Fall 2017

Life History and Management Methods for Blueline Tilefish (Caulolatilus microps) from the United States Mid-Atlantic Region

Michael A. Schmidtke
Old Dominion University, michael.schmidtke32@gmail.com

Recommended Citation

Schmidtke, Michael A.. "Life History and Management Methods for Blueline Tilefish (Caulolatilus microps) from the United States Mid-Atlantic Region" (2017). Doctor of Philosophy (PhD), Dissertation, Ocean & Earth Sciences, Old Dominion University, DOI: 10.25777/rydf-t328
https://digitalcommons.odu.edu/oeas_etds/11

LIFE HISTORY AND MANAGEMENT METHODS FOR BLUELINE TILEFISH (*CAULOLATILUS MICROPS*) FROM THE UNITED STATES MID-ATLANTIC REGION

by

Michael A. Schmidtke
B.S. May 2011, North Carolina State University
M.S. December 2014, Old Dominion University

A Dissertation Submitted to the Faculty of Old Dominion
University in Partial Fulfillment of the Requirements for the
Degree of

DOCTOR OF PHILOSOPHY

OCEANOGRAPHY

OLD DOMINION UNIVERSITY
December 2017

Approved by:

Cynthia M. Jones (Director)

Alexander Bochdansky (Member)

Norou Diawara (Member)

Thomas Miller (Member)

ABSTRACT

LIFE HISTORY AND MANAGEMENT METHODS FOR BLUELINE TILEFISH (CAULOLATILUS MICROPS) FROM THE UNITED STATES MID-ATLANTIC REGION

Michael A. Schmidtke
Old Dominion University, 2017
Director: Dr. Cynthia M. Jones

Blueline tilefish (*Caulolatilus microps*) have recently undergone an increase in landings from waters off the US Mid-Atlantic region (Virginia-Massachusetts). Currently, life history characteristics of blueline tilefish in this area are undescribed. Additionally, appropriate indices of population abundance are not available for this region, necessitating the use of data-limited management methods to set restrictions for harvest of blueline tilefish in this area. This dissertation describes growth and reproductive characteristics of blueline tilefish caught off the coast of Virginia and explores improvements of data-limited management methods that have been applied to blueline tilefish in the US Mid-Atlantic. Blueline tilefish from the US Mid-Atlantic show a similar growth pattern to fish caught off North and South Carolina prior to the rise of commercial fishing for blueline tilefish in this area. This growth pattern is dissimilar from that observed more recently for fish from the US South Atlantic (east coast of Florida-North Carolina), with fish in the US Mid-Atlantic growing more slowly to larger maximum lengths. Blueline tilefish from the US Mid-Atlantic are gonochoristic multiple batch spawners that spawn, on average, every 1.44 days from May-November, with annual fecundities up to 50,000,000 oocytes for the largest individuals. Batch fecundity is correlated with total length in a similar fashion as seen previously in the US South Atlantic. These studies are the first to characterize growth and reproduction of blueline tilefish from the US Mid-Atlantic. Sensitivities of modal estimates and variabilities of total allowable catch (TAC), from data-limited management procedures (MP) applicable to blueline tilefish, to input values were analyzed to determine which inputs most impacted TAC. Natural mortality and mean maximum length from the von Bertalanffy growth equation had the greatest impacts. A composite MP was developed to incorporate distributions of TAC estimates from multiple MPs into a single, combined distribution that could be used to estimate a composite TAC. These analyses provide information that may enhance decision-making for management of data-limited fish stocks, such as blueline tilefish.

This dissertation is dedicated to my grandparents, Robert and Marilyn Schmidtke. My education would not have been possible without your generous love and support. Since my earliest interest in marine life, you have encouraged me and provided me with opportunities to pursue my dreams. You are the people I want to be like when I grow up. I aspire to the lifestyle you’ve lived, full of faith, love, and generosity, garnering the respect of all who have known you. I admire the children you’ve raised, providing me with a father and uncles (and with them, a mother and aunts) who have been role models for my entire life. I am blessed to be spoiled as one of your grandchildren.

I’d also like to express my sincerest love and appreciation for my wife, Genese, and son, Joseph. You are the inspiration that has motivated me, especially within the last year, to finish what I started. I prayerfully look forward to our lives after this chapter and all of the adventures in store for us!

ACKNOWLEDGMENTS

The research described in this dissertation would not have been possible without contributions from so many people. First, I would like to thank my advisory and dissertation committee: Dr. Cynthia M. Jones, Dr. Alexander Bochdansky, Dr. Norou Diawara, and Dr. Thomas Miller, for their lessons taught in the classroom, as well as guidance and patience throughout the dissertation process.

My advisor, Dr. Cynthia M. Jones, has never shied away from pushing me to my limits, and for that, I am forever grateful. Academically, she has provided me with training that has tested my mental boundaries but has also been absolutely essential to my career in fisheries science and management. Professionally, she has volunteered me to participate in stock assessments and management activities that have provided me opportunities to serve on two stock assessment data work groups, one stock assessment panel, and as an analytical lead in making blueline tilefish harvest recommendations to the Mid-Atlantic Fishery Management Council. These experiences have been invaluable to my development as a scientist and helped me earn my current position with the Atlantic States Marine Fisheries Commission.

I would like to thank my fellow graduate students and colleagues at the Center for Quantitative Fisheries Ecology: Dr. Kristen Anstead, Kathleen Kirch, James Davies, Jessica Gilmore, Antranik Kajajian, Dr. Jason Schaffler, Dr. Stacy Baez, Dr. Hongsheng Liao, Dr. Renee Reilly, Jason Ferguson, Dr. Julian Ashford, and Dr. William Persons, as well as several lab technicians. Between otolith-sectioning instructions, sample collection trips, discussions of science and other invigorating topics, and so much more, each of them made my graduate education a fun and fulfilling experience!

I would also like to thank Dr. Joey Ballenger and the research staff at the South Carolina Department of Natural Resources Marine Research Institute for hosting me for my first otolith aging workshop, along with several others that followed, and for generally collaborating with me to address blueline tilefish life history questions. I would specifically like to thank Joey for writing the proposal to fund this research and conducting the first couple years of sample collections.

I would like to thank my collaborators for their assistance and comments while completing my chapters, including: Dr. Amy Schueller for developing and providing code for the selectivity bias growth adjustments in Chapter 2, Dr. Susan Lowerre-Barbieri and the staff at the Florida Fish and Wildlife Research Institute for processing reproductive samples and providing comments and methods text for Chapter 3, and Dr. Thomas Miller for working with me through my crash course in DLMTool and helping me to develop research questions addressed in Chapter 4.

Finally, this work would not have been possible without funding and support from the Virginia Marine Resources Commission (VMRC), including sampling and processing efforts by Myra Thompson, Chris Williams, and Richard Hancock, and sample contributions donated by recreational anglers through the Virginia Marine Sportfish Collection Project.

TABLE OF CONTENTS

Page

Chapter

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

BACKGROUND

Throughout the history of fisheries management, models depicting the status of fished populations have expanded in complexity through the incorporation of longstanding population surveys, a variety of biological data, and advanced statistical techniques. Most often, the latest and most complex models are reserved for widely popular and financially lucrative commercial fisheries that have operated (in the United States) on a large scale for hundreds of years. However, the extensive and precise information required to operate such models may not be available for fisheries that have emerged more recently, causing fishery managers and scientists to assess newer fisheries using more basic models with modest amounts of information that have been collected for a relatively short time period. Within the last 50 years, fisheries for several demersal species found in depths of approximately 100-300 m off the US Atlantic coast have emerged and grown in popularity (Parker and Mays 1998). One such fishery is that for blueline tilefish (*Caulolatilus microps*).

Blueline tilefish is a species that inhabits bottom waters along the continental shelf edge and slope off the Atlantic coast of North America. Though previously thought to range from Campeche, Mexico through the Gulf of Mexico and up the US Atlantic coast as far north as Cape Henry, Virginia (Dooley 1978), recent encounters through commercial and recreational fisheries and rare catches in offshore fishery-independent surveys indicate a more northerly extent to their range, with encounters as far north as Georges Bank off Massachusetts (Southeast Data Assessment and Review (SEDAR) 2017).

Historical landings of blueline tilefish along the US Atlantic coast have come primarily from the US South Atlantic region (east coast of Florida-North Carolina) (National Marine Fisheries Service (NMFS[1])). In this region, blueline tilefish have typically been caught indirectly during targeted fishing for more popular co-occurring species such as snowy grouper (Ross and Huntsman 1982). In the mid-2000s, a relatively small, primarily recreational fishery for blueline tilefish began to develop in the US Mid-Atlantic region (Virginia-Massachusetts) (Mid-Atlantic Fishery Management Council (MAFMC) 2016a). In Virginia, this evolved into a directed fishery in which blueline tilefish became one of the primary targets for headboat trips out of Virginia Beach, VA. These trips targeted blueline tilefish near the offshore submarine canyons at the continental shelf edge, most often Norfolk Canyon, roughly 70 miles east of Virginia Beach. This increase of harvest in Virginia motivated the proposal for research described in Chapters 2 and 3 to characterize life history information of blueline tilefish being targeted by this fishery.

Blueline tilefish have historically been managed as two stocks, one in the Gulf of Mexico and the other along the US Atlantic coast. Blueline tilefish along the US Atlantic coast have a relatively short, but quite complicated management history. In response to a large increase in tilefish landings during the early 1980s (although there is still some question about the species composition of these landings) the South Atlantic Fishery Management Council (SAFMC) (with jurisdiction in US South Atlantic federal waters, 3-200 nautical miles offshore) began managing blueline tilefish harvest in 1983 (SEDAR 2017). Minimal landings occurred north of North

---

[1] National Marine Fisheries Service (NMFS), Fisheries Statistics Division, 2017. Personal commun.

Carolina, but these were not large enough to motivate large-scale management actions. As the US Mid-Atlantic recreational fishery for blueline tilefish began to develop in the mid-2000s, Virginia established state landings restrictions in 2007 for both commercial and recreational fisheries for tilefish and grouper species; Maryland instituted similar limits in 2010.

In 2013, a stock assessment was conducted for US South Atlantic blueline tilefish. The assessment determined that the stock was overfished and overfishing was occurring (SEDAR 2013). This assessment considered use information presented in Chapters 2 and 3 of this dissertation during the assessment, but did not have enough information to evaluate landings north of North Carolina. Following this assessment's status determination (overfished and overfishing), the SAFMC increased restrictions on harvest of blueline tilefish in the US South Atlantic. In the first year of increased restrictions, 2014, commercial harvest of blueline tilefish from the US Mid-Atlantic increased to over ten times the average of the previous ten years (NMFS). Recognizing that this increase was clearly driven by movement of fishers and fishing activity northward rather than movement of the stock, the MAFMC instituted emergency restrictions on harvest and established an annual harvest limit for the US Mid-Atlantic, using data-limited management methods (MAFMC 2016b). The most recent assessment of the entire Atlantic stock indicates that the portion of the stock located south of Cape Hatteras, NC, is not overfished and overfishing is not occurring (SEDAR 2017). Due to data limitations, stock status is currently not able to be determined for the portion of the stock located north of Cape Hatteras, so only suggested data-limited methods for determining harvest levels in this region are provided by the assessment (SEDAR 2017).

Previous research on blueline tilefish is limited in the amount, type, and geographic range of studies. Two studies have been conducted that have investigated age, growth, and reproductive characteristics of blueline tilefish caught off North and South Carolina (Ross and Huntsman 1982; Ross and Merriner 1983; Harris et al. 2004). One study observed blueline tilefish, via submersible, constructing burrows in seafloor sediments off Florida and South Carolina (Able et al. 1987). Another study observed, via submersible, blueline tilefish along with other species occupying these burrows off North Carolina (Parker and Ross 1986). More recent research and monitoring efforts for blueline tilefish have focused on annual aging and reproductive characteristics in the US South Atlantic and genetic connectivity throughout the species range in the US Atlantic (McDowell 2016; O'Donnell and Darden 2016; Klibansky 2017; SEDAR 2017). No studies prior to those described in this dissertation have investigated age, growth, or reproductive characteristics of blueline tilefish from the US Mid-Atlantic.

Studies on aging of blueline tilefish using otoliths are limited to two studies that both sampled off North and South Carolina (Ross and Hunstman 1982; Harris et al. 2004). Otoliths are calcified structures located in the inner ear of teleost fishes that are used for balance and hearing. Otoliths are metabolically inert and grow throughout a fish's entire lifetime, allowing them to be considered as one of the most reliable methods for determining age in fishes, through the counting of periodic (often assumed or proven to be annual) rings formed by the periodic and sequential incorporation of translucent and opaque material into the otolith over time (Campana 1999). The size and chemical composition of accumulated materials varies depending on environmental (such as temperature) and biological (such as spawning) conditions, analogous to the counting of annual rings to determine the age of trees, which have actually been used to validate otolith ages (Campana and Jones 1992; Campana 1999; Campana and Thorrold 2001; Black et al. 2005). However, counting of annual rings (annuli) is subject to errors of interpreting distinct rings, and precision of counts is often measured by comparing multiple independent reads of the same otolith. Ross and Hunstman (1982) did not report any measures of aging precision in their study of blueline tilefish, but Harris et al. (2004) reported difficulties with distinguishing annuli, resulting in low precision between independent reads of the same otoliths (24% agreement, 64% within 2 years). These difficulties are typical for deeper water demersal fishes and have also been observed for several species that co-occur with blueline tilefish, such as blackbelly rosefish (*Helicolenus dactylopterus*) (White et al. 1998), snowy grouper (*Epinephelus niveatus*) (Wyanski et al. 2000), barrelfish (*Hyperoglyphe perciformis*) (Filer and Sedberry 2008), and golden tilefish (*Lopholatilus chamaeleonticeps*) (Lombardi-Carlson 2012). Difficulty in aging these species stems from limited environmental seasonality due to the depths at which these species live (blueline tilefish range from 48 to 236 m (Harris et al. 2004)), resulting in opaque zones of annuli that are not easily distinguishable from translucent zones (see Figure 3 in Chapter 2, for example).

Previous studies of reproductive characteristics for blueline tilefish caught off North and South Carolina indicate that blueline tilefish are multiple batch spawners (Ross and Merriner 1983; Harris et al. 2004). One of these studies observed prematurational sex reversal in four juvenile individuals (Ross and Merriner 1983), but no similar observations have been made in more recent studies. Therefore, blueline tilefish are generally considered gonochoristic. Blueline tilefish from the US South Atlantic are prolific spawners with females spawning, on average, every two days from March through October, resulting in annual fecundity estimates in excess of 10,000,000 oocytes per female (Harris et al. 2004). Neither of the previous studies investigated blueline tilefish in the US Mid-Atlantic or acquired a substantial number of specimens during November through February, presumably the non-spawning season.

Movement of the Gulf Stream offshore near Cape Hatteras, NC, creates differing environmental conditions between the north and south of this biophysical boundary and the potential for local adaptations within populations whose ranges span it. Although genetic information indicates genetic connectivity north and south of Cape Hatteras in blueline tilefish (McDowell 2016; O'Donnell and Darden 2016; SEDAR 2017), a lack of data that would indicate large-scale movement during the adult life stage suggests this connectivity is more likely to occur during the larval or juvenile life stage. Additionally, recent drifter data suggests a potential mechanism for connectivity among blueline tilefish from the Gulf of Mexico, US South Atlantic, and US Mid-Atlantic via transport of pelagic larvae (Klibansky 2017; SEDAR 2017). This indicates that there is still potential for local adaptations after settlement that would impact the productivity and health of local populations within the Atlantic stock. Physical differences have been observed between populations north and south of Cape Hatteras in golden tilefish, such as US Mid-Atlantic golden tilefish mean maximum length estimated as 9-10 cm longer than those of US South Atlantic and faster growth to maximum size in the US South Atlantic (Katz et al. 1983; Northeast Fisheries Science Center (NEFSC) 2014; SEDAR 2011b), which exemplify the potential for non-uniform population characteristics among regions.

Due to a lack of reliable abundance indices, the Atlantic stock of blueline tilefish, particularly the portion of the stock located north of Cape Hatteras, is considered data-limited and must rely on data-limited models for suggested harvest levels (SEDAR 2017). The Data-Limited Methods Toolkit (DLMTool) is an R program that includes over 80 data-limited management procedures (MP) that can be used to set harvest limits (Carruthers and Hordyk 2017). DLMTool has been used previously to set harvest limits for blueline tilefish in the US Mid-Atlantic by evaluating methods applicable to the available data for US Mid-Atlantic blueline tilefish through a management strategy evaluation (MSE), selecting MPs that fit MSE performance criteria, and averaging the medians of total allowable catch distributions from the selected MPs when applied to US Mid-Atlantic blueline tilefish data (MAFMC 2016b).

Practical use of data-limited management procedures likely requires subjective decision-making at several points during the modeling process. Knowledge of sensitivities and tendencies of MPs with respect to input values and output TACs could be useful for informing these decisions. For example, in most probability distributions, narrowness of values surrounding a mode would be considered indicative of a decreased level of uncertainty. However, if the data informing a distribution with a narrow peak is based on limited data (such as only catch) or is highly sensitive to data with a high degree of inherent uncertainty that is not portrayed through model inputs (such as age for deepwater species like blueline tilefish), this distribution may be less informative than a wider distribution that incorporates several types of data and is more sensitive to inputs with less uncertainty.

Previous applications of DLMTool to demersal stocks in the US Mid-Atlantic have used a combination of MPs to make harvest limit recommendations (Miller et al. 2015; MAFMC 2016b). In both applications, medians from simulated TAC distributions were averaged to calculate the final harvest limit recommendation. While this method certainly makes use of information from multiple plausible methods, it does so in a way that ignores the central tendency of TAC distributions around modal values as well as the variability associated with the TACs for the MPs being combined. An alternative method would be to: 1) use modes of TAC distributions as estimates, and 2) develop a composite distribution that incorporates information from multiple MPs to calculate TAC as the central tendency of multiple method simulations.

OBJECTIVES

The objectives of research presented in this dissertation are to:

1. Characterize age and growth characteristics of blueline tilefish from off the coast of Virginia and compare these to age and growth characteristics observed in the US South Atlantic.
2. Characterize reproductive characteristics of blueline tilefish from off the coast of Virginia to determine whether and to what level spawning is occurring in this region.
3. Explore methods to improve the practical application of data-limited fishery management procedures within the context of US Mid-Atlantic blueline tilefish.

Each of these objectives will be addressed in the following chapters. In Chapter 2, I will describe aging of blueline tilefish using otoliths, model growth of blueline tilefish collected off the coast of Virginia using the von Bertalanffy growth model (von Bertalanffy 1938), and compare growth model parameter estimates to those estimated for blueline tilefish in the US South Atlantic. In Chapter 3, I will describe reproductive characteristics of blueline tilfish collected off the coast of Virginia using sex ratio, spawning seasonality, and estimates of fecundity. This information will be used to assess whether blueline tilefish located off Virginia reproduce locally and what reproductive productivity levels are for this portion of the Atlantic stock. In Chapter 4, I will analyze data-limited fishery management procedures (MP) from the Data-Limited Toolkit (DLMTool) to determine the level of influence inputs have on MPs and develop a composite MP that will allow incorporation of information from multiple MPs in the form of a combined distribution. Analyses from Chapter 4 will all be completed using the mode of total allowable catch distributions rather than the median, which has been used most often in previous applications of DLMTool. Finally, in Chapter 5, I will summarize the major conclusions of this dissertation and its impacts as the first research describing life history of blueline tilefish in the US Mid-Atlantic and providing analytical methods that can inform data-limited management decisions through the use of modal total allowable catch (TAC) estimates, impacts of input parameters from US Mid-Atlantic blueline tilefish on TAC distributions, and a composite TAC distribution.

CHAPTER 2

AGE AND GROWTH OF BLUELINE TILEFISH (*CAULOLATILUS MICROPS*) OFF VIRGINIA WITH SPATIAL AND TEMPORAL COMPARISONS TO THE US SOUTH ATLANTIC

INTRODUCTION

The blueline tilefish (*Caulolatilus microps*) is a demersal teleost that inhabits the North American outer continental shelf, shelf break, and slope at depths ranging from 48 to 236 m (Dooley 1978; Ross and Hunstman 1982; Harris et al. 2004). Although once thought to range from Cape Charles, Virginia, to Campeche, Mexico (Dooley 1978), more recent data show that blueline tilefish have been commercially caught in waters further north off the Mid-Atlantic region of the United States (from Virginia through Massachusetts; hereafter: US Mid-Atlantic) and landed as far north as Massachusetts (National Marine Fisheries Service (NMFS)[1]).

Blueline tilefish, as well as other malacanthids, occupy crevices or burrows as a method of predator avoidance (Able et al. 1987). Therefore, they can occupy a wide range of bottom substrates ranging from hard substrates, such as rocks and boulders, to sandy sediments in which they construct burrows in the seafloor (Parker and Ross 1986; Able et al. 1987). Previous research outside the US Mid-Atlantic suggests blueline tilefish are long-lived and slow growing, with observed ages of up to 43 years and low Brody growth parameter estimates of

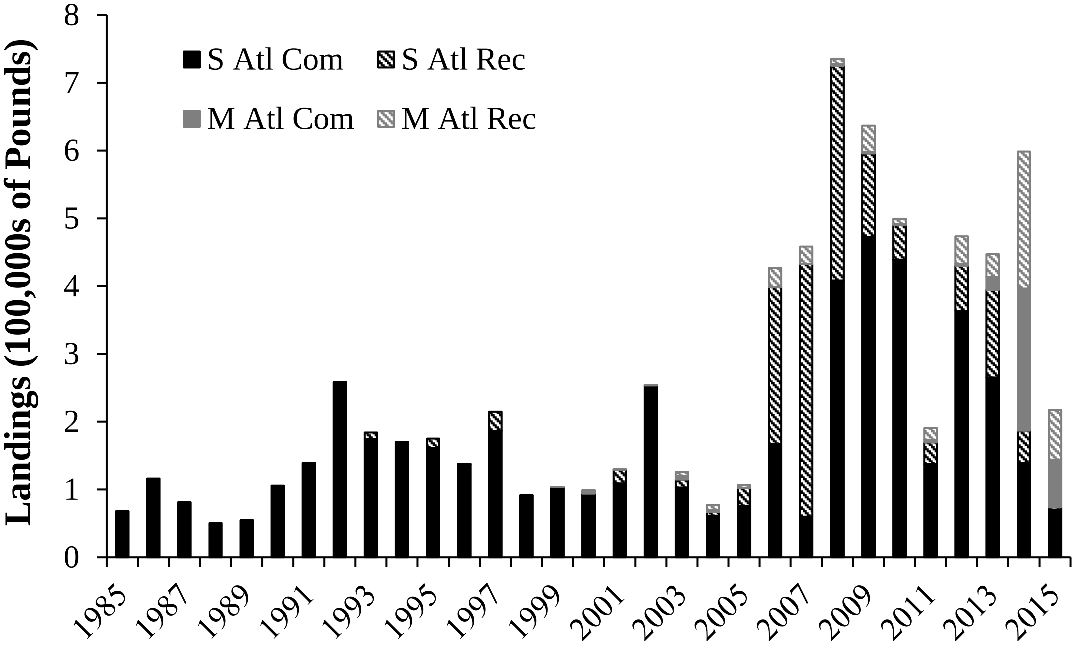
---

[1] National Marine Fisheries Service (NMFS), Fisheries Statistics Division, 2017. Personal commun.

approximately 0.1 (Ross and Huntsman 1982; Harris et al. 2004). Similar longevities and growth rates have been observed in other species occupying the cool, stenothermal bottom waters of the North American continental shelf and slope (Stefánsson et al. 1971; Ruzecki et al. 1972), such as snowy grouper (*Epinephelus niveatus*) (Wyanski et al. 2000), golden tilefish (*Lopholatilus chamaeleonticeps*) (Lombardi-Carlson 2012), and Warsaw grouper (*E. nigritus*) (Manooch and Mason 1987).

Blueline tilefish are currently managed as 2 stocks: a Gulf of Mexico stock and a US Atlantic coast stock, although a stock identification workshop associated with the most recent assessment did not observe genetic differences in blueline tilefish between these regions (McDowell 2016; O'Donnell and Darden 2016; Southeast Data Assessment and Review (SEDAR) 2017). The Atlantic stock has been harvested commercially and recreationally since the 1970s. Historical landings were almost entirely from the commercial fishery off the South Atlantic region of the United States (from the Atlantic coast of Florida through North Carolina; hereafter: US South Atlantic). Similar to other deepwater reef fisheries in this area, annual landings peaked during the 1980s (Parker and Mays 1998; SEDAR 2017). However, proportions of commercial landings have shifted northward since the 1980s, with increasing proportions of commercial landings coming from North Carolina and states further north in the US Mid-Atlantic (NMFS). Though increasing since the early 2000s, annual US Mid-Atlantic commercial landings were relatively modest until 2014, when they suddenly increased to nearly ten times the average from the previous ten years (NMFS) (Figure 1). This increase coincided with stricter harvest regulations in the US South Atlantic, resulting from a 2013 benchmark stock assessment that reported overfished and overfishing statuses for the fishery in that region (SEDAR 2013).

Figure 1. Landings of blueline tilefish along the US Atlantic coast, divided by region as US South Atlantic (S Atl) (Florida-North Carolina) and US Mid-Atlantic (M Atl) (Virginia north) and fishing sector as commercial (Com) or recreational (Rec). Landings records prior to 1985 did not specify blueline tilefish as an individual species, but grouped multiple species of "tilefish" into a single assemblage. These are not shown here, but estimates of more historical landings may be referenced via Parker and Mays (1998) or Southeast Data, Assessment, and Review (SEDAR) (2017). Sources: Commercial landings and US South Atlantic recreational landings: Personal communication, National Marine Fisheries Service (NMFS) Fisheries Statistics Division 2017; US Mid-Atlantic recreational landings: Mid-Atlantic Fishery Management Council (MAFMC) 2016a.



The increase in US Mid-Atlantic catch has introduced concerns over how blueline tilefish should be managed outside of the US South Atlantic. Little is known about the Mid-Atlantic portion of the stock, including whether the fishery in this area is newly exploiting a pre-existing resource or a new resource introduced through a shift in the stock's range. Stock differences among other demersal, co-occurring species that span the biophysical boundary created by movement of the Gulf Stream offshore at Cape Hatteras, North Carolina, such as black sea bass (*Centropristis striata*) (SEDAR 2011a; Northeast Fisheries Science Center (NEFSC) 2012) and golden tilefish (*Lopholatilus chamaeleonticeps*) (NEFSC 2014; SEDAR 2011b), as well as differences in landings (i.e., fishing mortality) histories between regions preclude assumptions of uniformity in population dynamics throughout the Atlantic stock of blueline tilefish. A lack of data from the US Mid-Atlantic has forced managers to use data-limited methods that are highly reliant on catch time series and life history parameters to estimate appropriate catch limits for this region (Carruthers 2014; Miller et al. 2015). Recreational harvest, in particular, has been scarcely recorded for blueline tilefish in the US Mid-Atlantic via conventional methods, like the Marine Recreational Information Program, and had to be estimated using a Delphi survey process (Mid-Atlantic Fishery Management Council (MAFMC) 2016a), further increasing uncertainty of historical catch estimates in this region.

Spatial or temporal variations in growth and other life history characteristics can occur due to differences or changes in selective pressures. Fishing, specifically, can impact growth by selectively removing individuals from certain groups (such as larger or faster-growing fish) within a population, thus altering the population's growth and reproductive dynamics. Blueline tilefish in waters off the US South Atlantic have undergone longer and more intense exploitation than those further north (NMFS) (Parker and Mays 1998; MAFMC 2016a), potentially leading to growth differences between these areas. Spatial variation in growth parameters may be indicative of structuring within a population and can impact the calculation of management reference points (Law 2000; Heino et al. 2013; Maunder et al. 2016). Spatial differences in size-at-age are also indicative of a lack of adult connectivity between areas. Changes in size-at-age due to an increase in fishing pressure for one area would be reflected throughout a stock if population subunits were connected during the adult stage, as adults from multiple subunits (with multiple growth patterns) would be caught while moving between areas. Although no studies have directly investigated blueline tilefish movement between populations north and south of Cape Hatteras, NC, their burrow-constructing behavior indicates a fair degree of investment in their local environments (Able et al. 1987). Additionally, a tagging study of a similar species, golden tilefish, showed evidence of a sedentary adult lifestyle that is not conducive to connectivity of widely-spaced population subunits during this life stage (Able et al. 1982; Grimes et al. 1983). Physical and genetic differences have been observed between populations north and south of Cape Hatteras in golden tilefish (Katz et al. 1983; NEFSC 2014; SEDAR 2011b). Although this was the case for golden tilefish, differences in physical characteristics that reflect a lack of adult connectivity do not necessarily reflect a lack of genetic connectivity. Blueline tilefish show no genetically distinct population sub-units throughout their range along the US Atlantic coast (McDowell 2016; O’Donnell and Darden 2016; SEDAR 2017). Genetic connectivity between subunits with different sizes-at-age may be maintained at the larval or juvenile stage, with distinctions in growth due to spatial differences in resources or mortality not being observable until after recruitment. Pelagic drifter data indicate that interactions of pelagic larvae with the Gulf Stream in the Atlantic and Loop Current in the Gulf of Mexico would provide a potential mechanism for genetic connectivity of blueline tilefish from the Gulf of Mexico, US South Atlantic, and US Mid-Atlantic during the larval life stage (Klibansky 2017; SEDAR 2017).

