Design and Model-based Approaches for Estimating Abundance of American Horseshoe Crab

Chad C. Wong

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science
In
Fisheries and Wildlife Sciences

Yan Jiao, Chair
Eric Hallerman, Co-Chair
Francesco Ferretti

December 12th, 2023
Blacksburg, Virginia

Keywords: Horseshoe crab, channeled whelk, knobbed whelk, summer flounder, *Limulus polyphemus, Busysotypus canaliculatus, Busycon carica, Paralichthys dentatus,* abundance, hurdle model, CPUE standardization, spatio-temporal kriging

Design and model-based approaches for estimating abundance of American Horseshoe Crab

Chad C. Wong

ABSTRACT

The American horseshoe crab (HSC), *Limulus polyphemus*, is one of four species of horseshoe crabs found throughout the world, and the only one found in North America. It is an economically and ecologically important species throughout its native range from Maine to the Yucatan Peninsula. Harvested for fertilizer and livestock feed in the 19th century, the species is now harvested as bait for whelk and eel fisheries, and for their blood by the biomedical industry. The Atlantic States Marine Fisheries Commission (ASMFC) started to formally manage HSC in 1998 with its Interstate Fisheries Management Plan (IFMP). Unique emphasis and harvest limits have been placed on the Delaware Bay stock, as it is commercially exploited and a critical food source for the threatened red knot, *Calidris canutus rufa.* Previously, estimates of relative and total abundance of HSC in the Delaware Bay area were based on a design-based approach using a stratified random sampling design. In Chapter 1 of this work, I developed hurdle models for each of the six HSC demographic groups to standardize catch-per-unit-effort (CPUE) and estimate relative abundance using a model-based approach. It was determined that while the two approaches resulted in mostly convergent estimates of relative abundance, external factors such as month, time-of-day, and average depth have major effects on the observed CPUE of all demographic groups. Chapter 2 involved the development of hurdle models for the three species of bycatch frequently caught in our trawls, channeled whelk (*Busysotypus canaliculatus*), knobbed whelk (*Busycon carica*), and summer flounder (*Paralichthys dentatus*). It was found that channeled whelk relative abundance has been at a historical low since 2016, while summer flounder has been at a consistent high. Recent estimates of knobbed whelk relative abundance

have been less variable than previously seen, with estimates since 2016 being similar to those seen before 2012. These results provide the first estimates for whelk population trends in the mid-Atlantic region and add to the growing knowledge of summer flounder relative abundance in the area. In Chapter 3, I applied the hurdle models developed in Chapter 1 to estimate the total abundance of HSC in the Delaware Bay area. For this work, I developed two spatio-temporal variograms to estimate bottom temperature and bottom salinity at unmeasured cells per month in the time series. The results showed that night estimates of total abundance were consistently higher than daytime estimates, and estimates from September or November resulted in the highest estimated catch for all demographic groups. The results suggest that when comparing September model-based estimates at night to those of the design-based approach, nearly a third of all previous design-based estimates significantly underestimated the total abundance of HSC in the Delaware Bay area. This result suggests that the ASMFC can recommend increased harvest limits for mature individuals if that action aligns with the goals of their adaptive resource management (ARM) framework.

Design and model-based approaches for estimating abundance of American Horseshoe Crab

Chad C. Wong

GENERAL AUDIENCE ABSTRACT

The American horseshoe crab (HSC), *Limulus polyphemus*, is one of four species of horseshoe crabs found throughout the world, and the only one found in North America. It is an economically and ecologically important species throughout its native range from Maine to the Yucatan Peninsula. Harvested by the millions in the 19th century, the species is still harvested as bait for whelk and eel fisheries, and for their blood by the biomedical industry, on the order of hundreds of thousands. Formal management of HSC by the Atlantic States Marine Fisheries Commission (ASMFC) began in 1998, and allowable catch and landings have decreased since 1999. A strong focus has been placed on the Delaware Bay population, as it is the center of HSC abundance and provides a critical food source for the threatened red knot, *Calidris canutus rufa*, while also being commercially exploited. To effectively manage the species, it is important to accurately estimate relative and total abundance so that proper harvest limits can be set. Previously, estimates of relative and total abundance of HSC in the Delaware Bay area were based on a design-based approach using a stratified random sampling design. In Chapter 1 of this work, I developed hurdle models (a generalized linear model that models the probability of observations and the observed positive counts using two separate regression models that are then combined) for each of the six HSC demographic groups to remove the effect of external factors (year, latitude, longitude, depth [inshore/offshore], topography, average trawl depth, time-of-day, month, bottom temperature, bottom salinity, and distance from shore) on our observed catch-per-unit-effort (CPUE) and estimate relative abundance using a model-based approach. It was determined that while the two approaches resulted in mostly convergent estimates of relative

abundance, factors like month, time-of-day, and average depth had major effects on the observed CPUE of all demographic groups. Chapter 2 involved developing similar hurdle models for three species of bycatch frequently caught in our trawls, i.e., channeled whelk (*Busysotypus canaliculatus*), knobbed whelk (*Busycon carica*), and summer flounder (*Paralichthys dentatus*). It was found that channeled whelk relative abundance has been at a historical low since 2016, while summer flounder has been at a consistent high. Recent estimates of knobbed whelk relative abundance have been less variable than previously seen, with estimates since 2016 being similar to those seen before 2012. These results provide the first estimates for whelk population trends in the mid-Atlantic region and add to the growing knowledge of summer flounder relative abundance in the area. In Chapter 3, I applied the hurdle models developed in Chapter 1 to estimate the total abundance of HSC in the Delaware Bay area. To do this, the bottom temperature and salinity had to be estimated for each geographic cell. This was accomplished by developing two spatio-temporal variograms which allowed me to estimate either variable at an unmeasured point and time based on its spatial and temporal distance from a measured value in a process known as spatio-temporal kriging. The results showed that night estimates of total abundance were consistently higher than daytime estimates and that estimates from September or November resulted in the highest estimated total abundance for all demographic groups. The results suggest that when using September model-based estimates at night to compare against the design-based approach, nearly a third of all previous design-based estimates significantly underestimated the total abundance of HSC in the Delaware Bay area. This outcome could justify ASMFC increasing recommended harvest limits for mature individuals if that action aligns with the goals of their adaptive resource management (ARM) framework.

ACKNOWLEDGMENTS

I would like to thank Drs. Eric Hallerman and Yan Jiao for entrusting me with the responsibility for this project and providing me with guidance throughout the entire process. This work would also not have been possible without our field crew, Jeff Eutsler and Phil Schantz, who worked tirelessly throughout autumn to gather the data necessary for this study. I would also like to thank the Atlantic States Marine Fisheries Commission and especially Kristen Anstead for her willingness to help on the project. Special thanks to Dr. Vanessa Hull, Dr. Miguel Acevedo, and Travis Tuten, whose initial guidance and encouragement played an enormous role in me landing this position in the first place.

It takes a village to complete a graduate degree, and I would not have been able to do it without the unwavering love and support of my family and friends. While I wish I could list everyone, Nico Hazlett, Katie Syzmanski, and Leila Hazlett deserve special recognition for being the best-chosen family that I could have asked for during this process and for putting up with me these past two years.

Table of Contents

Chapter 1

Comparison of design and model-based approaches for estimating relative abundance of American Horseshoe Crab in the Delaware Bay area

Chad C. Wong[1], Yan Jiao[1], and Eric Hallerman[1]

[1] Department of Fish and Wildlife Conservation, Virginia Tech, 310 West Campus Drive, Blacksburg, VA 24061, USA

Abstract

The American horseshoe crab (HSC) *Limulus polyphemus* is harvested as bait for whelk and eel fisheries and for their blood by the biomedical industry. The Atlantic States Marine Fisheries Commission (ASMFC) started managing HSC in 1998 with unique emphasis and harvest limits placed upon the Delaware Bay area stock, as that subpopulation is the center of abundance for horseshoe crabs along the Atlantic coast and provides a critical food resource for the threatened red knot *Calidris canutus rufa*. Our group has performed trawl surveys of the lower Delaware Bay and mid-Atlantic coast in these areas since 2003 to estimate relative abundance. This study is the resulting analysis of the cumulative data from these surveys. While previous assessments have used only design-based estimators, this study provides not only design-based estimators, but also model-based estimators of relative abundance. Model analysis revealed that longitude, average depth, distance from shore, time-of-day, and month affected the catch-per-unit-effort of all demographic groups. Design-based and model-based estimates of relative abundance were relatively convergent except for immature males in 2005 and 2011, immature females in 2021, and mature females in 2020. Model-based estimates were on average more precise than the design-based approach in terms of the two approaches' uncertainty intervals of any given year. Relative abundance varied greatly throughout the time series, although previously reported downward trends of newly-mature females since 2016 were confirmed. We suggest that model-based estimates should be used for informing management going forward. Future work should consider redesigning the stratification of the survey design based on average depth and distance from shore and try to sample more evenly between months and times-of-day to validate the results of this study.

Introduction

With fossil lineages extending back 445 million years (Rudkin and Young 2009), the horseshoe crab lineage has survived five mass extinction events. However the four extant species of horseshoe crab are currently under threat from ongoing anthropogenic impacts to the environment and to the species themselves. American horseshoe crabs (*Limulus polyphemus*) are distributed from Maine to the Yucatán Peninsula, with winter temperatures appearing to be the factor limiting their northern distribution (Sekiguchi and Shuster 2009; Smith et al. 2017). Originally harvested and ground up for fertilizer and livestock feed, American horseshoe crab populations have dramatically decreased since individuals were harvested in the millions during the late 19th and early 20th centuries (Shuster and Botton 1985; Millard et al. 2015). Following the widespread adoption of chemical fertilizers, HSC populations stabilized, but declined again during the 1990s due to rising biomedical and bait harvests (Millard et al. 2015). With the implementation of the Interstate Fisheries Management Plan for Horseshoe Crab in 1998, harvest limits were set and reporting became mandatory for all states with an active HSC fishery, and as a result, HSC landings have drastically decreased since 1999 (Walls et al. 2002; ASMFC 2019). Total harvest has not exceeded one million individuals since 2003, with an average annual landing of 752,886 between 2004 – 2017 (ASMFC 2019). Harvest limits are decided through the application of the ASMFC's adaptive resource management model (ARM) which yields a recommendation for a state-by-state annual quota, with an average limit of 100,000 – 500,000 male HSC harvested annually in the Delaware Bay area (ASMFC 2019), a coastline which includes New Jersey, Delaware, Maryland, and Virginia. Overall, stock status for HSC is assessed by dividing the Atlantic coast into four regions (Northeast, New York, Delaware Bay, Southeast), with each stock assigned a population status based on the most recent relative abundance estimates compared to the 1998 reference point (ASMFC 2019). According to the most recent horseshoe crab stock assessment (ASMFC 2019), the Delaware Bay regional stock status is assessed as "neutral", suggesting that there has not been a significant increase or decrease in overall population size since the program's implementation.

Today, two major stakeholders harvest horseshoe crabs, commercial bait fishers, and the biomedical industry. Commercial fishers use horseshoe crabs as bait for American eel (*Anguilla rostrada*) and whelk/conch species (*Busycon* spp. and *Sinistrofulgur* sp.) (Walls et al. 2002; ASMFC 2019). Although whelk fishers do not prefer either sex, American eel fishers prefer female HSC (Kreamer and Michels 2009), so much so, that in the 1990s, almost entirely female HSC were used for bait in Delaware Bay eel pot fisheries (Kreamer and Michels 2009). The biomedical industry uses the lymphocytes isolated from HSC blood to produce *Limulus* amebocyte lysate (LAL), which contains an agglutinating agent that can be used to detect gram-negative bacterial endotoxins in everything from vaccines and injectable pharmaceuticals to implantable medical devices and contaminated food products (Upmann and Bonaparte 1999; Walls et al. 2002). Due to the female HSC's larger size and hence larger blood volume, the biomedical industry also has a female-skewed harvest (Rutecki et al. 2004; James-Pirri et al. 2012; ASMFC 2019). Natural mortality occurs from predation by species such as loggerhead sea turtles (*Caretta caretta*), leopard sharks (*Triakis semifasciata*), and some ray and skate species (Walls et al. 2002).

Although the bait and biomedical industries are the major extractive stakeholders, a major consideration of ASMFC's ARM model (McGowan et al. 2011b) is the demographic recovery of the threatened red knot (*Calidris canutus rufa*). The Delaware Bay is a stopover area for many species of shorebirds making their spring migration from overwintering areas in South

America to their breeding grounds in the Arctic, including red knots (ASMFC 2019). This shorebird migration corresponds with the arrival of HSC to their intertidal breeding grounds for spawning from their overwintering locations offshore. Breeding coincides with warming water temperatures (Watson et al. 2009) and usually lasts between April through July in the mid- and northern coasts of the Atlantic Ocean, with peak spawning in the Delaware Bay area occurring during new and full moons between May and June (Shuster and Botton 1985; Michels et al. 2008). During this time, there is a skewed operational sex ratio (OSR) of males to females, due to both female-biased natural mortality and harvest and males returning to the beach to mate more frequently than females (Brockmann and Penn 1992; Smith et al. 2002; Hallerman et al. 2022). In the Delaware Bay, this OSR averaged 3.1 - 4.7 males to one female between 1999 and 2008 (Michels et al. 2008; Wong et al. 2023). The synchronization of HSC breeding and shorebird migration is due in part to the availability of high-energy nutrients from abundant HSC eggs. Red knots on this flyway use HSC eggs as their primary food source to gain weight in advance of their flight to the Arctic (Botton et al. 2003; Mizrahi and Peters 2009). While the number of eggs needed per bird to fly to their final migratory destination is debated, the link between HSC egg availability and red knot survival is well-accepted (Karpanty et al. 2006; McGowan et al. 2011a; ASMFC 2022). A recent integrated population model (IPM) of Delaware Bay area populations showed a significant correlation between HSC abundance and red knot survival, but not between HSC abundance and red knot recruitment (ASMFC 2022). The most recent revision of the ARM framework implements an Approximate Dynamic Programing (ADP) approach to simultaneously maximize the harvest of HSC and the abundance of migrating red knots (ASMFC 2022).

With these factors in mind, the ASMFC is actively working to rebuild the HSC population coast-wide, with special consideration given to the Delaware Bay area. The Delaware Bay is considered the center of abundance for HSC (Shuster 2015) due to its high abundance relative to other populations along the Atlantic coast. The most recent population estimates of the coastal Delaware Bay area included 11.4 million mature females and 19.9 million mature males (Wong et al. 2023). Reaching sexual maturity between eight and ten years and with a life span of fourteen to twenty years, long-term monitoring is essential for HSC (Shuster 1955; Sekiguchi et al. 1988). It could take decades for the effects of management to become detectable while harvest continues, especially with the uncertainty regarding natural and biomedical mortality rates of HSC. The 1998 HSC fishery management plan (FMP) established a biomedical mortality threshold of 57,000 crabs which has been exceeded every year since 2007 with the exception of 2016, although no management action has been taken occurred in response to that (ASMFC 2019). Due to confidentiality on the level of biomedical harvest, mortality estimates from bleeding can be only roughly estimated and are taken as 15% for management purposes (95% C.I. 4 – 30%) (ASMFC 2019). Established in 2008, this generalized mortality rate was adopted based on a peer review performed by the ASMFC of available literature on biomedical bleeding mortality rates from multiple states and studies, and remained in use following an updated meta-analysis in 2019 (ASMFC 2008, 2019). This uncertainty around biomedical fishery mortality means that the best way to monitor relative changes in population abundance is through sampling and modeling studies such as those reported here. As with any species' management, it is important to accurately estimate population size, range, relative abundance, migration patterns, and geographic distribution of genetic variation to support effective conservation of the species (King et al. 2015, as cited in Hallerman et al. 2022). This study seeks to advance the understanding of HSC relative abundance through continuation of previous design-based estimates and the development of novel model-based estimates for HSC in the Delaware Bay area. In this study catch-per-tow, or catch-per-unit-effort (CPUE), is used to represent the original survey observation data of the number of horseshoe crabs caught per standard 15-minute tow. Relative abundance is the standardized CPUE, which is the estimated year effect plus the effect of the other explanatory variables using averaged values (see Maunder and Punt 2004). The year effect is the estimated coefficient of the year categorical variable in the hurdle model.

Previous design-based estimates of horseshoe crab relative abundance were based upon a stratified random sampling design developed from a pilot trawl-survey study performed in 2001 (Hata and Berkson 2004). The advantages of design-based sampling schemes and associated estimators of abundance include cost-effectiveness, replicability, and the ability to weight observations based on the proportion of demographic classes represented in the population (Yu et al. 2012; Lohr 2019 chap. 3). This last aspect is the basis for stratified random sampling, in which populations are assumed to be randomly distributed and homogenous within strata, and not influenced by external factors. While design-based methods allow us to weight CPUE and estimate relative and total abundance for a single sampling period, they assume that catch rate is directly proportional to abundance (Maunder and Punt 2004). This is one of the main assumptions of design-based estimators and also its greatest weakness, as an observed decrease in CPUE would be interpreted as a direct decrease in abundance. Such a decrease is not always the case. Underlying external factors or time-series effects could lead to variations in CPUE that were not caused by changes in abundance, resulting in biased design-based estimates and an overall decrease in estimate precision (Maunder and Starr 2003).

Model-based estimators – in the form of generalized linear models, generalized additive models, and other regression-based approaches – seek to remediate this problem by allowing the inclusion of other explanatory variables in addition to effort and effects of strata to explain CPUE. This approach includes the possibility of using spatial, temporal, and environmental factors across the entire time-series, as model-based estimators use all data across survey years to best inform variable selection and model fitting. Design-based estimators assume that each year is independent, estimating abundance on that year's data alone. While a model-based approach complicates the estimators and relies heavily on the statistical distributions and predictor variables considered, it should more accurately estimate relative abundance for the time series if modeled properly. This approach, also known as CPUE standardization, allows model-based estimators to break the linearity assumption between catch and abundance to assess which external factors, in addition to true changes in abundance, have led to observed changes in CPUE over the time-series (Olsen et al. 1999; Maunder and Punt 2004). This type of approach also allows the use of two-component mixture models, such as hurdle or zero-inflated models, known for their ability to handle data with a large number of zero observations (Aubry and Francesiaz 2022). For these types of models, zero counts and positive counts are either modeled separately, building two models that are later combined (hurdle models) or modeled simultaneously, solving for a joint likelihood function (zero-inflated models). The use of these models could further increase the accuracy of relative abundance estimates when compared to traditional single-stage regression models (Potts and Elith 2006), or a design-based approach. After a model is built and validated, inferential statistics on the population can be estimated regarding the effects of these external parameters on horseshoe crab relative abundance. In contrast, design-based approaches can only make estimations on the proportion of the population that was sampled (those with a non-zero probability of being caught), of which investigators can only determine descriptive statistics such as stratified mean, stratified variance, and relative changes in abundance based on the sampling design (Sterba 2009). If the model built is supported by cross-validation testing, the inferential statistics estimated can also be used to predict CPUE if given the necessary explanatory variables. This approach can then be used to estimate total population abundance in an area and predict future changes in relative abundance.

A model-based approach will also help determine the validity of the initial sampling design and help redesign the survey if necessary. Although the initial stratified random design turned out to be significantly better than simple random sampling, Hata and Berkson (2004) acknowledged that it is not certain that the respective strata were optimally defined. With a model-based approach, sample design can be re-evaluated to optimize effort and enhance estimation accuracy. This can be especially important when investigating the effect of seasonality and locality on CPUE. For example, if certain months lead to significantly different estimates of CPUE than others, previous design-based estimates could be unintentionally under- or over-estimating relative abundance. The objectives of this study were to develop hurdle models for each of the six horseshoe demographic groups to standardize CPUE, determine the effect of external factors on CPUE, estimate relative abundance using a model-based approach to remove the effect of these external factors on our estimates, and compare these estimates to previous design-based estimates. The practice of using both design-based and model-based approaches not only increases the validity of the estimators, but also allows investigators to compare relative abundance estimates and the uncertainties of each to assess whether one approach is better suited to the data than the other.

Methods

Sampling Sites and Sampling Method

The survey area (Figure 1) included two sampling areas, the Lower Delaware Bay (LDB) and the Delaware Bay Area (DBA). The LDB extended from the mouth of the bay to a line created between Egg Island Point, NJ and Kitts Hammock, DE. The DBA sampling area consists of the Atlantic coast extending from 0-12 nautical miles out from shore and from the northern limit of Atlantic City, NJ (39˚ 20' N) south to Wachapreague, VA (37˚ 40' N). The area is stratified based on bottom topography (trough, non-trough) and distance from shore (0-3nm, inshore; 3-12nm, offshore). This stratification scheme was determined from the results of a 2001 pilot survey of the Delaware Bay area (Hata and Berkson 2004) and has been used since. Troughs are considered areas that are at least 2.4 m deep, no more than 1.8 km wide, and more than 1.8 km long, determined from NOAA nautical charts. These charts were georeferenced into ArcGIS Pro (ERSI 2023) to create a digital map used to design the sampling scheme. The sample area was divided into cells of 1' of latitude by 1' of longitude using a World Geodetic System (WGS) 1984 coordinate system. The LDB consists of 196 cells (~528 km^2), all being considered inshore, and the DBA consists of 1890 cells (~5127 km^2). Sampling has occurred annually since 2002 in the DBA and 2010 in the LDB, with a gap in sampling between 2013 through 2015. The number of trawls performed in each stratum was based on both the proportional area that each stratum represents of the entire sampling area and the historical variation of CPUE observed among strata from previous years' sampling. For this study, only the 697 trawls within the DBA (Table 1) were analyzed, as previously only the Delaware Bay area estimates have been used to inform management decisions by the ASMFC. Trawls were performed using a two-seam flounder trawl with an 18.3-m headrope and 24.4-m footrope, rigged with a Texas Sweep of 13-

mm link chain and a tickler chain. The net body consists of 15.2-cm (6-in) stretched mesh, and the bag consisted of 14.3-cm (5 5/8-in) stretched mesh. Tows are aimed to have 15 minutes bottom time, but occasionally were shorter to avoid fishing gear (e.g., gill nets, crab and whelk pots) or vessel traffic.

Data collected/Environmental Factors

This study considered the HSC population as divided into six demographic groups, based on sex (male or female) and maturity (immature, newly-mature, or mature). Mature females are easiest to identify due to their larger size, likely due to an extra molt in development (Smith et al. 2009a), mating scars from male amplexus (Hata and Hallerman 2022), lack of specialized grasping appendages, darker color, and buildup of fouling organisms on the carapace (Botton et al. 2003). Newly mature and mature males are distinguished by the presence or absence, respectively, of the "thumb" on the boxing glove-like grasping appendage. Age is not used as a demographic characteristic because there is no current procedure to determine age. Therefore, the current best estimates for age are based on overall size/carapace width and finding newly mature individuals. Unfortunately, size-of-maturity varies across locations (Riska 1981; Faurby et al. 2011), and while the age of maturity is estimated as being between 9 – 11 years (Sekiguchi et al. 1988; Shuster and Sekiguchi 2003), there is no definitive way to tell age-at-maturity. Along with this demographic information, data on environmental variables also was collected during the time of sampling. These variables included date, beginning and ending trawl depth, bottom temperature, bottom salinity, air and surface temperatures, and general weather conditions. A summary of the percentage of trawls performed during the day and night is shown in Table 2, and the different variable cross-correlations can be seen in a matrix plot in Appendix A (Figure S1). Due to a possible salinity measurement error in 2020 and 2022, and salinity measurements in 2020-2022 being originally recorded in specific gravity, two salinity scenarios were tested in this analysis, resulting in an unadjusted- and adjusted-salinity scenario. The former uses the original measured bottom salinity data, while the latter uses the original data plus an added increase to salinity measurements between 2020-2022. This increase was calculated based on previously recorded salinities of historical trawls that were spatially and temporally (using ordinal date) close to those performed in 2020-2022. Procedures for the two salinity-scenarios tested and dealing with missing data are presented in Appendix A.

Design-based models used

Two distributions were considered to estimate the stratified mean CPUE (relative abundance) and later population mean (total abundance) using a design-based approach: a normal-distribution and a delta-lognormal distribution. The normal distribution estimates for mean and associated variance can be found below (Formula 3, Appendix B) (Lohr 2019 chap. 3). The delta distribution was introduced by Aitchison and Brown (1957) to address highly skewed data by considering the proportion of positive values and using separate estimators for positive and zero values when estimating the mean and variance (Formula 4, Appendix B). This distribution assumes that the positive values follow a log-normal distribution. The below equations were used to estimate mean CPUE per year, and similar equations were used to estimate the total population of horseshoe crabs.

Model-based approach

For the DBA, a hurdle model (also referred to as a delta-lognormal model) was used for its ability to handle excess zero observations in large datasets (Lo et al. 1992; Maunder and Punt 2004) and similarity to the delta-lognormal design-based approach. The assumption of a hurdle model is that the process that results in the high frequency of zeroes is not necessarily the same process that results in the positive data (Cameron and Trivedi 1998). Preliminary analysis suggested that this could be the case for all six demographic groups of horseshoe crabs, but especially for newly-mature males and females (Table 3, Figure 2). This type of model uses a two-stage approach that models the probability of occurrence using a binomial distribution and, for this study, a lognormal distribution to model positive observations (Maunder and Punt 2004). These processes are assumed to be independent, and the parameters of interest are estimated by solving for the maximum-likelihood estimate of each stage separately.

Lognormal correction factor

When using a lognormal distribution, there are many methods in the literature to correct the bias presented when back-transforming a lognormal estimate. Clifford et al. (2013) go into great detail about these different estimates in their comparison of lognormal correction factors using a simulation analysis. One correction factor that is commonly implemented is the residual maximum likelihood (REML) estimate, $\hat{\mu}_{Est}(\chi_0) = exp\left(\frac{1}{2}s^2\right) * exp\left(\chi_0^{\tau}\hat{\beta}\right)$, where s^2 is the estimated residual variance and $\chi_0^{\tau}\hat{\beta}$ is the $E(Y_0|\chi_0)$. Another, from Bradu and Mundlak (1970), is referred to by Clifford et al. as the uniform minimum variance unbiased (UMVU), $\hat{\mu}_{Est}(\chi_0) = {}_0F_1\left(\frac{m}{2};\frac{m(1-\nu(\chi_0))}{4}s^2\right) * exp\left(\chi_0^{\tau}\hat{\beta}\right)$, where m is the degrees of freedom, and $\nu(\chi_0) = \chi_0^{T}(X^{T}X)^{-1}\chi_0$. While the REML method treats the correction factor as a constant, the UMVU approach includes uncertainty around the $\hat{\beta}$ estimate as covariates change, which for this study resulted in a different correction factor per year. Clifford et al. (2013) concluded that the UMVU estimate was the only “truly unbiased” estimate, although others, not including REML, resulted in similar performance in terms of bias, mean square errors, and mean square predictive error. Lo et al. (1992) was one of the first studies in fisheries science to use a hurdle model, referred to as a delta-lognormal model, to model relative abundance through CPUE standardization. In that study, the authors wrote that a standardized and unbiased measure of anchovy in positive flights using a lognormal distribution during year j is $d_j = exp\left(\beta_0 + b_{ij} + \sigma^2/2\right)$, where $\hat{d}_j = exp\left(\hat{\beta}_0 + \hat{b}_{ij}\right)\psi_{dj}$ in which ψ_d is a correction for bias. This correction bias and associated formula are attributed to Bradu and Mundlak (1970) in Appendix B. These two estimates (d_j and where $\hat{d}_j$) represent the REML and UMVU methods in Clifford et al. (2013) respectively, and suggest that either approach should provide approximately equal correction factors for the lognormal estimate. For the six demographic group models in this study, the two correction factors were similar across all years within a demographic group. For the purposes of this study, all estimates from the lognormal distribution used the UMVU correction factor.

Variable selection

A forward stepwise regression based on Akaike's (1973) information criterion (*AIC*) was performed to choose the best combination of explanatory variables for both stages of the hurdle model independently. Variables that were initially selected but were both non-significant (*p*-value <0.05) and did not decrease the *AIC* by at least 3 were removed, resulting in the final best-fitting models. The explanatory variables considered for all models were: year, latitude, longitude, depth (inshore/offshore), topography (trough/non-trough), average trawl depth, time-of-day (day/night), month (August, September, October, November), bottom temperature, bottom salinity, and distance from shore. Latitude, longitude, and distance from shore were all based upon the midpoint of a given trawl.

Estimating year effect

The year effect can be extracted by multiplying the expected values from each stage together after back-transformation and adding a bias correction to the lognormal stage estimates (Formula 6, Appendix B). The mean value of the continuous variables and a weighted average of the categorical variables based on relative frequency were input into the models to estimate the year effect. The weighted average proved especially important, as the final year effect estimate was sensitive to the time (day/night) and month categorical variables for both immature and mature individuals.

A 1000-iteration bootstrap analysis with replacement was performed for all demographic groups to build quantile intervals around the year effect estimates. This process involved randomly sampling the data with replacement up to the total sample size, refitting the model, and estimating the year effect from this updated model. These quantile intervals were used to estimate the uncertainty around each year-effect estimate and determine the significance in the change of relative abundance compared to previous years. To determine the effect of each explanatory variable on CPUE, we used the 1000-iteration bootstrap estimates from above. During each iteration, we calculated the relative abundance of all demographic groups while changing one explanatory variable's input value and keeping all others at their respective means for continuous variables, and weighted means for categorical variables. For continuous variables, we tested all values between the observed minimum and maximum of each variable over the time series, dividing each range to have approximately 150 values. For testing categorical variables, we set all categories within the selected explanatory variable to zero except the category of interest, which was set to one.

Predictive ability

Out-of-sample metrics to examine the models' predictive power on unseen data (not used to train the model) were tested using a *k*-fold cross validation (CV). For this type of CV, the models were refit using *k*-1 of the folds (training data), and then the CPUE of the retained *k* data (testing data) was predicted using the refit models. This was repeated *k* times until all unique group combinations had been tested. Ten-fold, 5-fold, and 3-fold cross-validation schemes were tested. When using a hurdle model, it is important that when splitting the data into folds, the proportion of zeros in each fold remains the same. This is because the proportion may affect the overall fit of each stage, possibly resulting in drastically different coefficient estimates from those estimated when initially fitting the models with all available data. Stratified cross-validation ensured these proportions by stratifying the data based on zero or non-zero CPUE and randomly sampling the correct proportion from each stratum, keeping the proportion approximately the same in all folds (Kohavi 1995). A stratified 10-fold cross-validation was performed in addition to the three other traditional cross-validations.

Results

Variable selection

Explanatory variables selected by each model varied by demographic group and modeling stage. Comparing the two salinity scenarios, adjusting 2020-2022 salinities did not change the explanatory variables selected except for the immature male model, which gained bottom salinity in the binomial stage. For all demographic groups, average depth, distance from shore, and longitude were the continuous variables selected in at least one stage of all the respective hurdle models (Table 4). Bottom temperature was considered significant for both immature and mature individuals in at least one modeling stage, as was bottom salinity if considering the adjusted-salinity scenario. Latitude was selected in both mature male and female models in at least one modeling stage. Of the four categorical explanatory variables considered, only month and time-of-day were selected for all demographic groups in at least one stage (Table 4). Topography (trough/non-trough) was selected for newly-mature females in the lognormal stage and immature males in the binomial stage. Depth (inshore/offshore) was never selected as an explanatory variable.

Although variable selection differed among demographic groups, variables shared among demographic groups had the same positive or negative relationship with CPUE. From testing the effects of individual explanatory variables, we found that the effect of topography for newly-mature females and immature males was not significantly different between trough and non-trough strata, although trough estimates did have a higher estimated mean (Figure 3). Time-of-day affected the estimated CPUE of all demographic groups, each having a higher mean and upper 95% quartile range during the night, although significantly higher only for immature individuals (Figure 4). A similar pattern was observed when assessing the month variable, as all demographic groups except newly-mature males had the highest mean CPUE and upper quartile range during September and lowest in August (Figure 5). The estimated CPUE for September was significantly higher than any other month for immature females and significantly higher than August for immature males and mature individuals. Newly-mature males had their highest mean CPUE in November, although nearly identical to the CPUE for September and not significantly different from that of any other month. Bottom temperature and salinity were negatively correlated with CPUE for all immature and mature individuals, with bottom temperature having a larger relative effect than bottom salinity (Figure 6). While average depth had a positive correlation with CPUE for all demographic groups, distance from shore was negatively correlated with CPUE for all demographic groups. Longitude was also negatively correlated with

CPUE for all groups, with mature males appearing to be most affected. Latitude had a positive correlation with mature individuals' CPUE. Compared to the unadjusted-salinity scenario, the adjusted-salinity scenario did not noticeably change the trend or magnitude of any variables, except for the addition of bottom salinity for immature males, which had a negative correlation with CPUE (Figure 7).

The design-based and model-based approaches showed similar estimates of relative abundance between both design-based distributions (Figure 8). All estimated values from the design and model-based approaches, along with their associated statistics, can be found in Appendix C. In the unadjusted-salinity scenario, only mature females in 2020 and immature males in 2005 and 2021 had significantly different model-based estimates from the normal or delta-lognormal design-based estimates. These same years and demographic groups were also significantly different in the adjusted-salinity scenario, along with immature females estimates in 2021. In this study, a significant difference between relative abundance estimates was defined as the design-based 95% confidence interval and the model-based 95% quantile interval not overlapping for a given year. Although only a few model-based estimates of relative abundance were significantly different from the design-based approach, 62% of all female and 56% of all male model-based estimates were more precise than the design-based estimates when comparing the uncertainty interval ranges from each estimation approach.

Comparing the unadjusted and adjusted models' year effects, the estimated year effects of each scenario were not significantly different in any year (Figure 9). Among immature females and mature individuals, the adjusted-salinity scenario year effect estimate was lower than the unadjusted scenario in all years except 2020 – 2022. During these years, the adjusted-scenario year effect quantile intervals were higher, although not significantly so, than the unadjusted quantile intervals, notably for mature males and mature females. For immature males, the adjusted-scenario year effect mean was slightly, non-significantly, higher in 2003, 2005-2008, 2010, 2016, and 2022. While the adjusted-salinity scenario resulted on average in more precise estimates (smaller 95% quantile intervals) than the unadjusted-salinity scenario, the improvements were minor and not consistent across demographic groups or years. Relative abundance of newly-mature individuals did not exhibit differences between the adjusted and unadjusted salinity scenarios, as bottom salinity was not selected as an explanatory variable in either model.

Relative abundance

Overall, relative abundance has varied through the time series for all demographic groups. Immature individuals have been highly variable over time with no apparent trend. Mature males and females had relatively consistent estimates of relative abundance between 2003-2011, although they have been more variable since 2018, with mature females showing a possible positive trend in recent years. The relative abundances of newly-mature individuals have been highly variable, although, in recent years, newly-mature females appear to have a small declining trend. Neither of the model-based scenarios consistently estimated higher or lower year effects when compared to either design-based estimator.

Discussion

Abundance trends/year effect

We found that both design-based and model-based approaches resulted in similar estimates of relative abundance for all demographic groups. The only demographic group models and years that differed significantly in estimates of relative abundance between the model-based approach from those of either of the design-based approaches were immature-males in 2005 and 2021, immature females in 2021, and mature females in 2020. The most dramatic differences were found in 2020 and 2021 among mature males and females, likely due to these years having record high numbers of these individuals in single trawls. Relative abundance has varied throughout the time series for all demographic groups, although some short-term trends were apparent. Low estimates of newly-mature females have been noted in reports to the ASMFC (Wong et al. 2023) based on a design-based approach and were again noted in these model-based estimates. These reports also suggested a recent minor positive trend for newly-mature males that is evident also in the most recent design-based estimate, although not as apparent in the model-based approach. This decline of newly mature females could become an area of concern for recruitment if estimates of newly-mature individuals remain exceptionally low over a run of years. In recent years, very few newly-mature females have been caught during the sampling period, with zero caught in 2021. While this could be due to a true decrease in relative abundance from previous years, there could be other reasons for this observation. One reason could be due to the time-lag for reaching maturity. The estimated age of maturity for female horseshoe crabs in the Delaware Bay is ten to twelve years (Smith et al. 2009a). When comparing estimates of immature females between 2007-2011 to mature females ten to thirteen years later, similar patterns of low relative abundance compared to previous years can be seen. It is important to note that this is not a closed system (Hallerman et al. 2022) and that migration and predation will affect the overall recruitment rate. Another possibility could be due to gear catchability and the area sampled. Lastly, the lack of newly mature females may be due to an observation error from classifying newly mature females as mature females, resulting in a low-biased estimate of abundance.

Spatial Effects

One of the major benefits of standardizing catch rate is to determine whether and which external factors affect CPUE and to what degree. Average depth, distance from shore, and longitude were spatial factors shown to affect CPUE for all demographic groups. Latitude was also selected as a significant explanatory variable for mature individuals in at least one modeling stage. The estimated effect of these spatial variables upon CPUE agrees with the justification for the stratified design of this study. Part of the original stratification design was dividing the study area into inshore and offshore areas because horseshoe crab abundance decreases with depth (Hata and Berkson 2004, citing Botton and Ropes 1987). While no models for any demographic group selected the depth categorical variable as influencing CPUE, average depth and distance from shore were selected for all demographic groups in both stages of the hurdle model, except for newly-mature females in the log-normal stage. However, these two explanatory variables had opposite effects on CPUE, despite being correlated (0.694) in the Delaware Bay area. These outcomes show that increasing distance from shore decreases CPUE, while increasing average depth increases CPUE. This agrees with the stratification of inshore vs offshore, but not the justification of lower depths resulting in fewer crabs. This initial justification could instead be attributed to crabs moving offshore during the time of sampling, rather than an effect of average depth. Additionally, the effects of both variables are confounded with that of the depth (inshore/offshore) categorical variable (Figure 10a). These reasons likely resulted in the models not selecting depth as an explanatory variable. Topography was the second stratification adopted for the trawl survey, as commercial fishermen stated that crabs are more abundant in troughs (Hata and Berkson 2004). Among all hurdle models, only those for newly-mature females and immature males selected topography as a significant variable to explain differences in CPUE.

Further, topography was selected in only one stage of each model, with the individual explanatory variable effects analysis showing no significant difference between trough and non-trough for either group. This outcome suggests that it may not be topography causing the different CPUE values, but the increased depth associated with a trough that results in higher CPUE. Topography was not found to be confounded with average depth or any other explanatory variables (Figure 10b). These results suggest that trawls in deeper waters near shore will have higher catch rates than those farther offshore. Latitude and longitude were heavily correlated within the Delaware Bay (0.91) due to the shape of the coastline. This confoundment may partly explain why longitude was selected as an explanatory variable in the models for all demographic groups and latitude only for mature individuals models. For mature males and females, latitude had an opposite, positive relationship with CPUE compared to longitude. When considering the sample area, generally, the higher the latitude of the trawl, the closer it is to the mouth of the bay. This positive correlation between latitude and mature individuals CPUE would make sense, as the bay itself acts as a highly-productive estuarine nursery and holding system, with abundant food sources (Anderson and Shuster 2003). Longitude was not strongly correlated with average depth (0.146) or distance from shore (0.256), but did have the same negative relationship to CPUE as distance from shore. This relationship between longitude and CPUE could be due to a variety of environmental variables that are heterogeneous across space, including temperature, salinity, and water currents. It is also possible that the relationship of CPUE with longitude and latitude cannot be fully explained by a linear relationship as modeled here and would require a polynomial function or spline to parse out a more complex, non-linear relationship.

Temporal and Environmental Effects

September had the highest mean CPUE estimates for all demographic groups except for newly-mature males. This effect is likely a reflection of post-breeding seasonal migration. It is thought that until around the age of 8, immature individuals stay within the Delaware Bay, after which females migrate to the continental shelf to mature, while males tend to remain in the bay to mature (Smith et al. 2009a). Both sexes tend to move deeper into the water as they age, with larger immature horseshoe crabs having been found along the continental shelf (Anderson and Shuster 2003). Between May and June in the Delaware Bay region, mature individuals migrate inshore to breed and then migrate offshore to the continental shelf post-breeding, although some remain in nearby embayments (Shuster and Botton 1985; Michels et al. 2008; Smith et al. 2009b, 2017). This migration could explain the higher CPUE estimates in September if trawls intercepted newly-mature and mature crabs migrating offshore. It is also possible that our sampling area is not large enough to truly capture the offshore migration and that more crabs are caught during these months due to higher nutrient and prey item availability, leading to increased activity that makes horseshoe crabs more vulnerable to being caught. This could also help explain why the CPUE of immature individuals is estimated to be higher in these months when they themselves do not participate in the breeding migration.

Month was not found to be confounded with any spatial variables such as distance from shore or longitude. Trawls performed at night exhibited higher mean estimated abundances than those during the day, though significantly so only for immature individuals. The explanation is less clear, as a few laboratory (Watson et al. 2009; Dubofsky et al. 2013) and telemetry (Watson and Chabot 2010) studies suggested that, while there are observations of both diurnal and nocturnal activity patterns in horseshoe crabs, there is not a consistent preference. That said, this result does agree with findings from the pilot study that found trawls at night had higher catches on average than those during the day (Hata and Berkson 2004). Hata and Berkson suspected that this was due to horseshoe crabs burrowing during the day and being active at night, making them more susceptible to the trawl. Time-of-day had no confounding variables.

Bottom salinity and bottom temperature were the only measured environmental parameters considered and both were selected as a significant explanatory factor in at least one stage of the hurdle model for immature and mature individuals when considering the adjusted salinity scenario. Bottom salinity had little effect on CPUE, likely as adult horseshoe crabs are known to tolerate an extremely wide range of salinity concentrations (Towle and Henry 2003), with only embryonic and larval development being affected by salinity (Botton et al. 2003). Interestingly, these models suggested that CPUE of immature and mature individuals increased with cooling temperatures. This was not expected, as Watson et al. (2009) found that there is decreased behavior in colder temperatures outside of the breeding season. Bottom salinity and bottom temperature were partly confounded by the month categorical variable, suggesting that the inclusion of month could partly mask the effects of temperature and salinity (Figure 10a). One reason that both variables may have been selected despite having the same relationship to CPUE and month having been selected by all demographic groups, is that they have opposite correlations with time; as it gets later in the year, bottom temperature decreases while bottom salinity increases. Month would not be able to fully account for these opposite relationships, which may suggest why models for immature and mature crabs still selected these explanatory variables.

Cross-validation and trawl removal

From the cross-validation analysis, there were minor differences between the two salinity scenarios' predictive ability. These results showed that increasing salinity had a negligible effect on all demographic groups' models' ability to predict CPUE based on unseen environmental parameters. Models for both mature sexes had higher mean-square-error (MSE) values in the cross-validation than newly-mature and immature individuals of the same sex (Table 5a). This is largely due to the CPUE of mature individuals being the highest and most variable among all maturity groups on average. However, we suspect that these high MSE values are also due, in part, to atypical trawls observed in some years. For this study, atypical trawls were defined as tows with over 1000 individuals of a specific demographic group caught within a single trawl. For mature males, this consisted of four trawls and for mature females one trawl (Figure 2). The atypical mature female observation occurred concurrently in one of the four atypical mature male trawls. In total, these data accounts for four distinct trawls in the entire time series, representing about 0.5% of our entire dataset. To try to account for these atypical trawls, two weighted hurdle models were tested, one based on absolute residuals and the other on squared residuals. The concept behind both approaches is that each trawl is assigned a weight based on the residuals from the original fitted model, with trawls having larger residuals being weighed less heavily in a new, refitted model. The effects of removing the trawls from the cross-validation entirely were also tested. This is equivalent to setting the weights of the four atypical trawls to zero in a weighted model. When comparing the standard hurdle model to the two weighted hurdle models, there was little difference in predictive performance (Table 5b). Removal of these four trawls resulted in a large decrease in MSE, going from tens of thousands to thousands for mature females and hundreds of thousands to tens of thousands for mature males. These results suggest that these atypical trawls had a considerable influence on model-fitting and that the models for mature demographic groups were unable to accurately predict these extreme events. Further investigation is needed to determine how best to handle these atypical events.

