel configuration were documented by Stauffer (2004) and Politis et al. (2014). A contour map of depth in the GOM (Fig. 5.2) was made using depth data collected by NOAA National Centers for Environmental Information (Boyer et al. 2018).

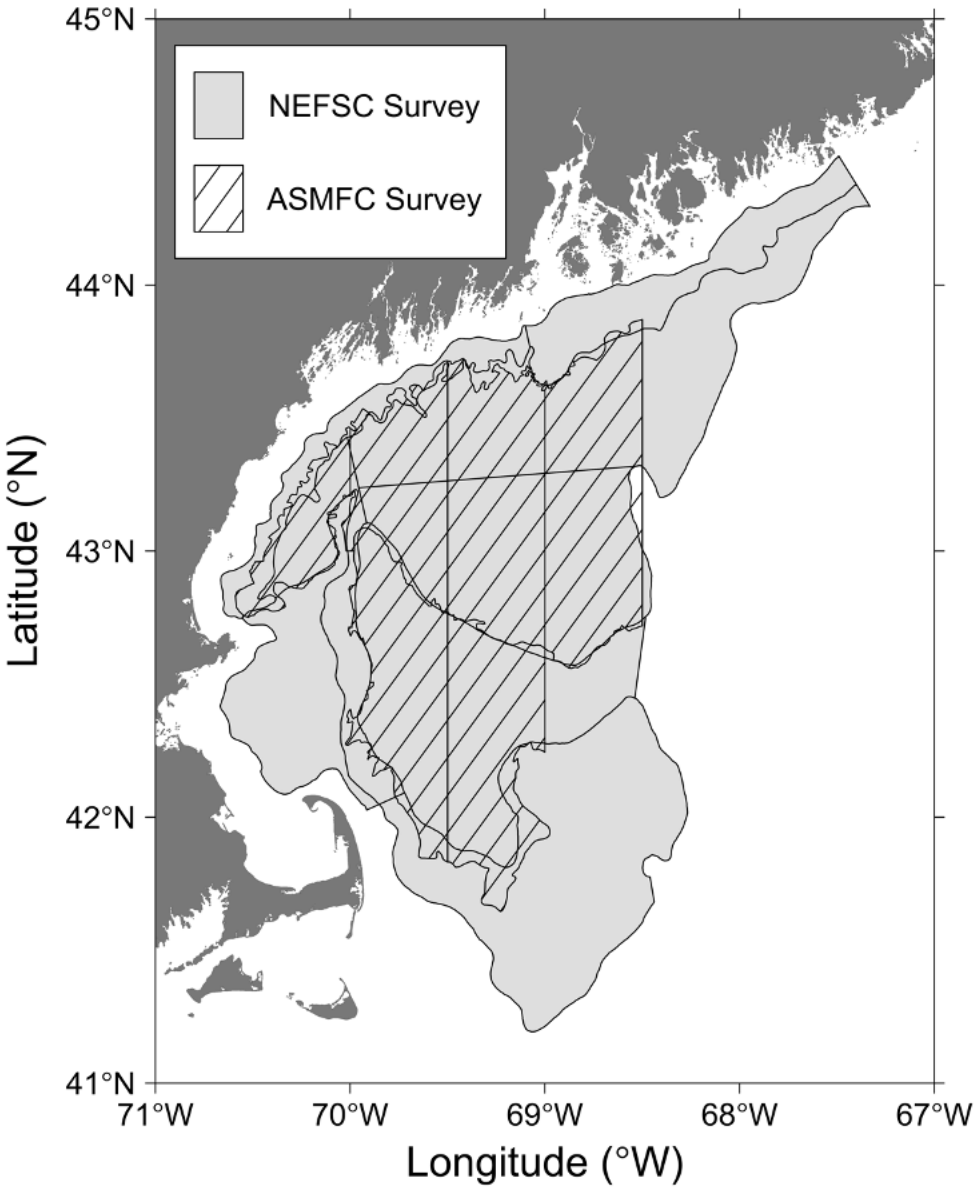


Figure 5.1. Sampling areas of Atlantic States Marine Fisheries Committee (ASMFC) summer shrimp bottom trawl surveys in 1984-2019 and Northeast Fisheries Science Center (NEFSC) fall bottom trawl surveys in 1991-2018.

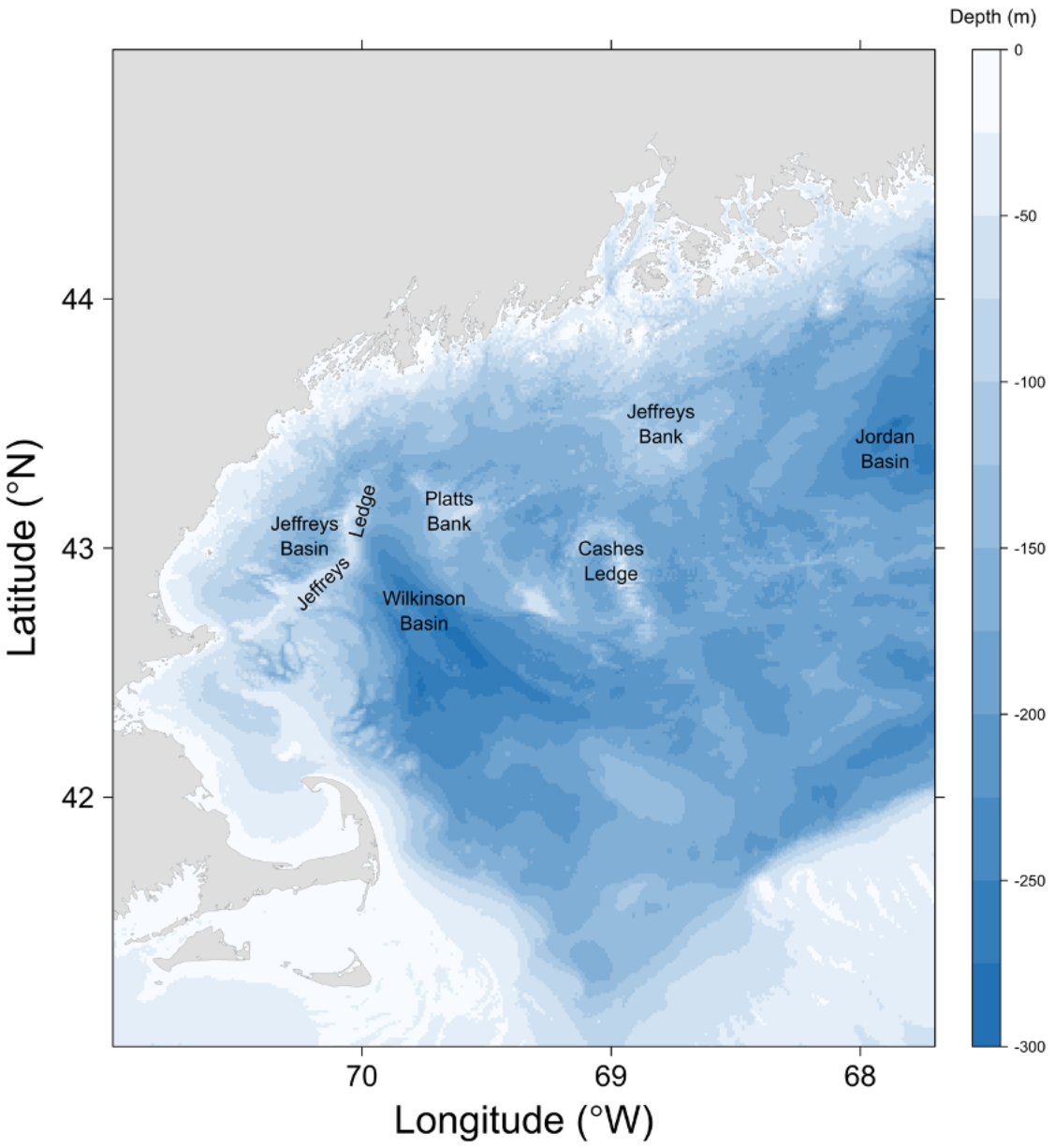


Figure 5.2. A map of bathymetry of the Gulf of Maine. (Data resources: NOAA National Centers for Environmental Information)

Biological data including carapace length and life-stage were collected from NEFSC fall surveys and ASMFC summer surveys. male, ovigerous female, female I, and female II data were used in this study. Male shrimp were further classified into recruits and mature males based on a

cutoff carapace length for recruitment in summer and fall each year (Hunter personal comm.). Because samples of shrimp actually undergo the transition from males to females in winter, relatively few shrimp at transitional stage were available in summer and fall. Therefore, data on transitional stage shrimp were not used. female I shrimp were defined as mature females that had not yet spawned, which can be distinguished morphologically by sharply pointed abdominal spines that are lost after the first spawning (McCrary 1971). Ovigerous (or, egg-bearing) females were defined as shrimp carrying eggs on their pleopods. Female II shrimp were identified by their vestigial abdominal spines, which indicate that they have spawned at least once before (McCrary 1971). Spawning occurs in deep waters offshore in late summer and fall, therefore most females sampled in the fall surveys were ovigerous, and we did not include the few female I and female II shrimp captured in our analysis. For each tow, a maximum of 2-kg shrimp were taken as subsamples. Data of life-stage and carapace length composition were collected for these subsamples. Data of the remainder of samples (unstaged) collected in a tow were expanded for each tow based on carapace length distribution and proportion of life-stage of staged data collected from that tow.

If a tow caught no shrimps in a stratum where any shrimp (season and life-stage specific) was present, the tow was considered a zero-catch tow. Tows in strata where shrimp has never been caught were removed. Datasets were processed season and life-stage specifically.

Catch per unit effort (CPUE) for each life stage was standardized as number of shrimp per 20 minutes tow duration. The average standardized CPUE of a stratum was weighted by stratum area (nmi^2) to derive the overall stratified mean (Reuchlin-Hugenholtz et al. 2015):

$$\bar{CPUE}_{k,s} = \frac{\left(\frac{\sum_{n=1}^{n_s} CPUE_{k,s,n}}{n_{k,s}}\right)\frac{A_{k,s}}{\sum A_{k,s}}}{\sum\left[\left(\frac{\sum_{n=1}^{n_s} CPUE_{k,s,n}}{n_{k,s}}\right)\frac{A_{k,s}}{\sum A_{k,s}}\right]}$$

The stratified bottom temperature data of a stratum was weighted by stratum area:

$$BOTTEMP_k = \frac{\sum_{s=1}^{n_s} \overline{BOTTEMP}_{k,s}\, A_{k,s}}{\sum_{s=1}^{n_s} A_{k,s}}$$

where $\overline{BOTTEMP}$ = average bottom temperature in a stratum, k = year, s = stratum, A = area

Spatial aggregation of each life-stage was evaluated using modified Camargo's index (Camargo 1995, Payne et al. 2005):

$$E = 1 - \left[\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(\frac{\left|CPUE_i - CPUE_j\right|}{n}\right)\right]$$

where E=Camargo's index, n = number of tows, $CPUE_i$ and $CPUE_j$ = CPUE at i^{th} location and j^{th} location. This evenness index value ranges from 0 to 1, indicating patchy to even distribution. This indicator is relatively unaffected by different types of sampling (e.g., random and non-random) and is able to generate unbiased estimates (Payne et al. 2005).

Each tow is classified into low and high CPUE. We defined low CPUE as CPUE < 25% quantile of the entire time series (season and life-stage specific), and high CPUE as CPUE > 25% quantile of the entire time series based on preliminary analyses described below. Proportion of low/high CPUE tows in a stratum is considered low/high density area (LDA/HDA). These density areas (DAs) are weighted by CPUE and area of strata (Reuchlin-Hugenholtz et al. 2015):

$$DA_k = \frac{\sum_{s=1}^{k_s} DA_s A_s \bar{CPUE}_s}{\sum_{s=1}^{k_s} A_s \bar{CPUE}_s}$$

A preliminary analysis was conducted to determine the boundaries of low/high CPUE. LDAs and HDAs were calculated using a range of quantiles (0%, 10%, 20%, 25% and 33%), and the time series at each threshold examined graphically. These LDAs and HDAs had similar trends. However, at the 10% quantile LDAs were 0 for at least a few years, and HDAs were 0 for several years after the stock collapsed. DAs at 20%, 25%, and 33% were highly correlated, and 20% and 25% almost overlapped. Therefore, 25% was selected. Zero density areas (ZDA) were almost constant over years for most life-stages in the two seasons except for Summer Young Male. Yet, ZDAs of Summer Young Male are highly correlated with LDAs at 25%. Therefore, ZDA was not used.

Based on DDHS, we hypothesized that LDAs would increase and HDAs would decrease when the population was declining, as areas with optimal conditions were selected by shrimp and the population density in suboptimal areas would decrease.

The center of gravity (CG) and its variance (termed inertia) were estimated for each life-stage in each season to examine possible changes in distribution. The CG was defined as (Woillez et al. 2009):

$$CG_k = \frac{\sum\left[\left(\frac{\sum_{n=1}^{n_s} CG_{k,s,n} CPUE_{k,s,n}}{\sum_{n=1}^{n_s} CPUE_{k,s,n}}\right) A_{k,s} \bar{CPUE}_{k,s}\right]}{\sum A_{k,s} \bar{CPUE}_{k,s}}$$

And inertia (I) (Woillez et al. 2009) as

$$I_k = \frac{\sum_{n=1}^{n_s} \left(L_{k,s,n} - CG_k\right)^2 A_{k,s} \bar{CPUE}_{k,s}}{\sum A_{k,s} \bar{CPUE}_{k,s}}$$

where L is the nth location in stratum s in year k. I_k consists of variances of x and y axes (longitude and latitude). The variability around CG was defined as the product of the squared root of I_k multiplied by p as:

$$I_{CG} = \pi\sqrt{I_x I_y}$$

5.3.2 Data analysis

Time series of estimated annual Camargo's evenness indices, proportions of LDA and HDA, latitude and longitude of CG, and inertia were examined graphically, and linear regression models and segmented linear models were used to explore trends of temporal changes and change points indicating possible regime shifts (Muggeo 2003; 2017; Friedland et al. 2020). Change points of segmented linear models were estimated using functions in the "strucchange" R-package (Zeileis et al. 2002).

The relationships between the spatial distribution indicators and abundance indices and abundance-weighted bottom temperature were further examined using multiple linear regression models where the spatial distribution indicators were response variables, and bottom temperature and abundance were predictor variables in each model. Before fitting the models, variance inflation factor (VIF) analyses were conducted for evaluating possible multicollinearity. Only predictors with VIFs <3 were included in the linear regression models (Schmiing et al. 2013; Brosset et al. 2019). A predictor was excluded from the model if it was not significant ($P > 0.05$). Once the significant predictor variables were identified, a set of autoregressive integrated moving average (ARIMA) models were examined for residual structures, including multiple linear regression (not correcting for residual structures), ARIMA (0, 0, 1), ARIMA (1, 0, 0), ARIMA (1, 0, 1), ARIMA (0, 1, 1), ARIMA (0, 1, 0), and ARIMA (1, 1, 0) where the first number in the parentheses is AR order, the second is degree of differencing, and the third is MA order. Autocorrelation of residuals was not found for the best models (see Results), higher order of lags were therefore not used for the residual structure. Model selection for ARIMA residual structures were based on Akaike information criterion (AIC), R-squared (R^2), and root mean squared error (RMSE). Models with the highest R^2 and lowest AIC and RMSE were considered the best. Normality and autocorrelation of residuals were graphically examined using QQ-plots and temporal autocorrelation plots for the time series. All analyses were performed using R version 4.0.3 (R Core Team 2020).

Accumulated percentage of stations and percentage of northern shrimp (life stage specific) at ranked bottom temperature were used to further examine if northern shrimp had a preference for colder water temperature. A logistic model was used to describe the relationship between accumulative percentage of stations (or shrimp) and bottom temperature:

$$P(t) = \frac{1}{1 + e^{-k(t-t_{50})}}$$

where t is bottom temperature; P(t) is accumulative percentage of stations (or shrimp) at a given bottom temperature t, k is the steepness of the logistic curve, t50 is the bottom temperature where 50% of the stations (or shrimp) had bottom temperature lower (or higher) than t50, i.e. the bottom temperature of the midpoint of the curve.

The t50 for number of stations, the bottom temperature where 50% of the stations had lower values of bottom temperature, was then considered unweighted bottom temperature. The t50 for number of shrimp, the bottom temperature where 50% of the shrimp were found had lower values of bottom temperature, was then considered abundance weighted bottom temperature. The values of t50 for each year were used to evaluate the differences between unweighted and abundance weighted bottom temperature (difference = abundance weighted – unweighted).

5.4 Results

5.4.1 Surveys

The numbers of stations sampled for each survey (including zero-catch) are given graphically by year in Fig. 5.3. The NEFSC fall bottom trawl surveys had higher numbers of stations sampled due to its larger survey area (Fig. 5.1). The bottom temperatures for all life stages in summer and fall and the estimated abundance index were shown in Fig. 5.4. The

bottom temperatures were relatively higher in the past 15 years (Fig. 5.4a). The estimated abundance index was relatively stable before 2004, and after reaching a high peak in 2006 it started decreasing until the population collapsed in 2012 (Fig. 5.4b). The time series of mean day of year for summer and fall bottom trawl surveys were shown in Fig. 5.5. The average mean day of year for summer surveys ranged 210-220 during 1984-2000, and it ranged 200-210 during 2000-2019 (Fig. 5.5a). For fall surveys, the average mean day of year ranged 290-305 during 1991-2010, and it ranged 305-320 after 2010 (Fig. 5.5b).

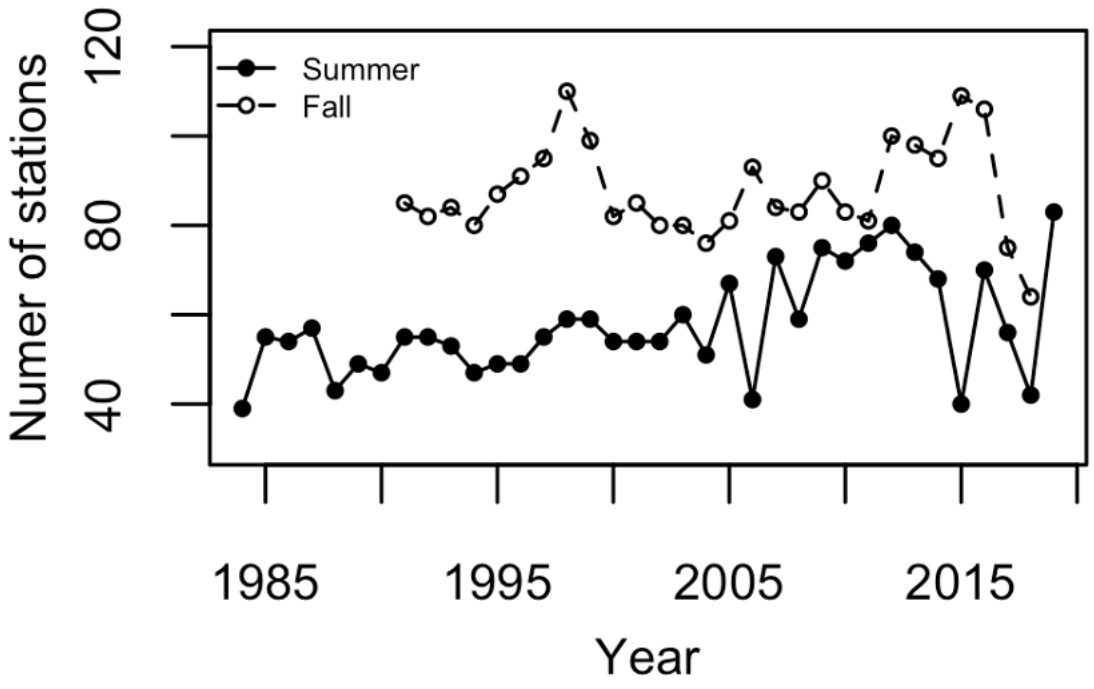


Figure 5.3. Number of stations sampled by Atlantic States Marine Fisheries Committee (ASMFC) summer shrimp bottom trawl surveys in 1984-2019 and Northeast Fisheries Science Center (NEFSC) fall bottom trawl surveys in 1991-2018.

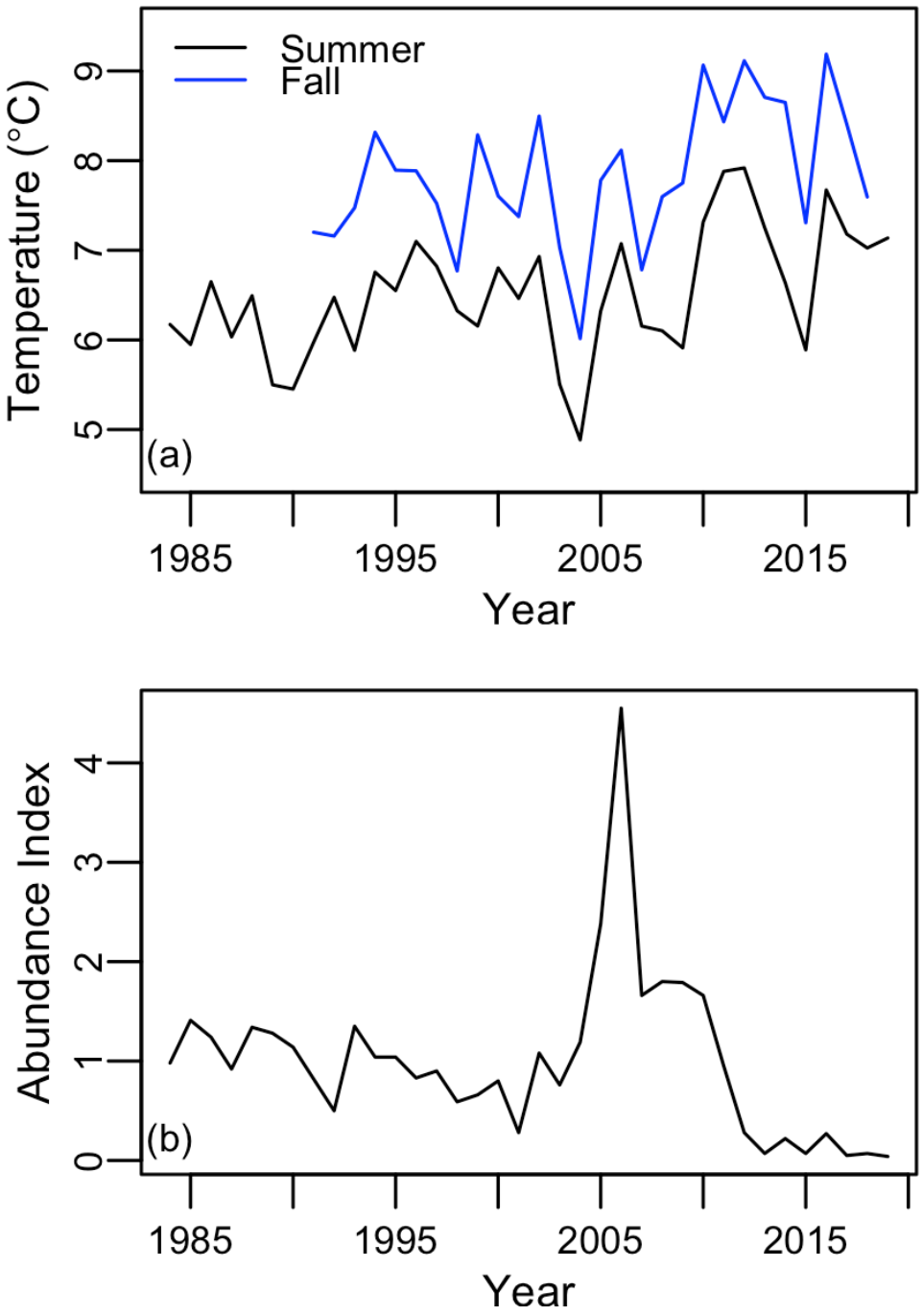


Figure 5.4. Time series of (a) bottom temperature in summer (black line) and fall (blue line) and (b) abundance index of the Gulf of Maine northern shrimp population.

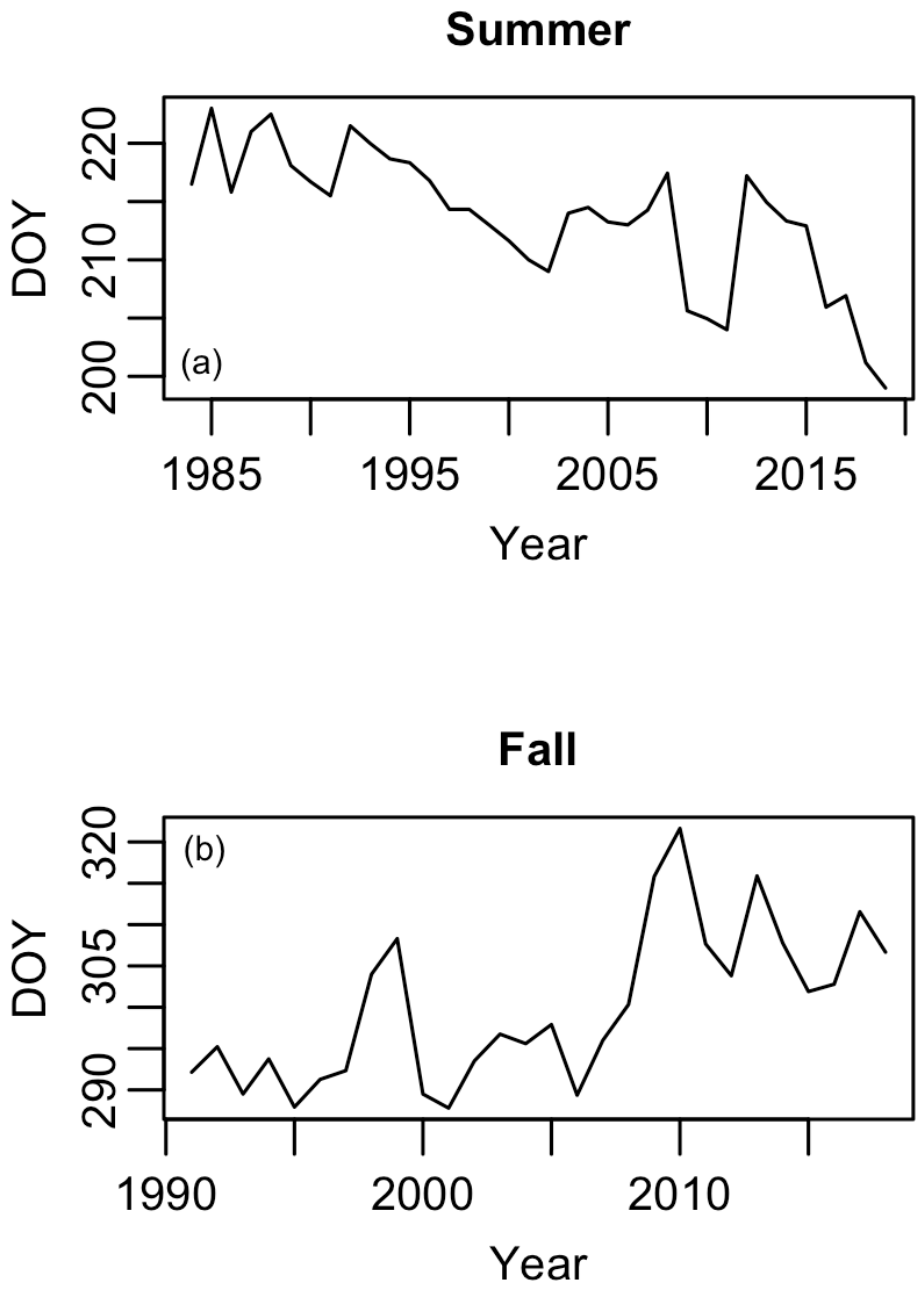


Figure 5.5. Time series of mean day of year(DOY) for Atlantic States Marine Fisheries Committee (ASMFC) summer shrimp bottom trawl surveys in 1984-2019 and Northeast Fisheries Science Center (NEFSC) fall bottom trawl surveys in 1991-2018.

5.4.2 Summer

The temporal changes of CG are shown in Fig. 5.6. The CG of summer mature groups concentrated in the Platts Bank, shifting toward the shore over time (roughly decadal, Fig. 5.6). While the summer recruits had different distributions from the mature groups, which concentrated in the Jeffreys ledge in the western GOM (Fig. 5.6).

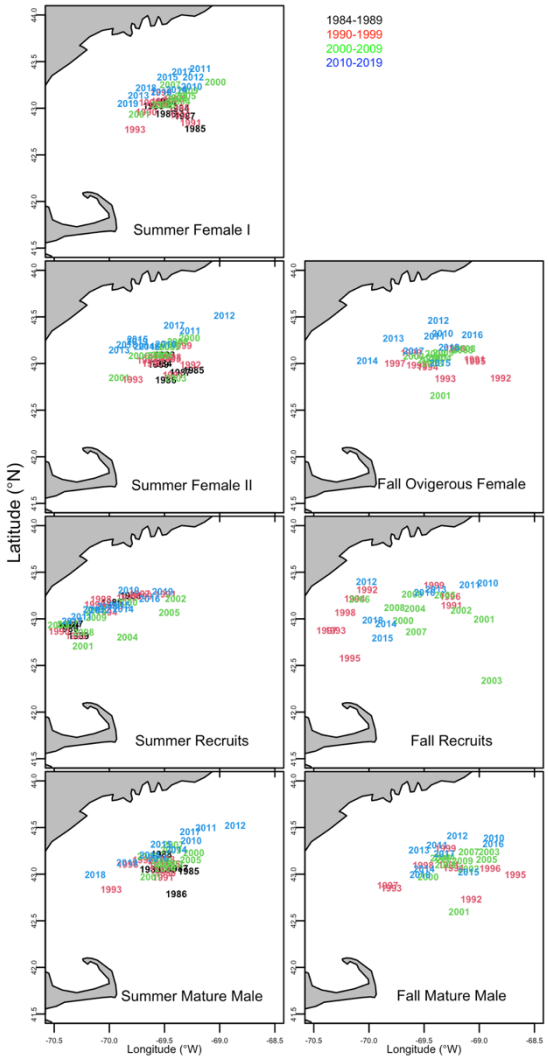


Figure 5.6. Maps of center of gravity (CG) of each year for different life stages in summer and fall.

The summer mature groups (female I, female II, and mature males) had similar patterns and trends in the time series of Camargo's evenness index and latitude of CG with significant linear trends (negative slopes for Camargo's evenness index and positive slopes for latitude of CG, Figs. 5.7-5.9), indicating patchier distribution and northward shifts of CG over time for all summer mature groups.

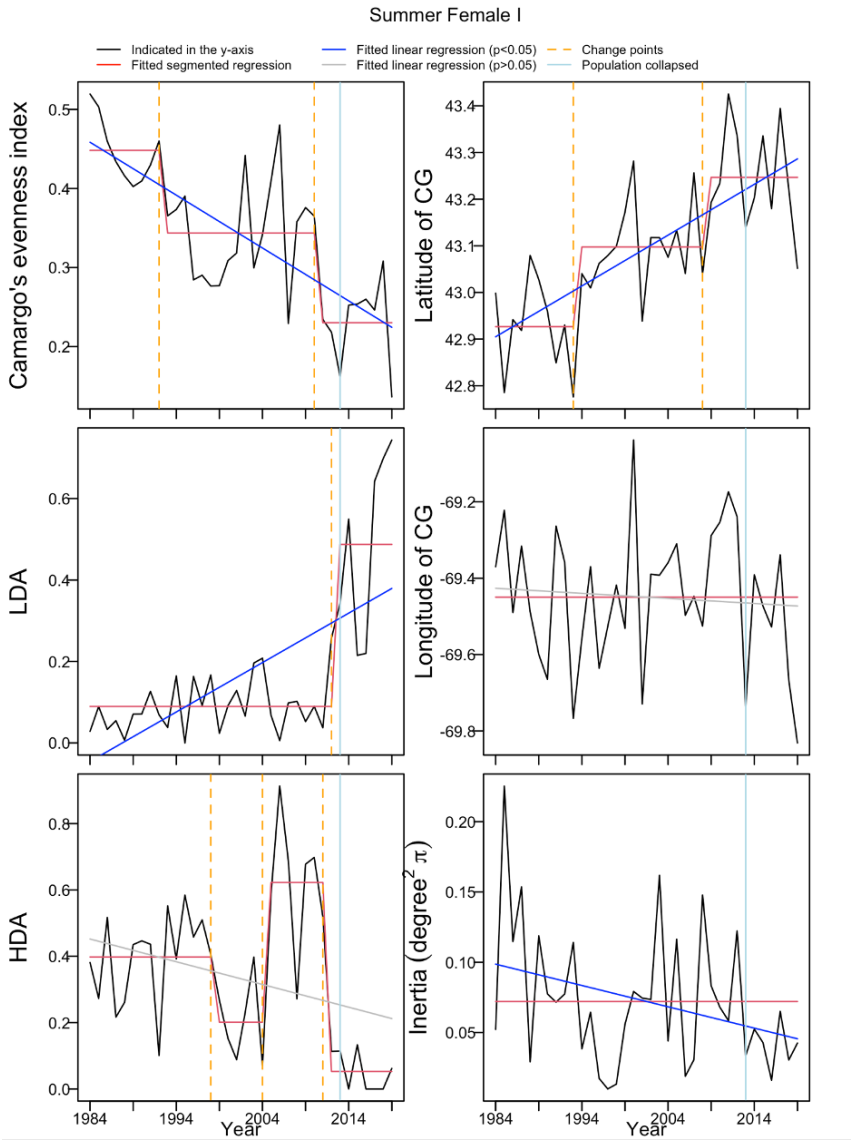


Figure 5.7. Time series of Camargo's evenness index, proportion of LDA and HAD (low and high density areas), latitude and longitude of CG (center of gravity), and Inertia for summer female I of the Gulf of Maine northern shrimp. The blue (significant, $p<0.05$) and gray (nonsignificant, $p>0.05$) lines were fitted linear regression models. The orange dashed lines denoted change points identified by the segmented regression models. The lightblue lines denoted the year (2013) when the northern shrimp population collapsed.

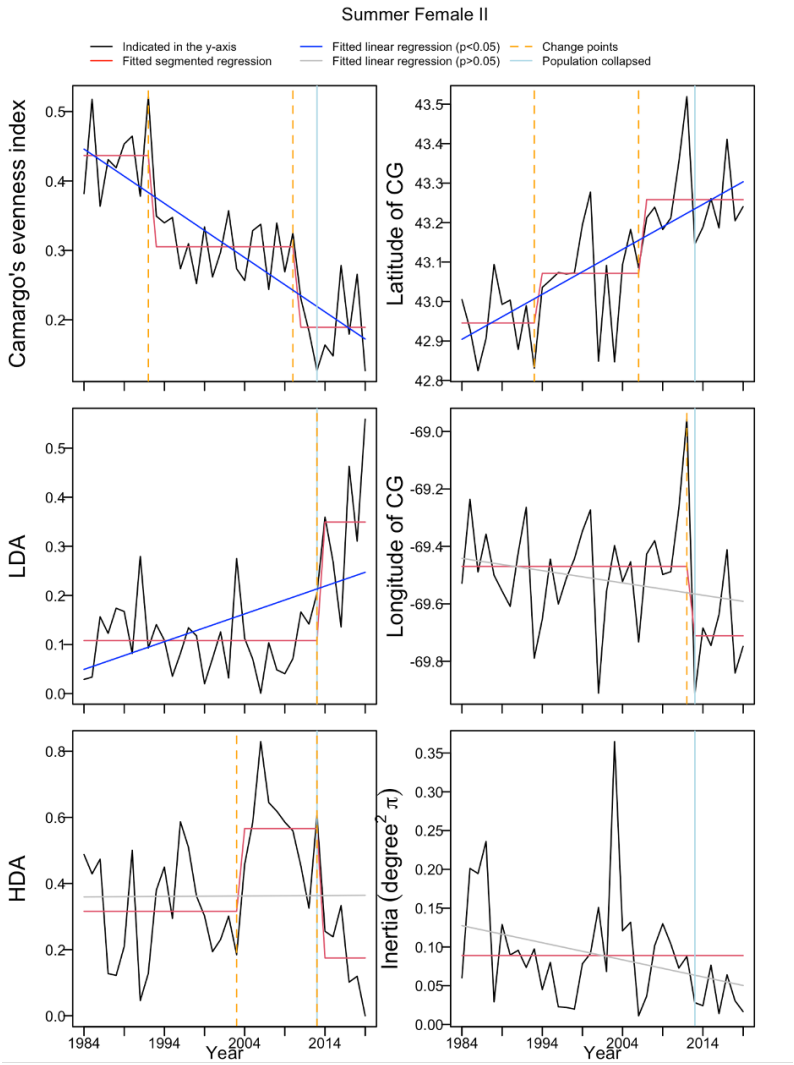


Figure 5.8. Time series of Camargo's evenness index, proportion of LDA and HAD (low and high density areas), latitude and longitude of CG (center of gravity), and Inertia for summer female II of the Gulf of Maine northern shrimp. The blue (significant, $p<0.05$) and gray (nonsignificant, $p>0.05$) lines were fitted linear regression models. The orange dashed lines denoted change points identified by the segmented regression models. The lightblue lines denoted the year (2013) when the northern shrimp population collapsed.