Studies of life history characteristics provide vital information for stock assessment models, such as age and growth data, and can also indicate non-genetic structuring of adult populations targeted by fisheries (Ihssen et al. 1981; Begg et al. 1999; McBride 2014). Previous research on blueline tilefish life history has focused on the populations from the South Atlantic, and current databases of age and size information for fish in this region are maintained by the National Oceanic and Atmospheric Administration (NOAA) Southeast Fisheries Science Center and the South Carolina Department of Natural Resources. However, characteristics of the population from the US Mid-Atlantic are currently undescribed, and the need for such data is now amplified by the drastic increase in US Mid-Atlantic catches in 2014. Considering the uncertainty associated with current catch estimates and general lack of data for blueline tilefish in the Mid-Atlantic, reliable estimation of life history parameters is vital to inform current data-limited management models for blueline tilefish in this region. This study will characterize the age and growth of blueline tilefish off the coast of Virginia, in the southern portion of the US Mid-Atlantic, and compare them with past and present data for blueline tilefish from the US South Atlantic, to determine whether and how this stock may vary across space and time.

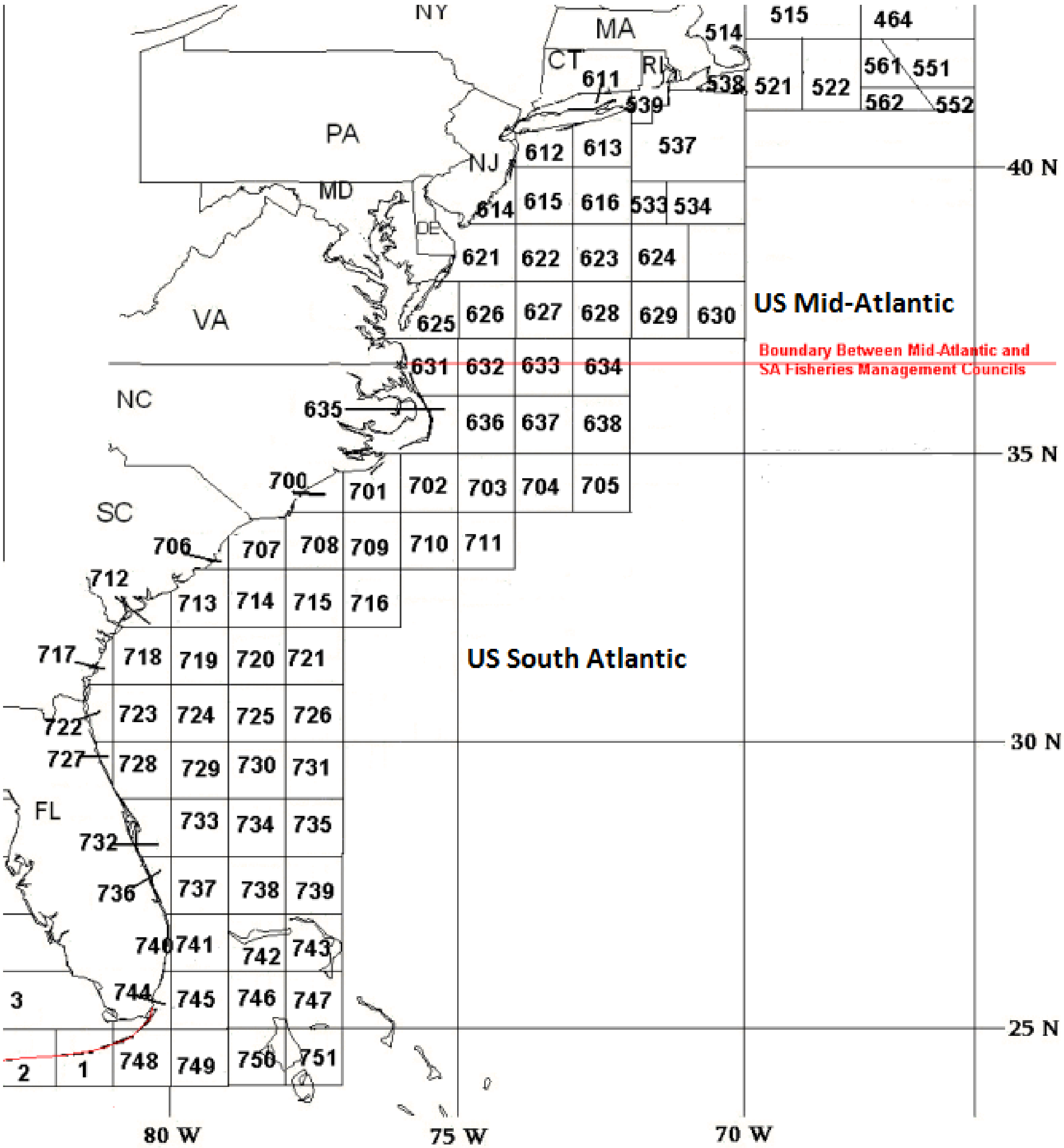
MATERIALS AND METHODS

Sample Collection and Processing

Blueline tilefish specimens caught off the Virginia coast were collected from 2009-2012. Collection methods included purchases of whole fish from commercial fishermen, donations of whole fish or carcasses from recreational anglers, and quasi-fishery independent collections by scientists from the Virginia Marine Resources Commission (VMRC) and the Old Dominion University Center for Quantitative Fisheries Ecology (CQFE) aboard recreational charter and head boat vessels (hereafter referred to as “special charters” because while all fish caught during these collections were kept, regardless of size or any other characteristic, fishing locations were representative of the recreational fishery and not selected randomly). Samples were collected throughout the year to have an adequate number of samples within each month for otolith marginal increment analysis (MIA).

Total and fork lengths (mm) were measured for all specimens. Catch locations reported by fishermen were identified within NMFS statistical areas (Figure 2). Sex was determined as male, female, or juvenile (indeterminate) whenever possible through macroscopic examination of reproductive organs within the body cavity by CQFE and VMRC scientists. Saggital otoliths were removed and stored in coin envelopes. One otolith from each specimen (randomly selected between left and right) was embedded in epoxy resin. A transverse section (0.4 mm thick) was made through the core using a Buehler Isomet low-speed saw, and sections were mounted on glass slides using Flotexx. Slides were viewed under a microscope at 20-40x magnification using transmitted light.

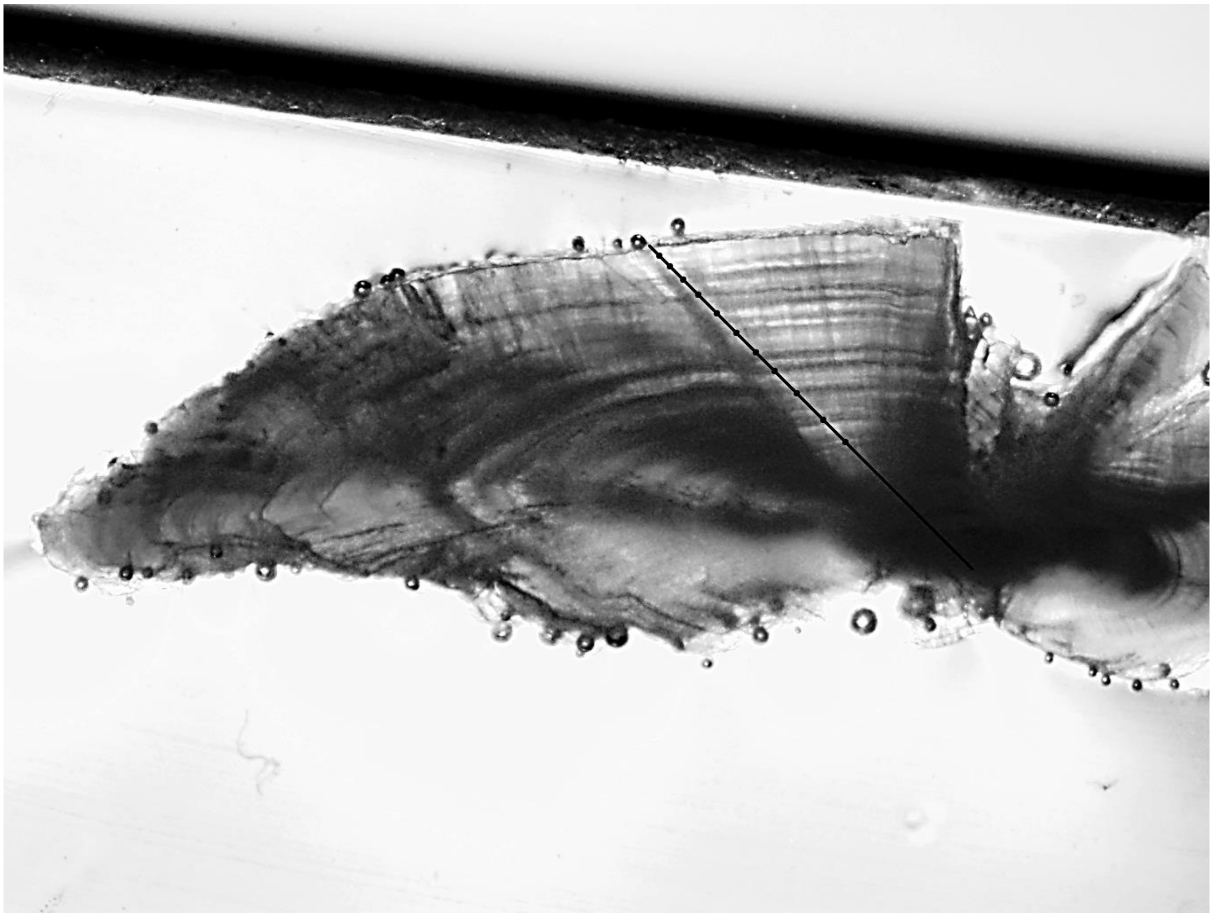
Figure 2. National Marine Fisheries Service statistical areas for US South Atlantic and US Mid-Atlantic regions.



Aging

Aging was attempted for all specimens collected from 2009-2011. To reduce time and costs of processing, the 2009-2011 data was used to proportionally allocate (based on total length) a subsample of the 2012 specimens for age analyses (Quinn and Deriso 1999). In preparation for SEDAR 32, an aging protocol was established by age readers from CQFE and other agencies throughout the US South Atlantic to ensure consistency of aging methods throughout the Atlantic coast (SEDAR 2013). More recently, a study of bomb radiocarbon levels in older blueline tilefish otoliths, in which one of the age readers from the present study participated, indicated that aging methods used in this study were comparable to those of other labs along the US Atlantic coast and, depending on the reference chronology used, are either consistent with or slightly under-aging relative to the true age of blueline tilefish (Ballenger et al., unpublished). Increments consisted of one translucent and one opaque zone (Figure 3) and were primarily counted along a ventral axis of the section. Occasionally, the dorsal region of the section was counted if the ventral was unclear, and when possible, both regions were counted and compared for additional age verification. Discontinuous opaque or translucent areas were common, which made aging difficult. When possible, increments were confirmed to extend from the sulcal groove to the distal edge of the section. Increments were counted independently by two CQFE readers without knowledge of fish size or time of capture. If independent counts differed, the slide was recounted by both readers until a consensus age could be agreed upon. If no age could be agreed upon, the specimen was discarded from age analyses. Precision between readers was evaluated using bias plots, average percent error (APE), and percent agreement of initial readings. A paired t-test was used to determine whether average biases from initial readings deviated from zero.

Figure 3. Eleven year old blueline tilefish (*Caulolatilus microps*) otolith section, with a line indicating axis of measurement for marginal increment analysis and markers denoting opaque zones of annuli.



Marginal Increment Analysis

Periodicity of increment formation was investigated using MIA. Increment widths were measured for a stratified (by month) random sample of the aged dataset. Increment widths were measured from the otolith nucleus to the edge along an axis roughly 45 degrees proximal to the ventral axis (Figure 3). This axis was used because this region of the otolith was where increments were most consistently visible and distinguishable. Increments in other portions of otolith sections often showed splitting or were visibly faded, including along the distal edge, a more common axis for increment measurement. All increments were measured using Image-Pro Plus vers. 6.2.0.424 (Media Cybernetics, Inc.). An index of completion was calculated by multiplying the marginal increment width by 100 and then dividing by the width of the last complete increment (Hyndes et al. 1992). Monthly average indices were calculated and plotted against the calendar year to determine timing and periodicity of increment formation.

Age and Length Frequencies

Overall and sex-specific age and length distributions were developed, and sex-specific age and length distributions were compared using Kolmogorov-Smirnov tests. Mean lengths and ages for sexually distinguishable adults were compared among sample collection methods (fishery-dependent, special charters), capture locations (NMFS statistical areas), and sexes using a multivariate analysis of variance (MANOVA). Significant factors were further compared using Student's t-tests. Overall and sex-specific means were also compared with those of previous studies using Student's t-tests.

Virginia Growth Analyses

Length at age was modeled using the von Bertalanffy (VB) growth function (von Bertalanffy 1938):

$$L_t = L_\infty[1 - e^{-\kappa(t-t_0)}], \qquad \text{(Equation 1)}$$

where L_∞ is the horizontal asymptote representative of the mean maximum length, κ is the Brody growth parameter representative of how quickly maximum length is achieved, and t_0 is the theoretical age at length=0. Since previous studies have found evidence of sexually dimorphic growth in US South Atlantic blueline tilefish (Ross and Huntsman 1982; Harris et al. 2004), sex-specific growth models were compared by a likelihood ratio test (LRT) to determine whether this characteristic is also seen in fish caught off Virginia (Kimura 1980). This pairwise test uses categorical combined VB models to simultaneously describe the two data sets being compared. A test statistic with a chi square distribution is calculated using a ratio of the likelihood for a combined VB model that satisfies some linear constraint (H_{01}: $L_{\infty 1} = L_{\infty 2}$; H_{02}: $\kappa_1 = \kappa_2$; H_{03}: $t_{01} = t_{02}$) divided by the likelihood of an unconstrained model. A large test statistic indicates that the constrained model does not adequately represent both data sets; thus, the parameter being constrained is different for each of the compared data sets. As this test can evaluate individual parameter estimates, rather than only entire sets of parameter estimates, it can indicate if growth differences are attributable to different asymptotic sizes or relative growth rates using tests with constraints on L_∞ and κ, respectively.

Sample sizes of long-lived species, such as blueline tilefish, can be relatively small at older, rarer ages. A lack of fish at older ages causes the VB model to base estimation of the asymptote, L_∞, more on the growth trajectory of more abundant, younger fish whose growth has not yet reached an inflection point, rather than less abundant, older fish whose growth is tending towards an asymptote. The lack of older fish weakens those data points' ability to "pull down" the model towards its actual asymptote, leading to potential overestimation of L_∞. Increasing the relative weights of data points in these rarer age classes, can allow them greater influence in estimation of L_∞, making this parameter more representative of fish that have reached their maximum lengths. We accomplished this by weighting observed total lengths by the inverse of the sample sizes for each age and group (Beckman et al. 1990; Hyndes et al. 1992). Weighted models can still be compared by LRTs, because under assumed normally distributed errors within each age, the ratio of likelihoods reduces to a ratio of the variances (Kimura 1980). The only adjustment to the LRT is that variance is now weighted for the full and reduced models:

$$\hat{\sigma}_w^2 = \frac{\sum_{i=1}^{2} \sum_{j \in A_i} \sum_{k=1}^{n_{ij}} w_{ijk}(L_{ijk} - \hat{L}_{ij})^2}{\sum_{i=1}^{2} \sum_{j \in A_i} \sum_{k=1}^{n_{ij}} w_{ijk}},$$

where i indicates group 1 or 2 of the two groups being modeled, j=ages of group i belonging to the set of numbers A_i, A_i=the set of ages in group i for which n_{ij} is a positive integer, n_{ij}=number of fish of age j in group i, k=observed fish of age j in group i, $w_{ijk}=\frac{1}{n_{ij}}$=weight for the kth fish of age j in group i, L_{ijk}=observed total length of the kth fish of age j in group i, and $\hat{L}_{ij}$=estimated length for fish of age j in group i. As the parameter L_∞ estimates mean asymptotic length, we would expect this value to be near or less than the maximum observed length for a sample containing fish at the oldest possible ages for that species. Thus, weighted regression was used when L_∞ from unweighted least squares regression exceeded the maximum observed length within a given dataset and when comparisons to such a dataset were made using LRTs.

Regional Growth Comparisons

We made two comparisons of growth among groups of blueline tilefish that varied across space and time. The first comparison was among current Virginia blueline tilefish and US South Atlantic blueline tilefish capturing during 3 past time periods: 1972-77 (Ross and Huntsman 1982) 1982-87 (Harris et al. 2004), and 1996-99 (Harris et al. 2004). Although these previous studies listed overall mean total lengths and ages with standard deviations, complete raw data sets were unavailable from which to estimate variances of length at age estimates. Therefore, we represented growth for these time periods using estimates of VB growth parameters from the previous studies (Ross and Huntsman 1982; Harris et al. 2004). An important distinction between these studies was that Ross and Huntsman (1982) aged whole otoliths, while Harris et al. (2004) aged sections similar to those used in our aging. These studies fit VB growth models to unweighted mean total lengths at age, so we fit a VB model for current Virginia fish to unweighted mean total lengths at age for this comparison. Historical estimates for t_0 are highly variable, potentially confounding comparability of these models. As t_0 is a positioning parameter that holds no biological meaning, we refit the Virginia model with t_0 fixed at each of the values estimated by historical studies. We compared fits of models with equal t_0 values using variance ratio tests (VRT; Zar 1996). Although the LRT is more powerful than the VRT (Cerrato 1990), LRTs require length at age data (that was unavailable for the historical data sets) for both models being compared to estimate variance for full and reduced models. A VRT forms an F-statistic from the ratio of the residual sum of squares of an alternative (in this case, historical) model plotted against a data set (current Virginia blueline tilefish) divided by the residual sum of squares of a model fit to that data set. A large F-value indicates that the alternative model fits that data much worse than the fitted model and is not likely representative of the population from which the plotted data were sampled.

To make comparisons among modern blueline tilefish from different geographical regions, the data collected during our sampling was supplemented by concurrent data collected by the NOAA Beaufort Laboratory from 2003-2011, which included samples caught in waters off Florida through Virginia. Some of these samples only had length measurements for either fork or total length. To use as much data as possible, a linear relationship between fork and total lengths was estimated based on individuals from both (our own and NOAA's) data sets that had both measurements, and missing values were imputed.

The second comparison of growth was between modern (caught during 2003 and later) blueline tilefish that were caught north and south of Cape Hatteras, North Carolina. We selected Cape Hatteras as our boundary for comparison under the assumption that connectivity would be more likely to occur on either side of, rather than across, this biophysical boundary. Cape Hatteras is located within NMFS statistical area 635 (Figure 2), with increasing numbers to the east and south. Therefore, blueline tilefish with reported catch locations were divided into northern (statistical areas less than 635) and southern (statistical areas greater than or equal to 635) groups. VB growth models were regressed upon total lengths at age and compared using likelihood ratio tests (Kimura 1980). To increase the probability of differences among parameters being attributable to true differences in growth rather than sampling variability, significance levels for regional comparisons were set at the $\alpha=0.01$ level.

Selectivity Bias Adjustment

A bias in size at age may be introduced by the fact that the majority of blueline tilefish collected during the present study were sampled through donations by recreational fishermen, while the majority of fish collected in the NOAA sample were acquired through samples of the commercial fishery. Discards are minimal for both sectors, reducing potential bias from high grading (SEDAR 2013). However, gear differences may have some impact on sizes of fish caught in each sector, setting a de facto minimum size limit on both fisheries (as hook size would limit the number of fish with smaller gape) and maximum size limit on the commercial fishery (as this sector is more likely concerned with total catch weight rather than individual fish size). We used AD Model Builder (Fournier et al. 2012) to address fishery sector and regional selectivity biases, at both small and large sizes, by re-fitting the VB model using a truncated length error distribution and then comparing parameter estimates with and without this adjustment (McGarvey and Fowler 2002; Schueller et al. 2014; for simulated example of methods see Appendix). To define ages subject to minimum length truncation due to gear selectivity patterns, we defined the minimum length limit as the minimum length observed for the first age of full selection. For each region, the first age of full section was defined as the age exhibiting maximum numbers-at-age plus one. All ages younger than the first age of full selection were subject to the minimum size limit. The maximum size limit was defined as the smaller of the unadjusted, region specific L_∞ estimates. Once defined, the maximum size limit was applied to all fully selected ages for the region exhibiting the smaller L_∞ estimate. We only applied the maximum size limit to the region with the smaller unadjusted L_∞ estimate because smaller asymptotic size in a region could be due to bias from reduced selectivity of larger sizes. We make the assumption that such a bias would be less evident in the region with the greater L_∞ estimate.

RESULTS

Sampling

A total of 2,104 blueline tilefish were collected by CQFE from 2009-2012, with at least 34 fish collected in each month of the calendar year. Blueline tilefish were caught at depths of around 50-200 m, typically in hard-bottomed areas. All fish, except five from a commercial trawl, were caught using rod-and-reel. Specimens were caught in the submarine canyons along the edge of the continental shelf east of the Virginia coast, most often Norfolk Canyon in NMFS statistical area 632 (Figure 1). The majority of our sample (n=1,752) came from fishery dependent sampling via donations by recreational anglers. Blueline tilefish collected by special charters (n=296) constituted 14% of our total sample, and 56 fish were purchased from the commercial fishery.

Length Frequency

Specimens ranged from 283 to 892 mm total length with an overall mean of 538 mm. No significant differences were found in mean length between fishery dependent and special charter samples or capture locations (Table 1), so all samples were pooled for analyses of length. Length frequencies were significantly different between sexes (Kolmogorov-Smirnov test: D=0.24, N=[Males:1108, Females:818], *P*<0.0001) (Figure 4). The mean total length of males (578 mm, SD=132) was significantly larger than that of females (508 mm, SD=123; Student's T-test (unequal variances): t=12.03, df=1825, *P*<0.0001). Overall mean length (539 mm) was smaller than mean lengths of fish caught off North Carolina (554 mm) and South Carolina (609 mm) in the 1970s (Ross and Huntsman 1982). Mean length differed significantly from mean lengths of blueline tilefish collected off the Carolinas in 1982-1987 (591 mm, SD=79; t=13.08, df=2447, *P*<0.0001) and 1996-1999 (524 mm, SD 72; t=3.57, df=1952, *P*=0.0004) (Harris et al. 2004).

Table 1. Multivariate analysis of variance (MANOVA) results for mean age and total length of sexually distinguishable adult blueline tilefish (*Caulolatilus microps*) collected off Virginia during 2009-2012, accounting for effects of collection method (fishery dependent or special charter), location, and sex (n=1341). ***Significant difference at $\alpha=0.05$. Student's t-tests for differences due to collection method had *P*-values of 0.2325 for total length and <0.0001 for age, and t-tests for differences due to sex had *P*-values of <0.0001 for total length and 0.9891 for age.

	MANOVA		
Effect	Levels	F	*P*-value
Collection Method	2: Fishery-Dependent, Special Charters	7.83	0.0004***
Location	3: NMFS Statistical Areas 625, 626, 632	2.86	0.0579
Sex	2: Female, Male	183.98	<0.0001***

Figure 4. a) Total length (n=1926) and b) age (n=1351) frequencies for male and female blueline tilefish (*Caulolatilus microps*) captured off Virginia from 2009-2012.

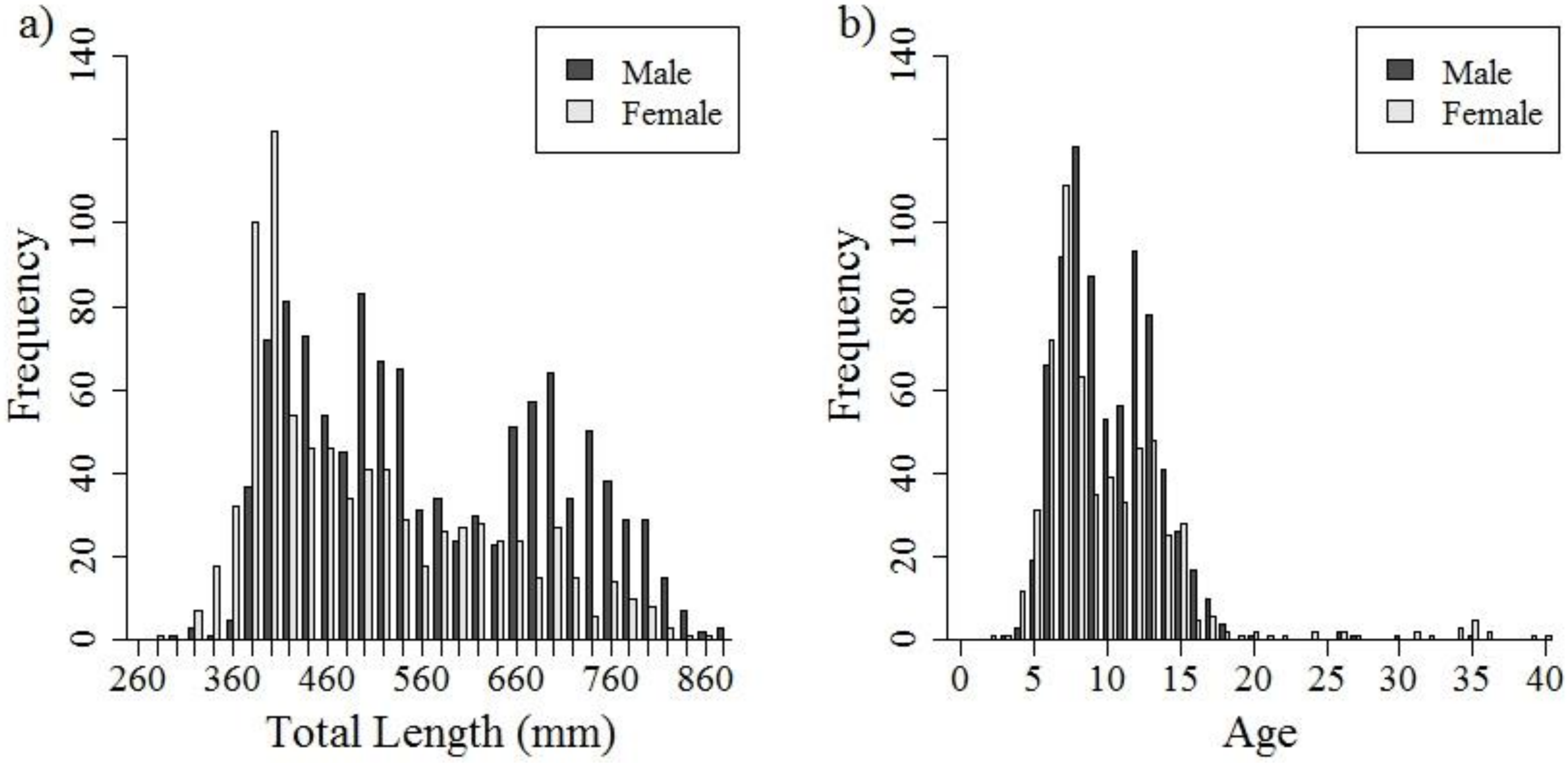
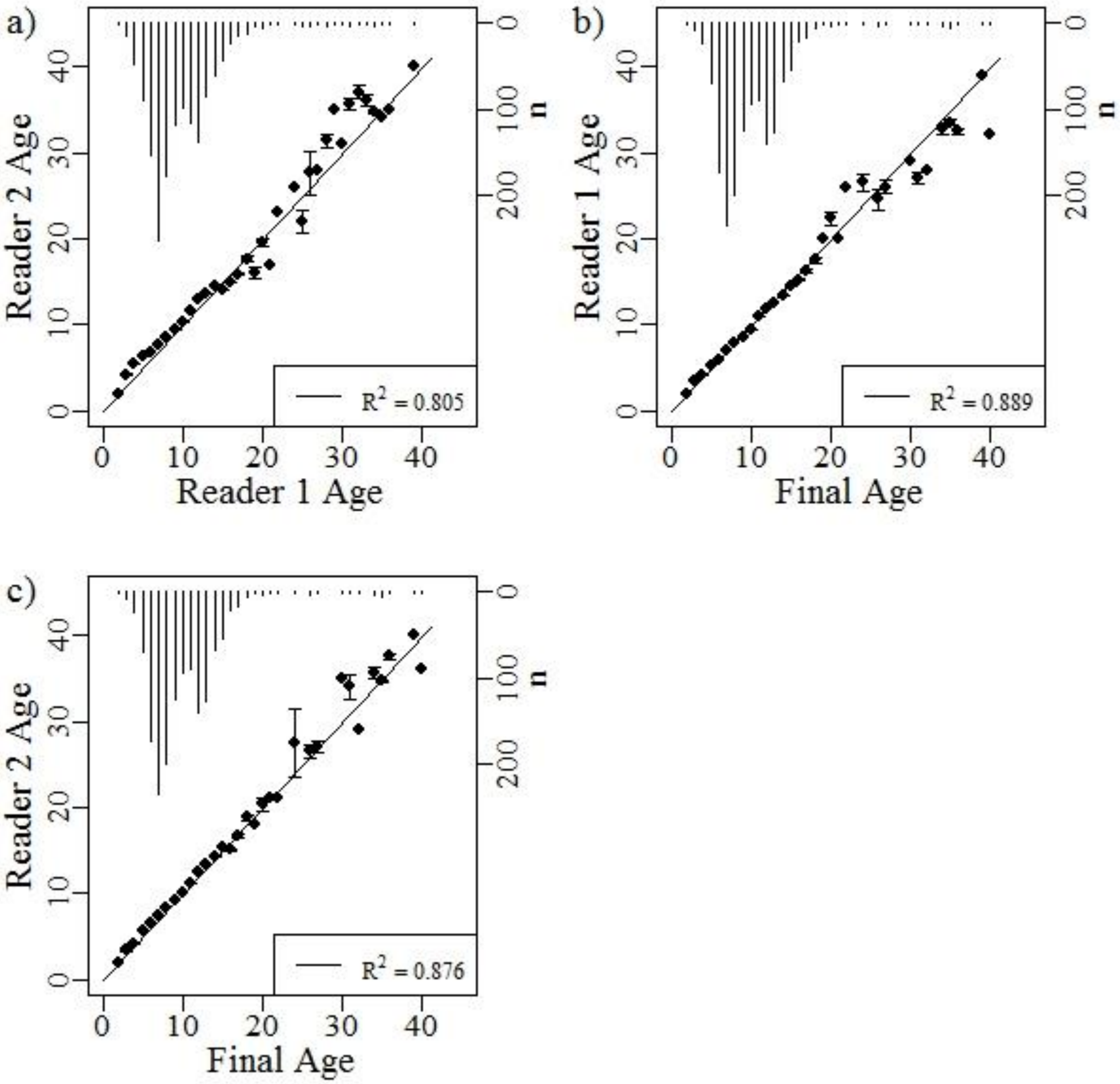


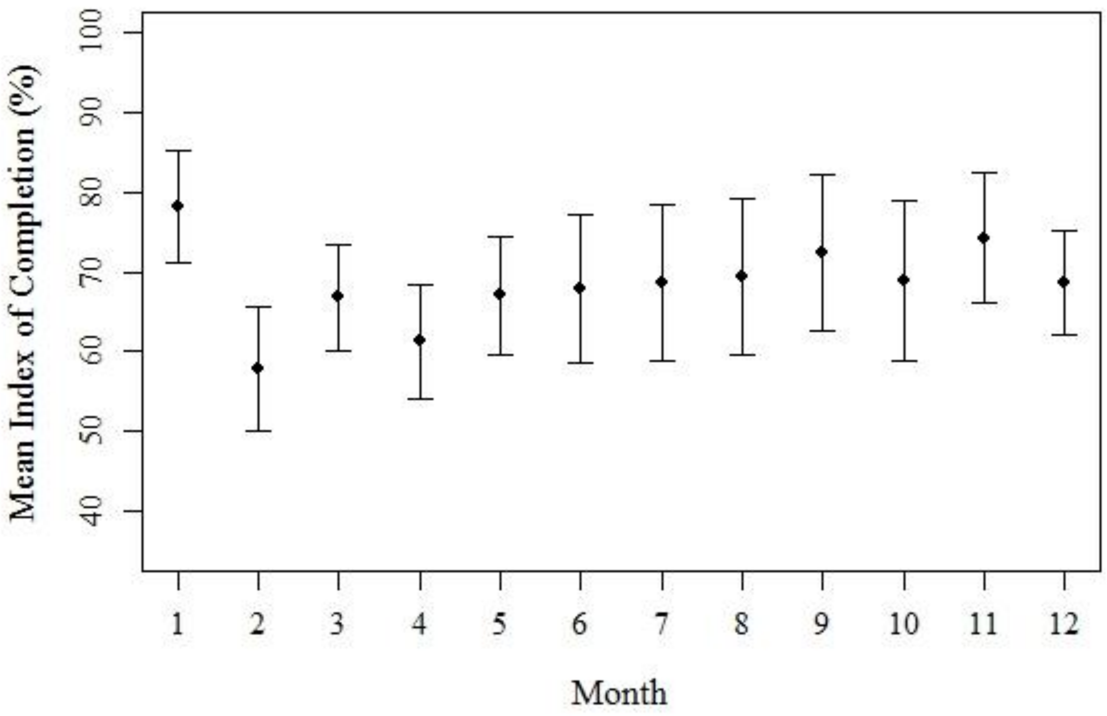
Figure 5. Bias plots of pairwise comparisons for independent and final age readings of blueline tilefish (*Caulolatilus microps*) captured off Virginia from 2009-2012. Diagonal lines represent 1:1 relationships between ages, with R^2 values for these relationships fit to data depicted in the legend. Error bars represent standard errors about mean ages assigned by the read depicted on the y-axis for all fish assigned an age by the read depicted on the x-axis. Vertical bars show numbers of fish (n) at age according to the read depicted on the x-axis.