Adjusted vs Unadjusted Salinity Scenarios

Due to a suspected error in salinity measurements during the three most recent years, 2020-2022, a two-scenario approach was used to investigate if this error would lead to differences in variable selection, estimates of relative abundance, and model performance. There was negligible effect on variable selection and predictive ability. Regarding model performance, 59% of all female relative abundance estimates and 59% of all male estimates from the adjusted-salinity scenario were more precise than the unadjusted-salinity scenario, although the improvement was minimal. For females, the 95% quantile interval range decreased by an average of 2.7% from the unadjusted-salinity scenario to the adjusted scenario across all years, and for males, there was an average decrease of 8.0%, greatly due to estimates of mature males in recent years. There were additional differences in the mean relative abundance estimates of any given year between the two scenarios. In either scenario, all models that selected salinity as a significant explanatory factor estimated a negative relationship between bottom salinity and CPUE. This outcome would suggest that in a linear model like those used here, if all other factors remained the same while salinity increased, relative abundance would decrease, resulting in a lower estimated year effect. When increasing the salinity measurements between 2020 and 2022 from the unadjusted scenario, this was indeed the effect for most years in the adjusted-salinity scenario. Although not significantly different, among immature females and mature individuals, the adjusted-salinity scenario year effect mean was higher than the unadjusted between 2020 – 2022, especially so for mature males. For immature males, the adjusted-scenario year effect mean was slightly higher in 2003, 2005-2008, 2010, 2016, and 2022. The reason for this unexpectedly higher value is how the adjusted-models fit the new data and how parameter coefficients were estimated to maximize the log-likelihoods of the models. The goal of a multiple linear regression analysis is to determine the best linear relationship among explanatory variables to describe any overall patterns in the data. In the adjusted scenario, salinity was not weighted as heavily as other explanatory variables to fit the overall patterns in the data. This approach resulted in adjusted salinity having less of an effect on the year estimates than in the unadjusted scenario compared to other explanatory variables, resulting in these unexpected values. For immature individuals and mature males, this led to all year coefficients increasing, with years that had larger differences between adjusted and unadjusted relative abundance having increased more than others (Table 6) (Appendix A, Figures S3, S4). For mature females, this coefficient increase was found only between 2020-2022. Mature males had the largest difference in estimated year effect between 2020 and 2022. This demographic group is unique because the mature male model was the only model that selected bottom salinity as a significant explanatory factor in the lognormal stage of the adjusted-salinity scenario. This resulted in the relatively small difference in the log-scale year effect estimates to be exponentially increased after back-transforming the estimate to the normal scale (Appendix A, Figure S5). Additionally, the lognormal correction factor also increased for mature males in the adjusted-salinity scenario between 2020–2022 when compared to the unadjusted-salinity scenario (Appendix A, Figure S6), further increasing the difference between year-effect estimates. These changes appear to be at least part of the cause in the difference observed in the final year-effect estimates between the two salinity scenarios.

Future recommendations

Our results suggest that the design-based approach to modeling the relative abundance of horseshoe crabs is informative and useful for estimation, showing similar trends in population dynamics to the model-based estimates with only a few years' estimates being significantly different between the two approaches. This in mind, the model-based approach was more precise on average than the design-based approach when comparing the uncertainty intervals of each approach around a given year's relative abundance estimate. For stock assessment purposes, the model-based estimates from the unadjusted-salinity scenario should be used for investigating trends in relative abundance, as they account for external factors not considered by the design-based approach.

Through catch rate standardization, the model-based approach revealed that temporal variables such as time and month, as well as spatial variables, such as average depth and distance from shore, had effects on CPUE. The current stratification design allows the design-based estimates to capture changes in relative abundance with similar precision, although on average lower, to the model-based approach presented here even with external influences. However, a redesign of the trawl survey may reveal more nuanced information on how external factors are affecting CPUE and increase the accuracy of total abundance estimates using a designed-based approach. Stratification by true depth (e.g., 20-40ft, 40-60ft, etc.) and distance from shore (e.g., 3-6nm, 6-9nm, 9-12nm) may prove more informative than the coarser stratification currently in place. As shown in this study, average depth and distance from shore have opposite effects on CPUE and cannot be represented by one categorical depth (inshore/offshore) variable. Topography may still prove to be useful, as it is not confounded with average depth and was selected as a significant explanatory variable by the newly-mature female model, a key demographic group of interest. In a design-based approach, total abundance is estimated by extrapolating the estimated mean relative abundance of each stratum to the proportion it represents in the total study area. Assuming that horseshoe crabs are randomly distributed and homogenous within these new strata, a redesign would allow for more accurate extrapolation of the sampled areas to the total area when estimating total abundance due to its finer scale using information already readily available.

Crew availability and weather have determined the timing of sampling in the past. If possible, sampling should occur more evenly between the current sampling months of August and November and the time-of-day (day/night) that sampling takes place. While CPUE standardization helps determine the effect of these external factors, it cannot do this precisely without thorough sampling. Of the 697 trawls performed in the DBA over the time-series, only 13% and 12% took place in August and November respectively, and 35% and 40% in September and October. Having more even sampling would allow our estimates of relative abundance to be more robust and confirm that the low estimates seen in August for all demographic groups are due to a seasonality or post-breeding migration effect, rather than due to a lack of data. The same consideration can be applied regarding the effect of time-of-day on relative abundance. While trawls at night were expected to have higher mean CPUEs for all demographic groups, 70% of the trawls analyzed in this study occurred during the day, with only 30% at night. In the pilot study, Hata and Berkson (2004) stated that night-time sampling resulted in more horseshoe crabs per trawl and was more precise than daytime sampling. Our results further support this finding and emphasize the need for more even sampling between the night and day. If a previous sampling year had consisted mainly of daytime samples, these results suggest that the relative and total abundance estimates from the design-based approach may be underestimating the true abundance. More consideration should be placed on these factors when planning sampling schedules if weather and crew availability permit.

Considering the long-term management of horseshoe crabs, one of the most important aspects of population recovery is recruitment. With the already largely male-skewed operational sex ratio (OSR) in the Delaware Bay area, and high relative catches of all male demographic groups, males are not of major concern currently. A primary focus of the ASMFC is the relative abundance of newly-mature females in the Delaware Bay area. This demographic group acts as a marker for the overall trends in abundance of the population, as recruitment of newly-mature females determines whether enough immature females are reaching sexual maturity to maintain the population growth rate at or above a minimum replacement level. While there has been an estimated decrease in the relative abundance of newly-mature females since 2018, this may not be representative of the population. This could be due to the aforementioned effects of sampling month and time-of-day, or that our results are underestimating the true relative abundance of newly-mature females, as there has not been an associated decrease observed in relative abundance estimates of mature females. But with this in mind, relative abundance estimates of mature females are much less sensitive to changes in abundance than that of newly-mature females. This difference is because the newly-mature classification constitutes only one year of a twenty-year lifespan while those of immature and mature classifications consist of eight to ten year-classes. If newly-mature recruitment is truly low, it may take years for it to have a noticeable change in the relative abundance of mature females. This single-year classification works for and against management, as newly-mature females could act as an early warning sign for population decline if they are accurately sampled at a proportion representative of the entire population. This aspect raises the final concern of observation error. As the maturity classifications are based greatly on the external visual indicators, with some horseshoe crabs also being pierced with an awl to check for eggs, there is a chance that some newly-mature females are being classified as mature females. While we do not suspect this to be the case, it is a possibility to consider. In conclusion, management should continue to err on the side of caution regarding the female HSC population in the Delaware Bay. Due to female-biased harvest and the importance of HSC eggs for maintaining the threatened red-knot population, a moderately conservative approach would allow future estimates of relative abundance to confirm or deny any apparent demographic trends without causing major long-term consequences to the female component of the population.

Literature Cited

Aitchison, J. 1955. On the Distribution of a Positive Random Variable Having a Discrete Probability Mass at the Origin. Journal of the American Statistical Association 50(271):901–908.

Aitchison, J., and J. A. C. Brown. 1957. The Lognormal Distribution with Special Reference to Its Uses in Economics. Cambridge University Press, Cambridge, U.K.

Anderson, L. I., and C. N. Shuster. 2003. Throughout Geologic Time: Where Have They Lived? Pages 189–223 *in* C. N. Shuster, H. J. Brockmann, and R. B. Barlow, editors. The American Horseshoe Crab. Harvard University Press, Cambridge, MA.

ASMFC. 2008. Horseshoe Crab Technical Committee Report. Washington, D.C. Available: https://www.asmfc.org/uploads/file/july08TCReport.pdf

ASMFC. 2019. 2019 Horseshoe Crab Benchmark Stock Assessment. Atlantic States Marine Fisheries Commission, Arlington, VA. Available: https://asmfc.org/uploads/file/63d2ed62HSCAssessment_PeerReviewReport_May2019.pdf

ASMFC. 2022. 2021 Horseshoe Crab Adaptive Resource Management Revision Overview. Atlantic States Marine Fisheries Commission, Arlington, VA. Available: https://www.asmfc.org/uploads/file/63d2eb79HSC_ARM_RevisionOverview_Jan2022.pdf

Aubry, P., and C. Francesiaz. 2022. On Comparing Design-based Estimation Versus Model-based Prediction to Assess the Abundance of Biological Populations. Ecological Indicators 144:109394.

Botton, M. L., B. A. Harrington, N. Tsipoura, and D. Mizrahi. 2003. Synchronies in Migration: Shorebirds, Horseshoe Crabs, and Delaware Bay. Pages 5–32 *in* C. N. Shuster, H. J. Brockmann, and R. B. Barlow, editors. The American Horseshoe Crab. Harvard University Press, Cambridge, MA.

Botton, M. L., and J. W. Ropes. 1987. Populations of Horseshoe Crabs, *Limulus polyphemus*, on the Northwestern Atlantic Continental Shelf. Fishery Bulletin 85(4):805–812.

Bradu, D., and Y. Mundlak. 1970. Estimation in Lognormal Linear Models. Journal of the American Statistical Association 65(329):198–211.

Brockmann, H. J., and D. Penn. 1992. Male Mating Tactics in the Horseshoe Crab, *Limulus polyphemus*. Animal Behaviour 44(4):653–665.

Cameron, A. C., and P. K. Trivedi. 1998. Regression Analysis of Count Data. Page Hammond, Peter and Alberto Holly, editors. Cambridge University Press, Cambridge, MA.

Clifford, D., N. Cressie, J. R. England, S. H. Roxburgh, and K. I. Paul. 2013. Correction Factors for Unbiased, Efficient Estimation and Prediction of Biomass from Log-Log Allometric Models. Forest Ecology and Management 310:375–381.

Cragg, J. G. 1971. Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods. Econometrica 39(5):829.

Dubofsky, E. A., S. D. Simpson, C. C. Chabot, and W. H. Watson. 2013. Patterns of Activity Expressed by Juvenile Horseshoe Crabs. The Biological Bulletin 225(1):42–49.

ERSI. 2023. ArcGIS Pro Version 3.1.0. Environmental Systems Research Institute, Inc., Redlands, CA.

Faurby, S., K. S. K. Nielsen, S. Bussarawit, I. Intanai, N. van Cong, C. Pertoldi, and P. Funch. 2011. Intraspecific Shape Variation in Horseshoe Crabs: The Importance of Sexual and Natural

Selection for Local Adaptation. Journal of Experimental Marine Biology and Ecology 407(2):131–138.

Hallerman, E. M., D. N. Hata, M. S. Eackles, and T. L. King. 2022. Population Genetics and Movement Show Metapopulation Dynamics of Mid-Atlantic Region Horseshoe Crabs. Pages 41–57 *in* J. T. Tanacredi, M. L. Botton, P. K. S. Shin, Y. Iwasaki, S. G. Cheung, K. Y. Kwan, and J. H. Mattei, editors. International Horseshoe Crab Conservation and Research Efforts: 2007-2020. Springer International Publishing, Cham, Switzerland.

Hata, D., and J. Berkson. 2004. Factors Affecting Horseshoe Crab *Limulus polyphemus* Trawl Survey Design. Transactions of the American Fisheries Society 133(2):292–299.

Hata, D. N., and E. M. Hallerman. 2022. Relative Abundance of Horseshoe Crabs in the Delaware Bay Region: A Critical Factor for Adaptive Resource Management. Pages 415–433 *in* J. T. Tanacredi, M. L. Botton, P. K. S. Shin, Y. Iwasaki, S. G. Cheung, K. Y. Kwan, and J. H. Mattei, editors. International Horseshoe Crab Conservation and Research Efforts: 2007- 2020. Springer International Publishing, Cham, Switzerland.

James-Pirri, M.-J., P. A. Veillette, and A. S. Leschen. 2012. Selected Hemolymph Constituents of Captive, Biomedically Bled, and Wild Caught Adult Female American Horseshoe Crabs (*Limulus polyphemus*). Marine and Freshwater Behaviour and Physiology 45(4):281–289.

Karpanty, S. M., J. D. Fraser, J. Berkson, L. J. Niles, A. Dey, and E. P. Smith. 2006. Horseshoe Crab Eggs Determine Red Knot Distribution in Delaware Bay. The Journal of Wildlife Management 70(6):1704–1710.

Kohavi, R. 1995. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. Pages 1137–1143 Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

Kreamer, G., and S. Michels. 2009. History of Horseshoe Crab Harvest on Delaware Bay. Pages 299–313 *in* J. T. Tanacredi, M. L. Botton, and D. R. Smith, editors. Biology and Conservation of Horseshoe Crabs. Springer, Boston, MA.

Lo, N. C., L. D. Jacobson, and J. L. Squire. 1992. Indices of Relative Abundance from Fish Spotter Data based on Delta-Lognormal Models. Canadian Journal of Fisheries and Aquatic Sciences 49(12):2515–2526.

Lohr, S. L. 2019. Sampling: Design and Analysis. CRC Press, Boca Raton, FL.

Maunder, M. N., and A. E. Punt. 2004. Standardizing Catch and Effort Data: A Review of Recent Approaches. Fisheries Research 70(2–3):141–159.

Maunder, M. N., and P. J. Starr. 2003. Fitting Fisheries Models to Standardised CPUE Abundance Indices. Fisheries Research 63(1):43–50.

McGowan, C. P., J. E. Hines, J. D. Nichols, J. E. Lyons, D. R. Smith, K. S. Kalasz, L. J. Niles, A. D. Dey, N. A. Clark, P. W. Atkinson, C. D. T. Minton, and W. Kendall. 2011a. Demographic Consequences of Migratory Stopover: Linking Red Knot Survival to Horseshoe Crab Spawning Abundance. Ecosphere 2(6):1–22.

McGowan, C. P., D. R. Smith, J. A. Sweka, J. Martin, J. D. Nichols, R. Wong, J. E. Lyons, L. J. Niles, K. Kalasz, J. Brust, M. Klopfer, and B. Spear. 2011b. Multispecies Modeling for Adaptive

Management of Horseshoe Crabs and Red Knots in the Delaware Bay. Natural Resource Modeling 24(1):117–156.

Michels, S. F., D. Smith, and S. Bennett. 2008. Horseshoe Crab Spawning Activity in Delaware Bay: 1999–2007: Report to the Horseshoe Crab Technical Committee. Atlantic States Marine Fisheries Commission, Arlington, VA.

Millard, M. J., J. A. Sweka, C. P. McGowan, and D. R. Smith. 2015. Assessment and Management of North American Horseshoe Crab Populations, with Emphasis on a Multispecies Framework for Delaware Bay, U.S.A. Populations. Pages 407–431 *in* R. H. Carmichael, M. L. Botton, P. K. S. Shin, and S. G. Cheung, editors. Changing Global Perspectives on Horseshoe Crab Biology, Conservation and Management. Springer International Publishing, New York City, New York.

Millero, F. J., and A. Poisson. 1981. International One-Atmosphere Equation of State of Seawater. Deep Sea Research Part A. Oceanographic Research Papers 28(6):625–629.

Mizrahi, D. S., and K. A. Peters. 2009. Relationships Between Sandpipers and Horseshoe Crab in Delaware Bay: A Synthesis. Pages 65–87 *in* J. T. Tanacredi, M. L. Botton, and D. Smith, editors. Biology and Conservation of Horseshoe Crabs. Springer, Boston, MA.

NOAA NCEI. 2022. ETOPO 2022 15 Arc-Second Global Relief Model. NOAA National Centers for Environmental Information. Available: https://doi.org/10.25921/fd45-gt74.

Olsen, A. R., J. Sedransk, D. Edwards, C. A. Gotway, W. Liggett, S. Rathbun, K. H. Reckhow, and L. J. Young. 1999. Statistical Issues for Monitoring Ecological and Natural Resources in the United States. Environmental Monitoring and Assessment (54):1–45.

Pennington, M. 1983. Efficient Estimators of Abundance, for Fish and Plankton Surveys. Biometrics 39(1):281–286.

Potts, J. M., and J. Elith. 2006. Comparing Species Abundance Models. Ecological Modelling 199(2):153–163.

Reid, D. F. 2008. NOAA Technical Memorandum GLERL-139 - Conversion of Specific Gravity to Salinity for Ballast Water Regulatory Management. Great Lakes Environmental Research Laboratory, Ann Arbor, MI.

Riska, B. 1981. Morphological Variation in the Horseshoe Crab *Limulus polyphemus*. Evolution 35(4):647.

Rudkin, D. M., and G. A. Young. 2009. Horseshoe Crabs – An Ancient Ancestry Revealed. Pages 25–44 *in* J. T. Tanacredi, M. L. Botton, and D. Smith, editors. Biology and Conservation of Horseshoe Crabs. Springer, Boston, MA.

Rutecki, D., R. H. Carmichael, and I. Valiela. 2004. Magnitude of Harvest of Atlantic Horseshoe Crabs, *Limulus polyphemus*, in Pleasant Bay, Massachusetts. Estuaries 27(2):179–187.

Sekiguchi, K., H. Seshimo, and H. Sugita. 1988. Post-Embryonic Development of the Horseshoe Crab. The Biological Bulletin 174(3):337–345.

Shuster, C. N. 1955. On Morphometric and Serological Relationships within the *Limulidae*, with Particular Reference to *Limulus polyphemus* (L.). Ph.D., New York University, Ann Arbor.

Shuster, C. N. 2015. The Delaware Bay Area, U.S.A.: A Unique Habitat of the American Horseshoe Crab, *Limulus polyphemus*. Pages 15–39 *in* R. H. Carmichael, M. L. Botton, P. K. S.

Shin, and S. G. Cheung, editors. Changing Global Perspectives on Horseshoe Crab Biology, Conservation and Management. Springer International Publishing, Cham, Switzerland.

Shuster, C. N., and M. L. Botton. 1985. A Contribution to the Population Biology of Horseshoe Crabs, *Limulus polyphemus* (L.), in Delaware Bay. Estuaries 8(4):363.

Shuster, C. N., and K. Sekiguchi. 2003. Growing Up Takes About Ten Years and Eighteen Stages. Pages 103–132 *in* C. N. Shuster, J. H. Brockmann, and R. B. Barlow, editors. The American Horseshoe Crab. Harvard University Press, Cambridge, MA.

Smith, D. R., H. J. Brockmann, M. A. Beekey, T. L. King, M. J. Millard, and J. Zaldívar-Rae. 2017. Conservation Status of the American Horseshoe Crab, (*Limulus polyphemus*): A Regional Assessment. Reviews in Fish Biology and Fisheries 27(1):135–175.

Smith, D. R., M. T. Mandt, and P. D. M. MacDonald. 2009a. Proximate Causes of Sexual Size Dimorphism in Horseshoe Crabs (*Limulus polyphemus*) of the Delaware Bay. Journal of Shellfish Research 28(2):405–417.

Smith, D. R., M. J. Millard, and R. H. Carmichael. 2009b. Comparative Status and Assessment of *Limulus polyphemus* with Emphasis on the New England and Delaware Bay Populations. Pages 361–386 *in* J. T. Tanacredi, M. L. Botton, and D. Smith, editors. Biology and Conservation of Horseshoe Crabs. Springer, Boston, MA.

Smith, D. R., P. S. Pooler, B. L. Swan, S. F. Michels, W. R. Hall, P. J. Himchak, and M. J. Millard. 2002. Spatial and Temporal Distribution of Horseshoe Crab (*Limulus polyphemus*) Spawning in Delaware Bay: Implications for Monitoring. Estuaries 25(1):115–125.

Sterba, S. K. 2009. Alternative Model-Based and Design-Based Frameworks for Inference from Samples to Populations: From Polarization to Integration. Multivariate Behavioral Research 44(6):711–740.

Towle, D. W., and R. P. Henry. 2003. Coping with Environmental Changes: Physiological Challenges. Pages 224–244 *in* C. N. Shuster, J. H. Brockmann, and R. B. Barlow, editors. The American Horseshoe Crab. Harvard University Press, Cambridge, MA.

Upmann, M., and C. Bonaparte. 1999. Rapid Methods for Food Hygiene Inspection. Pages 1887–1895 *in* Robinson, R. K., Batt, Carl A., and Patel, P.D., editors. Encyclopedia of Food Microbiology. Elsevier, San Diego, CA.

Walls, E. A., J. Berkson, and S. A. Smith. 2002. The Horseshoe Crab, *Limulus polyphemus* : 200 Million Years of Existence, 100 Years of Study. Reviews in Fisheries Science 10(1):39–73.

Watson, W. H., and C. C. Chabot. 2010. High Resolution Tracking of Adult Horseshoe Crabs *Limulus polyphemus* in a New Hampshire Estuary Using Fixed Array Ultrasonic Telemetry. Current Zoology 56(5):599–610.

Watson, W. H., S. Y. Schaller, and C. C. Chabot. 2009. The Relationship Between Small- and Large-Scale Movements of Horseshoe Crabs in the Great Bay Estuary and *Limulus* Behavior in the Laboratory. Pages 131–147 *in* J. T. Tanacredi, M. L. Botton, and D. Smith, editors. Biology and Conservation of Horseshoe Crabs. Springer, Boston, MA.

Wessel, P., and W. H. F. Smith. 1996. A Global, Self-Consistent, Hierarchical, High-Resolution Shoreline Database. Journal of Geophysical Research: Solid Earth 101(B4):8741–8743.

Wong, C., Y. Jiao, and E. Hallerman. 2023. Results of the 2022 Horseshoe Crab Trawl Survey: Report to the Atlantic States Marine Fisheries Commission Horseshoe Crab and Delaware Bay Ecology Technical Committees. Atlantic States Marine Fisheries Commission, Arlington, VA.

Yu, H., Y. Jiao, Z. Su, and K. Reid. 2012. Performance Comparison of Traditional Sampling Designs and Adaptive Sampling Designs for Fishery-Independent Surveys: A Simulation Study. Fisheries Research 113(1):173–181.

Table 1: Number of trawls performed per year in each stratum within the Delaware Bay area (DBA), for a total of 697 trawls. The four strata are inshore trough (IN/TR), inshore non-trough (IN/NT), offshore trough (OFF/TR), and offshore non-trough (OFF/NT).

	DBA				
Year	IN/TR	IN/NT	OFF/TR	OFF/NT	
2003	12	10	5	13	
2004	9	7	5	18	
2005	12	7	8	12	
2006	8	5	5	16	
2007	14	6	10	15	
2008	16	8	8	13	
2009	11	8	7	15	
2010	14	9	12	18	
2011	12	9	11	18	
2012	6	6	0	0	
2016	11	7	9	17	
2017	11	9	5	17	
2018	6	7	13	15	
2019	8	8	8	21	
2020	10	10	15	9	
2021	15	9	12	7	
2022	1	6	7	26	
Total:	176	131	140	250	**=697**

Table 2: Percentage of trawls performed annually in the Delaware Bay area during the day and at night between 2003 and 2022.

Year	Day Trawls	Night Trawls
2003	55%	45%
2004	38.46%	61.54%
2005	17.95%	82.05%
2006	67.65%	32.35%
2007	57.78%	42.22%
2008	6.67%	93.33%
2009	95.12%	4.88%
2010	98.11%	1.89%
2011	98%	2%
2012	0%	100%
2016	65.91%	34.09%
2017	97.62%	2.38%
2018	78.05%	21.95%
2019	86.67%	13.33%
2020	81.82%	18.18%
2021	95.35%	4.65%
2022	77.5%	22.5%

Table 3: Percentage of trawls performed during the study period (2003 – 2022) that captured zero or at least one horseshoe crabs of that demographic group in the Delaware Bay area (DBA). The six demographic groups are immature female (FI), newly-mature female (FN), mature female (FM), immature male (MI), newly-mature male (MN), and mature male (MM).

DBA	% ZERO	% POSITIVE
ANY	6.89%	93.11%
FI	24.82%	75.18%
FN	58.97%	41.03%
FM	18.51%	81.49%
MI	34%	66%
MN	55.09%	44.91%
MM	15.49%	84.51%
Total Trawls: 697		

Table 4: Explanatory variables selected in the binomial (B) and log-normal (LN) stages of the hurdle models for each demographic group from a forward stepwise regression. The top table represents variables selected in the unadjusted-salinity scenario, and the bottom table the adjusted-salinity scenario. Plus and minus symbols represent either a positive or negative relationship of that variable to CPUE. For categorical variables, the underlined variables are those set as the baseline.

UNADJUSTED	FI		FN		FM		MI		MN		MM	
	B	**LN**	**B**	**LN**	**B**	**LN**	**B**	**LN**	**B**	**LN**	**B**	**LN**
AVGDEP	+	+	+	+	+	+	+	+	+	+	+	+
BOTSAL	--				--							--
BOTTEMP	--	--				--	--	--				--
DEPTH (<u>IN</u>/OFF)												
Distance from shore	--	--	--		--	--	--	--	--	--	--	--
MIDLAT					+	+						+
MIDLON	--	--	--	--	--	--	--	--	--	--		--
MONTH (<u>8</u>/9/10/11)	+	+		+	+	+	+	+	+	+	+	+
TIME (<u>Day</u>/Night)	+	+	+		+	+	+	+	+	+	+	+
TOPO (<u>NT</u>/TR)				+			+					

ADJUSTED	FI		FN		FM		MI		MN		MM	
	B	**LN**	**B**	**LN**	**B**	**LN**	**B**	**LN**	**B**	**LN**	**B**	**LN**
AVGDEP	+	+	+	+	+	+	+	+	+	+	+	+
BOTSAL	--				--		--					--
BOTTEMP	--	--				--	--	--				--
DEPTH (<u>IN</u>/OFF)												
Distance from shore	--	--	--		--	--	--	--	--	--	--	--
MIDLAT					+	+						+
MIDLON	--	--	--	--	--	--	--	--	--	--		--
MONTH (<u>8</u>/9/10/11)	+	+		+	+	+	+	+	+	+	+	+
TIME (<u>Day</u>/Night)	+	+	+		+	+	+	+	+	+	+	+
TOPO (<u>NT</u>/TR)				+			+					

Table 5a: Results of *k*-fold cross-validation on best models for each of the six demographic group models among the two scenarios, unadjusted bottom salinity and adjusted salinity. Results presented are mean-square-error (MSE), root-mean-square-error (RMSE), and mean-absolute-error (MAE).

		Unadjusted			Adjusted		
	MATGRP	**MSE**	**RMSE**	**MAE**	**MSE**	**RMSE**	**MAE**
10-Fold	**FI**	3010.62	52.9	27.43	3010.34	52.9	27.43
	FN	148.08	9.45	3.64	148.08	9.45	3.64
	FM	51595.83	120.13	30.63	51589.94	120.11	30.63
	MI	1738.93	40.95	20.43	1737.12	40.94	20.42
	MN	587.8	18.6	6.82	587.8	18.6	6.82
	MM	195597.3	312.23	76.3	195762.8	312.92	76.44
5-Fold	**FI**	2750.52	52.38	26.82	2749.96	52.38	26.82
	FN	146.7	10.57	3.58	146.7	10.57	3.58
	FM	51933.76	151.76	30.7	51928.41	151.75	30.69
	MI	1807.47	42.27	20.77	1803.7	42.23	20.75
	MN	583.13	19.29	6.83	583.13	19.29	6.83
	MM	198391.1	364.21	77.57	198779.4	366.21	77.77
3-Fold	**FI**	2784.66	52.46	26.92	2784.13	52.46	26.91
	FN	150.46	11.7	3.57	150.46	11.7	3.57
	FM	52957.47	176.03	31.22	52951.37	176.01	31.21
	MI	1514.19	38.11	19.47	1512.63	38.09	19.46
	MN	610.29	23.13	7.13	610.29	23.13	7.13
	MM	190298.8	399.28	73.61	189976.5	399.05	73.58

Table 5b: Results of the 10-fold cross-validation for the four methods tested to alleviate the effect of atypical trawls on predictive performance in the unadjusted-salinity scenario, where ε represents the model residuals upon which weighting was based.

	MATGRP	MSE	RMSE	MAE
Standard Hurdle (*n*=697)	**FM**	51595.83	120.13	30.63
	MM	195597.3	312.23	76.3
Weighted Hurdle $\lvert\varepsilon\rvert$ (*n*=697)	**FM**	52497.18	123.81	31.81
	MM	196617.5	292.58	73.83
Weighted Hurdle $\lvert\varepsilon^2\rvert$ (*n*=697)	**FM**	51830.23	118.50	28.42
	MM	189259.5	277.04	66.50
Without atypical (*n*=693)	**FM**	1798.24	40.66	18.74
	MM	23413.95	129.05	47.78

Table 6: Difference between adjusted-salinity scenario estimates minus the unadjusted-salinity scenario estimates for the year explanatory variable coefficient (right) and the mean relative abundance (left). Bolded values indicate years in which the adjusted-salinity scenario relative abundance is higher than the unadjusted-salinity scenario.

	YEAR EXPLANATORY VARIABLE COEFFICIENT				MEAN RELATIVE ABUNDANCE (CPUE)			
	FI	**FM**	**MI**	**MM**	**FI**	**FM**	**MI**	**MM**
2003	-1.13	0.58	**-13.38**	-2.27	0.37	0.06	**-0.28**	1.55
2004	-1.11	0.61	-13.19	-2.25	0.22	0.06	0.02	0.82
2005	-1.12	0.61	**-13.24**	-2.26	0.39	0.2	**-0.05**	0.92
2006	-1.12	0.6	**-13.39**	-2.27	0.19	0.34	**-0.31**	1.98
2007	-1.12	0.58	**-13.58**	-2.28	0.62	0.63	**-1.8**	2.5
2008	-1.12	0.6	**-13.46**	-2.27	0.39	0.09	**-0.38**	2.13
2009	-1.11	0.64	-13.12	-2.24	1.34	0.42	1.21	1.19
2010	-1.12	0.58	**-13.52**	-2.27	0.2	0.04	**-0.53**	1.9
2011	-1.1	0.66	-12.87	-2.23	0.46	0.45	0.51	2.85
2012	-1.12	0.63	-13.05	-2.25	0	0.26	0	1.45
2016	-1.12	0.59	**-13.44**	-2.27	0.45	0.11	**-0.47**	3.65
2017	-1.11	0.65	-12.99	-2.24	0.81	0.29	0.51	3.99
2018	-1.1	0.67	-13.01	-2.24	0.51	0.38	0.39	3.82
2019	-1.1	0.66	-12.94	-2.24	0.71	0.47	0.47	3.95
2020	**-2.5**	**-0.5**	-13.04	**-2.81**	**-3.94**	**-2.55**	0.24	**-83.4**
2021	**-2.52**	**-0.5**	-13.17	**-2.82**	**-3.56**	**-3.05**	0.01	**-46.4**
2022	**-2.49**	**-0.48**	**-13.19**	**-2.79**	**-4.35**	**-4.1**	**-0.01**	**-11.7**

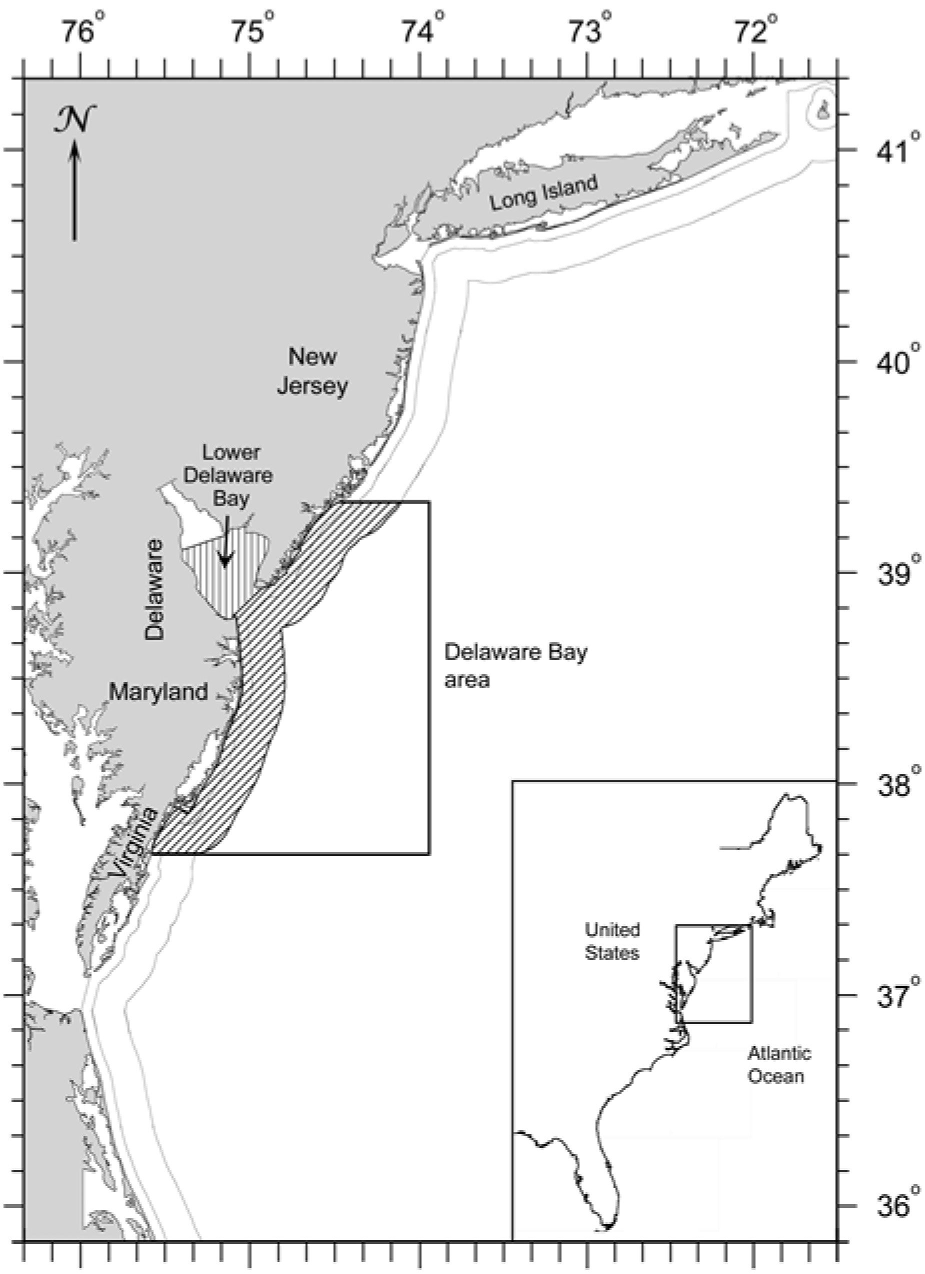


Figure 1: Horseshoe crab trawl survey sampling areas. Contours indicate three and 12 nautical miles from shore.

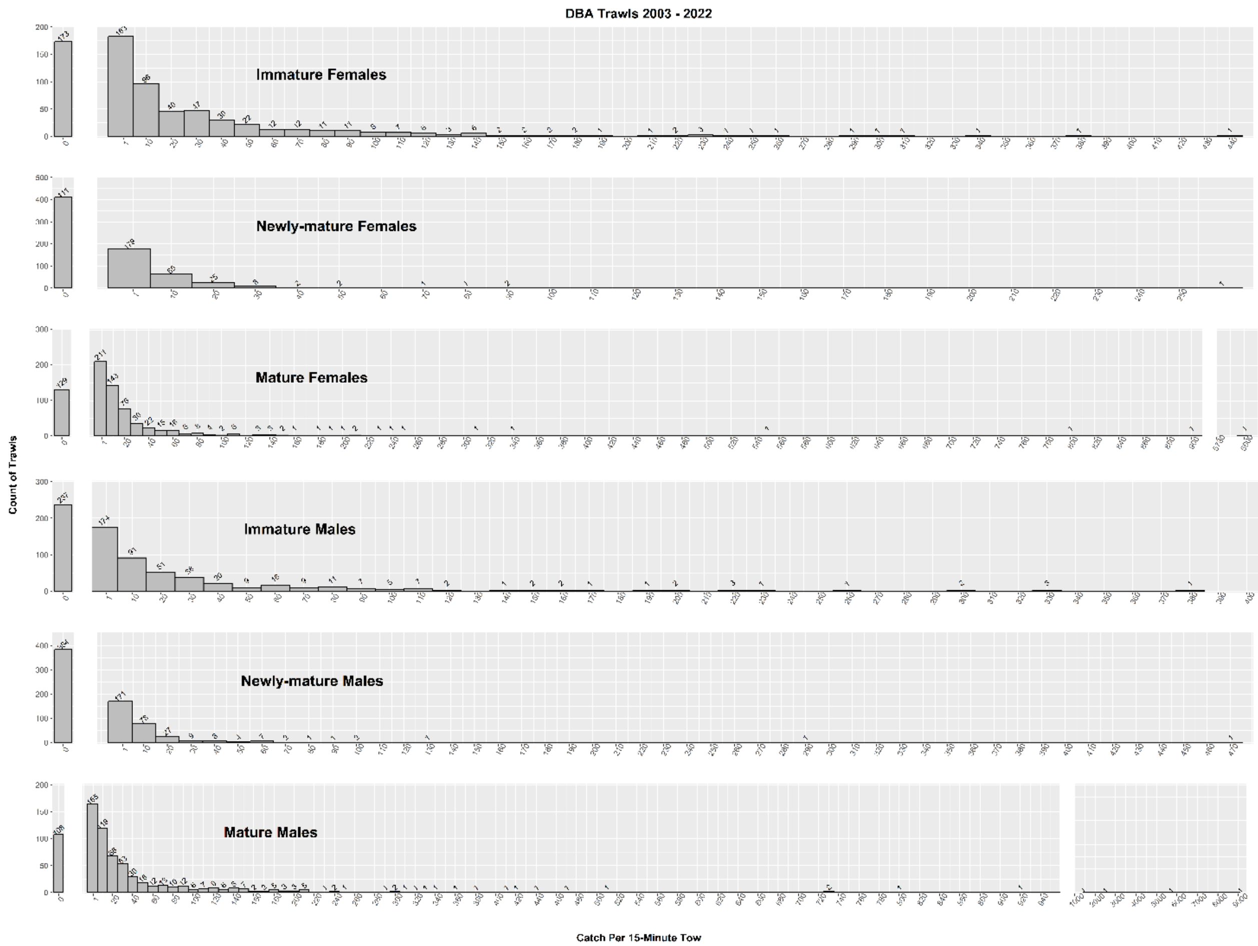


Figure 2: Counts of Delaware Bay trawls binned by CPUE per demographic group, with zeros separated on the left and positive counts on the right. Note the horizontal and vertical scale differences.

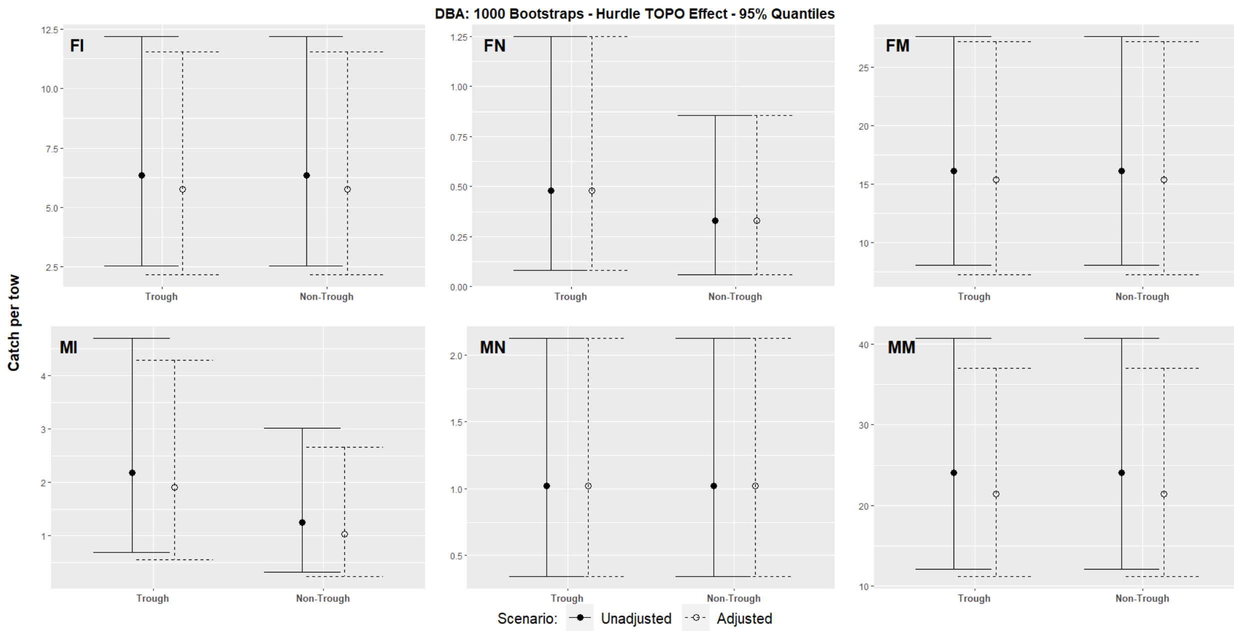


Figure 3: Relative effect of the topography categorical variable on catch per 15-minute tow of horseshoe crabs in the coastal Delaware Bay area shown by the mean and 95% quantile intervals. Solid circles and lines indicate the unadjusted bottom salinity scenario while open circles with dashed lines show the adjusted bottom salinity scenario.

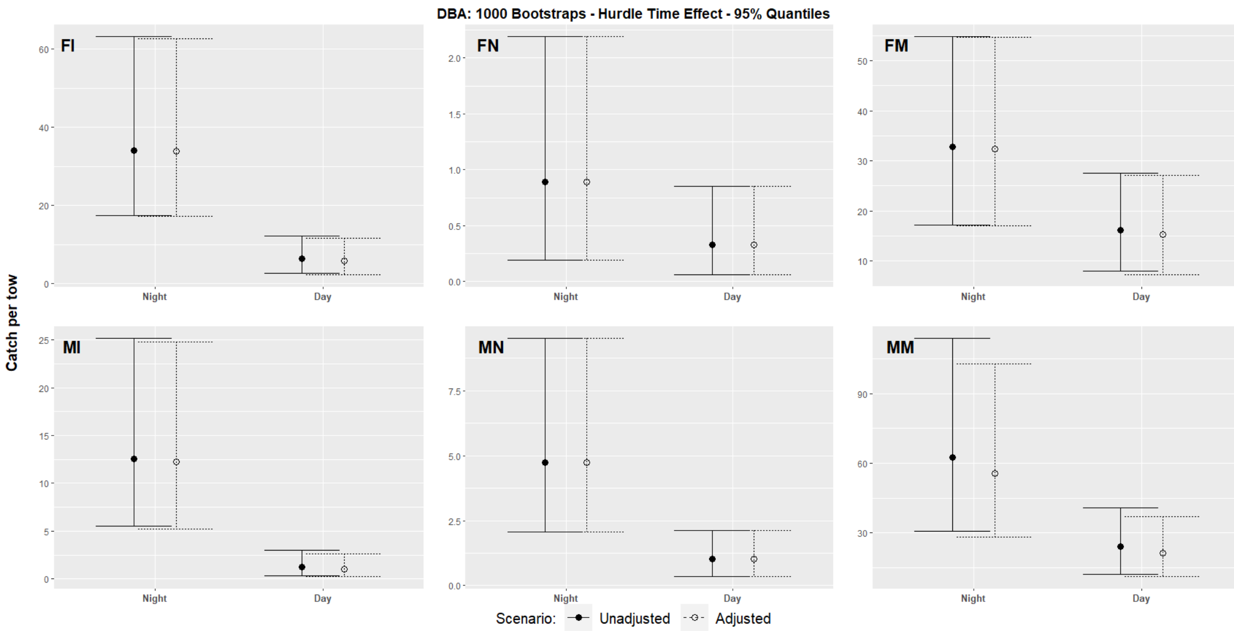


Figure 4: Relative effect of the time-of-day categorical variable on catch per 15-minute tow of horseshoe crabs in the coastal Delaware Bay area shown by the mean and 95% quantile intervals. Solid circles and lines indicated the unadjusted bottom salinity scenario while open circles with dashed lines show the adjusted bottom salinity scenario.