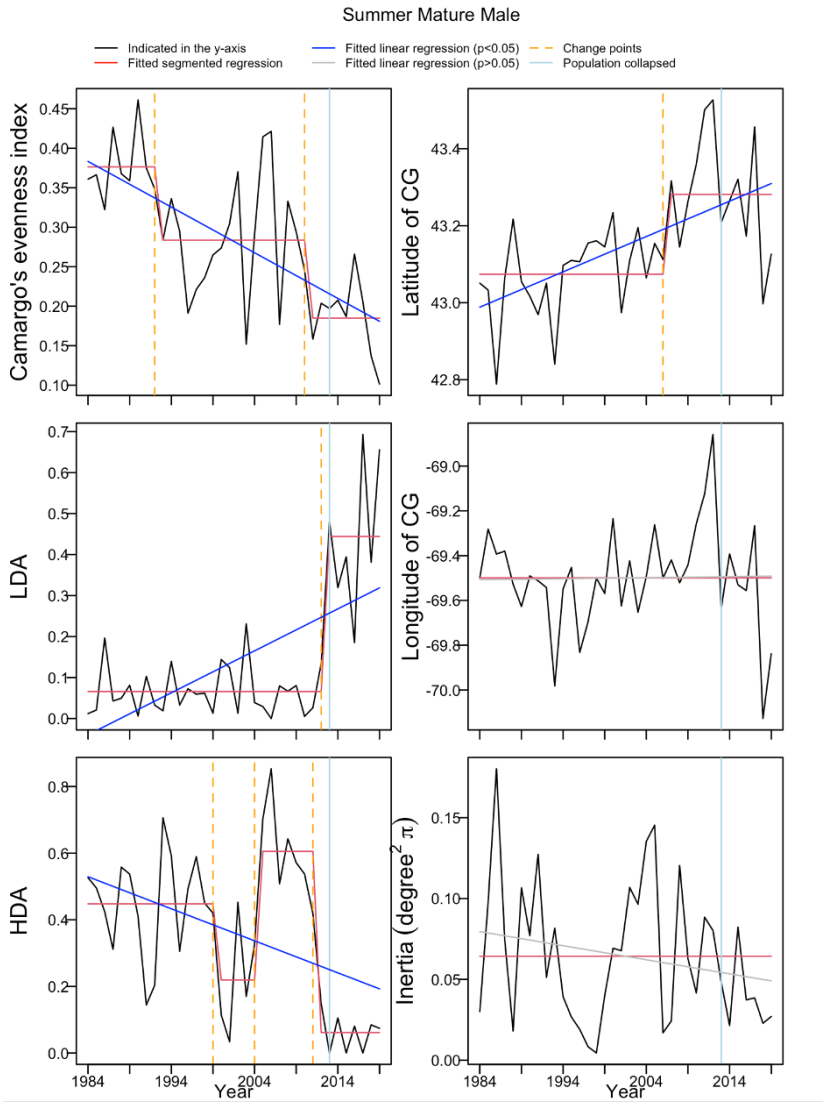


Figure 5.9. Time series of Camargo's evenness index, proportion of LDA and HAD (low and high density areas), latitude and longitude of CG (center of gravity), and Inertia for summer mature males of the Gulf of Maine northern shrimp. The blue (significant, $p<0.05$) and gray (nonsignificant, $p>0.05$) lines were fitted linear regression models. The orange dashed lines denoted change points identified by the segmented regression models. The lightblue lines denoted the year (2013) when the northern shrimp population collapsed.

The best models of each index for each life stage were shown in Table 5.1. The models suggested that the changes in Camargo's evenness index was positively correlated with population abundance for all summer groups, meaning the patchier distribution was associated with lower population abundance. The effects of bottom temperature and population abundance on Camargo's evenness index of summer recruits were not significant, and no patterns or trends were found in the time series of Camargo's evenness index for summer recruits.

Table 5.1. Best linear models of each spatial distribution indicators for each life stage in summer and fall. Log(Abund) and Bottemp are estimated parameters for log(Abundance index) and Bottom temperature. *: $p<0.05$; **: $p<0.01$; ***: $p<0.001$. ARIMA structure is the auto regressive integrated moving average structure of residuals for each model. AR = estimated auto regressive parameter; MA = estimated moving average parameter, RMSE = root mean squared error. OviF = Ovigerous Female.

	Log(Abund)	Bottemp	ARIMA structure	AR	MA	R^2	RMSE
Summer female I							
Camargo's	0.055***		(1, 0, 0)	0.510***	NA	0.572	0.062
LDA	–0.140***		(0, 0, 0)	NA	NA	0.695	0.105
HDA	0.154***		(0, 0, 0)	NA	NA	0.564	0.154
Latitude		0.044	(0, 1, 1)	NA	–0.722***	0.44	0.116
Longitude	0.050*		(0, 0, 0)	NA	NA	0.104	0.168
Inertia		–0.079*	(0, 0, 0)	NA	NA	0.124	0.145
Summer female II							
Camargo's	0.037**		(1, 1, 0)	-0.701***	NA	0.679	0.056
LDA	–0.083***		(0, 0, 0)	NA	NA	0.597	0.078
HDA	0.085**		(1, 0, 0)	0.407**	NA	0.416	0.149
Latitude		0.066*	(0, 1, 1)	NA	–0.755***	0.458	0.12
Longitude	0.050***		(1, 0, 1)	0.558***	–1.000***	0.308	0.166

Inertia		–0.155***	(1, 0, 0)	0.282	NA	0.318	0.162
Summer mature male							
Camargo's	0.049***		(0, 0, 1)	NA	0.352*	0.464	0.066
LDA	–0.135***		(0, 0, 0)	NA	NA	0.774	0.083
HDA	0.158***		(0, 0, 1)	NA	0.481**	0.746	0.117
Latitude		0.055	(0, 1, 1)	NA	–0.740***	0.315	0.131
Longitude	0.069*	0.114*	(0, 0, 0)	NA	NA	0.158	0.209
Inertia		–0.079*	(0, 0, 1)	NA	0.466*	0.244	0.133
Summer Recruits							
LDA	–0.175***		(0, 0, 1)	NA	–0.234	0.7	0.138
HDA	0.181***		(0, 0, 0)	NA	NA	0.51	0.201
Latitude		0.134**	(1, 0, 1)	–0.570*	0.858***	0.295	0.133
Fall ovigerous female							
Camargo's	0.02***		(0, 0, 0)	NA	NA	0.348	0.031
LDA	–0.046***		(0, 0, 0)	NA	NA	0.344	0.072
HDA	0.053	–0.002	(1, 0, 0)	0.727***	NA	0.571	0.155
Latitude		0.114***	(0, 0, 1)	NA	0.395*	0.443	0.12
Inertia	0.113***	–0.103*	(0, 0, 0)	NA	NA	0.476	0.171
Fall mature male							
HDA	0.101***	0.039	(1, 0, 0)	0.472**	NA	0.548	0.132
Latitude		0.083*	(0, 0, 1)	NA	0.625***	0.414	0.142
Inertia	0.091*		(0, 0, 0)	NA	NA	0.174	0.226
Fall Recruits							
LDA	–0.086**		(0, 0, 0)	NA	NA	0.238	0.176
Latitude		0.247***	(1, 0, 1)	0.512**	–1.000***	0.397	0.243

The changes in latitude of CG were significantly positively correlated with bottom temperature for summer female II and recruits, indicating the northward shifts of CG were associated with warmer temperature. Although no clear trends were found in the time series of longitude of CG for all summer groups (Figs. 5.7-5.10) and only one change point was identified for female II at 2012 (Fig. 5.8), the changes in longitude of CG were positively correlated with population abundance for summer mature groups (Table 5.1). The models suggested the

westward shifts of CG of summer mature groups were associated with lower population abundance.

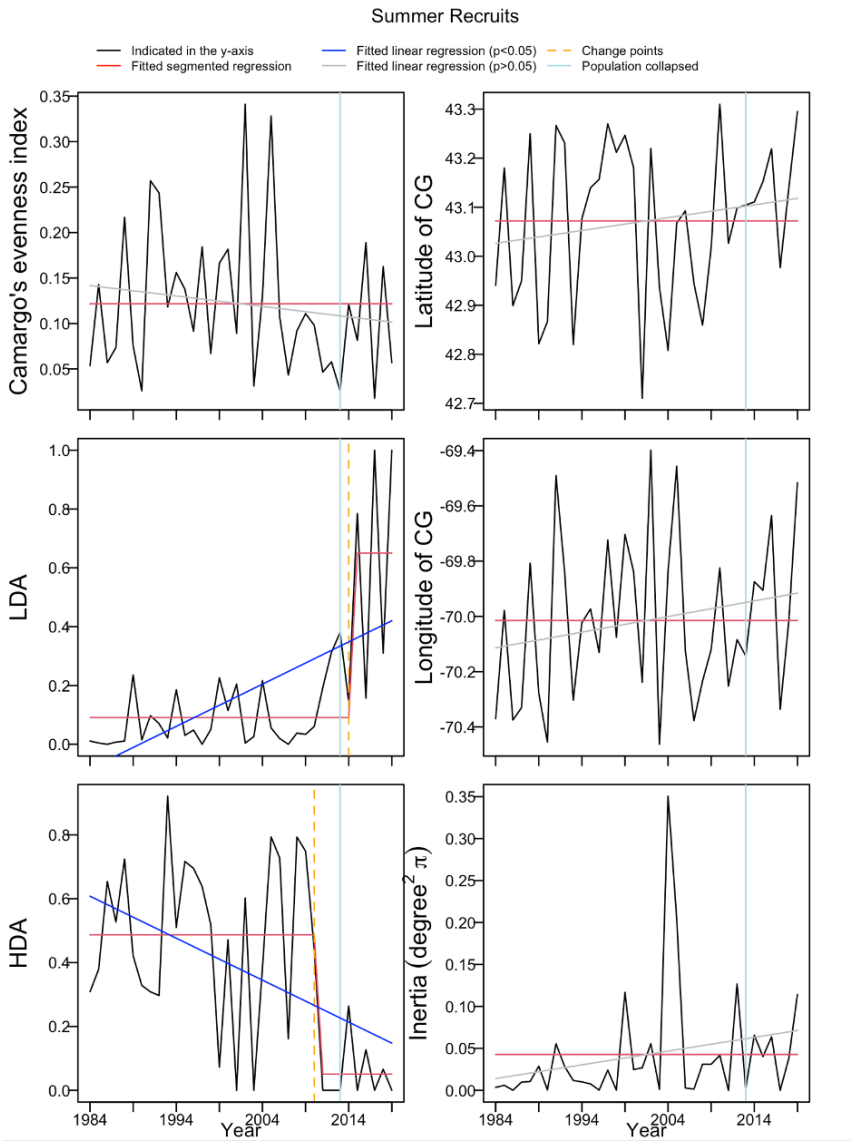


Figure 5.10. Time series of Camargo's evenness index, proportion of LDA and HAD (low and high density areas), latitude and longitude of CG (center of gravity), and Inertia for summer recruits of the Gulf of Maine northern shrimp. The blue (significant, $p<0.05$) and gray (nonsignificant, $p>0.05$) lines were fitted linear regression models. The orange dashed lines

denoted change points identified by the segmented regression models. The lightblue lines denoted the year (2013) when the northern shrimp population collapsed.

Two change points were identified by segmented regression models for the time series of Camargo's evenness index of the summer mature groups in 1992 and 2010 (Figs. 5.7-5.9). Before 1992, the Camargo's evenness index remained at 0.38-0.46 for the summer mature groups. During the period of 1993-2010, the averages of Camargo's evenness index of these three life stages decreased to the level of 0.28-0.34. After 2011, the averages of Camargo's evenness index of these three life stages decreased to an even lower level of 0.18-0.22. However, no trends or change points were found in the Camargo's index time series for summer recruits (Fig. 5.10). A summary of change points of the time series of the indices was shown in Fig. 5.11.

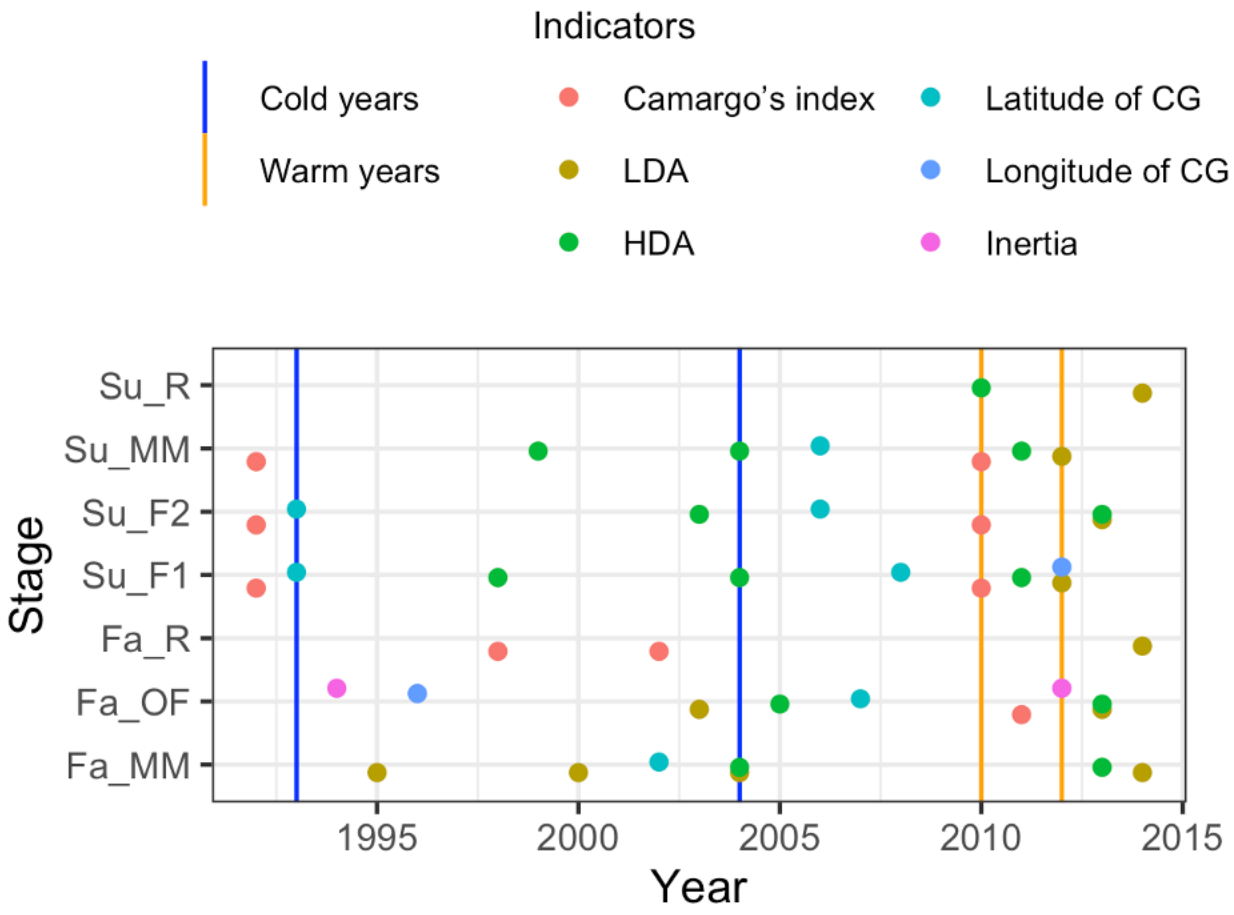


Figure 5.11. A summary plot of all the change points of each spatial distribution indicator identified by the segmented regression models for each life stage of the Gulf of Maine northern shrimp in summer and fall. The blue lines denoted two coldest years over the time series, and the orange lines denoted two hottest years over the time series.

Two change points were identified for the time series of latitude of CG of the two summer female groups (Figs. 5.7-5.8). The latitude of CG was at around 42.9°N before the first change point in 1993, it then increased to around 43.1°N before the second change point in 2008 and 2006 for summer female I and II, respectively (Figs. 5.7-5.8). It shifted northward to around 43.25°N since the second change points of the two summer female groups. One change point

was identified for summer mature male in 2006 (Fig. 5.9)-the latitude of CG was around 43.08°N before the first change point, and shifted northward to around 43.25°N since 2007.

As for Inertia (variability around CG), the decreasing linear trends of Inertia were significant only for summer female I with no change points identified (Fig. 5.7), indicating the variability around the CG decreased over time for summer female I. Although the patterns of Inertia were not clear and no linear trends or change points identified for other summer groups, it was noted that the Inertia of summer female II and recruits remained at low levels over almost the entire time series (<0.25 for female II and <0.15 for recruits) but reached a high peak around 0.35 in 2004 which was the coldest year in the time series (Figs. 5.4, 5.8 and 5.10). Furthermore, the model suggested that the changes in Inertia for summer mature groups were negatively correlated with bottom temperature (Table 5.1), with warmer temperature associated with lower variability around the CG.

The linear trends were significantly positive for all LDA time series of summer groups (Figs. 5.7-5.10), indicating more and more proportions of LDA over time for all summer groups. On the other hand, the linear trends of HDA time series were significantly negative for summer male groups but not for female groups (Figs. 5.7-5.10), indicating less and less proportion of HDA for summer male groups over time.

For all summer groups, the models suggested that the changes in LDA were negatively correlated with population abundance and the changes in HDA were positively correlated with population abundance (Table 5.1), with decreasing population abundance associated with more LDA and fewer HDA. While the bottom temperature effects were not significant on both LDA and HDA (Table 5.1).

The LDA of summer groups remained at low level with low variations until the change points in 2012-2014 where the population collapsed, while after the change points the LDA increased dramatically from below or around 0.1 to above 0.3-0.6 (Figs. 5.7-5.10).

Three change points were identified for summer female I and mature male HDAs (Figs. 5.7 and 5.9). The HDAs remained at around 0.4 before the first change point in 1998 and 1999 for summer female I and mature male, respectively. The HDAs of these two life stages decreased to around 0.2 during the first two change points, and increased to around 0.6 during 2005-2011 before it dropped to a very low level below 0.1 after 2012.

No autocorrelations were found in the residuals of all summer models after accounting for the ARIMA structures (Figs. B.1-B.4).

In summary, the distributions of summer mature groups were getting patchier and shifting northward over time. The changes in distribution evenness, proportions of LDA and HDA were associated with population abundance, while the shifts of CG were associated with bottom temperature for summer groups.

5.4.3 Fall

In fall, the mature groups (ovigerous female and mature male) concentrated in the Platts Bank and Cashes Ledge areas (Fig. 5.6). While the recruits migrate from Jeffreys ledge in summer to more offshore areas toward Platts Bank. The shifts of CG toward the shore over time were also observed for fall mature groups (Fig. 5.6).

Similar to summer mature groups, significant northward shifts of CG over time (Figs. 5.12-5.13) were also found for the fall mature groups. The changes in the latitude of CG were positively correlated with abundance weighted bottom temperature for fall mature groups (Table

5.1), with warmer temperature associated with northward shifts of CG. One change point was identified for fall ovigerous female and mature male in 2007 and 2002, respectively (Figs. 5.12-5.13). Before the change point, the latitude of CG for fall mature groups was around 43°N, it shifted to around 43.2°N after the change point. No significant linear trends were found in the time series of longitude of CG for fall mature groups, although one change point was identified in 1996 for ovigerous female.

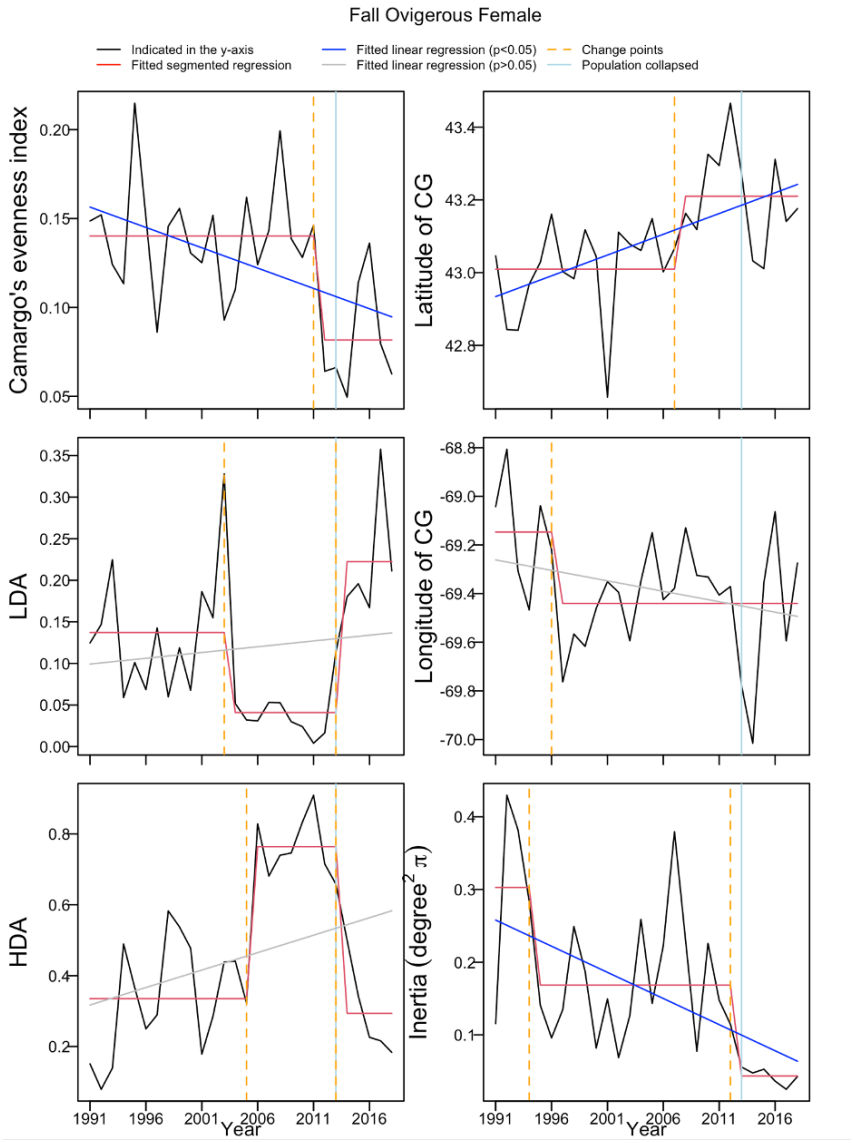


Figure 5.12. Time series of Camargo's evenness index, proportion of LDA and HAD (low and high density areas), latitude and longitude of CG (center of gravity), and Inertia for fall ovigerous female of the Gulf of Maine northern shrimp. The blue (significant, $p<0.05$) and gray (nonsignificant, $p>0.05$) lines were fitted linear regression models. The orange dashed lines denoted change points identified by the segmented regression models. The lightblue lines denoted the year (2013) when the northern shrimp population collapsed.

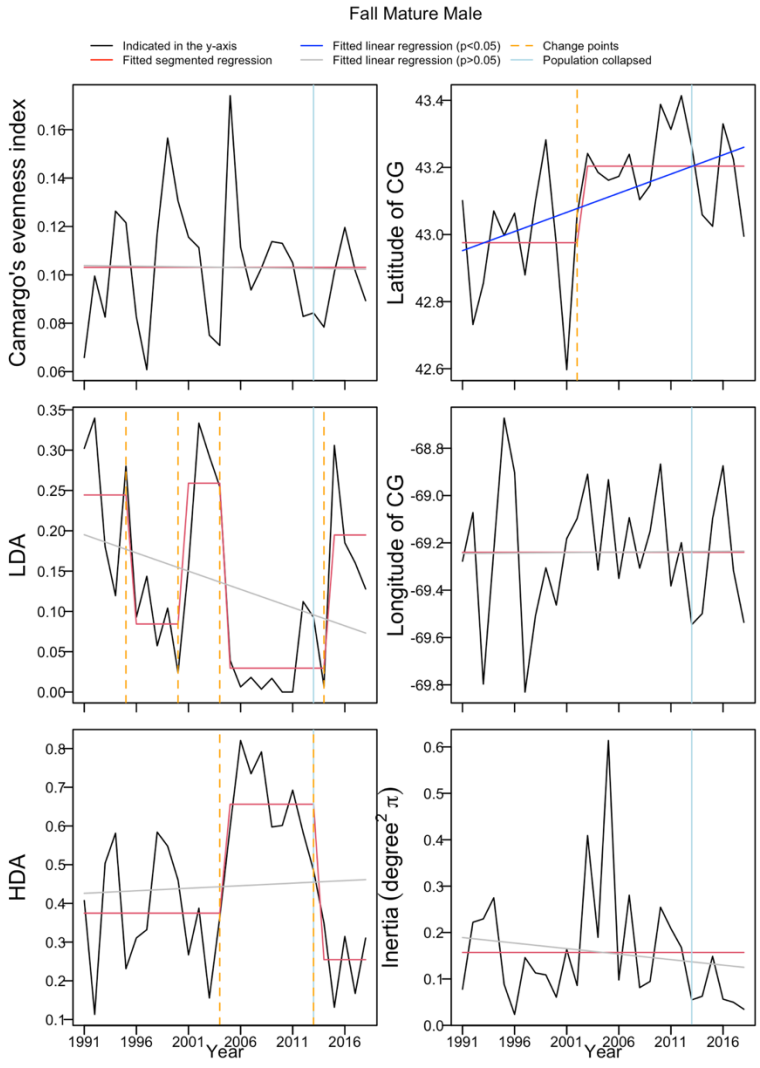


Figure 5.13. Time series of Camargo's evenness index, proportion of LDA and HAD (low and high density areas), latitude and longitude of CG (center of gravity), and Inertia for fall mature male of the Gulf of Maine northern shrimp. The blue (significant, $p<0.05$) and gray (nonsignificant, $p>0.05$) lines were fitted linear regression models. The orange dashed lines denoted change points identified by the segmented regression models. The lightblue lines denoted the year (2013) when the northern shrimp population collapsed.

With regards to Inertia, a significant decreasing linear trend was found for fall ovigerous female, indicating the variability around the CG decreased over time for ovigerous female (Fig. 5.12). Two change points in 1994 and 2012 were identified for the time series of Inertia of ovigerous female (Fig. 5.12). The Inertia were at moderate level around 0.17 and decreased to a very low level around 0 after 2013. Although no linear trends or change points were found in the time series of Inertia for summer mature males (Fig. 5.13), the models suggested that the changes in Inertia were positively correlated with population abundance for fall mature groups (Table 5.1), with lower variability around CG associated with lower population abundance. In addition to abundance effects, the model indicated that the Inertia of ovigerous females was negatively correlated with bottom temperature (Table 5.1), with warmer temperature associated with lower variability around CG.

For distribution evenness in the fall, significantly decreasing Camargo's evenness index (i.e. patchier distribution) over time was found for ovigerous female (Fig. 5.12), which was associated with declining population abundance (Table 5.1). One change point was identified in 2011 for the time series of Camargo's evenness index of fall ovigerous female (Fig. 5.12). After 2011, the Camargo's evenness index of fall ovigerous female decreased from 0.14 to below 0.10.

No significant linear trends were found for the LDA time series of fall mature groups. Similar to summer groups, the models suggested that the changes in LDA of fall ovigerous females and recruits were negatively correlated with population abundance (Table 5.1), with higher proportions of LDA associated with lower population abundance.

The linear trends of HDA time series were not significant for fall mature groups. In contrast to the significant negative effects of population abundance on LDA for ovigerous

female, both effects on HDA were not significant for fall ovigerous female after accounting for the ARIMA structure for the residuals (Table 5.1). As for fall mature males, the changes in HDA were positively correlated with population abundance (Table 5.1), with higher proportions of HDA associated with higher population abundance. The bottom temperature effects on HDA of fall mature male were positive but not significant after accounting for the ARIMA structure for the residuals (Table 5.1).

Two to four change points were identified for LDA and HDA of fall mature groups (Figs. 5.12-5.13). The time series of LDA and HDA of fall mature groups showed a roughly reversed pattern that LDA was at a lower level and HDA was at a higher level during 2006-2012 (Figs. 5.12-5.13).

As for fall recruits, although one and two change points were identified for Camargo's evenness index and LDA, respectively, the patterns of the time series of these indices were not clear and no linear trends were found in these time series (Fig. 5.14).

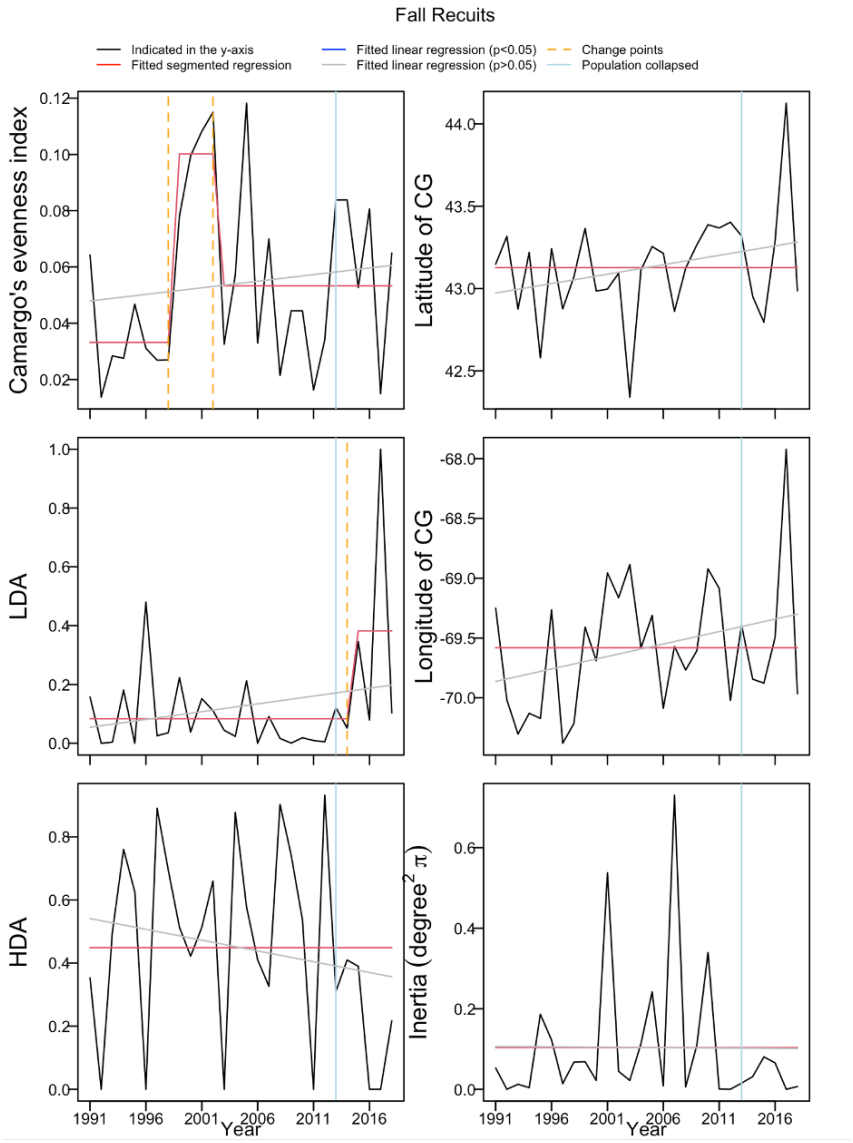


Figure 5.14. Time series of Camargo's evenness index, proportion of LDA and HAD (low and high density areas), latitude and longitude of CG (center of gravity), and Inertia for fall recruits of the Gulf of Maine northern shrimp. The blue (significant, $p<0.05$) and gray (nonsignificant, $p>0.05$) lines were fitted linear regression models. The orange dashed lines denoted change points identified by the segmented regression models. The lightblue lines denoted the year (2013) when the northern shrimp population collapsed.

No autocorrelations were found in the residuals of all fall models after accounting for the ARIMA structures (Figs. B.5-B.7).

In summary, similar to summer mature groups, the CG of fall mature groups were also shifting northward. However, the distribution was getting patchier only for ovigerous females. The northward shifts of CG were correlated with bottom temperature, while the changes in distribution evenness, proportions of LDA and HDA were more associated with population abundance,

The unweighted and abundance weighted bottom temperatures were shown in Fig. 5.15. The abundance weighted bottom temperature was generally lower than the unweighted bottom temperature for all life stages in summer and fall (Fig. 5.15).

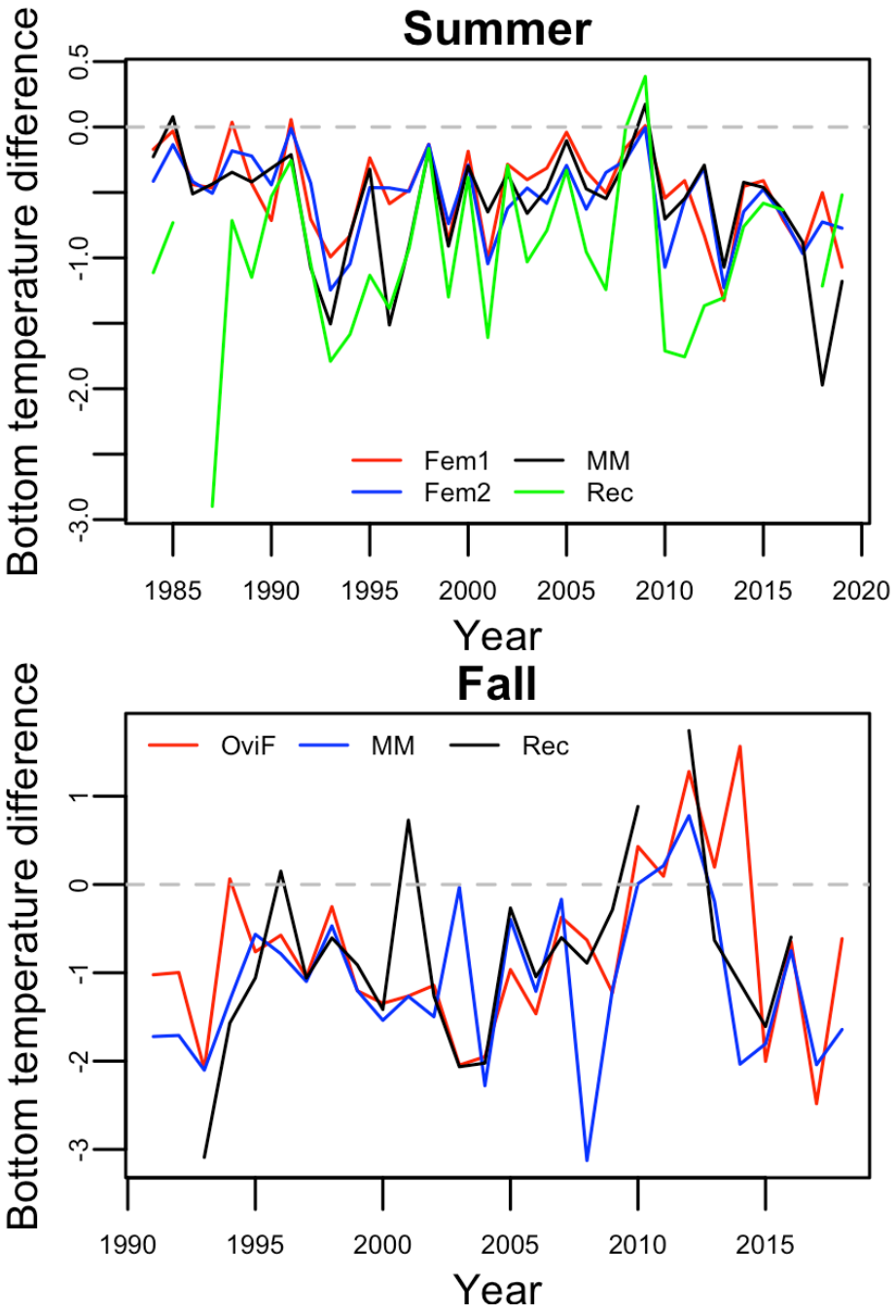


Figure 5.15. Differences of unweighted and abundance weighted bottom temperatures for summer and fall(differences = abundance weighted - unweighted). Fem1 = female I, Fem2 = female II, MM = mature male, Rec = recruits, OviF = ovigerous female.