Aging

Ages were determined for 967 of 983 fish collected from 2009-2011 and 517 of 1121 fish collected during 2012. Five otoliths were discarded from age analyses due to readers' inability to agree upon a consensus age. Ages ranged from 2 to 40 years with an overall mean of 10 years. Percent agreement between independent readings was 26%, with 60% and 81% of independent readings within 1 and 2 years of each other, respectively. APE between independent readings was 16%. Average bias between independent readings was significantly greater than 0 but significantly less than 1 (0.604; 95% Confidence Interval: [0.505, 0.703]). Variability among independent and final ages was high, but linear relationships among ages were well approximated as 1:1 (Figure 5). No significant differences were found in mean age among capture locations (Table 1). Mean age differed between fishery dependent (10 years) and special charter (9 years) samples (t=-3.56, df=389, *P*=0.0004), but while statistically significant, a 1 year discrepancy in mean age holds little biological significance for a fish with a maximum lifespan of over 40 years. Therefore, all fish were grouped across location and source of capture for analyses of age.

Figure 6. Monthly mean indices of completion with 95% confidence intervals for marginal increments of blueline tilefish (*Caulolatilus microps*) captured off Virginia from 2009-2012, (n=337).



Marginal Increment Analysis

Marginal increments for 337 fish captured during all months of the year across all years sampled were analyzed to validate periodicity of increment formation. Monthly samples ranged between 25 and 30 otoliths. Monthly mean marginal increments showed a great deal of variability, with the smallest mean indices of completion being observed in February and April (56.6% and 61.2%, respectively) (Figure 6). The limited range of observed values precludes strong conclusions about increment formation periodicity. However, we do note an overall increasing trend in monthly mean indices throughout a 1 year period, with mean indices at the beginning of that period, February and April, being significantly less than the mean index in the last month, January, according to 95% confidence intervals. Furthermore, the timing of minimum mean indices in February and April coincides with the timing of annual increment formation reported by Ross and Huntsman (1982). Therefore, we continued with age analyses under the assumption of 1 increment formed per year.

Virginia Growth Analyses

Age frequencies were significantly different between sexes (D=0.15, N=[Males:770, Females:581], P<0.0001), likely due to a greater number of females older than 20 years (Figure 4). However, mean age did not differ between sexes (10 years, t=-0.06, df=910, P=0.95). The lack of males at older ages resulted in an unrealistically high estimate of L_∞ of 1001 mm (maximum observed total length: 892 mm) when sex-specific VB models were regressed without weights on total lengths at age (Figure 7a). To increase the relative effect of individuals from rarer ages, sex-specific growth was remodeled using inverse frequency weighted lengths at age (Figure 7b). Estimates of t_0 were significantly different under both weighting scenarios (P<0.0001), potentially confounding comparability of biologically relevant parameters, L_∞ and κ. Therefore, models were refit to weighted lengths at age with t_0 fixed at the mean of the estimates from the original weighted regressions, -1.32 (Figure 7c). Males had significantly greater L_∞ than females and a similar κ (Table 2), indicating a faster growth rate.

Figure 7. Von Bertalanffy growth curves regressed upon a) unweighted total lengths (TL) at age (t), b) weighted total lengths at age, and c) weighted total lengths at age with t_0 fixed, for male (n=770) and female (n=579) blueline tilefish (*Caulolatilus microps*) captured off Virginia from 2009-2012. When applied, weights were calculated as the inverse of the sample size for a given age and sex.

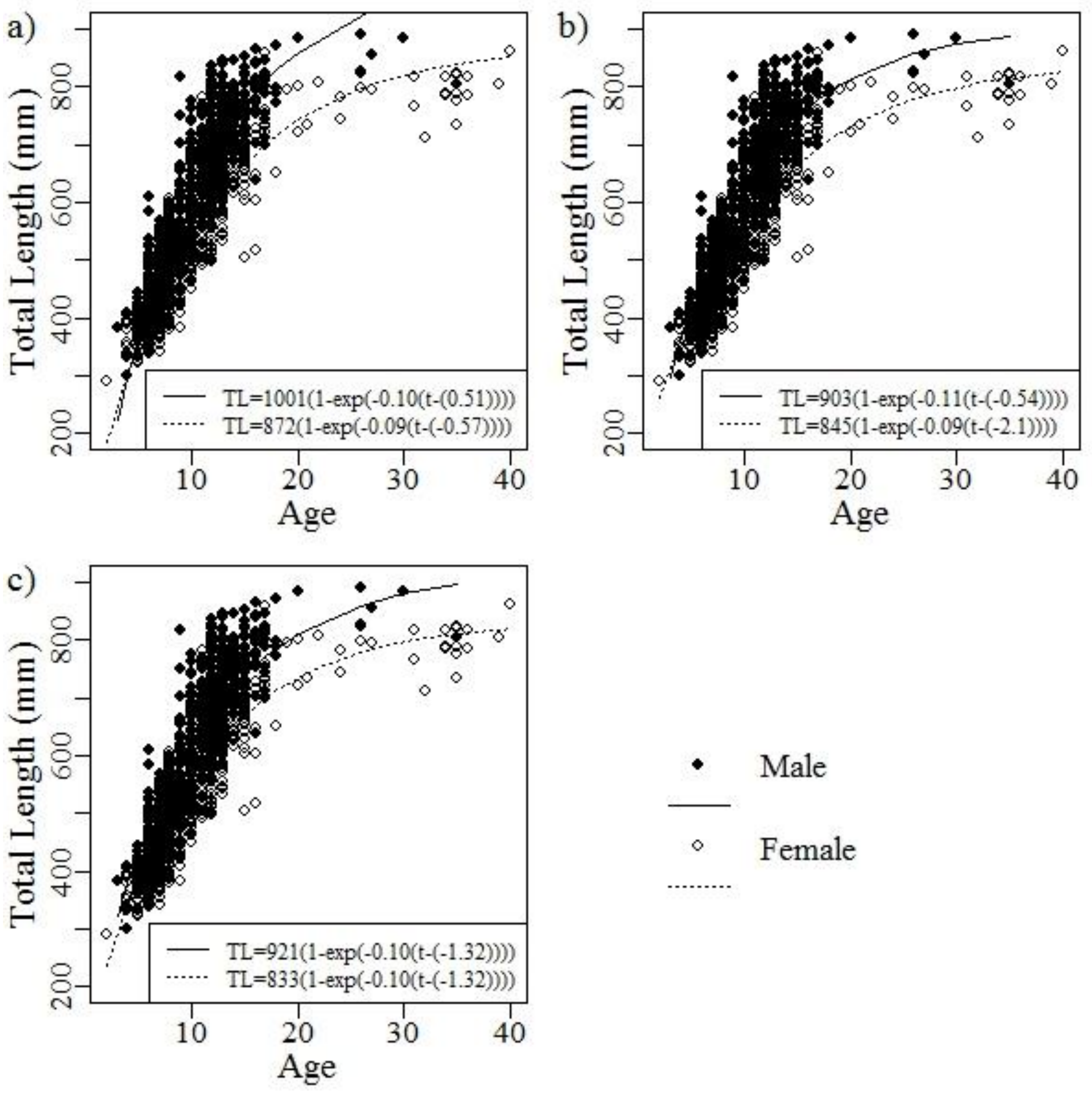


Table 2. Von Bertalanffy parameter estimates, with fixed t_0, of weighted total lengths at age for blueline tilefish (*Caulolatilus microps*) collected off Virginia during 2009-2012. Weights were calculated as the inverse of the sample size for a given age and sex. Chi square test statistics and *P*-values resulting from likelihood ratio tests of equality between parameter estimates are shown in the bottom rows. ***Significant difference at α=0.05.

	L_∞	κ	t_0	n
Male	921	0.10	-1.32	770
Female	833	0.10	-1.32	579
χ^2	111	0.14		
P-value	<0.0001***	0.7131		

Regional Growth Comparisons

Mean age of blueline tilefish from the present study (10, SD=4.34) was less than that reported for fish from the Carolinas during the 1980s (16.9, SD=7.9; t=19.43, df=631, *P*<0.0001) but did not differ from fish collected during the 1990s (10, SD=5.8) (Harris et al. 2004). The VB model fit to unweighted mean lengths at age for our study was not significantly different from that of US South Atlantic blueline tilefish in the 1970s (Ross and Huntsman 1982). Our model did differ from those estimated for US South Atlantic blueline tilefish during the 1980s and 1990s (Harris et al. 2004), with Virginia fish growing slower at young ages and to larger maximum sizes (Figure 8; Table 3).

Figure 8. Pairwise comparisons of Von Bertalanffy growth curves regressed upon unweighted mean lengths at age for blueline tilefish (*Caulolatilus microps*) captured off Virginia from 2009-2012 (n=1482) and off the US South Atlantic during a) 1972-77 (Ross and Huntsman 1982), b) 1982-87 (Harris et al. 2004), and c) 1996-99 (Harris et al. 2004). Fits to Virginia data are shown with t_0 variable (original) and fixed at the value estimated by each of the US South Atlantic models (adjusted). Error bars represent standard errors about the mean lengths at age.

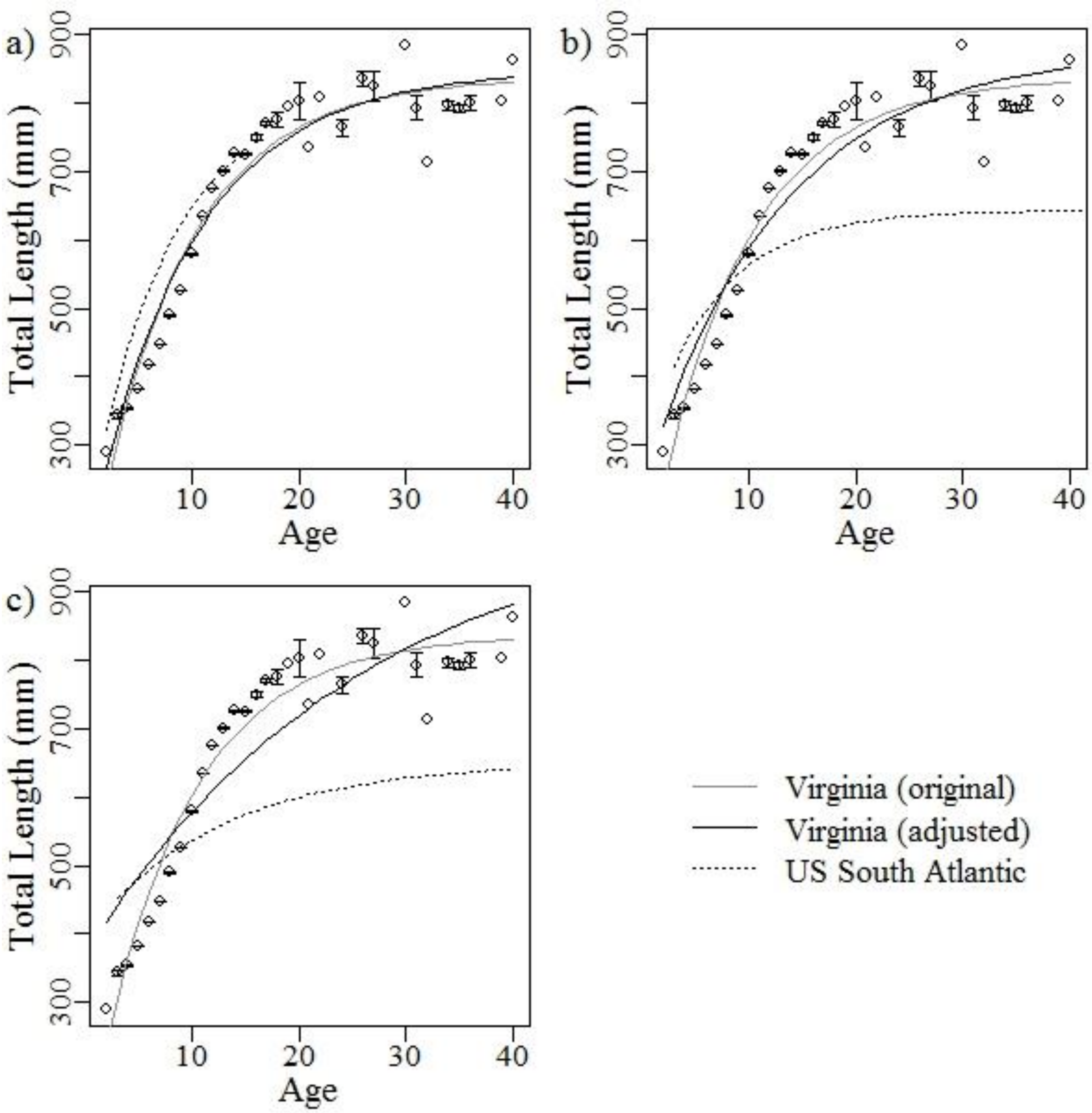
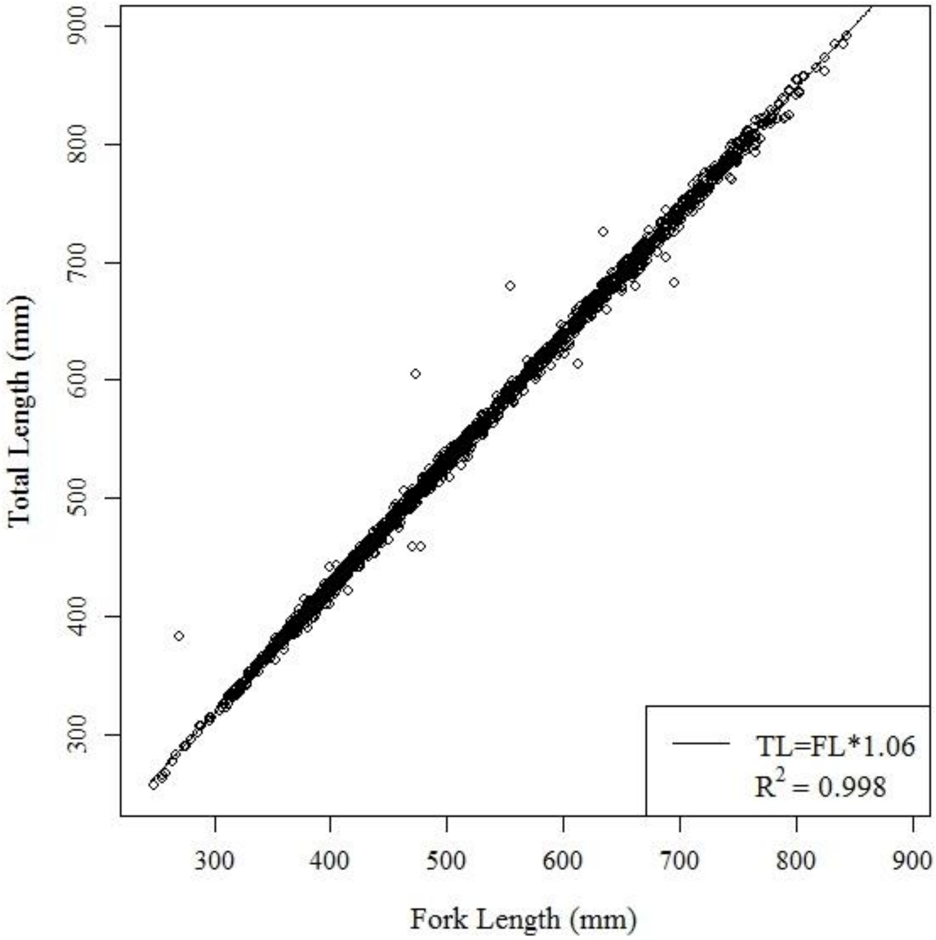


Table 3. Parameter estimates for von Bertalanffy growth models of unweighted mean total lengths at age for blueline tilefish (*Caulolatilus microps*) collected off Virginia during 2009-2012 (n=1482) and previous results for studies of blueline tilefish collected off the US South Atlantic (with sampling periods shown below each study). Models were compared by variance ratio tests between each of the past models and a current Virginia model regressed with t_0 fixed at the value estimated by each past model.

Study; Time Period	Parameter Estimates			Virginia Parameter Estimates (t_0 fixed)				
	L_∞	κ	t_0	L_∞	κ	t_0	F (df=31)	P
Present Study; 2009-2012	837	0.12	-0.84					
Ross and Huntsman 1982; 1972-1977	810	0.14	-1.64	848	0.10	-1.64	1.91	0.0769
Harris et al. 2004; 1982-1987	643	0.15	-3.88	879	0.08	-3.88	7.77	<0.0001
Harris et al. 2004; 1996-1999	651	0.08	-11.77	1026	0.04	-11.77	4.49	<0.0001

Figure 9. Linear regression of total and fork lengths for blueline tilefish (*Caulolatilus microps*) captured from 2003-2012 (n=2277).



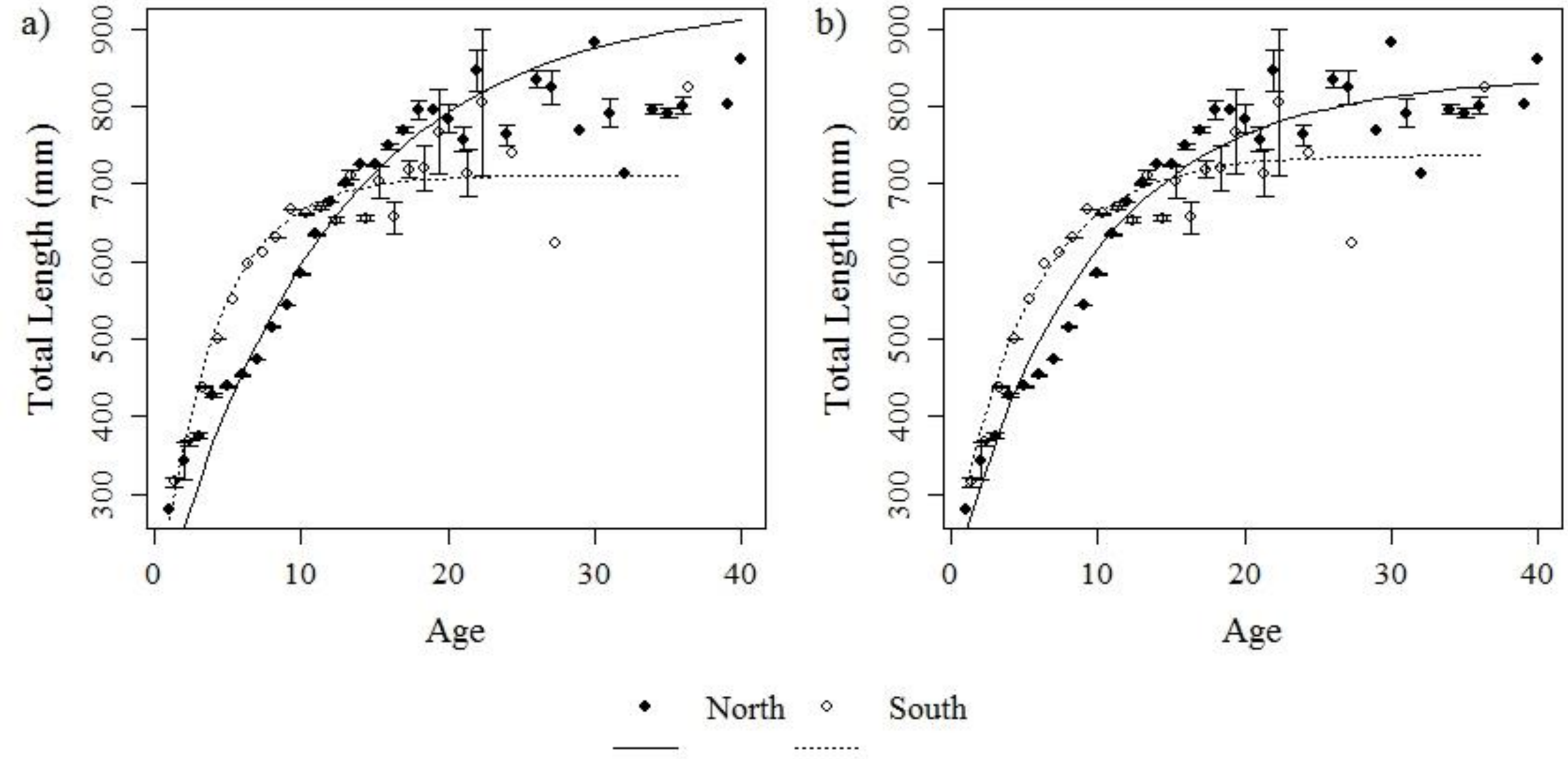
We observed a strong linear relationship between fork and total lengths for blueline tilefish collected from 2003-2012 (R^2=0.998; Figure 9), and used this relationship to impute missing values for modern data sets. Lengths at age for blueline tilefish collected from 2003-2012 varied between regions north and south of Cape Hatteras (Table 4, Figure 10). The unweighted regression of the VB model produced an L_∞ estimate for northern fish (936 mm) that was greater than the maximum length observed in that region (913 mm) (Table 4, Figure 10a), so

Table 4. Parameter estimates for Von Bertalanaffy growth models of unweighted and weighted total lengths at age for blueline tilefish (*Caulolatilus microps*) caught north (n=1737) and south (n=2627) of Cape Hatteras, NC, from 2003-2012. When applied, weights were calculated as the inverse of the sample size for a given age and sex. Chi square test statistics and *P*-values resulting from likelihood ratio tests of equality between parameter estimates are shown in the bottom rows. ***Significant difference at α=0.01.

	L_∞	κ	t_0
		Unweighted	
North	936	0.09	-1.76
South	711	0.26	-0.85
χ^2	123	128	7.30
P-value	<0.0001***	<0.0001***	0.0069***
		Weighted	
North	839	0.11	-2.31
South	739	0.19	-1.85
χ^2	262	142	4.64
P-value	<0.0001***	<0.0001***	0.0326

both regional models were refit to inverse frequency weighted lengths at age. The difference in t_0 estimates for the weighted regressions was less than 0.5 and not significant at the α=0.01 level. Therefore, we did not refit models with a fixed t_0. Blueline tilefish caught north of Cape Hatteras, NC, had significantly greater L_∞ and lesser κ estimates than those caught to the south (Table 4, Figure 10).

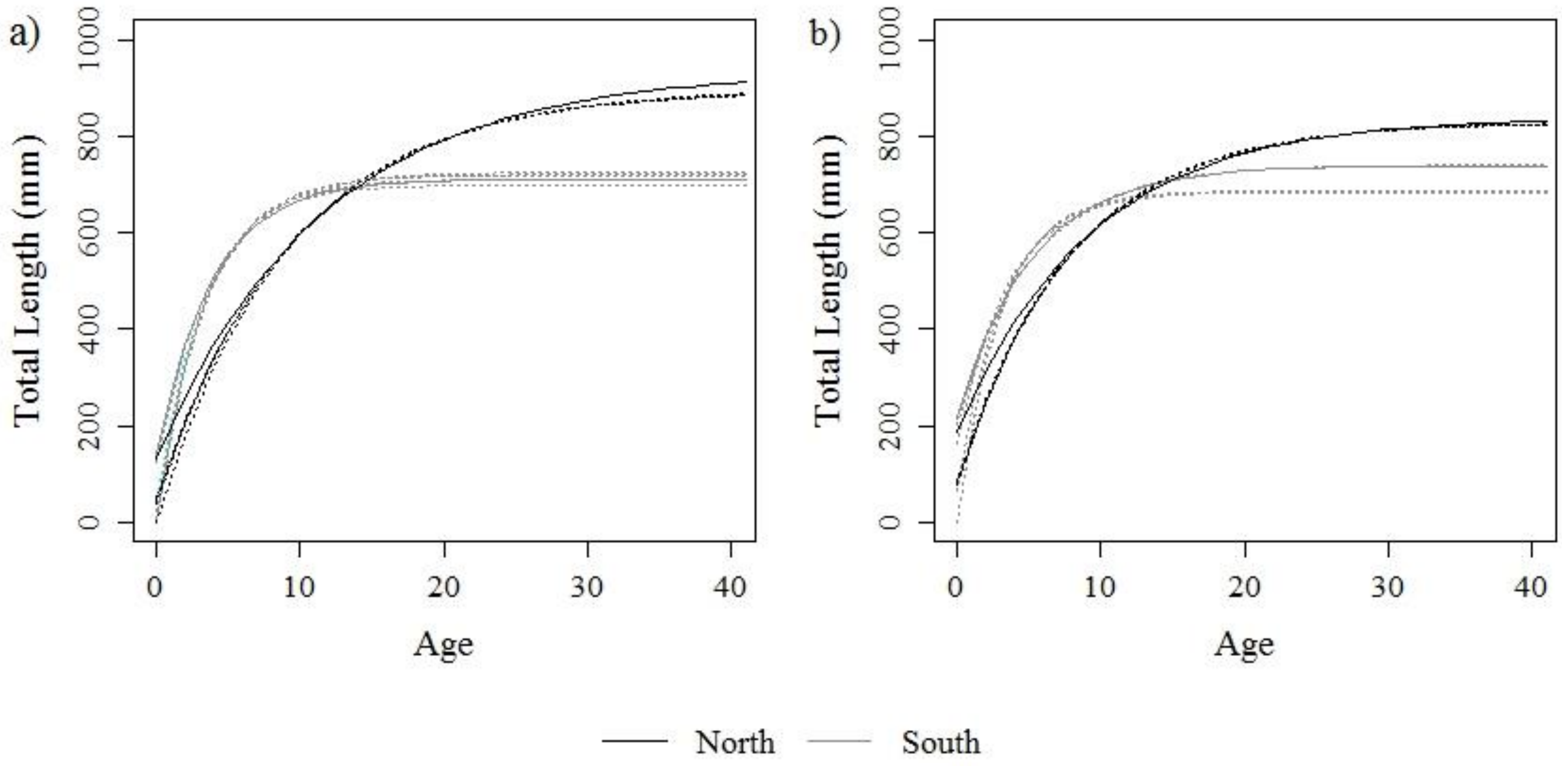
Figure 10. Von Bertalanffy growth curves regressed upon a) unweighted total lengths at age and b) weighted total lengths at age, for blueline tilefish (*Caulolatilus microps*) captured north (n=1737) and south (n=2627) of Cape Hatteras, North Carolina. When applied, weights were calculated as the inverse of the sample size for a given age and region. Data are shown as mean total lengths at age with standard errors about the means.