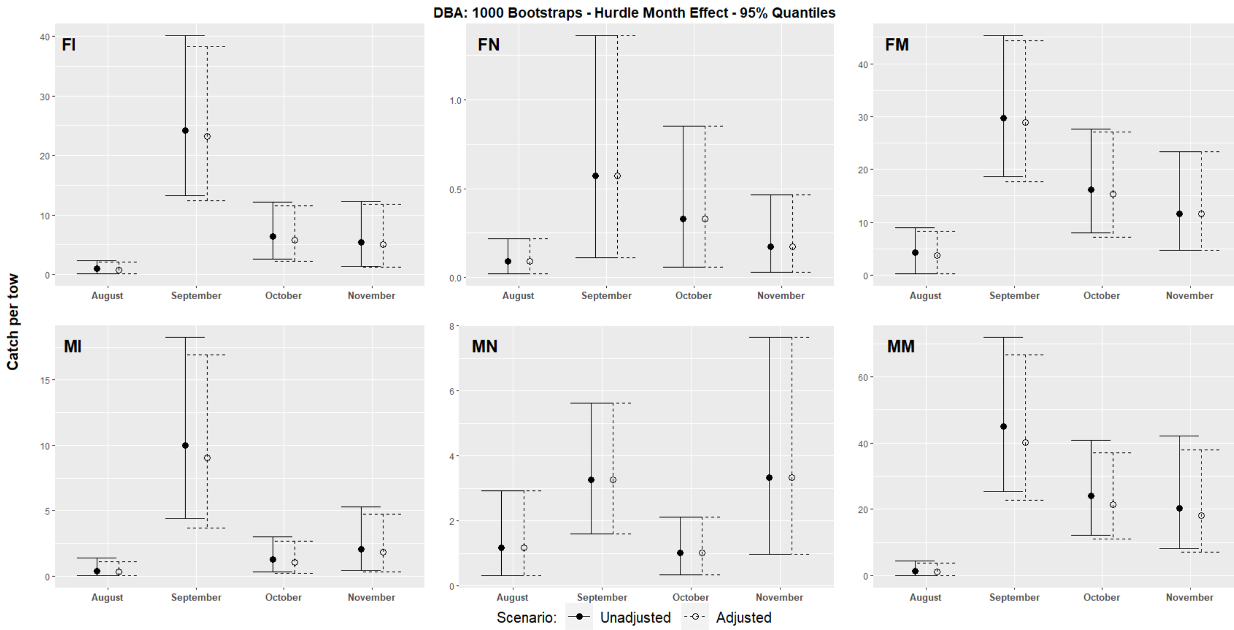


Figure 5: Relative effect of the month categorical variable on catch per 15-minute tow of horseshoe crabs in the coastal Delaware Bay area shown by the mean and 95% quantile intervals. Solid circles and lines indicated the unadjusted bottom salinity scenario while open circles with dashed lines show the adjusted bottom salinity scenario.

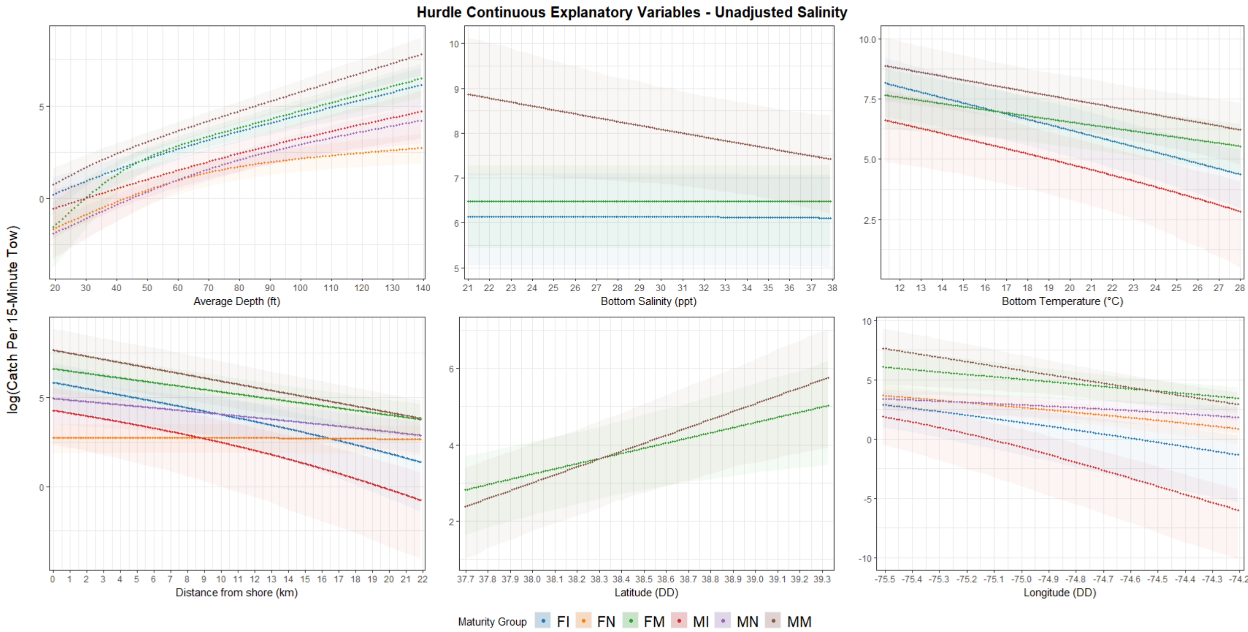


Figure 6: Individual explanatory variable effects analysis showing the mean value and 95% quantile intervals of 1000 iteration bootstrap analysis when changing the value of a continuous explanatory variable input in each demographic unadjusted model independently to see the results on the catch per 15-minute tow, shown in the log-scale for better visibility.

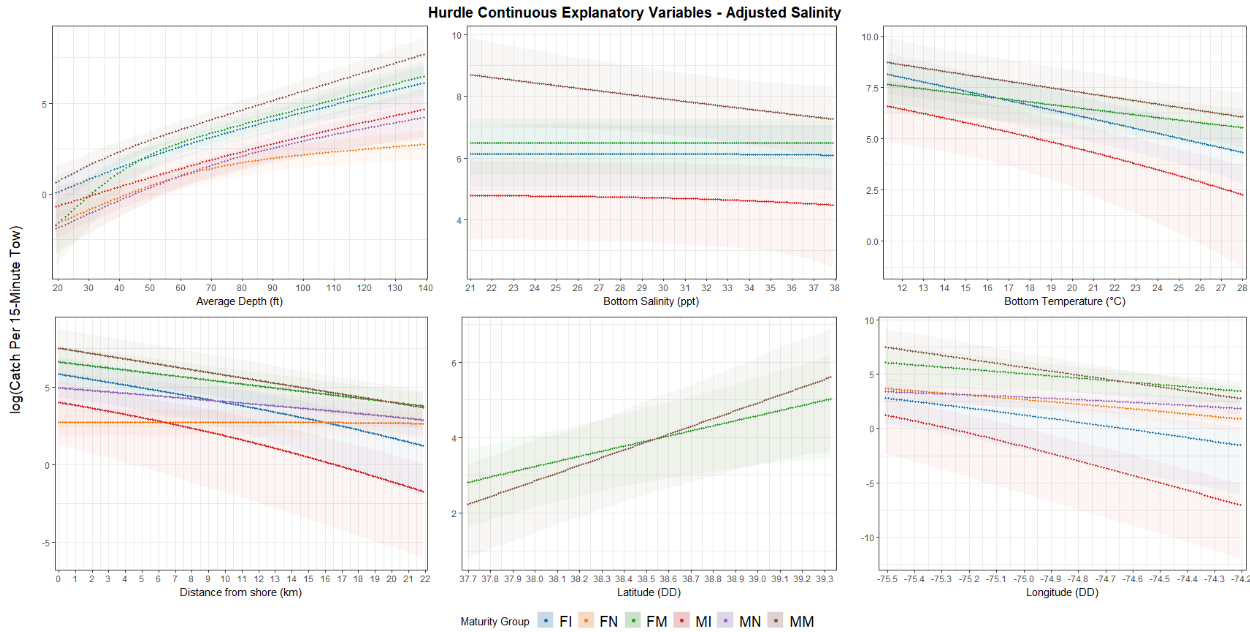
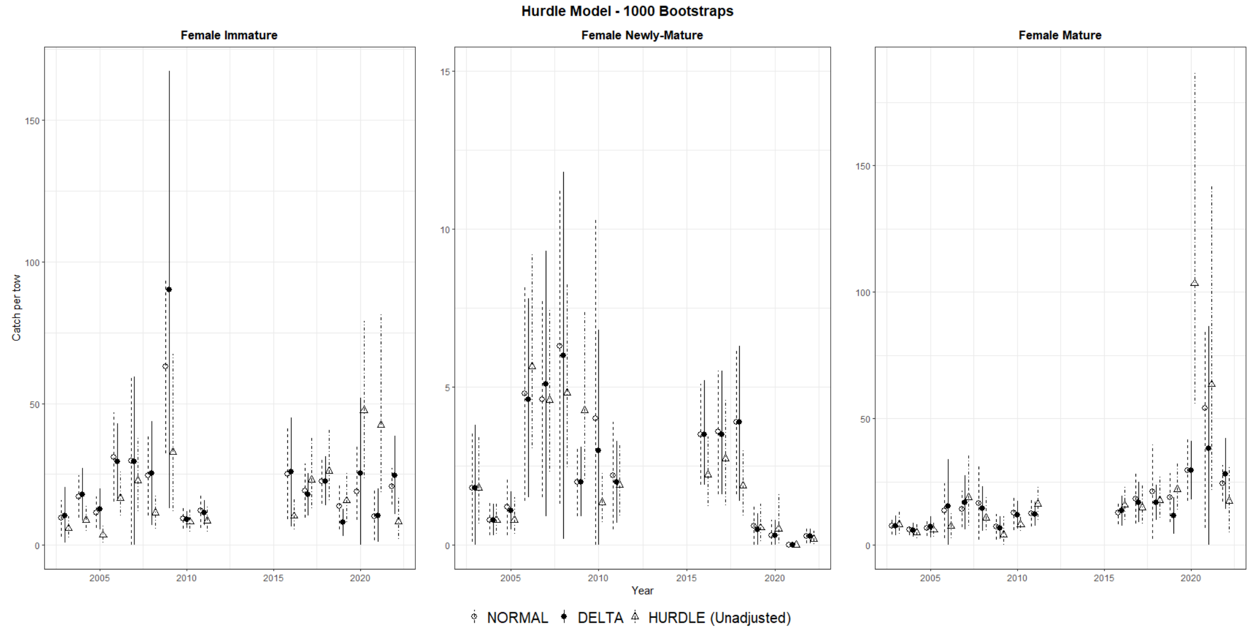


Figure 7: Mean value and 95% quantile intervals of 1000 iteration bootstrap analysis when changing the value of a continuous explanatory variable input in each demographic adjusted model independently to see the results on the catch per 15-minute tow, shown in the log-scale for better visibility.



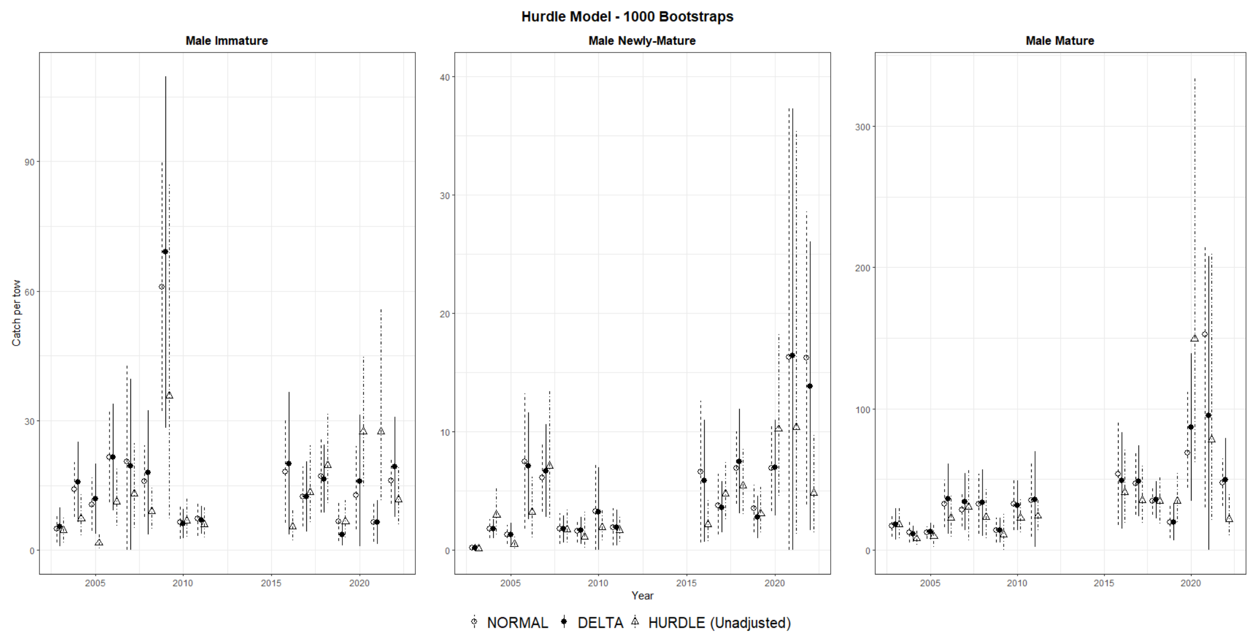


Figure 8: Estimated mean year effect for male horseshoe crabs and associated 95% confidence limits for design-based approach and 95% quantile intervals for the model-based approach. Open circles with dashed lines indicate the design-based method assuming a normal distribution, solid circles with solid lines assuming delta approach, and open triangles with dot dashed lines indicating the hurdle model approach.

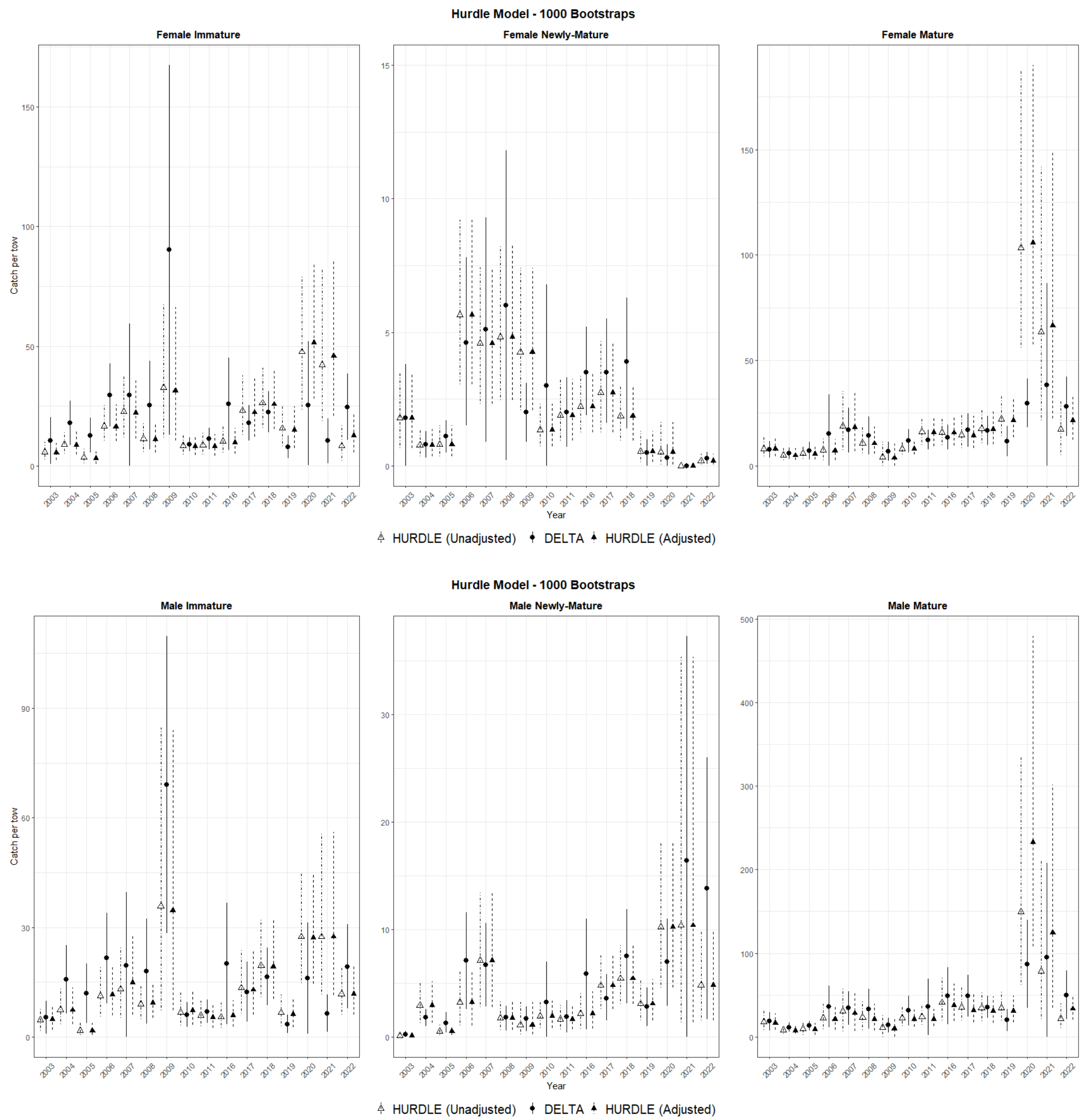


Figure 9: Comparing the unadjusted-salinity scenario and adjusted-salinity scenario model-based estimates of mean year effect and 95% quantile intervals of male horseshoe crabs with results of the design-based delta approach included for comparison.

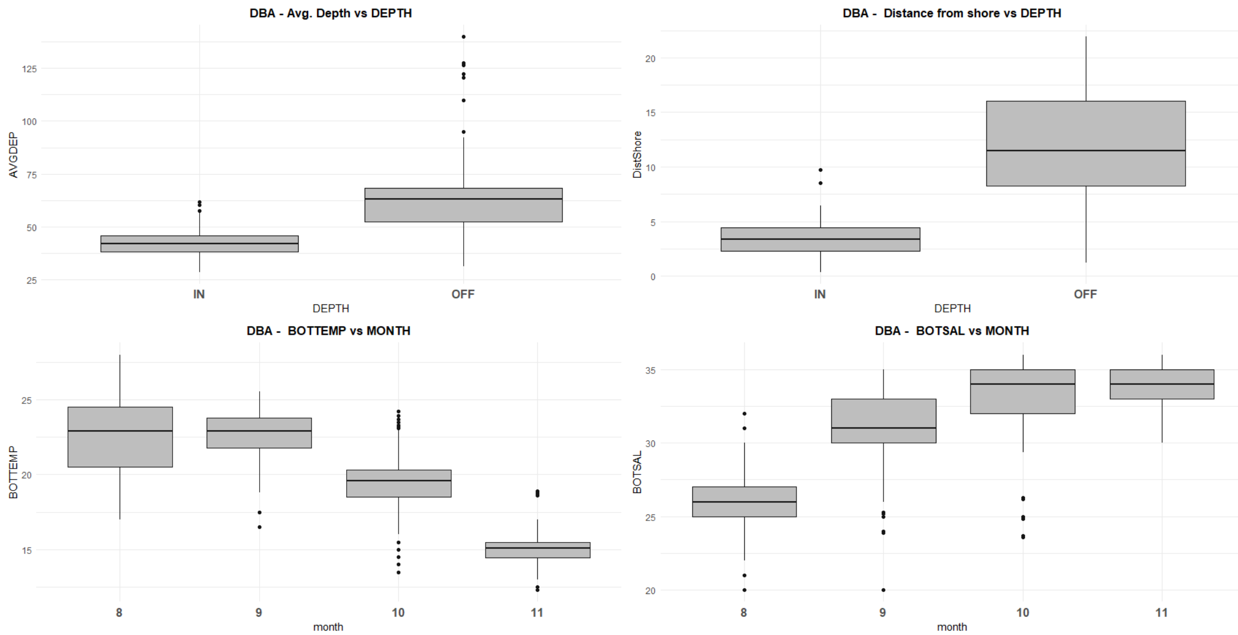


Figure 10a: Confounding variables found to occur over the study period within the Delaware Bay area, (DBA) showing first and third quartiles and mean within the box, and whiskers extending from the box representing 1.5 * IQR (interquartile range).

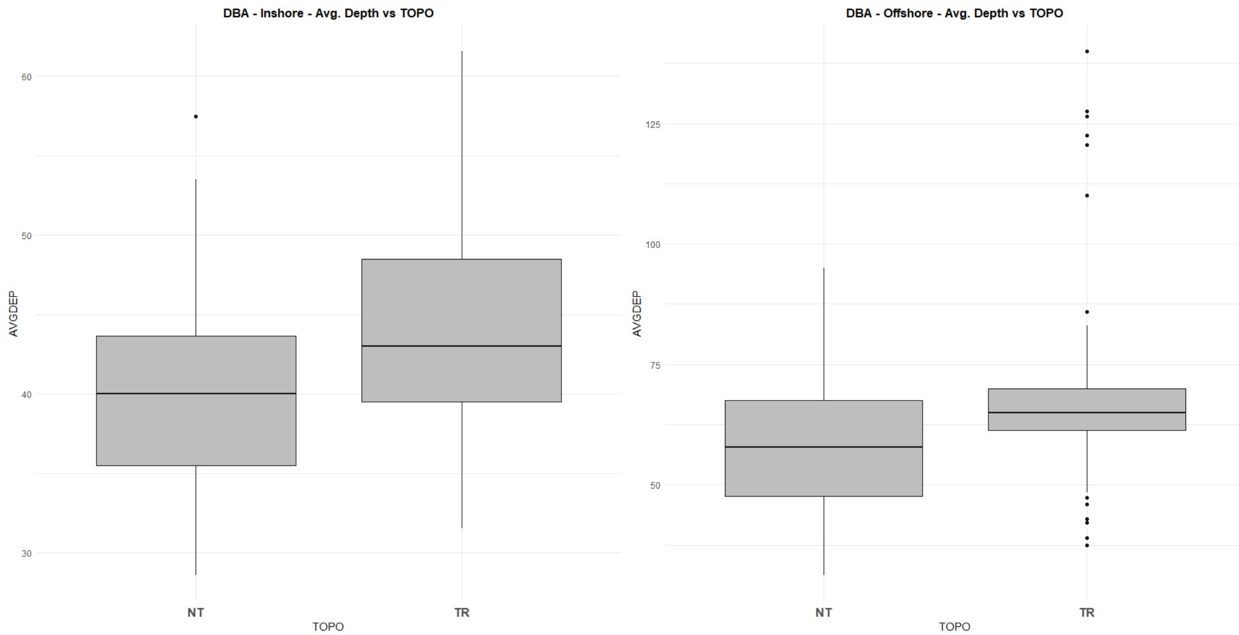


Figure 10b: Relationship between the topography categorical variable (non-trough/trough) and average depth (ft) within the Delaware Bay area (DBA) showing the first and third quartiles and mean within the box and the whiskers extending from the box representing 1.5 * IQR (interquartile range).

Chapter 2

Extracting relative abundance of channeled whelk, knobbed whelk, and summer flounder from inshore trawls within the Delaware Bay area

Chad C. Wong[1], Yan Jiao[1], and Eric Hallerman[1]

[1] Department of Fish and Wildlife Conservation, Virginia Tech, 310 West Campus Drive, Blacksburg, VA 24061, USA

Abstract

Channeled whelk (*Busysotypus canaliculatus*), knobbed whelk (*Busycon carica*), and summer flounder (*Paralichthys dentatus*) are commercially harvested and ecologically important species along the U.S. mid-Atlantic coast. However, very few stock assessments have been performed for channeled and knobbed whelk, with none along the mid-Atlantic coast. Summer flounder is managed and assessed as a single stock from North Carolina to the U.S.-Canada border. This study provides regional whelk population trends and adds to current knowledge on summer flounder abundance by using a hurdle model to analyze bycatch in a trawl survey for horseshoe crabs (*Limulus polyphemus*) in the Delaware Bay area from 2003-2022. We found that both species of whelks followed similar patterns in relative abundance, although channeled whelk are caught in smaller numbers on average. Since peaks of relative abundance in 2012, estimates of both species have been more consistent starting in 2016, although the relative abundance of channeled whelk has been at a historical low for the time series. Relative abundance for summer flounder was highly variable before 2016, with recent years exhibiting higher, more consistent catches than previous years. Summer flounder estimates of relative abundance were not similar to NOAA's coast-wide Age Structured Assessment Program estimates, but were similar to the NorthEast Area Monitoring and Assessment Program's spring survey Age Indices estimates between 2009 and 2012. These results suggest that there is high spatial heterogeneity in coastwide population trends of summer flounder over time.

Introduction

Channeled whelk (*Busysotypus canaliculatus*) and knobbed whelk (*Busycon carica)* are two gastropod species found along the Atlantic coast from Cape Cod, Massachusetts to central Florida (Angell 2018). Originally harvested to stop predation upon clams and oysters, whelk harvest has increased dramatically since the 1960s due to an increase in demand as a food item and the displacement of lobster fishermen due to the declining lobster stock (Leavitt and Burt 2000; Power et al. 2009; Fisher and Rudders 2017; Angell 2018). Although both species are harvested, channeled whelk has, and continues to be, the species of focus for commercial and recreational fisheries and for management (Fisher and Rudders 2017; Angell 2018). Whelk fisheries along the Atlantic coast, however, are only lightly regulated by management authorities, and biological characteristics of interest for management are still being researched. Because they are distributed within three miles of shore, channeled whelk are managed at the state level, with harvest restricted primarily by permitting, although numbers harvested and reporting requirements vary greatly among states along the East coast (Askin and Fisher 2021). In the mid-Atlantic region, New Jersey, Delaware, Maryland, and Virginia all have varying minimum legal size (MLS) regulations placed on channeled whelk. Some states also have seasonal closures, permit requirements, gear regulations, or other means of limiting harvest (Fisher and Rudders 2017; Askin and Fisher 2021). Among mid-Atlantic states, only Delaware has MLS regulations in place for knobbed whelk.

Estimated to live up to twenty years with slow growth and sexual dimorphism, both channeled and knobbed female whelks are estimated to reach maturity around 8-9 years, with channeled whelk males around 6-7 years, and knobbed whelk males around 5-6 years, although age-of-maturity varies based on location and among published studies (Peemoeller and Stevens

2013; Angell 2018; Nelson et al. 2018). Currently, there are no published stock assessments for either channeled or knobbed whelk from mid-Atlantic states. Massachusetts and Rhode Island have performed stock assessments on whelk fisheries, with the results of both assessments (Nelson et al. 2018; Angell 2020) suggesting overfishing in the respective areas, with the Massachusetts population being considered overfished. Angell (2020) found that channeled whelk abundance is highly correlated with fishing mortality rate based on a biomass dynamic model. While managed at the state level, results from a population genetic analysis of channeled whelk revealed seven distinct populations along the Atlantic coast (Askin 2020; Askin et al. 2022). Five regional management units were suggested based on these results, including one in the mid-Atlantic region represented by collections from Ocean City, MD, and Chincoteague, Eastern Shore, and Light Tower, VA, suggesting the need for interstate management.

Summer flounder (*Paralichthys dentatus*) is a highly-important commercial and recreational fish species, with combined landings of around 16.8 million pounds in 2019 (Ardini et al. 2022). Maturing generally by age two for both sexes and surviving up to seventeen years, sexual dimorphism is present, with females having a faster growth rate, and both sexes reaching similar maximum lengths (Ardini et al. 2022). The species is found along the entire Atlantic coast, but is most abundant in the Mid-Atlantic region, with spawning occurring in the fall and winter in the open ocean (Ardini et al. 2022). There is minimal genetic differentiation coastwide (Wirgin et al. 2022), suggesting that the entire Atlantic Coast contains a single genetic population. The National Oceanic and Atmospheric Administration (NOAA) manages and assesses summer flounder as one stock, ranging from the southern border of North Carolina to the US-Canada border (Ardini et al. 2022; MAFMC and NMFS 2022). From the initial establishment of state-based commercial and recreational fishing allocations in the mid-1990s, subsequent allocations changed for the first time in 2022 following changes in federal estimates resulting in a higher estimate of recreational catch than historical estimates. This amendment resulted in a shift in allocation of total acceptable biological catch (ABC) of 5% from commercial fishermen to recreational fishermen (55%:45%). These allocations previously were based on historical landings from the recreational and commercial sectors, but are now based upon catch. This is an important distinction as catch-based allocations include both landings and estimates of dead discards. For 2023, a 15.27 million-pound commercial quota and 10.61 million-pound recreational quota were set coastwide (MAFMC and NMFS 2022), with ABCs being divided among states including and north of North Carolina based on local abundance estimates. In addition to these harvest limits, there are also minimum and maximum size limits, seasonal closures, permit requirements, minimum mesh sizes for nets, and accountability measures to better manage the species. The Northeast Fisheries Science Center (NEFSC), as part of NOAA Fisheries, preforms stock assessments for all summer flounder ranging between North Carolina and the U.S.-Canada border. This Age Structured Assessment Program (ASAP) estimates spawning stock biomass (SSB), mature biomass, and estimated number of recruits of summer flounder, and are now based on annual commercial and recreational fishery catches, which includes landings and discards at age (Northeast Fisheries Science Center 2019). The most recent stock assessment by NOAA for summer flounder found that the population is not overfished, but overfishing is occurring, and recruitment has been below average since 2011 (ASMFC 2023; NEFSC 2023). The ASAP assessment, while based on catch, is calibrated using data from separate fishery-independent surveys, including the Virginia Institute of Marine Sciences (VIMS) NEAMAP survey along the coastal area of the mid-Atlantic (Figure 1).

NEAMAP is a multi-state, cooperative state and federal fishery-independent study implemented in 2006 to monitor populations of multiple species of fishes from Cape Hatteras, NC to Cape Cod, MA, including summer flounder (Chouinard et al. 2009; NEAMAP 2020). Performing two trawl surveys annually in spring and fall, the sample area is broken into 15 regions with a total of 150 stations. The study uses a stratified random design based on depth and 1.5 x 1.5 minute cells, with two stations per stratum and about one station per 30 nm^2 (Chouinard et al. 2009; Bonzek et al. 2009). The depths of these trawls vary depending upon the region being sampled, with a minimum sampling depth of 6.1 m (20 ft) and a maximum of 36.6 m (120 ft). Between Montauk, NY and Cape Hatteras, NC, sampling occurs inshore between the 6.1 m (20 ft) and 18.3 m (60 ft) depth contours. Trawls are conducted using a four-seam, three-bridle, 400 x 12 cm net with a cookie sweep and 2.54-cm knotless liner at the cod end with Thyboron Type IV 66" doors on a 90’ commercial trawler (VIMS 2023a). This data is then analyzed to produce fall and spring indices of abundance for all species of interest. Abundance indices are given in a single estimate for all age-class data combined, as well as separate estimates for individual age-classes, if aging is possible (VIMS 2023b).

The data set analyzed in this study is from the Virginia Polytechnic Institute and State University’s fall horseshoe crab (*Limulus polyphemus*) trawl survey. The study area ranges from Atlantic City, NJ to slightly north of Wachapreague, VA to sample horseshoe crabs (Figure 2). The survey area is stratified by distance from shore (0-3 nm, 3-12 nm) and bottom topography (trough, non-trough), performing 15-minute bottom trawls using a two-seam flounder trawl with an 18.3-m headrope and 24.4-m footrope, rigged with a Texas Sweep of 13-mm link chain and a tickler chain. While focusing on horseshoe crab, three bycatch species are recorded for occurrence and total length. These include knobbed whelk (*Busycon carica*), channeled whelk

(*Busycotypus canaliculatus)*, and summer flounder (*Paralichthys dentatus*). Sampling has occurred from 2003 to 2023, with a gap between 2013 through 2015 due to lack of funding. The goal of this study was to use a hurdle-model to estimate the relative abundance of the three bycatch species caught within the Virginia Tech survey between 2003 and 2022. These results provide some of the first estimates of relative abundance for channeled and knobbed whelk in the mid-Atlantic region. These results add to the growing knowledge of summer flounder and allow comparison of results from a small regional survey to broader coastal results from NEFSC ASAP and VIMS NEAMAP. The VIMS NEAMAP estimates reported here were extracted from the NEAMAP Abundance Index database (VIMS 2023d) and NEFSC ASAP estimates from NOAA's Stock Smart database (NOAA Fisheries 2023a).

Methods

The survey area (Figure 2) included two sampling areas, the Lower Delaware Bay (LDB) and the Delaware Bay Area (DBA). The LDB extended from the mouth of the bay to a line created between Egg Island Point, NJ and Kitts Hammock, DE. The DBA sampling area consisted of the Atlantic coast extending from 0-12 nautical miles out from shore and from the northern limit of Atlantic City, NJ (39˚ 20' N) south to Wachapreague, VA (37˚ 40' N). The area was stratified based on bottom topography (trough, non-trough) and distance from shore (0-3 nm, inshore; 3-12 nm, offshore). This stratification scheme was determined from the results of a 2001 pilot survey of the Delaware Bay area (Hata and Berkson 2004) and has been used annually since. Troughs were considered areas that are at least 2.4 m deep, no more than 1.8 km wide, and more than 1.8 km long, as determined from NOAA nautical charts. These charts were georeferenced into ArcGIS Pro (ERSI 2023) to create a digital map used to design the sampling scheme. The sample area was divided into cells of 1' of latitude by 1' of longitude using a World

Geodetic System (WGS) 1984 coordinate system. The DBA consisted of 1890 cells (~5127 km^2). The number of trawls performed in each stratum was based on both the proportional area that each stratum represents of the entire sampling area and the historical variation of CPUE observed among strata from previous years' sampling. Sampling occurred annually since 2002 in the DBA, with a gap in sampling between 2013 through 2015. For this study, we analyzed only the 697 trawls within the DBA (Table 1).

Trawls were performed using a two-seam flounder trawl with an 18.3-m headrope and 24.4-m footrope, rigged with a Texas Sweep of 13-mm link chain and a tickler chain. The net body consisted of 15.2-cm (6-in) stretched mesh, and the bag consisted of 14.3-cm (5 5/8-in) stretched mesh. Catchability has not been evaluated. Tows were aimed to have 15-minutes bottom time, but occasionally were shorter to avoid fishing gear (e.g., gill nets, crab and whelk pots) or vessel traffic. Sampling depths within the Delaware Bay area ranged from a minimum of 8.69 m (28.5 ft) to 42.7 m (140 ft) maximum, with an average of 16.06 meters (52.7 ft). Three species of bycatch were recorded when trawling for horseshoe crabs, knobbed whelk (*Busycon carica*), channeled whelk (*Busycotypus canaliculatus*), and summer flounder (*Paralichthys dentatus*). Counts and individual total lengths of these species were recorded. Along with this count data, environmental variables were also collected during the time of sampling. These variables included date, beginning and ending trawl depth, bottom temperature, bottom salinity, air and surface temperatures, and general weather conditions.

Hurdle Model

For the DBA, a hurdle model was used for its ability to handle excess zero observations in large datasets (Lo et al. 1992; Maunder and Punt 2004). A hurdle model assumes that the process that results in the high frequency of zeroes is not necessarily the same process that results in the positive data (Cragg 1971; Cameron and Trivedi 1998) (Formula 5). Preliminary analysis suggested that this could be the case for all three species of bycatch (Table 2). A hurdle model uses a two-stage approach that models the probability of occurrence using a binomial model and, for this study, a lognormal model to model positive observations (Maunder and Punt 2004). These processes are assumed to be independent, and the parameters of interest are estimated by separately solving for the maximum likelihood estimate of each modeling stage. A lognormal distribution was chosen for the second stage, as a histogram of CPUE over the time series shows a heavily right-skewed distribution (Figure 3). A uniform minimum variance unbiased (UMVU) estimate (Bradu and Mundlak 1970; Lo et al. 1992; Clifford et al. 2013) was used as the bias correction factor for all log-normal estimates.

Variable selection

A forward stepwise regression based on Akaike's (1973) information criterion (*AIC*) was performed to choose the best combination of explanatory variables for both stages of the hurdle model independently. Variables that were initially selected but were both non-significant (p-value <0.05) and did not decrease the *AIC* by at least 3 were removed, resulting in the final best-fitting models. The explanatory variables considered for all models were: year, latitude, longitude, depth (inshore/offshore), topography (trough/non-trough), time-of-day (day/night), month (August, September, October, November), bottom temperature, bottom salinity, and distance from shore. Latitude, longitude, and distance from shore were all based upon the midpoint of a given trawl.

Estimating year effect

The year effect can be extracted by multiplying the expected values from each modeling stage together after back-transformation and adding a bias correction factor to the lognormal modeling stage estimates (Formula 1). The mean value of the continuous variables and a weighted average of the categorical variables based on relative frequency were input into the models to estimate the year effect. A 1000-iteration bootstrap analysis with replacement was performed for all demographic groups to build quantile intervals around the year effect estimates in R using a custom script written by the authors. This process involved randomly sampling the data with replacement up to the total sample size, refitting the model, and estimating the year effect from this updated model. These quantile intervals were used to estimate the uncertainty around each year-effect estimate and to determine the significance of the change of relative abundance compared to previous years. To determine the individual effect of each explanatory variable on CPUE, we ran the 1000-iteration bootstrap from above results. During each iteration, we calculated the relative abundance of all demographic groups while changing one explanatory variable's input value and keeping all others at their respective means for continuous variables, and at their weighted means for categorical variables. For continuous variables, we tested all values between the observed minimum and maximum of each variable over the time series, dividing each range to have approximately 150 values. For testing categorical variables, we set all categories within the selected explanatory variable to zero except the category of interest, which was set to one.

Results

Variable selection and individual variable effects analysis

Significant explanatory variables that were selected varied depending on the hurdle modeling stage and species being modeled (Table 3). The best-supported model for explaining

CPUE variation of knobbed whelk estimated negative coefficients for distance from shore and longitude in both modeling stages, and average depth in the lognormal stage. Latitude had a positive estimated coefficient in both stages, and bottom temperature only in the log-normal stage of the model. Topography was the only categorical variable selected and was estimated to have non-significantly higher mean catches of knobbed whelk in troughs than in non-troughs (Figure 4). While estimated CPUE was not highly sensitive to latitude, all other continuous spatial variables (longitude, average depth, and distance from shore) had strong effects (Figure 5).

The best-supported channeled whelk model selected two continuous variables, distance from shore in the log-normal stage and longitude in the binomial stage (Table 3). Both variables were estimated to be negatively related to CPUE (Figure 5). The hurdle model for channeled whelk selected the most categorical variables between the two whelk species, with month, time-of-day, and topography all selected. Within each of these categorical variables, none of the potential categories resulted in significantly different effects on CPUE than others (Figure 4). That said, trawls at night, in troughs, and in months other than August were estimated to give rise to higher catches than the other possible categories for each variable.

Distance from shore, longitude, and topography were selected as significant explanatory variables in both the channeled and knobbed whelk models in either stage. Each had the same positive or negative correlation with CPUE for each species' models, but at varying levels of effect. Increasing longitude and distance from shore decreased estimated catch for both species, although knobbed whelk CPUE declined at a faster rate with either variable when compared to channeled whelk CPUE. Longitude and distance from shore were not strongly correlated (0.256) (Figure 6). Differences in topography did not significantly affect the estimated CPUE of either species, with troughs resulting in a higher mean CPUE relative to non-troughs for the two species.

Summer flounder had positive coefficient estimates for average depth in both modeling stages, and positive estimates for bottom temperature and latitude in the binomial stage. The only negative parameter coefficient estimate was distance from shore, which was selected in the lognormal stage. From the individual explanatory variable effects analysis, we found that the effects of bottom temperature and latitude on CPUE, while present, were minor (Figure 5). Average depth and distance from shore had much stronger effects on CPUE. Despite these two variables being positively correlated (0.694) (Figure 6), they had opposite effects on CPUE, suggesting that deeper, inshore trawls resulted in higher catches of summer flounder. The only categorical variable selected was time (day/night), with higher mean CPUEs estimated during the day, although not significantly different from those at night (Figure 4).

Year effect

Estimated year effects varied over the study period for all three species of bycatch, although they have been more consistent since 2016 for all species (Figure 7). Knobbed and channeled whelk have exhibited similar patterns of relative abundance since 2003, although estimates for channeled whelk were more variable in years prior to 2012. Both species have had consistent estimates of relative abundance since 2016. While recent estimates of knobbed whelk are similar to those seen earlier in the study, the relative abundance of channeled whelk has been at a historical low since 2016, with an estimated low for the time series in 2018. Knobbed whelk has had significantly higher estimates of relative abundance on average than channeled whelk (Figure 8). The relative abundance of summer flounder has been highly variable over the study period, with no apparent trend. No summer flounder were caught in 2004, 2007, and 2008, resulting in a zero estimate for year effect. In recent years, summer flounder relative abundance has been consistently high and at higher levels than seen in earlier years, with an estimated high for the time series in 2021, although not significantly higher than the previous three years.

NEAMAP & NOAA

VIMS NEAMAP's summer flounder indices of abundance are published both cumulatively for all ages, as well as for individual age-classes for spring and fall surveys (VIMS 2023d). Compared to NEAMAP's all-ages indices for the spring and fall, our hurdle model estimates of summer flounder relative abundance were similar to NEAMAP's indices between 2009 and 2012 for the spring survey estimates (Figure 9, 10). Among NEAMAP's induvial age estimates, no age group in the spring or fall had consistently similar trends of relative abundance when compared to our hurdle model estimates. However, when looking at smaller periods of time in NEAMAP's spring estimates, there were some similarities. Hurdle model estimates between 2009 and 2012 were similar to NEAMAP's spring estimates for ages one, two, three, and seven and older individuals. That sharp, small, downward trend in the spring also can be seen in earlier years for other age-classes, starting with age four in 2008, age five in 2009, and age six in 2010. Otherwise, no other similarities in trends of relative abundance can be found among NEAMAP's spring or fall estimates. Although sampling for the horseshoe crab trawl survey occurs in the fall (August-November), all of NEAMAP's fall age-class estimates are more negatively correlated with our hurdle model estimates than any spring age-class estimates, other than age-4 individuals.

Compared to NOAA's ASAP assessment estimates, there were no similarities in trends of relative abundance between the hurdle model estimates and NOAA's estimated mature stock biomass, spawning stock biomass, or number of recruits besides the slight positive trend seen in spawning stock biomass between 2017 and 2019. Looking at the overall correlation between the hurdle and NOAA's estimates, there were only negative correlations between our hurdle model estimates and NOAA's estimates.