The differences (abundance weighted – unweighted) ranged -1.97-0.17 for summer mature groups (Table 5.2 and Fig. 5.15). The differences were larger for summer recruits, with a mean of difference of -0.97 (Table 5.2 and Fig. 5.15), meaning the abundance weighted bottom temperature was lower than unweighted bottom temperature by around 1 °C for summer recruits.

Table 5.2. Summary of differences of unweighted and abundance weighted bottom temperature for each life stage

(SD = standard deviation).

Summer	Minimum	Maximum	Mean	SD
Female I	-1.325	0.057	-0.484	0.345
Female II	-1.245	-0.006	-0.537	0.321
Mature male	-1.972	0.173	-0.595	0.452
Recruits	-2.899	0.386	-0.966	0.630
Fall				
Ovigerous female	-2.482	1.564	-0.803	0.962
Mature male	-3.126	0.779	-1.112	0.881
Recruits	-3.090	1.742	-0.869	1.016

The differences between unweighted and abundance weighted bottom temperatures for fall groups were larger than summer groups (Table 5.2 and Fig. 5.15). The abundance weighted bottom temperatures were lower than unweighted bottom temperature by from 0.8 °C (for fall ovigerous females) to 1.1 °C (for fall mature males) (Table 5.2 and Fig. 5.15). Some estimates of the parameters were missing for some years as the models did not converge.

5.5 Discussion

Our results suggest that evenness, LDA and HDA are correlated with abundance for almost all life stages in summer and Fall, while shifts in CG were more bottom temperature related. These indicators provided different information on the spatial structure of the distribution. The Camargo's evenness index evaluates the evenness or patchiness of the distribution, implying how aggregating a species or a life stage may be. LDA and HDA can also be indicators of distribution evenness, providing different information on spatial structure (proportion of low and high density areas). It is intuitive that the LDAs are negatively correlated with abundance and HDAs are positively correlated with abundance as, based on DDHS, the proportion of LDA would increase when the population decline and the proportion of HDA would increase when the population abundance increase. The indicators of CG and Inertia quantify the possible distributional shifts and variation of the population distribution.

Our results showed that the distributions were getting patchier over time for mature groups except fall mature males. The changes in Camargo's evenness index could be explained by the DDHS for the mature groups (except fall mature males) as the patchier distributions were associated with declining population abundance, not bottom temperature. Intuitively, the declining population abundance also led to changes in LDA and HDA, consequently resulting in patchier distribution.

The information of LDA and HDA together is consistent with DDHS theory predictions that when population abundance is low, individuals occupy the optimal habitat; but when population abundance is high, individuals spread into suboptimal habitat (Thorson et al. 2016). Our results showed that the temporal changes of LDA and HDA were roughly a counterpart of each other for fall mature groups. The higher levels of LDA and lower levels of HDA during

2006-2012 reflected higher levels of population abundance during 2005-2010. The beginning of this period occurred right after the coldest year (2004) over the time series (Fig. 5.11), and the higher levels of HDA ended when it entered the so-called warm years after 2010 (hottest in 2012 over the time series). Change points of HDA time series were also identified around these cold and warm years for mature groups in both summer and fall (Fig. 5.11), implying that the HDA entered into a different level after the change points. Nevertheless, for summer mature groups, the temporal changes of LDA and HDA had different patterns and were not a reversed counter part of each other. The LDA of summer mature groups remained at low levels and started to surge to high levels when the population started to collapse in 2013, and change points of LDA were also identified one year before 2013 for summer female I and mature males (Figs. 5.7 and 5.9). Therefore, comparing to HDA, LDA of summer mature groups may be less sensitive to population abundance changes.

The differences between unweighted and abundance weighted bottom temperature showed that northern shrimp at different life stages had a preference for colder water temperature. However, it should be noted that the differences between unweighted and abundance weighted bottom temperature were positive values for fall ovigerous females in 2010-2014 which were the years with warmer temperature. i.e. more than half of the ovigerous females stayed in the locations where water temperatures were higher than the average temperature during the warm years (2010-2014). Furthermore, after the stock collapsed, the HDA of fall ovigerous female only dropped to a similar level as before the high HDA period (2006-2012), but a surge of LDA was still observed. It was thus hypothesized that fall ovigerous females may have certain requirement for habitat possibly for reproduction-when the population collapsed, individuals aggregate in the optimal habitat, which result in a delayed decline in HDA but rapid

surge in LDA. The results of Camargo's evenness index and Inertia also supported this hypothesis as the distribution of fall ovigerous female became patchy and the variability around the CG decreased to a very low level after the population collapsed.

In addition to patchier distribution, the CG of the mature groups shifted around 33km (~42.9-43.2°N) northward over time, which was correlated with warming temperature. The bottom temperatures in summer and fall were relatively higher in the past 15 years, as a result of greater influxes of Warm Slope Water relative to Cold Labrador Slope Water and/or Scotian Shelf Water on top of a baseline of warming waters globally (Townsend et al., 2015; Fig. 5.4). The temperature could have indirect effects on the northern shrimp population. Richards and Hunter (2021) reported that changes in distribution and migration phenology of predatory longfin squid during the marine heatwave of 2012 may have been major factors in the collapse of the GOM northern shrimp. The collapse of the population could result in changes in spatial structures such as CG shifts due to declined population abundance. However, the northward shifting of CGs started in the 1980s, which could not be explained by the invasion of longfin squid. Combining with the results of CG, I, Camargo's evenness index, LDA, and HDA, it is likely the warming temperature in summer in the southern areas of their current habitat prevents adult northern shrimp from migrating further southward. The distances of inshore-offshore migration therefore become shorter and shorter over time as the shoreline runs northeast to southwest, which means the adult shrimp remain increasingly closer to the shore over the years and their distribution become patchier. Dow (1981) reported that normal inshore migration of ovigerous females did not occur in the mid-1950s when temperatures were high in the GOM. However, Haynes and Wigley (1969) reported incomplete inshore migration of ovigerous females during low temperatures.

The northward shifting CG may also be explained by the changes in survey timing. The timing of summer surveys had been earlier by 10-20 days compared to the 1980s. Most mature groups migrate to offshore areas during spawning season which takes place in late summer and fall. The northward shifting CG might reflect the earlier timing of survey as shrimp are on their offshore migration prior to spawning season. On the other hand, the timing for fall surveys had been later by around 15 days. After spawning season, ovigerous females start their inshore migration. Therefore, the northward shifting CG might reflect later survey timing as ovigerous females migrate to inshore areas after spawning season. Nevertheless, the northward shifting CG was also found for fall mature males.

No matter what role water temperature plays in the shifting of CG of adult shrimp, our results showed that the CG of mature northern shrimp population concentrated in Platts Bank and was slowly shifting northward over time rather than leaving the GOM. Greater winter heat flux (both sensible and latent [evaporative] heat flux) from surface waters in the western Gulf, result from cold, dry air outbreaks nearer the North American continent. This leads to deeper winter convective mixing, and, in addition to less tidal mixing of deep waters in the western Gulf, results in generally colder bottom waters throughout much of western GOM (Townsend et al., 2006). Apollonio et al. (1986) indicated that the Jeffreys Basin in the western GOM was considered a refuge for northern shrimp, providing suitable temperature conditions to support offshore life stages of northern shrimp. In contrast to the western GOM, the stronger turbulence in the eastern GOM produces greater vertical instability and more through mixing of surface and bottom waters throughout the year, the bottom temperature in the eastern GOM is therefore slightly higher than the western GOM (Apollonio et al. 1986). Moreover, the higher water temperature in the southern GOM prevents northern shrimp from extending its geographic range

toward Cape Cod and Georges Bank (Haynes and Wigley 1969). Therefore, the distribution of northern shrimp is effectively limited to deeper waters (c. 120-150 m) in the western GOM.

The CG results in this study showed that summer recruits were found close to the shore around Jeffreys ledge and migrate toward Platts Bank and Cashes Ledge in the fall. These migrating recruits are likely immature males which mature in offshore waters (Apollonio et al. 1986). However, the fall recruits in our data may not be representative as they were only a small portion of young males in the population, which were not fully recruited to the survey areas. As a result, in some years the numbers of observations of fall recruits were less than 10 as the majority of young males remain in inshore waters. This may also be one of the reasons that the patterns of spatial distribution for fall recruits were found unclear or nonsignificant, and our observations of these recruits may reflect their ontogenetic migration.

This study aimed to examine the temporal changes in spatial distribution indicators which incorporated information from bottom trawl survey, associating the changes in spatial structure with population abundance and bottom temperature. These spatial indicators were correlated with population abundance but could provide different information on spatial structure of the population. Nevertheless, some phenomena remained unexplained such as the distribution getting patchier and shifting slowly northward since the 1980s. Plausibly, the patchier distribution might simply be a result of shifting CG which was associated with changes in bottom temperature. Despite the models suggestion that the shifts of CG were associated with bottom temperature, the mechanisms behind the correlation were not clear. Therefore, further investigations are needed for better understanding the driving mechanisms.

CHAPTER 6. HABITAT SUITABILITY MODELS USING SURVEY AND FINITE-VOLUME COMMUNITY OCEAN MODEL BOTTOM TEMPERATURE DATA FOR THE GULF OF MAINE NORTHERN SHRIMP (*PANDALUS BOREALIS*)

6.1 Abstract

The Gulf of Maine northern shrimp (*Pandalus borealis*) population collapsed in 2013, and a moratorium has been imposed on the fisheries since 2014 due to low population abundance and perceived continuous recruitment failures. The collapse of the population has been hypothesized to be associated with warming water temperature in the past decade. This study examines temperature effects on the quality of habitat for adult shrimp in summer and fall when spawning take place using habitat suitability index (HSI) as well as the relationships between HSI and spawning stock biomass (SSB) index for adult male and female. The HSI in this study described habitat selection for depth and bottom temperature. The results showed that the quality of habitat had declined significantly for adult shrimp in both summer and fall, especially over the past decade. Furthermore, the increasing proportion of low quality habitat in summer and fall was correlated with declining SSB index at a 2-year lag. This study provides an alternative hypothesis for the collapse of the fishery for the GOM northern shrimp.

6.2 Introduction

Northern shrimp (*Pandalus borealis*) is one of the ecologically and economically important demersal species in the Gulf of Maine (GOM) once supported a significant commercial fishery that targeted egg-bearing female shrimps in the New England states. The annual commercial landings fluctuated from less than 100 mt to over 12,000 mt, and the fishery

has experienced three major collapses over the past five decades (ASMFC NSTC 2019). The fishery has been on moratorium since 2014 due to a steep decline of abundance and consecutive recruitment failures for several years (ASMFC NSTC 2019). The low abundance and recruitment failures have been perceived to be related to unfavorably high water temperature (ASMFC NSTC 2019).

Northern shrimp are protandric hermaphrodites; they start their lives as males, maturing and mating at around 2.5 years old before they transfer to females for the remainder of their lives (Shumway et al. 1985, Richards et al. 2012). Young males start their offshore migration during their first year of life (Shumway et al. 1985; Apollonio et al. 1986). Spawning takes place in late summer and fall and egg-bearing females migrate inshore to hatch their eggs in winter. This historically observed seasonal inshore-offshore migration was perceived to be water temperature-induced and makes spatial distribution vary by seasons and life stages (Haynes and Wigley 1969; Apollonio et al. 1986). The northern shrimp seemed to be less tolerant of warm temperature at certain stage in their life span, and egg-bearing females and hatching larvae preferred colder bottom water temperature which drives inshore migration for them (Apollonio et al. 1986).

There has been increasing concern about the ecological impact of climate change and variability on the marine ecosystem and fisheries (Mills et al. 2013; Kleisner et al. 2017). Climate-induced changes in marine ecosystem includes the variability in temperature, sea level, ocean CO2 uptake, and extreme weather (IPCC 2014), which could have significant effects on distributional shift of habitat, growth, and reproduction of marine species (Hare and Able 2007; Richards et al. 2012; Hare et al. 2016). The sea surface temperature in the Gulf of Maine between 1982-2013 had increased faster than 99 % of the global oceans (Pershing et al. 2015).

The bottom water temperature in the GOM has also been increasing especially over the past decade, which could have impact on the benthic species (Kavanaugh et al. 2017).

The GOM northern shrimp are considered very susceptible to environmental variability as they are at the southern end of their distribution in the North Atlantic Ocean (Richards et al. 2012). It is essential to assess the relationship between the recent collapse in shrimp population and potential quality habitat loss due to climate-induced effects. Therefore, this study aims to develop habitat suitability index (HSI) models, quantifying the relationship between the quality of habitat and environmental factors and providing essential information on the spatiotemporal distribution and habitat preference for effective habitat-related fishery management for northern shrimp. Estimating the effects of climate change on spatial variation of habitat quality will help address how fisheries management can adapt to climate change.

6.3 Materials and Methods

6.3.1 Data

Shrimp catch data were collected by the Atlantic States Marine Fisheries Commission (ASMFC) northern shrimp summer surveys in July and August during 1984-2019, and by the Northeast Fisheries Science Center (NEFSC) fall bottom trawl surveys in October and November during 1991-2018. The ASMFC shrimp summer surveys were designed specifically for northern shrimp to estimate population abundance and provide essential input information for stock assessment for the GOM northern shrimp (ASMFC NSTC 2018; Cao et al. 2017). The NEFSC fall bottom trawl surveys were designed for multispecies including demersal fish species and invertebrate species for providing relative abundance of each species in the GOM (Politis et al. 2014). Both surveys followed consistent sampling protocols with stratified random sampling

designs. For the NEFSC fall bottom trawl surveys, the station locations were allocated throughout the western GOM which was divided into several strata based on depth and geographical location (Politis et al. 2014). The ASMFC summer northern shrimp surveys also considered historical fishing pattern for stratification (Cao et al. 2017; Clark, 1989). The sampling locations were randomly sampled from each stratum, and the number of stations within a stratum was generally proportional to the area of the strata and the overall variation of multispecies distribution among strata for fall surveys (Politis et al. 2014); the importance of the stratum to the northern shrimp assessment was considered for summer survey (Cao et al. 2017). Details of the survey sampling design and vessel configuration were documented by Stauffer (2004) and Politis et al. (2014).

Both surveys collected biological data including carapace length and life-stage. Life stages were classified into male, female I, female II, transition, and ovigerous female. Transition stage shrimp were not used as samples of shrimp at transition stage were relatively few and the time series was incomplete. Male shrimp were further classified into recruits and mature male based on carapace length of recruitment cutoffs for each year (DMR). Female I shrimp were defined as mature females that have not yet reached spawning age, possessing sharply pointed abdominal spines (McCrary 1971). Female II shrimp were defined as having vestigial abdominal spines and which have spawned at least once before (McCrary 1971). Female I and female II in summer were combined as Female. For summer data, female and mature male data were used in this study. As for fall data, ovigerous female and mature male data were used as only a few female I and female II were caught in fall.

Unstaged data were expanded for each tow based on carapace length distribution and proportion of life-stage of staged data collected from that tow. If a tow caught no shrimps in a stratum where any shrimp (season and life-stage specific) was present, the tow is considered a zero-catch tow. Carapace length distributions of males were used to determine length of recruitment cutoffs in fall and summer of each year, male shrimp were then further classified into Recruits and mature male based on the cutoffs.

Abiotic data included water temperature and sampling depth. During the NEFSC fall surveys since 1991 and during the ASMFC shrimp surveys since 1984, bottom temperatures were measured with expendable bathythermographs (XBTs) and with a CTD (conductivity, temperature and depth profiler).

6.3.2 Analysis

Datasets are processed season- and life-stage specifically. Survey catch per unit effort (CPUE) for each life stage was standardized as number of shrimp per 20 minutes tow duration.

6.3.2.1 Suitability index (SI)

Water temperature and depth were used for calculating suitability index (SI) and consequent habitat suitability index (HSI) based on literature review and expert opinion (Apollonio et al. 1986; Richards et al. 2012; M. Hunter, personal communication). Apollonio et al. (1986) indicated that water temperature may be a driver for inshore-offshore migration for the GOM northern shrimp. The success of recruitment and strength of year classes were also considered to be affected by water temperature (Richards et al. 2012).

Each variable was classified into 10 bins using Jenk's natural breaks classification (Bivand et al. 2020). The suitability index for each variable was calculated as:

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

where i = variable, k = bin, $CPUE_{i,k}$ = the average CPUE in k^{th} bin of i variable, the $CPUE_{i,min}$ and $CPUE_{i,max}$ are the bins with minimum and maximum average CPUEs of i variable. The SI ranged from 0 to 1, with 0 for the lowest CPUE and 1 for the highest CPUE.

In a preliminary analysis, the SI of bottom temperature for fall groups showed a bimodal pattern (see Results section). Therefore, each of fall ovigerous female and mature male were further subset into small ovigerous Female (DCL ≤ 25 mm), large ovigerous female (DCL > 25 mm), small mature male (DCL ≤ 20 mm), and large mature male (DCL > 20 mm).

6.3.2.2 Habitat suitability index (HSI)

Habitat suitability index (HSI) was estimated as the geometric mean (GMM) of SIs of all the variables:

$$HSI = \left(\prod_{i=1}^{n} SI_i\right)^{1/n}$$

The GMM was used as it assumes that if an SI in an area is zero then the HSI in that area is rated zero. The arithmetic mean was not considered in this study as it allows a non-zero HSI value even if one of the SIs was zero in an area, which may not be biologically meaningful (Draugelis-Dale 2008).

A Generalized Additive Model (GAM; Hastie and Tibshirani 1990) with lognormal distribution was used to quantify the relationships between HSI and candidate habitat variables. A variance inflation factor (VIF) analysis was conducted to identify multicollinearity before fitting models to the data. A threshold VIF of 3 was set for evaluating possible collinearity between predictors in the data set (Schmiing et al. 2013; Brosset et al. 2019). The form of the GAMs was:

$$g(y) \sim \alpha + \sum_{i=1}^{k} f_i(x_i) + \varepsilon_i$$

where g() is the link function, y is HSI, f_i is the i^{th} smooth function based on thin plate regression splines, x_i is the i^{th} explanatory variable, and ε_i is residual errors.

The VIF analysis suggested that multicollinearity was not an issue in modeling the dataset as all explanatory variables (bottom temperature, depth, longitude, and latitude) had VIFs < 3. Longitude and latitude were included in the base model in order to catch significant spatial trends not captured in other variables. Models of all combinations of other variables (bottom temperature and depth) were built using maximum likelihood methods (Fisher et al. 2018).

Model selection was based on full-subsets information theoretic approaches (Anderson and Burnham 2002; Fisher et al. 2018), using Akaike's information criterion (Akaike 1973), AIC weights (ωAIC), deviance explained by the model, and graphical inspection. Root mean squared

error (RMSE) estimated with leave-one-out cross-validation (LOOCV) was used to evaluate predictive performance of models (Zuur et al. 2009; Arlot and Celisse 2010).

Models with ΔAIC <2 were considered best models (Anderson and Burnham 2002). Models with the lowest AICc scores were refit with REML (Zuur et al. 2009; Wood 2017) and effect graphs of these models were presented. Biological relevance and the relationships between response variables and predictors were graphically examined. Relative importance of predictors was calculated as the sum of AICc weights of models in which a predictor is present (Anderson and Burnham 2002; Fisher et al. 2018).

The residuals of the optimal model were graphically examined with QQ-plots to inspect any patterns in the residual errors (Zuur et al. 2009). Semivariance of Pearson residual errors from the most optimal model was examined for the presence of spatial autocorrelation (Cressie 1993; Pebesma 2004; Gräler et al. 2016). Autocorrelation (AC) and partial autocorrelation (PAC) plots of Pearson residuals were used for inspecting temporal autocorrelation (Stoffer and Shumway 2017). If temporal autocorrelation was suggested by the plot, then year was included in the model in order to account for the autocorrelation between years.

Depth data collected by NOAA National Centers for Environmental Information (NCEI) and bottom temperature data interpolated from ASMFC and NEFSC surveys data were used for hindcasting spatiotemporal variation in HSI for northern shrimp in summer and fall. Seasonal mean values of bottom temperature simulated from Finite-Volume Community Ocean Model (FVCOM) were also used for hindcasting the expected HSI. The predicted HSI using the two data sources of bottom temperature was qualitatively compared. Linear regression models were fit to hindcasted HSI time series for examining temporal trends.

The estimated HSI were classified into low (< 0.25 quantile of HSI), medium (> 0.25 quantile and < 0.75 quantile), and high (> 0.75 quantile) quality groups for each year and each life stage. The proportions of high and low quality HSI were cross correlated with estimated spawner biomass (SSB) indices (ASMFC NSTC 2019). The SSB was estimated based on the size-distribution of the population and length-weight relationship. The spawners were defined as females larger than 22 mm (dorsal carapace length) (ASMFC NSTC 2019). All statistical analyses were performed using R 4.0.3 statistical software (R Core Team 2020).

6.4 Results

The maps of distribution of CPUE (number of shrimp per 20 minutes tow) were shown in Fig. 6.1. In summer, high densities of mature female and male shrimp were found in Jeffrey's Ledge, Platts Bank, Jeffrey's Bank, and Cashes Ledge. In fall, small ovigerous female were more concentrated in Jeffrey's Ledge. While large ovigerous female, small and large mature male were more spread out toward Jeffrey's Bank and Wilkinson Basin.

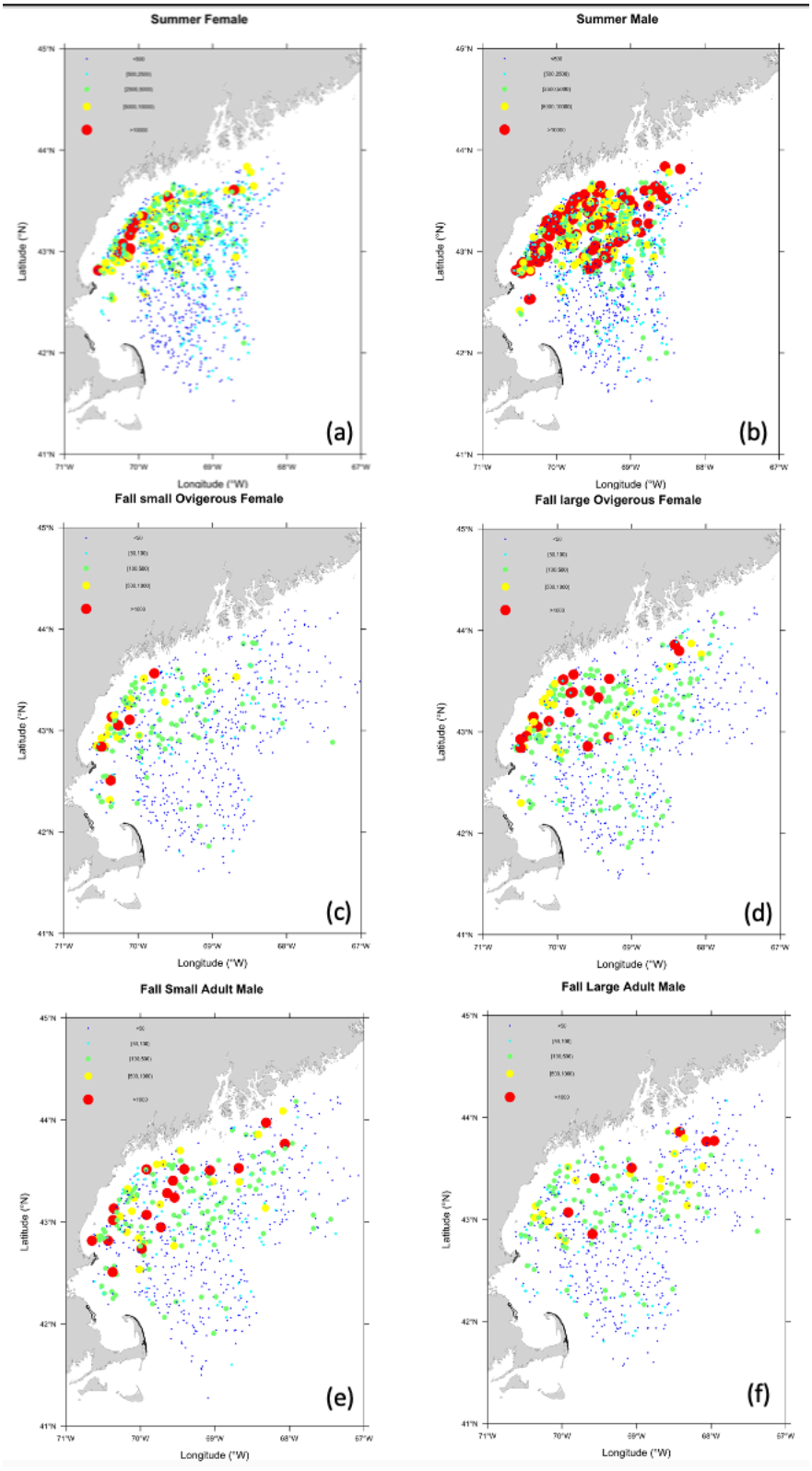

Summer Female
Summer Male
Fall small Ovigerous Female
Fall large Ovigerous Female
Fall Small Adult Male
Fall Large Adult Male
(a)
(b)
(c)
(d)
(e)
(f)
Latitude (°N)
Longitude (°W)
45°N
44°N
43°N
42°N
41°N
71°W
70°W
69°W
68°W
67°W

Figure 6.1. Maps of distribution of survey CPUE (number of shrimp per 20 minute tow time) for (a) summer female and (b) summer female collected from Atlantic States Marine Fisheries Commission (ASMFC) summer northern shrimp survey, and (c) fall small ovigerous female, (d) fall large ovigerous female, (e) fall small mature male, and (f) fall large mature male collected from Northeast Fisheries Science Center (NEFSC) fall bottom trawl survey.

The suitability indices of bottom temperature and depth for summer and fall female and male northern shrimp were shown in Fig. 6.2. The suitability indices for both summer female and male decreased with increasing bottom temperature. The suitability indices for depth had similar trends in summer and fall for both female and male, with the highest suitability indices around 120-130 m. However, the trends of suitability indices of bottom temperature for both female and male in the fall were not clear, showing a bimodal distribution with one mode around 6-7°C and one mode around 9.5°C. Each of the fall groups (ovigerous female and mature male) were thus further subset into small and large sizes.

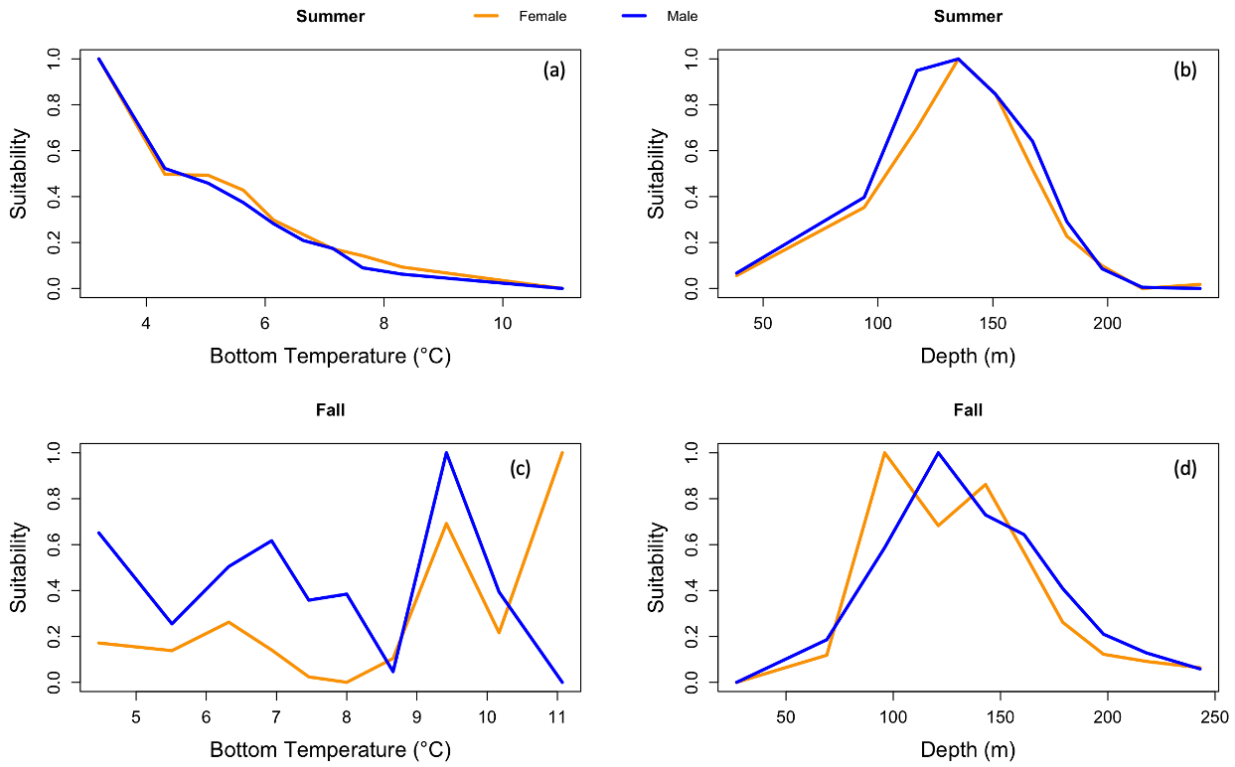


Figure 6.2. The relationships between suitability index and bottom temperature (a and c) and depth (b and d) for female (orange lines) and male (blue lines) northern shrimp in summer and fall.

The SI of small ovigerous female and mature male have higher values when bottom temperature < 7°C, while all four fall groups have a higher level of SI with bottom temperature ~ 9.5°C (Fig. 6.3). The SI of depth for fall small ovigerous female and mature male have similar patterns with a single mode at around 100 m for small ovigerous female and at around 120 m for small mature male. While the SI of depth for both large ovigerous female and mature male have a bimodal distribution with a second mode at around 150 m (Fig. 6.3).

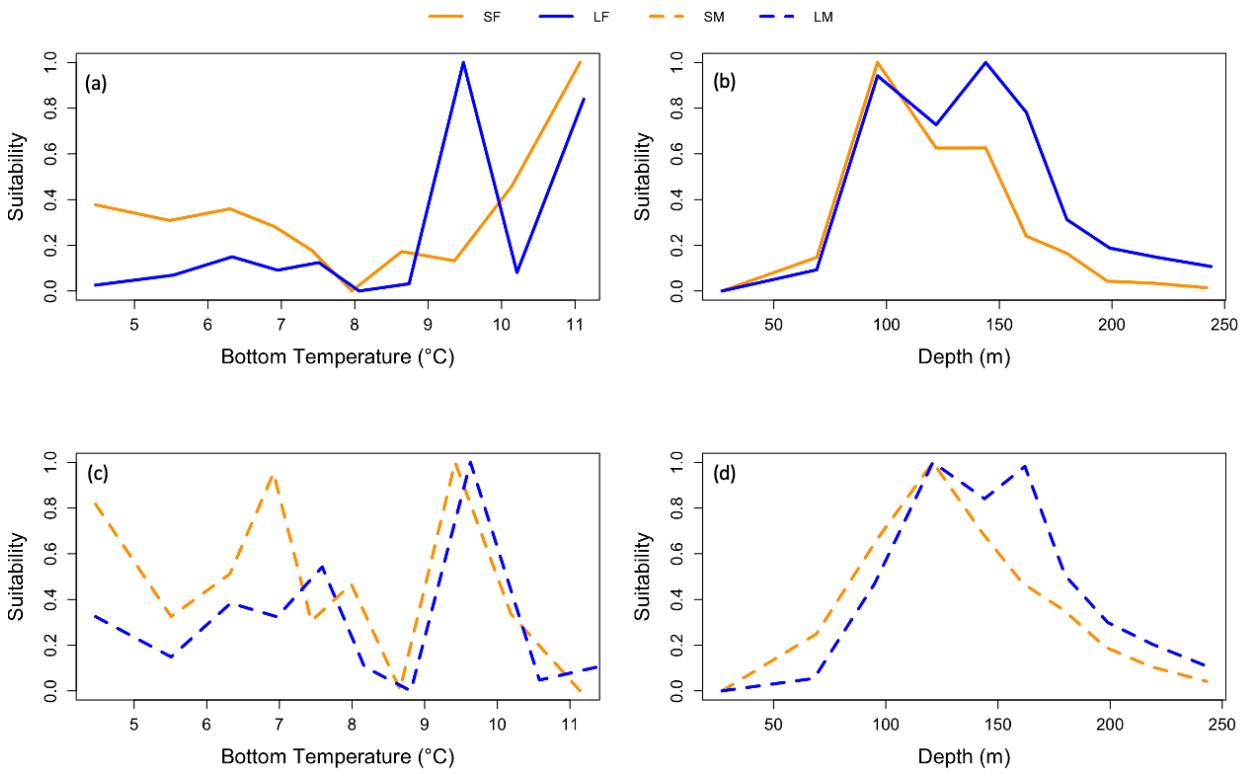


Figure 6.3. The relationships between suitability index and bottom temperature (a and c) and depth (b and d) for fall small ovigerous female (SF, orange solid lines), fall large ovigerous female (LF, blue solid lines), fall small mature male (SM, orange dashed line), and fall large mature male (LM, blue dashed lines).

6.4.1 Models

For both summer and fall groups, the full model was selected as the best model with the lowest AIC value (Table 6.1) for each group. Moreover, the best models outcompete the rest of models with ΔAIC>10 and ωAIC=1, suggesting that bottom temperature and depth are two important variables and showing low level of model uncertainties. The models were able to explain 96.1% and 96.6% of the variation in HSI for female and male, respectively (Table 6.1).

Table 6.1. The best models with the lowest AIC values for each group in summer and fall. N=number of observations, DE = deviance explained, EDF=effective degrees of freedom, Adj R^2=adjusted R^2, LOOCV RMSE=leave-one-out cross-validation root mean squared error, LON=longitude, LAT=latitude, BOTTEMP=bottom temperature.

Models		N	DE	EDF	Adj R^2	LOOCV RMSE
Summer female	s(LON, LAT) + s(BOTTEMP) + s(DEPTH)	1767	0.961	18.9	0.961	0.090
Summer male	s(LON, LAT) + s(BOTTEMP) + s(DEPTH) + YEAR	1767	0.966	17.8	0.965	0.041
Fall small ovigerous female	s(LON, LAT) + s(BOTTEMP) + s(DEPTH)	1149	0.866	23.0	0.877	0.074
Fall large ovigerous female	s(LON, LAT) + s(BOTTEMP) + s(DEPTH)	1226	0.801	20.8	0.797	0.099
Fall small male	s(LON, LAT) + s(BOTTEMP) + s(DEPTH) + YEAR	1184	0.726	17.5	0.716	0.134
Fall large male	s(LON, LAT) + s(BOTTEMP) + s(DEPTH)	1162	0.838	21.6	0.835	0.095

The models explained 72.6-86.6% of the variation for fall groups (Table 6.1). The LOOCV RMSEs (predictive errors, i.e. the average bias between prediction and observation) ranged from 0.041 to 0.134 among the six groups. However, the AC and PAC plots showed temporal autocorrelation at lag 1 for summer male and at lag 4 for fall small mature male. Year was therefore included in the models of these two groups for accounting for temporal autocorrelation. Spatial and temporal autocorrelations were not suggested by semivariance (Fig. 6.4), AC and PAC plots (Fig. 6.5) for models of other groups.