Selectivity Bias Adjustment

Weighted and unweighted regional models were refit using a truncated normal likelihood for partially selected ages with first age at full selection being defined as 7 years for fish caught south of Cape Hatteras (S) and 8 years for fish caught north of Cape Hatteras (N). Each region's model was refit several times using all combinations of age at full selection (7 or 8) and minimum length limits defined as minimum total lengths for first age at full selection (Minimum total length at Age 7: 393 mm (S), 340 mm (N); Age 8: 426 mm (S), 380 mm (N)). Maximum length limits were defined as the unadjusted L_∞ estimates, 711 mm and 740 mm, for unweighted and weighted runs, respectively, and applied to adjusted fits of US South Atlantic models. Adjusted models closely estimated unadjusted models, indicating minimal sampling bias for the original fits (Figure 11). This amount of bias would not explain differences in the regional growth curves under either weighting scenario. Thus, we conclude that the unadjusted growth curves and comparisons are representative of growth for blueline tilefish north and south of Cape Hatteras.

Figure 11. Von Bertalanffy growth curves regressed upon a) unweighted total lengths at age and b) weighted total lengths at age, for blueline tilefish (*Caulolatilus microps*) captured north (n=1737) and south (n=2627) of Cape Hatteras, North Carolina, with assumed normal (solid line) and truncated-normal (dashed lines) error distributions. When applied, weights were calculated as the inverse of the sample size for a given age and region. Model runs with truncated error distributions for both regions had lower length limits of 393 mm, 340 mm, 426 mm, and 380 mm for fish at least 7, 7, 8, and 8 years old, respectively. Model runs with truncated error distributions for the southern region included runs with and without maximum length limits of a) 711 mm and b) 740 mm.



DISCUSSION

Virginia blueline tilefish were long-lived and slow growing, as seen for other areas and times. Ages from 2 to 40 years were observed, with a similar range as samples of US South Atlantic blueline tilefish since the 1980s (Harris 2004). Slow growth was indicated by low estimates of κ, which was characteristic of US South Atlantic blueline tilefish during the 1970s (Ross and Huntsman 1982) and other deepwater reef fishes such as snowy grouper (Wyanski et al. 2000) and golden tilefish (Lombardi-Carlson 2012).

While the limited range of indices of completion disallows definitive conclusions concerning the periodicity of increment formation, mean marginal increments were smallest during February and April (Figure 5). This coincides with previously observed results that indicated annual increment formation between February and April (Ross and Huntsman 1982). The difficulty in measuring marginal increments is common for long-lived species as marginal increments typically decrease in size with increasing age, exaggerating errors due to processing or increment measurement (White et al. 1998; Filer and Sedberry 2008; Andrews et al. 2012). Alternative methods such as radiocarbon dating of historical samples would be better suited to validate aging techniques for blueline tilefish and other deepwater species that are similarly difficult to age. Preliminary results of a recent radiocarbon validation study of blueline tilefish ages from otoliths indicate that aging techniques used in this study either aged blueline tilefish accurate to or slightly under the true age (Ballenger et al., unpublished). Imprecision of aging methods is common among deeper, demersal species that experience little environmental seasonality (White et al. 1998; Wyanski et al. 2000; Filer and Sedberry 2008; Lombardi-Carlson 2012), and this holds true for blueline tilefish, as well. However, despite an APE of 16%, this study improves on previous precision levels for blueline tilefish, with initial readings for this study agreeing on 26% of samples aged and being within 2 years for 81% of samples aged (previous study: 24% agreement, 64% within 2 years; Harris et al. 2004).

Similar to blueline tilefish in the US South Atlantic (Ross and Hunstman 1982; Harris et al. 2004), blueline tilefish caught off Virginia exhibited sexually dimorphic growth (Figure 6), with males growing to significantly greater asymptotic total lengths than females, indicated by differences in L_∞ estimates. A lack of difference in κ estimates indicates that asymptotic length is approached at a similar relative rate between sexes, thus males grow at a faster absolute rate than females. Male mean total lengths at age were consistently greater than females for ages 5 and older and males were present in larger numbers for all size classes over 500 mm total length (Figure 4a). Previous research suggests that most US South Atlantic female blueline tilefish are sexually mature by this size (Ross and Merriner 1983; Harris et al. 2004), so the divergence in growth around this length is likely reflective of female energy reallocation from somatic to reproductive growth.

Growth of blueline tilefish in the US South Atlantic has changed throughout the course of this fishery's history. Asymptotic length, indicated by L_∞, has decreased in this region with increased catches during the 1980s and reduced catches maintained through the present (Ross and Huntsman 1982; Harris et al. 2004). The 1970s data set did not include any blueline tilefish older than 15 years (Ross and Huntsman 1982), which may be reflective of inconsistent aging (use of whole otoliths rather than sections) or sampling methods compared to other data sets investigated. This study did have a much smaller sample size (n=201) than other data sets, which may have limited the collection of rarer, old fish. Aging inconsistencies were less likely between the present study and that of Harris et al. (2004), as one of the calibration sets used during the SEDAR 32 Aging Workshop contained samples from this study (SEDAR 2013).

Growth of modern Virginia blueline tilefish was similar to that of blueline tilefish caught off the US South Atlantic during the 1970s, prior to the onset of larger-scale commercial fishing in this region. Although the 1970s data did not include the full age range for this species, growth was not asymptotic at 15 years old and followed a similar trajectory to that of the modern Virginia data (Figure 8a). Ross and Hunstman (1982) did observe slightly faster growth at the ages they collected. Considering the use of whole otolith aging by Ross and Huntsman (1982), ages from this study may be biased low due to reduced visibility of smaller increments near the edge of the otolith. Increased ages to account for such a bias would shift the growth curve from Ross and Hunstman (1982) towards that of the Virginia data, potentially increasing similarity between these curves. Growth curves of modern Virginia blueline tilefish differed significantly from US South Atlantic blueline tilefish collected since the 1980s, after the onset of commercial fishing, with notable differences in L_∞ (Figure 8b and c). This suggests that in the absence of substantial fishing mortality, blueline tilefish are able to reach significantly greater asymptotic sizes. The similarity between models also indicates that growth observed for blueline tilefish in the Virginia was characteristic of a lightly-exploited population, which corresponds with this region’s catch history. The sudden rise in US Mid-Atlantic commercial catches since our sampling period may have similar effects on growth as the historical maximum catch off the US South Atlantic during the 1980s, which resulted in smaller maximum sizes. The portion of the stock off the US Mid-Atlantic should continue to be monitored and managed to prevent negative effects of overfishing, including reduced abundance of larger individuals that are especially prized by the recreational fishery.

Due to our sample composition, several of the unadjusted growth models in this study showed concerning patterns in the fits of the data to the models, including residuals that tend in the same direction for several consecutive years (see Figure 8a) and L_∞ values that overestimate presumably asymptotic lengths observed at the oldest ages (see Figures 7a and 10a). We recognize that the directionality of residuals may be impacted by sampling of only the largest fish in partially-recruited young age classes. Therefore, the data points and resultant mean lengths at age would likely be overestimates, and a sampling design that more comprehensively captures lengths at the youngest ages would likely help improve the fits in this aspect.

Lengths at age showed a fairly sharp inflection point around 20 years old, which resulted in L_∞ estimates from unadjusted models that overestimated lengths at age for fish older than 20 years. We adjusted for the overestimates of L_∞ by weighting our sample by the inverse of the sample sizes at age, implicitly assuming that the L_∞ values were, in fact, overestimates of mean asymptotic length. The sharp inflection point suggests a possibility that the largest sizes, which should be seen at ages beyond 20 years, are not fully susceptible to the sampling gear, meaning that the true value of L_∞ would sit above the largest observed values. Use of the truncated normal error adjustment with a maximum size limit, as applied to the modern south of Cape Hatteras data, would improve the L_∞ estimate in this case. Our assumption that the limited numbers of fish at older ages were due to a lack of availability rather than gear selectivity stems from the tendency of hook-and-line gear to be less restrictive towards large fish than smaller fish. Growth model fits to length at age data could also be improved through exploration of models other than the von Bertalanffy function. In this study, we adhered to a longstanding convention of assuming a von Bertalanffy pattern of growth. We explicitly chose this model because of the unmanaged state of Mid-Atlantic blueline tilefish at the time this research was conducted. The goal of characterizing growth in this region, for the purposes of this study, was to provide growth parameters that would be usable for a wide variety of currently available stock assessment and management models, which necessitates the use of the von Bertalanffy model. Since this baseline information is now in place for assessment and management purposes, future studies could investigate alternative models or a piece-wise regression approach that would more accurately depict blueline tilefish growth.

Proportions of commercial, recreational, and fishery independent samples for each of the compared data sets varied across sampling times and locations. Specifically, the current Virginia and 1970s US South Atlantic data sets were both largely composed of fishery independent or specially chartered and recreational samples, while more recent data from the US South Atlantic had greater proportions of commercial samples. This type of discrepancy is unavoidable for this stock, due to the ways in which this fishery developed in each region. The sample composition, relative to fishing sector, is likely similar between the current Virginia and 1970s US South Atlantic data sets because both samples were collected prior to the onset of larger-scale commercial fishing in their respective areas. Modern growth data from different fishing sectors could not be compared within regions due to great discrepancies in sample sizes. To ensure regional differences in size-at-age between modern data sets were attributable to differences in growth, rather than selectivity between fisheries, we refit the modern regional VB models using truncated error structures that eliminated the probability of observing individuals in non-fully selected age classes that were smaller or larger than minimum and maximum size thresholds, respectively. VB models fit using truncated normal error structures showed little deviation from those fit using normally distributed errors (Figure 11), indicating minimal sampling bias due to differences in size-based selectivity between the prominent fishing sectors for each region.

Fishing can decrease asymptotic size of individuals within fished populations by targeting large individuals and reducing abundance across the entire population. Since faster-growing individuals are recruited to the fishery and caught at younger ages, fewer large individuals survive to their maximum reproductive potential (at the largest sizes assuming no senescence, which has not been observed in this species), increasing the relative genetic contribution of smaller individuals to the total population's reproduction. Over time, prevalence of fish with smaller asymptotic sizes may become distinctive for heavily-fished subunits of patchily distributed populations that lack connectivity with less heavily-fished areas. Furthermore, regional differences in fishing pressure that induce changes to growth would result in a stock that is regionally sub-structured. This type of growth structure could lead to differing levels of fishing productivity and resiliency among subunits. It also may be accompanied by differences in other characteristics, such as morphology, behaviors, reproduction, or genetics (Enberg et al. 2012). Blueline tilefish have not shown any evidence of movement between patches to homogenize the stock during the adult stage, but have also shown no evidence of genetically distinct population sub-units within the Atlantic stock (SEDAR 2017). Therefore, connectivity is likely maintained through mixing or movement of pelagic larvae or juveniles (Dooley 1978). Few surveys have identified juvenile or larval blueline tilefish; thus, little is known about this life stage. Further research on juvenile or larval blueline tilefish would be vital to characterizing connectivity, growth, recruitment, and reproductive sourcing for the Atlantic stock.

The results of this study indicate growth variation between blueline tilefish north and south of Cape Hatteras, NC. Blueline tilefish off Virginia currently exhibit growth characteristics similar to lightly-fished US South Atlantic blueline tilefish caught during the 1970s. However, with a recent increase in exploitation, US Mid-Atlantic blueline tilefish may be susceptible to overfishing and a similar fishery-induced reduction in growth as those in the US South Atlantic.

Although blueline tilefish do not exhibit genetic structuring across Cape Hatteras, variation in growth between regions may alter management reference points between regions (Law 2000; Heino et al. 2013; Maunder et al. 2016), which should be considered when forming management strategies for this stock. Blueline tilefish growth should continue to be monitored off Virginia and investigated in other areas of the US Mid-Atlantic so that any fishery or geographically induced variations in size at age can be detected and accounted for in future assessments.

CHAPTER 3

REPRODUCTIVE BIOLOGY OF BLUELINE TILEFISH (*CAULOLATILUS MICROPS*) FROM THE NORFOLK CANYON

INTRODUCTION

Blueline tilefish (*Caulolatilus microps*) are a demersal fish species that is patchily distributed along the continental shelf and slope waters of the North American Atlantic and Gulf coasts (Dooley 1978). Though previously thought to range from Campeche, Mexico as far north as Cape Charles, VA, catches during recent years have indicated a more northerly extent to their range, with landings as far north as Massachusetts (National Marine Fisheries Service (NMFS)[1]). Blueline tilefish harvests off the US Mid-Atlantic (Virginia-Massachusetts) had been relatively miniscule compared to those caught off the US South Atlantic (Florida-North Carolina) prior to 2014, when US Mid-Atlantic landings increased to nearly ten times the average from the previous ten years (NMFS[1]; Mid-Atlantic Fishery Management Council (MAFMC) 2016a) (Figure 1). This increase occurred during the first year of increased harvest restrictions on blueline tilefish in the US South Atlantic resulting from determinations of the 2013 stock assessment, Southeast Data, Assessment, and Review (SEDAR) 32, that the Atlantic stock of blueline tilefish was overfished and overfishing was occurring (SEDAR 2013).

---

[1] National Marine Fisheries Service (NMFS), Fisheries Statistics Division, 2017. Personal commun.

Movement of the Gulf Stream offshore near Cape Hatteras, NC, creates differing environmental conditions to the north and south of this biophysical boundary and the potential for local adaptations within populations whose ranges span it. Differences in growth of blueline tilefish north and south of Cape Hatteras have been observed, with blueline tilefish to the north growing more slowly and to larger maximum sizes than those to the south (Chapter 2). Blueline tilefish are generally believed to exhibit high site fidelity that may limit connectivity among patches during the adult stage, evidenced by observed burrowing behavior showing energetic investment in a local habitat (Able et al. 1987), a tagging study of the taxonomically, behaviorally, and geographically similar golden tilefish, *Lopholatilus chamaeleonticeps* that showed little movement during the adult life stage (Able et al. 1982; Grimes et al. 1983), and a behavioral study of the partially co-occurring sand tilefish (*Malacanthus plumieri*), which shows males and females defending mating and feeding territories, respectively (Baird and Liley 1989). Variations in growth may be due to differences in regional fishing histories, but also indicate limited adult mixing of populations between regions. Connectivity among adults is not to be confused with genetic connectivity, which occurs throughout the blueline tilefish range along the US Atlantic and Gulf of Mexico coasts (McDowell 2016; O’Donnell and Darden 2016; SEDAR 2017), evidenced by genetically homogenous fish caught in several different areas. Due to the high site fidelity seen in similar species that presumably also occurs in adult blueline tilefish, this genetic connectivity is likely preserved through the larval or juvenile life stages, of which little is known for this species. Drifter data provide a potential mechanism of pelagic larval dispersal that would lead to the genetic connectivity that has been observed among regions (Klibansky 2017; SEDAR 2017).

Reproductive studies provide information on sex ratios, spawning cycles, and productivity that can inform and enhance the efficiency of stock assessment models and management plans. Reproductive biology has been described for blueline tilefish off the US South Atlantic throughout the development of the deepwater reef fishery in this region (Ross and Merriner 1983; Harris et al. 2004). Blueline tilefish off the US South Atlantic are multiple batch spawners that spawn, on average, every other day from spring through fall (Harris et al. 2004). Batch sizes up to 95,000 oocytes have been observed and do not decrease throughout the spawning season (Harris et al. 2004). Although reproductive characteristics have been described for blueline tilefish off the US South Atlantic (Ross and Merriner 1983; Harris et al. 2004), currently no studies describe these characteristics off the US Mid-Atlantic. Considering the observed spatial differences in size at age that are likely reflective of different exploitation histories or local adaptations between regions (Chapter 2), reproductive characteristics may not be uniform throughout the US Atlantic. The recent increase in commercial landings for a previously unstudied and unmanaged portion of this stock off the US Mid-Atlantic further amplifies the need for reproductive data that is representative of the entirety of this stock's range. This chapter will describe the reproductive characteristics of blueline tilefish caught in the Norfolk Canyon, off the southern portion of the US Mid-Atlantic, in terms of sex ratio, spawning seasonality, and fecundity and will compare these characteristics to those previously described for blueline tilefish off the US South Atlantic to determine whether these characteristics vary across space and time.

MATERIALS AND METHODS

Blueline tilefish were collected from the Norfolk Canyon during 2009-2014. Specimens were collected from commercial and recreational fisheries, as well as from collections by scientists from the Old Dominion University Center for Quantitative Fisheries Ecology and Virginia Marine Resources Commission (VMRC) aboard recreational charter vessels (hereinafter referred to as “special charters”). For special charters conducted during 2013-2014, bottom temperature was measured at fishing sites using a HOBO U12 Deep Ocean Temperature Logger lowered on an additional fishing line. This line was left on the bottom for a minimum of 5 minutes, and bottom temperature was identifiable on the data readout by an approximately constant temperature during the time that the logger was on the bottom. Recreational samples were primarily collected through the Virginia Marine Sportfish Collection Project, a freezer program conducted by VMRC, through which anglers donated carcasses to scientific research after filleting them at local cleaning stations. Total and fork length (TL and FL, respectively) measurements, sagittal otoliths, and macroscopic determinations of sex and reproductive phase according to the Virginia Finfish Sexual Maturity Index (Virginia Index) (Table 5) were taken for all fish collected. Total weight was measured for all whole fish, and gonads were extracted from fresh specimens, weighed, and fixed in 10% formalin for a minimum of 24 hours. All gonads used were preserved as fresh, never frozen specimens.

Transverse sections of sagittal otoliths were aged under a Nikon SMZ1000 microscope at 20-40x magnification, using transmitted light. Increments consisted of one translucent and one opaque zone, and were counted independently by 2 readers without knowledge of fish size or time of capture. If independent counts differed, the slide was recounted by both readers until a consensus age could be agreed upon. If no age could be agreed upon, the specimen was

Table 5. Virginia Marine Resources Commission (VMRC) Sexual Maturity Index used to macroscopically determine reproductive phases.

VMRC Sexual Maturity Index	Juvenile	Male	Female
0	Any fish which cannot be distinguished as male or female.		
1		Any fish which can be distinguished as a male. The testes will have few if any blood vessels and a flattened exterior side.	Any fish which can be distinguished as a female. The ovaries have many blood vessels and are tubular.
2		Any male with whitish testes. These usually have more form and are hard.	Any female with large ovaries. Usually the liquid within them is colored (species specific).
3		Any male with large white testes which have viable sperm. When testes are cut, the milt will flow out.	Any female with the presence of eggs. These may be small, granular looking inside the ovary.
4		Any male with large deflated testes. There may be some sperm remaining and an increase in blood vessels may occur.	Any female with ripe eggs. White and clear (species specific).
5			Any female with large deflated ovaries. Look for the presence of eggs remaining or a similarly colored liquid as seen in phase 2 females.

discarded from age from age analyses. Aging was attempted for all specimens collected from 2009-2011, as well as a proportionally allocated subsample of the 2012 specimens, based on the 2009-2011 data (Quinn and Deriso 1999).

After fixation, gonad samples were rinsed in water, preserved in 80% ethanol, and sent to the Florida Fish and Wildlife Conservation Commission's Fish and Wildlife Research Institute for histological sectioning and staining. Samples were embedded in glycol methacrylate, sectioned to 3–5-µm thickness, stained with periodic acid–Schiff's hematoxylin, and then counterstained with metanil yellow (Quintero-Hunter et al. 1991). Sections were viewed using transmitted light at 40-400x magnification under an Olympus BX41 microscope. Microscopic sex, germ cell developmental stages, reproductive state, and reproductive phases were assigned based on criteria from Brown-Peterson et al. (2011). The following histological indicators were used in females: primary growth (PG), cortical alveoli (CA), vitellogenic (Vtg1-3), and oocyte maturation (OM) stage oocytes and post-ovulatory follicles (POFs). Fish with secondary growth oocytes (SG) were considered mature (Lowerre-Barbieri et al. 2011). Actively spawning females were identified based on late OM, ovulation, or fresh POFs. Male histological indicators included the germinal epithelium (continuous or discontinuous), primary spermatocytes, secondary spermatocytes, primary spermatogonia, secondary spermatogonia, spermatids and spermatozoa. Males with spermatozoa present were considered mature.

Microscopic sex assignments were compared with macroscopic assignments to evaluate whether macroscopic sexes were reliable enough for use in calculating sex ratios. Sex ratios were calculated for the entire sample, as well as within months, ages, and 50 mm TL bins. Binomial tests were used to determine if proportions of males (p) significantly deviated from a 1:1 ratio with females (p=0.5).

Previous studies suggest that blueline tilefish are multiple batch spawners (Harris et al. 2004), thus annual fecundity is a function of the spawning frequency (number of spawning events per season) and batch fecundity. Spawning season was determined using reproductive phase frequencies. We also tested whether macroscopic phase criteria according to the Virginia Index (Table 5) were an appropriate proxy for related microscopic phase criteria according to the standardized terminology of Brown-Peterson et al. (2011) using sex-specific analyses of variance (ANOVA).

Batch fecundity was estimated using the hydrated oocytes method (Hunter et al. 1985) for gonads that were collected from freshly caught females and preserved in formalin. Three samples ranging from 32 to 77 mg were taken from random locations of preserved ovaries that had been histologically identified as containing hydrated oocytes. Hydrated oocytes were identified and counted for each sample. The average number of oocytes per mg was calculated and multiplied by the preserved weight of the gonad. Preserved gonad weight (PW; g) was calculated from fresh gonad weight (FW; g) using a conversion factor for scamp (*Mycteroperca phenax*) that has previously been used for US South Atlantic blueline tilefish (Harris et al. 2004): $PW=FW*0.897+1.148$. Log-transformed batch fecundity was linearly modeled against age and TL. One-way analyses of covariance (ANCOVA) were used to determine whether month of capture affected the slopes of these relationships. We used a variance ratio test (Zar 1996) to test whether the relationship for batch fecundity at TL differed from that observed in blueline tilefish off the US South Atlantic during the 1990s (Harris et al. 2004).

The presence and deterioration level of POFs indicated recent spawning activity at the time of death. Rates of deterioration can be impacted by water temperature, which we observed (on average) as approximately 13°C for areas where we captured blueline tilefish in Norfolk

Canyon during May, July, and August. POFs were aged as Day 0 (spawning within 24 hours of death), Day 1 (spawning 24-48 hours prior to death), or Day 2 (spawning >48 hours prior to death) according to criteria previously described for species that spawn in similar temperatures (10-16°C) (Ganias et al. 2014): *Merluccius hubbsi* (Macchi et al. 2004), *Engraulis anchoita* (Pájaro et al. 2009), *Brevoortia aurea* (Macchi and Acha 2000), *Ramnogasster arcuata* (Rodrigues et al. 2008), and *Percophis brasiliensis* (Militelli and Macchi 2001). Terminology describing components of annual fecundity of batch-spawning fishes was defined by Lowerre-Barbieri et al. (2011). Daily spawning fraction, or the proportion of females spawning daily (Hunter and Macewicz 1985; Murua et al. 2003; Stratoudakis et al. 2006; Ganias 2009; Murua et al. 2010) was estimated using the time-calibrated method (Fitzhugh et al. 1993; Wilson and Nieland 1994). This method assumes that POFs are identifiable up to 48 hours after spawning (Ganias et al. 2014) and that oocytes in the late stages of vitellogenesis will be hydrated and eventually spawned within the next 12 hours (Brown-Peterson et al. 1988; Nieland et al. 2002). Proportions of females collected during spawning months with each stage of POF (Day 0, Day 1, and Day 2) were calculated. These proportions were averaged to estimate the daily spawning fraction. The spawning interval, or the time between spawning events during the spawning season, was estimated as the reciprocal of the spawning fraction (Almatar et al. 2004; Murua and Motos 2006). Spawning frequency, or the number of spawning events within a spawning season, was estimated as the length of the spawning season divided by the spawning interval (Lowerre-Barbieri et al. 1996; Murua et al. 2003; Murua et al. 2006). Annual fecundity was calculated by multiplying batch fecundity by the spawning frequency.

RESULTS

We collected a total of 2,293 blueline tilefish ranging in TL from 283 mm to 892 mm, with an average of 533 mm. Of these 1,484 fish were aged. Ages ranged from 2 to 40 years, with an average of 10 years. Since approximately 83% of our total sample was collected from the freezer donation program, weights from whole fish and fresh gonads were available only for subsets of our data. Descriptive statistics of TLs and ages for our entire sample as well as these subsets are shown in Table 6.

Table 6. Descriptive statistics of ages and total lengths of blueline tilefish collected from the Norfolk Canyon from 2009-2014. Information includes sample types (Sample), sample sizes (n), minimum total length or age (Min.), maximum total length or age (Max.), and mean total length or age.

	Total Length (mm)				Age (years)			
Sample	n	Min.	Max.	Mean	n	Min.	Max.	Mean
Total	2293	283	892	533	1484	2	40	10
Whole Fish	354	289	884	547	261	2	34	9
Fresh Gonads	307	289	854	539	238	2	34	9

Table 7. Monthly sex ratios (M/F) of blueline tilefish collected from Norfolk Canyon from 2009-2014. Monthly proportions of males (M) were tested against a 1:1 correlation with females (F) (p=0.5) by a binomial test. ***Significant difference at α=0.05 level.

Month	n	M/F	*P*-value	
Jan.	244	3.6038	<0.0001	***
Feb.	205	0.9524	0.7800	
Mar.	135	1.4107	0.0579	
Apr.	60	1.0000	>0.999	
May	97	0.8302	0.4168	
June	251	1.0744	0.6137	
July	416	0.8824	0.2203	
Aug.	312	1.1224	0.3358	
Sept.	34	3.8571	0.0008	***
Oct.	128	1.0317	0.9296	
Nov.	59	1.1071	0.7948	
Dec.	163	3.9394	<0.0001	***

Macroscopic and microscopic sex assignments were highly correlated (98% agreement), so macroscopic sexes were used to calculate sex ratios (Tables 7-9). The overall sex ratio was male skewed (1.28 M: 1 F, n=2104, P<0.0001). Males were caught in significantly greater proportions during September, December, and January (Table 7). Sex ratio varied amongst ages with significantly more males at ages 8-16 years and more females at the youngest (4 years) and oldest ages (18+ years) (Table 8). Sex ratio varied among TLs with significantly more females at 300-400 mm and more males at TLs greater than 500 mm (Table 9).

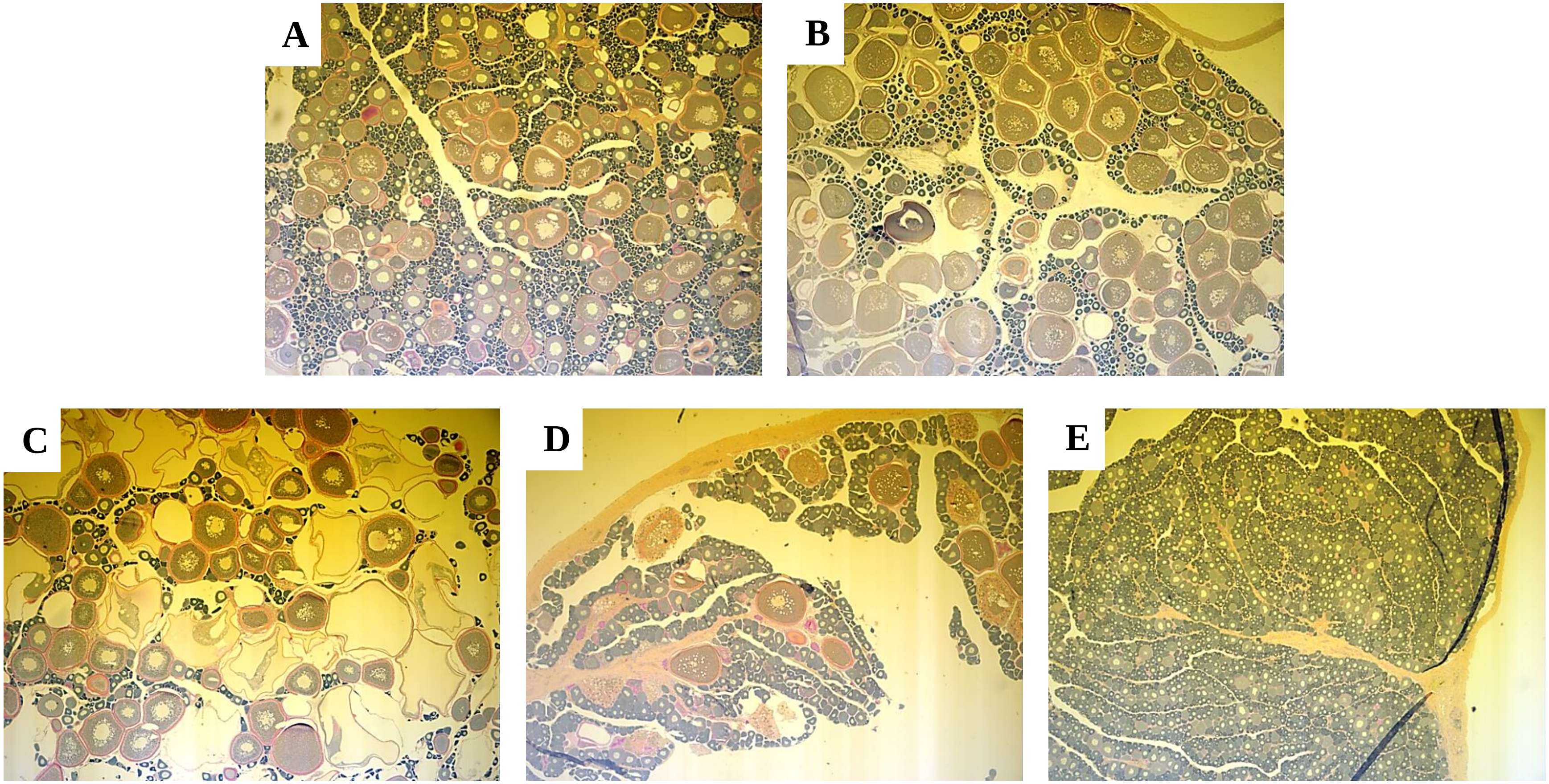
Table 8. Sex ratios (M/F) of blueline tilefish collected from Norfolk Canyon from 2009-2014 divided into 50 mm total length (TL) bins with minimum and maximum lengths of Min TL and Max TL, respectively. Bin proportions of males (M) were tested against a 1:1 correlation with females (F) (*p*=0.5) by a binomial test. ***Significant difference at α=0.05 level.