Discussion

Channeled and knobbed whelk are ecologically important species that are subject to both targeted fisheries and harvest as bycatch in trawls for other species in the mid-Atlantic region. Commercial whelk fisheries have grown dramatically since their establishment int the 1960s, with whole, live whelks being worth three times as much in the 2010s compared to the 1990s (Fisher 2015). Despite this value, there is little information in the literature regarding whelk biology, behavior, and abundance in the mid-Atlantic region (Askin et al. 2022). While knobbed and channeled whelk belong to the same family of sea snails, *Busyconidae* (Olsson 1967), and inhabit similar distributional ranges along the eastern United States from central Florida to Cape Cod, MA (Angell 2018), our models suggested that only a few spatial and environmental variables affect the expected CPUE of both species, and not equally. These shared significant explanatory variables include distance from shore, longitude, and topography. Troughs resulted in a higher mean catch of both species than non-trough trawls, but not significantly so. These results suggest that trawls in troughs, closer to shore, and at lower longitudes would have higher catches of both species than those not in troughs, offshore, and further north. While little is known about habitat selection, knobbed whelk fisheries usually are conducted in inshore, estuarine areas in and around Virginia, while channeled whelk fisheries are conducted both inshore and offshore, though predominantly offshore for states in the mid-Atlantic (Fisher and Rudders 2017; Askin and Fisher 2021). These fishery locations aligned with observations of the spatial distribution of trawls with whelk within the study period (Figure 11). Although increasing distance from shore resulted in smaller catches of both species, many trawls close to the 12-nm outer boundary of the study area only caught channeled whelk. Among these offshore trawls, those that caught both species generally had higher numbers of channeled whelk, especially south of Delaware Bay. Conversely, trawls near shore that caught both species almost always caught more knobbed whelk than channeled whelk. Both species were frequently caught together nearshore. For knobbed whelk, average depth was deemed a significant explanatory variable, with shallow depths being associated with higher CPUEs. Although distance from shore and average depth exhibited only a low correlation overall within the Delaware Bay area (0.694), the findings support the view that knobbed whelks are more abundant in inshore areas and help explain why knobbed whelk fisheries are conducted inshore. Other factors such as latitude and bottom temperature also affected the CPUE of knobbed whelk. This result suggests that trawls further north in areas or times with warmer water result in larger catches of knobbed whelk. Few temporal explanatory variables were deemed significant among the whelk models, with only month and time-of-day found to influence channeled whelk CPUE. While the month variable suggests a seasonality aspect, only the month of August resulted in a non-significant lower mean, while CPUEs for September through November were nearly identical, limiting the seasonality effect.

Relative abundance of both whelk species exhibited similar patterns over the time series, although channeled whelk relative abundance was more variable prior to 2012. In recent years (2016-2022), estimates of knobbed whelk relative abundance have been relatively constant and at similar levels to those seen before 2012. Channeled whelk estimates have also been relatively consistent, but at much lower levels than those estimated for earlier years in the study, with an all-time low estimated in 2018.

Currently, there are no whelk stock assessments in the mid-Atlantic region with which to compare our results. Worldwide, whelk fisheries have followed boom-and-bust cycles (Askin et al. 2022), with knobbed whelk in South Carolina, for example, exhibiting such a pattern in the United States (Power et al. 2009). Our results, however, did not show a boom-and-bust cycle in the study area, which could be a result of the limited time of the study, no sampling occurring between 2012-2016, not targeting the species directly, gear effectiveness, or the area sampled. This in mind, when comparing this study's estimates of knobbed whelk relative abundance to combined mid-Atlantic states landings, both had large peaks between 2010 and 2012, with recent years' estimates being similar to those seen prior to said peaks. Due to sampling not occurring between 2012 – 2016, it cannot be known for certain if this was a true "boom" in a boom-bust-cycle, especially with landings not being standardized for effort and other external factors (e.g., occurrence of tropical storms). However, a similar peak in relative abundance was observed in channeled whelk between 2010-2012 in both commercial landings and hurdle model estimates of relative abundance, with recent years having some of the lowest estimates for the entire time series. With the high gene flow among channeled whelk populations from Virginia Beach, VA to the southern border of Delaware in Ocean City, MD (Askin et al. 2022), the region at large may have gone through a boom-and-bust cycle, but we cannot be certain.

Although the hurdle model for summer flounder CPUE selected bottom temperature and latitude as explanatory variables, these factors had minimal effects on CPUE. Similarly, time-of-day was selected as an explanatory variable, but there were non-significant differences in CPUE between trawls performed in the day and at night. The two variables that most affected CPUE for summer flounder were average depth and distance from shore. An increase in average depth led to an increased estimated CPUE, while distance from shore was negatively related to CPUE.

This result was unexpected, as within the Mid-Atlantic Bight, summer flounder overwinter offshore on the continental shelf starting in July/August and return inshore during the spring, although this behavior varies with latitude (Sackett et al. 2007). Our data and model did not reflect seasonal migration in the month and distance from shore variables, but possibly so in the average depth variable. While at a large spatial scale, depth increases with distance from shore, this relationship was not strong in our study area. In the DBA sampling area, the average depth ranged from a minimum of 8.7 to 42.7 m, with distance from shore ranging from 0.36 to 2.0 km. The edges of the continental shelf transform into the continental slope in the Mid-Atlantic Bight at 100 to 200 km offshore, at depths of around 100 to 200 m (Stevenson et al. 2004). Due to our study area extending to less than one-quarter of the distance to the edge of the continental shelf, we may not have observed the overarching trends reported at a broader spatial scale.

Estimated relative abundance for summer flounder was highly variable prior to 2012, with more consistent estimates occurring since 2016. There were no similar trends between this study's estimates of relative abundance and NOAA's ASAP estimates of summer flounder (Northeast Fisheries Science Center 2019) for any life-stage. This result might be expected due to the difference in the respective spatial coverage, as NOAA's ASAP estimates use data from North Carolina to the US-Canada border and thereby estimate general trends along the broader northeastern Atlantic coast. The reason why this study's estimates, which occur between August through November, match a small time-period in NEAMAP's spring survey estimates and not in their fall survey estimates is unclear. Initially, the difference was thought to be caused by differences in sampling depths between the studies. However, trawls performed in this study were at similar depths to NEAMAP's sampling depth stratification below Montauk, NY (Figure 1), with a mean of 16.1 m, with a 95% quantile between 9.7 and 23.9 m (Figure 12). Likely, the major explanation for the lack of similarity is that NEAMAP estimates use data from a much larger study area, possibly obscuring any regional variation in population trends that could have been similar to this study's estimates for the mid-Atlantic/Delaware Bay.

From this study, we have estimated regional population trends of knobbed whelk, channeled whelk, and summer flounder in the mid-Atlantic region. This data can be used as a starting point for future whelk stock assessments in the area and to help better understand the asynchronous spatial-temporal dynamics of summer flounder in near shore regions of the mid-Atlantic compared to overarching trends estimated along the broader Atlantic coast from surveys like VIMS' NEAMAP.

Literature Cited

Aitchison, J. 1955. On the Distribution of a Positive Random Variable Having a Discrete Probability Mass at the Origin. Journal of the American Statistical Association 50(271):901–908.

Angell, T. E. 2018. Age, Growth, and Sexual Maturity of the Channeled Whelk *Busycotypus canaliculatus* (Linnaeus, 1758) and Knobbed Whelk *Busycon carica* (Gmelin, 1791) in Narragansett Bay, Rhode Island. Journal of Shellfish Research 37(1):207–219.

Angell, T. E. 2020. 2006-2019 Catch, Effort, and Fishery Trends in the Rhode Island Whelk Fishery and Recent Stock Status. Rhode Island Department of Environmental Management - Division of Marine Fisheries, Jamestown, RI.

Ardini, G., J. Beaty, D. C. Leaning, K. Coutre, K. Dancy, M. Ferguson, E. Keiley, S. Lewis, G. Shepherd, and M. Terceiro. 2022. Amendment 22 to the Interstate Fishery Management Plans for Summer Flounder, Scup, and Black Sea Bass: Commercial and Recreational Allocation. Atlantic States Marine Fisheries Commission, Arlington, VA.

Askin, S. E. 2020. An Assessment of the Population Genetic Structure and Evidence of Multiple Paternity in Channeled Whelk, *Busycotypus canaliculatus*, Along the U.S. Atlantic Coast. M.S. Thesis, College of William & Mary, Williamsburg, Virginia.

Askin, S. E., and R. A. Fisher. 2021. A Summary of Collaborative Discussions on Existing Management within the Channeled Whelk Fishery. Page 14. Virginia Institute of Marine Science, Gloucester Point, VA.

Askin, S. E., R. A. Fisher, E. E. Biesack, R. Robins, and J. R. McDowell. 2022. Population Genetic Structure in Channeled Whelk *Busycotypus canaliculatus* along the U.S. Atlantic Coast. Transactions of the American Fisheries Society 151(5):543–558.

ASMFC (Atlantic States Marine Fisheries Commission). 2023. Program Overview - Summer Flounder. https://www.asmfc.org/species/summer-flounder.

Bonzek, C. F., J. Gartland, J. D. Lange, and R. J. Latour. 2009. Northeast Area Monitoring and Assessment Program (NEAMAP) - Data Collection and Analysis in Support of Single and Multispecies Stock Assessments in the Mid-Atlantic: Northeast Area Monitoring and Assessment Program Near Shore Trawl Survey. Virginia Institute of Marine Science, Gloucester Point, Virginia.

Bradu, D., and Y. Mundlak. 1970. Estimation in Lognormal Linear Models. Journal of the American Statistical Association 65(329):198–211.

Cameron, A. C., and P. K. Trivedi. 1998. Regression Analysis of Count Data. Cambridge University Press, Cambridge, MA.

Chouinard, G., D. Beutel, and M. Wilkins. 2009. Atlantic State Marine Fisheries Commission - Terms of Reference and Advisory Report of the NEAMAP Nearshore Trawl Survey Peer Review. Atlantic States Marine Fisheries Commission, 09–01, Virginia Beach, VA.

Clifford, D., N. Cressie, J. R. England, S. H. Roxburgh, and K. I. Paul. 2013. Correction Factors for Unbiased, Efficient Estimation and Prediction of Biomass from Log-Log Allometric Models. Forest Ecology and Management 310:375–381.

Cragg, J. G. 1971. Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods. Econometrica 39(5):829.

ERSI. 2023. ArcGIS Pro Version 3.1.0. Environmental Systems Research Institute, Inc., Redlands, CA.

Fisher, R. A. 2015. Coherent Approach to *Busycon/Busycotypus* Fishery Management Along the US Atlantic Seaboard - Proceedings from the 16th International Conference on Shellfish Restoration (ICSR), December 10-13, 2014, Charleston S.C. Virginia Institute of Marine Science, Gloucester Point, VA.

Fisher, R. A., and D. B. Rudders. 2017. Population and Reproductive Biology of the Channeled Whelk, *Busycotypus canaliculatus* , in the US Mid-Atlantic. Journal of Shellfish Research 36(2):427–444.

Henderson, M. J., and M. C. Fabrizio. 2013. Detecting Noncompliance in the Summer Flounder Recreational Fishery Using a Mark–Recapture Growth Model. North American Journal of Fisheries Management 33(5):1039–1048.

Kohavi, R. 1995. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. Pages 1137–1143 Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2. Morgan Kaufmann Publishers, Inc., San Francisco, CA, USA.

Leavitt, Dale F. and Burt, William P. 2000. Control of Predators on Cultured Shellfish: Exclusion Strategies. University of Massachusetts Dartmouth, 00–007, North Dartmouth, MA.

Lo, N. C., L. D. Jacobson, and J. L. Squire. 1992. Indices of Relative Abundance from Fish Spotter Data based on Delta-Lognormal Models. Canadian Journal of Fisheries and Aquatic Sciences 49(12):2515–2526.

Lohr, S. L. 2019. Sampling: Design and Analysis. CRC Press, Boca Raton, FL.

MAFMC and NMFS (Mid-Atlantic Marine Fisheries Commission and National Marine Fisheries Service). 2022. Summer Flounder, Scup, and Black Sea Bass 2023 Specifications - Supplemental

Information Report (SIR) and Fishery Specifications Document. National Marine Fisheries Service, Gloucester, MA.

Mark Gibson. 2017. Stock Assessment of Whelk in Rhode Island and Recommendations for Research and Management. Rhode Island Department of Environmental Management - Division of Marine Fisheries, Jamestown, RI.

Maunder, M. N., and A. E. Punt. 2004. Standardizing catch and effort data: a review of recent approaches. Fisheries Research 70(2–3):141–159.

Millero, F. J., and A. Poisson. 1981. International one-atmosphere equation of state of seawater. Deep Sea Research Part A. Oceanographic Research Papers 28(6):625–629.

Morse, W. W. 1981. Reproduction of the Summer Flounder, *Paralichthys dentatus* (L.). Journal of Fish Biology 19(2):189–203.

Morson, J. M., E. A. Bochenek, E. N. Powell, and J. E. Gius. 2012. Sex at Length of Summer Flounder Landed in the New Jersey Recreational Party Boat Fishery. North American Journal of Fisheries Management 32(6):1201–1210.

NEFSC. 2023. Summer flounder - 2023 Management Track Assessment Report. U.S. Department of Commerce, National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, Massachusetts.

Nelson, G. A., S. H. Wilcox, R. P. Glenn, and T. L. Pugh. 2018. Massachusetts Division of Marine Fisheries Technical Report TR-66 - A Stock Assessment of Channeled Whelk (Busycotypus canaliculatus) in Nantucket Sound, Massachusetts. Massachusetts Division of Marine Fisheries, Gloucester, MA.

NEAMAP. 2020. Northeast Area Monitoring and Assessment Program Management Plan 2012-2016. University of Maine Digital Commons 25.

NOAA Fisheries. 2023a. Stock SMART data records. NOAA Fisheries Office of Science and Technology. Available: https://apps-st.fisheries.noaa.gov/stocksmart.

NOAA Fisheries. 2023b. Fisheries One Stop Shop - Landing Query. NOAA Fisheries Office of Science and Technology. Available: www.fisheries.noaa.gov/foss.

NOAA NCEI. 2022. ETOPO 2022 15 Arc-Second Global Relief Model. National Oceanic and Atmospheric National Centers for Environmental Information. Available: https://doi.org/10.25921/fd45-gt74.

Northeast Fisheries Science Center. 2019. 66th Northeast Regional Stock Assessment Workshop (66th SAW) Assessment Summary Report. Woods Hole, MA.

O'Leary, C. A., C. C. Stawitz, and J. A. Nye. 2019. Detecting Somatic Growth Trends for Summer Flounder (*Paralichthys dentatus*) Using a State-Space Approach. Canadian Journal of Fisheries and Aquatic Sciences 77(5):917-930.

Olsson, A. 1967. Some Tertiary Mollusks from South Florida and the Caribbean. Ithaca. Paleontological Research Institution, Ithaca, New York.

Packer, D. B., Griesbach, S. J., Berrien, P. L., Zetlin, C. A., Johnson, D. L., and Morse, W. W. 1999. Essential Fish Habitat Source Document: Summer Flounder, *Paralichthys dentatus*, Life History and Habitat Characteristics. Diane Publishing, Woods Hole, MA.

Peemoeller, B.-J., and B. G. Stevens. 2013. Age, Size, and Sexual Maturity of Channeled Whelk (*Busycotypus canaliculatus*) in Buzzards Bay, Massachusetts. Fishery Bulletin 111(3):265–278.

Pennington, M. 1983. Efficient Estimators of Abundance, for Fish and Plankton Surveys. Biometrics 39(1):281–286.

Power, A. J., C. J. Sellers, and R. L. Walker. 2009. Growth and Sexual Maturity of the Knobbed Whelk, *Busycon carica* (Gmelin, 1791), From a Commercially Harvested Population in Coastal Georgia. Occasional Papers of the University of Georgia Marine Extension Service 4.

Sackett, D. K., K. W. Able, and T. M. Grothues. 2007. Dynamics of Summer Flounder, *Paralichthys dentatus*, Seasonal Migrations Based on Ultrasonic Telemetry. Estuarine, Coastal and Shelf Science 74(1–2):119–130.

Smith, R. W., and F. C. Daiber. 1977. Biology of the Summer Flounder, *Paralichthys denatatus*, in Delaware Bay. Fishery Bulletin 75(4):823–830.

Stevenson, D., L. Chiarella, D. Stephan, R. Reid, K. Wilhelm, J. McCarthy, and M. Pentony. 2004. Characterization of the Fishing Practices and Marine Benthic Ecosystems of the Northeast U.S. Shelf, and an Evaluation of the Potential Effects of Fishing on Essential Fish Habitat. NOAA Technical Memorandum NMFS-NE-181, U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA.

Tigano, A., and V. L. Friesen. 2016. Genomics of local adaptation with gene flow. Molecular Ecology 25(10):2144–2164.

VIMS. 2023a. Virginia Institute of Marine Science: NEAMAP Gear. https://www.vims.edu/research/units/programs/multispecies_fisheries_research/neamap/gear/index.php.

VIMS. 2023b. Virginia Institute of Marine Science: Abundance Indices - Methodology. https://www.vims.edu/research/departments/fisheries/programs2/multispecies_fisheries_research/interaction/abundance_indices/index.php.

VIMS. 2023c. Virginia Institute of Marine Science: What is NEAMAP? https://www.vims.edu/research/units/programs/multispecies_fisheries_research/neamap/index.php.

VIMS. 2023d, May. Multispecies Research Group NEAMAP Abundance Indices. Virginia Institute of Marine Science. Available: https://fluke.vims.edu/mrg/neamap_abundance/summerflounder.htm.

Wessel, P., and W. H. F. Smith. 1996. A Global, Self-Consistent, Hierarchical, High-Resolution Shoreline Database. Journal of Geophysical Research: Solid Earth 101(B4):8741–8743.

Wirgin, I., L. Maceda, J. Stabile, and J. Waldman. 2022. Genetic Population Structure of Summer Flounder *Paralichthys dentatus* Using Microsatellite DNA Analysis. Fisheries Research 250:106270.

Table 1: Number of trawls performed per year in each stratum within the Delaware Bay area (DBA), for a total of 697 trawls.

	Stratum				
Year	IN/TR	IN/NT	OFF/TR	OFF/NT	
2003	12	10	5	13	
2004	9	7	5	18	
2005	12	7	8	12	
2006	8	5	5	16	
2007	14	6	10	15	
2008	16	8	8	13	
2009	11	8	7	15	
2010	14	9	12	18	
2011	12	9	11	18	
2012	6	6	0	0	
2016	11	7	9	17	
2017	11	9	5	17	
2018	6	7	13	15	
2019	8	8	8	21	
2020	10	10	15	9	
2021	15	9	12	7	
2022	1	6	7	26	
Total:	176	131	140	250	**=697**

Table 2: Percentage of trawls performed during the study period (2003 – 2022) that captured zero or at least one of the four bycatch species: summer flounder (SFL), channeled whelk (CW), knobbed whelk (KW), and lightning whelk (LW).

DBA	% ZERO	% POSITIVE
SFL	34.3%	65.7%
CW	31.9%	68.1%
KW	18.4%	81.6%
LW	97.0%	03.0%
Total Trawls: 697		

Table 3: Explanatory variables selected in the binomial (B) and log-normal (LN) stages of the hurdle models for each species of bycatch from a forward stepwise regression. Plus and minus symbols represent either a positive or negative relationship of that variable with CPUE. For categorical variables, the underlined variables are those set as the baseline. Explanatory variables include average depth, bottom salinity, bottom temperature, depth, distance from shore, mid-latitude, mid-longitude, month, time-of-day, and topography.

	SFL		KW		CW	
	B	**LN**	**B**	**LN**	**B**	**LN**
AVGDEP	**+**	**+**		**--**		
BOTSAL						
BOTTEMP	**+**			**+**		
DEPTH (<u>IN</u>/OFF)						
Distance from shore		**--**	**--**	**--**		**--**
MIDLAT	**+**		**+**	**+**		
MIDLON			**--**	**--**	**--**	
MONTH (8/<u>9/10/11</u>)					**+**	
TIME (Day/<u>Night</u>)	**--**				**+**	**+**
TOPO (NT/<u>TR</u>)				**+**	**+**	**+**

Equations

Formula 1: Structure of a hurdle model (Cragg 1971; Maunder and Punt 2004).

$$\Pr(Y = y) = \begin{cases} w, & y = 0 \\ (1 - w) f(y), & y > 0 \end{cases}$$

Where:

w is the probability of a zero observation, and $f(y)$ is the mean value of the log-normal model.

Formula 2: Extracting year effect from our hurdle model

$$\widehat{d_j} = \pi * \left(exp\left(\widehat{\beta_0} + \widehat{\beta_{ij}}\right) * \psi_{dj}\right)$$

Where:

$\widehat{d_j}$ is the year effect for year j, π is the probability of having a positive observation, ψ_{dj} is the correction bias when assuming a lognormal error structure.

$$\psi = g_m\left[\frac{m+1}{2m}\left(\widehat{\sigma^2} - \widehat{\sigma_\eta^2}\right)\right]$$

$$\hat{\eta} = \widehat{\beta_0} + \widehat{\beta_{ij}}$$

$$g_m(t) = \sum_{p=0}^{\infty}\left[\frac{m^p(m+2p)}{m(m+2)\ldots(m+2p)}\left(\frac{m}{m+1}\right)^p \frac{t^p}{p!}\right]$$

Where:

$\widehat{\sigma^2}$ is the residual variance for the lognormal model, m is the residual variance degrees of freedom.

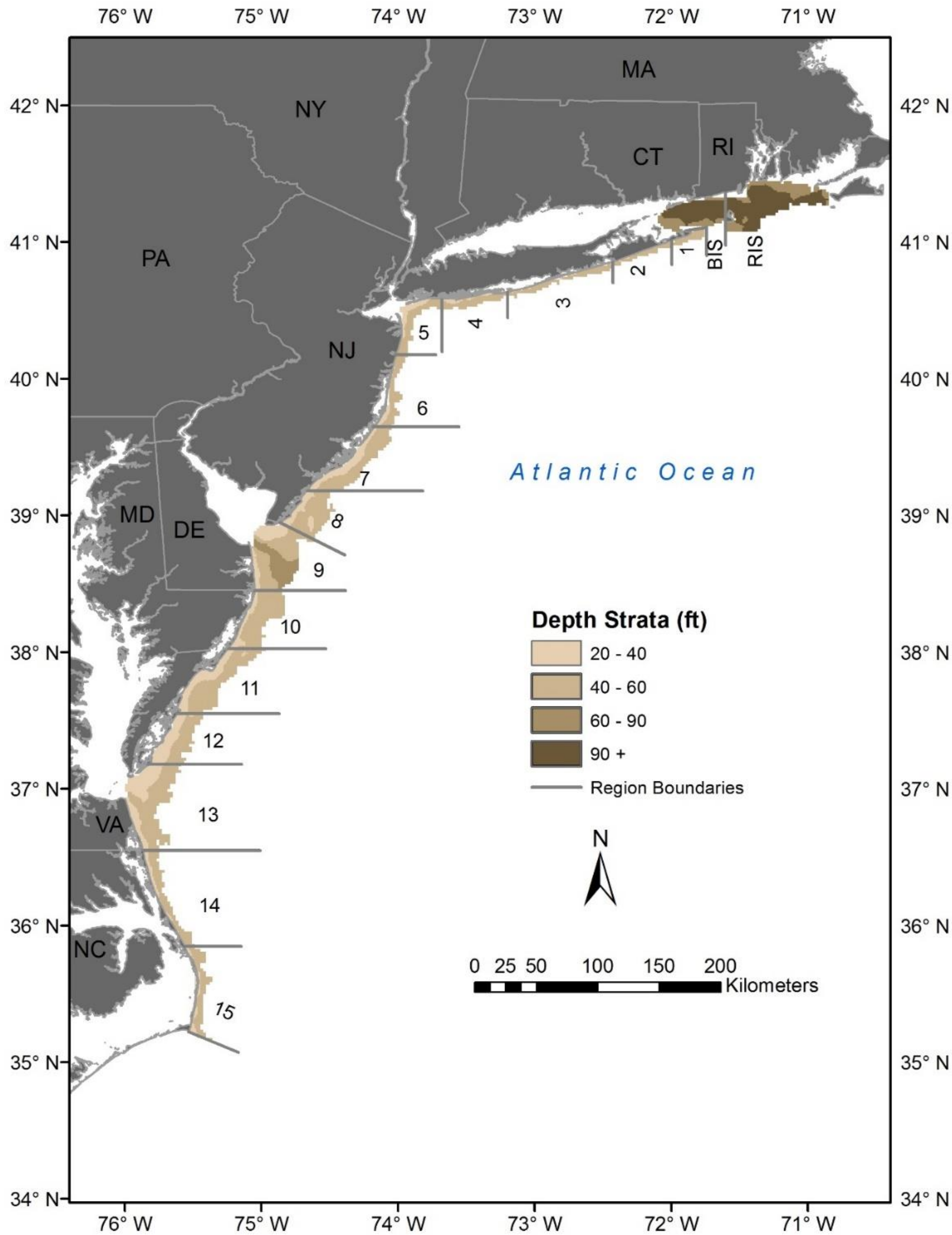


Figure 1: NEAMAP stratified sampling map showing the fifteen regions sampled from Cape Hatteras, NC to Cape Cod, MA (VIMS 2023c).

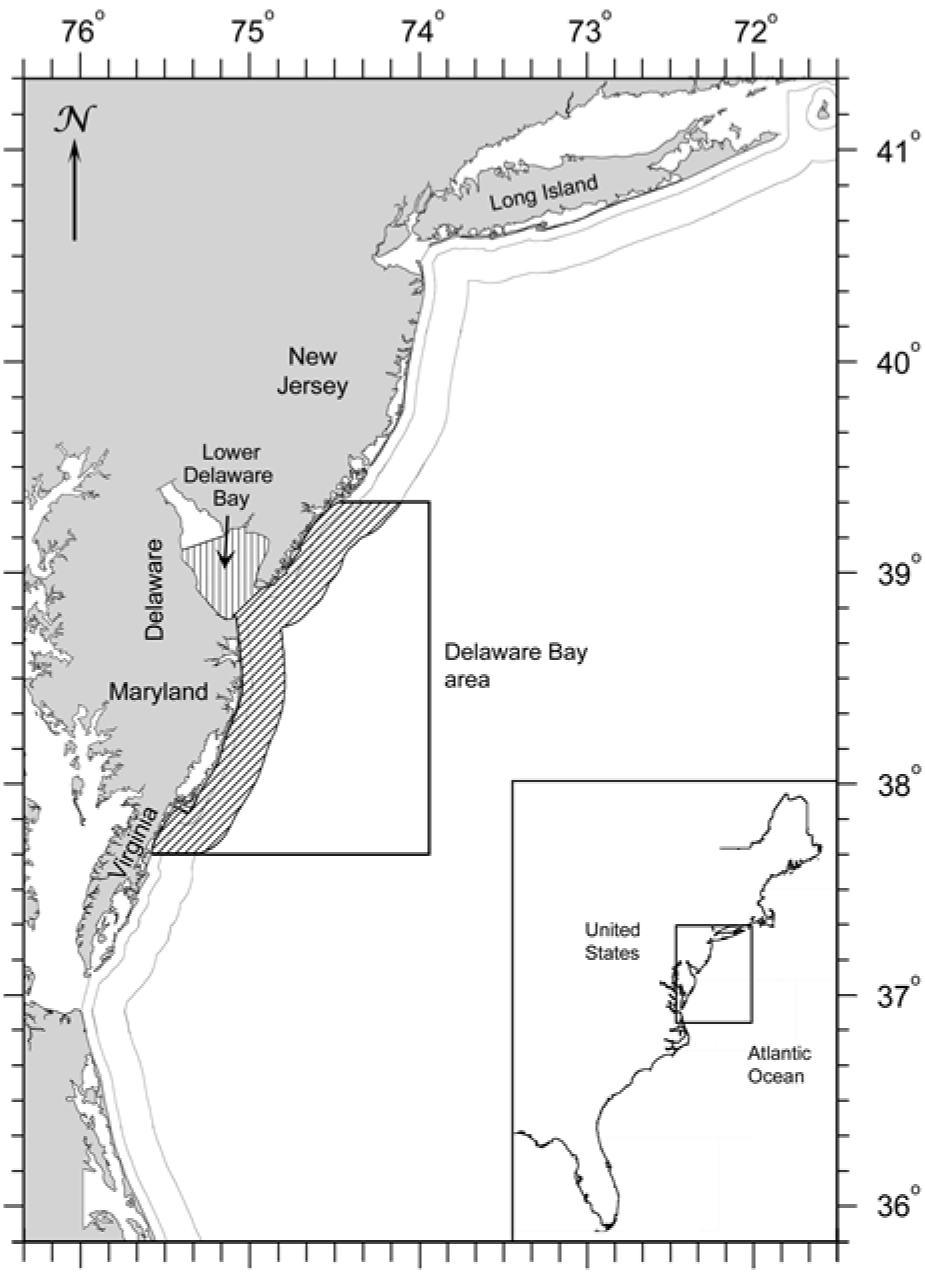


Figure 2: Horseshoe crab trawl survey sampling areas. Contours indicate three and 12 nautical miles from shore.

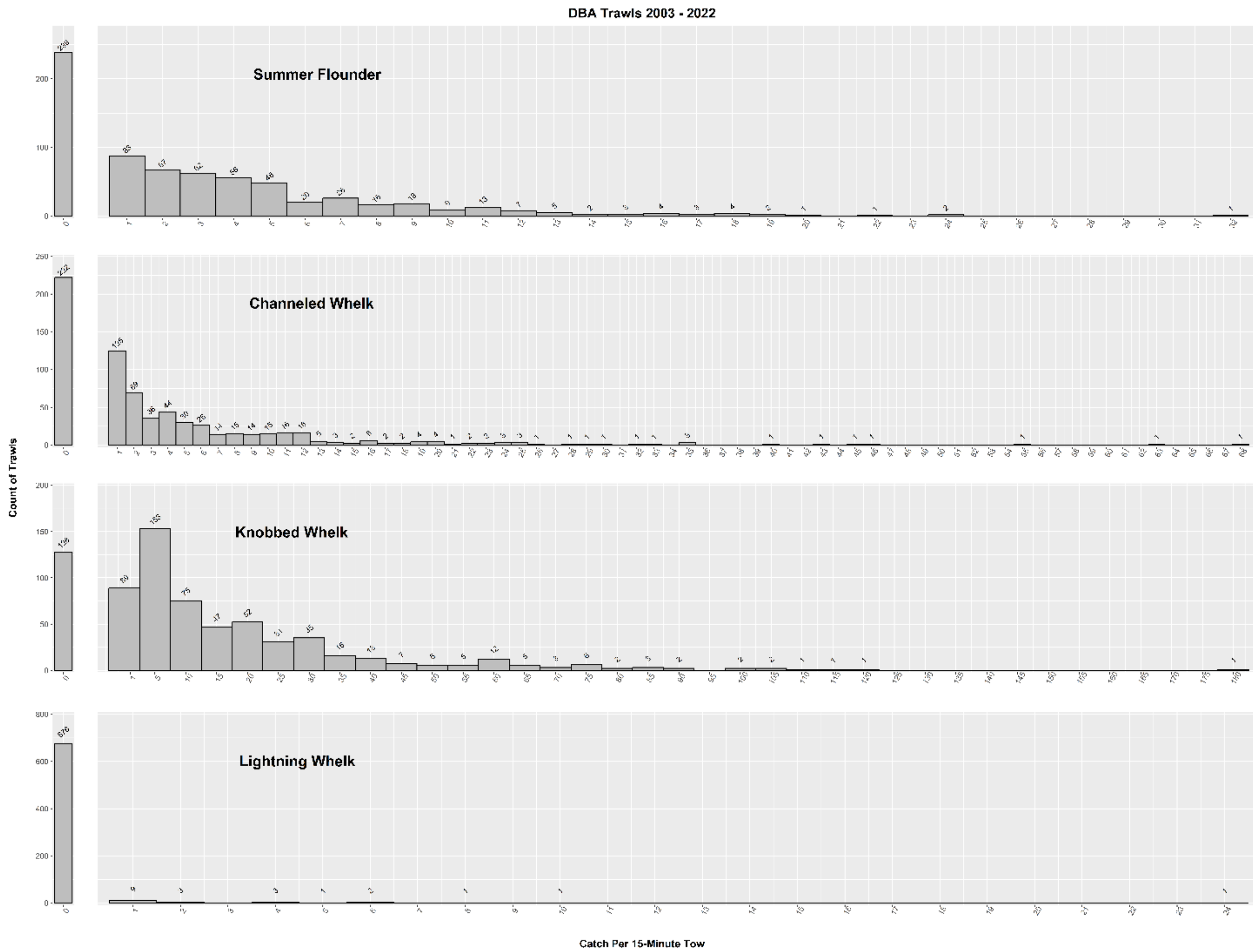


Figure 3: Counts of trawls performed in the Delaware Bay binned by CPUE for each species of bycatch, with zeros separated on the left and positive counts on the right. Note the horizontal and vertical scale differences.

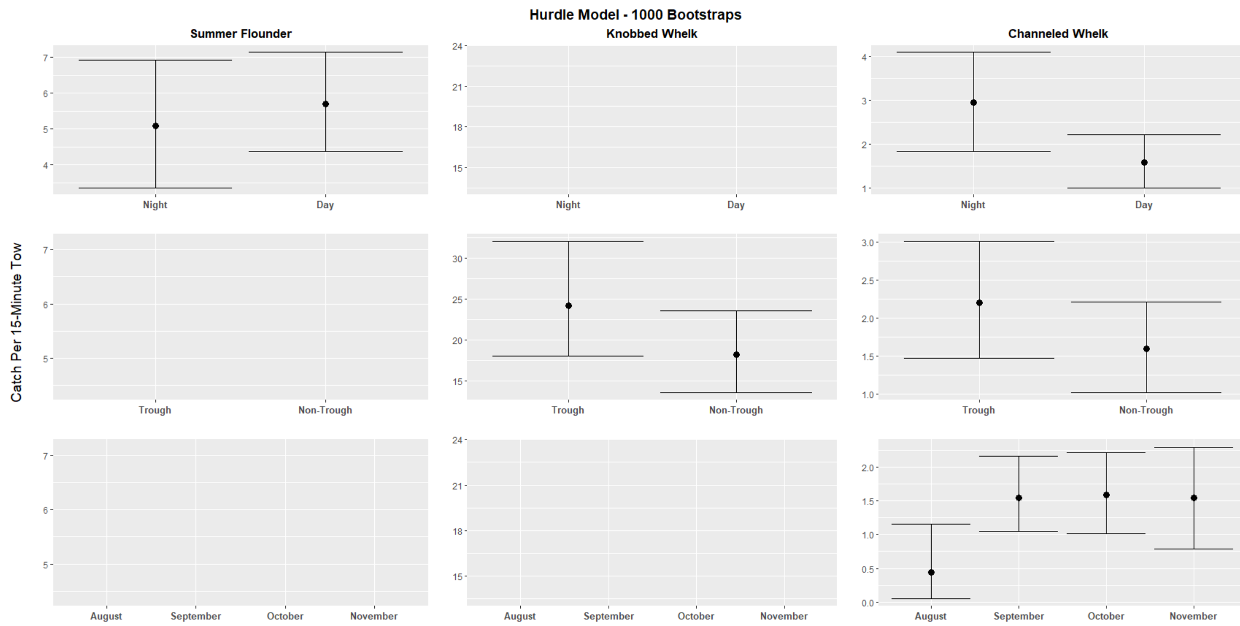


Figure 4: Individual explanatory variable effects analysis of categorical explanatory variables of summer flounder, knobbed whelk, and channeled whelk from the best hurdle model using a 1000-iteration bootstrap analysis. Explanatory variables not selected in specific species' models are shown as blank graphs.

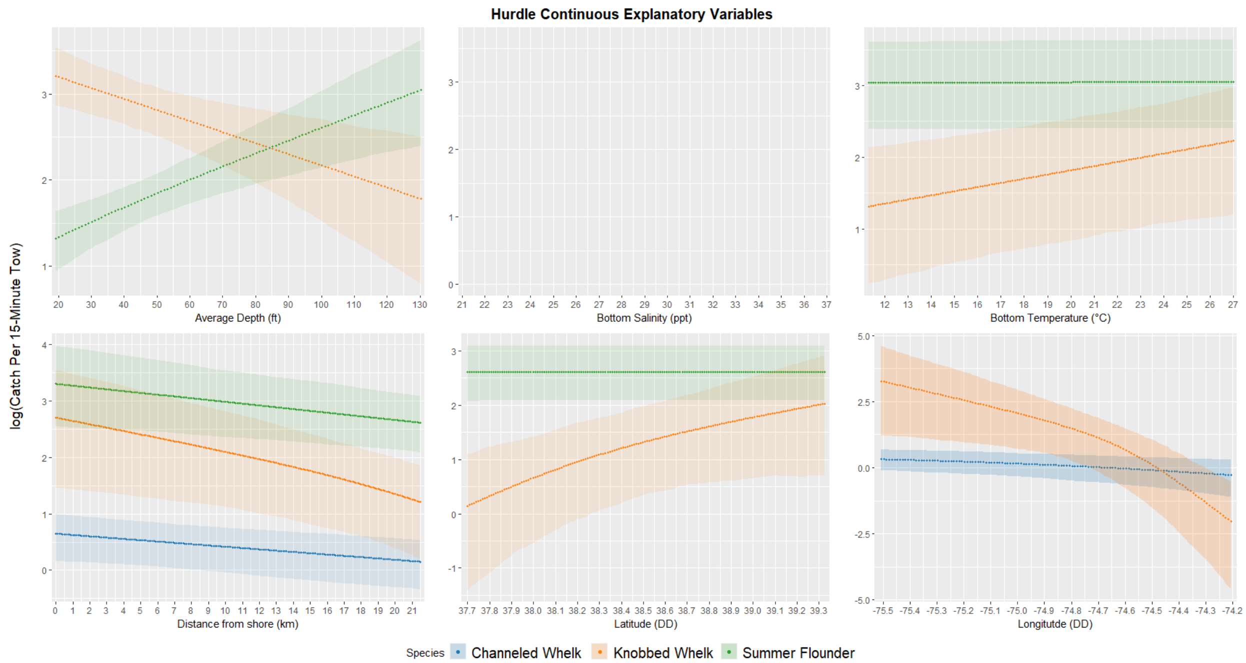


Figure 5: Individual explanatory variable effects analysis showing the mean value and 95% quantile intervals of 1000 iteration bootstrap for summer flounder when changing the value of each continuous explanatory variable in the hurdle model independently to see the results on the catch per 15-minute tow.

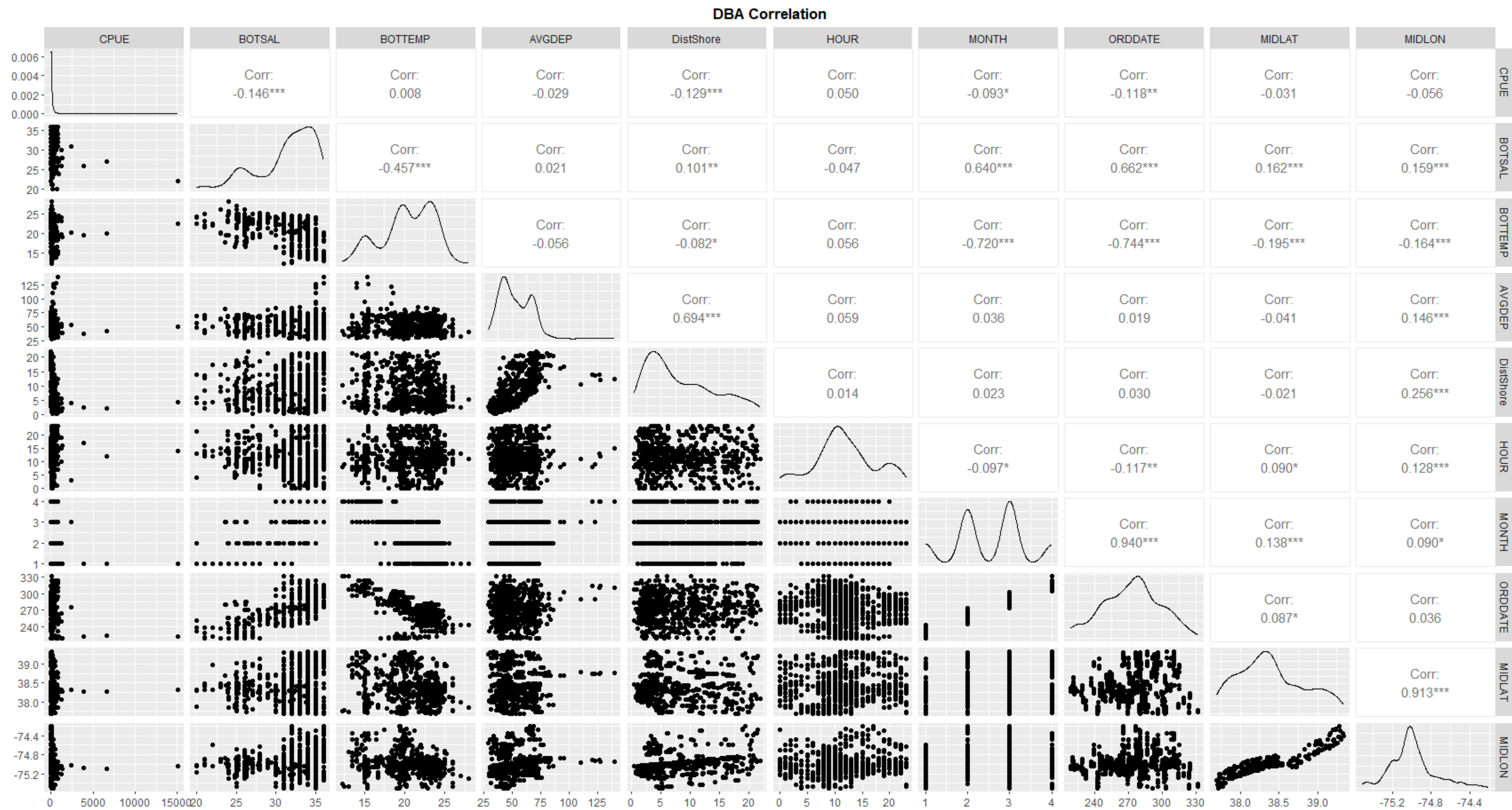


Figure 6: Pairs plot showing Pearson's correlation values and their associated *p*-values (H_a = true correlation is not equal to 0), density plots along the diagonal, and scatter plots for the 697 trawls performed between 2003 and 2022 in the Delaware Bay area after estimating missing environmental factors.

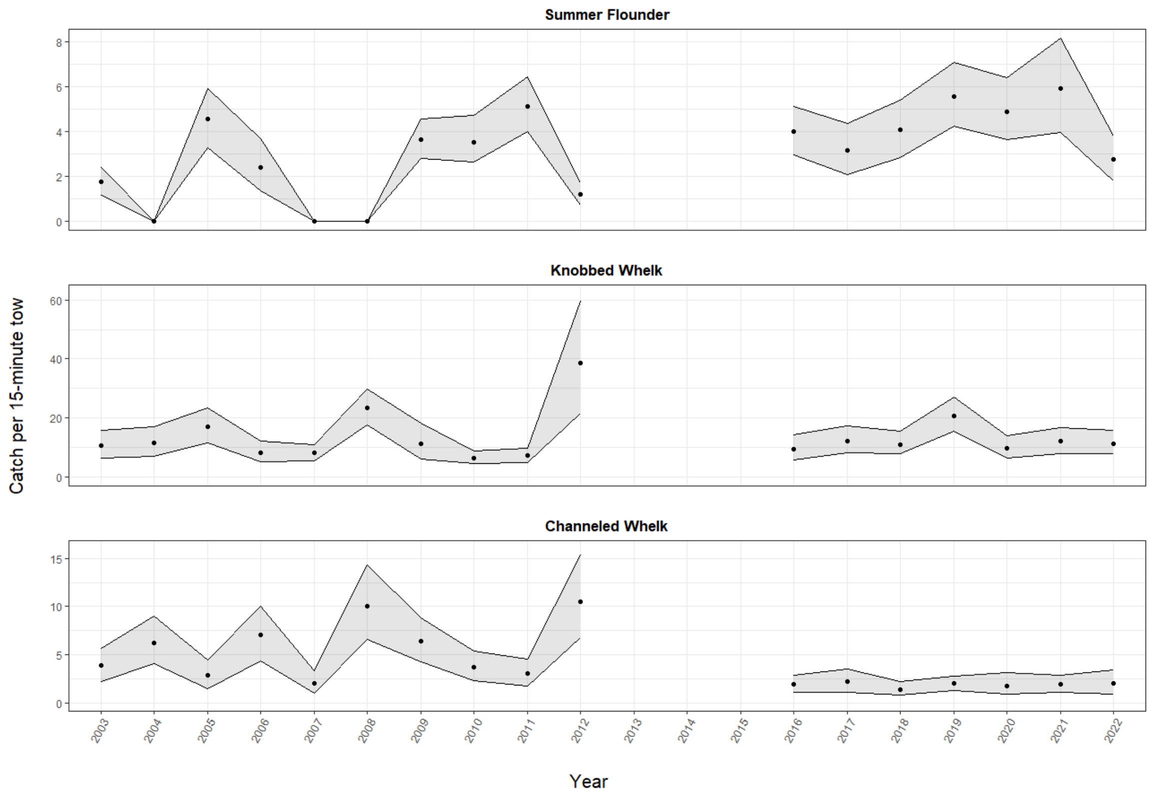


Figure 7: Mean year effect and 95% quantile estimates from 1000 iteration bootstrap analysis for the three species of bycatch modeled over the time series (2003-2022). No summer flounder were caught in 2004, 2007, and 2008, resulting in a zero estimate. The survey did not occur between 2013-2015.