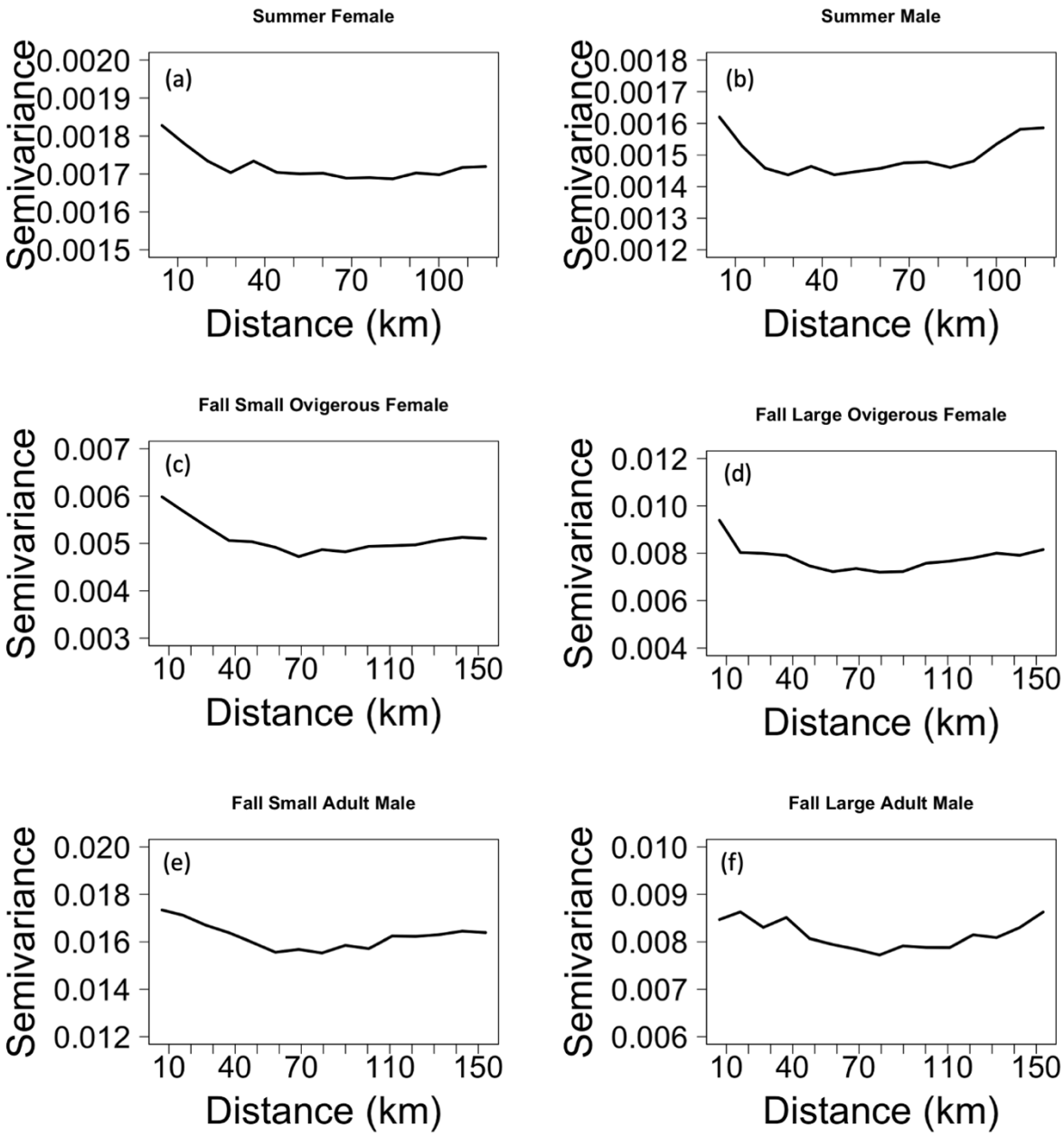


Figure 6.4. Variogram computed with Pearson residuals of the best models with lowest AIC values for (a) summer female, (b) summer male, (c) fall small ovigerous female, (d) fall large ovigerous female, (e) fall small adult male, and (f) fall large adult male.

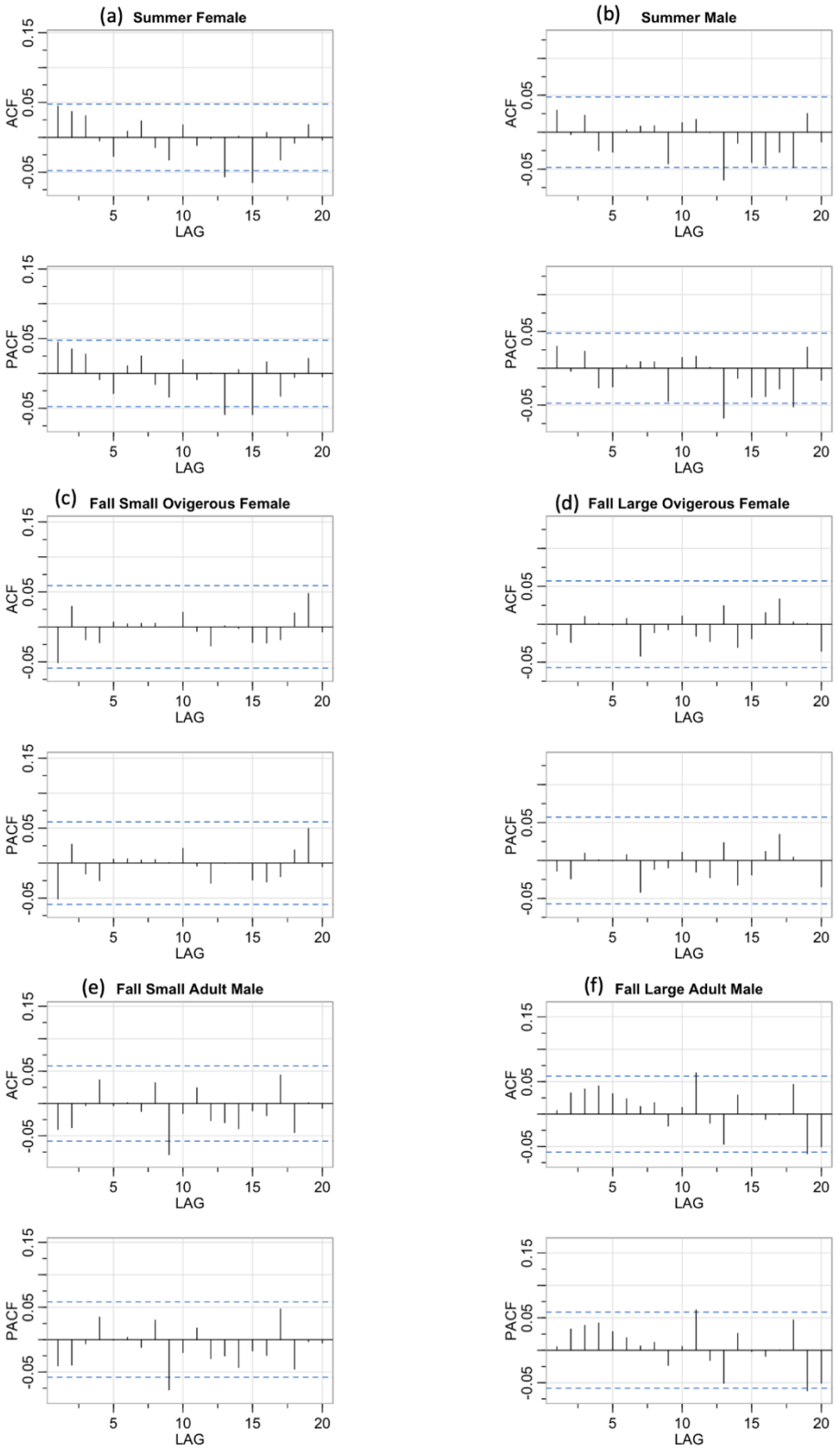

(a) Summer Female
(b) Summer Male
(c) Fall Small Ovigerous Female
(d) Fall Large Ovigerous Female
(e) Fall Small Adult Male
(f) Fall Large Adult Male
ACF
PACF
LAG
0.15
0.05
-0.05
5
10
15
20

Figure 6.5. Autocorrelation (AC) and partial autocorrelation (PAC) plots of Pearson residuals of the best models with lowest AIC values for (a) summer female, (b) summer male, (c) fall small ovigerous female, (d) fall large ovigerous female, (e) fall small adult male, and (f) fall large adult male.

The partial effects of bottom temperature on HSI for each group were shown in Fig. 6.6. For summer female and male, HSI decreased with increasing bottom temperatures, although the uncertainty increased when bottom temperature >10°C.

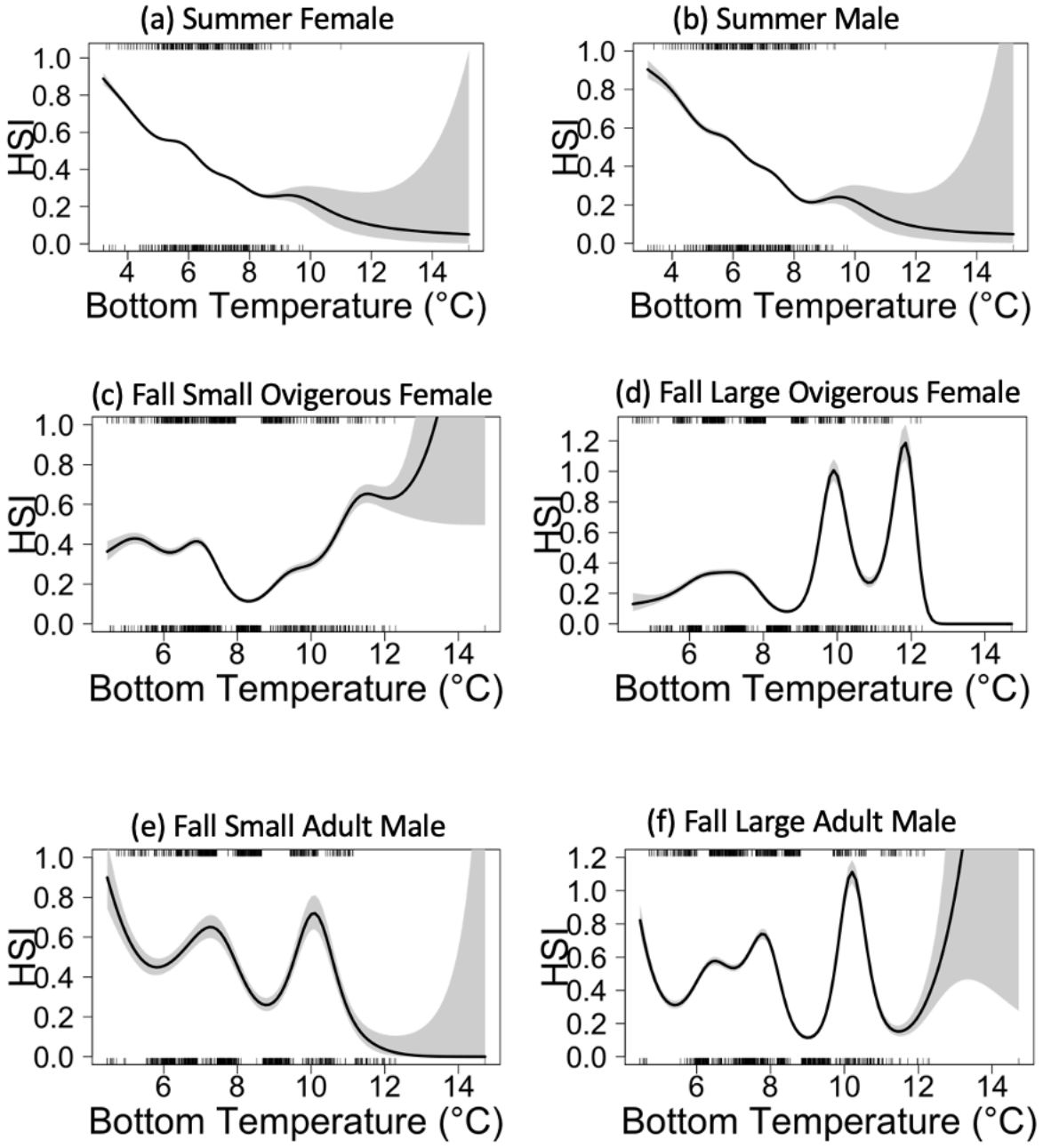


Figure 6.6. Partial effects of bottom temperature on habitat suitability index (HSI) of the best models with lowest AIC values for (a) summer female, (b) summer male, (c) fall small ovigerous female, (d) fall large ovigerous female, (e) fall small adult male, and (f) fall large adult male.

In contrast to summer groups, the bottom temperature effects on HSI for fall groups did not show clear trends (Fig. 6.6c-f). For small ovigerous female, the HSI increased when bottom temperature > 8°C, while the HSI of large ovigerous female was at higher levels at bottom temperatures of around 10 and 12°C (Fig. 6.6c-d). For fall male groups, both small and large mature male had higher HSI values at bottom temperatures of around 7-8°C and 10°C (Fig. 6.6e-f).

The partial effects of depth on HSI for each group were shown in Fig. 6.7. For summer groups, both female and male had the highest HIS when depth ranged 120-170 m (Fig. 6.7a-b). The fall groups had similar patterns except the HSI was higher at deeper depth for both large ovigerous female and adult male (Fig. 6.7c-f).

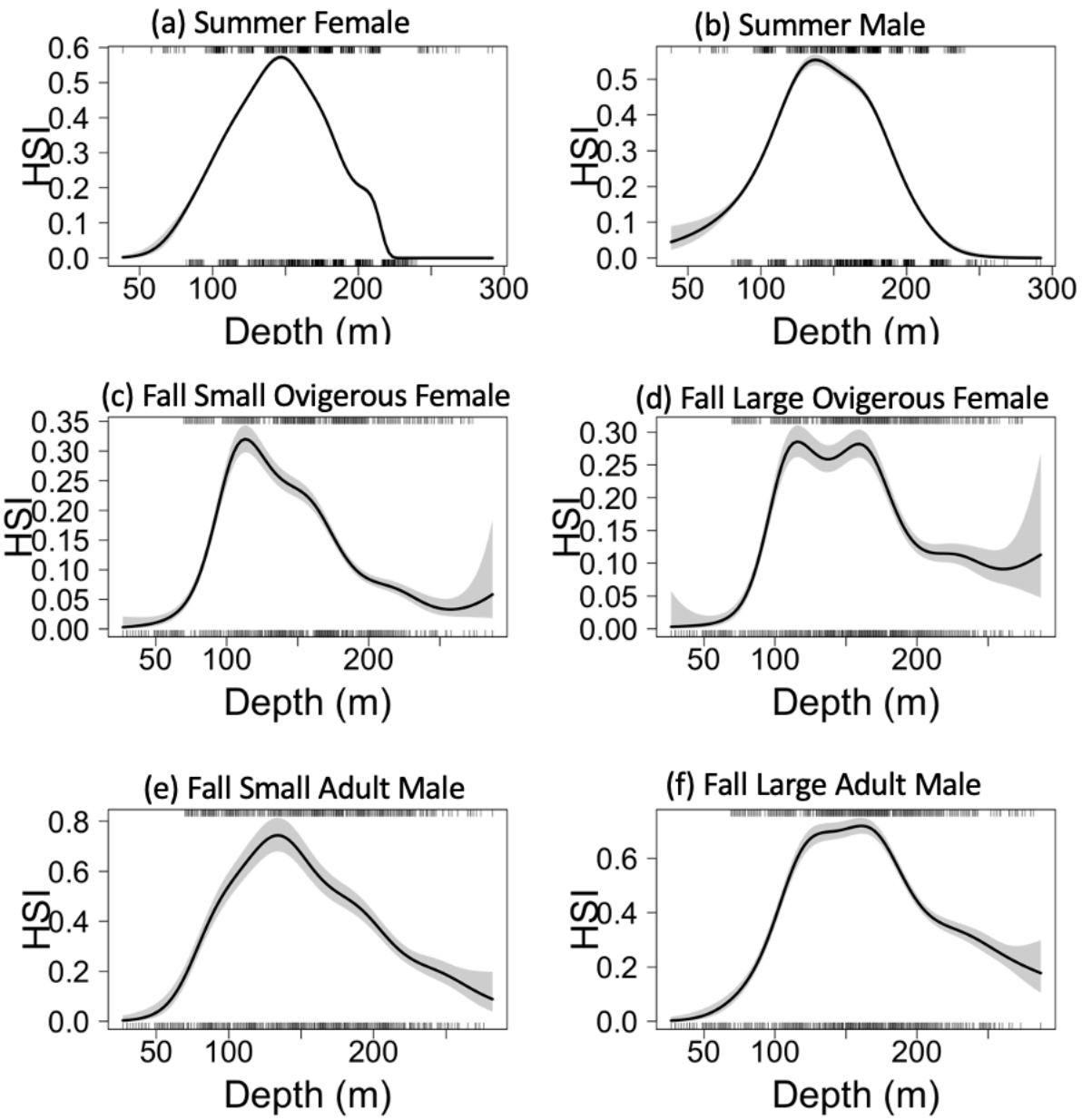


Figure 6.7. Partial effects of depth on habitat suitability index (HSI) of the best models with lowest AIC values for (a) summer female, (b) summer male, (c) fall small ovigerous female, (d) fall large ovigerous female, (e) fall small adult male, and (f) fall large adult male.

6.4.2 Interpolated survey bottom temperature

The maps of hindcasted HSI using interpolated survey bottom temperature data for Summer Female and Male were shown in Figs. 6.8-6.9. The HSI distributions had similar patterns for summer female and male.

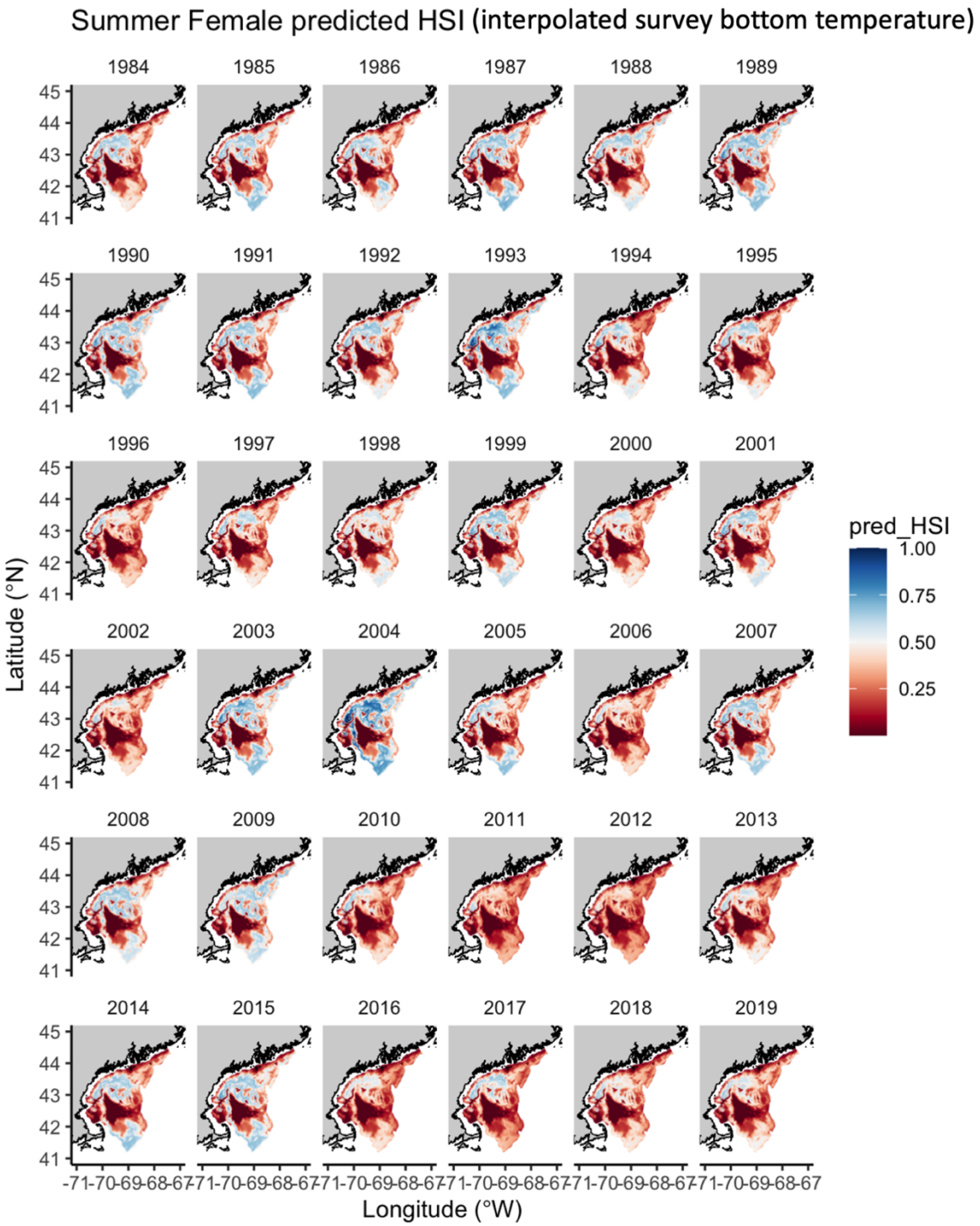


Figure 6.8. Maps of habitat suitability index (HSI) hindcasted with interpolated bottom temperature data for summer female from 1984 to 2019.

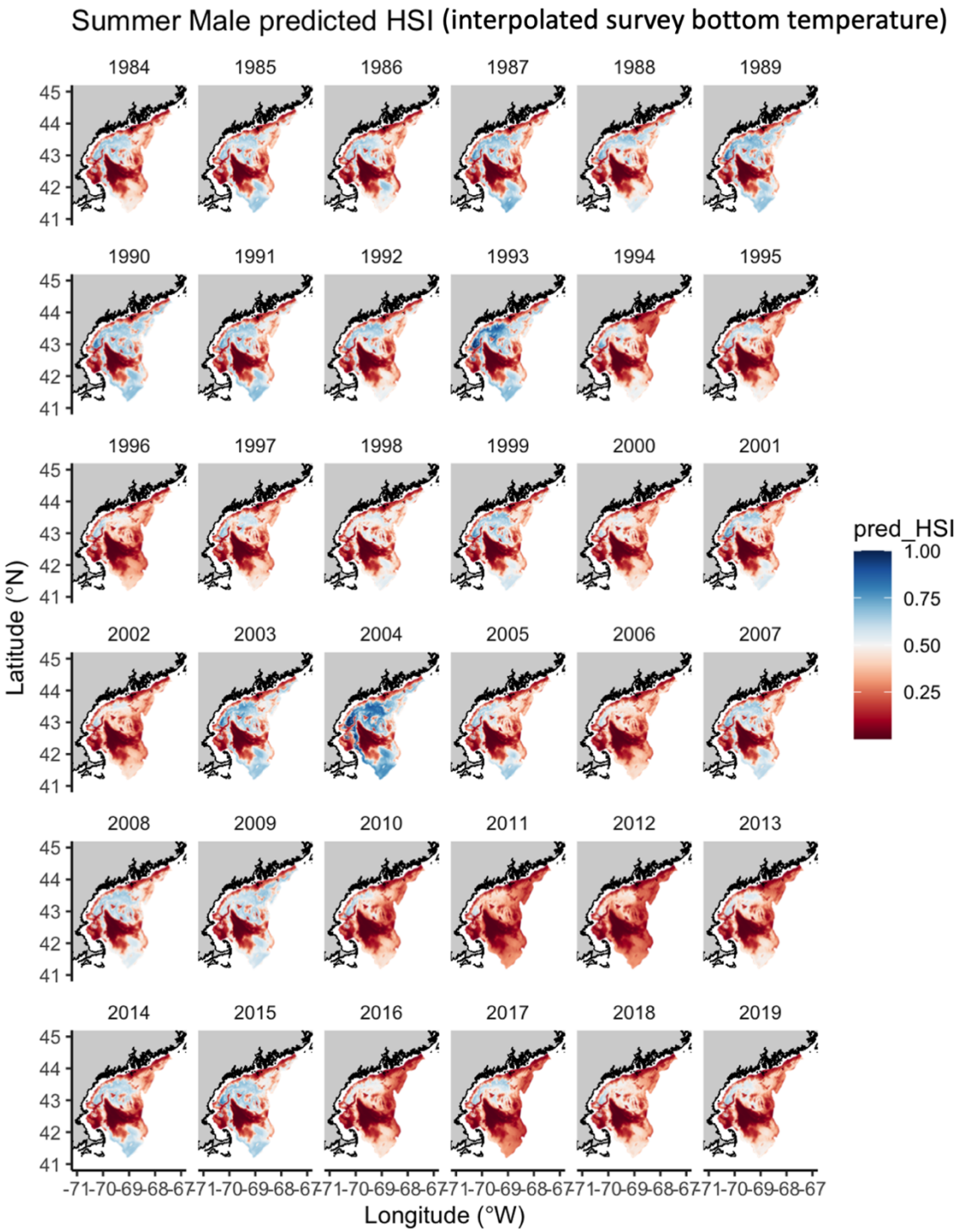


Figure 6.9. Maps of habitat suitability index (HSI) hindcasted with interpolated bottom temperature data for summer male from 1984 to 2019.

The HSI for female and male in the areas where most female and male distributed in the summer, i.e. Jeffrey's Ledge, Platts Bank, and Jeffreys Bank, were generally above 0.5 for most years before 2010. However, the HSI for both female and male in these areas were around or below 0.5 for most years after 2010, especially in 2010-2013.

As for HSI distribution for fall groups (Figs. 6.10-6.13), the small and large ovigerous female had similar patterns. However, the small ovigerous female tended to have higher HSI values, especially prior to 2008. The HSI for both small and large ovigerous female were relatively higher in Jeffrey's Ledge, Platts Bank, and Cashes Ledge for most years before 2008. However, the HSI in these areas were at low level (HSI<0.25) for most years after 2010. Similar to fall ovigerous female, The high HSI areas were mostly in Jeffrey's Ledge, Platts Bank, and Cashes Ledge for most years before 2008, with generally higher HSI values for small mature male than large mature male in these areas.

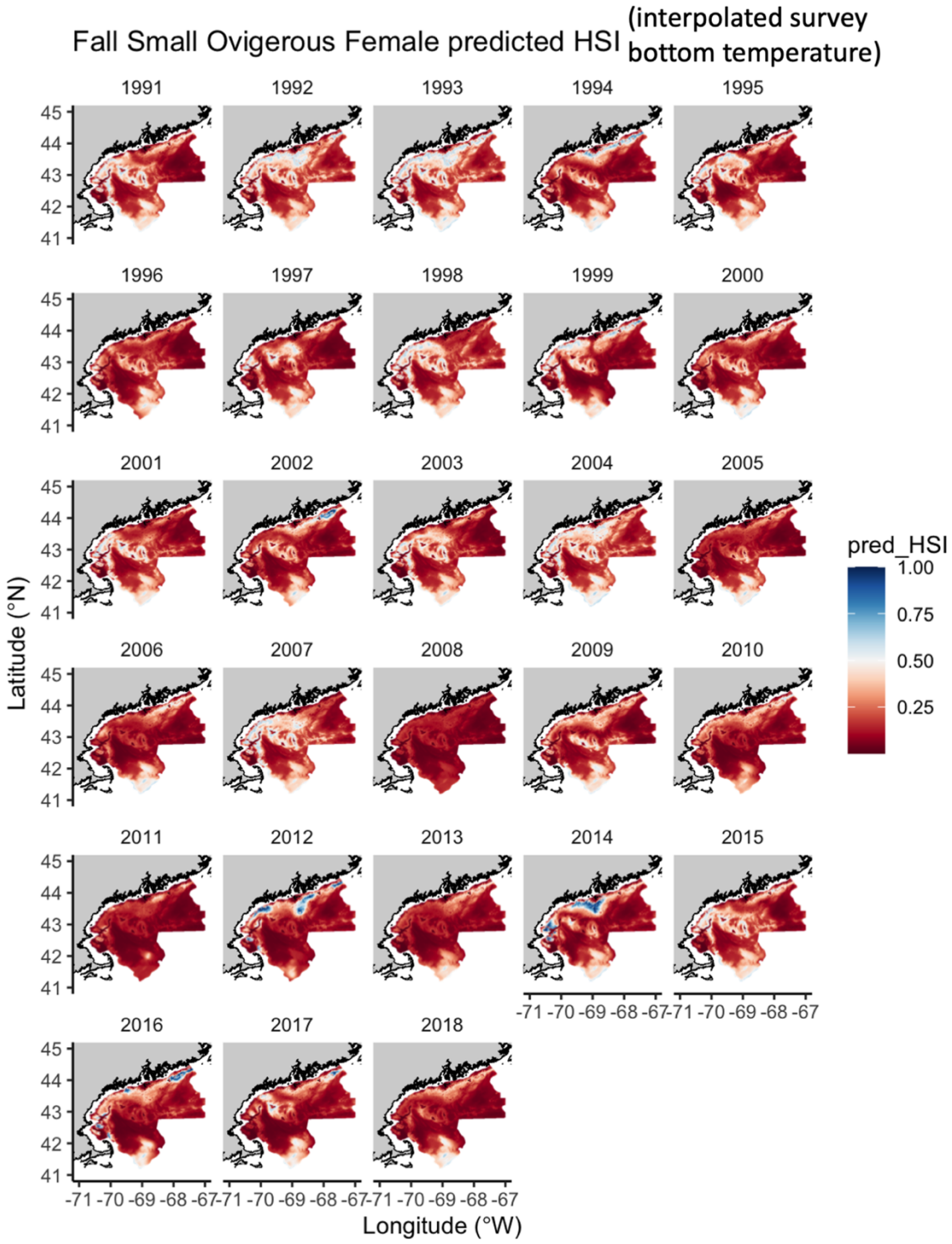


Figure 6.10. Maps of habitat suitability index (HSI) hindcasted with interpolated bottom temperature data for fall small ovigerous female from 1991 to 2018.

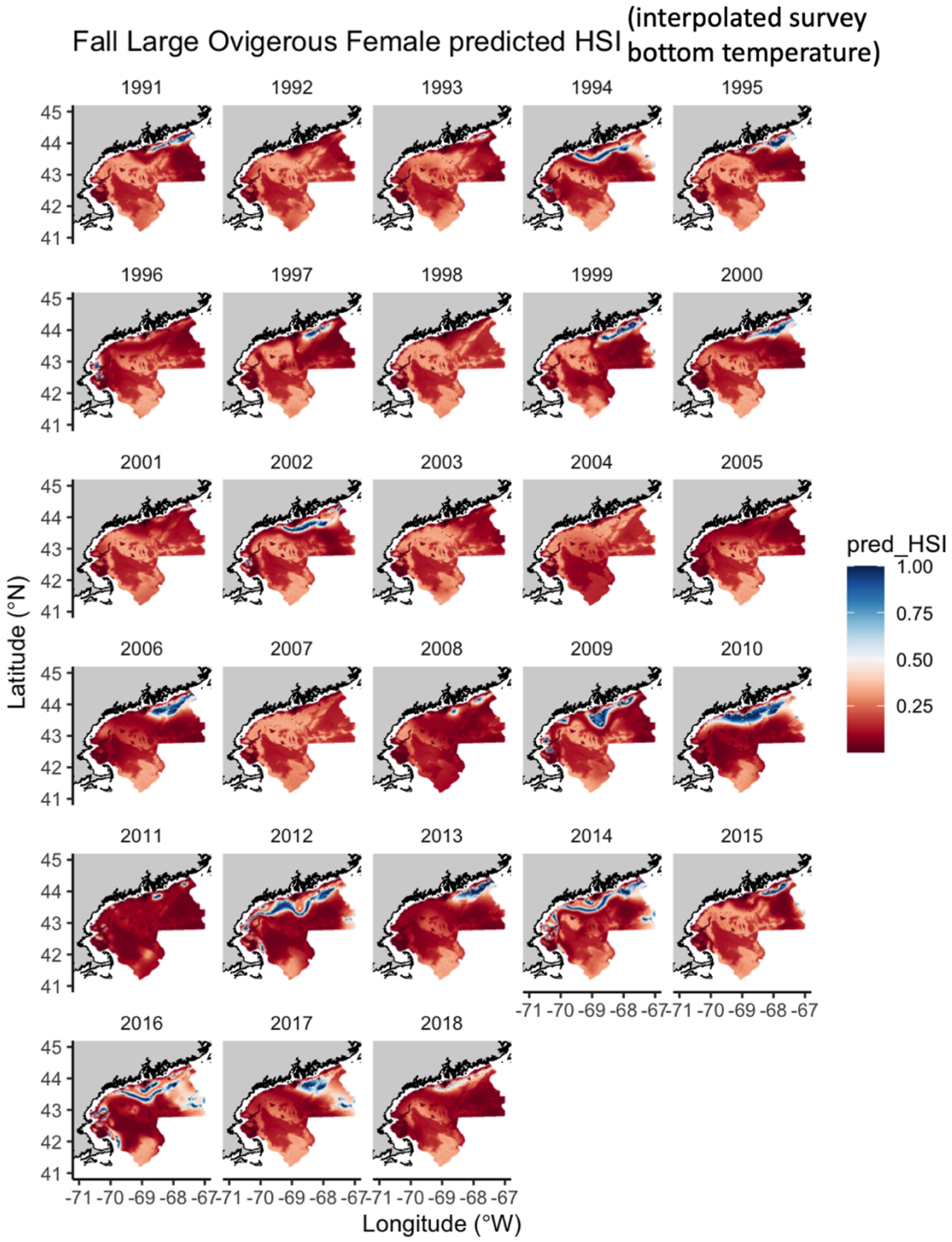


Figure 6.11. Maps of habitat suitability index (HSI) hindcasted with interpolated bottom temperature data for fall large ovigerous female from 1991 to 2018.

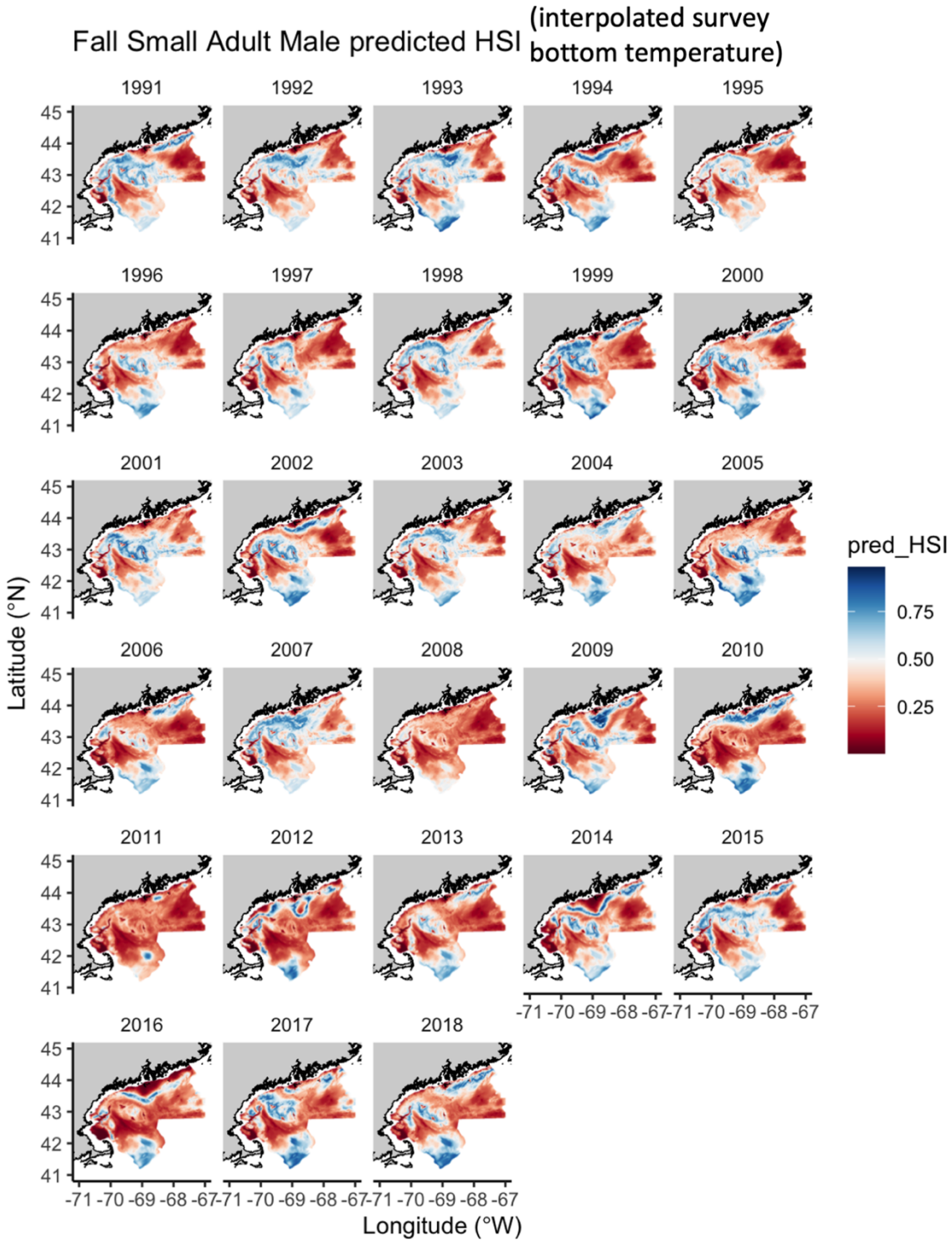


Figure 6.12. Maps of habitat suitability index (HSI) hindcasted with interpolated bottom temperature data for fall small adult male from 1991 to 2018.