Min TL (mm)	Max TL (mm)	n	M/F	*P*-value	
251	300	1	0	>0.999	
301	350	28	0.4000	0.0357	***
351	400	239	0.3352	<0.0001	***
401	450	424	0.9186	0.4091	
451	500	264	1.2000	0.1568	
501	550	303	1.6814	<0.0001	***
551	600	159	1.8393	0.0002	***
601	650	145	1.2308	0.2449	
651	700	175	2.4314	<0.0001	***
701	750	175	2.5714	<0.0001	***
751	800	121	3.6538	<0.0001	***
801	850	59	3.9167	<0.0001	***
851	900	9	3.5000	0.1797	

Table 9. Sex ratios (M/F) of blueline tilefish collected from Norfolk Canyon from 2009-2014 divided into age bins. Bin proportions of males (M) were tested against a 1:1 correlation with females (F) (*p*=0.5) by a binomial test. ***Significant difference at α=0.05 level.

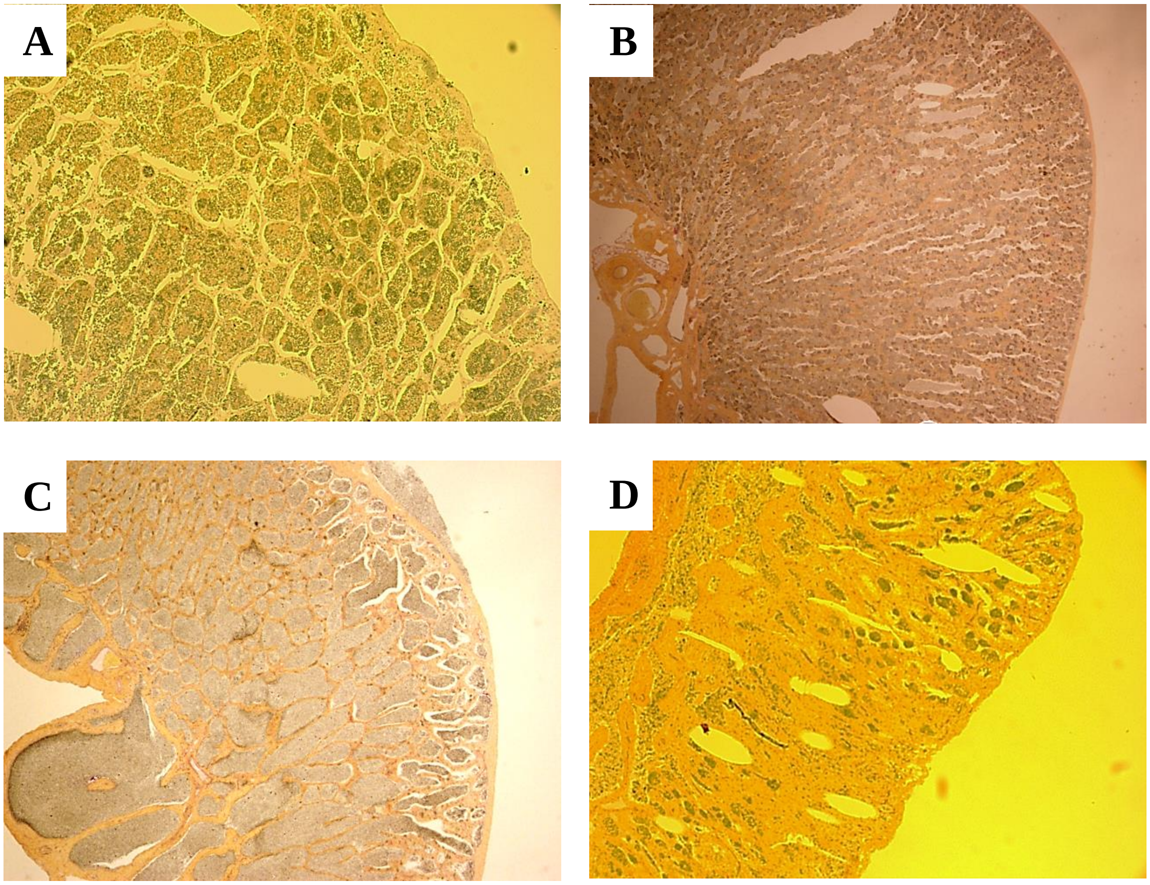
Age	n	M/F	*P*-value	
2	1	0	>0.999	
3	2	1.0000	>0.999	
4	16	0.2308	0.0213	***
5	50	0.6667	0.2026	
6	137	0.9296	0.7327	
7	201	0.8440	0.2590	
8	181	1.8730	<0.0001	***
9	122	2.4857	<0.0001	***
10	92	1.3590	0.1750	
11	89	1.6970	0.0192	***
12	139	2.0217	<0.0001	***
13	126	1.6250	0.0095	***
14	66	1.6400	0.064	
15	55	0.9642	>0.999	
16	22	3.4000	0.0169	***
17	16	1.6667	0.4545	
18+	37	0.3704	0.0076	***

Figure 12. Histological sections of female blueline tilefish gonads collected from Norfolk Canyon during 2009-2014. Phases according to terminology used by Brown-Peterson et al. (2011) are: A) Developing, B) Spawning Capable, C) Spawning Capable-Actively Spawning Sub-phase, D) Regressing, E) Regenerating.



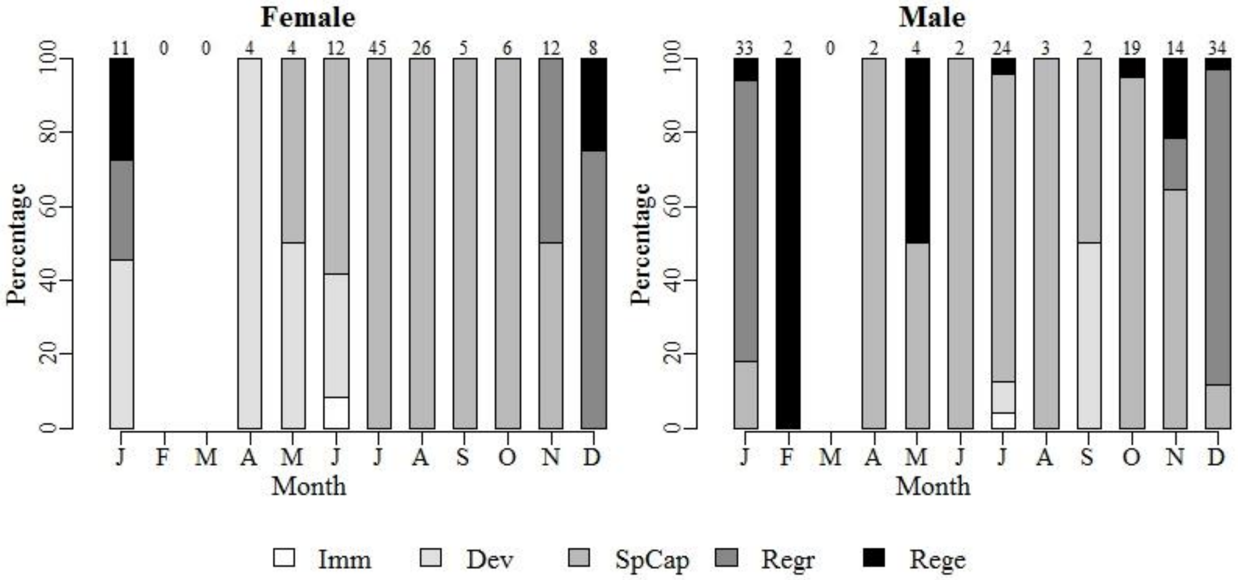
All microscopic reproductive phases were observed in both females (Figure 12) and males (Figure 13). Spawning capable females were microscopically identified from May-November (Figure 14). Proportions of spawning capable females increased from the previous month in May and June, were at one hundred percent from July through October, and decreased to 50% in November. Most of the microscopically-examined females collected in December were regressing, and most of the microscopically-examined females collected in January and April were developing. No females caught in February or March were examined

Figure 13. Histological sections of male blueline tilefish gonads collected from Norfolk Canyon during 2009-2014. Phases according to terminology used by Brown-Peterson et al. (2011) are: A) Developing, B) Spawning Capable, C) Regressing, D) Regenerating.



microscopically. Spawning capable males were microscopically identified in January and April-December (Figure 14). Both of the two microscopically-examined males caught in February were regenerating, and no males caught in March were examined microscopically.

Figure 14. Monthly percentages and sample sizes (above) of microscopic female and male reproductive phases for blueline tilefish captured in Norfolk Canyon from 2009-2014, determined according to phases and criteria defined by Brown-Peterson et al. (2011) (Juv=Juvenile, Dev=Developing, SpCap=Spawning Capable, Regr=Regressing, Rege=Regenerating).



Both macroscopic and microscopic phase assignments were available for 130 females and 134 males, and these groups are the subjects of the following analyses. Macroscopic reproductive phase assignments of females were significantly correlated with microscopic reproductive phases (F=19.51, *P*-value<0.0001, df=1). Most notably, 94 (96.9%) of 97 spawning capable females (determined microscopically), were identified macroscopically as phase 3 or 4 by the Virginia Index (Table 10). Females that were macroscopically identified as phase 3 and 4 by the Virginia Index were microscopically identified as spawning capable for 80 (81.6%) of 98 phase 3 females and 14 (87.5%) of 16 phase 4 females. Twelve (12.2%) of 98 females macroscopically identified as phase 3 were regressing. These twelve females were all captured during November, December, or January. No females captured in December or January were microscopically identified as spawning capable, and of twelve females captured in November with both macro- and microscopic reproductive phase assignments (all macroscopically identified as phase 3), 6 (50%) of them were microscopically identified as regressing (Table 10,

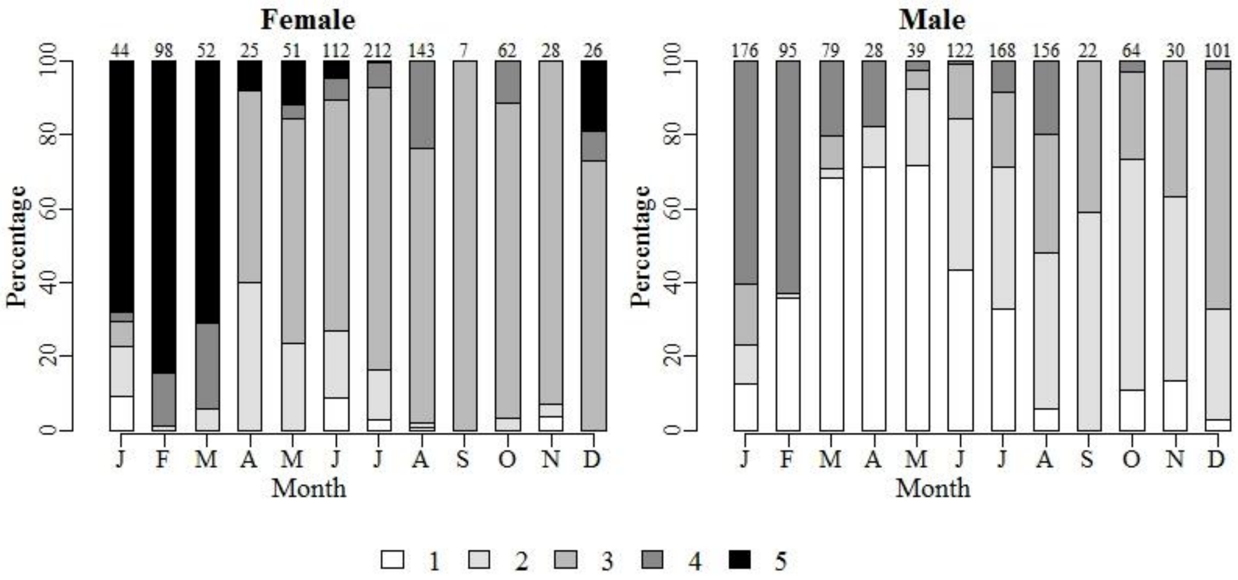
Table 10. Numbers of macroscopic and microscopic reproductive phase assignments for blueline tilefish collected from Norfolk Canyon from 2009-2014 that were assessed using both methods. Macroscopic phase assignments (1-5) were determined according to the Virginia Finfish Sexual Maturity Index. Microscopic phase assignments (Juvenile-Regenerating) were determined according to terminology and criteria from Brown-Peterson et al. (2011).

	Female							
		Macroscopic						
	Phase	0	1	2	3	4	5	Total
Microscopic	Juvenile	0	0	1	0	0	0	1
	Developing	0	4	0	5	0	3	12
	Spawning Capable	0	0	2	80	14	1	97
	Regressing	0	0	0	12	1	2	15
	Regenerating	0	0	0	1	1	3	5
	Total	0	4	3	98	16	9	130
	Male							
		Macroscopic						
	Phase	0	1	2	3	4		Total
Microscopic	Juvenile	0	0	1	0	0		1
	Developing	0	1	2	0	0		3
	Spawning Capable	0	10	40	12	0		62
	Regressing	0	0	10	46	0		56
	Regenerating	0	7	4	1	0		12
	Total	0	18	57	59	0		134

Figures 14 and 15). Macroscopic reproductive phase assignments of males were also significantly correlated with microscopic reproductive phases (F=4.46, *P*-value=0.037, df=1). Of 57 males macroscopically identified as phase 2, 40 (70.1%) were identified as spawning capable (Table 10). Of 59 males macroscopically identified as phase 3, 46 (78.0%) were identified as regressing (Table 10).

While spawning capable females, spawning capable males, and regressing males could be predicted with reasonable accuracy via macroscopic assessments of gonads, other reproductive phases were less identifiable from macroscopic assessments. Additionally, male macroscopic phases did not correspond as clearly to a defined season as females, with spawning capable males being microscopically identified in 10 of 11 months sampled (Figure 15). Males macroscopically identified as phase 2 using the Virginia Index composed less than half of all males collected in all months except October (61.5%). Therefore, female macroscopic phase assignments were used with female microscopic assignments to estimate the spawning season. Spawning months were estimated as months when female macroscopic reproductive phase assignments identified as phase 3 or 4, according to the Virginia Index, composed at least 50% of female macroscopic phase assignments and at least one female was microscopically identified as spawning capable. Within spawning months, the start and end date of the spawning season was determined by the earliest and latest dates when females macroscopically identified as reproductive phase 3 or 4 were collected. Under these definitions, the spawning season is estimated to extend from May 5 through November 30 (209 days). Microscopically identified spawning capable males composed at least 50% of microscopically-assessed males collected during May-November (Figure 14).

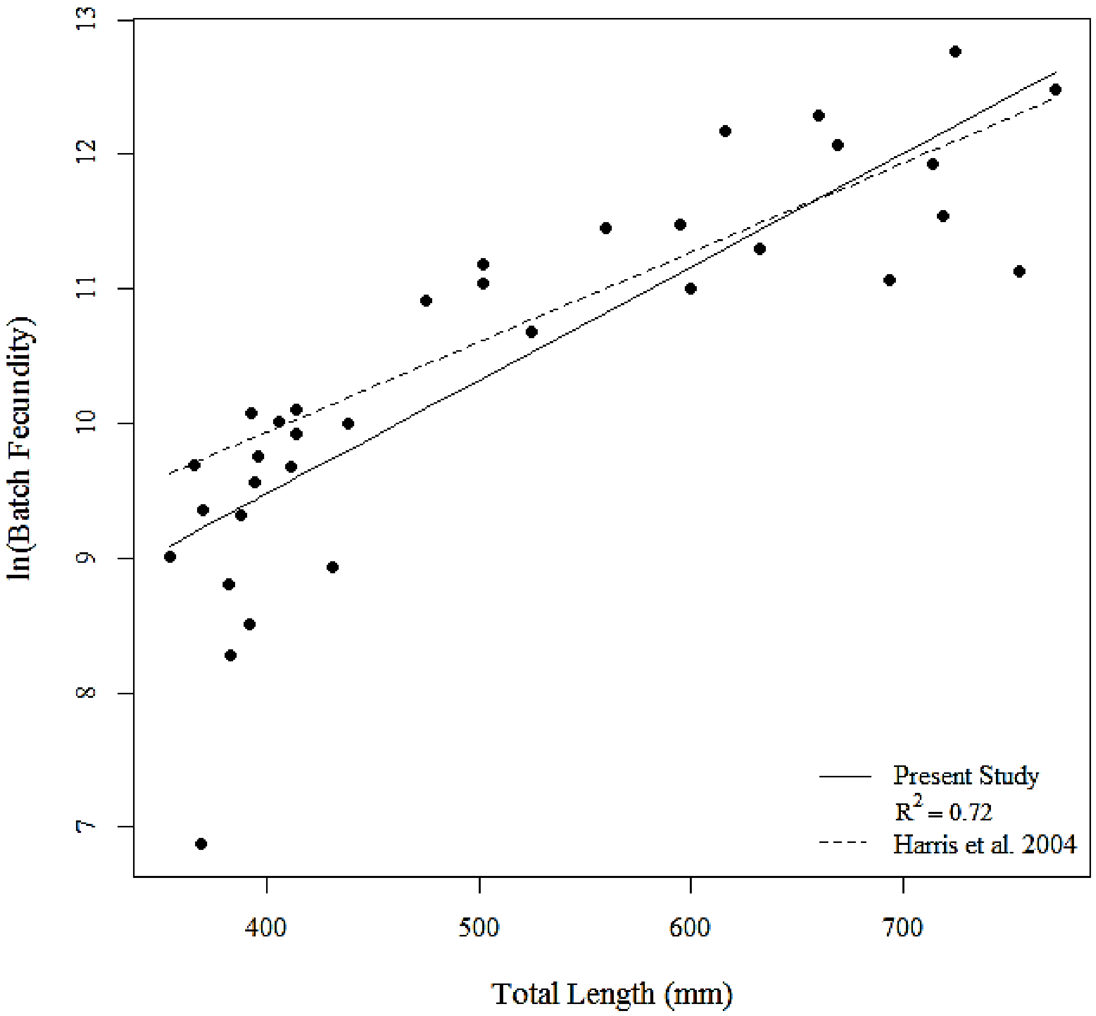
Figure 15. Monthly percentages and sample sizes (above) of macroscopic female and male reproductive phases for blueline tilefish captured in Norfolk Canyon from 2009-2014, determined according to phases and criteria defined by the Virginia Finfish Sexual Maturity Index. Phase 0 fish, or fish that cannot be distinguished as male or female (typically juveniles), are not included in these graphs.



Due to a lack of younger and smaller fish among the specimens that had fresh gonads extracted, we were unable to estimate age and length at first maturity. We observed one immature female that was 4 years old and 322 mm TL. The youngest mature female observed was 3 years old (382 mm TL) and the smallest mature female was 332 mm TL (4 years old). Gonads from both fish contained hydrated oocytes. We observed one immature male that was 5 years old and 395 mm TL. The youngest mature male was 3 years old (384 mm TL) and the smallest mature male was 379 mm TL (7 years old). Both fish showed signs of ongoing spermatogenesis.

Similar to those off the US South Atlantic, blueline tilefish from the Norfolk Canyon are multiple batch spawners, as indicated by the simultaneous presence of POFs and multiple stages of yolked oocytes in female gonads collected throughout the spawning season (Figure 12)

Figure 16. Batch fecundities at total lengths (TL) for blueline tilefish captured in Norfolk Canyon from 2009-2014. The regression of batch fecundity at total length from Harris et al. (2004) (dashed line), originally fit to female blueline tilefish ranging in size from 366-629 mm TL, was extrapolated to the size range observed in the present study (354-772 mm TL) for comparison.



(Ganias et al. 2015). Batch fecundity (BF) ranged from 965 (369 mm TL, Age 6) to 351,718 (725 mm TL, Age 13) oocytes and varied in a log-linear fashion with TL [ln(BF)=6.10+TL*8.43E-3; R^2=0.72] and age [ln(BF)=7.57+Age*0.32; R^2=0.66]. Month of sample collection did not have a significant effect on either of these relationships (ANCOVA: TL: F=0.33, *P*-value=0.57, df=2; Age: F=0.12, *P*-value=0.73, df=2). Batch fecundity at length was not significantly different from that previously observed for US South Atlantic blueline tilefish (F=1.25, df=33, *P*-value=0.53; Harris et al. 2004) (Figure 16). The mean proportion of females with Day 0 POFs (0.837) sampled during the spawning season was higher than those with Day 1 (0.704) or Day 2 (0.673) POFs by over 10%, potentially indicative of oversampling Day 0 females, as has been observed in other species (Ganias et al. 2014). Therefore, we adjusted the proportion of Day 0 POFs by setting it equal to the proportion with Day 1 POFs (0.704) (Picquelle and Stauffer 1985; Ganias et al. 2014). The average daily spawning fraction was 0.694. The average spawning interval was 1.441 days, corresponding to approximately 145 spawning events during the 209 day spawning season. Thus, annual fecundity ranged from 139,902 to 51,006,236 oocytes.

DISCUSSION

Blueline tilefish captured off the coast of Virginia from 2009-2014 had a sex ratio that deviated significantly from 1:1. The overall ratio was male-skewed, while various skews were observed within size groups, ages, and months of capture. Males were more prevalent at the largest sizes and intermediate ages (8-16), reflecting sexually dimorphic growth that has previously been observed in this species off the US South Atlantic (Ross and Huntsman 1982; Harris et al. 2004). The sex ratio observed off the US South Atlantic during the late 1990s (1.18

M: 1 F) was also male-skewed, though this deviation from a 1:1 ratio was not statistically significant. Although we observed a significantly greater overall proportion of males, we also observed a greater proportion of females at the oldest ages, meaning a substantial proportion of females have been able to grow to near their maximum size and presumably reach their full reproductive potential. Sex ratio only deviated from 1:1 during three months: January, September, and December, with males being more prevalent in all three. In January and December, most females caught were in a regressing or regenerating phase of their reproductive cycle, possibly reducing feeding behavior and susceptibility to capture. Females in other species have been caught more frequently while actively spawning, and this may be the case for blueline tilefish, as well (Ganias et al. 2014). Variations in sex ratio throughout the year may also be impacted by behavior. A taxonomically and geographically similar burrowing fish species, sand tilefish, settles in aggregations in which they form polygynous mating units (Baird and Liley 1989). If similar aggregating behaviors occur in blueline tilefish, sex ratio at lower sampling levels (i.e., across a small number of trips or sampling locations, as was the case for September [34 fish over four trips]) could be attributable to where the sample was collected. If the sample happened to be taken from an area near a female aggregation, females would be observed in greater proportions. At the same time, female aggregations would suggest large amounts of area with few to no females, where males would be more prevalent.

A lack of immature fish in our histological sample prevented estimation of age or length at first maturity. However, the ages and sizes of the 2 immature fish we did observe and smaller or younger mature individuals indicate that age at first maturity is likely 4 or younger in females and 5 or younger in males and most fish from both sexes mature prior to reaching 400 mm TL. These results are consistent with those observed for 4 immature blueline tilefish collected off the

US South Atlantic during the 1980s and 90s (Harris et al. 2004). An earlier study of US South Atlantic blueline tilefish observed later ages and larger sizes at first maturity (Ross and Merriner 1983). The transition to earlier and smaller maturation was possibly impacted by the emergence of commercial fishing in that area, as commercial landings peaked during the 1980s. This level of fishing was not present during or prior to sampling for this study. Similar to golden tilefish (*Lopholatilus chamaleonticeps*), blueline tilefish adults are generally believed to exhibit high site fidelity, potentially indicating that inter-regional connectivity between blueline tilefish in the US Mid and South Atlantic would be more likely to occur during the larval or juvenile life stages. A lack of juvenile and larval blueline tilefish samples currently limits knowledge about mechanisms of connectivity within the Atlantic stock during these life stages and prohibits strong conclusions about the timing of maturation in this species.

Blueline tilefish in the Norfolk Canyon exhibited all phases of the male and female spawning cycles, showing that the population there is capable of local reproduction. Protogynous hermaphroditism has been suggested to occur in this species (Ross and Merriner 1983), but we did not observe any gonads with both male and female reproductive structures. The range of TLs for females was from 289 mm to 862 mm and for males was 283 mm to 892 mm, showing significant overlap and further supporting a gonochoristic reproductive strategy. Although adequate sample sizes of non-spawning, female microscopic reproductive phases were not available to evaluate the accuracy of non-spawning macroscopic reproductive phase assignments, female macroscopic phases did correspond to the spawning capable microscopic phase with a reasonable amount of predictability. Monthly proportions of macroscopic phase 3 or 4 females according to the Virginia Index (which corresponded to the spawning capable microscopic phase of Brown-Peterson et al. 2011) were obviously lowered from January through

March. While macroscopic phase 3 or 4 females were identified in December, zero of eight females that were also evaluated for microscopic reproductive phase were identified as spawning capable (six were regressing and two were regenerating). We interpreted these results as errors in the macroscopic phase assignments and concluded that blueline tilefish likely do not spawn in December. We observed a decline in spawning activity during November and potentially December, as well as a lack of spawning activity during January through March, indicated by low proportions of spawning capable males and females and high proportions of regressing and regenerating gonads. Previous studies have suggested regression and regeneration occurring in December and January based on limited numbers of histological samples. While our microscopic samples in those months are also low, our macroscopic phase assignments are greater and indicate high percentages of non-spawning females from December through March, supporting hypotheses of oocyte regression, regeneration, and development during these months.

Despite low numbers of histologically evaluated females caught in February-May, information from macroscopic reproductive phase assignments can inform conclusions about the beginning of the spawning season. Lowered proportions of spawning capable-corresponding macroscopic phases from January-March indicated little to no spawning during these months. In April, the proportion of spawning capable-corresponding macroscopic phases increased to around 43% of females collected. However, only one female collected in April was evaluated both macroscopically and microscopically for reproductive phase (macroscopic: 3; microscopic: developing), limiting our ability to definitively conclude whether spawning occurs during this month. Collectively, our results indicate a spawning season that most likely occurs from May through November. Considering the potential for macroscopic phasing errors and limits of sample sizes, the season may begin as early as April and end as late as December. Further histological sampling would be useful in evaluating spawning levels during these months.

The estimated spawning season of May-November is generally overlapped by that estimated for US South Atlantic blueline tilefish in previous studies. Samples collected during the 1970s off the Carolinas estimated a spawning season of April through October, with peaks in female spawning activity in May and September (Ross and Merriner 1983). Samples collected during the 1980s and 1990s off the Carolinas estimated a spawning season from March through October, with a peak in female spawning activity in May (Harris et al. 2004). Neither of these studies collected histological samples in November, December, or January, although Ross and Merriner (1983) did measure gonadosomatic index (Hunter and Macewicz 1985) of 3 females caught during November that indicated minimal spawning activity. Considering spawning females were captured in October for both previous studies of US South Atlantic blueline tilefish and very few fish were captured in November, it is possible that spawning in this region may extend into November, as well. Our results indicate a peak in female spawning activity for the US Mid-Atlantic (indicated by the greatest monthly proportions of microscopically identified spawning capable females) from July through October, when 100% of females assessed via histological samples were spawning capable. Based on results in other areas and potential errors due to monthly sample sizes, the actual peak month for US Mid-Atlantic blueline tilefish is likely in the earlier part of this range, in July or August.

Blueline tilefish from the Norfolk Canyon show a similar multiple batch spawning strategy as those from the US South Atlantic. Estimates for batch fecundity were consistent with previous estimates for US South Atlantic blueline tilefish, although spawning frequency and the timing and duration of the spawning season may differ. The relationship we observed between

TL and batch fecundity was not significantly different from that observed off the US South Atlantic during the 1980s and 90s (Harris et al. 2004). This relationship was maintained even though we observed a larger TL range (354-772 mm) than had previously been used to fit the Harris et al. (2004) model (366-629 mm). We also estimated more frequent spawning (spawning interval of 1.44 days in our study versus 2 days in Harris et al. 2004) over a shorter spawning season (209 days in our study versus 240 days in Harris et al. 2004), resulting in greater estimates of annual fecundity at length than those observed off the US South Atlantic during the 1980s and 90s. Differences in spawning frequency and annual fecundity may be exaggerated due to rounding differences between our study and Harris et al. (2004). Harris et al. observed a spawning fraction of 0.56, corresponding to a spawning interval of approximately 1.79 days, which was rounded up to 2 days. Using the raw spawning interval estimate would have resulted in approximately 134 spawning events over a 240 day season and closer estimates of annual fecundity to those observed in the US Mid-Atlantic.

Differences in the timing and frequency of spawning may be impacted by differences in environmental conditions between regions. Latitudinal variation in spawning season has been observed in marine fishes, with fish at higher latitudes typically beginning the spawning season later in the year and ending the season earlier than their lower-latitude counterparts (Conover 1992). Photoperiod, rather than temperature, has been suggested as an environmental cue for spawning in blueline tilefish because of minimal seasonal temperature fluctuations for the depths in which they live (Ross and Merriner 1983). However, Ross and Merriner (1983) observed two spawning peaks, shown by monthly percentages of ripe females and gonadosomatic index for males and females captured off South Carolina, in May and September, months that differ in photoperiod by approximately two hours of light per day. Our study shows a potentially later spawning season start date for blueline tilefish from the US Mid-Atlantic than those previously studied in the US South Atlantic, but does not show evidence of an earlier termination of the spawning season as would be expected from a pattern of countergradient latitudinal variation (Conover 1992). This indicates that factors other than photoperiod may impact blueline tilefish spawning, the spatial scale compared by current studies of blueline tilefish reproduction is too small to distinguish latitudinal effects, or further sampling in both regions is needed in February-May and November-December to more precisely define the spawning season.