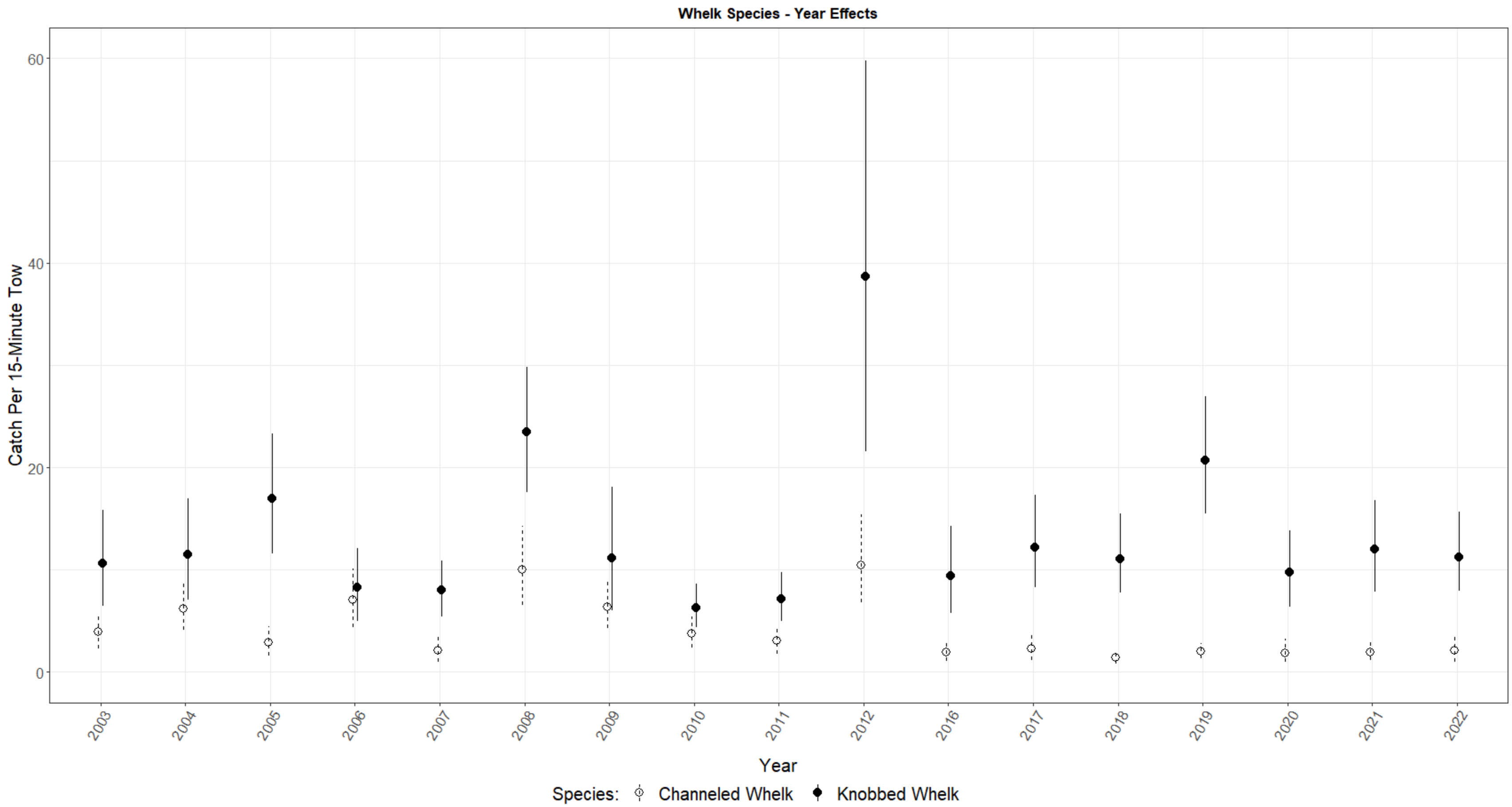


Figure 8: Mean year effect and 95% quantile estimates from 1000 iteration bootstrap of two species of whelk from above superimposed upon one another.

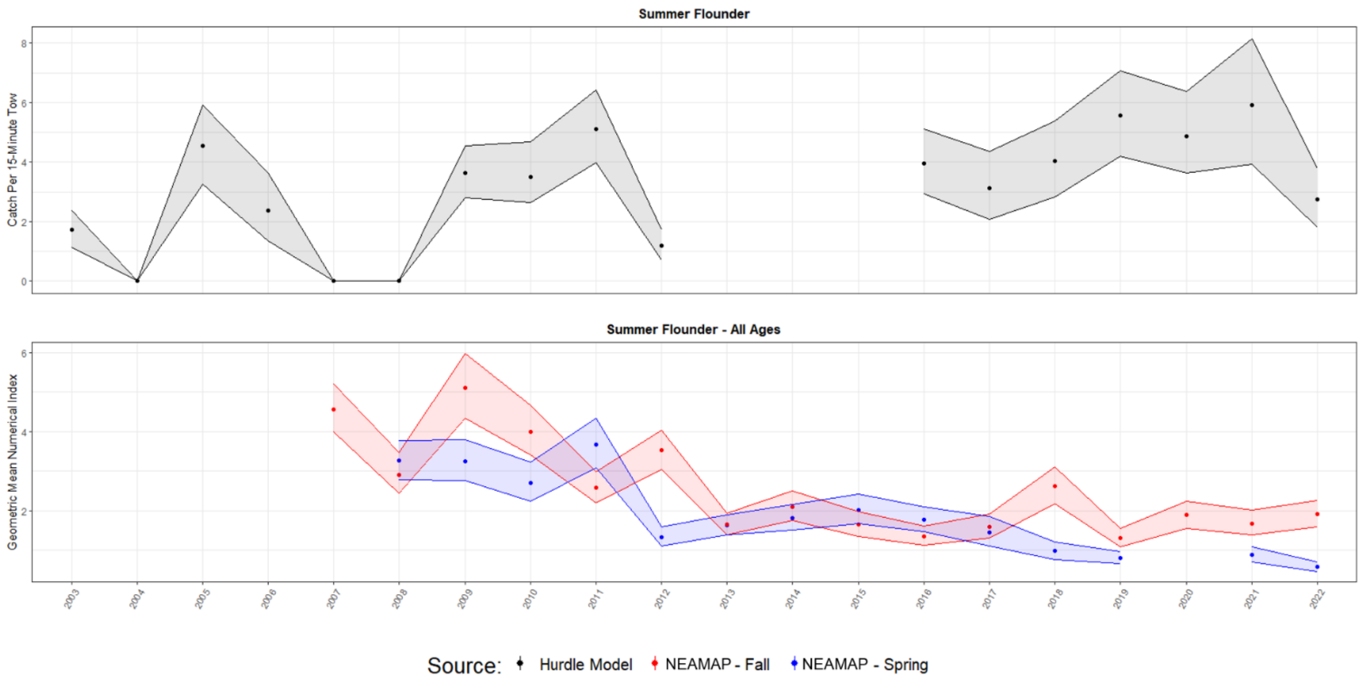


Figure 9: Comparing the relative abundance of summer flounder using the estimated year effect from our hurdle model (top) and NEAMAP's combined All Ages index estimates from their Fall and Spring trawl surveys (bottom). No summer flounder were caught in 2004, 2007, and 2008, resulting in a zero estimate.

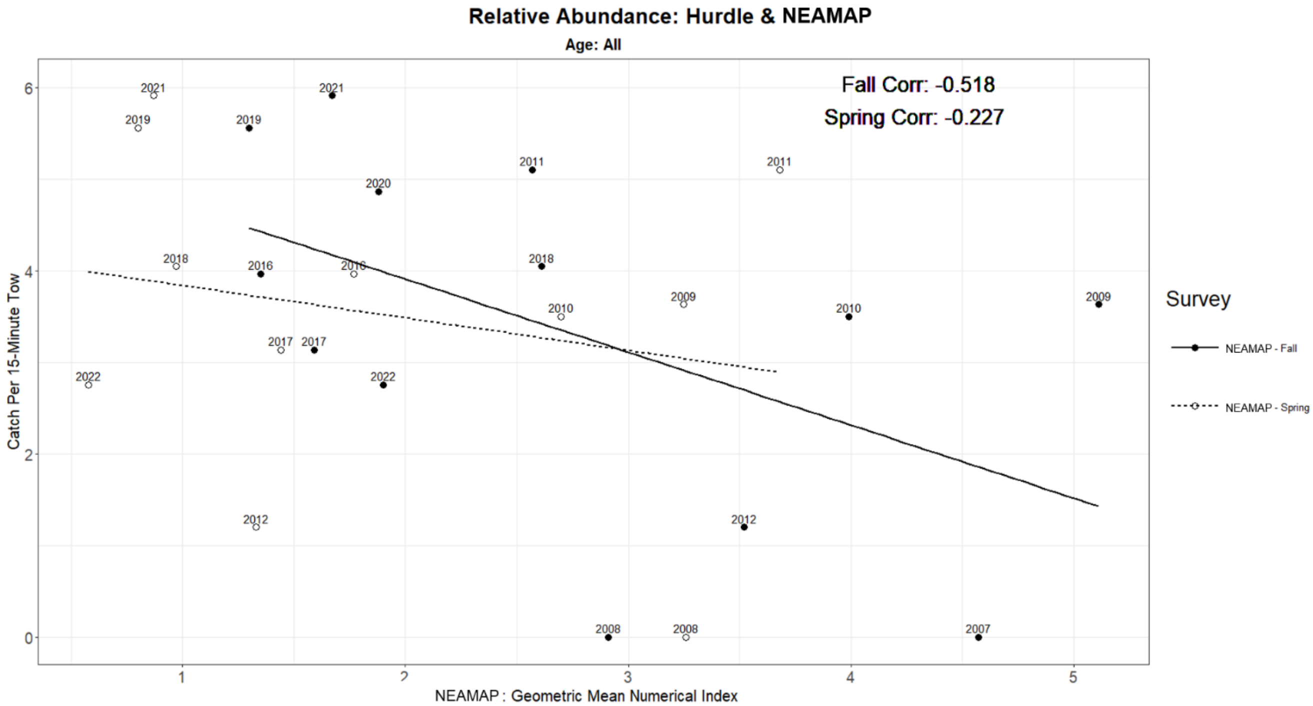


Figure 10: Correlation between hurdle model estimates (*y*) and NEAMAP combined All Ages index (*x*) for their Fall and Spring estimates. NEAMAP estimates were not available for 2003 through 2006, with gaps in the Spring estimates for 2007 and 2020.

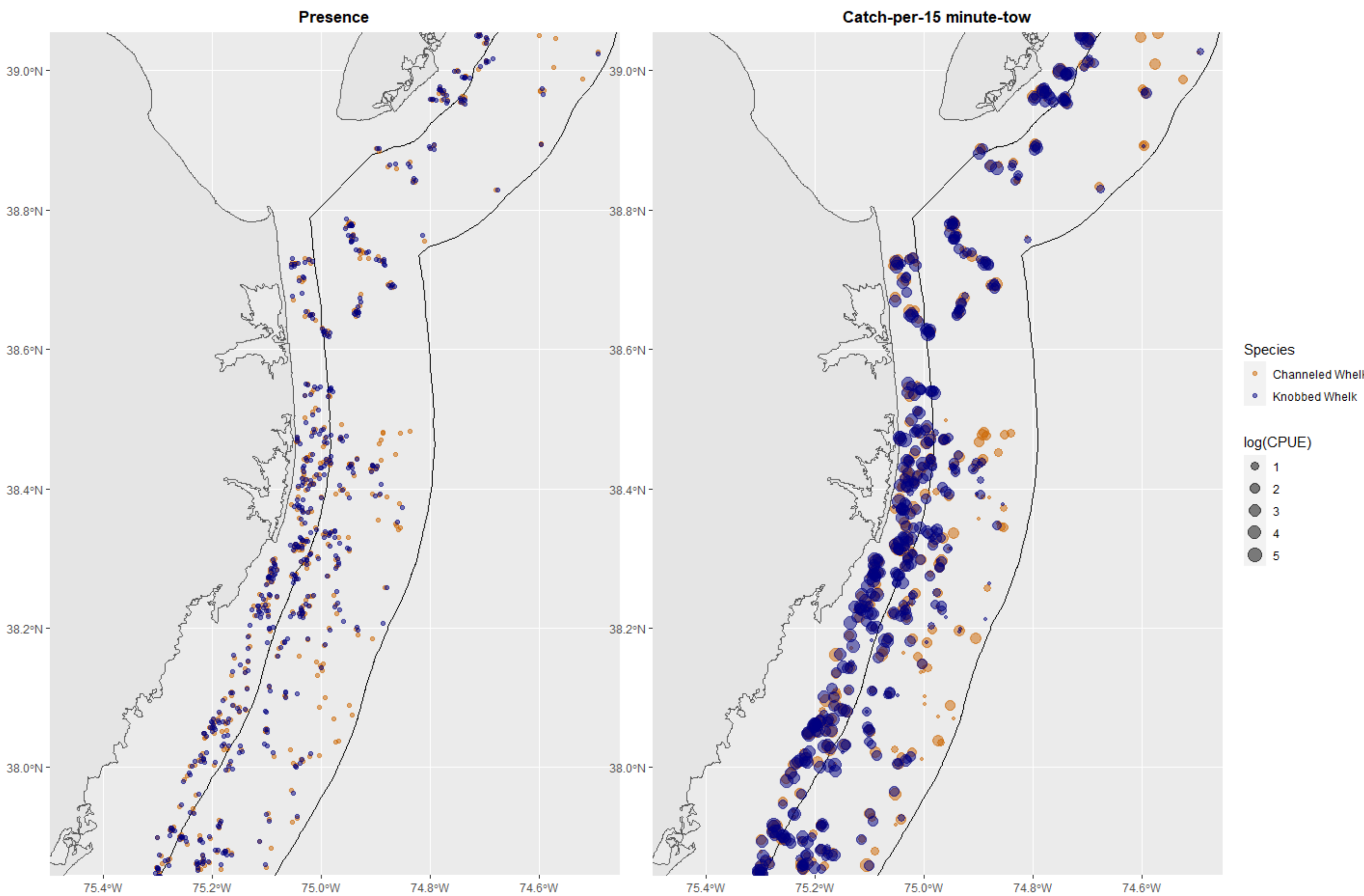


Figure 11: Spatial distribution of positive trawls of channeled and knobbed whelk between 2003 – 2022. The left map shows all positive trawl locations per species of whelk using one point size, while the right shows the same data with differing sizes to indicate log-scaled CPUE. Shown in log-scale for better viewing as knobbed whelk has a higher annual mean CPUE compared to that of channeled whelk.

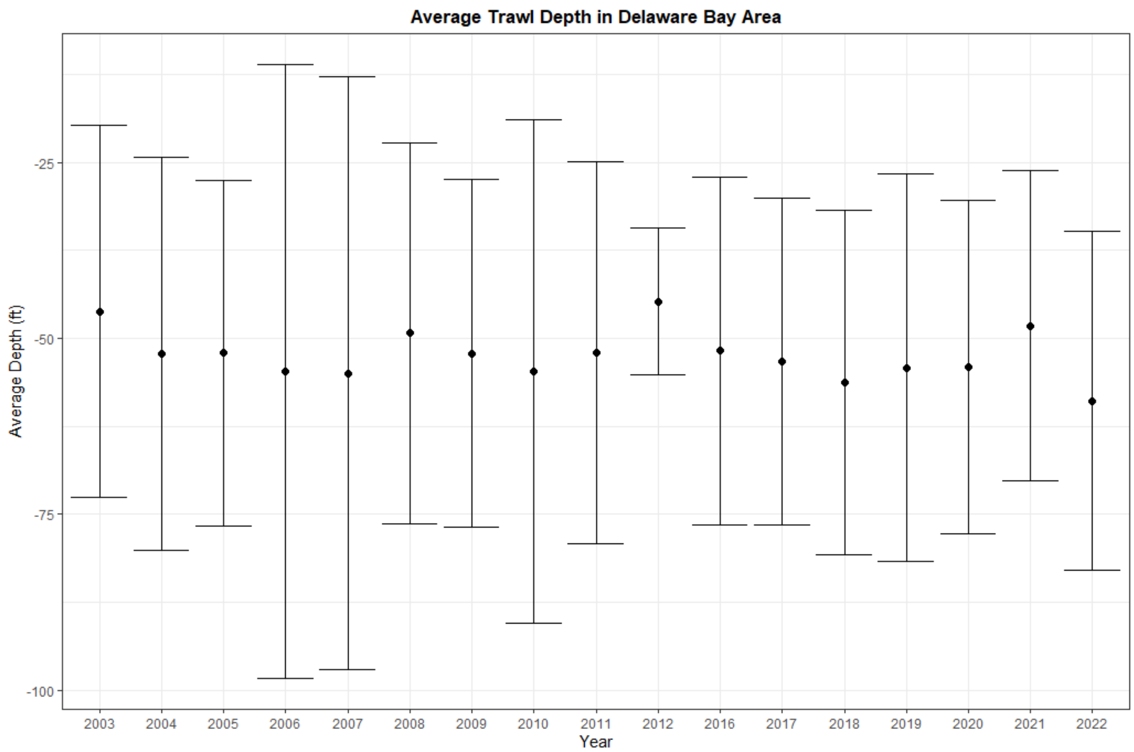


Figure 12: Average trawl depth for all trawls (*n*=697) performed within the Delaware Bay between 2003-2022.

Chapter 3

The effects of sampling time on total abundance estimates for American Horseshoe Crab using spatio-temporal kriging

Chad C. Wong[1], Yan Jiao[1], and Eric Hallerman[1]

[1] Department of Fish and Wildlife Conservation, Virginia Tech, 310 West Campus Drive, Blacksburg, VA 24061, USA

Abstract

The American horseshoe crab (HSC), *Limulus polyphemus*, is harvested as bait for whelk and eel fisheries and for their blood by the biomedical industry. The Atlantic States Marine Fisheries Commission (ASMFC) has managed HSC since 1998, with unique emphasis and harvest limits placed upon the Delaware Bay area stock, as that subpopulation is the center of abundance for horseshoe crabs along the Atlantic coast and provides a critical food resource for the threatened red knot, *Calidris canutus rufa*. Our group has performed trawl surveys of the lower Delaware Bay and mid-Atlantic coast in these areas since 2000 to estimate the relative and total abundance of HSC using a design-based approach. In a previous analysis of this data set, Wong et al. (Chapter 1) constructed six hurdle models to standardize catch rate and estimate relative abundance. While estimates of relative abundance allow managers to examine population trends, estimates of total abundance are used in the current stock assessment to help inform harvest limits. This study applies the hurdle models developed in Wong et al. (Chapter 1) to estimate the total abundance of HSC in the Delaware Bay area. To estimate unmeasured environmental parameters, we developed spatio-temporal variograms to provide monthly estimates of salinity and temperature between August and November. When considering the time-of-day categorical variable, we found that the mean total abundance estimates from trawls conducted at night were consistently higher than those estimated from trawls conducted during the day. Among night estimates, annual total abundance estimates varied by the month estimated, especially for immature and newly-mature females. Mean total abundance was consistently highest in September or November among all demographic groups, although the level of significance between different month estimates varied among demographic groups and year estimated. The results suggest that when using September model-based estimates at night to

compare against the design-based approach, nearly a third of all previous design-based estimates significantly underestimated the total abundance of HSC in the Delaware Bay area. Further, our results confirm how sampling time greatly affects both design and model-based estimates of total abundance and needs to be well considered in future sampling and estimates.

Introduction

With fossil lineages dating back to 445 million years ago (Rudkin and Young 2009), the horseshoe crab (HSC) lineage has survived five mass extinction events. Despite this history, the four extant species of horseshoe crab are not invulnerable to overwhelming ecological damage caused by humans. American horseshoe crabs (*Limulus polyphemus*) are distributed from Maine to the Yucatán Peninsula, with winter temperatures appearing to be the factor limiting their northern distribution (Sekiguchi and Shuster 2009; Smith et al. 2017). Originally harvested and ground up for fertilizer and livestock feed, American horseshoe crab populations have dramatically decreased since being harvested in the millions during the late 19th and early 20th centuries (Shuster and Botton 1985; Millard et al. 2015). Following the widespread adoption of chemical fertilizers, HSC populations stabilized, but declined again during the 1990s due to rising biomedical and bait harvests (Millard et al. 2015). With the implementation of the Interstate Fisheries Management Plan for Horseshoe Crab in 1998, harvest limits were set and reporting became mandatory for all states with an active HSC fishery, and as a result, HSC landings have decreased dramatically since 1999 (Walls et al. 2002; ASMFC 2019). Total harvest has not exceeded one million individuals since 2003, with an average annual landing of 752,886 between 2004 – 2017 (ASMFC 2019). Harvest limits are decided through the application of the ASMFC's adaptive resource management model (ARM), which yields a recommendation for a state-by-state annual quota, with an average limit of 100,000 – 500,000 male HSC harvested annually in the Delaware Bay area (ASMFC 2019), a coastline which includes New Jersey, Delaware, Maryland, and Virginia. More broadly across its range, stock status for HSC is assessed by dividing the Atlantic coast into four regions (Northeast, New York, Delaware Bay, Southeast), with each stock assigned a population status based on the most recent relative abundance estimates compared to the 1998 reference point (ASMFC 2019). According to the most recent horseshoe crab stock assessment (ASMFC 2019), the Delaware Bay regional stock status is assessed as “neutral”, suggesting that there has not been a significant increase or decrease in overall population size since the program’s implementation.

Today, two major stakeholders harvest horseshoe crabs, commercial bait fishers, and the biomedical industry. Commercial fishers use horseshoe crabs as bait for American eel (*Anguilla rostrada*) and whelk/conch species (*Busycon* spp. and *Sinistrofulgur* sp.) (Walls et al. 2002; ASMFC 2019). Although whelk fishers do not prefer either sex, American eel fishers prefer female HSC (Kreamer and Michels 2009), so much so that in the 1990s, almost entirely female HSC were used for bait in Delaware Bay eel pot fisheries (Kreamer and Michels 2009). The biomedical industry uses the lymphocytes isolated from HSC blood to produce *Limulus* amebocyte lysate (LAL), which contains an agglutinating agent that can be used to detect gram-negative bacterial endotoxins in everything from vaccines and injectable pharmaceuticals to implantable medical devices and contaminated food products (Upmann and Bonaparte 1999; Walls et al. 2002). Due to the female HSC’s larger size and hence larger blood volume, the biomedical industry also has a female-skewed harvest (Rutecki et al. 2004; James-Pirri et al. 2012; ASMFC 2019). Natural mortality of horseshoe crabs occurs from predation by species such as loggerhead sea turtles (*Caretta caretta*), leopard sharks (*Triakis semifasciata*), and some ray and skate species (Walls et al. 2002).

Although the bait and biomedical industries are the major extractive stakeholders, a major consideration of ASMFC’s ARM model (McGowan et al. 2011b) is the recovery of the threatened red knot (*Calidris canutus rufa*). The Delaware Bay is a stopover area for many species of shorebirds making their spring migration to their breeding grounds in the Arctic, including red knots (ASMFC 2019). This corresponds with the arrival of HSC to their intertidal breeding grounds for spawning from their overwintering locations offshore. Breeding for horseshoe crabs coincides with warming water temperatures (Watson et al. 2009) and usually lasts between April through July in the mid- and northern coast of the Atlantic Ocean, with peak spawning in the Delaware Bay occurring during new and full moons between May and June (Shuster and Botton 1985; Michels et al. 2008). During this time, there is a skewed operational sex ratio (OSR) of males to females, due to both female-biased harvest and males returning to the beach to mate more frequently than females (Brockmann and Penn 1992; Smith et al. 2002; Hallerman and Hata 2021). In the Delaware Bay, this OSR averaged 3.1 - 4.7 males to one female between 1999 and 2008 (Michels et al. 2008; Wong et al. 2023). The synchronization of HSC breeding and shorebird migration is due in part to the availability of high-energy nutrients from abundant HSC eggs. Red knots on this flyway use HSC eggs as their primary food source to gain weight in advance of their flight to the Arctic (Botton et al. 2003; Mizrahi and Peters 2009). While the number of eggs needed per bird to fly to their final migratory destination is debated, the link between HSC egg availability and red knot survival has been thoroughly examined and is well-accepted (Karpanty et al. 2006; McGowan et al. 2011a; ASMFC 2022). A recent integrated population model (IPM) of the Delaware Bay Area showed a significant correlation between HSC abundance and red knot survival, but not between HSC abundance and red knot recruitment (ASMFC 2022). The most recent revision of the ARM framework implements an Approximate Dynamic Programing (ADP) approach to simultaneously sustain the harvest of HSC while striving to reach a red knot stopover abundance threshold of 81,900 in the Delaware Bay area (ASMFC 2022). This red knot threshold value is a non-formal recovery target and is based on historical abundances previously estimated for the Delaware Bay.

With these factors in mind, the ASMFC is actively working to rebuild the HSC population coast-wide, with special consideration given to the Delaware Bay area. The Delaware Bay is considered the center of abundance for HSC (Shuster 2015) due to its high abundance relative to other populations along the Atlantic coast. Using a design-based estimation approach, the most recent population estimates for the coastal Delaware Bay area included a total of 11.4 million mature females and 19.9 million mature males (Wong et al. 2023). With a life span of fourteen to twenty years, and reaching sexual maturity between eight and ten years, long-term monitoring is essential for managing *L. polyphemus* (Shuster 1955; Sekiguchi et al. 1988). It could take decades for the effects of management to become detectable while harvest continues, especially with the uncertainty about natural and biomedical mortality rates for HSC. In 1998, the horseshoe crab fisheries management plan established a biomedical mortality limit of 57,000 crabs, although limit this has been exceeded every year since 2007, with the exception of 2016, with the associated data being confidential to the AMSFC (ASMFC 2019). Bait harvest limits are set each year based on recommendations from advisory committees implementing the ASMFC's ARM model, with a maximum recommendation of 500,000 males and 210,000 females since Addendum VIII (ASMFC 2023). This model uses a catch multiple survey analysis (CMSA) to estimate the total population of horseshoe crabs coastwide based on multiple regional total population estimates, natural mortality, and biomedical and bait harvest. HSC abundance estimates, in combination with red knot abundance estimates, allow the ASMFC to suggest bait harvest limits with consideration of HSC eggs as a forage item for migratory shorebirds.

Previously, Wong et al. (Chapter 1) developed six hurdle models to standardize the catch rate of HSC, estimate annual relative abundance, and determine the effects of external factors on the CPUE for all six HSC demographic groups. The data used in that analysis was from a horseshoe crab trawl survey of the coastal mid-Atlantic, performed annually each fall since 2003, with a gap in sampling between 2013 and 2015 (Figure 1). These demographic groups are defined by sex (male/female) and stage of maturity (immature, newly-mature, and mature). Applying these models, both sampling month (August, September, October, and November) and sampling time-of-day (day/night) were significant explanatory variables for CPUE for all six demographic groups. The sampling distribution of previous trawls has varied regarding time-of-day (Figure 2) and month (Figure 3) over the study period. External factors, such as distance from shore and average depth, also were found to be significant in explaining the variance in HSC relative abundance, although results varied among the hurdle stages and demographic groups being modeled (Table 1). Alongside the traditional design-based estimates, this study applies hurdle models to provide model-based estimates of total HSC abundance, and to determine how sampling time affects these estimates.

Methods

The survey area (Figure 1) included two sampling areas, the Lower Delaware Bay (LDB) and the Delaware Bay Area (DBA). The LDB extended from the mouth of the bay to a line created between Egg Island Point, NJ and Kitts Hammock, DE. The DBA sampling area consisted of the Atlantic coast extending from 0-12 nautical miles out from shore and from the northern limit of Atlantic City, NJ (39˚ 20' N) south to Wachapreague, VA (37˚ 40' N). For this study, total abundance was estimated only for the DBA, as trawls from this area only were used to fit the original hurdle models. The study area was stratified based on bottom topography (trough, non-trough) and distance from shore (0-3nm, inshore; 3-12nm, offshore). This stratification scheme was determined from the results of a 2001 pilot survey of the Delaware Bay area (Hata and Berkson 2004) and has been used annually since. Troughs were considered areas that are at least 2.4 m deep, no more than 1.8 km wide, and more than 1.8 km long, as determined from NOAA nautical charts. These charts were georeferenced into ArcGIS Pro (ERSI 2023) to create a digital map used to design the sampling scheme. The sample area was divided into cells of 1' of latitude by 1' of longitude using a World Geodetic System (WGS) 1984 coordinate system. The DBA consisted of a total of 1,935 cells in the study area, with inshore troughs representing 5.3% of cells in the study area, inshore non-troughs 22.0%, offshore troughs 17.0%, and offshore non-troughs 55.7%. The number of cells in this study differed slightly from the original values defined in the pilot study (*n*=1890), likely due to cells bordering the 12-nm boundary originally not being considered as part of the total cell count.

To estimate the total abundance of each demographic group, the hurdle models from Wong et al. (Chapter 1) were refitted to change the response variable from catch per fifteen-minute tow to catch density (catch/swept area) (Figure 4). For each tow, catch density was calculated from the product of tow distance and estimated net-spread, assuming that all fishing was done only by the net, and that there was no herding effect from the ground gear (sweeps):

$$catch/km^2 = catch/(tow\ distance\ (km) \times net\text{-}spread\ (km)).$$

No net-spread measurement device was used during sampling. Instead, net-spread was calculated using a net-spread regression relationship,

$$net\ spread\ (S,\ in\ meters)/tow\ speed\ (C,\ in\ KPH),$$

developed from trawl surveys earlier in the study period in which spread was measured $(S = 13.84 - 0.858 \times C)$. From this, the total abundance was estimated for each cell by inputting a cell's environmental parameters into each demographic group's model to estimate catch density, which then was multiplied by the cell area to estimate a cell's total HSC abundance. Total HSC abundance for the Delaware Bay area then was estimated by summing all cell abundance estimates. While the explanatory variables varied for each demographic group's model, the maximum number of variables considered in each were latitude, longitude, distance from shore, average depth of trawl, topography (trough/non-trough), depth (inshore/offshore), month, time-of-day (day/night), bottom temperature, and bottom salinity.

Latitude and longitude values for any given cell were based on the centroid point of the cell. Distance-from-shore values were calculated in R (R Core Team 2023) using the *sf* package's *st_distance* function (Pebesma 2018) and the NOAA NCEI's GSHHG shoreline database (Wessel and Smith 1996, 2018). The specific shape file, GSHSS_f_L1, contained the highest-resolution shoreline file available at a resolution of 1:250,000 (0.04km), last updated June 15, 2017. Using a World Geodetic System (WGS) 1984 coordinate system, distance from shore was calculated as the distance of the cell centroid to the nearest point of the shoreline. Cells were categorized as trough if they contained any trough within their cell. The depth categorical variable was determined by the 3-nm boundary from shoreline, with cells between 0-3 nm being considered inshore, and 3-12 nm considered offshore. If a cell was on the border of the 3-nm line, it was marked as inshore. Average depth was estimated using NOAA's ETOPO Global Relief Model's depth values from 2022 (NOAA NCEI 2022), which provide a 15 arc-second resolution depth map from the bathymetry data of multiple national sources. As the sample area was divided into cells of 1' of latitude by 1' of longitude, a total of 16 depth measurements were assigned to each cell, the mean of which was used as the single average depth value for the cell.

As month and time-of-day were shown to be significant explanatory variables for estimating CPUE for all demographic groups (Wong et al. Chapter 1), estimates of total abundance were estimated per month and time-of-day (Table 1). Bottom temperature and bottom salinity were the two unknown environmental variables selected in some of the hurdle models and are known to vary depending on time and location. Due to this variation, three kriging methods were considered to estimate these variables per cell: spatial kriging, temporal kriging, and spatio-temporal kriging. Spatial kriging is the process of estimating an unknown, continuous variable at a location based on the spatial correlation of previously measured values to one another determined through a covariance function based on distance (Wackernagel 1998, chap. 3). It assumes that the variable of interest is second-order stationary (mean and variance are constant across space) and isotropic in space (the same in all directions). Temporal kriging follows a similar rationale, but instead is based on temporal distance rather than physical distance and follows the same stationary and isotropic assumptions for time. Spatio-temporal kriging combines these concepts by considering the joint effect of spatial and temporal distances on the variable of interest based on a spatio-temporal covariance function (Gräler et al. 2016; Wikle et al. 2019). There are multiple types of spatio-temporal covariance models that change how the spatial and temporal components relate to one another (Gräler et al. 2016). For this study, we considered two covariance models, a separable model and a product-sum model. In a separable model (Appendix E, Formula 1), the spatio-temporal covariance function is the product of separate temporal and spatial covariance functions. In a product-sum model (Appendix E, Formula 2), the spatio-temporal covariance function is an extension of the separable model with an additional summation component between the temporal and spatial covariances, along with a weighting factor k for all three components. Spatio-temporal kriging using a product-sum model proved to be most accurate in predicting known bottom salinity, and a spatio-temporal separable model for predicting known bottom temperature, both based on a leave-one-out-cross-validation (LOOCV). Estimates for each cell were given per month in the time series (August, September, October, and November), and were assumed constant across years. Procedures and results for evaluating the three kriging methods, as well as fitting process for each model, are presented in Appendix E.

After estimating bottom temperature and bottom salinity, total HSC abundance was estimated for each demographic group, per year, month, and time-of-day. Hence, for each year, there were eight sets of estimates of HSC abundance, each representing a combination of month (8, 9, 10, 11) and time-of-day (day/night). Total abundance was estimated as the sum of all-cell abundance in the Delaware Bay area in a given temporal combination, resulting in eight distinct estimates per year. A 200-iteration bootstrap analysis with replacement was used to assess the uncertainty around these estimates in R using a custom script written by the authors. The random sampling for the bootstrap analyses was stratified by year, so that each year had the same number of trawls as the original data set. For each iteration, the hurdle models were refitted using the new data, the top-performing spatio-temporal semivariograms were rebuilt and the variograms refitted for bottom temperature and salinity. This procedure resulted in a new bottom salinity and temperature estimate for each cell and eight new total HSC abundance estimates for each iteration.

Results

Time-of-day effect

For this study, a significant difference between the eight model-based estimates of total abundance was defined as when the 95% quantile intervals between estimates for comparison did not overlap for a given year. When comparing the design and model-based approaches, a significant difference was defined as the design-based 95% confidence interval and the model-based 95% quantile interval not overlapping for a given year's estimate. The estimated effect of time-of-day on CPUE from the hurdle models was such that all estimates of mean total abundance were higher at night than day for estimates of the same month and year for all demographic groups. For immature individuals, almost all total abundance estimates at night for September (Males:13/16 estimates, Females:14/16) and October (M:13/16, F:13/16) were significantly higher than their respective day estimates (Table 2). Less of the total abundance estimates for immature individuals at night were significantly higher for August (M:3/16, F:6/16) and November (M:1/16, F:8/16). All other night-time estimates for immature individuals were not significantly different from their daytime counterparts. For mature individuals, there were no significant differences between day and night estimates for any year or month. Newly-mature individuals also had no significant differences between the different times-of-day for each month besides newly-mature males in 2004, 2007, and 2008, whose October night-time estimates were significantly higher. With these outcomes in mind, night-time model-based estimates will be used to compare against previous design-based estimates and comprise the values presented for the remainder of this study. This is because the night category for time-of-day had a higher estimated coefficient than the daytime category for all demographic groups, and Hata and Berkson (2003, 2004) found that night estimates were more precise and likely more accurate of true abundance than day estimates.

Month effect

For all model-based total abundance estimates at night, either September or November estimates had the highest mean values for any demographic group among all four months for all years (Figures 5,6). Conversely, August consistently had the lowest total abundance mean for all demographic groups except for newly-mature males, whose mean was lowest in October across all years estimated. Model-based estimates for September and November were for most years significantly higher than design-based estimates across all demographic groups, with 39% (37/96) being significantly higher for September and 31% (30/96) for November (Table 3). No demographic group had a particular month whose estimates were consistently significantly higher or lower than the design-based approach across all years.

Night-time mean total abundance estimates for immature females were consistently highest in November, for newly-mature females in September, and for mature females swapped between September and November as estimates from these two months were very similar (Figure 5). For all female maturity groups, most monthly estimates were either not significantly different or significantly higher than the design-based approach, with immature and newly-mature females having a handful of years in which August estimates were significantly lower than the design-based estimates. Immature females had the most single-month annual total abundance estimates that were significantly higher than the design-based estimates, with 75% (12/16) of November's model-based estimates being significantly higher, followed by September with 44% (7/16), and October with 25% (4/16). The August abundance estimates were not significantly different for any year except 2019, where it was significantly lower than the design-based approach. Estimates for mature females were similar, with September having the most years significantly higher at 31% (5/16), followed by October at 25% (4/16), and November at 19% (3/16). Newly-mature female estimates were rarely significantly different, with only August estimates for 2016 and 2018 being significantly-lower and September estimates for 2006 and 2009 being significantly higher than the design-based approach. All other month and year estimates for any female maturity group from the model-based approach were not significantly different than those from the design-based approach.

The highest mean total abundance estimates of immature males swapped between September and November depending on the year, while estimates for mature males were consistently highest in November, and newly-mature in September, except in 2003 where November had the highest mean for newly-mature males (Figure 6). For all male maturity groups, all estimates for any month were either not significantly different or significantly higher than the design-based approach. No model-based estimates were significantly lower than the design-based estimates for males. Like immature females, immature males had the most single-month total abundance estimates that were significantly higher than the design-based estimates, with 56% (9/16) of September's model-based estimates being significantly higher, followed by November with 31% (5/16), October with 19% (3/16), and lastly August with 13% (2/16). Half of the model-based total abundance estimates for mature males in both September and November were significantly higher than the design-based approach. Otherwise, 31% (5/16) of mature male estimates in October were significantly higher, and one August estimate in 2020. Model-based estimates of newly-mature males were not frequently significantly higher, with 38% (6/16) of September estimates being significantly higher than the design-based estimates. November estimates were only significantly higher in 2004 and 2019, and August and October only in 2019. All other month and year estimates for any male maturity group were not significantly different than the design-based approach.

Discussion

Wong et al. (Chapter 1) found that there was rarely a significant difference between the design-based approach and model-based approaches when estimating the relative abundance of mid-Atlantic horseshoe crabs. However, this was not the case for estimates of the total-abundance of horseshoe crabs. One explanation may be that to estimate relative abundance for each year, the mean of the continuous variables and a weighted average of the categorical variables were used in the model-based approach. This approach allowed us to consider a "mean" estimate of relative abundance among all categories to fit the central tendency among all observations of a demographic group, rather than have individual estimates for each combination of categorical variables. For estimating total abundance, while geographic each cell has its own set of unique, pre-defined continuous variables (e.g., longitude, latitude) and some individual categorical variables (e.g., topography and depth), there is no one correct way to assign temporal categorical variables to each cell. Although the total abundance of horseshoe crabs likely does not change considerably from month to month, the availability of crabs (P_m) to the sampling gear does vary. That is because catch is subject to variability in the number of crabs available in the survey area and gear selectivity at the sampling time, such that $Catch_{y,m} = N_y * S * Q \ * \ P_m$, where N_y is the true abundance of horseshoe crabs in year y, S is the gear selectivity or catch vulnerability, Q is catchability, and P_m is the availability of horseshoe crabs in the sampling area in month m. In this equation, time-of-day is included as part of Q (King 2007, chap. 4.2). This variability in P_m can be seen in the results above, with immature females and newly-mature females having the most annual variation in total abundance estimates between the respective months than any other demographic group. Other demographic group estimates, such as immature and newly-mature males and mature individuals, were not as strongly affected, with each of these demographic groups having two or three different month estimates in which the mean values and quantile intervals are remarkably similar. We suspect that the higher catches at night are due to increased horseshoe crab movement at night and burrowing during the day (Hata and Berkson 2004). There could be many reasons for the seasonality effect shown by month, but as suggested by Wong et al. (Chapter 1), it is likely a combination of post-breeding seasonal migration and increased activity of HSC due to higher nutrient availability from more prey items being available during these times.

For use in management, a single total abundance estimate for each demographic group is necessary to apply ASMFC's CMSA model and thereby suggest harvest limits. According to the fitted hurdle models from Wong et al. (Chapter 1), the night level for time-of-day and the September level for month resulted in the highest coefficient estimates of all levels for the two variables among all demographic group models. This outcome means that based on the observational data collected from the sampled population over the time series, trawls performed in September and at night resulted in a higher catch of all horseshoe crab demographic groups on average than other month and time combinations. It does not guarantee, however, that when predicting total abundance that estimates from September at night are going to consistently be the highest for all temporal combinations. This is because a coefficient estimates how the CPUE of a demographic group changes in response to that single variable changing, keeping all other factors constant. As the total abundance of each geographic cell was estimated based on time-of-day, month, and other explanatory variables that differ between cells, the summation of all the cells to estimate total abundance may not show night estimates in September to consistently be highest. Additionally, total abundance was estimated using a bootstrap analysis with replacement, meaning that the hurdle models were refitted each iteration with a "new" set of data, which resulted in different coefficient estimates for month and time for each iteration. This outcome was true of the spatio-temporal variogram model as well, which was refitted every bootstrap iteration, resulting in slightly different estimates of bottom salinity and bottom temperature being input to the associated hurdle models for each geographic cell. All of this in mind, as these original coefficient estimates were based on the true observational data, September estimates at night should be used as the model-based estimate of total abundance to help inform management. That said, the other estimates for total abundance should not be discounted, as they provide valuable insight into the temporal dynamics of the HSC population within the Delaware Bay area, and an estimated range of total abundance that we may expect in the area during a given month, time-of-day, and year.

In this case, all September-night model-based estimates of total abundance were higher than the design-based mean for all years among all demographic groups except for newly-mature females in 2010 and 2021 (Figures 7, 8). When considering all annual total abundance estimates across all demographic groups, 39% (37/96) of the model-based estimates were significantly higher than the design-based estimates. Immature males had the most annual total abundance estimates that were significantly higher than the design-based estimates, with 56% (9/16), and newly-mature females the lowest with 13% (2/16). Values from both the model-based and design-based approach can be found in Appendix F. Overall trends in estimated total abundance did not change for any demographic groups from the design-based approach (Figures 9, 10), although some of the variability seen in the design-based estimates between 2017 to 2020 was smoothed out in the model-based estimates. These results suggest that almost all previous design-based estimates have underestimated the total abundance of horseshoe crabs within the Delaware Bay, significantly so for almost a third of all estimates when using model-based estimates from September at night.

Conclusions

The results of this study show that sampling time, in terms of both month and time-of-day, has a significant effect on total abundance estimates for all horseshoe crab demographic groups. All demographic groups consistently had higher mean estimates of total abundance at night than during the day. Although certain groups were more affected by month than others, mean total abundance estimates at night and in September for all demographic groups were higher than the design-based means, except for two years among newly-mature females. Of these two years, zero newly-mature female HSC were caught in 2021, resulting in a zero estimate for both approaches. Previously, there had been concern about the abundance of newly-mature females, as previous model and design-based estimates of relative abundance, and design-based total abundance, had suggested a recent decline in these individuals since 2018 (Wong et al., Chapter 1). While still present, the September-night model-based estimates from this study suggest that the decline was not as dramatic as previously thought, with no significant difference in model-based estimates of total abundance since 2010, besides 2021 where no newly-mature females were caught.