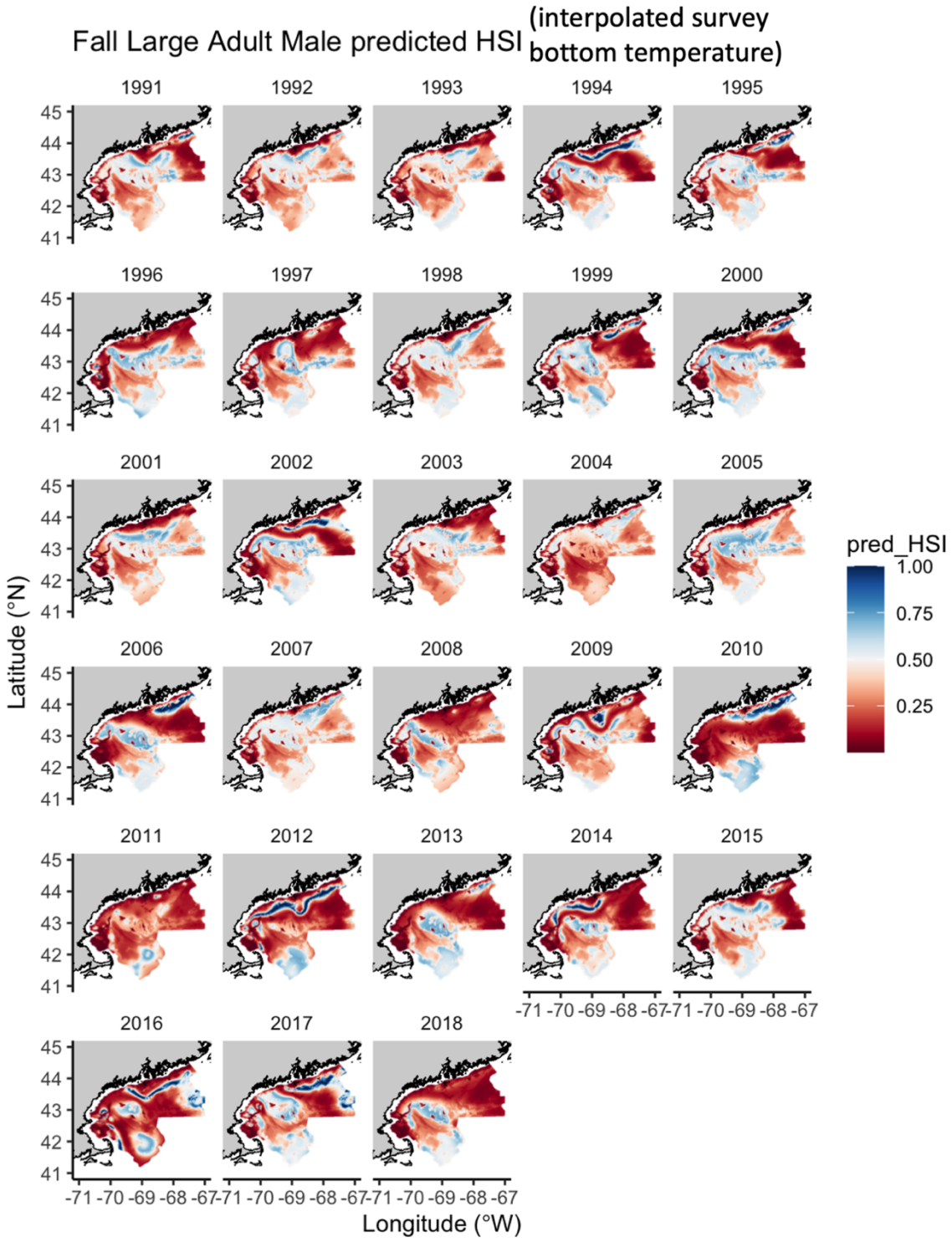


Figure 6.13. Maps of habitat suitability index (HSI) hindcasted with interpolated bottom temperature data for fall large adult male from 1991 to 2018.

The slopes of temporal changes for each group were shown in Fig. 6.14, and the corresponding p-values were shown in Fig. 6.15. The HSI for summer female and male were generally significantly decreasing over 1984-2019, especially in east and west fringe of Wilkinson Basin, Platts Bank, cashes Ledge, and Jeffreys Bank.

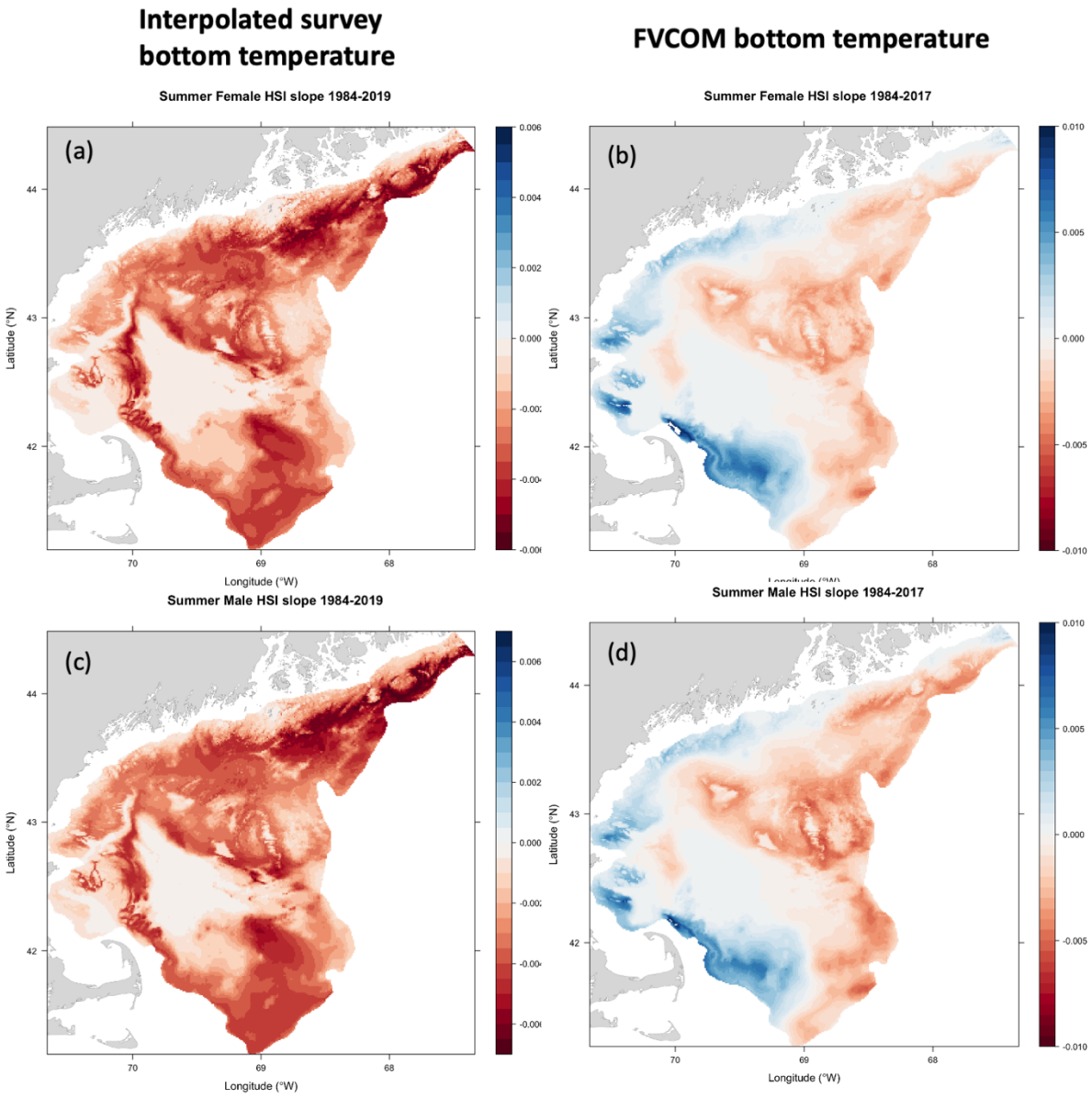


Figure 6.14. The maps of slopes of temporal changes in habitat suitability index (HSI) hindcasted with interpolated bottom temperature and FVCOM bottom temperature for summer female (a and b) and male (c and d).

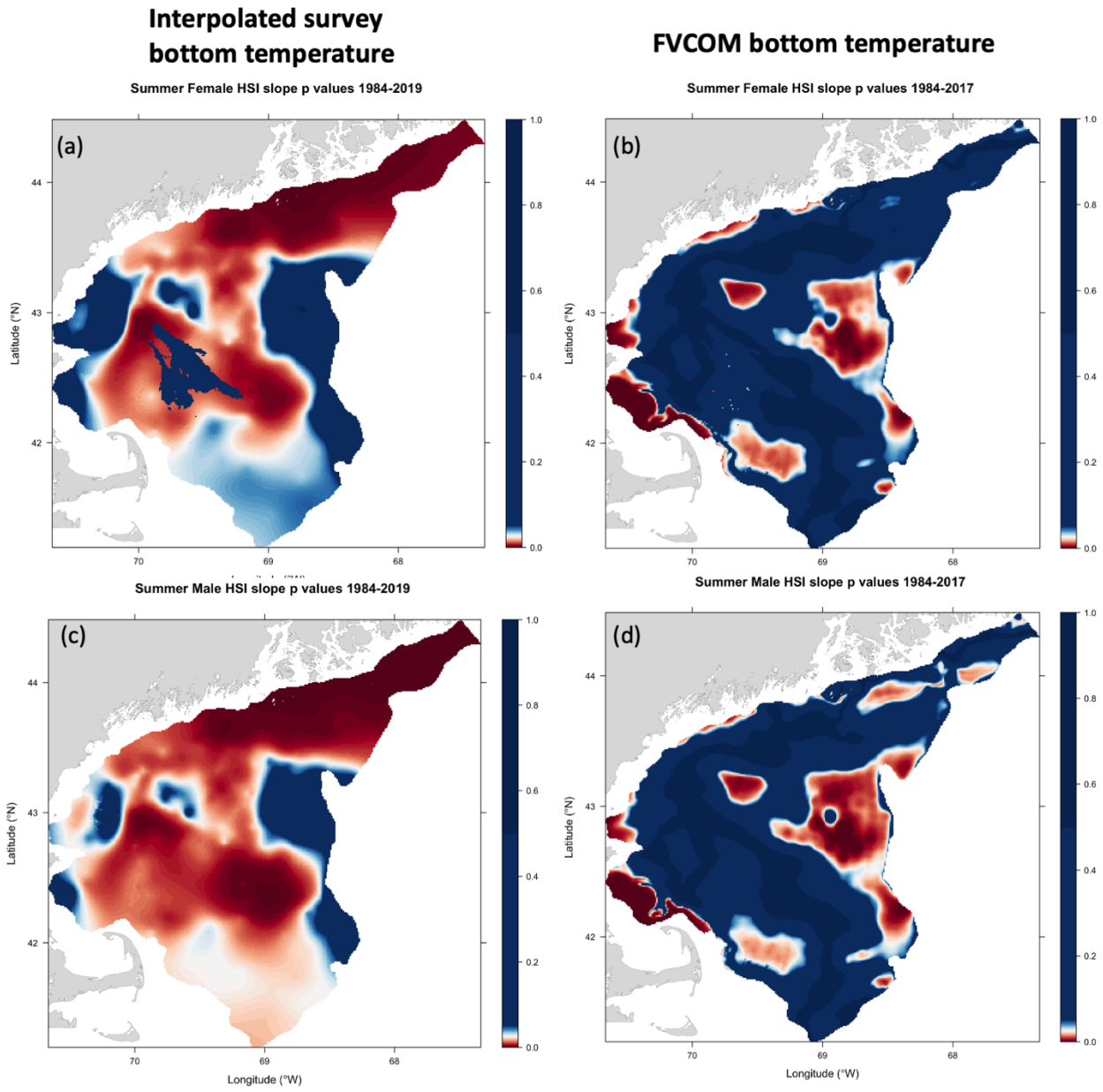


Figure 6.15. The maps of the p-values of temporal changes in habitat suitability index (HSI) hindcasted with interpolated bottom temperature and FVCOM bottom temperature for summer female (a and b) and male (c and d).

Similar patterns were also found for fall groups (Figs. 6.16-6.17); however, the HSI for fall groups in Jeffreys Bank were improving although not significantly (Figs. 6.18-6.19). Moreover, the decreasing trends of HSI in Cashes Ledge for both small and large fall ovigerous female over 1991-2018 were significant (Fig. 6.18). The changes in temporal trends for both small and large mature male in high density areas (Jeffrey's Ledge, Platts Bank, and Cashes Ledge) were not statistically significant (Fig. 6.19).

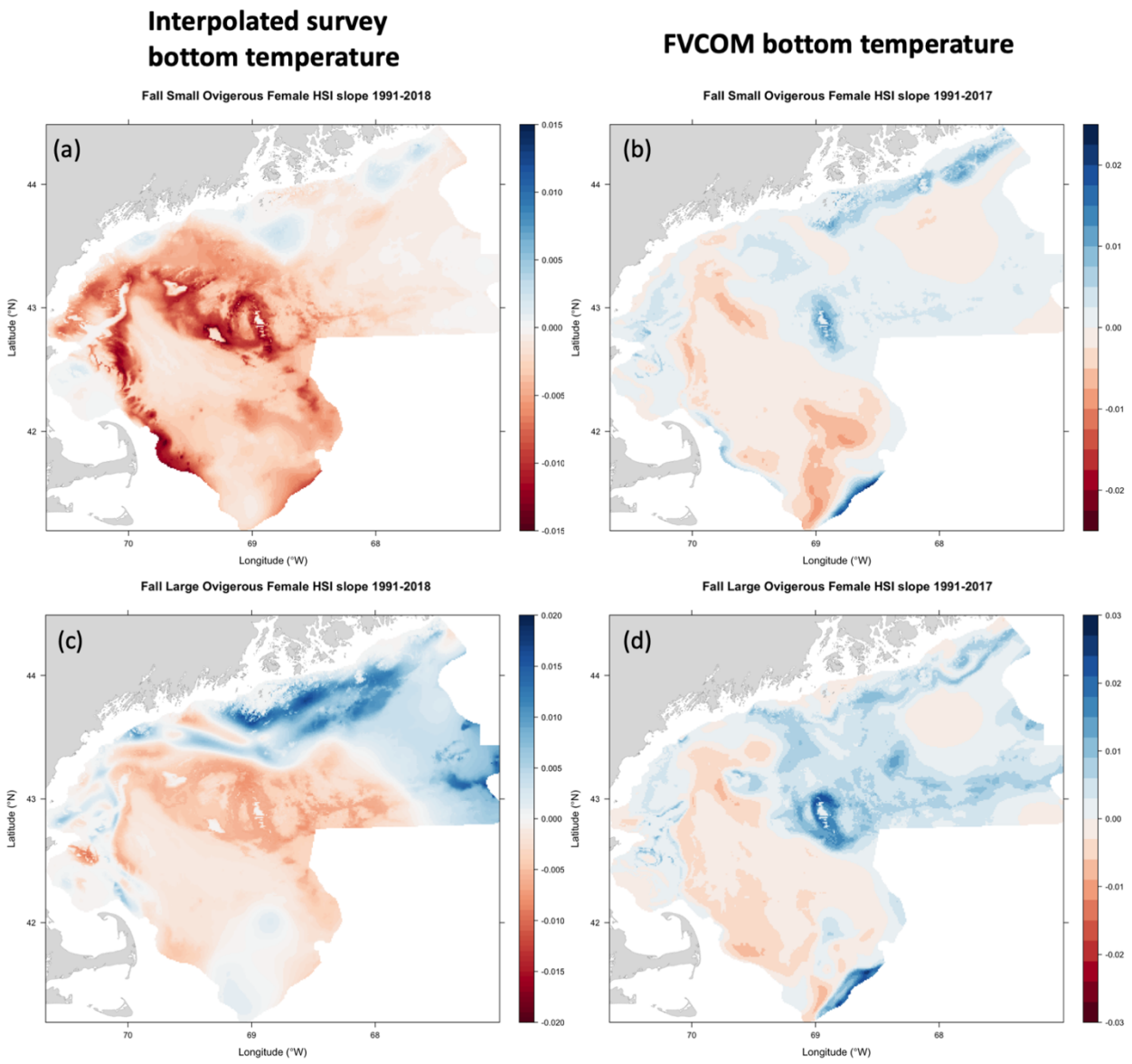


Figure 6.16. The maps of slopes of temporal changes in habitat suitability index (HSI) hindcasted with interpolated bottom temperature and FVCOM bottom temperature for fall small ovigerous female (a and b) and large ovigerous female (c and d).

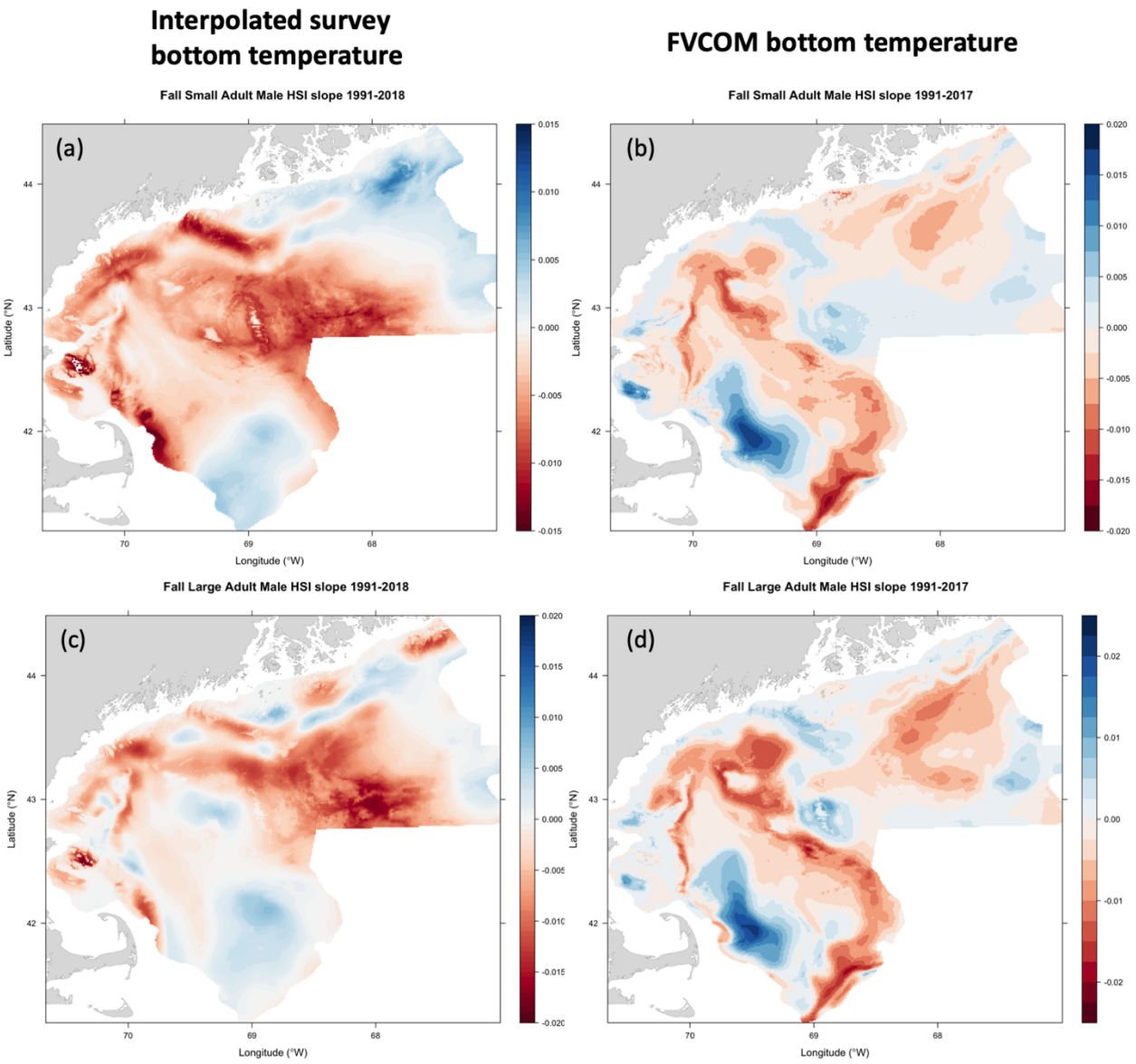


Figure 6.17. The maps of slopes of temporal changes in habitat suitability index (HSI) hindcasted with interpolated bottom temperature and FVCOM bottom temperature for fall small adult male (a and b) and large adult male (c and d).

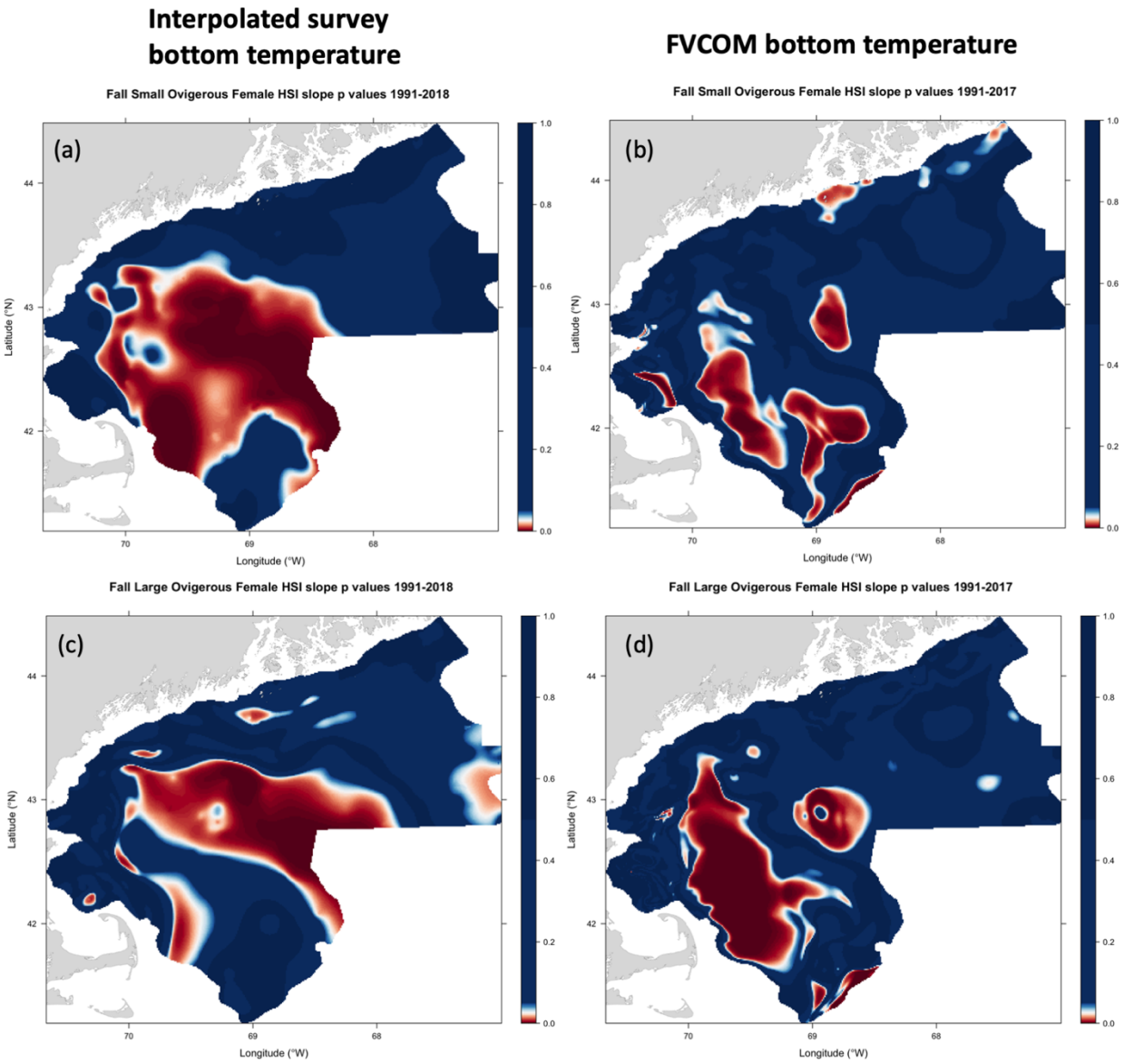


Figure 6.18. The maps of the p-values of temporal changes in habitat suitability index (HSI) hindcasted with interpolated bottom temperature and FVCOM bottom temperature for fall small ovigerous female (a and b) and large ovigerous female (c and d).

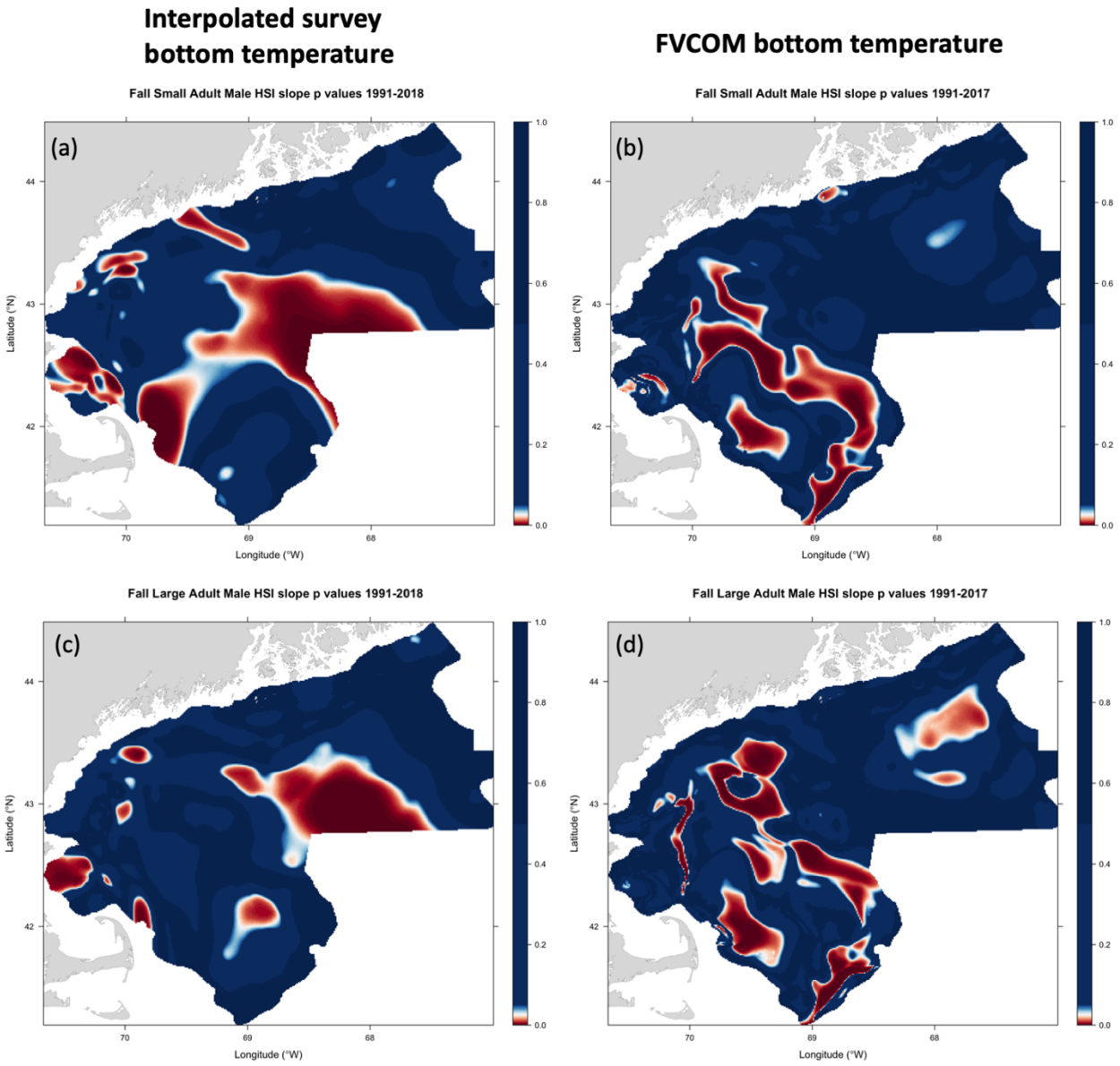


Figure 6.19. The maps of the p-values of temporal changes in habitat suitability index (HSI) hindcasted with interpolated bottom temperature and FVCOM bottom temperature for fall small adult male (a and b) and large adult male (c and d).

The proportions of low quality habitat were negatively correlated with SSB index at a 2-year lag for summer female (r=-0.55) and Summer Male (r=-0.58), meaning that if the proportion of low quality habitat increased, the SSB index would decrease two years later (Fig. 6.20-a and c). The proportions of low quality habitat were also negatively correlated with SSB index at a 2-year lag for fall large ovigerous female (r=-0.38, maximum r=-0.44 at lag 5), small mature male (r=-0.52), and large mature male (r=-0.56), but not significant for small ovigerous female (Fig. 6.20-e, g, I, and k).

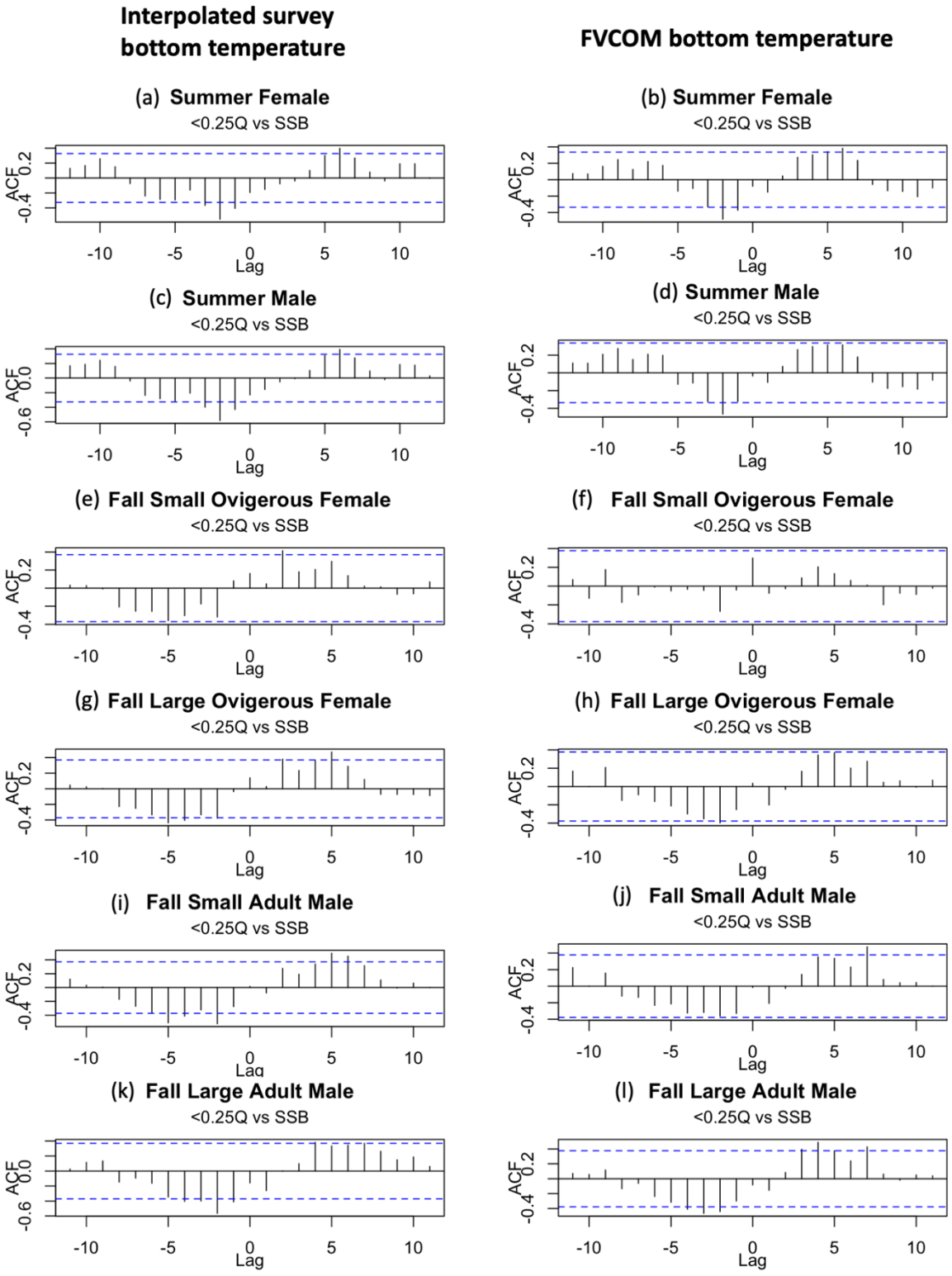

Interpolated survey bottom temperature
FVCOM bottom temperature
(a) Summer Female
(b) Summer Female
(c) Summer Male
(d) Summer Male
(e) Fall Small Ovigerous Female
(f) Fall Small Ovigerous Female
(g) Fall Large Ovigerous Female
(h) Fall Large Ovigerous Female
(i) Fall Small Adult Male
(j) Fall Small Adult Male
(k) Fall Large Adult Male
(l) Fall Large Adult Male
<0.25Q vs SSB
ACF
Lag

Figure 6.20. Cross correlation for proportions of low quality habitat (<0.25Q) and spawning stock biomass index (SSB) for adult life stages in summer and fall. Left panel plots are estimated with interpolated survey bottom temperature, and right panel plots are estimated with interpolated survey bottom temperature, and right panel plots are estimated with FVCOM bottom temperature data.

Conversely, the proportions of high quality habitat were positively correlated with SSB at a 2-year lag for summer female (r=0.49) and summer male (r=0.47), meaning that if the proportion of high quality habitat increased, the SSB index would increase two years later (Fig. 6.21-a and c). As for fall groups, the correlations between the proportion of high quality habitat and SSB were not significant for both small and large ovigerous females (Fig. 6.21-e and g). The proportions of high quality habitat were positively correlated with SSB at a 4-year lag for small mature male (r=0.41, maximum r=0.44 at lag 6) and large mature male (r=0.40) (Fig. 6.21-i and k).