Although regional growth patterns indicate population differences north and south of Cape Hatteras (Chapter 2), blueline tilefish in these regions have similar reproductive strategies. Similar fecundities at length with larger sizes and an increased spawning frequency indicate potentially greater productivity in the northern population. However, a similar age at maturity with a slower growth rate also makes this population susceptible to overfishing. The Atlantic stock, specifically in the US South Atlantic, has shown a reduction in age and size at maturity since the rise of the commercial fishery in the 1980s (Ross and Merriner 1983; Harris et al. 2004), potentially indicative of population juvenescence from overexploitation. This reduced age and size at maturity is maintained for the northern portion of the stock, in the US Mid-Atlantic, despite a relative lack of exploitation in this region prior to 2014.

The current catches of blueline tilefish in the US Mid-Atlantic, despite minimal historical catches suggest a resource that has been present but previously unexploited or a recent northward expansion in this species' range. With few historical records on this fishery, even within the US South Atlantic (sparse records of general tilefish landings, likely including multiple species of tilefish, date back to roughly the 1960s, although landings specifically of blueline tilefish were not identified until the 1980s), the answer to whether either of these scenarios is true becomes difficult to uncover. However, landings records and the few studies that have examined the biology of blueline tilefish (Ross and Huntsman 1981; Ross and Merriner 1983; Harris et al. 2004) in addition to the present study provide a basis for current and future management. With the recent increase in landings off the US Mid-Atlantic, this portion of the Atlantic stock should be monitored and managed to maintain numbers of large females seen in this region, which have the greatest fecundities. Further research on larval and juvenile transportation, as well as recruitment sourcing, would provide information to determine whether and to what extent this increased fecundity translates into increased regional contribution to the stock.

CHAPTER 4

SIMULATION ANALYSES OF MODAL TOTAL ALLOWABLE CATCH ESTIMATES FOR DATA-LIMITED STOCK ASSESSMENT MODELS AND A COMPOSITE MANAGEMENT PROCEDURE APPLIED TO US MID-ATLANTIC BLUELINE TILEFISH (*CAULOLATILUS MICROPS*)

INTRODUCTION

Fish stocks for which scientists are unable to develop a reliable index of abundance have been broadly categorized as "data-limited" and require the use of methods that do not rely upon such indices to characterize and set reference points for the fishery. Assessment methods for these stocks continue to be developed, adjusted, and refined in an effort to reduce the number of unmanaged stocks worldwide. These methods are often based on a time series of harvest, but can be further informed by auxiliary data inputs, such as state of depletion or growth, through the use of data that have been observed, "borrowed" from another species, or estimated by expert opinion (e. g. MacCall 2009; Smith et al. 2009; Berkson et al. 2011; Dick and MacCall 2011; Martell and Froese 2013). Blueline tilefish (*Caulolatilus microps*), found along the United States' Atlantic coast, are an example of such a data-limited stock that has been the topic of recent management discussions.

The blueline tilefish fishery has undergone many changes over the last 20 years, including changes in exploitation, biological sampling patterns, life history, and management, producing a scenario where a portion of the Atlantic stock of blueline tilefish off the Mid-Atlantic Bight of the United States (Virginia-Massachusetts; hereafter, US Mid-Atlantic) would be considered data-limited. Blueline tilefish have historically been found off the southeastern coast of the United States, being caught as far north as southern Virginia but most commonly being caught off the Carolinas (Dooley 1978; Parker and Mays 1998). Since the mid-2000s, blueline tilefish have been caught and landed in greater numbers at the northern end of this historical range and even further north, with landings as far north as Massachusetts (National Marine Fisheries Service (NMFS)[1]). In 2014, during implementation of increased catch restrictions off the South Atlantic region of the US (east coast of Florida-North Carolina; hereafter, US South Atlantic), commercial landings in the US Mid-Atlantic increased to a time series maximum for this region of over ten times the average of the previous ten years (NMFS; Mid-Atlantic Fishery Management Council (MAFMC) 2016a) (Figure 1). In 2016, the MAFMC enacted emergency restrictions on blueline tilefish harvest and tasked their Scientific and Statistical Committee (SSC) to recommend catch limits for the US Mid-Atlantic. Due to a lack of blueline tilefish in fishery independent surveys, a historically minimal commercial fishery, a questionable level of reporting historical recreational landings, and directed fishing efforts by a relatively small number of vessels that did report blueline tilefish landings in the US Mid-Atlantic, the SSC was unable to construct an appropriate index of abundance for blueline tilefish in this region (MAFMC 2016b). Thus, data-limited methods were investigated to estimate appropriate allowable catch limits.

The Data-Limited Methods Toolkit, programmed into the R package DLMTool, includes management strategy evaluation (MSE) and total allowable catch (TAC; a proxy for yield-based reference points, most often maximum sustainable yield) estimation methods, as well as other

---

[1] National Marine Fisheries Service (NMFS), Fisheries Statistics Division, 2017. Personal commun.

assessment analyses, for over 80 data-limited management procedures (MP) (Carruthers and Hordyk 2017). This package has been applied to make TAC recommendations for black sea bass and blueline tilefish along the US Atlantic coast (McNamee et al. 2015; MAFMC 2016b). While there are an abundance of MPs to choose from, and while the possibility of using many different types of data allows considerable flexibility in the number of species and types of data to which DLMTool can be applied, distinguishing the primary drivers of MPs that are applicable to a given data set can become more difficult, leading to less-informed estimations and explanations of TACs.

Sensitivity analyses that investigate the influence of changes to certain parameters on assessment results have become an important part of evaluating uncertainty and robustness of modern stock assessments and prioritizing areas of research that need to be emphasized to improve further use of a given model. DLMTool currently includes a function, Sense(), for a sensitivity analyses of TAC estimated as the median of a standardized relative frequency distribution of simulated TAC values. However, an alternative estimator could be the mode of the TAC standardized relative frequency distribution, or the value with greatest likelihood of being produced by the inputs. Given the often skewed shape of the TAC distribution, the median and mode often do not coincide and may differ in sensitivity to changes in certain parameters, and a sensitivity analysis function for a modal estimator is not currently built into DLMTool.

This chapter will discuss a simulation study in DLMTool that will evaluate trends in modal TAC estimates and uncertainties about those estimates for various MPs based on changes to input life history parameters. In addition, this chapter will introduce a composite MP, based on TAC estimates of multiple MPs. Composite MPs have been used in previous applications of DLMtool (Miller et al. 2015; MAFMC 2016b), with the intent of using a value representative of multiple MPs in data-limited situations where scientists and managers are unable to confidently evaluate and distinguish among TAC estimates from individual MPs. Unlike previous applications of composite values, because the composite estimate is derived from a combined distribution of simulated values across multiple methods, error can be estimated, at least in a relative fashion, using the combined standardized relative frequency distribution.

MATERIALS AND METHODS

Input values largely came from data compiled by the MAFMC SSC, including estimates of total removals from the commercial and recreational fisheries (Table 11) and life history information from studies of blueline tilefish from the US Mid-Atlantic (supplemented with information from the US South Atlantic or expert opinion when deemed necessary and appropriate by the SSC) (Tables 12-13) (Southeast Data Assessment and Review (SEDAR) 2013; MAFMC 2016b; Chapter 2). It should be noted that removals and harvest estimates in this chapter are in kilograms (kg) rather than pounds for unit consistency between model inputs and outputs. To help in distinguishing abbreviations for inputs, these abbreviations will be listed in the text in italics (e.g., the natural mortality input parameter will be shown as *Mort*). Two updates were made to input data to tailor it to the purposes of this study. The first was to estimate natural mortality (*Mort*) using the updated Hoenig equation from Then et al. (2014) (Table 13), rather than the original *Mort* estimation method of averaging of several point estimates from different estimation methods based on differing amounts and types of life history data. This change allowed use of a single, recently updated point estimate of natural mortality based on the strong correlation between *Mort* and maximum age, which has been recommended as preferable to other methods of estimating *Mort* based on life history data available in this study

(Kenchington 2014; Then et al. 2014). The second was to change all life history parameter coefficients of variation (CV) to the default value used in the MAFMC report, 0.2 (MAFMC 2016b). This change allows us to look at changes within a specified CV range across all parameters, eliminating effects of differing CV magnitudes across parameters on our results. In practical use of this program, CVs that have been measured or estimated, even in an ad-hoc manner, could provide a more accurate representation of uncertainty in the data.

Table 11. Time series of annual removals for blueline tilefish landed in the US Mid-Atlantic region. Source: MAFMC 2016b.

Year	Removals (kg)
1999	367
2000	1473
2001	780
2002	491
2003	3198
2004	26,531
2005	25,079
2006	36,070
2007	39,719
2008	29,004
2009	44,194
2010	28,227
2011	37,988
2012	49,504
2013	56,332
2014	148,119
2015	98,609

Table 12. Input proportions of annual catch at age for blueline tilefish landed in the US-Mid-Atlantic. Estimated from an aging study of blueline tilefish sampled primarily from the recreational fishery in Virginia.

Age	2010 (n=266)	2011 (n=614)
2	0.004	0.000
3	0.008	0.008
4	0.023	0.013
5	0.034	0.046
6	0.139	0.081
7	0.320	0.109
8	0.165	0.138
9	0.049	0.091
10	0.049	0.073
11	0.038	0.068
12	0.075	0.104
13	0.034	0.134
14	0.026	0.047
15	0.011	0.039
16	0.004	0.013
17	0.004	0.003
18	0.000	0.005
20	0.008	0.002
24	0.000	0.002
26	0.000	0.002
27	0.004	0.000
30	0.004	0.000
31	0.000	0.002
32	0.000	0.002
34	0.000	0.005
35	0.004	0.008
36	0.000	0.003
40	0.000	0.002

Table 13. DLMTool model inputs and definitions for blueline tilefish from the US Mid-Atlantic.

Input	Value
t: Duration t (years)	5
AvC: Average catch over time t (kg)	78110
Mort: Natural mortality; from Then et al. (2014)	0.167
L50: Length at 50% maturity (cm); halfway between the total lengths (TL) of the smallest mature female and the only immature female from Chapter 2.	32.7
LFC: Length at first capture (cm); smallest observed TL	26.0
LFS: Length at full selection (cm); modal TL	40.0
vbK: Von Bertalanffy K parameter (von Bertalanffy 1938); estimated from non-linear least squares (NLS) regression of TLs at ages	0.098
vbLinf: Von Bertalanffy Linf parameter (von Bertalanffy 1938); estimated from NLS regression of TLs at ages	92.6
vbt0: Von Bertalanffy t0 parameter (von Bertalanffy 1938); estimated from NLS regression of TLs at ages	-0.37
wla: Weight-length parameter *a* ($W=aL^b$; where W=weight in kg and L=total length in cm); estimated from NLS regression of whole weights at TLs	2.22E-06
wlb: Weight-length parameter *b* ($W=aL^b$); estimated from NLS regression of whole weights at TLs	3.39
steep: Steepness of stock-recruitment relationship; from SEDAR 32 (SEDAR 2013)	0.84
MaxAge: Maximum age (years); oldest observed age	40
CV_AvC: Coefficient of variation (CV) of most recent t years of data	0.52
CV_Mort, CV_L50, CV_LFC, CV_LFS, CV_vbK, CV_vbLinf, CV_vbt0, CV_wla, CV_wlb, CV_steep: Default value	0.2

We used the DLMTool's Can() function to identify MPs that were applicable to this stock, given the types of input information. We identified inputs and corresponding CVs that were required for application of the selected MPs, using DLMTool's Needed() function (Table 14). For each input or CV selected, we calculated a set of alternative values ranging from 50% to

Table 14. Required inputs for management procedures (MPs) selected by DLMtool for application to US Mid-Atlantic blueline tilefish data. Abbreviations for inputs can be found in Table 13.

MP	Required Inputs
AvC	*Catch (Cat)*
BK_CC	*Cat, Mort, LFC, CAA, vbK, vbLinf*
CC1	*Cat*
CC4	*Cat*
Fdem_CC	*Cat, Mort, CAA, vbK, vbLinf, vbt0, wla, wlb, MaxAge*
SPMSY	*Cat, L50, vbK, vbLinf, vbt0, steep, MaxAge*
YPR_CC	*Cat, Mort, LFS, CAA, vbK, vbLinf, vbt0, wla, wlb, MaxAge*

150% (in 5% increments) of the original input value (Table 13). This range of alternative values was chosen because beyond these limits, some input values became unreasonable with respect to correlated parameters, causing the TAC estimation process to fail. For example, more than a 50% decrease to the L_∞ parameter of the von Bertalanffy growth function (*vbLinf*) (Equation 1 from Chapter 2) becomes unreasonable with no change to either the κ (*vbK*) or t_0 (*vbt0*) parameters. For each combination of MP and input variation, we used the TAC() function to estimate a standardized relative frequency distribution of TAC outputs. The process of estimating TAC distributions involves repeated random draws of model parameters based on input values and CVs and calculation of TAC under conditions of the randomly drawn parameters and the selected MP. For a more detailed description of TAC estimation methods, please refer to Carruthers et al. (2014) and the DLMTool package reference manual and vignette at: https://cran.r-project.org/web/packages/DLMtool/index.html. Each TAC distribution was estimated by 1,000 of these random draws. For each TAC distribution estimated, up to one input was altered from its original value by a percent change ranging from -50% to 50% (with the 0% change value being the original input value), with all other inputs maintaining their original values. TAC distributions for all input variations were estimated for all MPs, regardless of whether a given input was required for a given MP. This provides several forms of “null” distributions that can be qualitatively compared to those of MPs that require the input being varied. Standardized relative frequency distributions were calculated by dividing frequencies of different TAC outputs by the greatest output frequency observed. Thus, the relative frequency of the mode of these distributions is always 1.

While the magnitudes of values in a standardized relative frequency distribution vary from those of a true probability distribution, the shape of the probability distribution is maintained and depicted through standardized relative frequencies. Thus, the mode of a standardized relative frequency distribution is also the value that was selected most often by a set of input parameters and underlying model (MP). Although standard relative frequency distributions of TAC can be output from DLMtool, to allow flexibility in calculation of TAC estimates and graphing results of multiple runs of the same MP, we estimated the TAC distributions outside of DLMTool using a gaussian kernel density estimation of the TAC results through R’s standard density() function. The resulting plotted distributions are identical to those produced by DLMTool’s plotOFL() function.

Previous applications of DLMTool MPs have used the medians of TAC distributions as point estimates for reference points (McNamee et al. 2015; Miller et al. 2015; MAFMC 2016b). TAC distributions are often right-skewed due to underlying log-normal distributions in some of the parameter estimation methods, thus the median and mode often do not coincide. We used the modes of TAC distributions as point estimates of reference points being estimated by each input-MP combination. TAC distributions are often non-uniform, with results from many repetitions surrounding a modal TAC, showing that certain output TAC values are preferably estimated by input-MP combinations. Using the TAC standardized relative frequency distribution as a proxy for a probability distribution of potential values for the reference point estimated by a given MP, we interpret the mode of that distribution as the most likely estimator of that reference point for a given input-MP combination, analogous to interpretation of results from bootstrapping techniques. The mode was calculated as the point where standardized relative frequency of our previously mentioned TAC distributions (plotted via kernel density estimation) was maximized. This does, admittedly, estimate the mode, which is discrete in nature, from a continuous distribution, making it an estimate derived from an estimated distribution rather than from the TAC results themselves. However, over a large number of repetitions in each distribution (we used 1,000), potential bias from this estimation technique is minimized.

Variability of observed or simulated values of a distribution about an estimate depicts uncertainty of that estimate being representative of that set of values or distribution. Often, this variability is calculated as a standard deviation about a mean, with smaller values being indicative of less uncertainty about the mean as a representative estimator for a set of data. For symmetric, unimodal probability distributions, the mean would be the same as the mode, and a small standard deviation would result in a narrow distribution of other values about the mean/mode. However, for non-symmetric distributions, such as those produced by many of the DLMTool MPs when estimating TAC, the mean and mode differ, resulting in standard deviation values about the mean that are not necessarily informative of variability about the mode. To measure variability about our modal TAC estimator we calculated an adjusted form of the sample standard deviation equation for each alternative input value-MP combination, using the modal TAC estimate rather than a mean estimate:

$$\mathrm{SD_{Mo}}=\sqrt{\frac{\sum_{i=1}^{n}(\mathrm{TAC_i}-\mathrm{TAC_{Mo}})^2}{n-1}},$$

where SD_{Mo} is the standard deviation about the mode, n equals our number of TAC repetitions for each alternative input value-MP combination (n=1,000), TAC_i is the TAC output from the ith repetition of that combination, and TAC_{Mo} is the mode of the TAC relative frequency distribution for that combination.

For each MP, percent differences of modal TAC point estimates resulting from alternative values of each input were plotted against percent modifications of alternative input values from the original input value, giving plots of percent change in TACs versus percent change in alternative input values for each MP. Similarly, percent differences of SD_{Mo}s resulting from alternative values of each input were plotted against percent modifications of alternative input values from the original input value, giving plots of percent change in SD_{Mo} versus percent change in alternative input values for each MP. To determine whether inputs were significant drivers of MPs, linear models were fit to plots of percent change in TAC versus percent change in alternative input values and percent change in SD_{Mo} versus percent change in alternative input values. Since all original input and output values have a percent variation of zero and to reduce potential confounding effects of a non-zero y-intercept estimate, we forced linear models to pass through the origin. To focus our discussion of inputs to the most important drivers of TAC estimates, variation of an input was determined to significantly affect TAC estimation for an MP if the slope of the percent change of TACs versus percent change of alternative input values was significantly different from zero, as determined by the *P*-value (α=0.05) of an analysis of variance (ANOVA), and if either the absolute slope or R^2 value was greater than or equal to 0.5. Correlations with slopes that were significantly different from zero for which the absolute slope and R^2 value were both less than 0.5 were considered to be more attributable to randomness of the model and sampling over a limited range of input values (21 values, ranging from 50% of the original value to 150% of the original value). Similarly, variation of an input was determined to significantly affect uncertainty about TAC estimation for an MP if the slope of the percent change of SD_{Mo}s versus percent change of alternative input values was significantly different from zero, as determined by the *P*-value (α=0.05) of an ANOVA, and if either the absolute slope or R^2 value was greater than or equal to 0.5. Due to the potential for non-linear relationships between input and output values, we recognize that linear models may fit observed data points poorly, resulting in lower R^2 values even when a strong correlation is evident. However, our use of multiple criteria for determining significance allows us to identify effects of inputs with such relationships through a minimum absolute slope threshold of 0.5, while allowing sufficient leeway for detecting important TAC-input relationships that do not exceed this threshold and show a fairly strong linear correlation (by an R^2 exceeding 0.5).

Comparisons of linear models can provide some perspective on the relative impacts of different inputs, which is the intent of this study, but relationships between inputs and outputs can be subject to non-linearity and species differences. Therefore, we do not intend estimated linear parameters to be descriptive of absolute relationships nor to necessarily be applicable generally across species. Rather, through this analysis we intend to provide a method that can allow scientists and managers who make decisions about inputs to easily know TAC and uncertainty implications of increasing or decreasing any single input value for any MP in the DLMTool program.

In previous applications of DLMTool, some have found it useful to combine information from multiple MPs to develop a composite estimate for TAC (Miller et al. 2015; MAFMC 2016b), due to a lack of knowledge on whether any individual MP depicts population dynamics of the data-poor species in question better than others. Previously, this has been done by an averaging of multiple TAC estimates from individual MPs. We included a similar composite TAC estimation method in our analyses. However, rather than averaging TAC estimates from several individual MPs, we estimated a composite distribution by combining all "observed" TAC values across all applicable MPs for each input being varied in our study. This composite distribution was equally weighted among individual MPs and used all 1,000 repetitions from each MP that was included. Thus, distributions for the composite (Comp) MP were based on a larger number of TAC "observations" than those used for individual MPs [n_{Comp}=1000*(number of MPs included in Comp MP), where n_{Comp} is the number of repetitions included in the Comp MP TAC distributions]. We applied similar methods for analyzing the Comp MP as those for individual MPs, estimating a recommended TAC as the mode of the Comp distribution, estimating SD_{Mo} across input variations, and regressing linear models to identify trends between input variations and TAC mode estimates and SD_{Mo}s.

The Can() function identified eight MPs that were applicable to this stock, based on the types of input information. One of these MPs, NFref, presents a moratorium management strategy as a reference for comparison with other models. As the TAC for this MP would not change, regardless of parameter values (TAC would always be 0), this MP was not included in our analyses. The remaining seven MPs were: average historical catch (AvC), Beddington and Kirkwood's life-history method combined with catch curve analysis (BK_CC) (Beddington and Kirkwood 2005), two constant catch management procedures (CC1 and CC4; for both methods the time duration over which catches were averaged was 5 years) (Geromont and Butterworth 2014), a demographic F_{MSY} (instantaneous fishing mortality rate at maximum sustainable yield) method using catch-curve analysis to estimate recent instantaneous total mortality rate (Fdem_CC) (McAllister et al. 2001), catch trend surplus production MSY estimation (SPMSY)

(Martell and Froese 2013), and yield per recruit analysis to approximate F_{MSY} paired with catch curve estimation of recent instantaneous mortality rate (YPR_CC) (Carruthers and Hordyk 2017). Required inputs for each of these MPs are shown in Table 14. To prevent overweighting of catch-only methods (AvC, CC1, and CC4) in the Comp MP, this MP only combined simulated TACs produced by the CC1, BK_CC, Fdem_CC, SPMSY, and YPR_CC MPs. Simulated TACs from these five MPs were combined to construct TAC distributions for the Comp MP. CC1 was chosen over AvC for inclusion in the Comp MP because this MP averages only the most recent catches, and due to the novelty of the blueline tilefish fishery off the US Mid-Atlantic, an average over the entire time series would underestimate current harvest levels, potentially biasing TAC low. CC4 averages over the same timespan as CC1, but reduces TAC by 30%. We used CC1 rather than CC4 because none of the other MPs included in our Comp MP included a similar scalar adjustment.

RESULTS

Absolute estimates of TAC using original input values varied among MPs, and MPs generally fell within one of two categories, one with distinctly higher TAC estimates and the other with distinctly lower TAC estimates, except one MP, YPR_CC, that estimated TAC between these categories (Figure 17). Estimates of TAC from the BK_CC and Fdem_CC MPs were approximately four to five times those of AvC, CC1, CC4, SPMSY, or Comp, and estimates from the YPR_CC MP were approximately three times those of AvC, CC1, CC4, SPMSY, or Comp (Figure 17). TAC estimated by AvC was the least of all MPs investigated with an average modal estimate of 35,623 kg. SD_{Mo} estimates of TAC distributions resulting from original input values also varied among MPs. BK_CC, Fdem_CC, YPR_CC, and Comp

Figure 17. Modal estimates (gray dots) of total allowable catch (kg) for US Mid-Atlantic blueline tilefish from 19 base runs of data-limited management procedures (MPs; AvC, BK_CC, CC1, CC4, Fdem_CC, SPMSY, and YPR_CC; descriptions of each MP are found on pages 78-79) in the Data-Limited Methods Toolkit (DLMTool) (Carruthers and Hordyk 2017) and a composite MP (Comp) that includes estimates from the BK_CC, CC1, Fdem_CC, SPMSY, and YPR_CC MPs. Mean values of modal estimates for each MP are marked by black dashes.

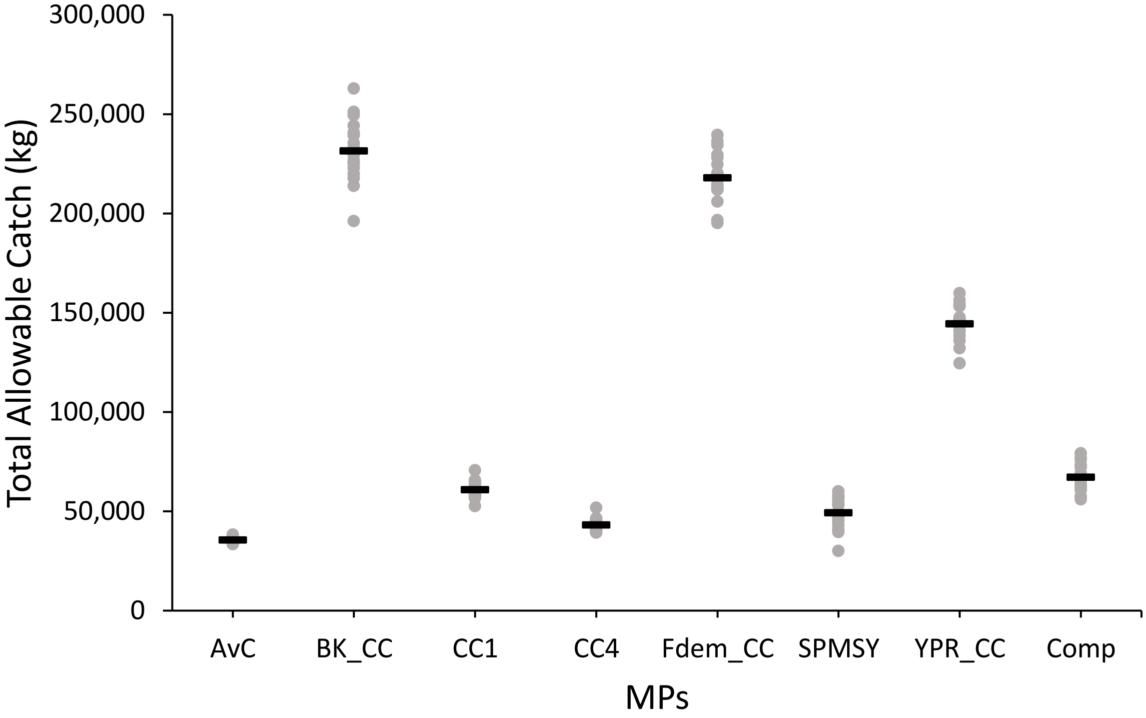
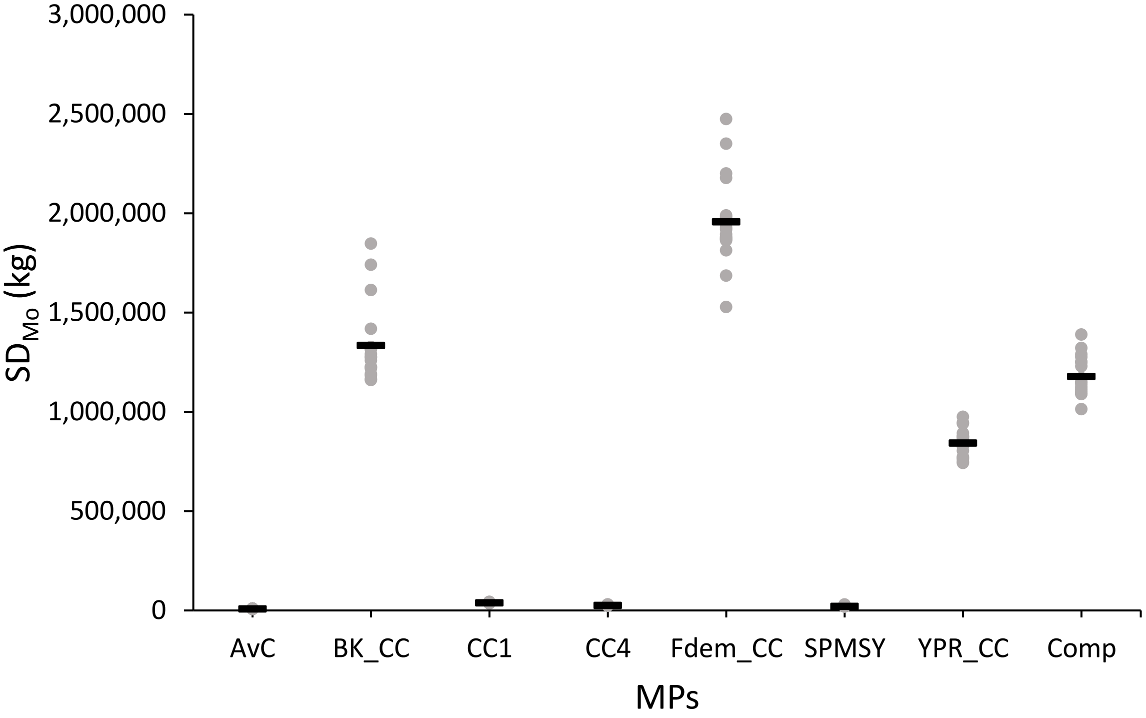


Figure 18. Estimates of standard deviation about the mode (SD_{Mo}; gray dots) of total allowable catch (kg) for US Mid-Atlantic blueline tilefish from 19 base runs of data-limited MPs in the Data-Limited Methods Toolkit (Carruthers and Hordyk 2017) and a composite MP (as defined for Figure 19). Mean values of SD_{Mo} are marked by black dashes.



estimates of SD_{Mo} were approximately 2 orders of magnitude greater than estimates from AvC, CC1, CC4, or SPMSY (Figure 18).

To focus discussion of changes to TAC estimates and variabilities in response to changes in input values, figures are only included for input-MP combinations that resulted in TAC or SD_{Mo} trends with a significant linear slope ($p<0.05$) or R^2 greater than 0.5. Additionally, figures of output distributions are only included for the two inputs that had significant impacts for almost all MPs tested, *Mort* and *vbLinf* (Figures 19 and 20, respectively). Distributions from

Figure 19. Distributions of total allowable catch (TAC) for Mid-Atlantic blueline tilefish, estimated by data-limited management procedures (MPs) from the Data-Limited Methods Toolkit for different values of natural mortality (*Mort*). The MPs shown are those for which the estimate or variability of TAC changed significantly across different *Mort* values: BK_CC, Fdem_CC, and YPR_CC, as well as a composite (Comp) MP that includes the BK_CC, CC1, Fdem_CC, SPMSY, and YPR_CC MPs (descriptions of all MPs except the Comp MP are on pages 78-79). Each MP's graph is scaled to its estimated TAC values to highlight differences in distributions within each MP that result from different input values for *Mort*. Therefore, graphs with very different TAC estimates may not have the same scale.