These results confirm the need to adjust our sampling scheme for future surveys, as suggested in Wong et al. (Chapter 1). If there was more even sampling between months (Figure 2) and the times-of-day (Figure 3), future research would be able to update these hurdle models and possibly bolster or change our understanding of how these variables affect CPUE while also improving the accuracy of the design and model-based estimates of total abundance. Simulation studies could also be performed to investigate which estimation method more accurately estimates the simulated abundance in the forms of relative and total abundance. A simulation study also could be used to determine what sampling scheme (number of trawls performed and sampling distribution among months and times-of-day) would result in the more accurate estimates of total abundance for both estimation approaches. This study also emphasizes the importance of estimating both the relative and total abundance of horseshoe crabs for management in the Delaware Bay area. Relative abundance estimates like those presented in

Wong et al. (Chapter 1) allow researchers to examine the overarching population trends of the six horseshoe crab demographic groups while removing the effect of external factors like month and time-of-day. This can help researchers determine how changes in management have affected the population relative to their implementation without concern about which month and time-of-day total abundance estimate most accurately represents the true population.

For management, these results suggest that a higher level of mature male and female horseshoe crab harvest may be justifiable for the mid-Atlantic region if it aligns with the goals of the ARM framework. As previously mentioned, horseshoe crab harvest limits are recommended based on both the total abundance of mature horseshoe crabs *and* the needs of migratory shorebirds (ASMFC 2023). It would be necessary for the ASFMC to see how these new estimates affect their CMSA model and, in turn, their ARM model and recommendations. It should also be noted that while the ASMFC has recommended the harvest of both male and female HSC annually, the Horseshoe Crab Management Board banned commercial bait harvest of female horseshoe crabs starting in 2013 and has continued to do so up to the upcoming 2024 season (ASMFC 2012, 2023). These results should reassure the general public that the already-conservative previous harvest recommendations of the ASMFC, if implemented by the Board, should not dramatically affect the population abundance of horseshoe crabs and or limit the availability of HSC eggs to migratory shorebirds.

Literature Cited

ASMFC. 2012. ASMFC Horseshoe Crab Board Sets Annual Specifications in Adaptive Multi-Species Management. Atlantic States Marine Fisheries Commission, Arlington, VA. Available: https://horseshoecrab.org/press/2012/10/pr55HSC_ARMspecifications.pdf

ASMFC. 2019. 2019 Horseshoe Crab Benchmark Stock Assessment. Atlantic States Marine Fisheries Commission, Arlington, VA. Available: https://asmfc.org/uploads/file/63d2ed62HSCAssessment_PeerReviewReport_May2019.pdf

ASMFC. 2022. 2021 Horseshoe Crab Adaptive Resource Management Revision Overview. Atlantic States Marine Fisheries Commission, Arlington, VA. Available: https://www.asmfc.org/uploads/file/63d2eb79HSC_ARM_RevisionOverview_Jan2022.pdf

ASMFC. 2023. Delaware Bay Horseshoe Crab Harvest Recommendation for 2024. Atlantic States Marine Fisheries Commission, Arlington, VA. Available: https://www.asmfc.org/uploads/file/6539327bDBETC__ARM_HSC_2024_HarvestRecommendationMemo.pdf

Botton, M. L., B. A. Harrington, N. Tsipoura, and D. Mizrahi. 2003. Synchronies in Migration: Shorebirds, Horseshoe Crabs, and Delaware Bay. Pages 5–32 *in* C. N. Shuster, H. J. Brockmann, and R. B. Barlow, editors. The American Horseshoe Crab. Harvard University Press, Cambridge, MA.

Brockmann, H. J., and D. Penn. 1992. Male Mating Tactics in the Horseshoe Crab, *Limulus polyphemus*. Animal Behaviour 44(4):653–665.

Dumelle, M., J. M. Ver Hoef, C. Fuentes, and A. Gitelman. 2021. A Linear Mixed Model Formulation for Spatio-Temporal Random Processes with Computational Advances for the Product, Sum, and Product–Sum Covariance Functions. Spatial Statistics 43:100510.

ERSI. 2023. ArcGIS Pro Version 3.1.0. Environmental Systems Research Institute, Inc., Redlands, CA.

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

Hallerman, E. M., D. N. Hata, M. S. Eackles, and T. L. King. 2022. Population Genetics and Movement Show Metapopulation Dynamics of Mid-Atlantic Region Horseshoe Crabs. Pages 41–57 *in* J. T. Tanacredi, M. L. Botton, P. K. S. Shin, Y. Iwasaki, S. G. Cheung, K. Y. Kwan, and J. H. Mattei, editors. International Horseshoe Crab Conservation and Research Efforts: 2007-2020. Springer International Publishing, Cham.

Hata, D., and J. Berkson. 2003. Abundance of Horseshoe Crabs (*Limulus polyphemus*) in the Delaware Bay area. NOAA Fishery Bulletin 101(4):933–938.

Hata, D., and J. Berkson. 2004. Factors Affecting Horseshoe Crab *Limulus polyphemus* Trawl Survey Design. Transactions of the American Fisheries Society 133(2):292–299.

James-Pirri, M.-J., P. A. Veillette, and A. S. Leschen. 2012. Selected Hemolymph Constituents of Captive, Biomedically Bled, and Wild Caught Adult Female American Horseshoe Crabs (*Limulus polyphemus*). Marine and Freshwater Behaviour and Physiology 45(4):281–289.

Karpanty, S. M., J. D. Fraser, J. Berkson, L. J. Niles, A. Dey, and E. P. Smith. 2006. Horseshoe Crab Eggs Determine Red Knot Distribution in Delaware Bay. The Journal of Wildlife Management 70(6):1704–1710.

King, M. G. 2007. Fisheries Biology, Assessment and Management. 2nd edition. Blackwell Publishing, Ames, Iowa.

Kreamer, G., and S. Michels. 2009. History of Horseshoe Crab Harvest on Delaware Bay. Pages 299–313 *in* J. T. Tanacredi, M. L. Botton, and D. R. Smith, editors. Biology and Conservation of Horseshoe Crabs. Springer, Boston, MA.

McGowan, C. P., J. E. Hines, J. D. Nichols, J. E. Lyons, D. R. Smith, K. S. Kalasz, L. J. Niles, A. D. Dey, N. A. Clark, P. W. Atkinson, C. D. T. Minton, and W. Kendall. 2011a. Demographic Consequences of Migratory Stopover: Linking Red Knot Survival to Horseshoe Crab Spawning Abundance. Ecosphere 2(6):1–22.

McGowan, C. P., D. R. Smith, J. A. Sweka, J. Martin, J. D. Nichols, R. Wong, J. E. Lyons, L. J. Niles, K. Kalasz, J. Brust, M. Klopfer, and B. Spear. 2011b. Multispecies Modeling for Adaptive Management of Horseshoe Crabs and Red Knots in the Delaware Bay. Natural Resource Modeling 24(1):117–156.

Michels, S. F., D. Smith, and S. Bennett. 2008. Horseshoe Crab Spawning Activity in Delaware Bay: 1999–2007: Report to the Horseshoe Crab Technical Committee. Atlantic States Marine Fisheries Commission, Arlington, VA. Available: https://www.delawarebayhscsurvey.org/statisticalreports

Millard, M. J., J. A. Sweka, C. P. McGowan, and D. R. Smith. 2015. Assessment and Management of North American Horseshoe Crab Populations, with Emphasis on a Multispecies

Framework for Delaware Bay, U.S.A. Populations. Pages 407–431 *in* R. H. Carmichael, M. L. Botton, P. K. S. Shin, and S. G. Cheung, editors. Changing Global Perspectives on Horseshoe Crab Biology, Conservation and Management. Springer International Publishing, New York, New York.

Mizrahi, D. S., and K. A. Peters. 2009. Relationships Between Sandpipers and Horseshoe Crab in Delaware Bay: A Synthesis. Pages 65–87 *in* J. T. Tanacredi, M. L. Botton, and D. Smith, editors. Biology and Conservation of Horseshoe Crabs. Springer, Boston, MA.

NOAA NCEI. 2022. ETOPO 2022 15 Arc-Second Global Relief Model. National Oceanic and Atmospheric Administration National Centers for Environmental Information. Available: https://doi.org/10.25921/fd45-gt74.

Pebesma, E. 2018. Simple Features for R: Standardized Support for Spatial Vector Data. The R Journal 10(1):439.

Pebesma, E. J. 2004. Multivariable Geostatistics in S: The gstat Package. Computers & geosciences 30(7):683–691.

R Core Team. 2023. R: A Language and Environment for Statistical Computing. R, R Foundation for Statistical Computing, Vienna, Austria.

Rudkin, D. M., and G. A. Young. 2009. Horseshoe Crabs – An Ancient Ancestry Revealed. Pages 25–44 *in* J. T. Tanacredi, M. L. Botton, and D. Smith, editors. Biology and Conservation of Horseshoe Crabs. Springer, Boston, MA.

Rutecki, D., R. H. Carmichael, and I. Valiela. 2004. Magnitude of Harvest of Atlantic Horseshoe Crabs, *Limulus polyphemus*, in Pleasant Bay, Massachusetts. Estuaries 27(2):179–187.

Sekiguchi, K., H. Seshimo, and H. Sugita. 1988. Post-Embryonic Development of the Horseshoe Crab. Biological Bulletin 174(3):337–345.

Sherman, M. 2011. Spatial Statistics and Spatio-Temporal Data: Covariance Functions and Directional Properties. John Wiley & Sons, New York, NY.

Shuster, C. N. 1955. On Morphometric and Serological Relationships within the *Limulidae*, with Particular Reference to *Limulus polyphemus* (L.). Ph.D., New York University, New York, NY.

Shuster, C. N. 2015. The Delaware Bay Area, U.S.A.: A Unique Habitat of the American Horseshoe Crab, *Limulus polyphemus*. Pages 15–39 *in* R. H. Carmichael, M. L. Botton, P. K. S. Shin, and S. G. Cheung, editors. Changing Global Perspectives on Horseshoe Crab Biology, Conservation and Management. Springer International Publishing, Cham, Switzerland.

Shuster, C. N., and M. L. Botton. 1985. A Contribution to the Population Biology of Horseshoe Crabs, *Limulus polyphemus* (L.), in Delaware Bay. Estuaries 8(4):363.

Smith, D. R., P. S. Pooler, B. L. Swan, S. F. Michels, W. R. Hall, P. J. Himchak, and M. J. Millard. 2002. Spatial and Temporal Distribution of Horseshoe Crab (*Limulus polyphemus*) Spawning in Delaware Bay: Implications for Monitoring. Estuaries 25(1):115–125.

Wackernagel, H. 1998. Multivariate Geostatistics, 2nd edition. Springer, Berlin and Heidelberg.

Walls, E. A., J. Berkson, and S. A. Smith. 2002. The Horseshoe Crab, *Limulus polyphemus* : 200 Million Years of Existence, 100 Years of Study. Reviews in Fisheries Science 10(1):39–73.

Watson, W. H., S. Y. Schaller, and C. C. Chabot. 2009. The Relationship Between Small- and Large-Scale Movements of Horseshoe Crabs in the Great Bay Estuary and *Limulus* Behavior in the Laboratory. Pages 131–147 *in* J. T. Tanacredi, M. L. Botton, and D. Smith, editors. Biology and Conservation of Horseshoe Crabs. Springer, Boston, MA.

Wessel, P., and W. H. F. Smith. 1996. A Global, Self-Consistent, Hierarchical, High-Resolution Shoreline Database. Journal of Geophysical Research: Solid Earth 101(B4):8741–8743.

Wikle, C. K., A. Zammit-Mangion, and N. A. C. Cressie. 2019. Spatio-temporal statistics with R. CRC Press, Boca Raton, FL.

Wong, C., Y. Jiao, and E. Hallerman. 2023. Results of the 2022 Horseshoe Crab Trawl Survey: Report to the Atlantic States Marine Fisheries Commission Horseshoe Crab and Delaware Bay Ecology Technical Committees. Atlantic States Marine Fisheries Commission, Arlington, VA.

Table 1: Explanatory variables selected in the binomial (B) and log-normal (LN) stages of the hurdle models for each demographic group from Wong et al.'s (Chapter 1) unadjusted-salinity models. Plus and minus symbols represent either a positive or negative relationship of that variable to CPUE. For categorical variables, the underlined variables are those set as the baseline.

	FI		FN		FM		MI		MN		MM	
Explanatory Variable	**B**	**LN**	**B**	**LN**	**B**	**LN**	**B**	**LN**	**B**	**LN**	**B**	**LN**
Average depth	+	+	+	+	+	+	+	+	+	+	+	+
Bottom salinity	--				--							--
Bottom temperature	--	--				--	--	--				--
Depth (<u>IN</u>/OFF)												
Distance from shore	--	--	--		--	--	--	--	--	--	--	--
Latitude					+	+						+
Longitude	--	--	--	--	--	--	--	--	--	--		--
Month (<u>8</u>/9/10/11)	+	+		+	+	+	+	+	+	+	+	+
Time-of-day (<u>Day</u>/Night)	+	+	+		+	+	+	+	+	+	+	+
Topography (<u>NT</u>/TR)				+			+					

Table 2: Percentage of total abundance estimates for each demographic group whose **night estimates** were not significantly different ("No Diff.") or significantly higher ("Higher") than the corresponding **day estimates** between 2002-2022. No night estimates were significantly lower than the corresponding day estimates.

Group	Month	No Diff.	Higher	Group	Month	No Diff.	Higher
Immature Females	Aug.	62.5%	37.5%	**Immature Males**	Aug.	81.25%	18.75%
	Sep.	12.5%	87.5%		Sep.	18.75%	81.25%
	Oct.	18.75%	81.25%		Oct.	18.75%	81.25%
	Nov.	50%	50%		Nov.	93.75%	6.25%
Newly-mature Females	Aug.	100%		**Newly-mature Males**	Aug.	100%	
	Sep.	100%			Sep.	100%	
	Oct.	100%			Oct.	81.25%	18.75%
	Nov.	100%			Nov.	100%	
Mature Females	Aug.	100%		**Mature Males**	Aug.	100%	
	Sep.	100%			Sep.	100%	
	Oct.	100%			Oct.	100%	
	Nov.	100%			Nov.	100%	

Table 3: Percentage of total abundance estimates for each demographic group whose **model-based night estimates** were significantly lower (“Lower”), not significantly different (“No Diff.”), or significantly higher (“Higher”) than the **design-based normal estimates** between 2002-2022.

Group	Month	Lower	No Diff.	Higher	Group	Month	Lower	No Diff.	Higher
Immature Females	Aug.	6.25%	93.75%		Immature Males	Aug.		87.5%	12.5%
	Sep.		56.25%	43.75%		Sep.		43.75%	56.25%
	Oct.		75%	25%		Oct.		81.25%	18.75%
	Nov.		25%	75%		Nov.		68.75%	31.25%
Newly-mature Females	Aug.	12.5%	87.5%		Newly-mature Males	Aug.		93.75%	6.25%
	Sep.		87.5%	12.5%		Sep.		62.5%	37.5%
	Oct.		100%			Oct.		93.75%	6.25%
	Nov.		100%			Nov.		87.5%	12.5%
Mature Females	Aug.		100%		Mature Males	Aug.		93.75%	6.25%
	Sep.		68.75%	31.25%		Sep.		50%	50%
	Oct.		75%	25%		Oct.		68.75%	31.25%
	Nov.		81.25%	18.75%		Nov.		50%	50%

Equations

Formula 1: Separable (product) covariance formula function (Gräler et al. 2016; Dumelle et al. 2021).

$$C(\boldsymbol{h}_s, h_t) = C_s(\boldsymbol{h}_s)C_t(h_t)$$

Where:

$C_{[s/t]}(h_{[s/t]})$ represents the covariance between two spatial (s) or temporal (t) points at distance h

$C_s(\boldsymbol{h}_s)$ is the spatial covariance function; $C_t(h_t)$ is the temporal covariance.

Formula 2: Product-sum covariance formula function (Gräler et al. 2016; Dumelle et al. 2021).

$$C(\boldsymbol{h}_s, h_t) = k_1 C_s(\boldsymbol{h}_s)C_t(h_t) + k_2 C_s(\boldsymbol{h}_s) + k_3 C_t(\boldsymbol{h}_t)$$

Where:

$C_{[s/t]}(h_{[s/t]})$ represents the covariance between two spatial (s) or temporal (t) points at distance h

$C_s(\boldsymbol{h}_s)$ is the spatial covariance function; $C_t(h_t)$ is the temporal covariance function

k_1, k_2, k_3 are nonnegative weights.

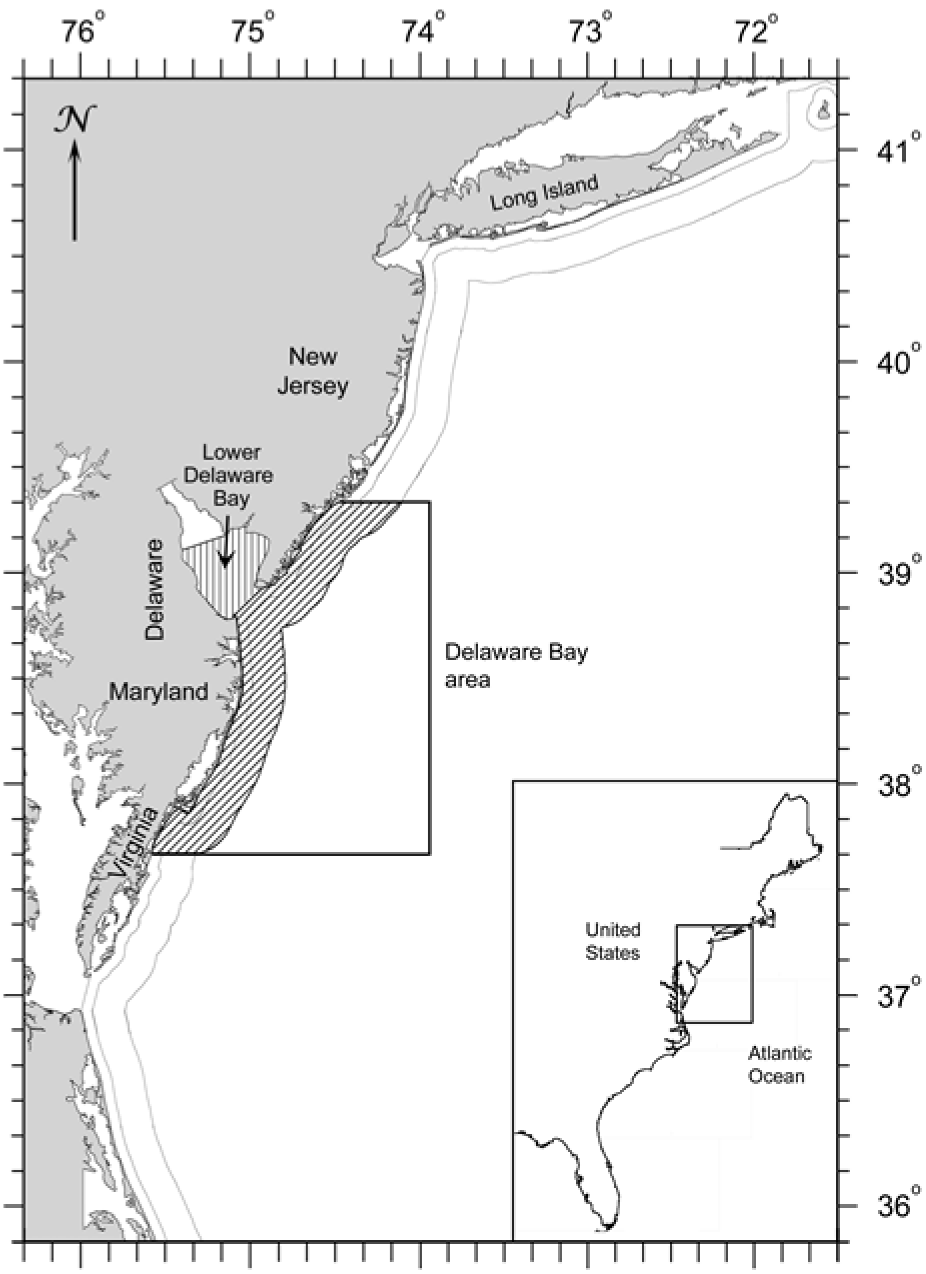


Figure 1: Horseshoe crab trawl survey sampling areas. Contours indicate three and 12 nautical miles from shore.

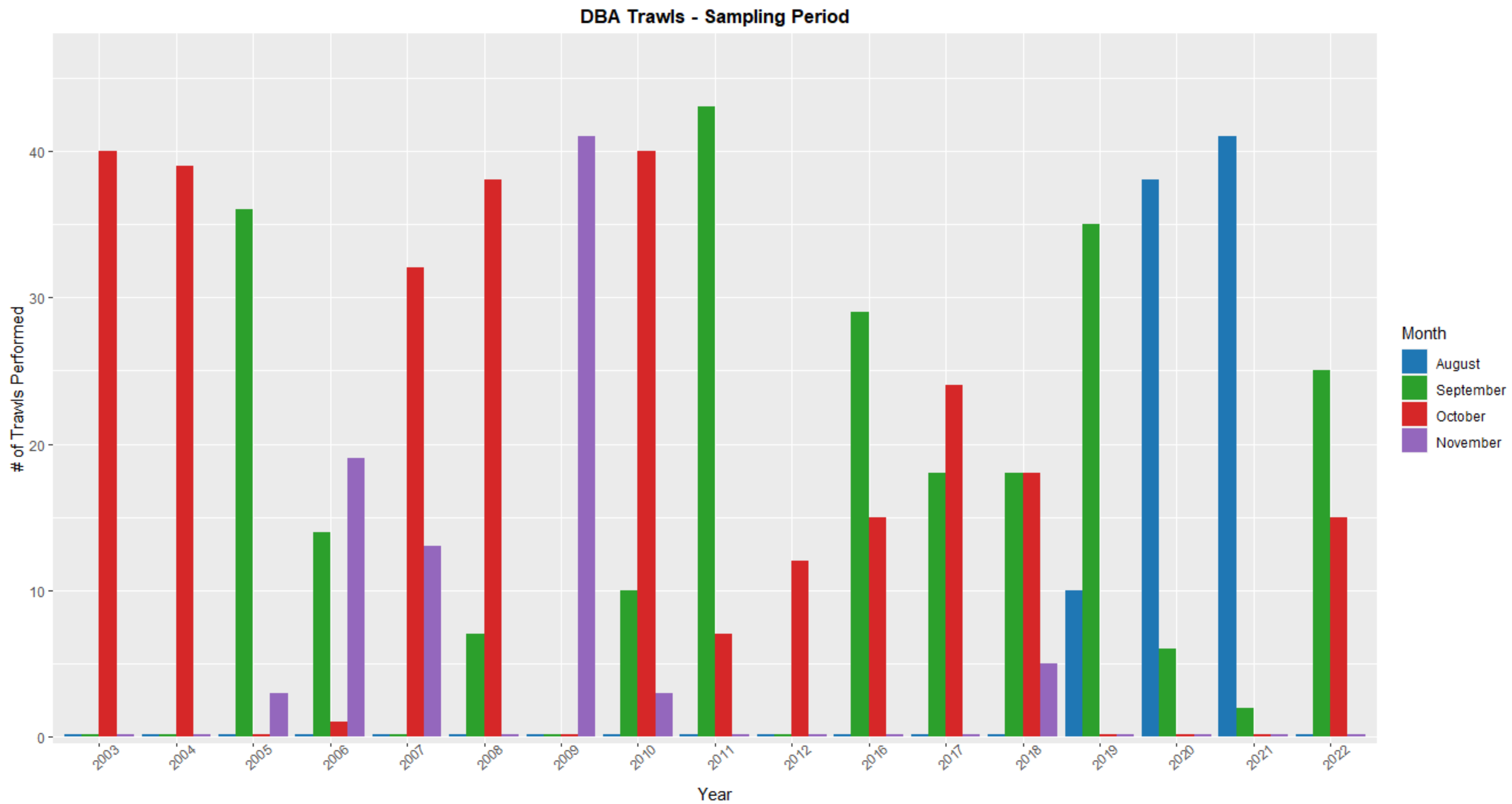


Figure 2: Counts of trawls performed in the Delaware Bay area for each month and year between 2003 and 2022 (*n*=697 trawls).

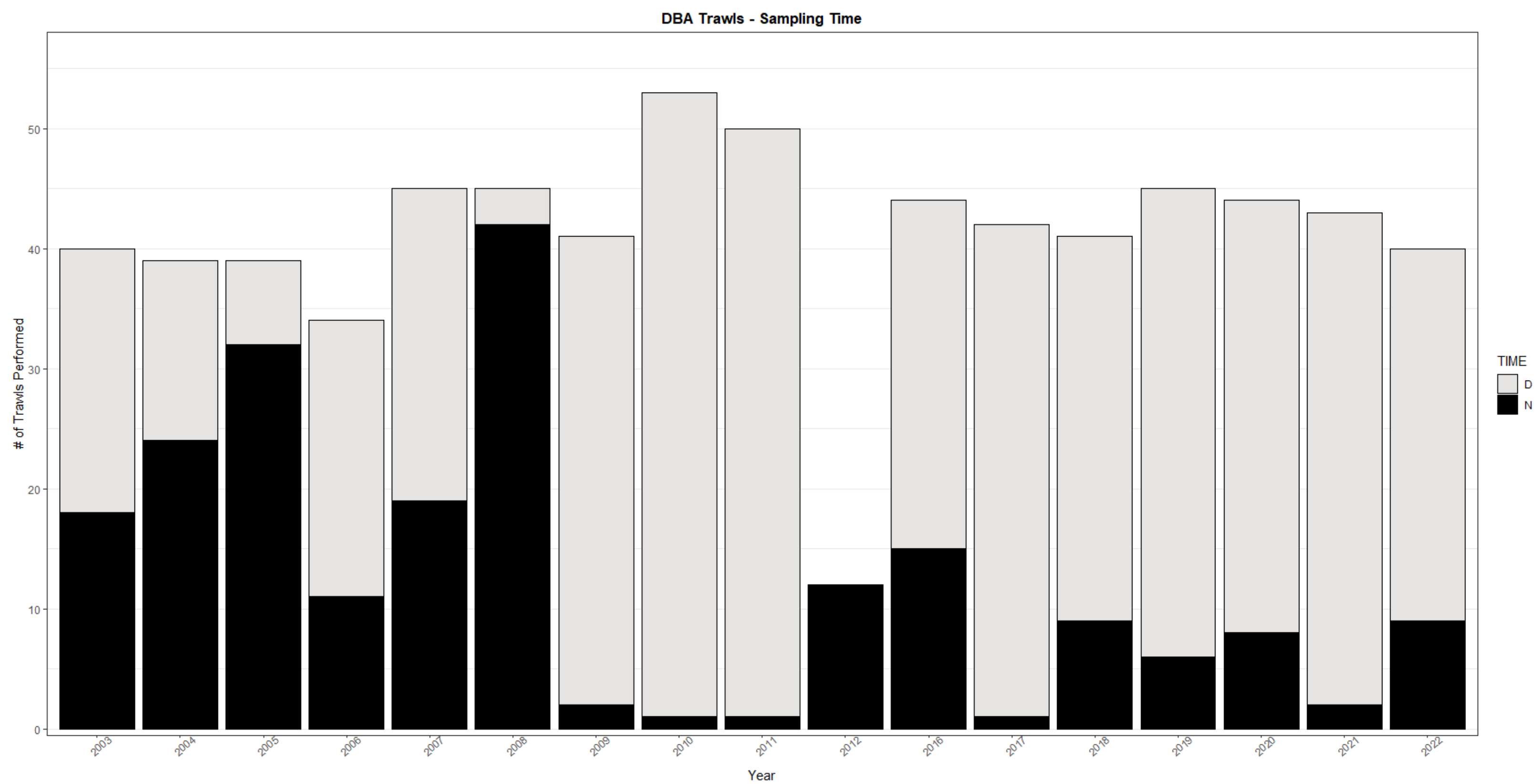


Figure 3: Stacked bar chart showing the number of trawls performed in the Delaware Bay area for each year during the day and at night between 2003 and 2022 (*n*=697 trawls).

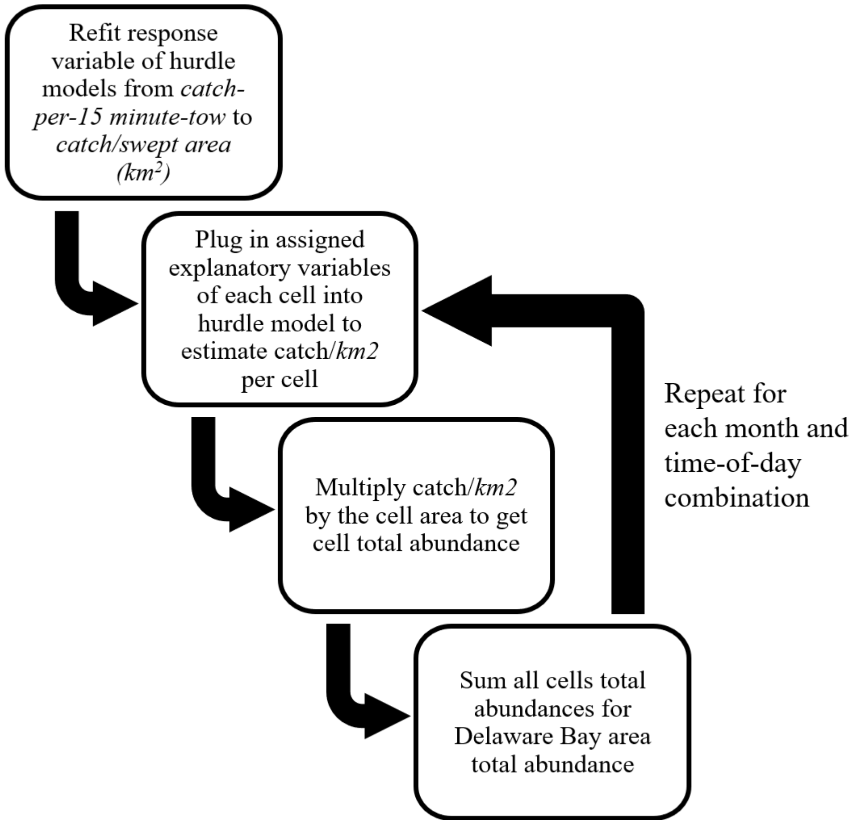


Figure 4: Process for estimating the total abundance of horseshoe crabs in the Delaware Bay area from the hurdle models developed in Wong et al. (Chapter 1).

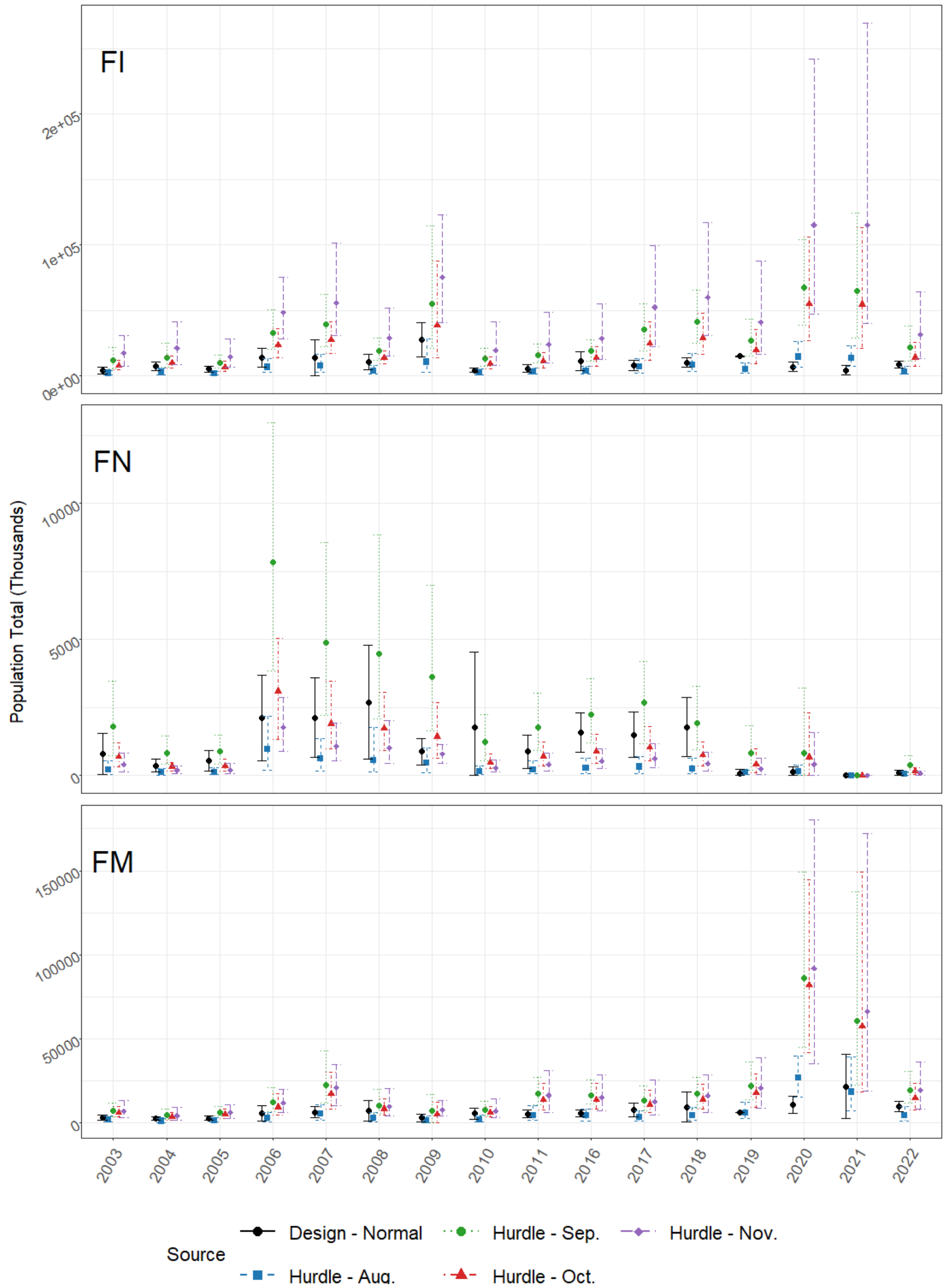


Figure 5: Annual model-based total abundance estimates of female HSC at **night** for each of the four months and their associated 95% quantile interval compared against annual design-based total abundance estimates and their 95% confidence intervals. From top to bottom: immature females, newly-mature females, and mature females.

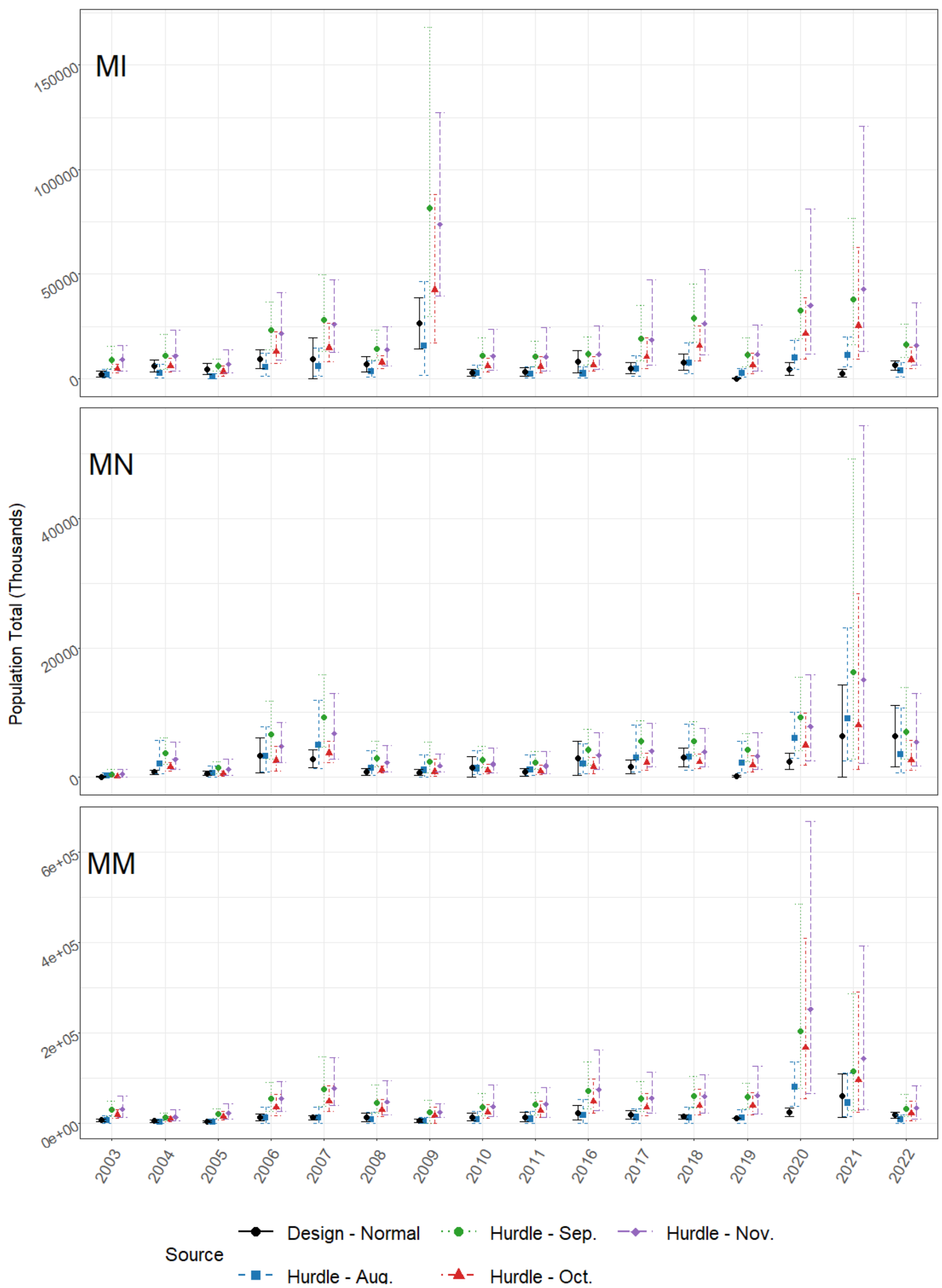


Figure 6: Annual model-based total abundance estimates of male HSC at **night** for each of the four months and their associated 95% quantile interval compared against annual design-based total abundance estimates and their 95% confidence intervals. From top to bottom: immature males, newly-mature males, and mature males.

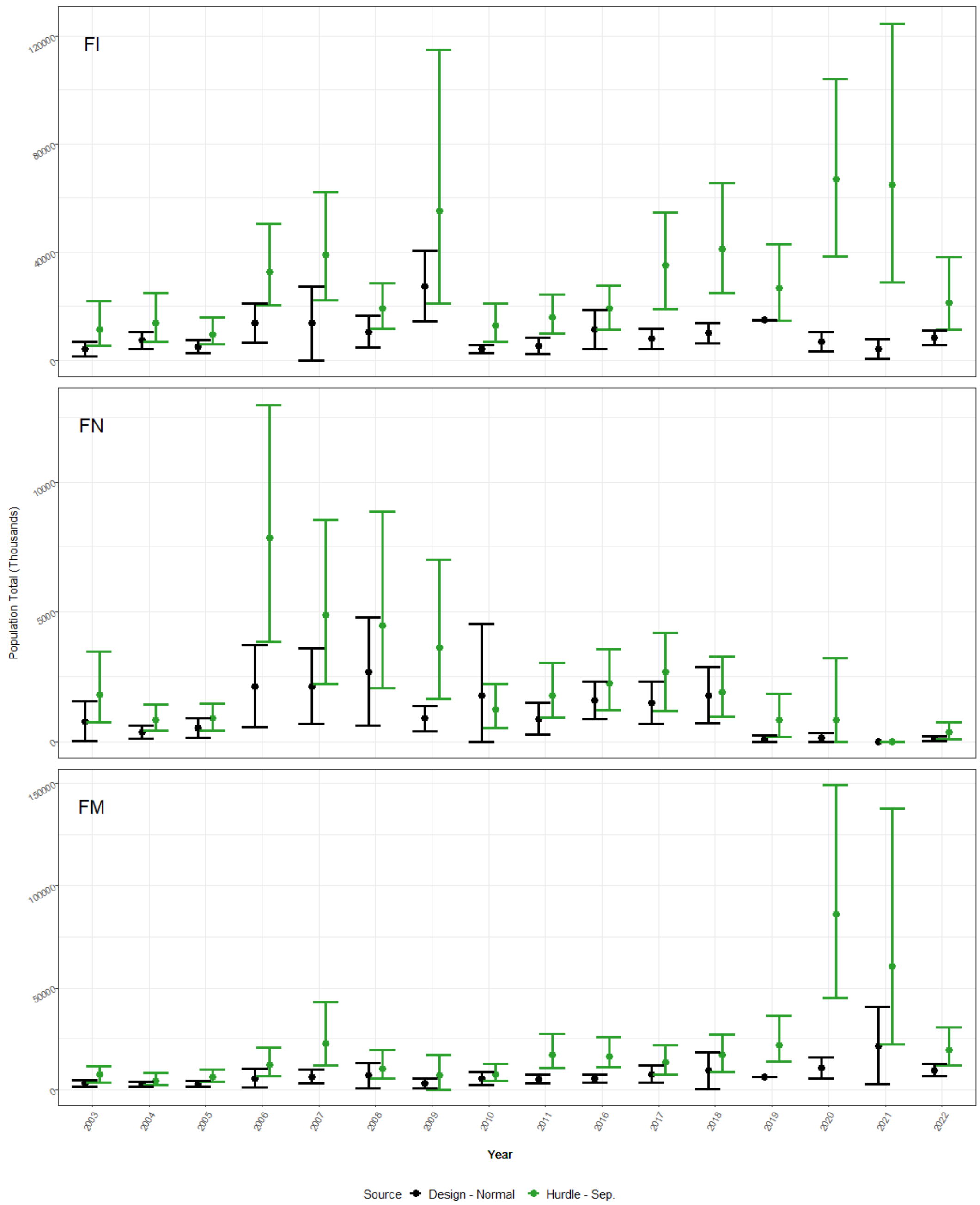


Figure 7: Annual model-based total abundance estimates of female HSC at **night** for **September** and their associated 95% quantile interval compared against annual design-based total abundance estimates and their 95% confidence intervals. From top to bottom: immature females, newly-mature males, and mature males.