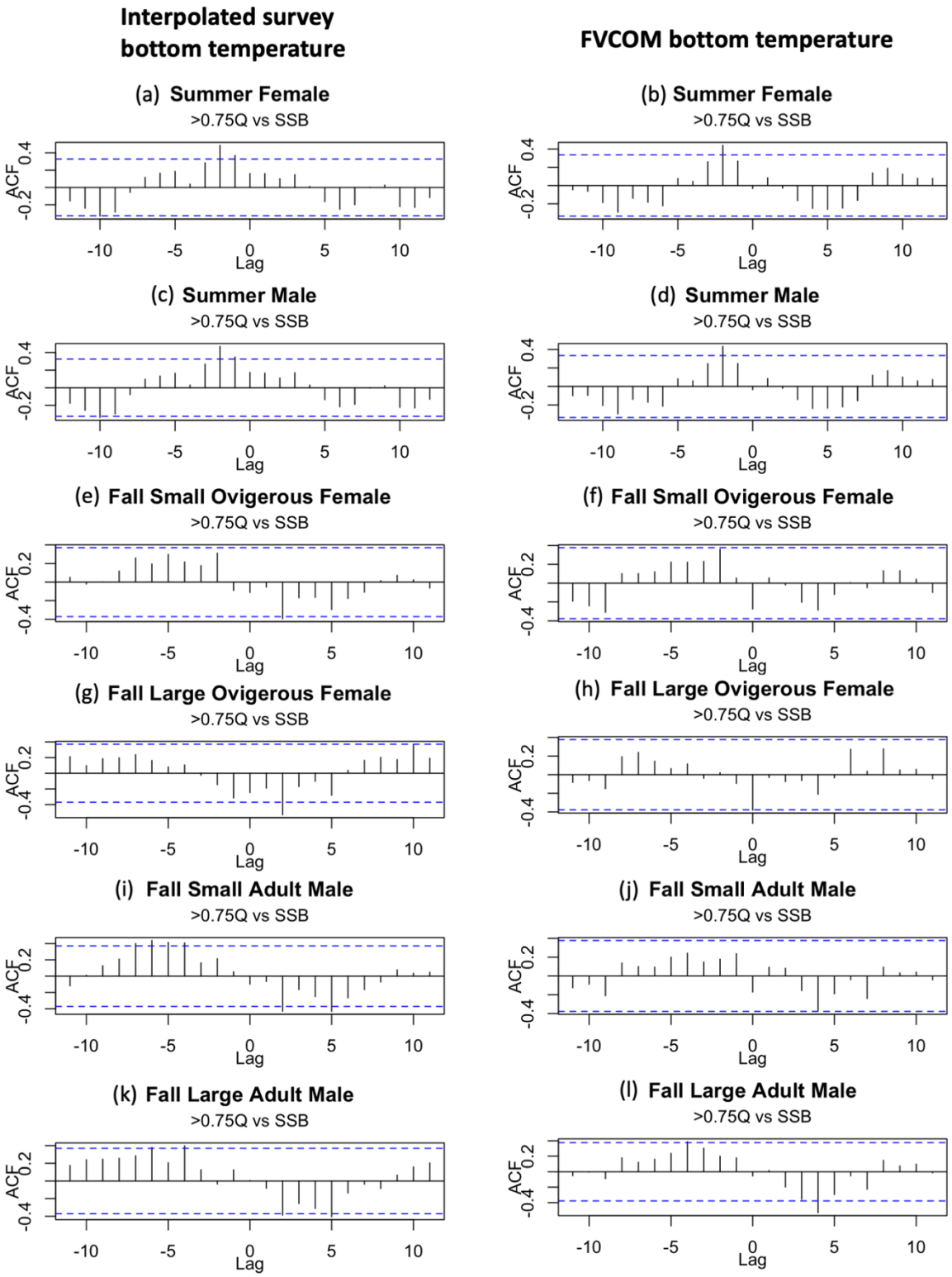

Interpolated survey bottom temperature
FVCOM bottom temperature
(a) Summer Female
(b) Summer Female
(c) Summer Male
(d) Summer Male
(e) Fall Small Ovigerous Female
(f) Fall Small Ovigerous Female
(g) Fall Large Ovigerous Female
(h) Fall Large Ovigerous Female
(i) Fall Small Adult Male
(j) Fall Small Adult Male
(k) Fall Large Adult Male
(l) Fall Large Adult Male
>0.75Q vs SSB
ACF
Lag

Figure 6.21. Cross correlation for proportions of high quality habitat (<0.75Q) and spawning stock biomass index (SSB) for adult life stages in summer and fall. Left panel plots are estimated with interpolated survey bottom temperature, and right panel plots are estimated with interpolated survey bottom temperature, and right panel plots are estimated with FVCOM bottom temperature data.

6.4.3 FVCOM bottom temperature

The HSI estimated with FVCOM data for summer female and male (Figs. 6.22-6.23) generally had lower HSI values prior to 2010 and higher HIS values after 2010 compared to those estimated with interpolated survey bottom temperature data. This is because the FVCOM yearly averaged bottom temperatures for the western GOM were generally higher than the yearly averaged survey bottom temperatures prior to 2010 for both summer and fall (Fig. 6.24). Moreover, the estimated HSI with FVCOM bottom temperature data had higher proportion of high quality HSI areas in the 2010s for fall large ovigerous female, small and large mature male, especially in 2012 (Figs. 6.25-6.28).

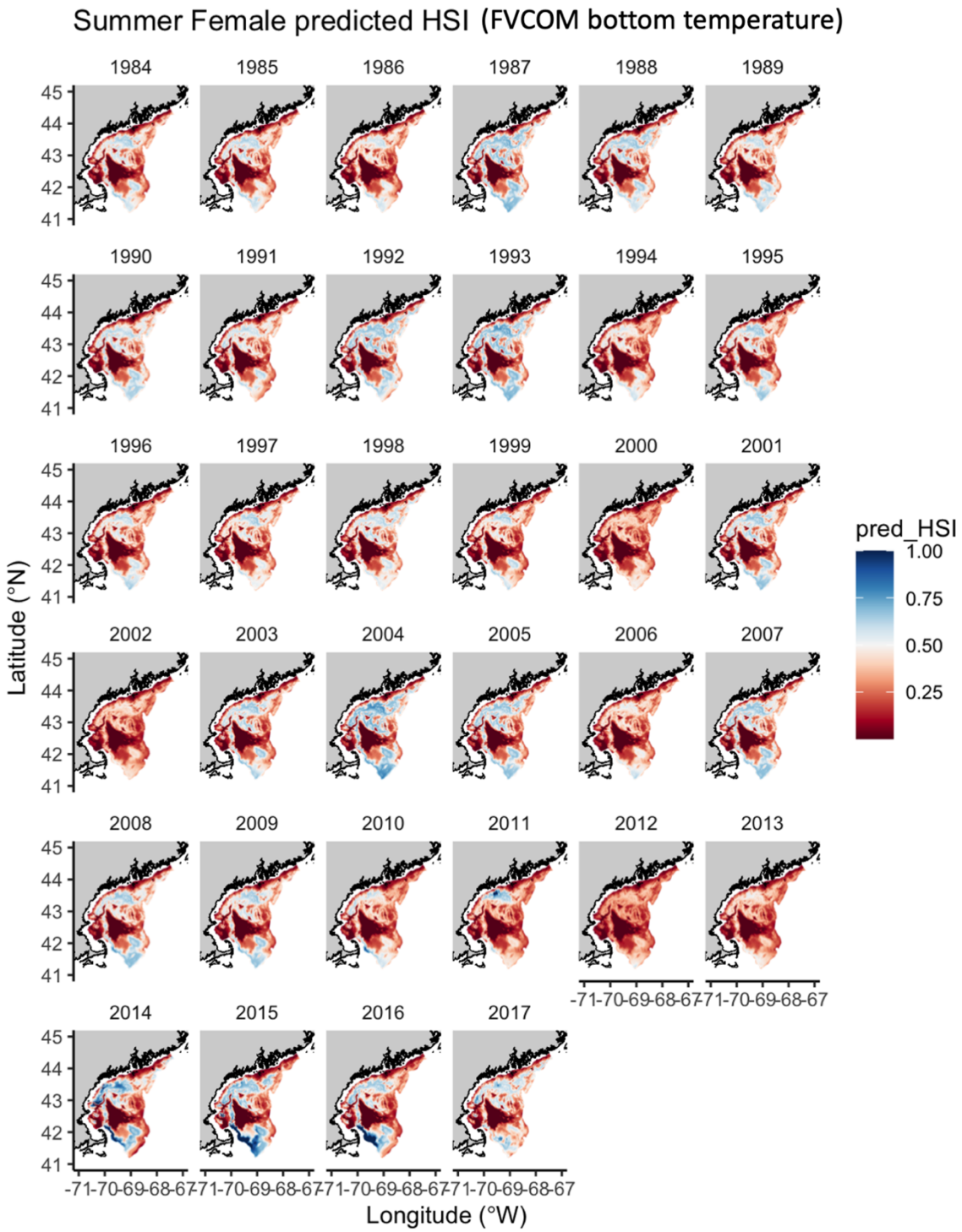


Figure 6.22. Maps of habitat suitability index (HSI) hindcasted with FVCOM bottom temperature data for summer female from 1984 to 2017.

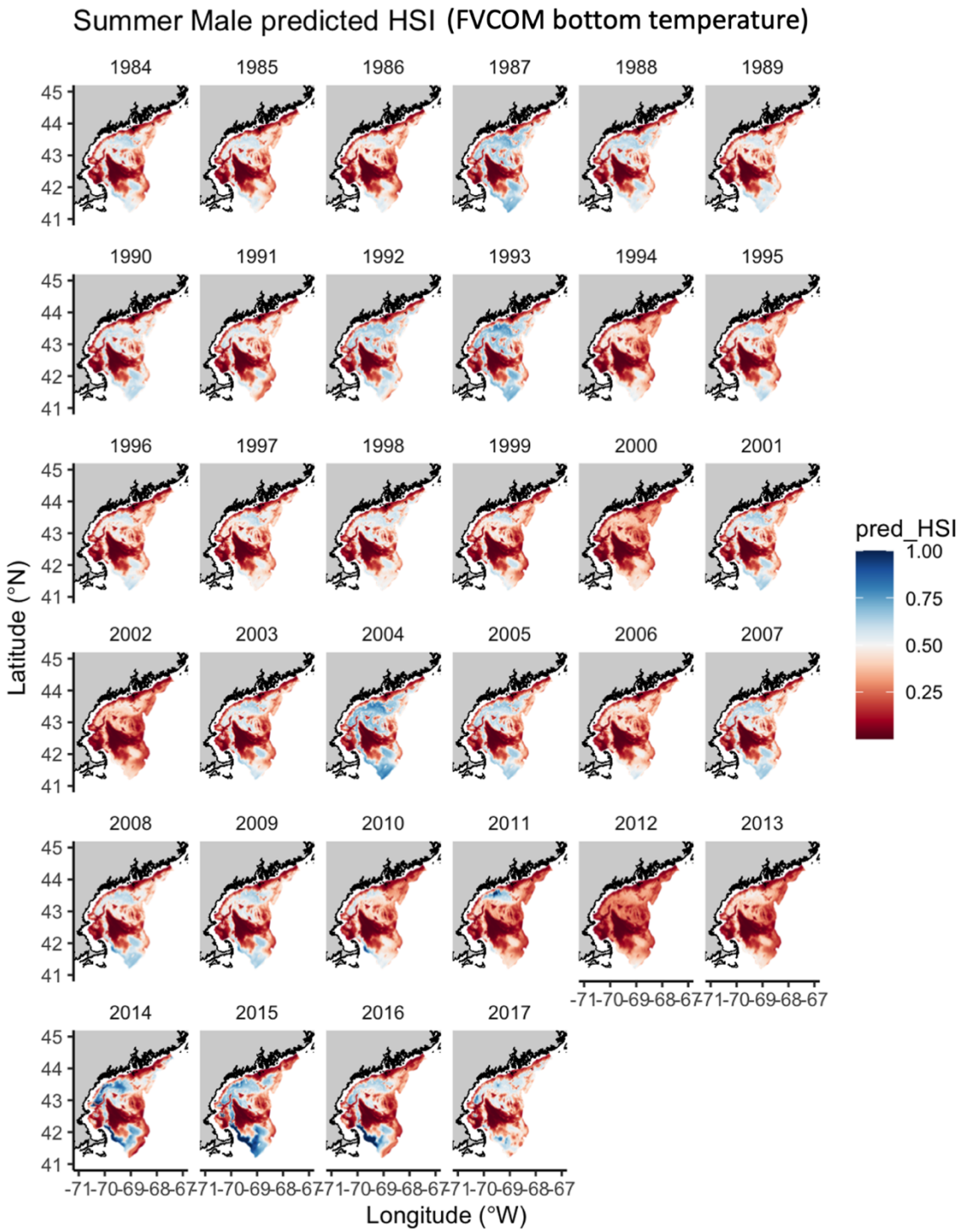


Figure 6.23. Maps of habitat suitability index (HSI) hindcasted with FVCOM bottom temperature data for summer male from 1984 to 2017.

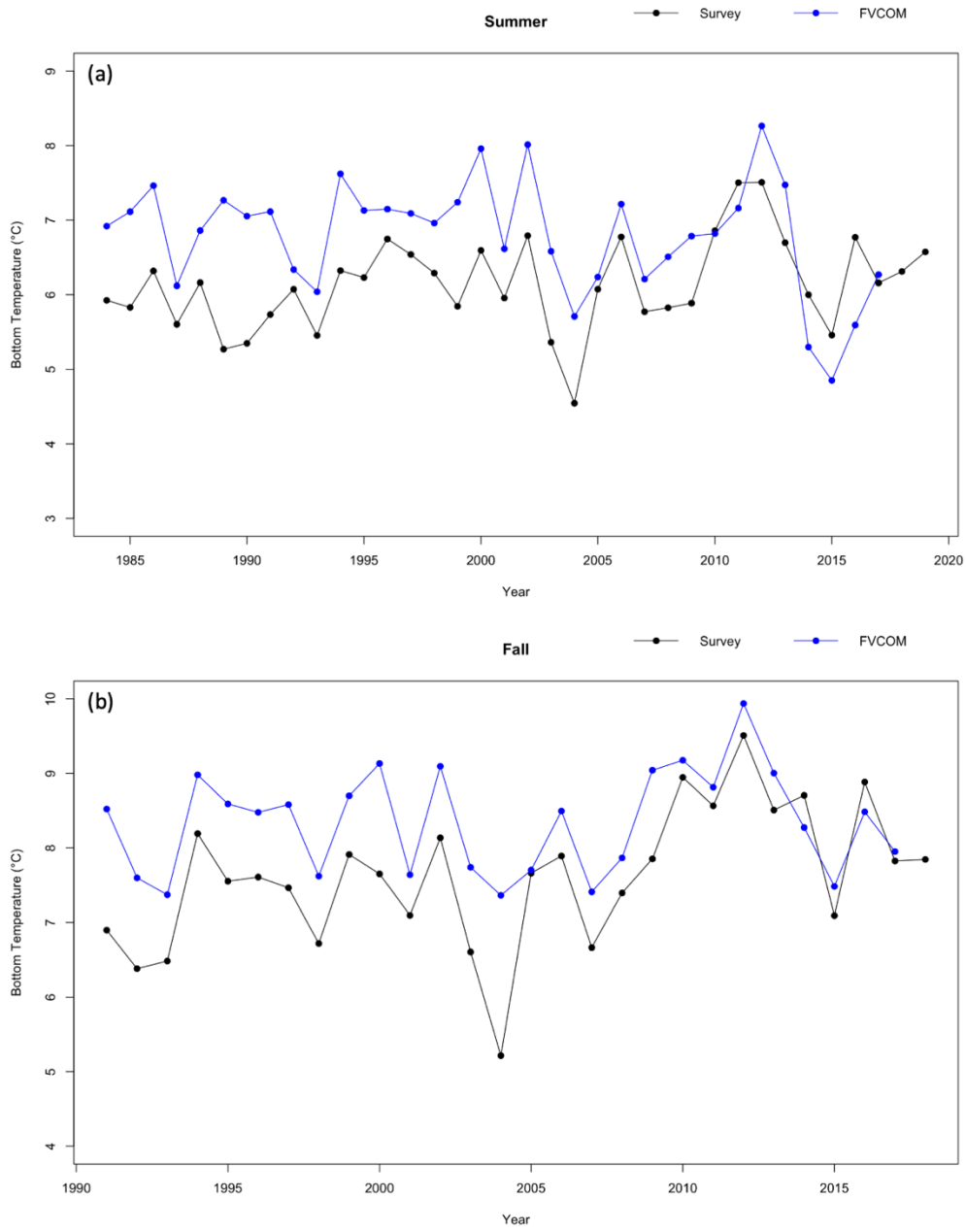


Figure 6.24. Annual average of bottom temperature of interpolated survey bottom temperature (survey, black lines) and Finite-Volume Community Ocean Model (FVCOM, blue lines) for (a) summer and (b) fall in the western Gulf of Maine.

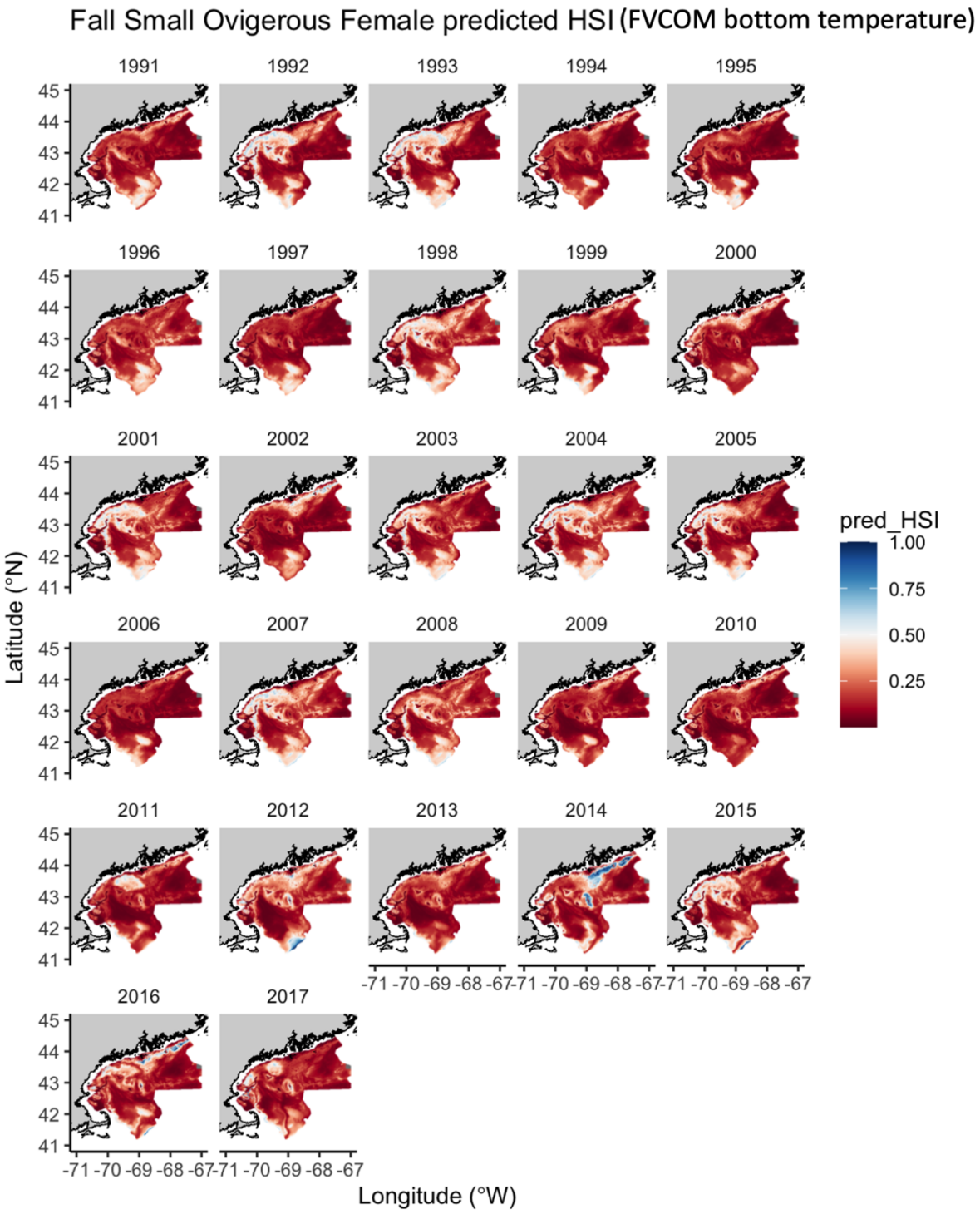


Figure 6.25. Maps of habitat suitability index (HSI) hindcasted with FVCOM bottom temperature data for fall small ovigerous female from 1991 to 2017.

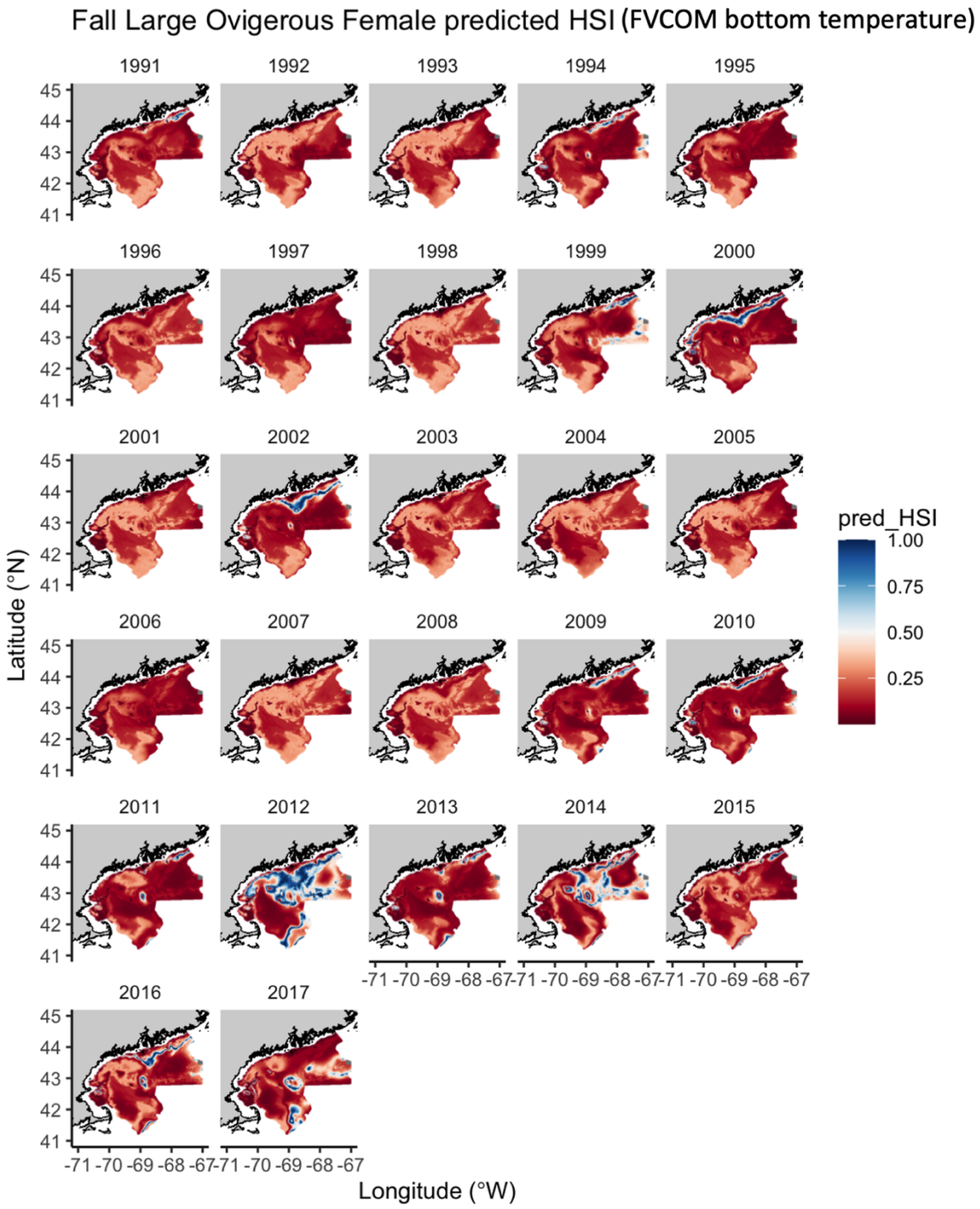


Figure 6.26. Maps of habitat suitability index (HSI) hindcasted with FVCOM bottom temperature data for fall large ovigerous female from 1991 to 2017.

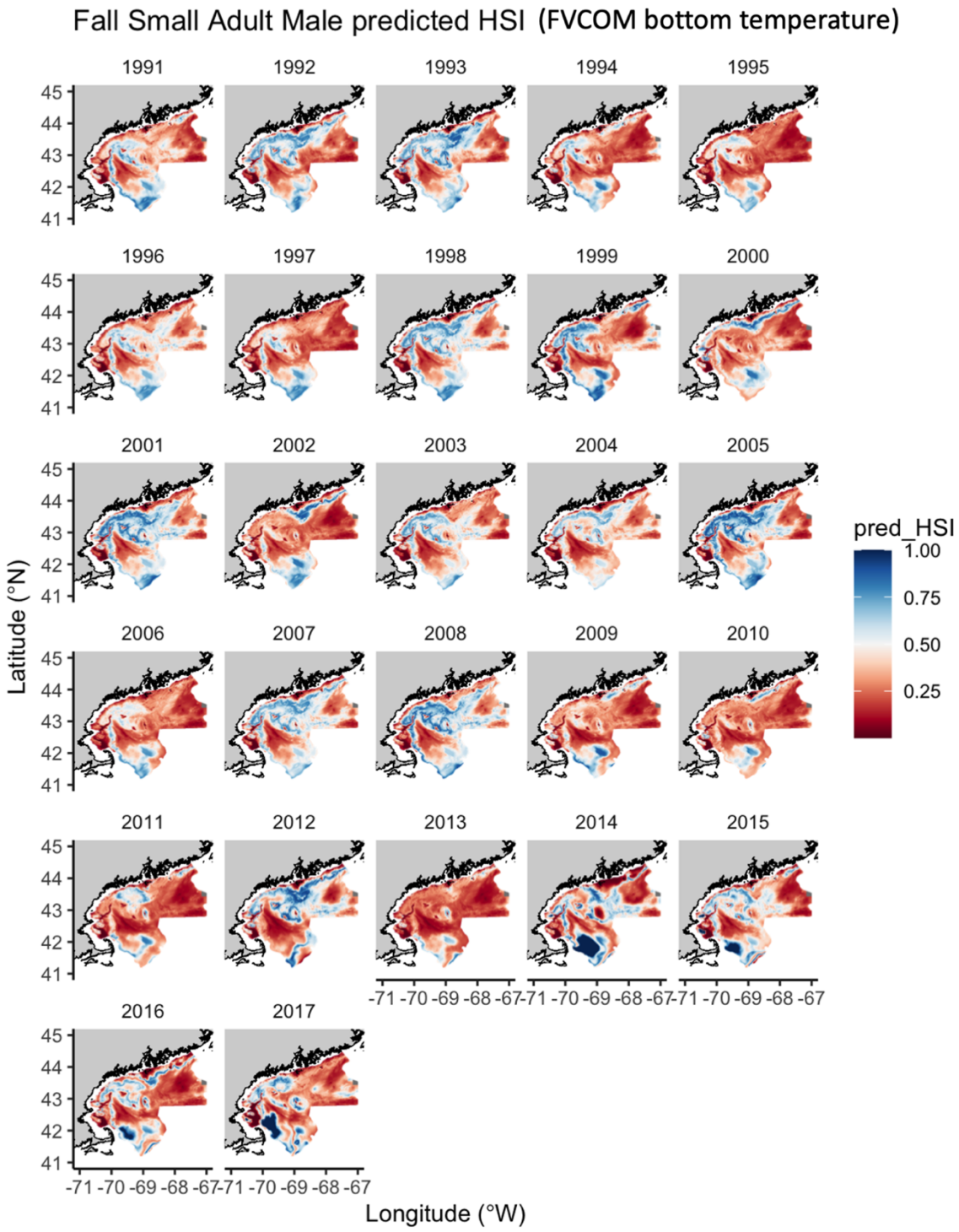


Figure 6.27. Maps of habitat suitability index (HSI) hindcasted with FVCOM bottom temperature data for fall small adult male from 1991 to 2017.

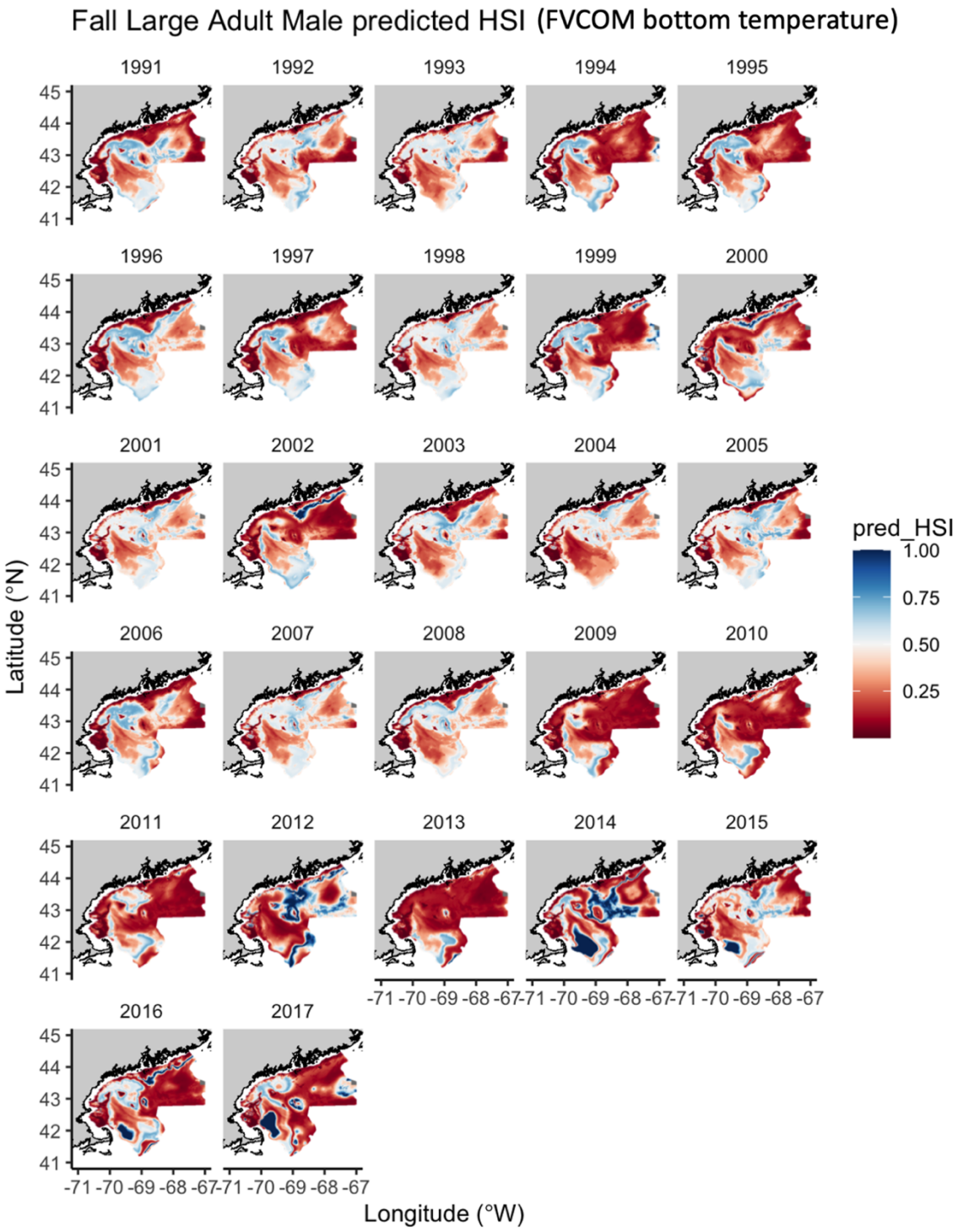


Figure 6.28. Maps of habitat suitability index (HSI) hindcasted with FVCOM bottom temperature data for fall large adult male from 1991 to 2017.

The slopes of temporal changes of HSI estimated with FVCOM data also showed a different pattern from those estimated with interpolated survey bottom temperature data. The slopes in Platts Bank and Cashes Ledge for summer female and male were still negative values and were statistically significant ($p<0.05$). However, the slopes of temporal changes for summer female and male in the Jeffreys Ledge were positive (Fig. 6.14-b and d) although they were not statistically significant ($p>0.05$, Fig. 6.15-b and d).

Discrepancies between the HSI estimated with the two data sources were also found for fall groups. The slopes estimated with FVCOM data were mostly positive values for fall groups in Jeffreys Ledge, Platts Bank, and Cashes Ledge (Figs. 6.16-6.17), although most of them were not statistically significant ($p>0.05$, Figs. 6.18-6.19).

The proportions of low quality habitat for summer groups estimated with FVCOM bottom temperature data had similar results as those estimated with interpolated survey bottom temperature data (Fig. 6.20). The proportions of low quality habitat were negatively correlated with SSB index at a 2-year lag for both Summer Female ($r=-0.48$) and Male ($r=-0.47$), meaning the increased proportion of low quality habitat would lead to decreasing SSB two years later . For fall groups, the correlations between the proportions of low quality habitat and SSB index were not significant ($p>0.05$) for small ovigerous female and small mature male (Fig. 6.20-f and j). While the proportions of low quality habitat were negatively correlated with SSB index for large ovigerous female ($r=-0.40$) at a 2-year lag and for large mature male ($r=-0.47$) at a 3-year lag (Fig. 6.20-h and l).

Similar to estimates with interpolated survey bottom temperature data, the proportions of high quality habitat estimated with FVCOM data were also positively correlated with SSB index at a 2-year lag for summer female and male (Fig. 6.21-b and d). However, it was negatively

correlated with SSB for fall large ovigerous female (r=-0.38) at no time lags (Fig. 6.21h), implying the higher the proportion of good habitat for large ovigerous female, the lower values of SSB would be found. The proportion of high quality habitat was positively correlated with SSB index for large mature male (r=0.39) at a 4-year lag, while the correlations between the proportions of high quality habitat and SSB were not statistically significant ($p>0.05$) for small ovigerous female and small mature male.

6.5 Discussion

Our study showed that the high density of northern shrimp was mostly found in Jeffreys Ledge, Platts Bank, Cashes Ledge, and Jeffreys Bank; the density became lower toward the sea in the deeper central Wilkinson Basin area. This is consistent with the observations of previous studies (Haynes and Wigley 1969; Apollonio et al. 1986), suggesting that these areas are important habitat for the GOM northern shrimp in summer and fall. Haynes and Wigley (1969) indicated that higher water temperature in the Georges Band and Cape-Cod areas may be one of the important factors that creating a thermal barrier which prevents the GOM northern shrimp from extending their distribution.

The selection of environmental factors included in the models was literature-based, and our analysis agreed on the suggestions of previous studies (Haynes and Wigley 1969; Apollonio et al. 1986). The models could explain 72.6-96.6% of the variation in HSI, and the full models that included bottom temperature and depth outcompeted the rest of the models for all life stages. Haynes and Wigley (1969) noted that the depth of northern shrimp distribution changed seasonally due to seasonal migration and may be an indirect association resulted from the

occurrence of suitable habitat of these depths. Therefore, the inclusion of depth and coordinate data was able to explain the variation in the space that was not explained by bottom temperature.

In this study, we used two sources of bottom temperature data (interpolated survey bottom temperature and FVCOM bottom temperature) for estimating HSI. It showed that the FVCOM bottom temperature tended to be higher than the interpolated survey bottom temperature in the western GOM, especially prior to 2010. The absolute difference between the yearly temperature of the two data ranged 0.04-1.99°C with a RMSE of 0.96 for summer and ranged 0.04-2.15°C with a RMSE of 0.92 for fall.

These differences might result from two possible reasons. One is the model predictive errors. Li et al. (2017) compared FVCOM bottom temperature data with Environmental Monitors on Lobster Traps (eMOLT) bottom temperature data and provided a RMSE of 2.28°C. However, the eMOLT bottom temperature sites were mostly in inshore areas, and may not be applicable of offshore areas bottom temperature data with confidence (Li et al. 2017). The second possible reason is due to the specific temperature preference of the GOM northern shrimp. Stickney and Perkins (1977) suggested that ovigerous female may gravitate toward cooler water when they encounter gradients. Apollonio et al. (1986) also indicated that older shrimp tended to avoid higher water temperature. Therefore, the ambient bottom temperature of where shrimp occurred might be lower than the average of the bottom temperature due to their preference.

More likely, the discrepancies between the HSI estimated with the two bottom temperature data were resulted from a combination of these two reasons stated above. The discrepancies of bottom temperature could lead to biased results of HSI. For example, the cross correlation between SSB index and proportion of high quality habitat estimated with interpolated survey bottom temperature for fall large ovigerous female was found not statistically significant.

However, the proportion of high quality habitat estimated with FVCOM bottom temperature for fall large ovigerous female was negatively correlated with SSB index at no time lags, implying the higher proportion of high quality habitat the higher SSB could be expected in a given year.

FVCOM data have been widely used for hindcasting distributions or HSI for many species (Decelles et al. 2015; Torre et al. 2018; Runnebaum et al. 2018; Mazur et al. 2019). These hindcasted models outputs were usually used for providing information for fisheries management or conservation. However, to our knowledge, the biases in estimated models outputs were rarely assessed or addressed. In this study, the differences in annual bottom temperature for summer and fall can be up to around 2°C. For the GOM northern shrimp which was perceived to be sensitive to temperature changes (ASMFC NSTC 2018), a difference of 2°C may lead to huge bias in estimated HSI. Furthermore, the biased estimation may not reflect the observations and result in inaccurate interpretations. Based on the existing data and our understanding on northern shrimp so far, we determined the HSI estimated with FVCOM was inaccurate and biased that may lead to incorrect inferences. We therefore considered the interpolated survey bottom temperature data more appropriate for hindcasting HSI for the GOM northern shrimp and are more likely to reflect the observations.