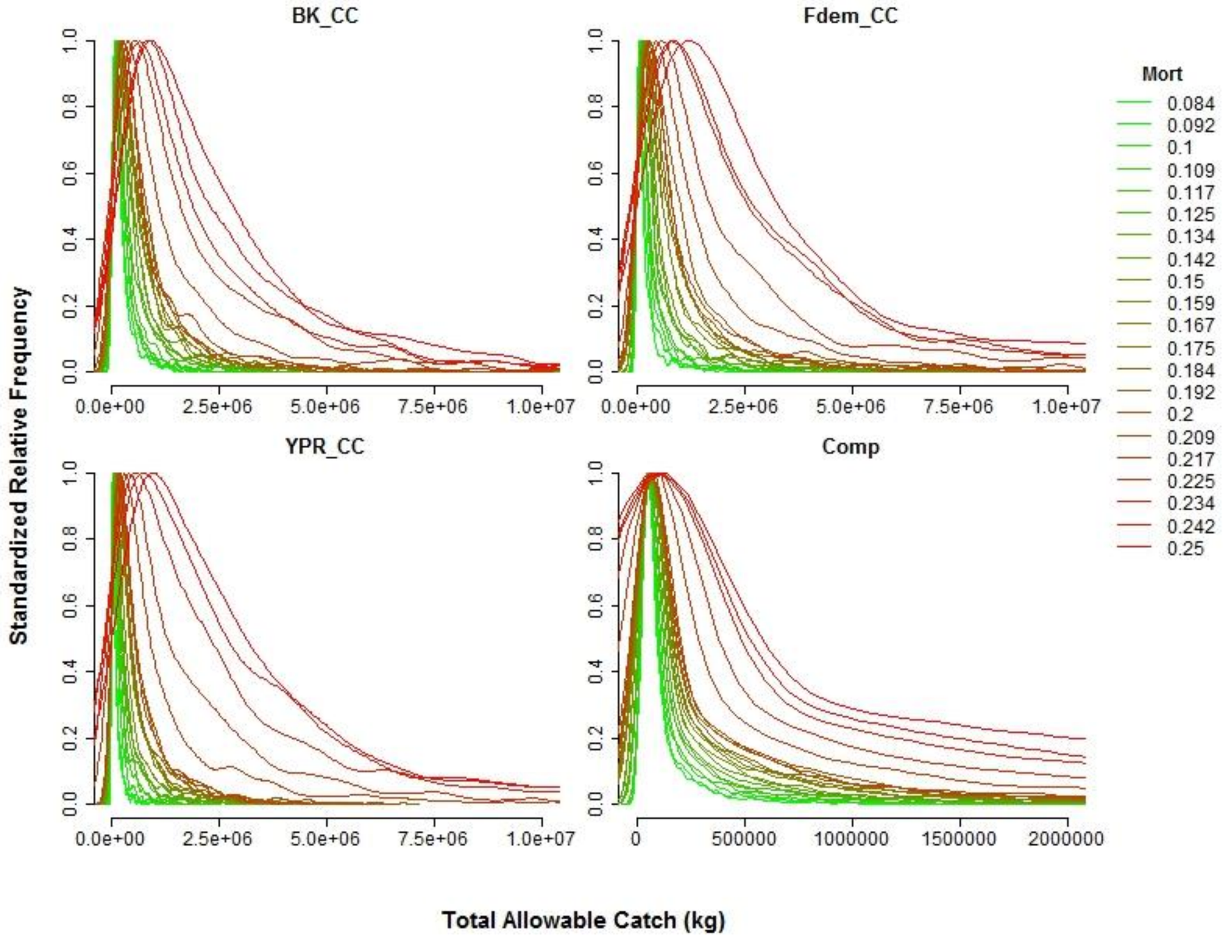
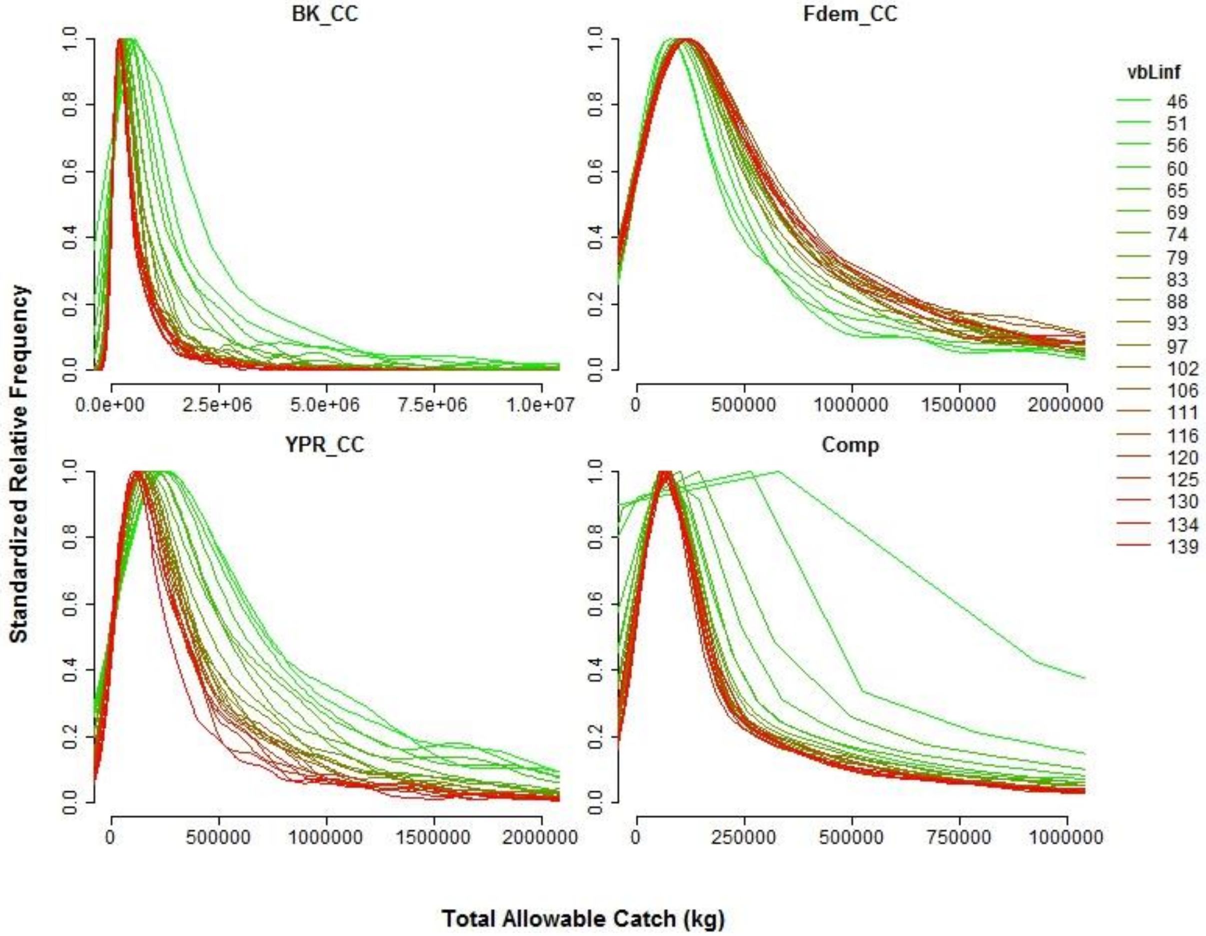


Figure 20. Distributions of total allowable catch (TAC) for Mid-Atlantic blueline tilefish, estimated by data-limited MPs (as defined for Figure 19) for different values of the L_∞ parameter (in cm) of the von Bertalanffy growth equation (von Bertalanffy 1938) (*vbLinf*). The MPs shown are those for which the estimate or variability of TAC changed significantly across different *vbLinf* values. Each MP's graph is scaled to its estimated TAC values to highlight differences in distributions within each MP that result from different input values for *vbLinf*. Therefore, graphs with very different TAC estimates may not have the same scale.



other inputs that had significant impacts are summarized through trend plots of TAC and SD_{Mo} estimates for each MP (Figures 21-22). Distribution plots for inputs other than *Mort* and *vbLinf*, are available upon request to the author.

Percent changes in TAC were more strongly correlated with percent changes of inputs required by each MP, though percent changes in TAC were not significantly correlated with all inputs required by each MP (to focus discussion of results, figures are only included for input-MP combinations with a significant [p<0.05] linear slope or R^2 value greater than 0.5 for percent change relationships between input and output values; contact author for complete results). Variation of inputs did not have a significant impact on TAC estimates for the AvC, CC1, CC4, or SPMSY MPs. For all individual (not composite) MPs that showed significant relationships between TAC and changes to input values, TAC estimates were most sensitive to changes in *Mort* (Figure 19). For the BK_CC MP, percent changes in TAC were were also significantly correlated with changes in (from greater to lesser absolute slope) *vbLinf* (Figure 20), *vbK*, *LFC*, and *CV_Mort* (Table 15, Figure 21). For the Fdem_CC MP, percent changes in TAC were also significantly correlated with changes in *steep*, *CV_Mort*, *vbK*, *wlb*, and *vbLinf* (Figure 20) (Table 15, Figure 21). For the YPR_CC MP, percent changes in TAC were also significantly correlated with changes in *vbLinf* (Figure 20), *CV_Mort*, and *LFS* (Table 15, Figure 21). For the Comp MP, percent changes in TAC were most sensitive to changes in *vbLinf* (Figure 20) and were also significantly correlated with changes in *Mort* (Figure 19) (Table 15, Figure 21).

Similarly, percent changes in SD_{Mo} were more strongly correlated with percent changes of inputs that are required by each MP, though percent changes in SD_{Mo} were not significantly correlated with all inputs required by each MP. Variation of inputs did not have a significant impact on SD_{Mo}s for the AvC, CC1, CC4, or SPMSY MPs. For the BK_CC MP, percent changes

Table 15. Slope estimates, *P*-values of slope estimates, and R^2 values of significant relationships between percent deviations of varied input values and total allowable catch (TAC) outputs from original input values and TAC outputs. Relationships were considered significant if the *P*-value for the estimate of linear slope was less than 0.05 and either the absolute slope estimate or R^2 was greater than 0.5. The Comp MP was a combination of BK_CC, CC1, Fdem_CC, SPMSY, and YPR_CC MPs. Descriptions of all MPs except Comp are on pages 78-79. Abbreviations for inputs and original input values can be found in Table 13.

MP	Varied Input	Slope	*P*-value	R^2
BK_CC	*Mort*	2.896	<0.0001	0.679
	vbLinf	-1.534	<0.0001	0.492
	vbK	1.066	<0.0001	0.961
	LFC	0.874	<0.0001	0.849
	CV_Mort	-0.437	<0.0001	0.746
Comp	*vbLinf*	-2.098	0.019	0.247
	Mort	0.934	<0.0001	0.695
Fdem_CC	*Mort*	3.957	<0.0001	0.63
	Steep	3.72	<0.0001	0.765
	CV_Mort	-0.536	<0.0001	0.884
	vbK	0.509	<0.0001	0.751
	Wlb	-0.45	<0.0001	0.858
	vbLinf	0.343	<0.0001	0.661
YPR_CC	*Mort*	4.53	<0.0001	0.588
	vbLinf	-1.074	<0.0001	0.731
	CV_Mort	-0.752	<0.0001	0.822
	LFS	0.856	<0.0001	0.955

in SD_{Mo} were most sensitive to changes in *vbLinf* (Figure 20), and were also significantly correlated with changes in (from greater to lesser absolute slopes) *Mort* (Figure 19), *LFC*, *vbK*, and *CV_Mort* (Table 16, Figure 22). For the Fdem_CC MP, percent changes in SD_{Mo} were most sensitive to changes in *Mort* (Figure 19), and were also significantly correlated with changes in *steep* and *CV_steep* (Table 16, Figure 22). For the YPR_CC MP, percent changes in SD_{Mo} were

Figure 21. Percent changes in total allowable catch (TAC) resulting from changes to original values (0% change) for input parameters (*LFC*, *LFS*, *Mort*, *steep*, *vbK*, *vbLinf*:, *wlb*, *CV_Mort*; descriptions of inputs are found in Table 13) describing Mid-Atlantic blueline tilefish using the Data-Limited Methods Toolkit's (DLMTool) BK_CC, Fdem_CC, and YPR_CC management procedures (MPs), as well as a composite (Comp) MP that includes the BK_CC, CC1, Fdem_CC, SPMSY, and YPR_CC MPs. Descriptions of all MPs except Comp are on pages 78-79. The inputs shown are those for which the TAC estimated by each MP changed significantly across different input values. Changes to some inputs resulted in markedly increased changes to TAC estimates for some MPs. These are shown in the same colors as described in the legend, but with open square markers instead of closed circles. Closed circle markers should be read according to the left vertical axis, and open square markers should be read according to the right vertical axis.

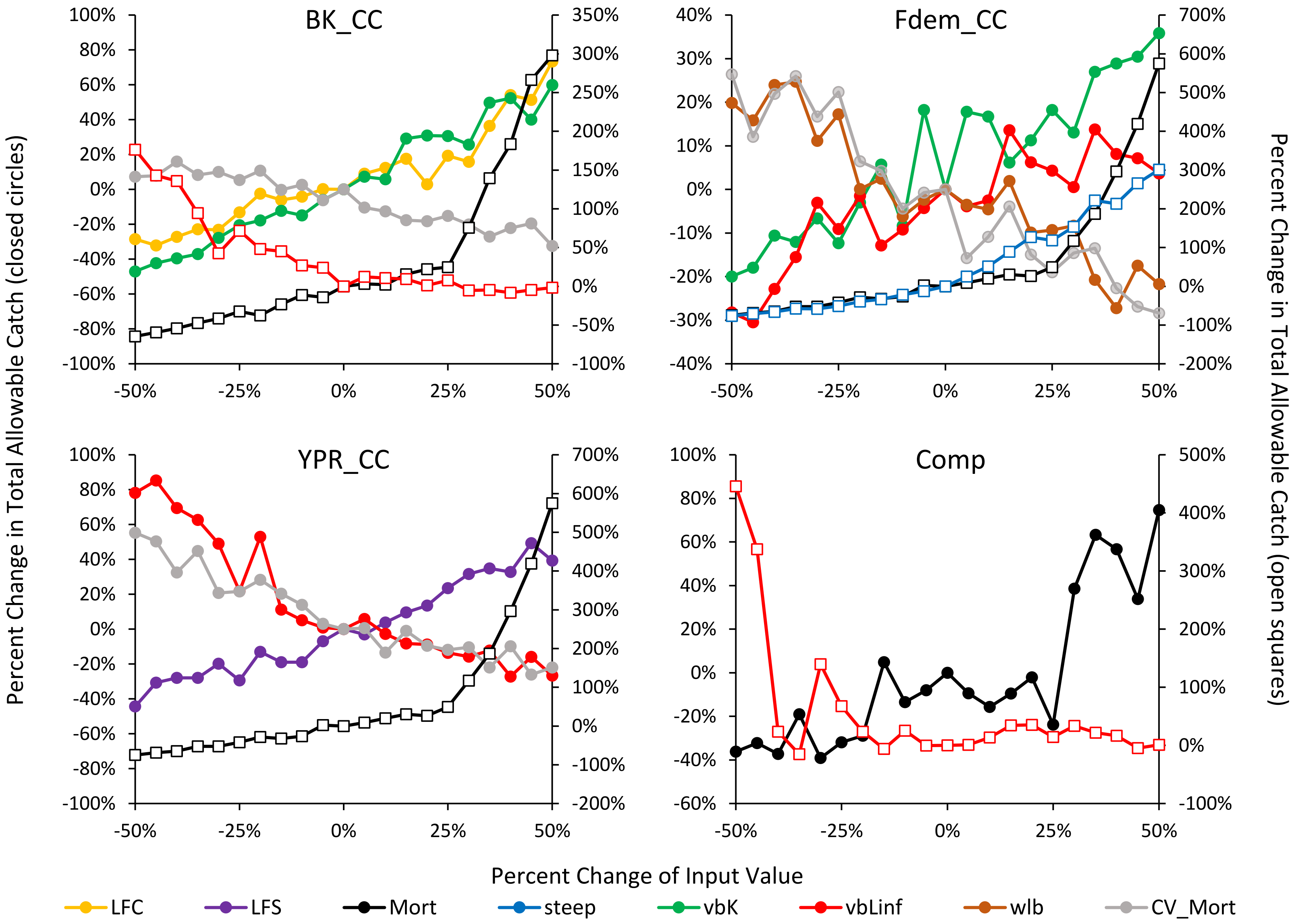

BK_CC
Fdem_CC
YPR_CC
Comp
Percent Change in Total Allowable Catch (closed circles)
Percent Change in Total Allowable Catch (open squares)
Percent Change of Input Value
100%
80%
60%
40%
20%
0%
-20%
-40%
-60%
-80%
-100%
350%
300%
250%
200%
150%
100%
50%
0%
-50%
-100%
40%
30%
20%
10%
0%
-10%
-20%
-30%
-40%
700%
600%
500%
400%
300%
200%
100%
0%
-100%
-200%
500%
-50%
-25%
0%
25%
50%
LFC
LFS
Mort
steep
vbK
vbLinf
wlb
CV_Mort

Figure 22. Percent changes in standard deviation about modal total allowable catch (TAC) resulting from changes to original values (0% change) for input parameters (*LFC*, *LFS*, *Mort*, *steep*, *vbK*, *vbLinf*, *CV_Mort*, and *CV_steep*; descriptions of inputs are found in Table 13) describing Mid-Atlantic blueline tilefish using the Data-Limited Methods Toolkit's (DLMTool) BK_CC, Fdem_CC, and YPR_CC management procedures (MPs), as well as a composite (Comp) MP that includes the BK_CC, CC1, Fdem_CC, SPMSY, and YPR_CC MPs. Descriptions of all MPs except Comp are on pages 78-79. The inputs shown are those for which the SD_{Mo} estimated by each MP changed significantly across different input values. Changes to some inputs resulted in markedly increased changes to SD_{Mo} estimates for some MPs. These are shown in the same colors as described in the legend, but with open square markers instead of closed circles. Closed circle markers should be read according to the left vertical axis, and open square markers should be read according to the right vertical axis.

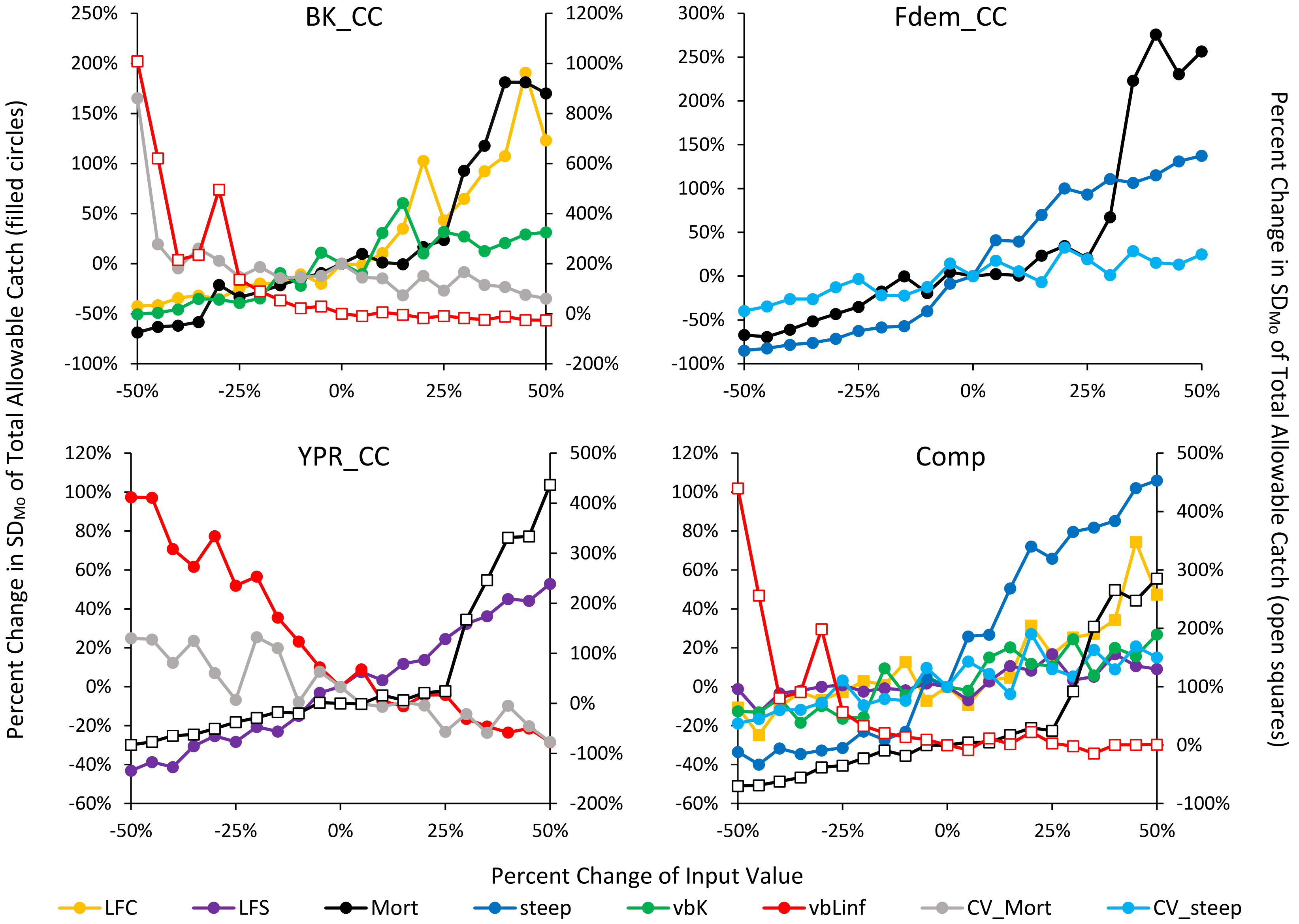

BK_CC
Fdem_CC
YPR_CC
Comp
250%
200%
150%
100%
50%
0%
-50%
-100%
1200%
1000%
800%
600%
400%
200%
0%
-200%
300%
250%
200%
150%
100%
50%
0%
-50%
-100%
120%
100%
80%
60%
40%
20%
0%
-20%
-40%
-60%
500%
400%
300%
200%
100%
0%
-100%
-200%
500%
400%
300%
200%
100%
0%
-100%
-50%
-25%
0%
25%
50%
Percent Change in SD$_{Mo}$ of Total Allowable Catch (filled circles)
Percent Change in SD$_{Mo}$ of Total Allowable Catch (open squares)
Percent Change of Input Value
LFC
LFS
Mort
steep
vbK
vbLinf
CV_Mort
CV_steep

Table 16. Slope estimates, *P*-values of slope estimates, and R^2 values of significant relationships between percent deviations of varied input values and standard deviations of modal total allowable catch outputs (SD_{Mo}) from original input values and SD_{Mo} estimates. Relationships were considered significant if the *P*-value for the estimate of linear slope was less than 0.05 and either the absolute slope estimate or R^2 was greater than 0.5. The Comp MP was a combination of BK_CC, CC1, Fdem_CC, SPMSY, and YPR_CC MPs. Descriptions of all MPs except Comp are on pages 78-79. Abbreviations for inputs and original input values can be found in Table 13.

MP	Varied Input	Slope	*P*-value	R^2
BK_CC	*Mort*	2.3452	<0.0001	0.772
	LFC	1.9245	<0.0001	0.723
	vbK	0.9325	<0.0001	0.763
	CV_Mort	-0.8008	0.0031	0.361
	vbLinf	-6.3052	0.0009	0.429
Comp	*Mort*	3.1803	<0.0001	0.684
	steep	1.6407	<0.0001	0.808
	LFC	0.6318	<0.0001	0.599
	vbK	0.4012	<0.0001	0.715
	CV_steep	0.3509	<0.0001	0.691
	LFS	0.1837	0.0001	0.526
	vbLinf	-2.5091	0.0023	0.38
Fdem_CC	*Mort*	3.0686	<0.0001	0.668
	steep	2.616	<0.0001	0.92
	CV_steep	0.5952	<0.0001	0.73
YPR_CC	*Mort*	4.2031	<0.0001	0.644
	LFS	0.9771	<0.0001	0.981
	CV_Mort	-0.4918	<0.0001	0.76
	vbLinf	-1.2831	<0.0001	0.716

most sensitive to changes in *Mort* (Figure 19), and were also significantly correlated with changes in *vbLinf* (Figure 20), *LFS*, and *CV_Mort* (Table 16, Figure 22). For the Comp MP, percent changes in SD_{Mo} were most sensitive to changes in *Mort* (Figure 19), and were also significantly correlated with changes in *vbLinf* (Figure 20), *steep*, *LFC*, *vbK*, *CV_steep*, and *LFS* (Table 16, Figure 22).

DISCUSSION

Differences in TAC estimation methods among MPs led to differences in modal TAC estimates and standardized relative frequency distributions. One way in which this occurred was through differing levels of stochasticity incorporated by each MP. BK_CC, Fdem_CC, and YPR_CC all include random sampling of multiple parameter values from log-normal distributions based on input parameter values and associated CVs for growth or mortality. On the other hand, the only stochasticity used in catch-only methods (AvC, CC1, and CC4) is resampling of a single value for catch (TAC) from a log-normal distribution based on input catch information. Additionally, SPMSY samples values from uniform distributions that are limited based on “estimation rules” determined by life history inputs (Martell and Froese 2013). Sampling from multiple log-normal distributions of life history inputs resulted in increased SD_{Mo} values for BK_CC, Fdem_CC, and YPR_CC. This increased variability (using SD_{Mo} as our metric) is not necessarily reflective of poorer estimates from these MPs. On the contrary, an increased level of variability should probably be expected of TAC estimates for data-limited stocks, and incorporation of data beyond catch may produce more informed TAC estimates. Therefore, managers should not place a high level of confidence in more “precise” estimation methods that are based on relatively small amounts of information (such as catch-only methods).

Methods that incorporate population characteristics, such as BK_CC, Fdem_CC, and YPR_CC MPs are likely able to provide more accurate and comparable measures of uncertainty than catch-only MPs.

A composite (Comp) MP was developed to compare results and refine a management application of DLMTool in which results from multiple MPs are incorporated into a single estimate of TAC. Previous applications of this strategy incorporated multiple MPs by setting criteria to select MPs that would be appropriate for the fishery based on probabilities of overfishing or biomass thresholds from MSE in DLMTool and averaging the TAC point estimates of these individual MPs (Miller et al. 2015; MAFMC 2016b), following a thought process that a true TAC point estimate would likely be within a range of estimates from different, appropriate estimation methods. No such criteria or MSE was performed in this analysis, but methods of this analysis could easily be applied to a suite of MPs that has been selected by MSE with associated criteria. The approach we present combines the results of individual random draws, which would be used to form distributions describing TAC for individual MPs, to form a single distribution from which a TAC point estimate can be made. This approach allows the variabilities of distributions from multiple MPs to be incorporated into the TAC point estimate and distribution. Additionally, depending on shapes and number of the distributions being combined, the Comp TAC point estimate does not necessarily sit directly in the middle of the point estimates of individual MPs. This allows the Comp TAC estimate to be at values that are most often selected as sustainable across multiple MPs rather than somewhere in between a potentially wide range of harvest levels selected as sustainable by individual MPs.

TAC estimates among MPs, using original inputs, generally fell within two groups: one, made up of estimates from BK_CC and Fdem_CC, with consistently higher estimates by approximately four to five times the magnitude of estimates from the lower group, made up of estimates from AvC, CC1, CC4, SPMSY, and Comp (Figure 17). TAC estimates from the YPR_CC MP were in between estimates from each of these two groups. For estimates from BK_CC, the increased values are likely due to this method estimating TAC at F_{Max}, the fishing mortality rate that maximizes yield per recruit, rather than F_{MSY} (Beddington and Kirkwood 2005). Managers considering BK_CC should be cautious of this distinction as TAC may be set at unsustainable levels by this MP, due to the tendency of F_{Max} to be greater than F_{MSY} (Gabriel and Mace 1999). However, BK_CC may have utility in characterizing some form of combined estimate as an upper bound. Fdem_CC estimates F_{MSY} as one half of the population's maximum per capita population growth rate (*r*). High TACs resulting from Fdem_CC could be due to a combination of factors. First, the *r*-estimation method used by Fdem_CC, the Euler-Lotka equation (Lotka 1907), can be biased high relative to other *r*-estimation methods (McCallister et al. 2001), leading to a potential high bias in an estimate of MSY. Second, *r* as used in the logistic population growth model and estimated by this method assumes a lack of density-dependent effects, thus it assumes a population that has undergone some level of depletion (McCallister et al. 2001). Considering the catch time series for blueline tilefish, particularly the recent increase from very low historic catch levels, this assumption may not be valid for Mid-Atlantic blueline tilefish and may lead to an overestimation of how many fish can be sustainably harvested from this population. Finally, this method assumes a constant catchability, which may not be the case for a developing fishery like Mid-Atlantic blueline tilefish. Catchability was likely low early in the time series, as fishermen likely had to work longer to locate areas where patchily distributed blueline tilefish are found, but grew as fishermen begin to identify, share, and copy locations where others had caught blueline tilefish, resulting in greater harvests for similar or less effort.