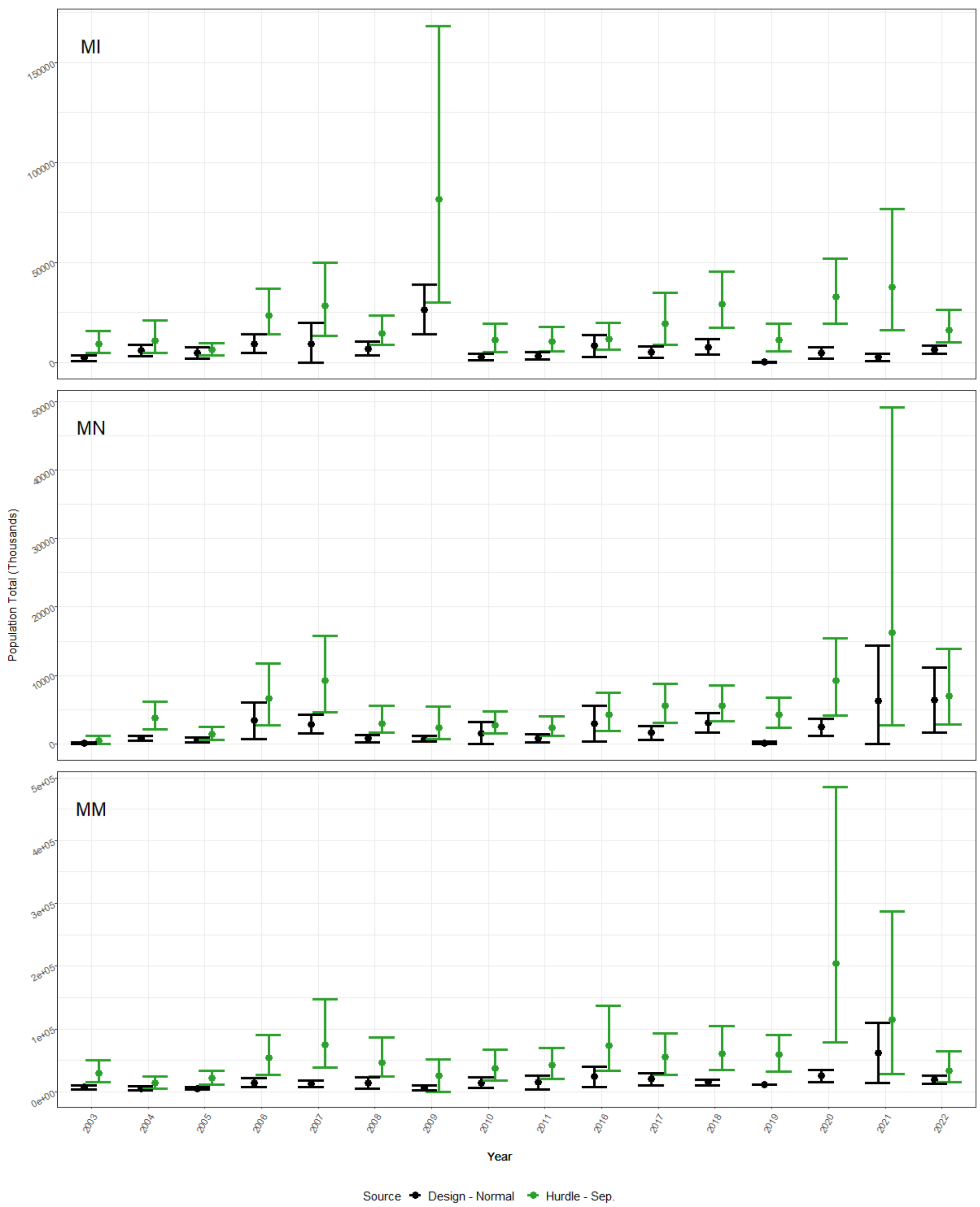


Figure 8: Annual model-based total abundance estimates of male HSC at **night** for **September** and their associated 95% quantile interval compared against annual design-based total abundance estimates and their 95% confidence intervals. From top to bottom: immature males, newly-mature males, and mature males.

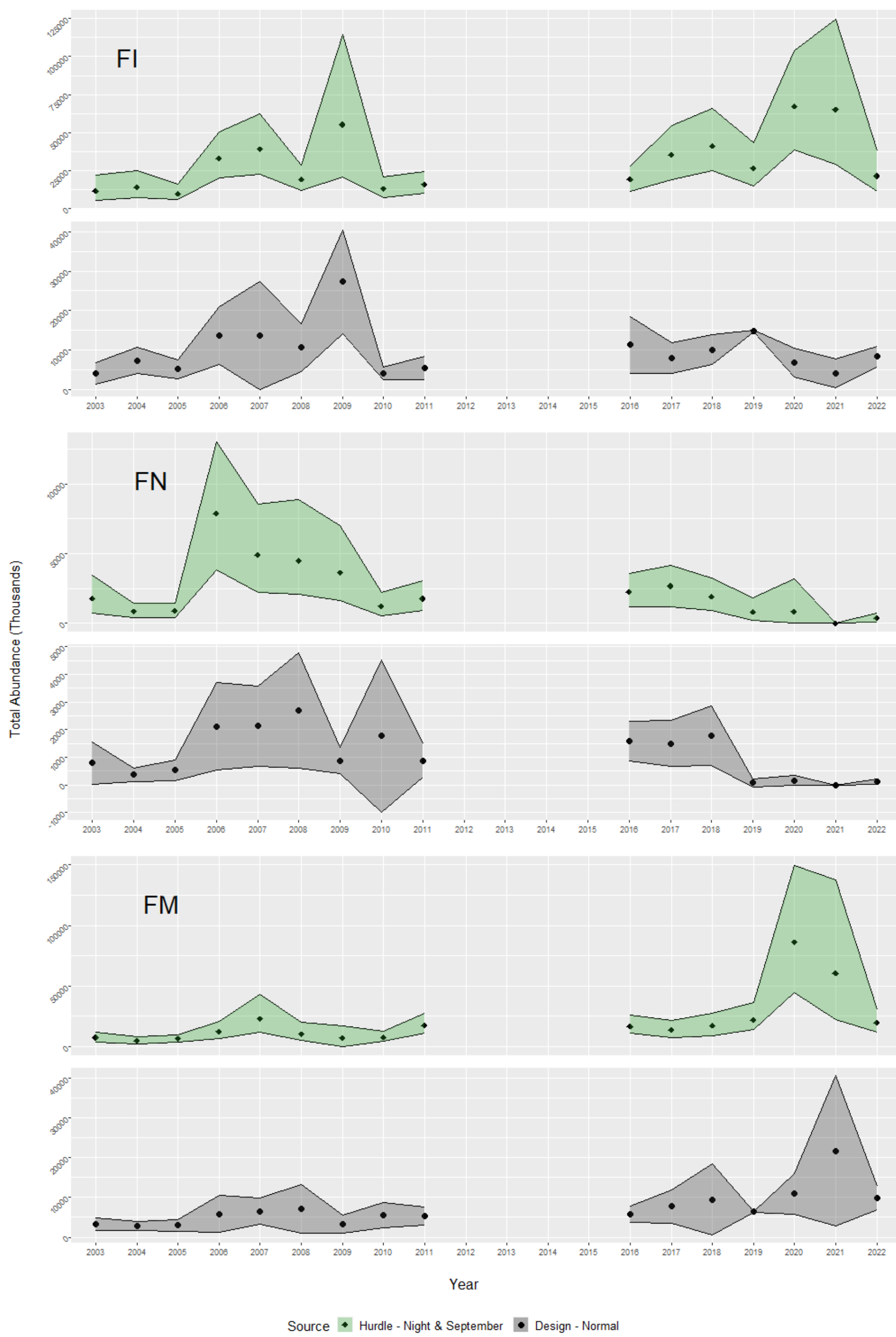


Figure 9: Annual model-based total abundance estimates of female HSC at **night** for **September** and their associated 95% quantile interval compared against annual design-based total abundance estimates and their 95% confidence intervals. From top to bottom: immature females, newly-mature females, and mature females.

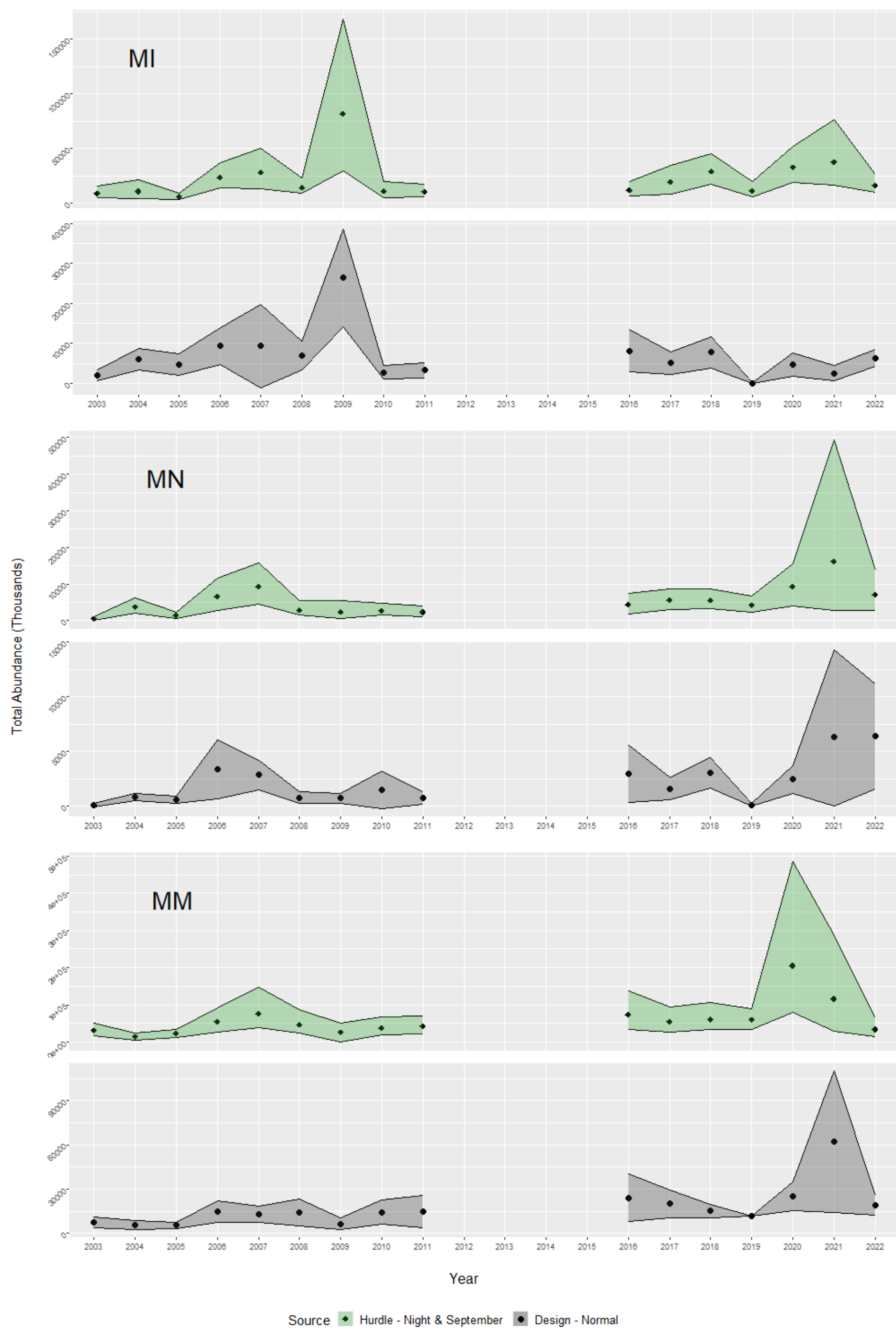


Figure 10: Annual model-based total abundance estimates of male HSC at **night** for **September** and their associated 95% quantile interval compared against annual design-based total abundance estimates and their 95% confidence intervals. From top to bottom: immature males, newly-mature males, and mature males.

CONCLUSIONS

Harvest of American horseshoe crabs (*Limulus polyphemus*) in the Delaware Bay area has gained more attention in recent years as whelk fisheries grow in the mid-Atlantic, synthetic alternatives for LAL become more widely available, and concerns grow around promoting the recovery of the threatened red knot (*Calidrus canutus rufa*) population. In Chapter 1 of this thesis, I showed that while design and model-based estimates of relative abundance of horseshoe crabs were relatively consistent, external spatial, temporal, and environmental factors did have significant impacts on observed catch-per-unit-effort (CPUE) over the time series. These results support the importance of standardizing CPUE in horseshoe crab fisheries management to reflect factors affecting estimates of abundance and to remove them to obtain unbiased estimates of abundance. This chapter also presented one approach for handling highly skewed, zero-inflated data using two-stage hurdle models. Future research and management should focus upon improving model fitting and accuracy by minimizing bias from sampling and observational errors, as well as testing new model designs. If the sampling design was updated using the stratification and sampling time scheme mentioned in Chapter 1, it could improve the accuracy of future abundance estimates and validate the effects of external factors upon CPUE estimated in the current hurdle models. An unknown observational error in classifying newly-mature individuals as mature individuals has been brought up as an area of concern by the Atlantic States Marine Fisheries Commission and researchers. This uncertainty should be further investigated. Future research should also investigate the use of non-linear models, such as generalized additive models (GAMs), to examine the effects of external factors, like latitude and longitude, that likely do not follow the assumed linear relationship with CPUE that the current model suggests.

In Chapter 2, I built hurdle models for three species of bycatch: channeled whelk (*Busysotypus canaliculatus*), knobbed whelk (*Busycon carica*), and summer flounder (*Paralichthys dentatus*). I found that relative abundance estimates for both species of whelk have been less variable since 2016, although recent estimates of channeled whelk are at a historical low, while knobbed whelk were estimated to occur at levels of abundance similar to those seen before 2012. Unfortunately, due to a lack of sampling between 2012 and 2016, I could not determine whether the sampled population followed a similar boom-and-bust trend observed in other whelk populations outside the mid-Atlantic region. The relative abundance of summer flounder since 2016 was estimated to be both more consistent and higher than prior to 2012. There was little similarity in trends of summer flounder abundance among my results and those of NEAMAP's Age Indices or NOAA's ASAP indices. This disparity could be due to a variety of reasons, including differences in sampling area, time, and gear. While the species is managed as a coastwide unit and acts as one genetic population, regional estimates like that presented in Chapter 2 might reveal local fluctuations in population abundance that are obscured in large-scale studies like those of NEAMAP and NOAA. That said, catchability has not been evaluated for this study, and these three species were not directly targeted in our efforts. Although these results provide an initial estimate of whelk relative abundance in the mid-Atlantic and add to the growing knowledge of summer flounder population dynamics along the coast, separate surveys targeting these species may be necessary for more accurate estimates of relative abundance.

The third chapter of this work returned the focus to horseshoe crabs as I used the hurdle models developed in Chapter 1 to estimate the total abundance of horseshoe crabs in the Delaware Bay area. While relative abundance allows managers to investigate trends in the population over time, total abundance estimates of newly-mature and mature individuals are necessary to inform the setting of harvest limits. Total abundance was estimated as the summation of all individual cell abundances in the study area, which required inputting the environmental parameters of each cell into the hurdle models. Due to the need for this information, bottom temperature and salinity were estimated at unmeasured cells and months using spatio-temporal kriging.

As month and time-of-day were shown to be significant explanatory variables for all six demographic groups in Chapter 1, eight total abundances were estimated per year based on a combination of time-of-day and month. From these results, I was able to show that night estimates were always higher than the respective day estimates for all maturity groups, and that night-time estimates for September or November were always higher than for the design-based approach. While all these estimates are based on observational data, answering the question of which of the eight combinations would act as the single "true" total abundance for management was not easy. I decided to regard night-time-September estimates as the "true" total abundance because these two factors had the highest coefficient estimates of all possible levels within the time-of-day and month categories, respectively. Using these values resulted in a third of all model-based estimates being significantly higher than the corresponding design-based estimates. As harvest limits are recommended by the ASMFC based on both the total abundances of newly-mature and mature horseshoe crabs *and* the needs of migratory shorebirds, it would be necessary for these new model-based estimates to be input into the ASMFC's CMSA and ARM models before larger harvest limits could be suggested based on these results. Future simulation studies should be conducted to test what sampling scheme supports the most accurate estimates of relative and total abundance when using a design or model-based approach. This type of study would allow future researchers to determine which of the eight model-based total abundance estimates most accurately resembles the simulated total abundance.

In conclusion, the findings of these chapters and the thrust of my thesis more generally highlight the necessity for continued monitoring of American horseshoe crabs in the Delaware Bay area. The future research and design recommendations suggested here are aimed at minimizing error and bias in sampling and modeling. Through continued annual surveying and assessment, managers will not only achieve a better understanding of horseshoe crab population dynamics, but the additional data will allow researchers to further refine and improve the models developed in this work for achieving more accurate estimates of horseshoe crab abundance in this critical region.

APPENDIX A

All material associated with Chapter 1.

Dealing with missing environmental data

Of the 697 unique trawls performed in the DBA between 2003 and 2022, 25 (3.59%) were missing one or more of the following environmental factors: bottom temperature (BOTTEMP), bottom salinity (BOTSAL), or end depth. All trawls missing BOTTEMP were also missing BOTSAL. A significant correlation was found between BOTTEMP and BOTSAL with ordinal date both inshore (0.74, 0.68), and offshore (0.75, 0.65) respectively. A forward stepwise regression using a generalized linear model with a Gaussian distribution was used to determine what factors best determined the missing variables inshore and offshore. To consider temporal variability and spatial variability, this best-fitting model was refit using four different “neighborhood” datasets. The first data set considers only spatial proximity, selecting the five nearest trawls from the missing data trawl based on the midpoint. This procedure was narrowed temporally by selecting the five nearest trawls within ±7 ordinal days, ±14 ordinal days, and ±21 ordinal days. The best-fitting linear regression model was also compared to a model in which the environmental parameter was estimated using only the mean of each “neighborhood” dataset. The best model for estimation was selected based on the lowest percent root mean square error (%RMSE) and lowest percent coefficient of variation (%CV). For BOTSAL, both the linear regression and estimation based upon the mean resulted in %CV of less than 10% in all cases. Because the linear regression approach predicted values ±5 PPT than those measured in the area, we used the mean values of our best-fitting “neighborhood” datasets to estimate missing BOTSAL. As the best linear regression model for bottom temperature was based on bottom salinity, and all trawls missing bottom temperature were also missing bottom salinity, the mean value of the best-fitting nearby trawls was used.

Dealing with missing end depths

Three methods were considered for estimating missing end depth: 1) Using NOAA ETOPO data, 2) Using the end depth of nearby trawls, and 3) Using the starting depth of trawls. The NOAA ETOPO database is a global relief model of topographic and bathymetric data of 15 arc-second resolution (NOAA NCEI 2022). When compared to previously measured end depths, the mean absolute difference was 5.22 ft (SD: 4.97 ft). Using the end depth of the nearest neighboring trawl resulted in a mean difference between true and estimated of 4.21 ft (SD: 5.87 ft). Lastly, using the beginning depth measurement as the estimated end depth resulted in a mean difference of 3.32 ft (SD: 3.67 ft). As it had the lowest standard deviation, the beginning depth was used to fill in missing end depths.

Bottom salinities of 2022

In all years prior to 2022, BOTSAL was recorded in parts-per-thousand (ppt). During 2022 sampling, bottom salinity was recorded in units of specific gravity. By combining the international one-atmosphere equation of the state of seawater developed by Millero and Poisson (1981) (Formula 1) and the equation of specific gravity for saline water (Reid 2008), we were able to estimate the salinity of the trawls given known temperature and specific gravity. This resulted in a BOTSAL ranging from 23.61 – 31.76 ppt, with a mean of 25.80.

Bottom salinity adjustments

When looking at both inshore and offshore bottom salinity measurements in the DBA, there is a large difference between recorded values in 2020 – 2022 and those earlier in the study.

While this may be expected of areas within the lower Delaware Bay, this was surprising to find for the coastal Delaware Bay area. When dividing the DBA into inshore and offshore trawls and comparing to the nearest five neighboring trawls spatially and temporally, it was found that bottom salinity measurements during these years were consistently lower than previous trawls by an average of -6.54 ppt (Table S1). We suspect that this is possibly due to reading or equipment error as well as the transformation from specific gravity to salinity for 2022. Hence, the mean difference shown in Table S1 was added to each trawl between 2020 – 2022, grouped by year and depth, as a candidate correction factor to test how salinity changed the results of this study using an unadjusted and adjusted bottom salinity scenario (Figure S2).

Distance from shore measurements

Distance from shore values were calculated using ArcGIS Pro's *near* tool feature and NOAA NCEI's GSHHG shoreline database. The specific shape file, GSHSS_f_L1, contains the highest-resolution shoreline file available at a resolution of 1:250,000 (0.04km), last updated June 15, 2017 (Wessel and Smith 1996). Using a WGS 1984 coordinate system, the value was calculated as the distance of the trawl midpoint to the closest point on the shoreline. At this scale, geodesic and planar methods for calculating distance resulted in the same value.

Table S1: Number of trawls performed each year based on **DEPTH** (inshore/offshore) and the mean difference between the recorded bottom salinity values and their nearest five neighbors in units of parts per thousand (PPT).

	DBA		
Year	DEPTH	# of Trawls	Mean Diff.
2020	IN	20	-6.26
2020	OFF	24	-6.86
2021	IN	24	-6.20
2021	OFF	19	-7.18
2022	IN	7	-6.17
2022	OFF	23	-6.56

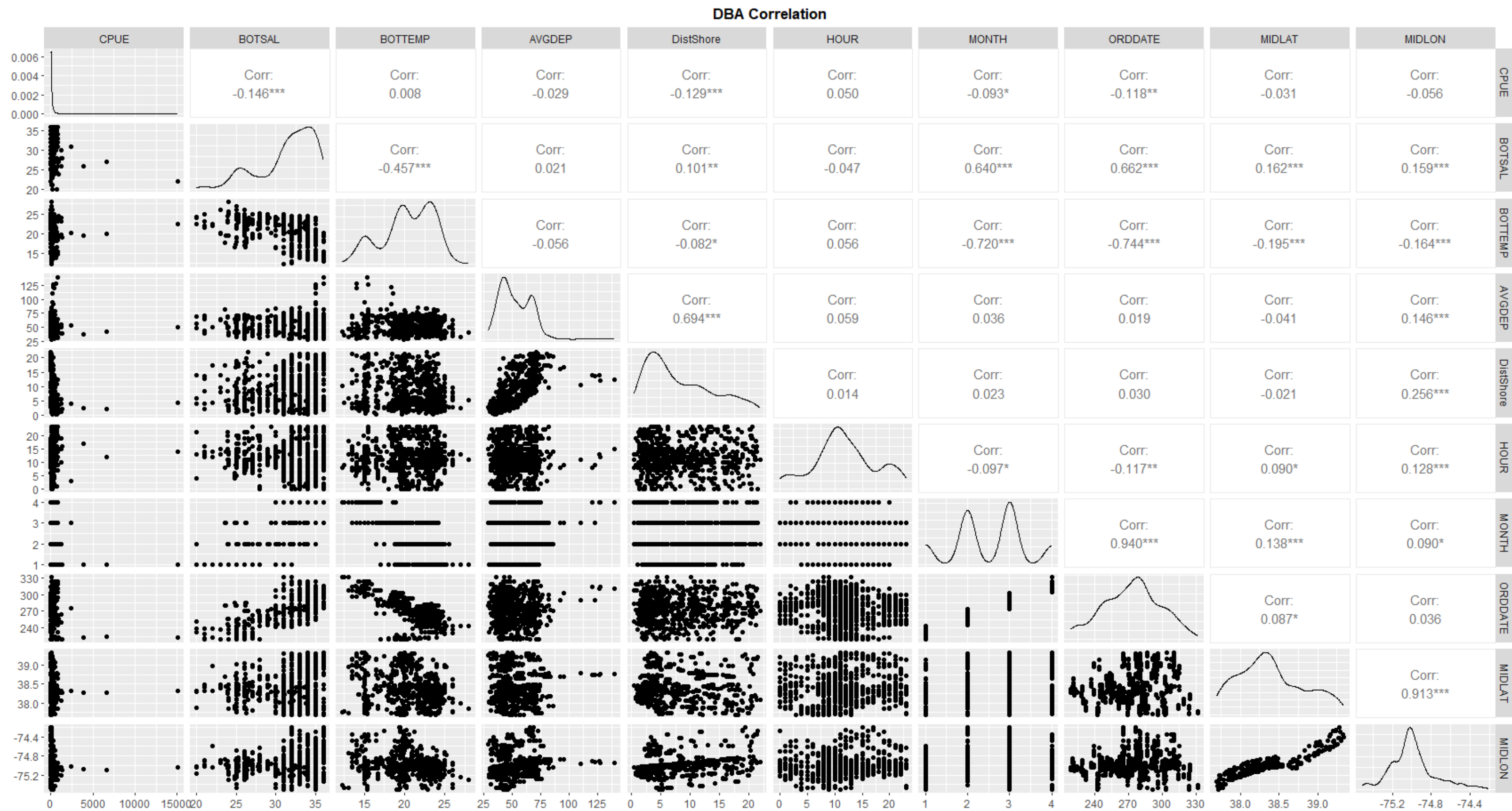


Figure S1: Pairs plot showing Pearson's correlation values and their associated p-values (H_a = true correlation is not equal to 0), density plots along the diagonal, and scatter plots for the 697 trawls performed between 2003 and 2022 in the Delaware Bay area both inshore and offshore after estimating missing environmental factors.

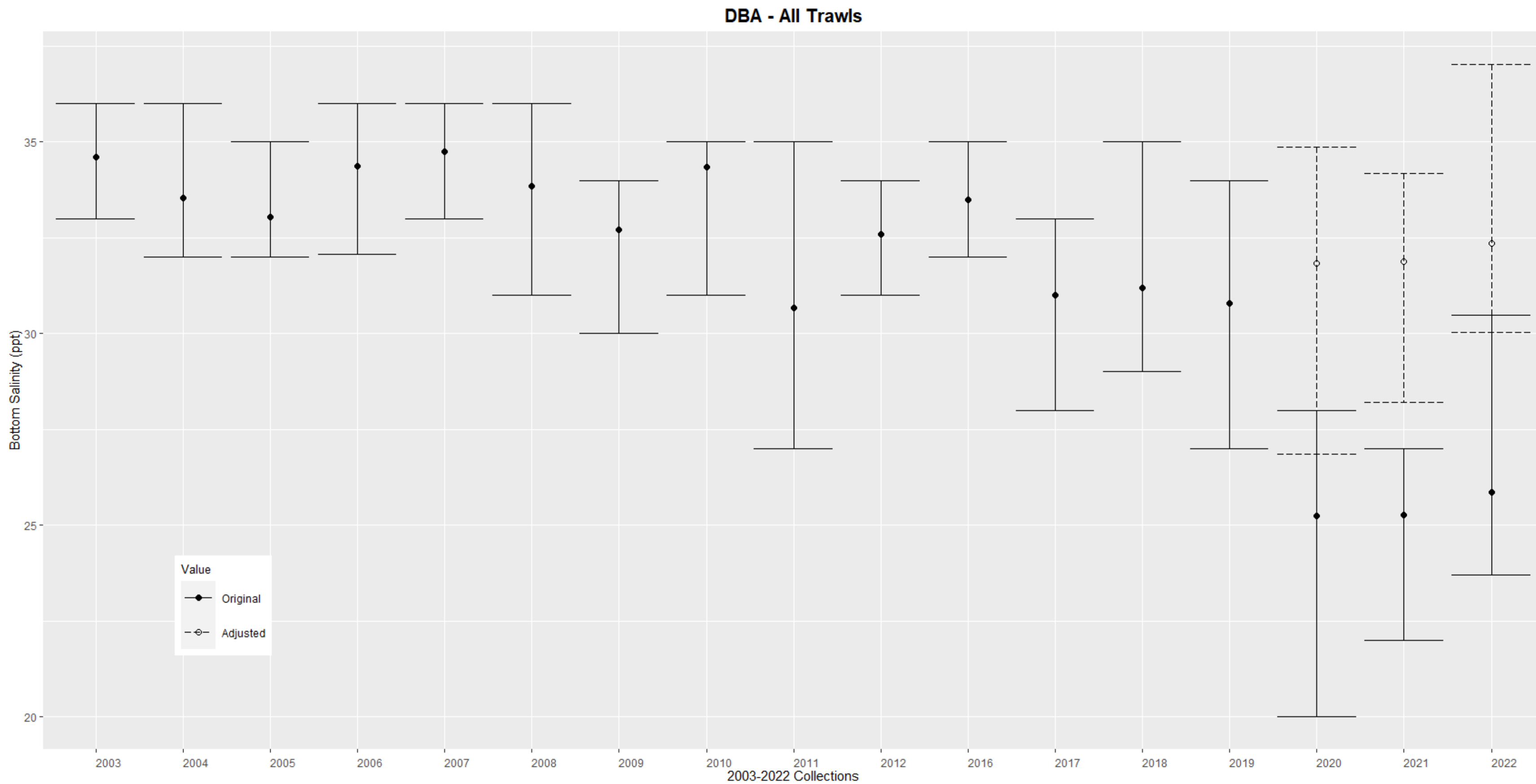


Figure S2: The 95% quantile ranges for bottom salinity measurements (ppt) per year for offshore trawls within the Delaware Bay area (DBA) across the study period and their associated means.

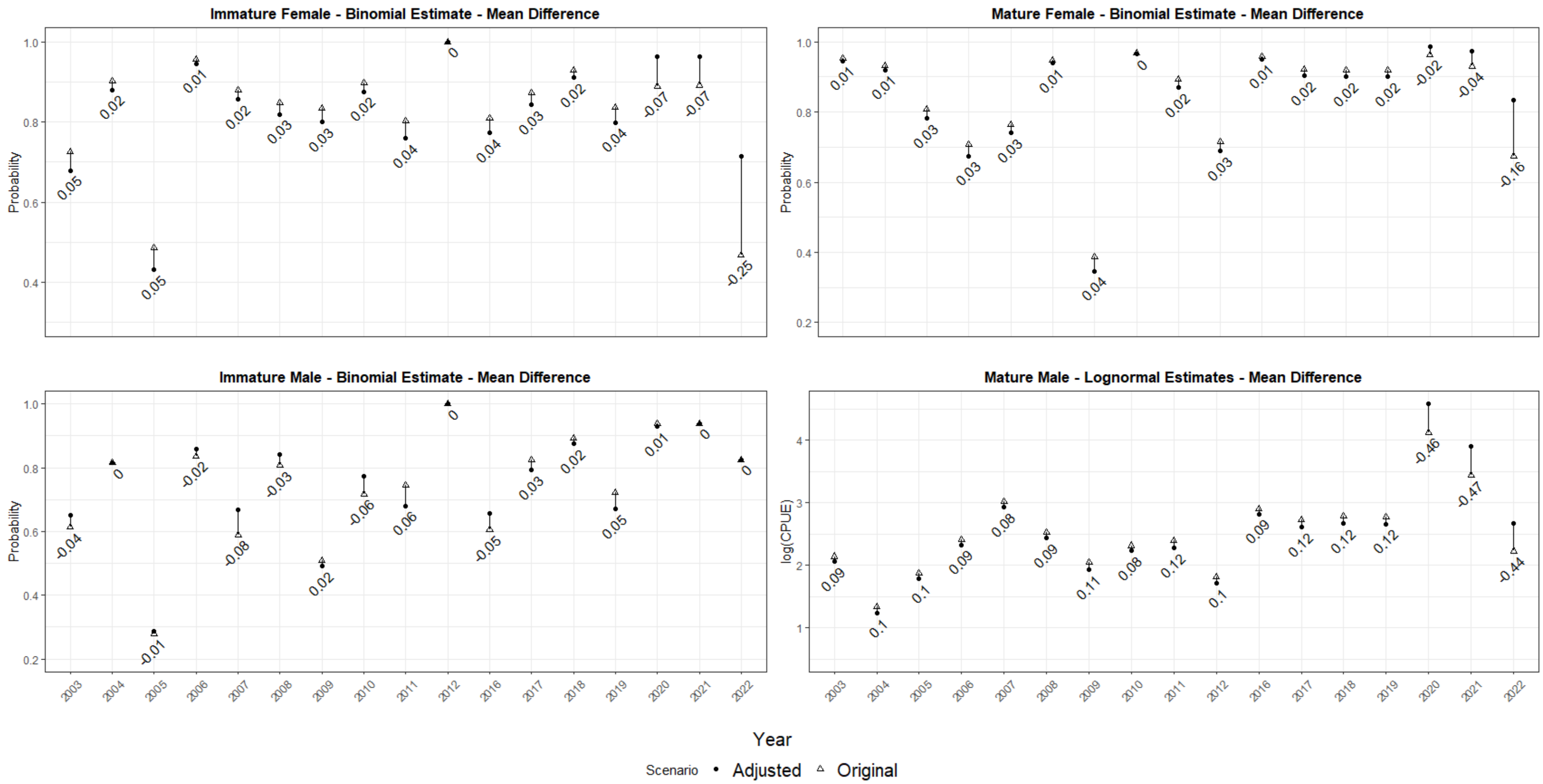


Figure S3: The four demographic groups and respective stages of each hurdle model that had bottom salinity as an explanatory variable. Values show the original estimated mean value for that stage minus the adjusted estimated mean value based on a 1000-iteration bootstrap analysis.

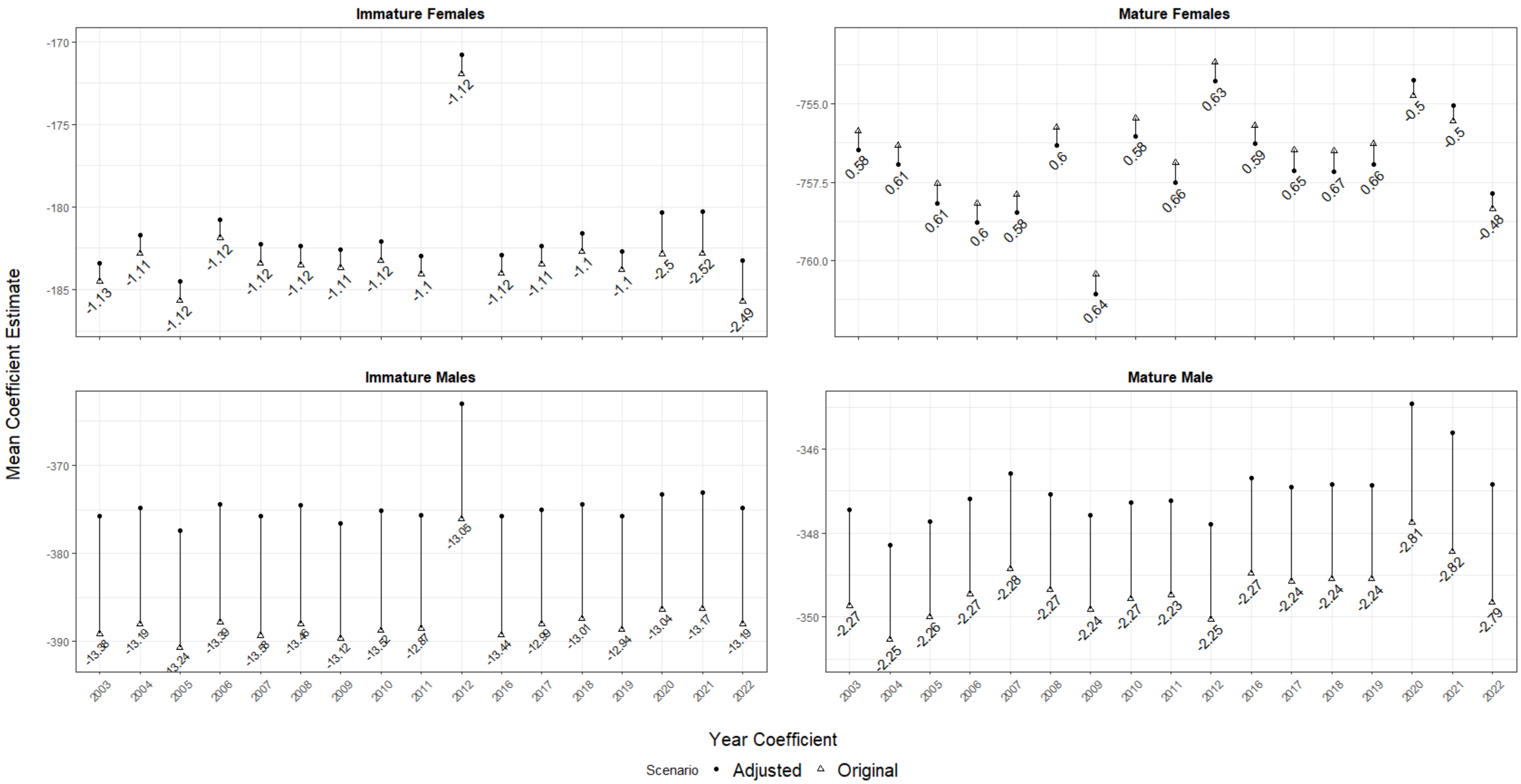


Figure S4: The four demographic groups and respective stages of each hurdle model that had bottom salinity as an explanatory variable. Values show the original estimated mean year coefficient values for that stage minus the adjusted estimated mean year coefficient value based on a 1000iteration bootstrap analysis.

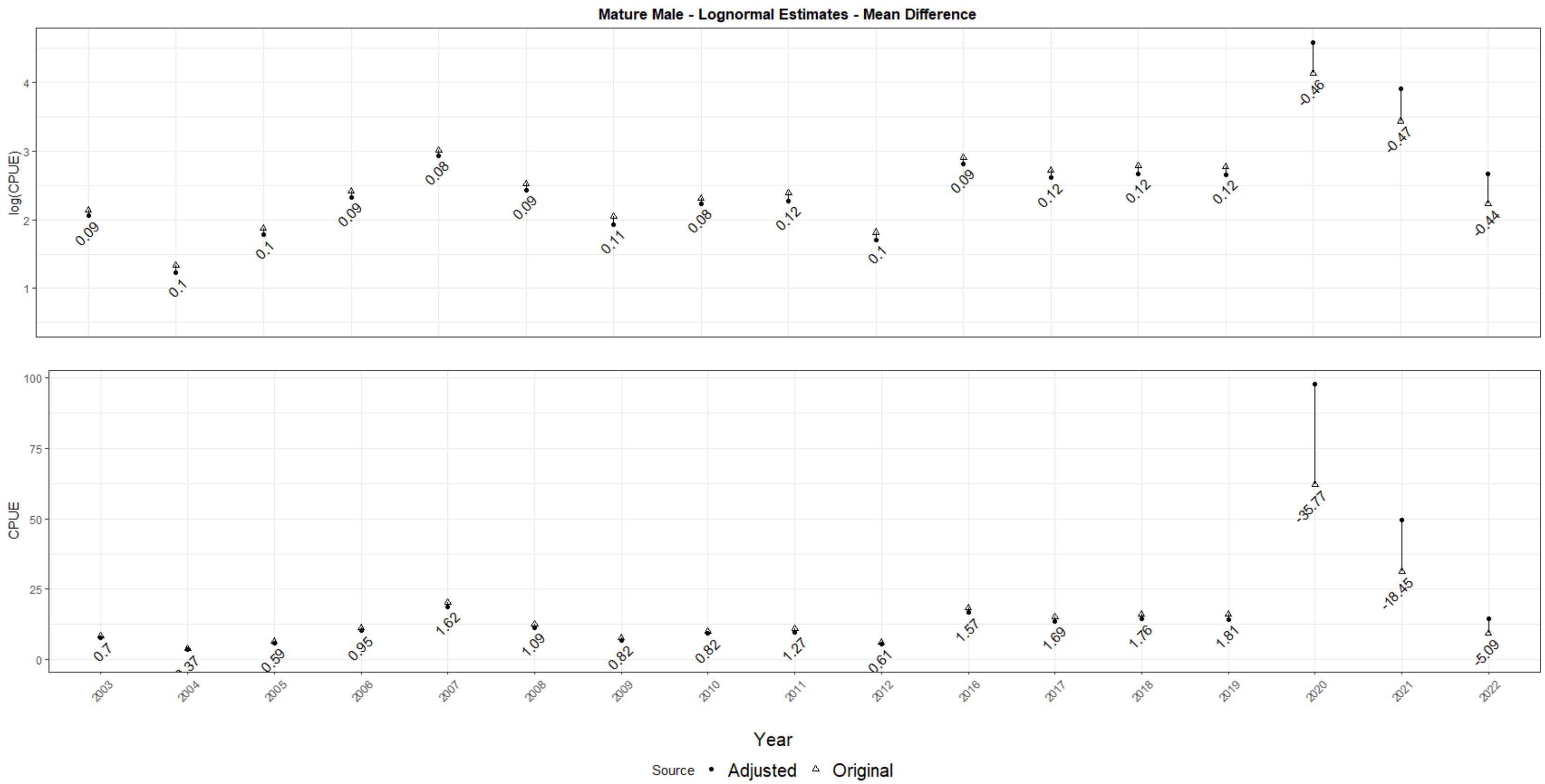


Figure S5: Graph showing the estimate from the lognormal stage of the hurdle model (without correction factor) for mature males in the original models and adjusted models to show the effects of back transforming on the difference between the scenarios. Values show the original correction factor per year minus the adjusted correction factor per year value based on a 1000-iteration bootstrap analysis.

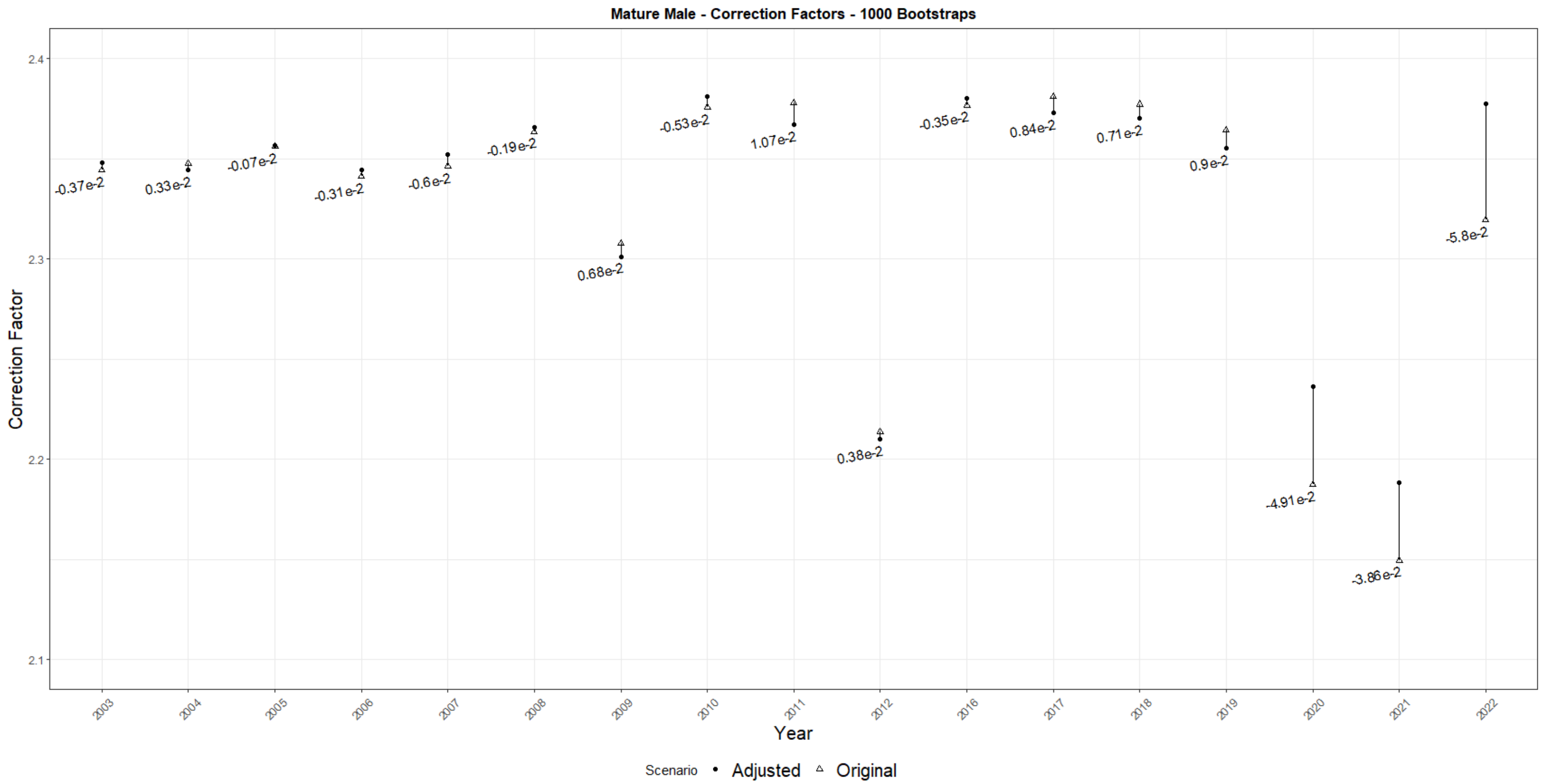


Figure S6: Graph showing the different correction factors applied to the lognormal stage per year for mature males to estimate the year effect in the original models and adjusted models. Values show the original correction factor per year minus the adjusted correction factor per year value based on a 1000-iteration bootstrap analysis.