The collapse of northern shrimp population has been perceived to be correlated with recruitment failures and declining population biomass resulting from unfavorably warm water temperatures (ASMFC NSTC 2018). Water temperature may have impacts on northern shrimp life cycle in various aspects. Water temperature effects on the incidence of parasites (white eggs) of the GOM northern shrimp were examined, yet the correlations between white eggs and environmental factors were found not significant (Chang et al. 2020). Chang et al. (unpublished) indicated that the annual population fecundity was found to be lower than those estimated four

decades ago, which might be due to environmental changes over the past four decades as well as different analysis methods used. The warming water temperature has shifted hatching timing and duration of the GOM northern shrimp (Richards 2012), although Richards et al. (2016) indicated that match-mismatch theory (Hjort 1914; Cushing 1990) did not seem able to explain shrimp survival at early life stages. Richards and Hunter (2021) examined the spatial overlaps in distributions of the GOM northern shrimp and their major predators, and hypothesized that the distributional shift of longfin squid (*Doryteuthis pealeii*) after the heatwave in 2012 might be the reason of shrimp population collapse due to spatial overlap and subsequent predation.

In this study, we provide an alternative hypothesis that the deterioration of habitat quality for adult shrimp in summer and fall (spawning season) might also be one of the reasons that the shrimp population collapsed. The hindcasted HSI maps showed that the HSI was at lower levels in the 2010s for most groups in summer and fall. The increase in the proportion of low quality habitat and the decrease in the proportion of high quality habitat were negatively and positively associated with SSB index at 2-4 year time lags, respectively. Our results suggest that the Jeffreys Ledge, Platts Bank, Cashes Ledge, and Jeffreys Bank are very important habitat for adult shrimp in summer and fall when the spawning takes place. The deterioration of these areas would have a negative impact on spawners.

It should be noted that the classification of small and large ovigerous female and mature male was based on the average sizes of fall ovigerous female and mature male, which could be somewhat arbitrary. Different thresholds used could influence the model outcome. Therefore, further investigation may be needed for a better understanding of northern shrimp distribution and gathering behaviors based on their life cycle.

BIBLIOGRAPHY

Aanes, S., and J. H. Vølstad. 2015. Efficient statistical estimators and sampling strategies for estimating the age composition of fish. Can. J. Fish. Aquat. Sci. **72**: 938-953.

Ahamed, F., and J. Ohtomi. 2011. Reproductive biology of the pandalid shrimp *Plesionika izumiae* (Decapoda: Caridea). J. Crust. Biol. 31, 441-449.

Akaike H. 1973. Maximum likelihood identification of Gaussian autoregressive moving average models. Biometrika. 60: 255-265.

Anderson, D. R., and K. P. Burnham. 2002. Avoiding pitfalls when using information-theoretic methods. J. Wildl. Manage. 66, 912–918.

Apollonio, S., and E. E. Dunton. 1969. The northern shrimp *Pandalus borealis* in the Gulf of Maine. Maine Department of Marine Research completion report: project 3-12 R, 81pp.

Apollonio, S., D. K. Stevenson, and E. E. Dunton. 1986. Effects of temperature on the biology of the northern shrimp, *Pandalus borealis*, in the Gulf of Maine. NOAA Tech. Rep. NMFS 42, Washington, DC, 22 pp.

Arlot, S., Celisse, A., 2010. A survey of cross-validation procedures for model selection. Stat. Surv. 4, 40-79.

ASMFC NSTC (Atlantic States Marine Fisheries Commission Northern Shrimp Technical Committee). 2018. Assessment report for Gulf of Maine northern shrimp.

ASMFC NSTC (Atlantic States Marine Fisheries Commission Northern Shrimp Technical Committee). 2019. Assessment report for Gulf of Maine northern shrimp.

Azarovitz T. R. 1981. A brief historical review of the Woods Hole Laboratory trawl survey time series. Can. Spec. Publ. Fish. Aquat. Sci. 58: 62-67.

Barneche, D. R., D. R. Robertson, C. R. White, and D. J. Marshall. 2018. Fish reproductive-energy output increases disproportionately with body size. *Science*, **360**: 642-645.

Barr DJ, R. Levy, C. Scheepers, and H. J. Tily. 2013. Random Effects Structure for Confirmatory Hypothesis Testing: Keep It Maximal. J Mem Lang. 68(3): 255-78. doi:10.1016/j.jml.2012.11.001

Bartoń K. 2019. MuMIn: Multi-model inference. R package version 1.43.6. Retrieved from http://cran.r-project.org/package=MuMIn.

Beninger, P. G., I. Boldina, and S. Katsanevakis. 2012. Strengthening statistical usage in marine ecology. *Journal of J. Exp. Mar. Biol. Ecol.* **426-427**: 97-108

Bigelow, H. B. 1927. Physical oceanography of the Gulf of Maine. Bull U.S. Bur Fish 40: 511-1027.

Bivand, R. 2020. classInt: Choose Univariate Class Intervals. R package version 0.4-3. https://CRAN.R-project.org/package=classInt

Bolker B. 2008. Ecological Models and Data in R. Princeton University Press.

Bolker B.M., M. E. Brooks, C. J. Clark, S. W. Geange, J. R. Poulsen, M. H. H. Stevens, and J. S. S. White. 2009. Generalized linear mixed models: a practical guide for ecology and evolution. Trends Ecol. Evol. 24: 127-135.

Boyer, T. P., O. K. Baranova, C. Coleman, H. E. Garcia, A. Grodsky, R. A. Locarnini, A. V. Mishonov, C. R. Paver, J. R. Reagan, D. Seidov, I. V. Smolyar, K. Weathers, M. M. Zweng. 2018. World Ocean Database 2018. A.V. Mishonov, Technical Ed., NOAA Atlas NESDIS 87. https://www.ncei.noaa.gov/sites/default/files/2020-04/wod_intro_0.pdf

Brillon, S., Y. Lambert, and J. Dodson. 2005. Egg survival, embryonic development, and larval characteristics of northern shrimp (*Pandalus borealis*) females subject to different temperature and feeding conditions. Mar. Biol. 147, 895–911.

Brosi, B. J., and E. G. Biber. 2009. Statistical inference, type II error, and decision making under the US Endangered Species Act. *Front Ecol Environ*. **7**: 487-494.

Brosset, P., H. Bourdages, M. Blais, M. Scarratt, and S. Plourde. 2019. Local environment affecting northern shrimp recruitment: a comparative study of Gulf of St. Lawrence stocks. ICES J. Mar. Sci. 76, 974−986.

Brown W. S., and R. A. Beardsley. 1978. Winter circulation in the western Gulf of Maine. Part 1. Winter cooling and water mass formation. J Physical Oce 8: 265-277.

Camargo, J. A. 1995. On measuring species evenness and other associated parameters of community structure. Oikos 74:538–542.

Cao, J., J. T. Thorson, R. A. Richards, and Y. Chen. 2017. Spatiotemporal index standardization improves the stock assessment of northern shrimp in the Gulf of Maine. Can. J. Fish. Aquat. Sci. 74:1781-1793.

Chang, H.-Y., and Y. Chen. 2020. Evaluating sampling strategies for collecting size-based fish fecundity data: an example of Gulf of Maine northern shrimp *Pandalus borealis*. J. Northw. Atl. Fish. Sci. 51, 33–43. doi:10.2960/J.v51.m730

Chang, H.-Y., R. Klose, and Y. Chen. 2020. Possible climate-induced environmental impacts on parasite-infection rates of northern shrimp *Pandalus borealis* eggs in the Gulf of Maine. Dis. Aquat. Organ. 140, 109-118. doi: 10.3354/dao03495. PMID: 32701067.

Clark, S. H. 1989. State-Federal northern shrimp survey. In: Proceedings of a workshop on bottom trawl surveys. Azarovitz, T. R., J. McGurrin and R. Seagraves (eds.). ASMFC Spec. Rept. 17: 27–29.

Clark, S. H., S. X.Cadrin, D. F. Schick, P. J. Diodati, M. P. Armstrong, and D. McCarron. 2000. The Gulf of Maine northern shrimp (*Pandalus borealis*) fishery: a review of the record. J. Northwest. Atl. Fish. Sci. 27, 193−226.

Clarke, A., C. C. E. Hopkins, and E. M. Nilssen. 1991. Egg size and reproductive output in the deepwater prawn Pandalus borealis Krøyer, 1838. Funct. Ecol. 5, 724–730.

Collins, L. A., A. G. Johnson, C. C. Koenig, and M. S. Baker. 1998. Reproductive patterns, sex ratio, and fecundity in gag, *Mycteroperca microlepis* (Serranidae), a protogynous grouper from the northeastern Gulf of Mexico. *Fish. Bull.* **96**: 415-427.

Cressie, N. A. C. 1993. Statistics for Spatial Data, Wiley

Cushing, D. H. 1990. Plankton production and year-class strength in fish populations: an update of the match/mismatch hypothesis. Adv. Mar. Biol. 26:249–293.

Dautov, S. S., L. I. Popova, and A. I. Begalov. 2004. Fecundity of the Grass Shrimp, *Pandalus kessleri* (Decapoda: Pandalidae), near the Southern Kuril Islands. Russ. J. Mar. Biol. 30, 199-203. https://doi.org/10.1023/B:RUMB.0000033956.38122.4a

de los Ríos, C., J. E. M. Watson, and N. Butt. 2018. Persistence of methodological, taxonomical, and geographical bias in assessments of species' vulnerability to climate change: a review. Glob. Ecol. Conserv. 15, e00412. https://doi.org/10.1016/j.gecco.2018.e00412

DeCelles, G., G. Cowles, C. Liu, and S. Cadrin. 2015. Modeled transport of winter flounder larvae spawned in coastal waters of the Gulf of Maine. Fish. Oceanogr., 24: 430-444. https://doi-org.wv-o-ursus-proxy02.ursus.maine.edu/10.1111/fog.12120

Di Stefano, J. 2003. How much power is enough? Against the development of an arbitrary convention for statistical power calculations. *Funct. Ecol.* **17**: 707-709.

Dow, R. L. 1977. Effects of climatic cycles on the relative abundance and availability of commercial marine and estuarine species. J Cons Cons Int Explor Mer 37:274-280.

Dow, R. L. 1981. Shrimp management in the Gulf of Maine. In T. Frady (editor), Proceedings of the International Pandalid Shrimp Symposium, Kodiak, Alaska, February 13-15, 1979, p. 59-61. Univ. Ataska, Sea Grant Rep. 81-3.

Draugelis-Dale, R. 2008. Assessment of effectiveness and limitations of habitat suitability models for wetland restoration: U.S. Geological Survey Open-File Report 2007-1254, 136 p.

Elliot, D. L. 1970. Fecundity of the northern shrimp, *Pandalus borealis*. BSc thesis, Bowdoin University, Brunswick, ME

Fisher, R., S. K. Wilson, T. M. Sin, A. C. Lee, and T. J. Langlois. 2018. A simple function for full-subsets multiple regression in ecology with R. Ecol Evol. 8, 6104-6113.

Friedland, K. D., R. E. Morse, N. Shackell, J. C. Tam, J. L. Morano, J. R. Moisan, and D. C. Brady. 2020. Changing Physical Conditions and Lower and Upper Trophic Level Responses on the US Northeast Shelf. Front. Mar. Sci. 7:567445. doi: 10.3389/fmars.2020.567445

Gräler B, E. Pebesma, and G. Heuvelink. 2016. Spatio-Temporal Interpolation using gstat. The R Journal 8(1): 204-218.

Grosslein M. G. 1969. Groundfish survey program of BCF at Woods Hole. Commer. Fish. Rev. 31: 22-35.

Hannah, R. W., S. A. Jones, and M. R. Long. 1995. Fecundity of the ocean shrimp (*Pandalus jordani*). Can. J. Fish. Aquat. Sci. **52**: 2098-2107.

Hare, J. A., and K. W. Able. 2007. Mechanistic links between climate and fisheries along the east coast of the United States: explaining population outbursts of Atlantic croaker (*Micropogonias undulatus*). Fisheries Oceanography 16:31-45.

Hare, J. A., W. E. Morrison, M. W. Nelson, M. M. Stachura, E. J. Teeters, and R. B. Griffis et al. 2016. A Vulnerability Assessment of Fish and Invertebrates to Climate Change on the Northeast U.S. Continental Shelf. PLoS ONE 11(2): e0146756. https://doi.org/10.1371/journal.pone.0146756

Hastie, T. J., and R. J. Tibshirani. 1990. Generalized additive models. Chapman & Hall, London

Haynes, E. B., and R. L. Wigley. 1969. Biology of the northern shrimp, *Pandalus borealis*, in the Gulf of Maine. *rans. Am. Fish. Soc.* **98**: 60-76

Hilborn, R., and C. J. Walters. 1992. Quantitative fisheries stock assessment, choice, dynamics and uncertainty. Chapman and Hall, London. doi:10.1007/978-1-4615-3598-0

Hixon, M. A., D. W. Johnson, and S. M. Sogard. 2014. BOFFFFs: on the importance of conserving old-growth age structure in fishery populations. *ICES J. Mar. Sci.* **71**: 2171-2185.

Hjort, J. 1914. Fluctuations in the great fisheries of northern Europe, viewed in the light of biological research. Rap. Proces. 20:1–228.

Hopkins, T. S., and N. Garfield III. 1979. Gulf of Maine intermediate water. J Mar Res 37: 103-139.

Hurlbert, S. H. 1984. Pseudoreplication and the design of ecological field experiments. Ecol. Monogr. 54, 187-211.

Hurvich, C. M., and C.-L. Tsai. 1989. Regression and time series model selection in small samples. Biometrika 76, 297–307. https://doi.org/10.1093/biomet/76.2.297

Hyndman, R., G. Athanasopoulos, C. Bergmeir, G. Caceres, L. Chhay, M. O'Hara-Wild, F. Petropoulos, S. Razbash, E. Wang, and F. Yasmeen. 2021. forecast: Forecasting functions for time series and linear models. R package version 8.14, https://pkg.robjhyndman.com/forecast/.

IPCC. 2014. Climate Change 2014: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change(Cambridge: Cambridge University Press)

Jónsdóttir I. G. 2018. Effects of changes in female size on relative egg production of northern shrimp stocks (*Pandalus borealis*). Reg. Stud. Mar. Sci. 24, 270-277

Jónsdóttir, I. G., Á. Magnússon, and U. Skúladóttir. 2013. Influence of increased cod abundance and temperature on recruitment of northern shrimp (*Pandalus borealis*). Mar Biol 160(5): 1203-1211.

Jorde, P.E., G. Søvik, J.-I. Westgaard, J. Albretsen, C. André, C. Hvingel, T. Johansen, A. D. Sandvik, M. Kingsley, and K. E. Jørstad. 2015. Genetically distinct populations of northern shrimp, *Pandalus borealis*, in the North Atlantic: adaptation to different temperatures as an isolation factor. Mol Ecol, 24: 1742-1757. https://doi.org/10.1111/mec.13158

Kavanaugh, M. T., J. E. Rheuban, K. M. A. Luis, and S. C. Doney. 2017. Thirty-three years of ocean benthic warming along the U.S. Northeast Continental Shelf and Slope: Patterns, drivers, and ecological consequences. J. Geophys. Res. 122, 9399–9414.

Koeller, P., C. Fuentes-Yaco, and T. Platt. 2007. Decreasing shrimp sizes off Newfoundland and Labrador—environment or fishing? *Fish. Oceanogr.* **16**: 105-115.

Kleisner, K. M., M. J.Fogarty, S. McGee, J. A. Hare, S. Moret, C. T. Perretti, and V. S. Saba. 2017. Marine species distribution shifts on the U.S. Northeast Continental Shelf under continued ocean warming. Prog. Oceanogr. 153, 24−36.

Lakens, D. 2017. Equivalence Tests: A practical primer for t tests, correlations, and meta-analyses. *Soc Psychol Personal Sci*. **8**: 355-362.

Lee R. F, A. N. Walker, S. C. Landers, T. L. Walters, S. A. Powell, and M. E. Frischer. 2019. Black spot gill syndrome in the northern shrimp, *Pandalus borealis*, caused by the parasitic ciliate *Synophrya* sp.. J. Invesl. Pathol. 161: 40-46.

Li, B., K. R. Tanaka, Y. Chen, D. C. Brady, and A.C. Thomas. 2017. Assessing the quality of bottom water temperatures from the Finite-Volume Community Ocean Model (FVCOM) in the Northwest Atlantic Shelf region. Journal of Marine Systems. 173: 21-30.

Luke S. G. 2017. Evaluating significance in linear mixed effects models in R. Behav. Res. Methods. 49(4):1494-1502.

Marshall, C. T., C. L. Needle, A. Thorsen, O. S. Kjesbu, and N. A. Yaragina. 2006. Systematic bias in estimates of reproductive potential of an Atlantic cod (*Gadus morhua*) stock: implications for stock-recruit theory and management. Can. J. Fish. Aquat. Sci. **63**: 980-994.

Martinez-Abrain, A. 2008. Statistical significance and biological relevance: a call for a more cautious interpretation of results in ecology. *Acta Oecol*. **34**: 9-11.

Mazur, M., B. Li, J.-H. Chang, and Y. Chen. 2018. Using an individual-based model to simulate the Gulf of Maine American lobster (*Homarus americanus*) fishery and evaluate the robustness of current management regulations. Can. J. Fish. Aquat. Sci. 76, 1709-1718. https://doi.org/10.1139/cjfas-2018-0122

Mazur, M., B. Li, J.-H. Chang, and Y. Chen. 2019. Contributions of a conservation measure that protects the spawning stock to drastic increases in the Gulf of Maine American lobster fishery. Mar Ecol Prog Ser 631:127-139. https://doi-org.wv-o-ursus-proxy02.ursus.maine.edu/10.3354/meps13141

McCrary J. A. 1971. Sternal spines as a characteristic for differentiating between females of some Pandalidae. J Fish Res Board Can 28:98–100.

Meyers T. R., D. V. Lightner, and R. M. Redman. 1994. A dinoflagellate-like parasite in Alaskan spot shrimp *Pandalus platycerosand* pink shrimp P. borealis. Dis Aquat Org 18:71–76.

Mills, K. E., A. J..Pershing, C. J. Brown, Y. Chen, F.-S. Chiang, D. S. Holland, S. Lehuta, J. A. Nye, J. C. Sun, A. C. Thomas, and R. A. Wahle, 2013. Fisheries management in a changing climate: Lessons from the 2012 ocean heat wave in the Northwest Atlantic. Oceanogr. 26, 191-195.

Morgan, M. J., A. Perez-Rodriguez, and F. Saborido-Rey. 2012. Does increased information about reproductive potential result in better prediction of recruitment? Can. J. Fish. Aquat. Sci. 68, 1361-1368. https://doi.org/10.1139/f2011-049

Muggeo, V. M. R. 2003. Estimating regression models with unknown breakpoints. Stat. Med. 22, 3055–3071. doi: 10.1002/sim.1545

Muggeo, V. M. R. 2017. Interval estimation for the breakpoint in segmented regression: a smoothed score-based approach. Aust. N.Z. J. Stat. 59, 311–322. doi: 10.1111/anzs.12200

Nakagawa, S., and I. C. Cuthill. 2007. Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biol.* **82**: 591-605.

Nakagawa S, P. C. D. Johnson, H. and Schielzeth. 2017. The coefficient of determination R2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. J R Soc Interface 14(134): 20170213.

Nally R. M, R. P. Duncan, J. R. Thomson, and J. D. L. Yen. 2018. Model selection using information criteria, but is the best model any good? J. Appl. Ecol. 55: 1441-1444.

Nunes, P. 1984. Reproductive and larval biology of northern shrimp *Pandalus borealis* (Krøyer) in relation to temperature. Ph.D. thesis, Univ. Alaska, Fairbanks, 195 p.

Nye, J. 2010. Climate change and its effects on ecosystems, habitats and biota. State of the Gulf of Maine report. Woods Hole: NOAA NMFS NEFSC.

O'Brien, L. 1999. Factors influencing the rate of sexual maturity and the effect on spawning stock for Georges Bank and Gulf of Maine Atlantic cod *Gadus morhua* stocks. *J. Northwest Atl. Fish. Sci.* **25**: 179-203.

Parkhurst, D. F. 2001. Statistical significance tests: equivalence and reverse tests should reduce misinterpretation. *Bioscience*. **51**: 1051-1057

Parsons, D. G., and G. E. Tucker. 1986. Fecundity of northern shrimp, *Pandalus borealis*, (Crustacea, Decapoda) in areas of the Northwest Atlantic. *Fish. Bull.* 84: 549-558.

Pasch B, B. M. Bolker, and S. M. Phelps. 2013. Interspecific Dominance via Vocal Interactions Mediates Altitudinal Zonation in Neotropical Singing Mice. Am. Nat. 182 (5): E161–E173. doi:10.1086/673263

Payne, L. X., D. E. Schindler, J. K. Parrish, and S. A. Temple. 2005. Quantifying spatial pattern with evenness indices. Ecological Applications, 15: 507-520. https://doi.org/10.1890/03-5029

Pebesma E. J. 2004. Multivariable geostatistics in S: the gstat package. Comput Geosci, 30: 683-691.

Pennington, M., L-M. Burmeister, and V. Hjellvik. 2002. Assessing the precision of frequency distributions estimated from trawl-survey samples. *Fish. Bull.* **100**: 74-80.

Pennington, M., and K. Helle. 2011. Evaluation of the design and efficiency of the Norwegian self-sampling purse-seine reference fleet. *ICES J. Mar. Sci.* **68**: 1764-1768.

Pereira, R. T., A. C. Almeida, G. M. Teixeira, A. A. D. P. Bueno, and A. Fransozo. 2017. Reproductive strategy of the shrimp *Nematopalaemon schmitti* (Decapoda, Caridea, Palaemonoidea) on the southeastern coast of Brazil. Nauplius 25, e2017003. Epub April 06, 2017.https://dx.doi.org/10.1590/2358-2936e2017003

Pershing, A. J, M. A. Alexander, C. M. Hernandez, L. A. Kerr, A. Le Bris, K. E. Mills, J. A. Nye, N. R. Record, H. A. Scannell, J. D. Scott, G. D. Sherwood, and A. C. Thomas. 2015. Slow adaptation in the face of rapid warming leads to collapse of the Gulf of Maine cod fishery. Science 350: 809-812.

Peterman, R. M. 1990. Statistical power analysis can improve fisheries research and management. Can. J. Fish. Aquat. Sci. **47**: 2-15.

Politis, P. J, J. K. Galbraith, P. Kostovick, and R. W. Brown. 2014. Northeast Fisheries Science Center bottom trawl survey protocols for the NOAA Ship Henry B. Bigelow. Northeast Fish Sci Cent Ref Doc. 14-06; 138 p. Online at: https://doi.org/10.7289/V5C53HVS

R Core Team. 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Ramirez-Llodra, E., P. A. Tyler, and J. T. P. Copley. 2000. Reproductive biology of three caridean shrimp, *Rimicaris ortunatea*, *Chorocaris chacei* and *Mirocaris ortunatea* (Caridea: Decapoda), from hydrothermal vents. J. Mar. Biolog. 80, 473-484.

Reuchlin-Hugenholtz E, N. L. Shackell, and J. A. Hutchings. 2015. The Potential for Spatial Distribution Indices to Signal Thresholds in Marine Fish Biomass. PLoS ONE 10(3): e0120500. https://doi.org/10.1371/journal.pone.0120500

Richards, R. A. 2012. Phenological shifts in hatch timing of northern shrimp *Pandalus borealis*. Mar. Ecol. Prog. Ser. 456, 149-158.

Richards, R. A., M. J. Fogarty, D. G. Mountain, and M. H. Taylor. 2012. Climate change and northern shrimp recruitment variability in the Gulf of Maine. *Mar. Ecol. Prog. Ser.* **464**:167-178

Richards, R. A., and M. Hunter. 2021. Northern shrimp *Pandalus borealis* population collapse linked to climate-driven shifts in predator distribution. PLoS ONE 16(7): e0253914. https://doi.org/10.1371/journal.pone.0253914

Richards, R. A., J. E. O'Reilly, and K. J. Hyde. 2016. Use of satellite data to identify critical periods for early life survival of northern shrimp in the Gulf of Maine. Fish. Oceanogr., 25: 306-319. https://doi.org/10.1111/fog.12153

Rinaldo R. G., and P. Yevich. 1974. Black spot gill syndrome of the northern shrimp *Pandalus borealis*. J. Invesl. Pathol. 24: 224-233.

Rogers, R., S. Rowe, R. M. Rideout, and M. J. Morgan. 2019. Fecundity of haddock (*Melanogrammus aeglefinus*) off southern Newfoundland. *Fish. Res.* 220: 105339.

Runnebaum, J., L. Guan, C. Jie, L. O'Brien, and Y. Chen. 2017. Habitat suitability modeling based on a spatiotemporal model: an example for cusk in the Gulf of Maine. Can. J. Fish. Aquat. Sci. 75(11): 1784-1797. https://doi.org/10.1139/cjfas-2017-0316

Schmiing, M., P. Afonso, F. Tempera, and R. S. Santos. 2013. Predictive habitat modelling of reef fishes with contrasting trophic ecologies. Mar Ecol Prog Ser 474:201-216. https://doi.org/10.3354/meps10099

Schneider C. A., W. S. Rasband, and K. W. Eliceiri. 2012. "NIH Image to ImageJ: 25 years of image analysis", Nature methods 9(7): 671-675, PMID 22930834

Shelton, A. O., S. B. Munch, D. Keith, and M. Mangel. 2012. Maternal age, fecundity, egg quality, and recruitment: linking stock structure to recruitment using an age-structured Ricker model. *Can. J. Fish. Aquat. Sci.* **69**: 1631-1641.

Shields, J. D. 1994. The parasitic dinoflagellates of marine crustaceans. Annu Rev Fish Dis 4:241-271

Shumway, S. E., H. C. Perkins, D. F. Schick, and A. P. Stickney. 1985. Synopsis of biological data on the pink shrimp, *Pandalus borealis* Krøyer, 1838. NOAA Tech Rep NMFS 30

Skuladottir, U., E. Jonsson, and I. Hallgrimsson. 1978. Testing for heterogeneity of *Pandalus borealis* populations at Iceland. ICES CM. 1978/K:27, 8 p

Smith, P. C. 1983. The mean and seasonal circulation off southwest Nova Scotia. J Phys Oceanogr 13: 1034-1054.

Smith, P. C. 1989. Seasonal and interannual variability of current, temperature and salinity off southwest Nova Scotia. Can J Fish Aq Sci 46(S1): s4-s20.

Smith, T. D. 2002. The Woods Hole bottom-trawl resource survey: development of fisheries-independent multispecies monitoring. *ICES Mar. Sci. Symp.* **215**:480-488.

Staples, K.W., Y. Chen, D. W. Townsend, and D. C. Brady, 2019. Spatiotemporal variability in the phenology of the initial intra-annual molt of American lobster (*Homarus americanus* Milne Edwards, 1837) and its relationship with bottom temperatures in a changing Gulf of Maine. Fish. Oceanogr. 28, 468–485. https://doi.org/10.1111/fog.12425

Staudinger, M. D., K. E. Mills, and K. Stamieszkin. 2019. It's about time: A synthesis of changing phenology in the Gulf of Maine ecosystem. Fish. Oceanogr. 28, 532–566.

Stauffer, G. 2004. NOAA Protocols for Groundfish Bottom Trawl Surveys of the Nation's Fishery Resources. U.S. Dep. Commerce, NOAA Tech. Memo. NMFS-F/SPO-65, 205 p.

Stickney, A. P. 1978. A previously unreported peridinian parasite in the eggs of the northern shrimp, *Pandalus borealis*. J. Invesl. Pathol. 32: 212-215.

Stickney, A. P. 1980. A characterization of the northern shrimp fishery of Maine. In: Walton CJ (ed.), Fisheries management and development, Vol. III, Element D: Characterization of the shellfisheries, p. 244-293. Completion report to the State Planning Office, Oct. 1, 1978-Sept. 30, 1979, Maine Department of Marine Resources, Augusta

Stickney, A. P. 1981a. The environmental physiology of the northern shrimp (*Pandalus borealis*) in the Gulf of Maine. In Frady T (ed.), Proc. IntI. pandalid shrimp symp., Kodiak, Alaska, 13-15 Feb. 1979. Univ. Alaska, Sea Grant Rep. 81-3:393

Stickney, A. P. 1981b. Laboratory studies on the development and survival of *Pandalus borealis* eggs in the Gulf of Maine. In Frady T (ed.), Proc. IntI. pandalid shrimp symp., Kodiak, Alaska, 13-15 Feb. 1979. Univ. Alaska, Sea Grant Rep. 81-3:395-40.

Stickney, A.P., and H. C. Perkins. 1977. Environmental physiology of commercial shrimp, *Pandalus borealis*. Project 3-202-R Completion Report, February 1, 1974 to January 31, I977, Dep. Mar. Resour., W. Boothbay Harbor, Maine, 78p.

Stoffer, D. S., and R. H. Shumway. 2017. Time Series Analysis and Its Applications: With R Examples. Germany: Springer International Publishing.

Thorson, J. T., and C. Minto. 2015. Mixed effects: a unifying framework for statistical modelling in fisheries biology. ICES J Mar Sci 72: 1245-1256. doi: 10.1093/icesjms/fsu213

Thorson, J. T., A. Rindorf, J. Gao, D. H. Hanselman, and H. Winker. 2016. Density-dependent changes in effective area occupied for sea-bottom-associated marine fishes. Proc. R. Soc. B.283: 20161853. 20161853 http://doi.org/10.1098/rspb.2016.1853

Torre, M.P., K. R. Tanaka, and Y. Chen. 2018. A spatiotemporal Evaluation of Atlantic Sea Scallop *Placopecten magellanicus* Habitat in the Gulf of Maine Using a Bioclimate Envelope Model. Mar Coast Fish. 10: 224-235. https://doi.org/10.1002/mcf2.10022

Townsend, D. W., A. C. Thomas, L. M. Mayer, M. Thomas, and J. Quinlan. 2006. Oceanography of the Northwest Atlantic Continental Shelf. pp. 119-168. In: Robinson, A.R. and K.H. Brink (eds). The Sea, Volume 14, Harvard University Press.

Townsend, D. W., N. R. Pettigrew, M. A. Thomas, M. G. Neary, and others. 2015. Water masses and nutrient sources to the Gulf of Maine. J Mar Res 73: 93–122.

Weltz, K., A. A. Kock, H. Winker, C. G. Attwood, and M. Sikweyiya. 2013. The influence of environmental variables on the presence of white sharks, Carcharodon carcharias at two popular Cape Town bathing beaches: a generalized additive mixed model. PLoS ONE 8: e68554. doi: 10.1371/journal.pone.0068554

Wieland, K., 2004. Length at sex transition in northern shrimp (*Pandalus borealis*) off West Greenland in relation to changes in temperature and stock size. Fish. Res. 69, 49-56.

Wieland, K., 2005. Changes in recruitment, growth, and stock size of northern shrimp (*Pandalus borealis*) at West Greenland: temperature and density-dependent effects at released predation pressure. ICES J. Mar. Sci. 62, 1454-1462.

Wieland, K., and H. Siegstad. 2012. Environmental factors affecting recruitment of northern shrimp Pandalus borealis in West Greenland waters. Mar Ecol Prog Ser 469: 297-306.

Wit, E., E. D. van Heuvel, and J.-W. Romeijn. 2012. 'All models are wrong..': An introduction to model uncertainty. Stat. Neerl. 66, 217–236.

Woillez, M., J. Rivoirard, and P. Petitgas. 2009. Notes on survey-based spatial indicators for monitoring fish populations. Aquat. Living Resour. 22(2), 155-164. doi:10.1051/alr/2009017

Wood, S. 2017. Generalized Additive Models: An Introduction with R, 2 edition. Chapman and Hall/CRC.

Worm, B., R. A. Myers. 2003. Meta-analysis of cod-shrimp interactions reveals top-down control in oceanic food webs. Ecology 84:162−173.

Zeileis, A., F. Leisch, K. Hornik, and C. Kleiber. 2002. "strucchange: An R Package for Testing for Structural Change in Linear Regression Models." Journal of Statistical Software, 7(2), 1–38. http://www.jstatsoft.org/v07/i02/.

Zimmermann, U., F. L. Carvalho, and F. L. Mantelatto. 2015. The reproductive performance of the red-algae shrimp *Leander paulensis* (ortmann, 1897) (decapoda, palaemonidae) and the effect of post-spawning female weight gain on weight-dependent parameters. Braz. J. Oceanogr. 63, 207–216.

Zuur, A. F., E. N. Ieno, N. J. Walker, A. A. Saveliev, and G. Smith. 2009. Mixed Effects Models and Extensions in Ecology with R. Springer, New York.