APPENDIX B

All material associated with Chapter 1

Formula 1: International one-atmosphere equation of state of seawater (Millero and Poisson 1981).

$$p = p_0 + AS + BS^{3/2} + CS$$

Where:

$$A = 8.24493 * 10^{-1} - 4.0899 * 10^{-3}t + 7.6438 * 10^{-5}t^2 - 8.2467 * 10^{-7}t^3 + 5.3875 * 10^{-9}t^4$$

$$B = -5.72466 * 10^{-3} + 1.0227 * 10^{-4}t - 1.6546 * 10^{-6}t^2$$

$$C = 4.8314 * 10^{-4}$$

$$p_0 = 999.842594 + 6.793952 * 10^{-2}t - 9.095290 * 10^{-3}t^2 + 1.001685 * 10^{-4}t^3 - 1.120083 * 10^{-6}t^4 + 6.536332 * 10^{-9}t^5$$

Where:

p is the density of seawater. S is salinity. p_0 is the density of pure water. t is temperature in degrees centigrade.

Formula 2: Relationship between specific gravity, salinity, and temperature at the ocean surface (Reid 2008)

$$\text{Specific Gravity} = \left(p_s(T)/p_0(15.56)\right) * 1000$$

Where:

$p_s(T)$ is the density of the water sample (S) at selected temperature (T).

$p_0(15.56)$ is the density of pure water (0) at reference temperature 15.56 °C.

Formula 3: Normal distribution estimates for mean and variance. (Lohr 2019 chap. 3)

$$y_{hj} = \sum \frac{y'_{hj}}{m_{hj}} * 15$$

Where:

y'_{hj} is the total number of crabs collected in trawl j in stratum h. m_{hj} is the total minutes trawled in trawl j in stratum h. y_{hj} is the standardized CPUE per 15-minute trawl.

$$\bar{y}_h = \frac{\sum y_{hj}}{n_h}$$

Where:

n_h is the total number of cells sampled in stratum h. $\bar{y}_h$ = mean CPUE of stratum *h*.

$$\bar{y}_{str} = \sum_{h=1}^{H} \frac{A_h}{A} \bar{y}_h$$

Where:

A_h is the total area of stratum h. A is the total sampled area. $\bar{y}_{str}$ is the stratified mean of CPUE.

Formula 4: Delta-lognormal distribution estimates for mean and variance (Aitchison 1955; Pennington 1983)

$$y_{hj} = \ln(\frac{y'_{hj}}{m_{hj}} * 15)$$

$$c = \begin{cases} \frac{m}{n} exp(\bar{y}) G_m\left(\frac{1}{2}s^2\right), & m > 1 \\ \frac{x_1^2}{n}, & m = 1 \\ 0, & m = 0 \end{cases}$$

$$\bar{y}_{st} = \sum_{h=1}^{H} \frac{A_h}{\sum A_h} * c$$

$$G_m(t) = 1 + \frac{n-1}{n} t + \sum_{j=2}^{\infty} \frac{(n-1)^{2j-1}}{n^j + (n+1)(n+3) \ldots (n+2j-3)} * \frac{t^j}{j!}$$

$$d = \begin{cases} \frac{m}{n} exp(2\bar{y}) \left\{G_m(2s^2) - \left(\frac{m-1}{n-1}\right) G_m\left(\frac{m-2}{m-1}s^2\right)\right\}, & m > 1 \\ \frac{x_1^2}{n}, & m = 1 \\ 0, & m = 0 \end{cases}$$

$$V(\bar{y}_{st}) = \sum_{h=1}^{H} \left(\left(\frac{A_h}{\sum A_h}\right)^2 * \left(\frac{d}{n_h}\right) * \left(1 - \frac{n_h}{N_h}\right) \right)$$

Where:

m is the number of positive trawls, n is the number of total trawls, $\bar{y}$ is the sample mean, s^2 is the sample variance of non-zero tows after log transforming, t is the number of positive trawls, x_1 is the number of crabs caught. G_m is defined as the correction function for the log-normal distribution below.

Formula 5: Structure of a hurdle model (Cragg 1971; Maunder and Punt 2004).

$$\Pr(Y = y) = \begin{cases} w, & y = 0 \\ (1-w)f(y), & y > 0 \end{cases}$$

Where:

w is your probability of a zero observation, and $f(y)$ is the mean value of the log-normal model.

Formula 6: Extracting year effect from our hurdle model

$$\widehat{d_J} = \pi * \left(exp(\widehat{\beta_0} + \widehat{\beta_{\iota J}}) * \psi_{dj}\right)$$

Where:

$\widehat{d_J}$ is the year effect for year j, π is the probability of having a positive observation, $\psi_{dj}$is the correction bias when assuming a lognormal error structure.

$$\psi = g_m\left[\frac{m+1}{2m}\left(\widehat{\sigma^2} - \widehat{\sigma_\eta^2}\right)\right]$$

$$\hat{\eta} = \widehat{\beta_0} + \widehat{\beta_{\iota J}}$$

$$g_m(t) = \sum_{p=0}^{\infty}\left[\frac{m^p(m+2p)}{m(m+2)\ldots(m+2p)}\left(\frac{m}{m+1}\right)^p\frac{t^p}{p!}\right]$$

Where:

$\widehat{\sigma^2}$ is the residual variance for the lognormal model, m is the residual variance degrees of freedom.

APPENDIX C

All material associated with Chapter 1

Table S2: Stratified mean catch-per-tow of horseshoe crabs in the **Delaware Bay Area** survey area in 2010-2022, with the mean, standard error (SE), and upper (UCL) and lower (LCL) limits of a 95% quantile interval from a bootstrap analysis calculated using a **hurdle model** based on the **unadjusted-scenario** dataset by demographic group.

	YEAR	MEAN	LCL	UCL	SE		YEAR	MEAN	LCL	LCL	SE
FI						MI					
	2003	5.87	2.56	10.04	1.89		2003	4.58	1.72	8.04	1.64
	2004	8.85	4.85	14.54	2.46		2004	7.36	3.44	13.46	2.6
	2005	3.52	0.89	7.18	1.64		2005	1.64	0.44	3.73	0.88
	2006	16.63	10.47	25.82	3.83		2006	11.24	5.57	19.16	3.5
	2007	22.67	11.95	37.67	6.62		2007	13.03	5.16	24.69	5.14
	2008	11.34	5.82	17.65	2.96		2008	9	4.87	14.41	2.48
	2009	32.79	11.95	67.45	14.5		2009	35.81	7.3	84.65	20.7
	2010	8.38	4.75	13.14	2.17		2010	6.74	3.07	12.01	2.37
	2011	8.66	4.64	13.71	2.32		2011	5.91	2.9	9.8	1.79
	2016	10.25	5.5	16.69	2.84		2016	5.42	2.31	9.29	1.77
	2017	23.08	12.92	37.65	6.33		2017	13.38	6.41	24.19	4.38
	2018	26.23	15.83	40.79	6.27		2018	19.6	10.85	32.04	5.33
	2019	15.68	8.01	25.62	4.62		2019	6.65	2.89	11.6	2.28
	2020	47.62	23.49	79.06	14.16		2020	27.42	14.72	44.78	7.73
	2021	42.32	18.62	82.15	16.44		2021	27.41	11.56	55.89	11.13
	2022	8.27	2	16.96	3.75		2022	11.7	6	19.42	3.38
FN						MN					
	2003	1.8	0.66	3.48	0.72		2003	0.12	0.01	0.35	0.09
	2004	0.79	0.34	1.38	0.28		2004	2.96	1.28	5.17	0.97
	2005	0.8	0.36	1.5	0.29		2005	0.51	0.15	1.11	0.25
	2006	5.65	3.06	9.21	1.6		2006	3.22	1.05	6.2	1.32
	2007	4.59	2.32	7.44	1.38		2007	7.1	2.75	13.4	2.67
	2008	4.82	2.47	8.25	1.55		2008	1.75	0.63	3.41	0.74
	2009	4.26	2.08	7.41	1.42		2009	1.13	0.21	3.24	0.83
	2010	1.34	0.73	2.31	0.39		2010	1.93	0.83	3.35	0.67
	2011	1.9	0.93	3.23	0.59		2011	1.66	0.69	2.95	0.57
	2016	2.22	1.24	3.44	0.56		2016	2.18	0.77	4.2	0.86
	2017	2.74	1.26	4.66	0.92		2017	4.79	2.6	7.58	1.3
	2018	1.87	0.93	2.98	0.53		2018	5.43	3.05	8.52	1.38
	2019	0.55	0.13	1.29	0.28		2019	3.08	1.5	5.32	0.98
	2020	0.52	0.05	1.63	0.45		2020	10.22	4.56	18.19	3.71
	2021	0	0	0	0		2021	10.38	1.35	35.36	9.94
	2022	0.19	0.03	0.44	0.11		2022	4.83	1.5	9.89	2.28
FM						MM					
	2003	8.16	4.34	13.92	2.43		2003	18.13	8.50	30.98	5.67
	2004	4.96	2.65	8.28	1.49		2004	8.13	3.71	14.57	2.87
	2005	5.97	3.16	9.64	1.63		2005	9.81	2.22	18.71	4.1

2006	7.42	2.76	12.83	2.48	2006	22.87	9.19	39.29	7.24
2007	18.9	7.47	35.37	6.98	2007	30.78	7.02	58.25	13.22
2008	10.69	5.84	18.72	3.34	2008	23.3	8.23	43.4	9.01
2009	4.14	0.00	11.27	3.02	2009	11.02	0.00	25.6	6.87
2010	8.07	5.45	11.39	1.56	2010	22.73	12.44	36.73	6.29
2011	16.19	9.98	23.03	3.39	2011	24.12	14.01	36.83	5.85
2016	15.81	9.83	23.18	3.43	2016	41.13	19.40	71.82	13.22
2017	14.76	8.38	23.96	4.05	2017	35.41	19.26	61.37	10.56
2018	17.76	10.62	26.84	4.03	2018	34.83	18.80	52.88	8.94
2019	22.09	13.93	33.08	4.9	2019	34.67	19.55	54.26	9.08
2020	103.38	55.93	187.35	34.54	2020	149.48	62.12	335.27	74.33
2021	63.47	21.61	141.68	31.68	2021	78.23	21.38	210.44	47.67
2022	17.41	4.92	30.86	6.23	2022	21.84	10.27	39.63	7.66

Table S3: Stratified mean catch-per-tow of horseshoe crabs in the **Delaware Bay Area** survey area in 2010-2022, with the mean, standard error (SE), and upper (UCL) and lower (LCL) limits of a 95% quantile interval from a bootstrap analysis calculated using a **hurdle model** based on the **adjusted-scenario** dataset by demographic group.

	YEAR	MEAN	LCL	UCL	SE		YEAR	MEAN	LCL	LCL	SE
FI						**MI**					
	2003	5.5	2.38	9.42	1.82		2003	4.86	1.89	8.53	1.68
	2004	8.63	4.63	14.25	2.46		2004	7.34	3.42	13.45	2.61
	2005	3.13	0.72	6.76	1.56		2005	1.68	0.43	3.79	0.89
	2006	16.44	10.3	25.66	3.82		2006	11.54	5.91	19.68	3.55
	2007	22.05	11.24	36.65	6.51		2007	14.83	6.09	27.48	5.7
	2008	10.95	5.44	17.14	2.96		2008	9.37	5.22	14.83	2.5
	2009	31.45	10.55	66.34	14.41		2009	34.59	7.29	83.88	20.35
	2010	8.19	4.53	12.73	2.14		2010	7.26	3.32	12.75	2.49
	2011	8.2	4.14	13.26	2.3		2011	5.41	2.38	9.24	1.77
	2016	9.8	4.98	15.81	2.76		2016	5.89	2.72	9.97	1.87
	2017	22.27	12.15	36.46	6.27		2017	12.87	6	23.47	4.34
	2018	25.72	15.39	40.36	6.26		2018	19.2	10.36	31.74	5.3
	2019	14.98	7.4	24.82	4.52		2019	6.17	2.48	10.96	2.2
	2020	51.57	27.83	83.93	14.64		2020	27.18	14.59	44.52	7.69
	2021	45.88	20.69	86.19	17.4		2021	27.4	11.45	56.09	11.12
	2022	12.62	5.38	21.44	4.11		2022	11.71	6.03	19.34	3.34
FN						**MN**					
	2003	1.8	0.66	3.48	0.72		2003	0.12	0.01	0.35	0.09
	2004	0.79	0.34	1.38	0.28		2004	2.96	1.28	5.17	0.97
	2005	0.8	0.36	1.5	0.29		2005	0.51	0.15	1.11	0.25
	2006	5.65	3.06	9.21	1.6		2006	3.22	1.05	6.2	1.32
	2007	4.59	2.32	7.44	1.38		2007	7.1	2.75	13.4	2.67
	2008	4.82	2.47	8.25	1.55		2008	1.75	0.63	3.41	0.74
	2009	4.26	2.08	7.41	1.42		2009	1.13	0.21	3.24	0.83
	2010	1.34	0.73	2.31	0.39		2010	1.93	0.83	3.35	0.67
	2011	1.9	0.93	3.23	0.59		2011	1.66	0.69	2.95	0.57
	2016	2.22	1.24	3.44	0.56		2016	2.18	0.77	4.2	0.86
	2017	2.74	1.26	4.66	0.92		2017	4.79	2.6	7.58	1.3
	2018	1.87	0.93	2.98	0.53		2018	5.43	3.05	8.52	1.38
	2019	0.55	0.13	1.29	0.28		2019	3.08	1.5	5.32	0.98
	2020	0.52	0.05	1.63	0.45		2020	10.22	4.56	18.19	3.71
	2021	0	0	0	0		2021	10.38	1.35	35.36	9.94
	2022	0.19	0.03	0.44	0.11		2022	4.83	1.5	9.89	2.28
FM						**MM**					
	2003	8.1	4.30	13.81	2.42		2003	16.59	8.11	27.76	4.93
	2004	4.9	2.54	8.25	1.48		2004	7.32	3.4	12.97	2.57
	2005	5.76	2.96	9.39	1.62		2005	8.9	2.1	16.63	3.64
	2006	7.08	2.51	12.55	2.46		2006	20.89	8.39	35.46	6.4
	2007	18.26	6.90	34.45	6.81		2007	28.28	6.87	53.81	11.82
	2008	10.6	5.71	18.53	3.32		2008	21.17	7.64	38.94	7.84
	2009	3.71	0.00	10.33	2.83		2009	9.83	0	23.36	6.18
	2010	8.03	5.41	11.34	1.56		2010	20.83	11.78	32.13	5.31
	2011	15.74	9.37	22.76	3.41		2011	21.28	12.47	32.78	5.36

2016	15.7	9.65	23.02	3.44	2016	37.48	18.55	63.71	11.47
2017	14.47	8.24	23.56	4.03	2017	31.42	17.16	53.85	9.54
2018	17.38	10.18	26.72	4.05	2018	31.01	16.57	47.95	8.23
2019	21.62	13.53	32.59	4.89	2019	30.72	17.27	48.96	8.36
2020	105.93	57.59	190.07	35.04	2020	232.89	107.64	479.15	92.45
2021	66.52	22.42	148.96	32.89	2021	124.64	34.33	301.77	69.27
2022	21.52	12.19	33.65	5.47	2022	33.54	20.31	49.97	7.64

Table S4: Stratified mean catch-per-tow of horseshoe crabs in the **Delaware Bay Area** survey area in 2010-2022, with the mean, standard error (SE), and upper (UCL) and lower (LCL) limits of a 95% confidence interval from a bootstrap analysis calculated using a **design-based approach** using a **normal distribution** by demographic group.

	YEAR	MEAN	LCL	UCL	SE		YEAR	MEAN	LCL	LCL	SE
FI						MI					
	2003	9.5	3	15.9	3.1		2003	4.9	1.8	8.1	1.5
	2004	17	9.5	24.5	3.6		2004	14	7.6	20.3	3.1
	2005	11.5	6.1	17	2.6		2005	10.6	4.4	16.7	2.9
	2006	31.1	15.3	46.9	7.5		2006	21.5	11.1	32	5
	2007	29.8	0	59.6	12.2		2007	20.5	0	43.2	9.3
	2008	24.6	10.3	38.9	6.6		2008	15.9	7.6	24.2	3.8
	2009	63.1	32.4	93.8	14.9		2009	61	32.1	89.8	14
	2010	9.4	5.7	13	1.8		2010	6.4	2.6	10.1	1.8
	2011	12.2	6	18.5	3		2011	7.3	3.3	11.2	1.9
	2016	25.1	9	41.1	7.7		2016	18.1	6.3	29.9	5.7
	2017	19.1	9.6	28.7	4.6		2017	12.4	5.5	19.3	3.3
	2018	22.5	14.5	30.6	3.8		2018	17.2	8.6	25.9	4.1
	2019	13.7	5.5	21.9	4.1		2019	6.6	2	11.1	2.2
	2020	18.8	8.7	35.4	6		2020	12.7	4.7	24	4.75
	2021	10.14	1.54	19.2	5.05		2021	6.39	1.83	10.99	2.66
	2022	20.69	14.22	27.17	3.83		2022	16.02	10.65	21.39	3.17
FN						MN					
	2003	1.8	0.1	3.6	0.8		2003	0.2	0	0.5	0.2
	2004	0.8	0.3	1.4	0.3		2004	1.8	1	2.6	0.4
	2005	1.2	0.3	2.1	0.4		2005	1.3	0.5	2.1	0.4
	2006	4.8	1.4	8.2	1.6		2006	7.5	1.8	13.2	2.7
	2007	4.6	1.5	7.7	1.5		2007	6.1	3.2	9.1	1.4
	2008	6.3	1.3	11.3	2.3		2008	1.8	0.5	3.1	0.6
	2009	2	0.9	3.1	0.5		2009	1.6	0.6	2.6	0.5
	2010	4	0	10.3	3		2010	3.3	0	7.2	1.9
	2011	2.2	0.5	3.9	0.8		2011	1.9	0.4	3.5	0.7
	2016	3.5	1.9	5.1	0.8		2016	6.6	0.6	12.6	2.9
	2017	3.6	1.6	5.5	1		2017	3.8	1.3	6.4	1.2
	2018	3.9	1.6	6.2	1.1		2018	6.9	3.9	10	1.5
	2019	0.6	0	1.2	0.3		2019	3.5	1.5	5.5	1
	2020	0.3	0	0.8	0.28		2020	6.9	3.3	10.6	2.1
	2021	0	NA	NA	0		2021	16.33	0	37.39	11.31
	2022	0.29	0.04	0.53	0.13		2022	16.21	3.85	28.57	7.24
FM						MM					
	2003	7.5	4.1	10.9	1.6		2003	17	9.4	24.7	3.6
	2004	6	3.7	8.3	1.1		2004	12.6	5.1	20.2	3.6
	2005	6.8	3.5	10	1.5		2005	12.3	7.8	16.7	2.1
	2006	13.5	2.7	24.2	4.2		2006	32.8	16.1	49.5	7.4
	2007	14.2	7.1	21.3	3.4		2007	28.4	16.8	39.9	5.6
	2008	16.5	2	31	6.8		2008	32.7	11.7	53.7	10
	2009	7.3	2.2	12.3	2.4		2009	14.2	5.5	22.9	4.1
	2010	12.7	5.7	19.7	3.3		2010	32.5	14.1	50.9	8.8
	2011	12.6	7.2	18.1	2.6		2011	35.4	9.5	61.4	11.5

2016	12.8	8.2	17.4	2.2	2016	53.9	17.8	90	16.2
2017	18.2	8.4	28	4.8	2017	47.2	25.1	69.3	10.8
2018	21.1	2.5	39.6	8.7	2018	34.9	24.9	44.9	4.8
2019	18.7	9	28.4	4.8	2019	19.7	8.4	31	5.6
2020	29.4	17.3	41.8	7.2	2020	68.8	44.1	111.7	14.7
2021	54.03	6.79	85.27	26.82	2021	152.63	30.01	215.49	69.66
2022	24.29	17.06	31.52	4.27	2022	47.84	31.01	64.67	9.9

Table S5: Stratified mean catch-per-tow of horseshoe crabs in the **Delaware Bay Area** survey area in 2010-2022, with the mean, standard error (SE), and upper (UCL) and lower (LCL) limits of a 95% confidence interval from a bootstrap analysis calculated using a **design-based approach** using a **delta distribution** by demographic group.

YEAR	MEAN	LCL	UCL	SE	YEAR	MEAN	LCL	LCL	SE
FI					**MI**				
2003	10.5	0.7	20.4	4.6	2003	5.4	0.9	9.9	2.1
2004	17.9	8.6	27.2	4.5	2004	15.7	6.4	25	4.5
2005	12.7	5.5	19.9	3.5	2005	11.9	3.8	20	3.9
2006	29.5	16.3	42.8	6.3	2006	21.6	9.2	33.9	5.4
2007	29.6	0	59.4	12.2	2007	19.5	0	39.6	8.2
2008	25.3	6.9	43.7	8.3	2008	18	3.6	32.4	6.3
2009	90.2	12.9	167.4	35.5	2009	69	28.3	109.7	19.8
2010	9	6.1	11.9	1.4	2010	6.1	2.8	9.5	1.6
2011	11.4	6.9	15.9	2.2	2011	6.9	3.7	10.1	1.6
2016	25.8	6.5	45.1	9.2	2016	20	3.5	36.6	7.9
2017	17.9	10.4	25.4	3.4	2017	12.3	4.2	20.5	3.3
2018	22.5	13.9	31.2	4.1	2018	16.5	8.7	24.4	3.7
2019	8	3.2	12.7	2.4	2019	3.5	1	6	1.2
2020	25.3	0.1	51.9	15.2	2020	16	0.8	31.3	9.1
2021	10.4	1.1	19.8	5.5	2021	6.4	1.3	11.5	3
2022	24.64	10.83	38.45	8.14	2022	19.26	7.68	30.85	6.9
FN					**MN**				
2003	1.8	0	3.8	0.9	2003	0.2	0	0.5	0.2
2004	0.8	0.3	1.3	0.2	2004	1.8	1	2.6	0.4
2005	1.1	0.5	1.7	0.3	2005	1.3	0.4	2.3	0.4
2006	4.6	1.5	7.8	1.4	2006	7.1	2.6	11.6	2.7
2007	5.1	0.9	9.3	2	2007	6.7	2.8	10.6	1.9
2008	6	0.2	11.8	2.7	2008	1.8	0.6	2.9	0.6
2009	2	0.9	3.1	0.5	2009	1.7	0.5	2.8	0.6
2010	3	0	6.8	1.8	2010	3.2	0	7	1.8
2011	2	0.7	3.3	0.6	2011	1.9	0.4	3.4	0.7
2016	3.5	1.9	5.2	0.8	2016	5.9	0.7	11	2.5
2017	3.5	1.6	5.5	0.9	2017	3.6	1.5	5.8	1
2018	3.9	1.4	6.3	1.2	2018	7.5	3.1	11.9	2.1
2019	0.5	0	1	0.2	2019	2.8	1	4.6	0.9
2020	0.3	0	0.8	0.3	2020	7	2.9	11	2.4
2021	0	NA	NA	0	2021	16.4	0	37.3	11.3
2022	0.29	0.06	0.52	0.13	2022	13.84	1.65	26.03	7.19
FM					**MM**				
2003	7.7	3.7	11.7	1.9	2003	18.4	7.3	29.6	5.2
2004	5.9	3.3	8.6	1.3	2004	11.4	5.7	17.1	2.8
2005	7.2	3	11.4	2	2005	13.2	7.3	19.1	2.8
2006	15.3	0	33.8	6.7	2006	36.2	11.4	60.9	10.1
2007	16.9	6.2	27.5	5.1	2007	34.3	14.3	54.4	9.7
2008	14.4	5.4	23.3	4.2	2008	33.5	9.8	57.2	11.2
2009	6.7	2.3	11.2	2.1	2009	14.1	5.3	22.8	4.2
2010	11.8	6.3	17.3	2.6	2010	31.5	13.8	49.2	8.6
2011	12.3	7.6	17.1	2.2	2011	36	2.2	69.8	14.7

	2016	13.5	7.6	19.5	2.9		2016	49.2	15.2	83.1	14.3
	2017	16.9	9	24.8	3.9		2017	48.9	23.9	74	12.2
	2018	16.8	9.9	23.7	3.3		2018	35.7	22.5	48.9	6.2
	2019	11.6	4.5	18.7	3.5		2019	20	6.8	33.3	6.6
	2020	29.6	18.1	41.2	6.9		2020	87	34.5	139.4	31.1
	2021	38.2	0	86.5	27.42		2021	95	0	207.8	64.1
	2022	28.22	14.1	42.33	8.28		2022	49.99	20.88	79.1	17.19

APPENDIX D

All material associated with Chapter 2.

Dealing with missing environmental data

Of the 697 unique trawls performed in the DBA between 2003 and 2022, 25 (3.59%) were missing one or more of the following environmental factors: bottom temperature (BOTTEMP), bottom salinity (BOTSAL), or end depth. All trawls missing BOTTEMP were also missing BOTSAL. A significant correlation was found between BOTTEMP and BOTSAL with ordinal date both inshore (0.74, 0.68), and offshore (0.75, 0.65) respectively (Figure 2 & 3). A forward stepwise regression using a generalized linear model with a Gaussian distribution was used to determine what factors best determined the missing variables inshore and offshore. To consider temporal variability and spatial variability, this best-fitting model was refit using four different "neighborhood" datasets. The first data set only considers spatial proximity, selecting the five nearest trawls from the missing data trawl based on the midpoint. This was narrowed temporally by selecting the five nearest trawls within ±7 ordinal days, ±14 ordinal days, and ±21 ordinal days. The best-fitting linear regression model was also compared to a model in which the environmental parameter was estimated using only the mean of each "neighborhood" dataset. The best model for estimation was selected based on the lowest percent root mean square error (%RMSE) and lowest percent coefficient of variation (%CV). For BOTSAL, both the linear regression and estimation based upon the mean resulted in %CV of less than 10% in all cases. Because the linear regression approach predicted values ±5 PPT than those measured in the area, we used the mean values of our best-fitting "neighborhood" datasets to estimate missing BOTSAL. As the best linear regression model for bottom temperature was based on bottom salinity, and all trawls missing bottom temperature were also missing bottom salinity, the mean value of the best-fitting nearby trawls was used.

Dealing with missing end depths

Three methods were considered for estimating missing end depth: 1) Using NOAA ETOPO data, 2) Using the end depth of nearby trawls, and 3) Using the starting depth of trawls. The NOAA ETOPO database is a global relief model of topographic and bathymetric data of 15 arc-second resolution (NOAA NCEI 2022). When compared to previously measured end depths, the mean absolute difference was 5.22 ft (SD: 4.97 ft). Using the end depth of the nearest neighboring trawl resulted in a mean difference between true and estimated of 4.21 ft (SD: 5.87 ft). Lastly, using the beginning depth measurement as the estimated end depth resulted in a mean difference of 3.32 ft (SD: 3.67 ft). As it had the lowest standard deviation, the beginning depth was used to fill in missing end depths.

Bottom salinities of 2022

In all years prior to 2022, BOTSAL was recorded in parts-per-thousand (ppt). During 2022 sampling, bottom salinity was recorded in units of specific gravity. By combining the international one-atmosphere equation of the state of seawater developed by Millero and Poisson (1981) and the equation of specific gravity for saline water (Reid 2008), we were able to estimate the salinity of the trawls given known temperature and specific gravity. This resulted in a BOTSAL ranging from 23.61 – 31.76 ppt, with a mean of 25.80.

Distance from shore measurements

Distance from shore values were calculated using ArcGIS Pro's *near* tool feature and NOAA NCEI's GSHHG shoreline database. The specific shape file, GSHSS_f_L1, contains the highest-resolution shoreline file available at a resolution of 1:250,000 (0.04km), last updated June 15, 2017 (Wessel and Smith 1996). Using a WGS 1984 coordinate system, the value was calculated as the distance of the trawl midpoint to the closest point on the shoreline. At this scale, geodesic and planar methods for calculating distance resulted in the same value.

APPENDIX E

All material associated with Chapter 3.

From our preliminary analysis, we found that bottom salinity and bottom temperature varied greatly from month to month, but not from year to year or between night and day trawls in a similar area. Due to this, traditional spatial-only kriging was hypothesized to be inappropriate as it would not capture the temporal variability when predicting unmeasured cells. Due to this, we wanted to test three different methods of kriging: spatial kriging (traditional), temporal kriging, and spatio-temporal kriging. The spatial-only and temporal-only approaches used ordinary kriging as it is a stochastic model and assumes an unknown mean. In ordinary kriging, there can only be one variable measurement per spatial location, or per cell in this study. To resolve this, the average value for bottom salinity or temperature of all trawls within a cell was used for cells with multiple trawls. Trawls that crossed through multiple cells were assigned to the cell where most of the trawl was performed based on trawl distance.

In the time series, 322 unique spatial cells were sampled from 2003 and 2022 from a total of 697 trawls. To determine the best-fitting variogram for bottom salinity, four different sets of cutoffs and widths were used to first fit the semivariogram (also referred to as the empirical variogram) (Supplemental Formula 1). This was performed in R using the *variogram()* function within the *gstat* package (Pebesma 2004; Gräler et al. 2016). This function uses a default cutoff of 1/3 the length of the diagonal of the spatial bounding box, and a default width of the cutoff divided by fifteen. Four different cutoffs were assessed, all based on the length of the diagonal bounding box: 1/2 the diagonal bounding box, 1/3 (default), 1/4, and 1/5. Four widths were also selected based on how the cutoff (C) value was divided: C/5, C/10, C/15 (default), C/20, resulting in sixteen different semivariograms. A variogram model was then automatically fit to each semivariogram through the *fit.variogram()* function in the *gstat* package. This function allows the user to input a semivariogram, specify the different variogram models to be considered (e.g., spherical, linear, exponential), and return the best-fitting model based on the lowest mean-square-error (MSE). From these sixteen fitted variogram models, the top three models based on MSE were selected.

Because the *gstat()* package does not include a function for temporal-only variogram fitting, time was transformed into a pseudo-spatial variable. This was done by taking the ordinal date of each trawl and while keeping the latitude the same among all trawls, transforming the longitude so that one ordinal day difference would be a change of 0.01 longitude. For example, an ordinal date of 315 would have a longitude of -73.15, and an ordinal date of 316 would have a longitude of -73.16. This resulted in one day spatially equaling about ~1km (0.8641482km). Similar to the spatial-only model, the temporal-only semivariogram can only have one measurement per temporal point, so the mean salinity was taken per ordinal date. Four cutoffs were considered: the entire time period (112 ordinal days), the default cutoff (~38 ordinal days), two months (two weeks or 56 days), and one month (four weeks or 28 days). Four widths were considered, corresponding with true possible time lags: days, weeks, months, and fourth months. These were all transformed into a spatial scale using the 1 day = 0.8641482km mentioned earlier. This process resulted in sixteen different temporal-only semivariograms. Variograms were fit to these using the same process as the spatial-only variograms, and the top three models were selected based on the lowest MSE.

The spatio-temporal fitting and model validation process is similar to spatial- or temporal-only kriging, starting with creating a spatio-temporal semivariogram (Supplemental Formula 2). This was done using *varigoramST()*, which allows for multiple measurements in a single cell, as long as they are different dates. To fit a spatio-temporal variogram, the user must define both the spatial cutoff and width, as well as the temporal lags and temporal units. Temporal lags and temporal unit arguments are similar to the spatial cutoff and width arguments, with temporal units defining the unit you will use to bin the data and temporal lags being the number of bins you will consider. Using the width and cutoff pairs from the top three spatial-only models, and the time lag and temporal unit pairs from the top three temporal-only models, nine spatio-temporal semivariograms were built for salinity.

The two variogram/covariance models considered to fit these semivariograms were a separable model and product-sum model. Both model types were fitted to each of the nine semivariograms using *gstat()*'s *vgmST()* function. Both models require two separate variogram models for the space and time component. These two variogram models input into the spatio-temporal variogram act as a starting point for fitting optimization and require the user to define the model type (e.g., linear, spherical, exponential) and initial values for the partial sill, range, and nugget for optimization. Two sets of initial values were tested for each spatio-temporal variogram: One using the parameters from the best fitting spatial-only and temporal-only variograms, and another based on methods by Wikle et al. (2019 chap. 4). This resulted in a total of 36 fitted spatio-temporal variogram models. Of the 36 fitted models, only the top three models from either the separable or product-sum models that converged were kept based on MSE.

With the top three models for each kriging method selected, predictive performance was compared between models of each environmental variable using a leave-one-out-cross-validation (LOOCV). For the cross-validation of the spatial- and temporal-only models, the left-out data set was either a single spatial cell or a single ordinal day respectively. For the spatio-temporal LOOCV, following similar methods to Gräler et al. (2016), the "one" data set left out consisted of all measurements in a single spatial cell. Overall predictive performance was evaluated based on root-mean-square-error (RMSE), mean-absolute-error (MAE), mean-error (ME), and correlation between the true and predicted values (COR). The kriging method and model with the lowest root-mean-square-error (RMSE) and highest correlation between the predicted and true value (COR) were chosen to estimate the unmeasured cells (Supplemental Table 1). This entire process was repeated for temperature following the same procedure (Supplemental Table 2).

As expected, the spatio-temporal variogram model outperformed both the spatial-only variogram and the temporal-only variogram models in predictive ability for both bottom salinity and temperature (Supplemental Tables 1 & 2). The temporal-only model had consistently lower RMSE and MAE than the spatial-only model, suggesting that time had a larger effect on bottom temperature and salinity when considering the area and time sampling occurred. For this study, as month was already selected as a temporal explanatory variable for all models, we estimated bottom temperature and bottom salinity values for each month sampling had previously occurred, August, September, October, and November.

Supplemental Table 1: Top three models from the spatial-only, temporal-only, and spatio-temporal **salinity** variogram models. The mean-square-error (MSE) is the result of fitting a variogram model to a semivariogram. The root-mean-square-error (RMSE), mean-absolute-error (MAE), mean-error (ME), and correlation (COR) are all results from the LOOCV and measure predictive performance. The top three spatio-temporal models listed here were all product-sum models.

Salinity Variograms	Model Parameters	MSE	RMSE	MAE	ME	COR
Spatial-Only	S: CO_1/2_width5 T:	0.145	3.3338	2.5782	-0.2784	0.326
	S: CO_1/3_width5 T:	0.005	3.3345	2.5794	-0.2841	0.324
	S: CO_1/6_width5 T:	0.253	3.4328	2.6646	-0.2209	0.227
Temporal-Only	S: T: Two_Month.Cut_weeks	0.237	2.443	1.853	-0.052	0.716
	S: T: Default(~5 weeks).cut_weeks	0.339	2.445	1.870	-0.032	0.716
	S: T: Entire(~16 weeks).cut_months	0.639	2.444	1.852	-0.053	0.716
Spatio-Temporal	S: CO_1/6_width5 T: Entire(~16 weeks).cut.months	0.1644	2.19	1.66	-0.11	0.79
	S: CO_1/6_width5 T: Two_Month.cut_weeks	0.3367	2.23	1.69	-0.1	0.78
	S: CO_1/6_width5 T: Default(~5 weeks).cut_weeks	0.2383	2.32	1.77	-0.12	0.76

Supplemental Table 2: Top three models from the spatial-only, temporal-only, and spatio-temporal **temperature** variogram models. The mean-square-error (MSE) is the result of fitting a variogram model to a semivariogram. The root-mean-square-error (RMSE), mean-absolute-error (MAE), mean-error (ME), and correlation (COR) are all results from the LOOCV and measure predictive performance. The top three spatio-temporal models listed here were all separable models.

Temperature Variograms	Model Parameters	MSE	RMSE	MAE	ME	COR
Spatial-Only	S: CO_1/4_width5 T:	0.473	3.0124	2.5084	0.2431	0.292
	S: CO_1/3_width5 T:	0.112	3.0125	2.51	0.2437	0.292
	S: CO_1/2_width5 T:	0.0172	3.0158	2.5105	0.2388	0.2883
Temporal-Only	S: T: One_Month.cut_weeks	0.00416	1.670	1.353	-0.082	0.847
	S: T: Entire(~16 weeks).cut_months	0.0566	1.674	1.348	-0.074	0.846
	S: T: Default(~5 weeks).cut_weeks	0.286	1.673	1.349	-0.074	0.846
Spatio-Temporal	S: CO_1/3_width5 T: One_Month.cut_weeks	0.0789	1.078	0.82	-0.007	0.9392
	S: CO_1/2_width5 T: Entire(~16 weeks).cut_months	0.4205	1.082	0.81	-0.005	0.9388
	S: CO_1/4_width5 T: Default(~5 weeks).cut_weeks	0.0637	1.12	0.86	0	0.93

Supplemental Equations

Supplemental Formula 1: Formula for calculating the traditional, spatial-only semivariogram, also referred to as the empirical variogram (Sherman 2011 pp. 45–47).

$$\hat{\gamma}(h) = \frac{1}{2|N(h)|} \sum_{N(h)} \left\{ \left[Z(s_i) - Z(s_j) \right]^2 \right\}$$

Where:

$Z(s_i)$ is a measured environmental parameter at sample location $\boldsymbol{s_i}$

$N(h) = \left\{ (s_i, s_j): s_i - s_j = h \right\}$ is the set of all point pairs within distance ($h = s_i - s_j$).

Supplemental Formula 2: Formula for calculating the spatio-temporal semivariogram, also referred to as the empirical variogram(Sherman 2011 pp. 125–127).

$$\hat{\gamma}(\boldsymbol{h}, u) = \frac{1}{2|N(\boldsymbol{h}, u)|} \sum_{N(\boldsymbol{h}, u)} \left\{ \left[Z(\boldsymbol{s}_i, \boldsymbol{t}_i) - Z(\boldsymbol{s}_j, \boldsymbol{t}_j) \right]^2 \right\}$$

Where:

$Z(s_i, \boldsymbol{t}_i)$ is a measured environmental parameter at sample location $\boldsymbol{s_i}$ and sample time $\boldsymbol{t}_i$.

$N(\boldsymbol{h}, \boldsymbol{u}) = \left\{ \left[(\boldsymbol{s}_i, \boldsymbol{t}_i), (\boldsymbol{s}_j, \boldsymbol{t}_j) \right]: s_i - s_j = h \; and \; t_i - t_j = u \right\}$

APPENDIX F

All material associated with Chapter 3.

Supplemental Table 1: Mean and 95% upper and lower quantile interval of **September, night-time female** total abundance estimates in thousands from the hurdle model-based approach. versus **design-based normal distribution** approach of mean, 95% confidence intervals.

		Model-Based			Design-Based		
	Year	**LCL**	**Mean**	**UCL**	**LCL**	**Mean**	**UCL**
Immature Females	2003	5303	11509	21857	1317	4089	6860
	2004	6883	13709	24930	4135	7376	10616
	2005	6037	9713	15842	2687	5104	7521
	2006	20339	32875	50372	6439	13714	20988
	2007	22308	38937	62109	48	13692	27335
	2008	11716	19202	28527	4612	10595	16578
	2009	20963	55180	114734	14232	27375	40519
	2010	6893	12894	20914	2497	4102	5706
	2011	9931	15879	24277	2420	5426	8433
	2016	11264	19097	27729	4144	11292	18441
	2017	18897	35224	54786	4077	7948	11818
	2018	24952	41072	65633	6391	10115	13839
	2019	14786	26570	43085	14682	14855	15027
	2020	38584	67123	103926	3106	6832	10559
	2021	28903	65058	124597	436	4053	7670
	2022	11457	21287	38162	5639	8328	11016
Newly-Mature Females	2003	741	1796	3469	26	787	1547
	2004	423	836	1442	120	367	613
	2005	431	888	1473	154	531	908
	2006	3850	7853	12977	540	2122	3705
	2007	2219	4887	8560	674	2129	3584
	2008	2072	4483	8853	613	2697	4780
	2009	1636	3631	7008	399	883	1366
	2010	536	1246	2230	0	1770	4532
	2011	932	1770	3033	269	882	1495
	2016	1197	2247	3569	863	1583	2304
	2017	1169	2672	4184	680	1502	2323
	2018	952	1916	3271	695	1780	2866
	2019	191	829	1843	0	77	225
	2020	0	840	3216	0	134	330

	2021	0	0	0	0	0	0
	2022	96	374	743	18	115	212
Mature Females	2003	3580	7472	11778	1764	3308	4851
	2004	2347	4602	8354	1615	2767	3919
	2005	4093	6443	9882	1592	2957	4323
	2006	6804	12561	20960	1218	5867	10517
	2007	11974	22769	43033	3243	6553	9864
	2008	5630	10342	19799	1008	7172	13336
	2009	174	7336	17140	936	3230	5523
	2010	4281	7584	12873	2478	5588	8698
	2011	10924	17270	27413	3147	5388	7629
	2016	11438	16558	25779	3700	5735	7770
	2017	7484	13556	21964	3537	7785	12033
	2018	8879	17413	27263	464	9463	18463
	2019	13968	21835	36251	6334	6420	6506
	2020	45016	86190	149461	5840	10927	16014
	2021	22372	60530	137753	2867	21766	40665
	2022	11972	19635	30651	6842	9839	12836

Supplemental Table 2: Mean and 95% upper and lower quantile interval of **September, night-time male** total abundance estimates in thousands from the hurdle model-based approach. versus **design-based normal distribution** approach of mean, 95% confidence intervals.

		Model-Based			Design-Based		
	Year	**LCL**	**Mean**	**UCL**	**LCL**	**Mean**	**UCL**
Immature Males	2003	4724	9109	15567	766	2114	3462
	2004	4650	10841	21159	3281	6033	8786
	2005	3487	6165	9505	1932	4673	7414
	2006	13956	23270	36909	4786	9378	13971
	2007	13064	28204	49869	-1035	9350	19735
	2008	8864	14335	23260	3350	6897	10443
	2009	29823	81616	168326	14140	26435	38730
	2010	5034	11077	19554	1139	2781	4423
	2011	5603	10487	17794	1382	3301	5219
	2016	6544	11772	19841	2858	8185	13512
	2017	8604	19318	35015	2335	5082	7829
	2018	17451	29029	45224	3882	7768	11653
	2019	5616	11340	19551	-104	66	236
	2020	19173	32590	51882	1679	4610	7540
	2021	16173	37797	76777	707	2548	4390
	2022	10174	16310	26266	4257	6359	8461
Newly-Mature Males	2003	0	471	1163	-66	78	222
	2004	2063	3758	6118	452	786	1120
	2005	604	1422	2404	233	580	927
	2006	2747	6567	11728	678	3377	6076
	2007	4628	9254	15829	1468	2841	4214
	2008	1571	2886	5601	237	776	1315
	2009	658	2346	5452	259	708	1157
	2010	1501	2680	4757	-252	1464	3180
	2011	1196	2300	3944	190	766	1343
	2016	1916	4299	7467	290	2939	5588
	2017	3050	5530	8795	557	1590	2623
	2018	3320	5512	8565	1663	3064	4466
	2019	2341	4221	6727	-43	112	267
	2020	4129	9283	15442	1184	2430	3676
	2021	2738	16243	49225	0	6308	14299
	2022	2827	6989	13845	1597	6370	11143
Mature Males	2003	16001	30029	49789	4082	7454	10827
	2004	5638	13493	23829	2297	5586	8875

2005	11736	22357	33170	3494	5408	7322
2006	26467	54641	90677	7188	14461	21734
2007	38437	75222	147786	7694	13100	18506
2008	24193	45918	86276	5247	14244	23240
2009	127	25552	51208	2383	6319	10255
2010	18385	36949	66932	6192	14396	22600
2011	20469	42107	69540	3825	14858	25890
2016	33092	73112	136848	7837	24017	40197
2017	26435	55141	93479	10724	19985	29245
2018	34379	60270	105179	10680	15264	19849
2019	32624	59024	90230	11497	11660	11824
2020	78568	205012	485463	15416	25200	34983
2021	28639	115672	287307	13877	61879	109880
2022	14916	33571	64722	12475	19032	25588