Appendix A.Chapter 4 Supplementary Data

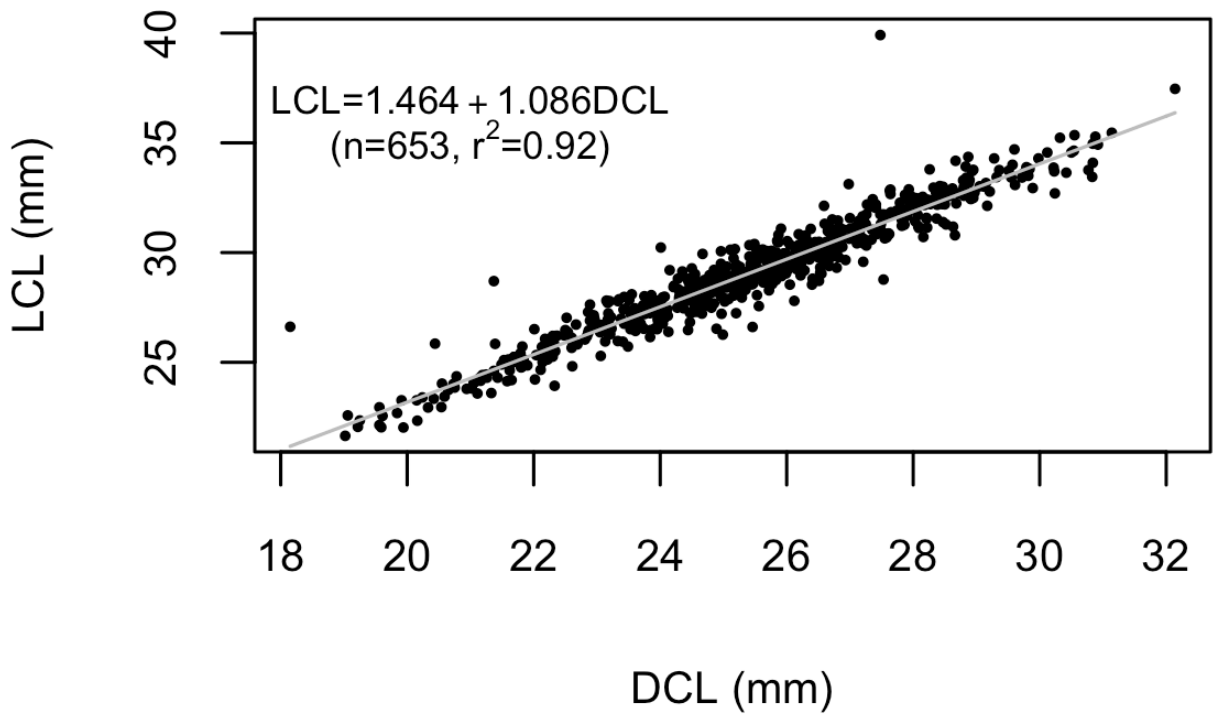


Figure A.1. The relationship between dorsal carapace length (DCL, mm) and lateral carapace length (LCL, mm) for the Gulf of Maine northern shrimp

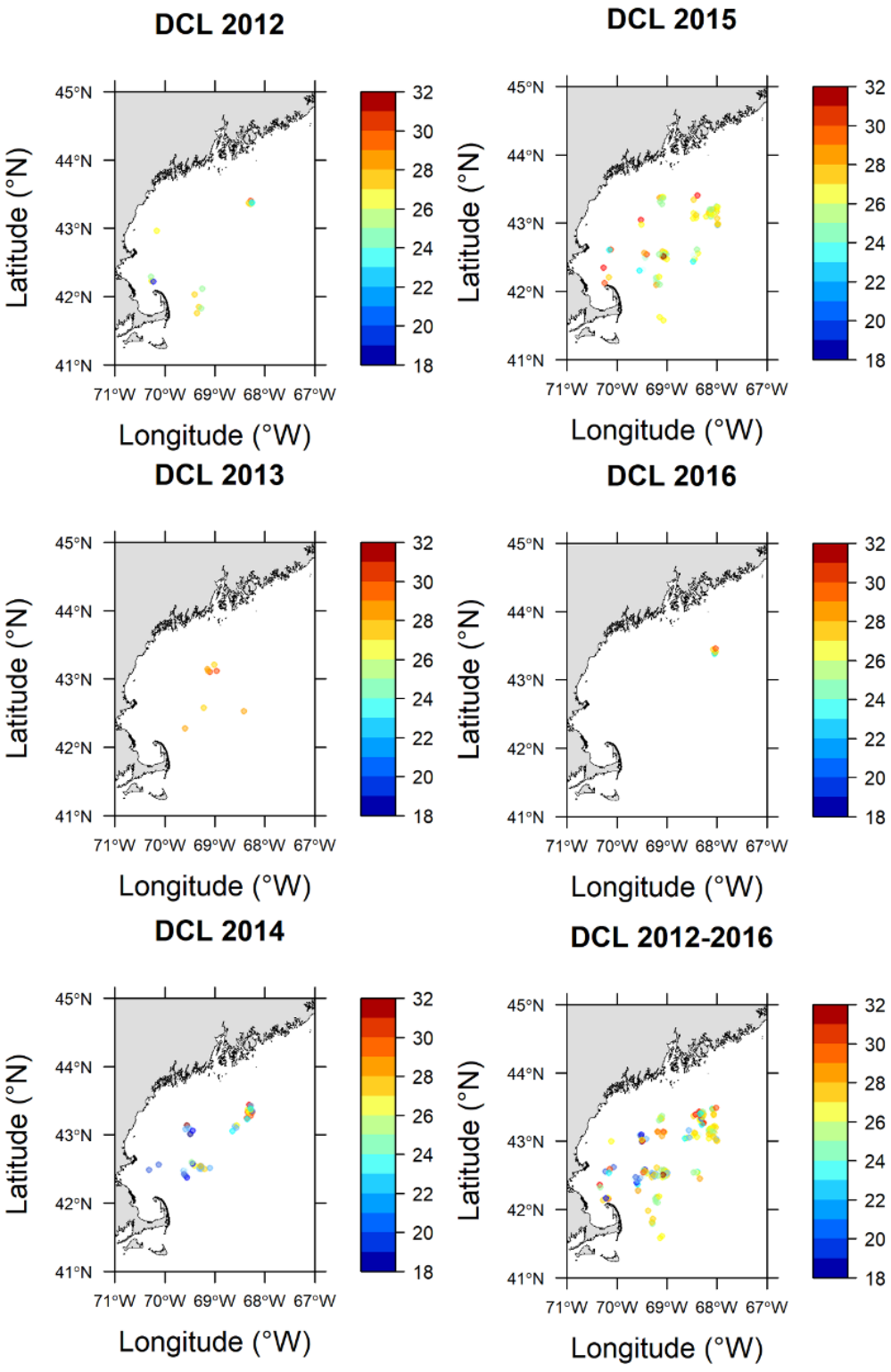


Figure A.2. Spatial distributions of dorsal carapace length (DCL, mm) in 2012-2016 and years pooled

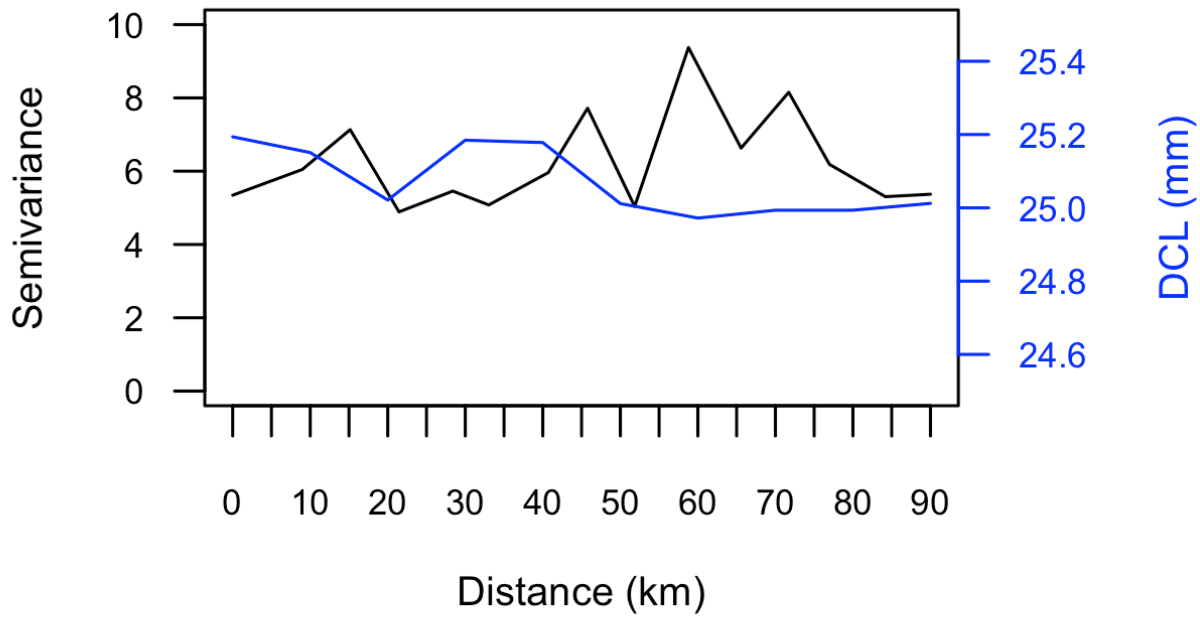


Figure A.3. Spatial variogram of dorsal carapace length (DCL, black line) and average DCL (blue line) over distances

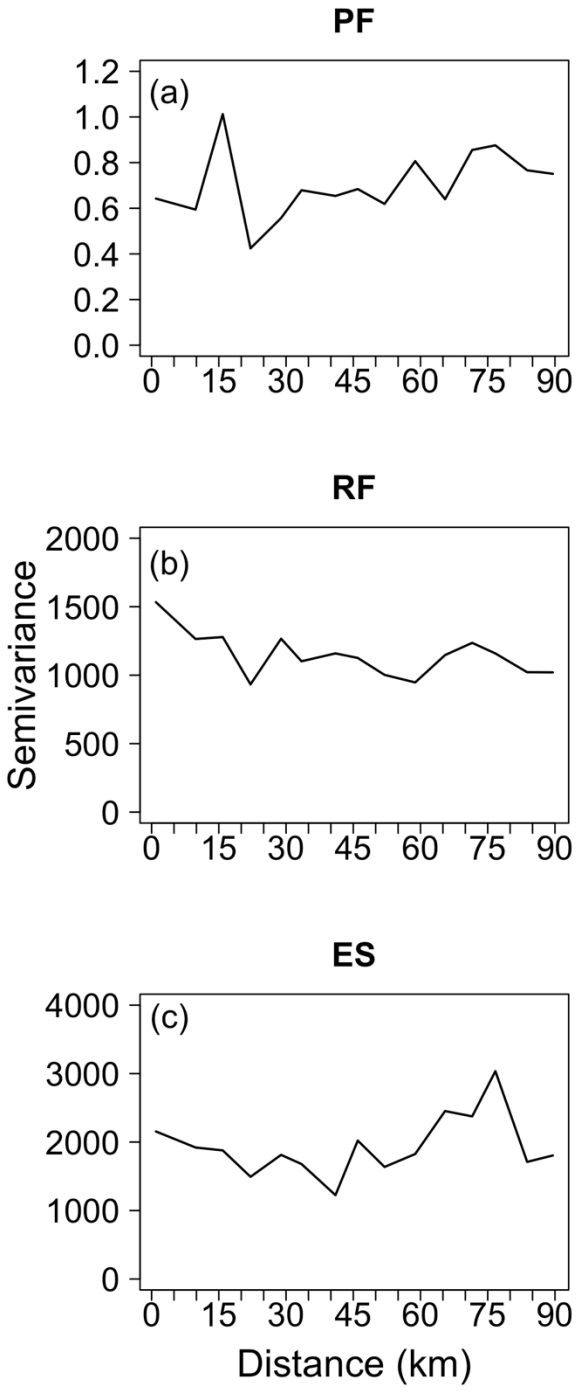


Figure A.4. Spatial variogram computed with Pearson residuals of the best models with lowest AICc for (a) potential fecundity (PF), (b) relative fecundity (RF), and (c) egg size (ES)

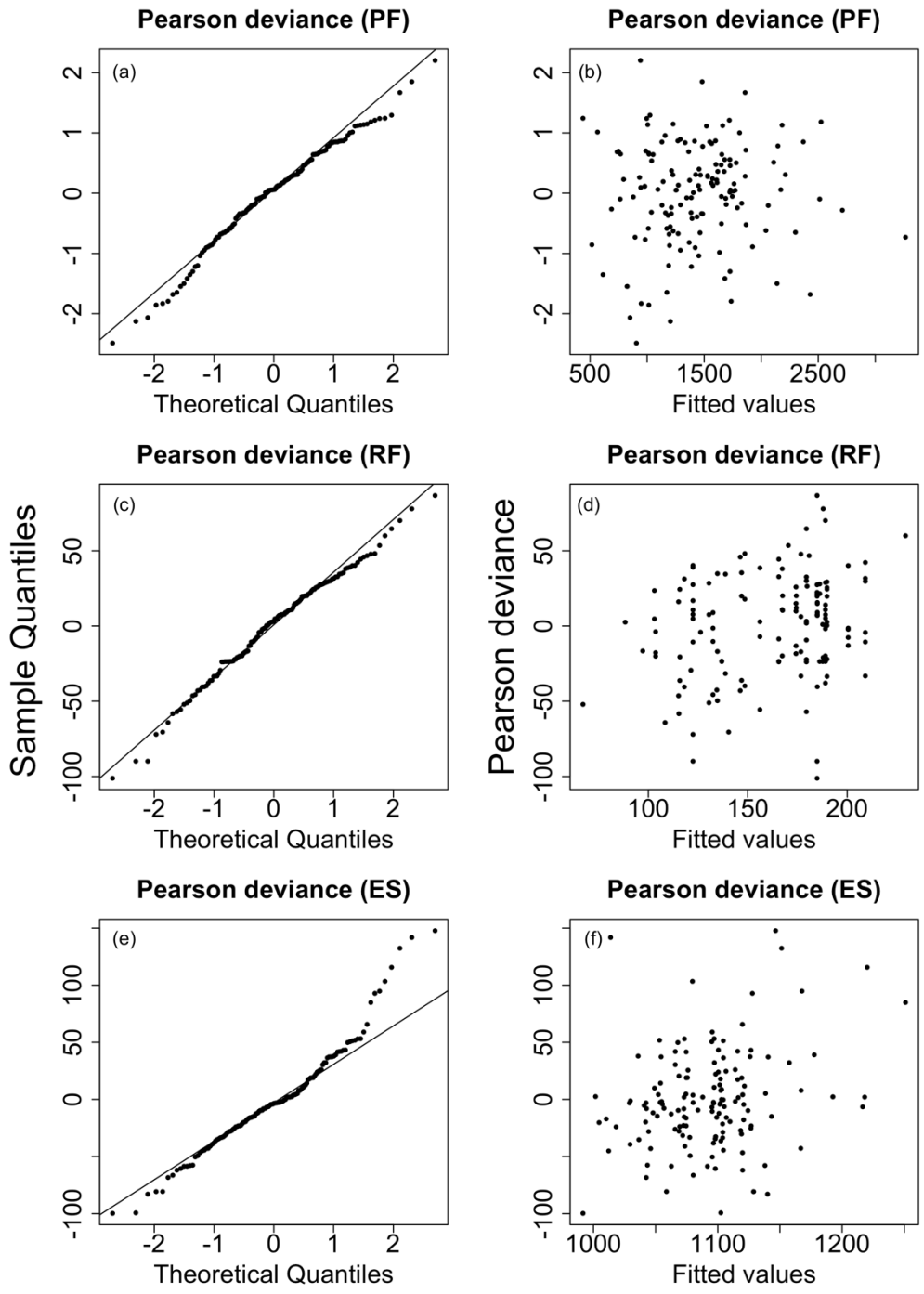


Figure A.5. QQ-plots and plot of Pearson deviance against fitted values of potential fecundity (PF, a-b), relative fecundity (RF, c-d), and egg size (ES, e-f)

Appendix B.Chapter 5 Supplementary Data

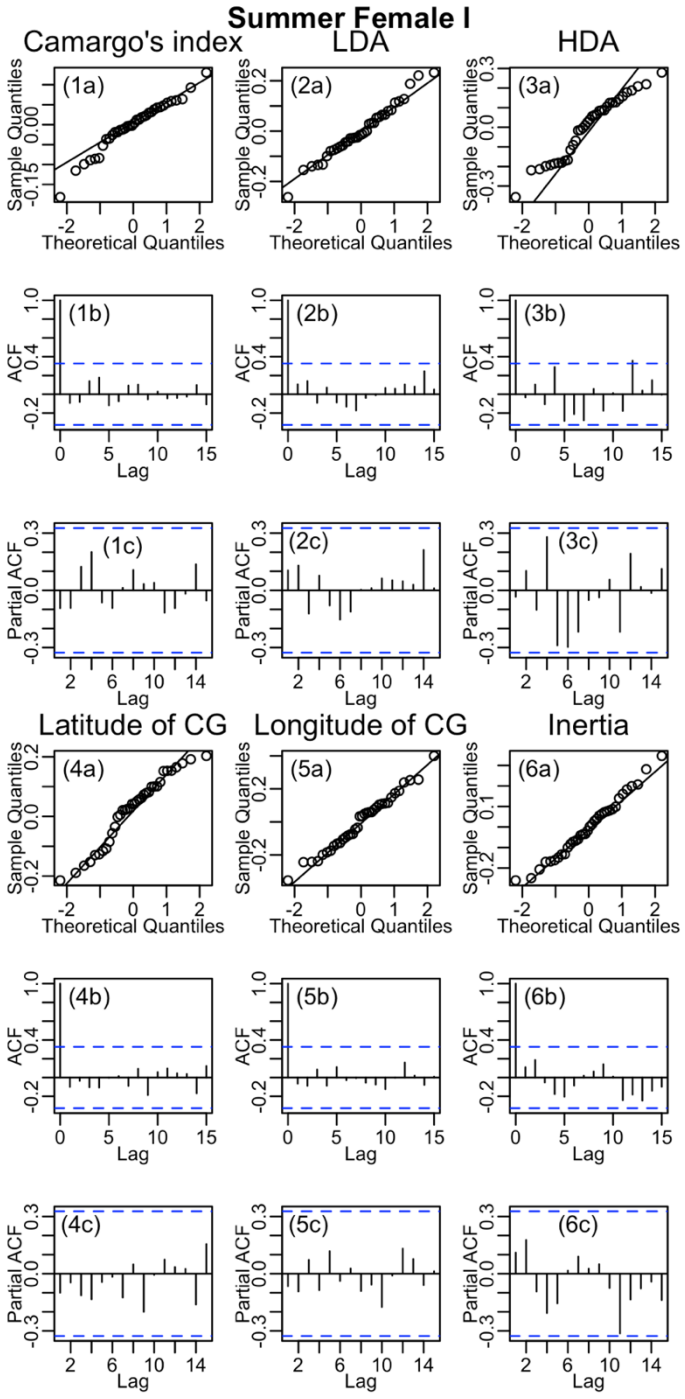


Figure B.1. (a) QQ-plot, (b) temporal autocorrelation plot, and (c) temporal partial autocorrelation plot of residuals of models for summer female I in Table 5.1.

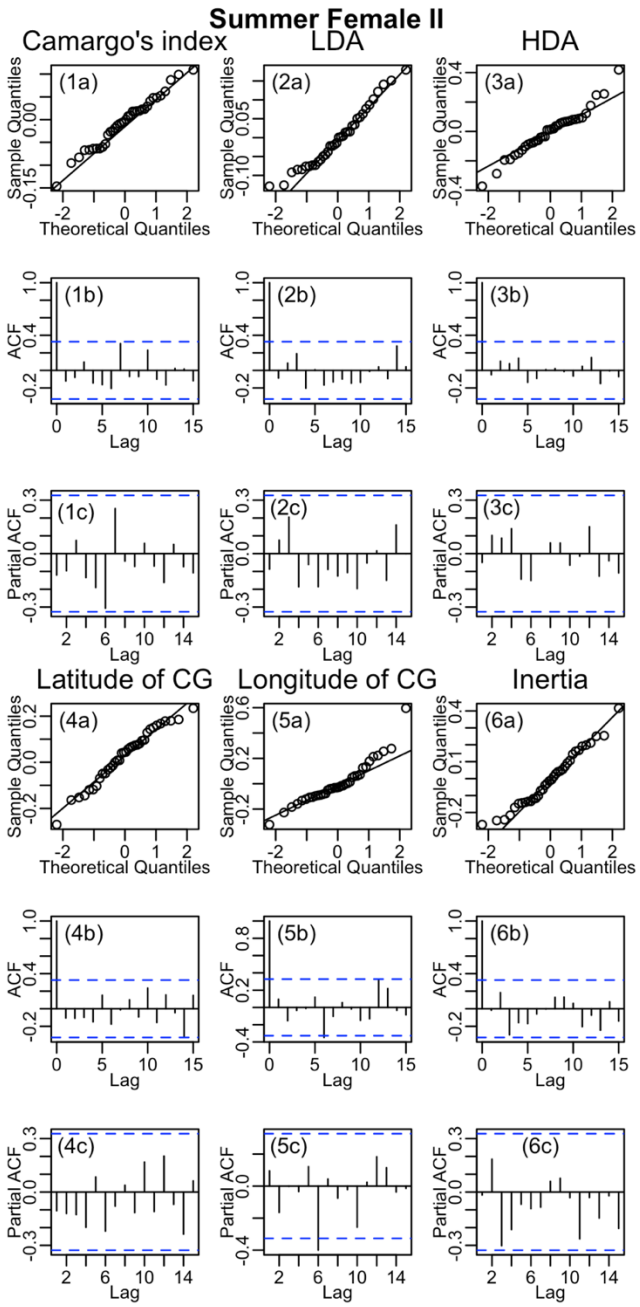


Figure B.2. (a) QQ-plot, (b) temporal autocorrelation plot, and (c) temporal partial autocorrelation plot of residuals of models for summer female II in Table 5.1.

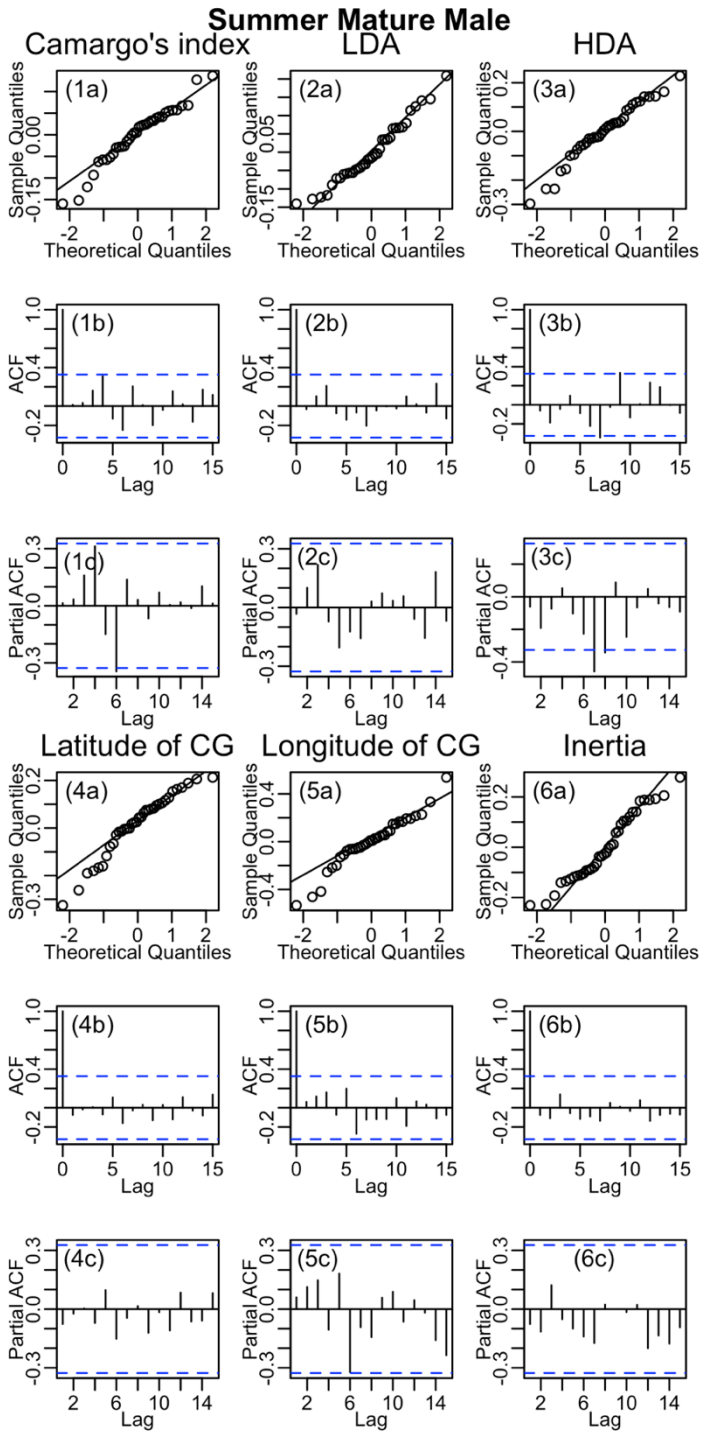


Figure B.3. (a) QQ-plot, (b) temporal autocorrelation plot, and (c) temporal partial autocorrelation plot of residuals of models for summer mature male in Table 5.1.

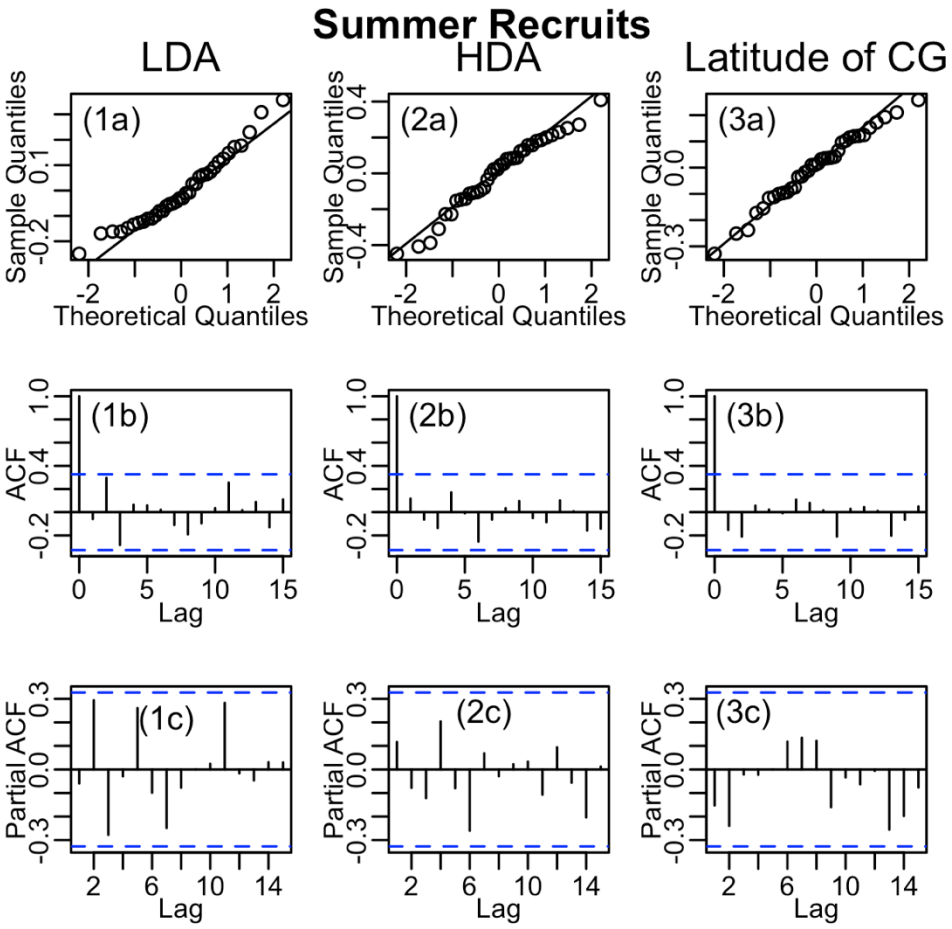


Figure B.4. (a) QQ-plot, (b) temporal autocorrelation plot, and (c) temporal partial autocorrelation plot of residuals of models for summer recruits in Table 5.1.

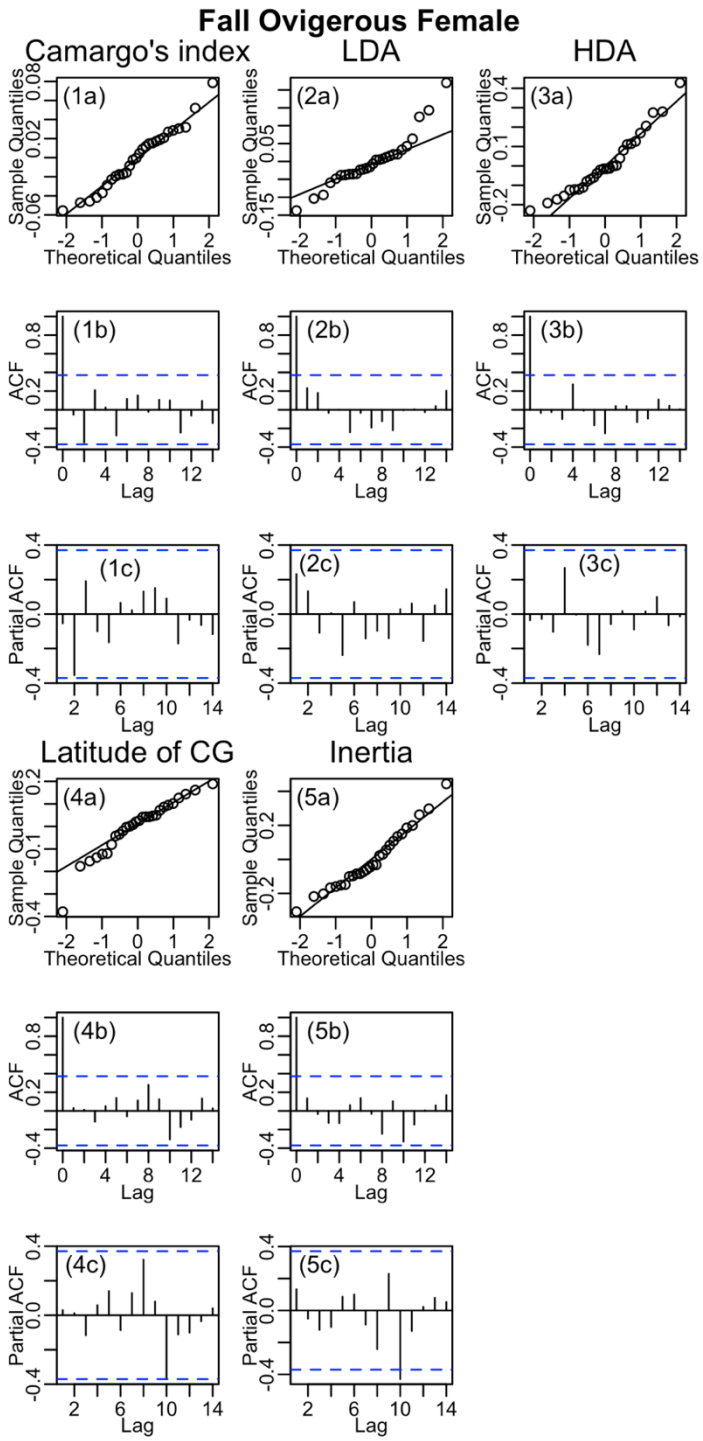


Figure B.5. (a) QQ-plot, (b) temporal autocorrelation plot, and (c) temporal partial autocorrelation plot of residuals of models for fall ovigerous female in Table 5.1.

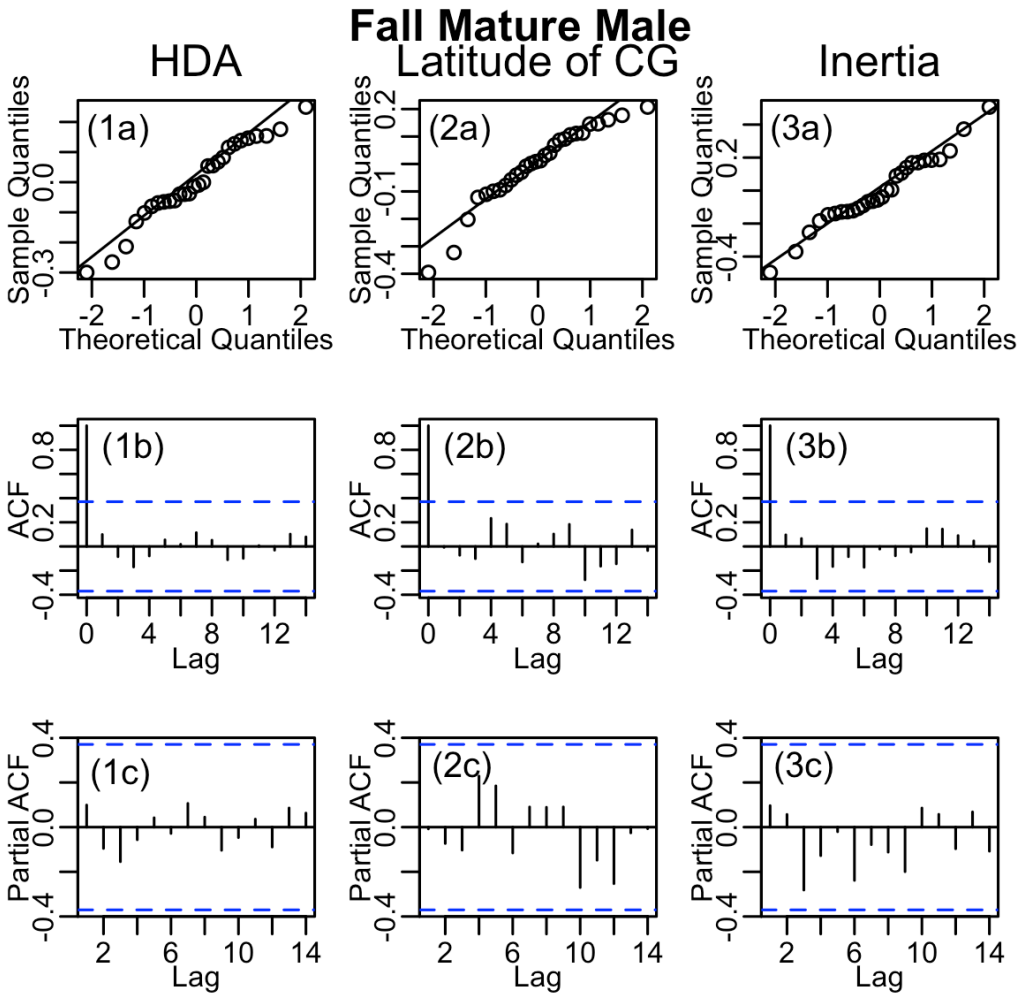


Figure B.6. (a) QQ-plot, (b) temporal autocorrelation plot, and (c) temporal partial autocorrelation plot of residuals of models for fall mature males in Table 5.1.

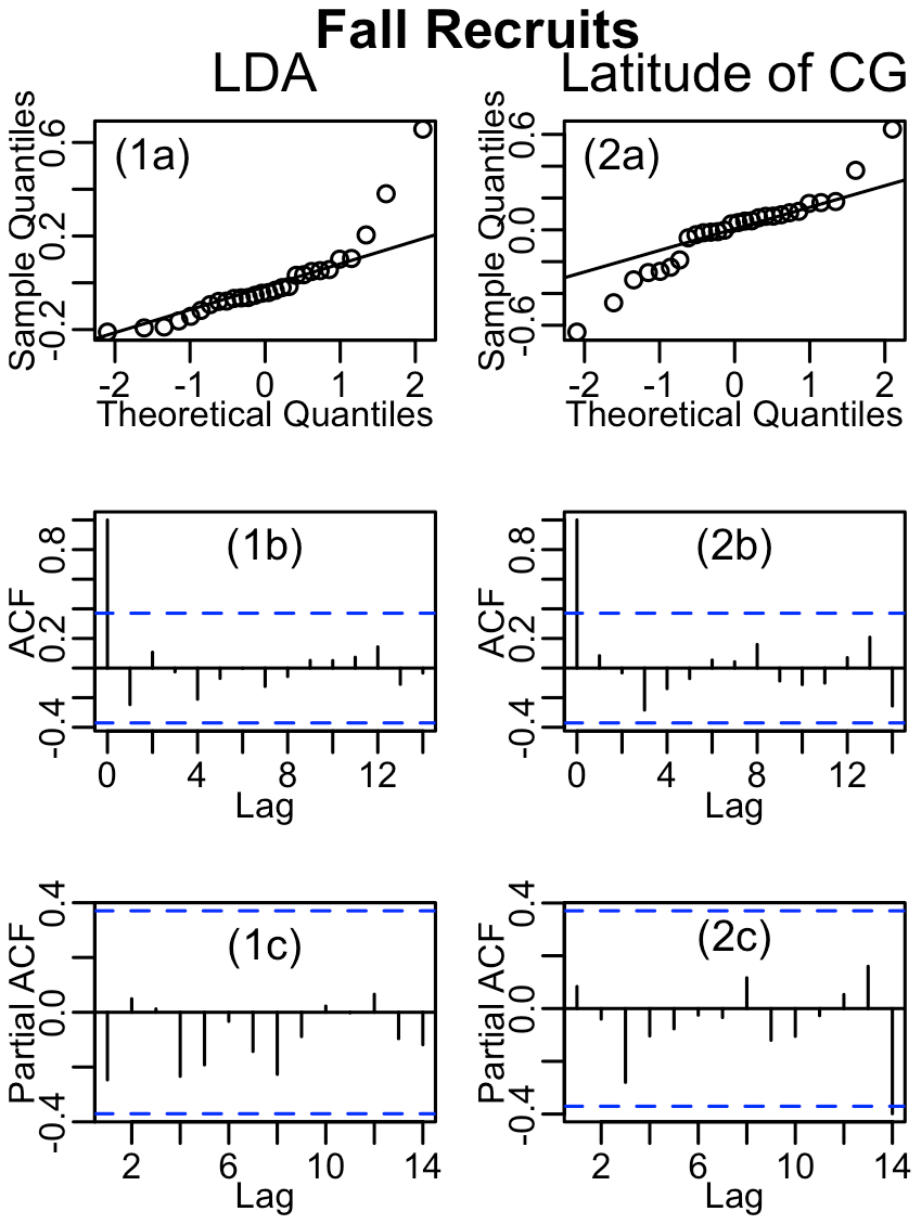


Figure B.7. (a) QQ-plot, (b) temporal autocorrelation plot, and (c) temporal partial autocorrelation plot of residuals of models for fall recruits in Table 5.1.

BIOGRAPHY OF THE AUTHOR

Hsiao-Yun Chang was born in Chiayi, Taiwan. She was raised in Kaohsiung and graduated from Kaohsiung Municipal Hsin-Chuang Senior High School. She then attended National Taiwan Ocean University and graduated in 2008 with a Bachelor's degree in Environmental Biology and Fisheries Science. She enrolled in graduate school the Institute of Oceanography at National Taiwan University (IONTU) after receiving her bachelor's degree, and graduated in 2011 with a Master degree in Science. She then remained at IONTU and served as a research assistant at Dr. Chi-Lu Sun's lab. She entered the School of Marine Sciences at the University of Maine in 2017. She is a candidate for the PhD degree in Ecology and Environmental Science from the University of Maine in August 2021.

Chapter 2 has been published as:

Chang, H.-Y., and Y. Chen. 2020. Evaluating sampling strategies for collecting size-based fish fecundity data: an example of Gulf of Maine northern shrimp *Pandalus borealis*. J. Northw. Atl. Fish. Sci. 51, 33–43. doi:10.2960/J.v51.m730

Chapter 3 has been published as:

Chang, H.-Y., R. Klose, and Y. Chen. 2020. Possible climate-induced environmental impacts on parasite-infection rates of northern shrimp *Pandalus borealis* eggs in the Gulf of Maine. Dis. Aquat. Organ. 140, 109-118. doi: 10.3354/dao03495. PMID: 32701067.

Chapter 4 has been submitted to Global Ecology and Conservation.