Several life history parameters can have significant impacts on TAC estimation by the MPs investigated in this study. The most drastic of these impacts were observed for changes to *Mort* and *vbLinf*, which were significantly correlated with changes in TAC estimates for the BK_CC, Fdem_CC, YPR_CC, and Comp MPs. Percent changes in *Mort* showed a relationship similar to a positive exponential with percent changes in TAC estimates from the BK_CC, Fdem_CC, and YPR_CC MPs. Percent changes in *vbLinf* showed strong, roughly linear relationships with percent changes in TAC estimates from the BK_CC (negative; R^2=0.492), Fdem_CC (positive; R^2=0.661), and YPR_CC (negative; R^2=0.731) MPs. Differences in the trends of TAC estimates for the individual MPs highlight how *vbLinf* is considered by each MP. BK_CC and YPR_CC consider *vbLinf* in context of maximizing the yield by maximizing the relative weight or biomass of each individual fish at the time of capture (Beddington and Kirkwood 2005; Carruthers and Hordyk 2017). In this context, lower values for *vbLinf* mean that the vast majority of fish are being caught at or above the largest size we would expect them to reach, thus TAC is maximized at lower *vbLinf* values. Alternatively, Fdem_CC considers *vbLinf* in the context of reproductive potential (McCallister et al. 2001). Larger fish are often able to produce a disproportionately greater number of eggs than smaller fish (Chapter 2 shows that this is the case for Mid-Atlantic blueline tilefish), thus isolated increases in *vbLinf* (that do not include typical accompanying changes to mortality or other von Bertalanffy parameters) would result in increases to the number of potential offspring, population growth rate, and MSY.

When multiple MPs are combined to form the Comp MP, the resultant relationship between *Mort* and TAC is similar in pattern to the relationships described previously by a positive exponential model; however, the magnitude of change seems to be dampened by the incorporation of other MPs with less significant *Mort*-TAC correlations. Using the Comp MP, the relationship between *vbLinf* and TAC appears to be almost knife-edged, with reductions to *vbLinf* beyond 40% resulting in drastic increases in TAC, but other variations of *vbLinf* showing minimal variation. This is due to a strong, negative relationship under the BK_CC MP, with effectively offsetting weaker relationships for the Fdem_CC and YPR_CC MPs. Considering the shapes of the TAC distributions resulting from changes to *vbLinf* in the Comp MP (Figure 20), the greatest percent reductions seem to drive a widening of the distribution, potentially implicating *vbLinf* values that are unreasonably low, given the values of other model parameters. This demonstrates the need to consider correlation among inputs if multiple values or ranges are being used to describe life history characteristics of a data-limited stock. A more in-depth study and discussion of impacts on TAC distributions from changes to multiple parameters that are explicitly correlated through models or have been observed to be correlated across a wide variety of species, such as *Mort* and life history parameters often used for its estimation (von Bertalanffy growth parameters and maximum age) or parameters with relationships described by the Beverton-Holt life history invariants (*Mort*, age and length at maturity, *vbK*, and *vbLinf*) (Charnov 1993), could be a more practical exercise than the single-parameter impacts shown here.

The only CV parameter that had a significant impact on TAC estimates was that of natural mortality, *CV_Mort*. *CV_Mort* showed negative linear relationships with TAC for the BK_CC, Fdem_CC, and YPR_CC MPs. This shows the strong reliance of these MPs on the certainty of *Mort* estimates. Less certain *Mort* estimates necessitate reduced TAC levels to prevent overfishing. Often, direct estimates of *Mort* are difficult to obtain, and much more so for data-limited stocks (Vetter 1988). Therefore, in practice, *Mort* estimates are often dependent upon empirically-derived relationships with life history parameters or, if age frequency data are available, catch curves, as used in this example. With these methods of estimation, uncertainty parameters, such as CV, are often estimated in an ad-hoc manner through expert opinion or some combination of multiple *Mort* estimation methods and can be quite large. Due to the important role that uncertainties about *Mort* estimates can have in TAC estimation, research that can provide improved information on *Mort* and uncertainty estimation techniques should be highly prioritized for data-limited stocks.

Uncertainty about modal TAC estimates was estimated by a descriptor of the narrowness of the peaks in the TAC distributions, SD_{Mo}. This descriptor of narrowness is not a true estimate of uncertainty about these estimates with respect to characteristics of the overall population and should not be compared among MPs with great differences in the amount, type, or use of input information. However, it does give a relative measure of variability among similar MPs and can portray impacts on variability of TAC estimates from a single MP when input parameters are altered.

Significant changes to SD_{Mo} were observed when input parameters were altered for the BK_CC, Fdem_CC, YPR_CC, and Comp MPs. Similar to TAC estimates, SD_{Mo} estimates were generally most impacted (determined by absolute slope of percent changes in SD_{Mo} at different input values) by changes to *Mort* and *vbLinf*. Similarities of influential parameters for estimates of both TAC and its variability resulting from alternative input values highlight a tendency of skewed, non-negative TAC distributions to widen (i.e., become more uncertain) as the modal TAC estimate increases. SD_{Mo} estimates for BK_CC were most impacted by changes to *vbLinf*, in a roughly negative exponential fashion, again alluding to unreasonable values with respect to other inputs at the lower end of the *vbLinf* range modeled. SD_{Mo} was most impacted by *Mort* for the Fdem_CC, YPR_CC, and Comp MPs (and was significantly impacted for the BK_CC MP), in a roughly positive exponential fashion, further illustrating the uncertainty of estimating large reference point values for stocks that experience a high level of natural mortality.

Estimates of SD_{Mo} from the Comp MP were significantly impacted by changes to all of the inputs that resulted in significant changes to SD_{Mo} using individual MPs. The correlations between inputs and SD_{Mo}s (slopes of linear models) were dampened due to inclusion of MPs that did not show significant relationships for some of these inputs, but never to the point that the slope was not significantly different from zero or that both R^2 and the absolute slope were both less than 0.5. Thus, as was discussed for the individual MPs, using the Comp MP, SD_{Mo} was most impacted by *Mort* and *vbLinf*.

Alterations of two CV inputs, *CV_steep* and *CV_Mort*, had significant impacts on SD_{Mo}. Using the Fdem_CC and Comp MPs, *CV_steep* had a moderate to low, but strongly linear (R^2=0.73 and 0.69, respectively), positive relationship with SD_{Mo}. Using the BK_CC and YPR_CC MPs, *CV_Mort* showed a negative relationship with SD_{Mo}, in an exponential shape for BK_CC and more modest linear shape for YPR_CC. A lack of significant impacts from many of the CV inputs demonstrates that trends in variability of TAC are likely driven more by input parameter estimates and relationships among parameters than input variability levels of individual parameters.

This study developed and compiled methods for estimation of a TAC point estimate as the mode of simulated TAC distributions estimated by DLMTool, comparisons of TAC distributions among MPs and under alternative input scenarios, and estimation of a composite MP that incorporates effects of multiple MPs into a single distribution with a single point estimate. This information can be used to help managers and scientists applying DLMTool methods to select appropriate MPs based not only on the type of data that they have, but also the levels of uncertainty associated with different data types. For example, aging of deeper water demersal species, like blueline tilefish, using otoliths can have a relatively high uncertainty compared to other fish species, due to the limited seasonal environmental fluctuations in their habitats (White et al. 1998; Wyanski et al. 2000; Filer and Sedberry 2008; Lombardi-Carlson 2012; Chapter 2). Length measurements, however, would likely have a much lower uncertainty. Von Bertalanffy growth parameters *vbK* and *vbLinf* are, respectively, more dependent upon age and length measurements. Therefore, managers and scientists may want to more strongly consider MPs in which *vbLinf* has a greater influence than *vbK*, such as YPR_CC rather than Fdem_CC. Additionally, if multiple MPs are equally plausible, TAC can be estimated based on multiple MPs, using the technique described in this chapter, from a combined distribution, which provides a TAC distribution and point estimate that incorporates variability of several MPs.

CHAPTER 5

CONCLUSIONS

Growth and reproduction of fished populations are two of the most basic, yet most essential, pieces of information to adequately model and manage fisheries. Local adaptations can result in non-uniform patterns of life history characteristics among areas within a fish populations, necessitating research that is representative of these characteristics throughout the population’s range. Two of my objectives were to extend the characterization of blueline tilefish growth and reproduction into the US Mid-Atlantic region, providing data for use in past and future assessments of blueline tilefish that is more representative of the range of the Atlantic stock.

In Chapter 2, I modeled growth of blueline tilefish caught off the coast of Virginia and compared this growth pattern to that observed in the US South Atlantic at different stages throughout the development of the blueline tilefish fishery in that region. Blueline tilefish off Virginia have sexually dimorphic growth, with males growing to larger maximum sizes than females, a pattern that matches previous studies of the US South Atlantic (Ross and Huntsman 1982; Harris et al. 2004). Current growth of blueline tilefish in the US Mid-Atlantic is similar to growth observed in the US South Atlantic during the 1970s, but different than that observed since an increase in commercial fishing for US South Atlantic demersal species during the 1980s and through the present. US Mid-Atlantic blueline tilefish grow more slowly and to greater maximum sizes than blueline tilefish from the US South Atlantic, indicating that: 1) Blueline tilefish in the US Mid-Atlantic through 2012 showed growth similar to that of a relatively unfished population in the US South Atlantic, and 2) Blueline tilefish in the US South Atlantic likely have undergone fishery-induced population juvenescence. This work provides a foundation for the use of management models that incorporate individual growth as well as a reference point against which future estimates of growth for blueline tilefish in the US Mid-Atlantic can be compared.

In Chapter 3, I characterized reproduction of blueline tilefish caught off Virginia through estimates of sex ratio, spawning seasonality, and fecundity to address the objective of determining whether these fish are capable of local reproduction. Sex ratios were male-skewed, but not so much as to inhibit productivity. Sex ratios at length were reflective of previously described sexually dimorphic growth, and sex ratios at size and length were supportive of no hermaphroditism in this population. US Mid-Atlantic blueline tilefish are multiple batch spawners, similar to the reproductive strategy observed in other areas. Female reproductive phase can be identified as spawning or not spawning through macroscopic assessments of gonads with a fair amount of accuracy relative to microscopic assessments, but error of macroscopic assessments increases towards the end of the spawning season. Spawning occurs from May through November, with increasing and decreasing spawning activity at the beginning and end of the season, respectively. Regression, regeneration, and development of new oocytes generally occurs from December through April. The relationship between batch fecundity and total length of US Mid-Atlantic blueline tilefish is similar to that of US South Atlantic blueline tilefish, and was maintained across a larger range of total lengths than previously observed. US Mid-Atlantic blueline tilefish spawn more frequently within a shorter timeframe than US South Atlantic blueline tilefish, resulting in likely similar levels of annual egg production per fish of similar size. US Mid-Atlantic blueline tilefish display all reproductive phases throughout the course of the year for males and females; thus, they are capable of local reproduction. This research is the first to use microscopic and macroscopic phase determinations to characterize spawning seasonality of blueline tilefish throughout the calendar year. Additionally, this research provides information that can inform future models of productivity for blueline tilefish in the US Mid-Atlantic.

The final objective of this dissertation was to explore ways to improve data-limited methods that have previously been used to recommend harvest levels for US Mid-Atlantic blueline tilefish. I addressed this objective in Chapter 4 through the use of an alternative statistic (the mode) as the point estimate of total allowable catch (TAC) frequency distributions, an analysis of sensitivities of data-limited management procedures, and development of a composite management procedure (MP) that incorporates patterns of variability and central tendency from multiple management procedures into a single, combined distribution. The use of the mode rather than the previously used median of TAC distributions will allow harvest recommendations to be made based on the harvest level that is most frequently determined to fit data inputs according to each MP, making input data more informative of recommended harvest levels. Sensitivity analyses of TAC outputs from MPs can inform decisions regarding which MPs are best suited for a particular stock, based on uncertainties of the input data and sensitivities of the MPs. For MPs applicable to US Mid-Atlantic blueline tilefish, TAC estimates were most sensitive to input values for natural mortality and the mean maximum length (L_∞) described by the von Bertalanffy growth equation (Equation 1 from Chapter 2). Finally, the development of a composite MP builds upon previous applications of data-limited methods that averaged TAC point estimates from multiple MPs to make harvest limit recommendations. The composite MP incorporates not just the point estimates of multiple MPs, but the entire distributions to form a combined distribution. This allows the variability of these estimates to be incorporated into the final recommendation, and also gives some relative indication of the uncertainty surrounding that composite point estimate through the shape of the combined distribution.

The work presented in Chapters 2-4 has been used in previous stock assessments and ongoing management discussions for blueline tilefish in this region. This work has laid the foundation in terms of biological information and built upon previous management methods to enhance recently-established management of blueline tilefish in the US Mid-Atlantic. Future management can be enhanced through improved precision of aging methods throughout the Atlantic coast, establishment of surveys that are useful in estimating abundance of blueline tilefish, and continued efforts to evaluate the practical effectiveness of data-limited management methods.

REFERENCES

Able, K. W., C. B. Grimes, R. A. Cooper, and J. R. Uzmann. 1982. Burrow construction and behavior of tilefish, *Lopholatilus chamaeleonticeps*, in Hudson Submarine Canyon. Environmental Biology of Fishes 7:199-205.

Able, K., D. Twichell, C. Grimes, and R. Jones. 1987. Tilefishes of the genus *Caulolatilus* construct burrows in the sea floor. Bulletin of Marine Science 40:1-10.

Almatar, S. M., K. P. Lone, T. S. Abu-Rezq, and A. A. Yousef. 2004. Spawning frequency, fecundity, egg weight and spawning type of silver pomfret, *Pampus argenteus* (Euphrasen) (Stromateidae), in Kuwait waters. Journal of Applied Ichthyology 20:176-188.

Andrews, A. H., E. E. DeMartini, J. Brodziak, R. S. Nichols, R. L. Humphreys, and C. T. Marshall. 2012. A long-lived life history for a tropical, deepwater snapper (*Pristipomoides filamentosus*): bomb radiocarbon and lead–radium dating as extensions of daily increment analyses in otoliths. Canadian Journal of Fisheries and Aquatic Sciences 69:1850-1869.

Baird, T. A., and N. R. Liley. 1989. The evolutionary significance of harem polygyny in the sand tilefish, Malacanthus plumieri: resource or female defence? Animal Behaviour 38(5):817-829.

Ballenger, J.C., A.R. Lytton, and K.R. Spanik. In Preparation. Validating Blueline Tilefish ages in the U.S. south Atlantic using bomb radiocarbon ($F^{14}C$). Journal of Fish Biology.

Beckman, D. W., A. L. Stanley, J. H. Render, and C. A. Wilson. 1990. Age and growth of black drum in Louisiana waters of the Gulf of Mexico. Transactions of the American Fisheries Society 119:537-544.

Beddington, J. R., and G. P. Kirkwood. 2005. The estimation of potential yield and stock status using life history parameters. Philosophical Transactions of the Royal Society B: Biological Sciences 360:163-170.

Begg, G. A., J. A. Hare, and D. D. Sheehan. 1999. The role of life history parameters as indicators of stock structure. Fisheries Research 43:141-163.

Berkson, J., L. Barbieri, S. Cadrin, S. Cass-Calay, P. Crone, P. Dorn, C. Friess, D. Kobayashi, T. J. Miller, W. S. Patrick, S. Pautzke, S. Ralston, and M. Trianni. 2011. Calculating acceptable biological catch for stocks that have reliable catch data only. National Oceanic and Atmospheric Administration (NOAA) Technical Memorandum NMFS-SEFSC-616. Available: http://safmc.net/wp-content/uploads/2016/06/Appendix20H_ORCS20Reports.pdf. (October 2017).

Black, B. A., G. W. Boehlert, and M. M. Yoklavich. 2005. Using tree-ring crossdating techniques to validate annual growth increments in long-lived fishes. Canadian Journal of Fisheries and Aquatic Sciences 62(10):2277-2284.

Brown-Peterson, N., P. Thomas, and C. R. Arnold. 1988. Reproductive biology of the spotted seatrout, *Cynoscion nebulosus*, in south Texas. Fishery Bulletin 86:373-388.

Brown-Peterson, N.J., D. M. Wyanski, F. Saborido-Rey, B. J. Macewicz, and S. K. Lowerre-Barbieri. 2011. A standardized terminology for describing reproductive development in fishes. Marine and Coastal Fisheries 3:52-70.

Campana, S. E. 1999. Chemistry and composition of fish otoliths: pathways, mechanisms and applications. Marine Ecology Progress Series 188:263-297.

Campana, S. E., and C. M. Jones. 1992. Analysis of otolith microstructure data. Pages 73-100 *in* D. K. Stevenson and S. E. Campana, editors. Otolith microstructure examination and analysis. Canadian Special Publication of Fisheries and Aquatic Sciences 117.

Campana, S. E., and S. R. Thorrold. 2001. Otoliths, increments, and elements: keys to a comprehensive understanding of fish populations?. Canadian Journal of Fisheries and Aquatic Sciences 58(1):30-38.

Carruthers, T. R., and A. Hordyk. 2017. DLMtool: Data-Limited Methods Toolkit. R package version 4.1. Available: https://CRAN.R-project.org/package=DLMtool. (October 2017).

Carruthers, T. R., A. E. Punt, C. J. Walters, A. MacCall, M. K. McAllister, E. J. Dick, and J. Cope. 2014. Evaluating methods for setting catch limits in data-limited fisheries. Fisheries Research 153:48-68.

Cerrato, R. M. 1990. Interpretable statistical tests for growth comparisons using parameters in the von Bertalanffy equation. Canadian Journal of Fisheries and Aquatic Sciences 47:1416-1426.

Charnov, E. L. 1993. Life History Invariants. Oxford University Press, New York, NY.

Conover, D. 1992. Seasonality and the scheduling of life history at different latitudes. Journal of Fish Biology 41:161-178.

Dick, E. J., and A. D. MacCall. 2011. Depletion-based stock reduction analysis: a catch-based method for determining sustainable yields for data-poor fish stocks. Fisheries Research 110:331-341.

Dooley, J. K. 1978. Systematics and biology of the tilefishes (Perciformes: Branchiostegidae and Malacanthidae), with descriptions of two new species. NOAA Technical Memorandum NMFS-CIRC-411. Available: http://spo.nmfs.noaa.gov/sites/default/files/legacy-pdfs/CIRC411.pdf. (October 2017).

Enberg, K., C. Jørgensen, E. S. Dunlop, Ø. Varpe, D. S. Boukal, L. Baulier, S. Eliassen, and M. Heino. 2012. Fishing-induced evolution of growth: concepts, mechanisms and the empirical evidence. Marine Ecology 33:1-25.

Filer, K., and G. Sedberry. 2008. Age, growth and reproduction of the barrelfish *Hyperoglyphe perciformis* (Mitchill) in the western North Atlantic. Journal of Fish Biology 72:861-882.

Fitzhugh, G. R., B. A. Thompson, and T. G. Snider. 1993. Ovarian development, fecundity, and spawning frequency of black drum *Pogonias cromis* in Louisiana. Fishery Bulletin 91:244-253.

Fournier, D. A., H. J. Skaug, J. Ancheta, J. Ianelli, A. Magnusson, M. N. Maunder, A. Nielsen, and J. Sibert. 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. Optimization Methods and Software 27:233-249.

Gabriel, W. L., and P. M. Mace. 1999. A review of biological reference points in the context of the precautionary approach. Proceedings of the Fifth National NMFS Stock Assessment Workshop: Providing Scientific Advice to Implement the Precautionary Approach Under the Magnuson-Stevens Fishery Conservation and Management Act. NOAA Technical Memorandum NMFS-F/SPO-40. Available: https://www.st.nmfs.noaa.gov/Assets/stock/documents/workshops/nsaw_5/gabriel_.pdf. (October 2017).

Ganias, K. 2009. Linking sardine spawning dynamics to environmental variability. Estuarine, Coastal and Shelf Science 84:402-408.

Ganias, K., S. K. Lowerre-Barbieri, and W. Cooper. 2015. Understanding the determinate-indeterminate fecundity dichotomy in fish populations using a temperature dependent oocyte growth model. Journal of Sea Research 96:1-10.

Ganias K., H. Murua, G. Claramunt, R. Dominguez-Petit, P. Gonçalves, F. Juanes, J. Keneddy, N. Klibansky, M. Korta, Y. Kurita, S. Lowerre-Barbieri, G. Macchi, M. Matsuyama, A. Medina, C. Nunes, G. Plaza, R. Rideout, S. Somarakis, A. Thorsen, A. Uriarte, and M. Yoneda. 2014. Chapter 4: Egg production. 109 pages *in* R. Domínguez-Petit, H. Murua, F. Saborido-Rey and E. Trippel, editors. Handbook of applied fisheries reproductive biology for stock assessment and management. Digital Consejo Superior de Investigaciones Cientificas (CSIC), Vigo, Spain. Available: http://hdl.handle.net/10261/87768. (October 2017).

Geromont, H. F., and D. S. Butterworth. 2014. Generic management procedures for data-poor fisheries: forecasting with few data. International Council for the Exploration of the Sea (ICES) Journal of Marine Science 72:251-261.

Grimes, C., S. Turner, and K. Able. 1983. A technique for tagging deep-water fish. Fishery Bulletin 81:663-666.

Harris, P. J., D. M. Wyanski, and P. T. Powers Mikell. 2004. Age, growth, and reproductive biology of blueline tilefish along the southeastern coast of the United States, 1982–1999. Transactions of the American Fisheries Society 133:1190-1204.

Heino, M., L. Baulier, D. S. Boukal, B. Ernande, F. D. Johnston, F. M. Mollet, H. Pardoe, N. O. Therkildsen, S. Uusi-Heikkilä, A. Vainikka, R. Arlinghaus, D. J. Dankel, E. S. Dunlop, A. M. Eikeset, K. Enberg, G. H. Engelhard, C. Jørgensen, A. T. Laugen, S. Matsumura, S. Nusslé, D. Urbach, R. Whitlock, A. D. Rijnsdorp, and U. Dieckmann. 2013. Can fisheries-induced evolution shift reference points for fisheries management? ICES Journal of Marine Science 70:707-721.

Hunter, J. R., and B. J. Macewicz. 1985. Measurement of spawning frequency in multiple spawning fishes. NOAA Technical Report NMFS 36. Available: https://swfsc.noaa.gov/publications/CR/1985/8544.PDF. (October 2017).

Hyndes, G., N. Loneragan, and I. Potter. 1992. Influence of sectioning otoliths on marginal increment trends and age and growth estimates for the flathead *Platycephalus speculator*. Fishery Bulletin 90:276-284.

Ihssen, P., H. Booke, J. Casselman, J. McGlade, N. Payne, and F. Utter. 1981. Stock identification: materials and methods. Canadian Journal of Fisheries and Aquatic Sciences 38:1838-1855.

Katz, S., C. Grimes, and K. Able. 1983. Delineation of tilefish, *Lopholatilus chamaeleonticeps*, stocks along the United States east coast and in the Gulf of Mexico. Fishery Bulletin 81:41-50.

Kenchington, T. J. 2014. Natural mortality estimators for information-limited fisheries. Fish and Fisheries 15:533-562.

Kimura, D. K. 1980. Likelihood methods for the von Bertalanffy growth curve. Fishery Bulletin 77:765-776.

Klibansky, N. 2017. Estimating dispersal of Blueline Tilefish (Caulolatilus microps) eggs and larvae from drifter data. SEDAR50-DW23. SEDAR (SouthEast Data Assessment and Review), North Charleston, SC. Available: http://sedarweb.org/docs/wpapers/S50_DW23_Klibansky_Drifter_1.25.2017.pdf. (October 2017).

Law, R. 2000. Fishing, selection, and phenotypic evolution. ICES Journal of Marine Science 57:659-668.

Lombardi-Carlson, L. A. 2012. Life history, population dynamics, and fishery management of the golden tilefish, *Lopholatilus chamaeleonticeps*, from the southeast Atlantic and Gulf of Mexico. University of Florida, Gainesville, FL.

Lotka, A. J. 1907. Studies on the mode of growth of material aggregates. American Journal of Science 141:199-216.

Lowerre-Barbieri, S. K., N. J. Brown-Peterson, H. Murua, J. Tomkiewicz, D. M. Wyanski, and F. Saborido-Rey. 2011. Emerging issues and methodological advances in fisheries reproductive biology. Marine and Coastal Fisheries. 3:32-51.

Lowerre-Barbieri, S. K., M. E. Chittenden, and L. R. Barbieri. 1996b. Variable spawning activity and annual fecundity of weakfish in Chesapeake Bay. Transactions of the American Fisheries Society 125:532-545.

MacCall, A. D. 2009. Depletion-corrected average catch: a simple formula for estimating sustainable yields in data-poor situations. ICES Journal of Marine Science 66:2267-2271.

Macchi, G. J., and E. M. Acha. 2000. Spawning frequency and batch fecundity of Brazilian menhaden, *Brevoortia aurea*, in the Río de la Plata estuary off Argentina and Uruguay. Fishery Bulletin 98:283-289.

Macchi, G. J., M. Pájaro, and M. Ehrlich. 2004. Seasonal egg production pattern of the Patagonian stock of Argentine hake (*Merluccius hubbsi*). Fisheries Research 67:25-38.

Manooch III, C. S., and D. L. Mason. 1987. Age and growth of the warsaw grouper and black grouper from the southeast region of the United States. Northeast Gulf Science 9:65-75.

Martell, S., and R. Froese. 2013. A simple method for estimating MSY from catch and resilience. Fish and Fisheries 14(4):504-514.

Maunder, M. N., P. R. Crone, A. E. Punt, J. L. Valero, and B. X. Semmens. 2016. Growth: theory, estimation, and application in fishery stock assessment models. Fisheries Research 180:1-3.

McAllister, M. K., E. K. Pikitch, and E. A. Babcock. 2001. Using demographic methods to construct Bayesian priors for the intrinsic rate of increase in the Schaefer model and implications for stock rebuilding. Canadian Journal of Fisheries and Aquatic Sciences 58(9):1871-1890.

McBride, R. S. 2014. The continuing role of life history parameters to identify stock structure. Pages 77-107 *in* S. X. Cadrin, L. A. Kerr, and S. Mariani, editors. Stock Identification Methods, 2nd ed. Academic Press, San Diego, CA.

McDowell, J. 2016. Summary of the results of a genetic-based investigation of blueline tile fish (*Caulolatilus microps*). SEDAR50-DW05. SEDAR, North Charleston, SC. Available: http://sedarweb.org/docs/wpapers/S50_DW05_McDowell_BLTgenetics_8.11.2016.pdf. (October 2017).

McGarvey, R., and A. J. Fowler. 2002. Seasonal growth of King George whiting (*Sillaginodes punctata*) estimated from length-at-age samples of the legal-size harvest. Fishery Bulletin 100:545-558.

McNamee, J., G. Fay, and S. X. Cadrin. 2015. Data Limited Techniques for Tier 4 Stocks: An Alternative Approach to Setting Harvest Control Rules Using Closed Loop Simulations for Management Strategy Evaluations. Available from https://static1.squarespace.com/static/511cdc7fe4b00307a2628ac6/t/55a661a5e4b060ebc9d03cf0/1436967333432/DLanalysis_bsb_final.pdf. (October 2017).

Mid-Atlantic Fishery Management Council (MAFMC). 2016a. Blueline Tilefish Catch Series. Available: https://static1.squarespace.com/static/511cdc7fe4b00307a2628ac6/t/56e86c5c4d088ec570afc5b2/1458072668398/BLT+Catch+MemoRev.pdf. (October 2017).

MAFMC. 2016b. Blueline Tilefish Subcommittee Report. Available from http://static1.squarespace.com/static/511cdc7fe4b00307a2628ac6/t/56fea92d27d4bd220b60b50a/1459530031307/BLT+Subcommittee+Report+20160322.pdf. (October 2017).

Militelli, M. I., and G. J. Macchi. 2001. Spawning frequency and batch fecundity of Brazilian flathead, *Percophis brasiliensis* in coastal waters off the Buenos Aires Province. Scientia Marina 65(2):169-172.

Miller, T. J., O. P. Jensen, J. Wiedenmann, and K. Drew. 2015. Review of the McNamee et al. "Data Limited Techniques for Tier 4 Stocks: An Alternative Approach to Setting Harvest Control Rules Using Closed Loop Simulations for Management Strategy Evaluations". Available: https://static1.squarespace.com/static/511cdc7fe4b00307a2628ac6/t/55f823a9e4b05c1dc7afece6/1442325417102/Review+of+McNamee+et+al+20150914.pdf. (October 2017).

Murua, H., L. Ibaibarriaga, P.´ Alvarez, M. Santos, M. Korta, M. Santurtun, and L. Motos. 2010. The daily egg production method: a valid tool for application to European hake in the Bay of Biscay? Fisheries Research 104:100-110.

Murua, H., G. Kraus, F. Saborido-Rey, A. Thorsen, P. Witthames, and S. Junquera. 2003. Procedures to estimate fecundity of wild collected marine fish in relation to fish reproductive strategy. Journal of Northwest Atlantic Fishery Science 33:33-54.

Murua, H., P. Lucio, M. Santurtún, and L. Motos. 2006. Seasonal variation in egg production and batch fecundity of European hake *Merluccius merluccius* (L.) in the Bay of Biscay. Journal of Fish Biology 69:1304-1316.

Murua, H., and L. Motos. 2006. Reproductive strategy and spawning activity of the European hake *Merluccius merluccius* (L.) in the Bay of Biscay. Journal of Fish Biology 69:1288-1303.

Nieland, D. L., R. G. Thomas, and C. A. Wilson. 2002. Age, growth, and reproduction of spotted seatrout in Barataria Bay, Louisiana. Transactions of the American Fisheries Society 131(2):245-259.

Northeast Fisheries Science Center (NEFSC). 2012. 53rd Northeast Regional Stock Assessment Workshop (53rd SAW) Assessment Report. US Department of Commerce NEFSC Reference Document 12-05. Available: https://www.nefsc.noaa.gov/publications/crd/crd1205/. (October 2017).

NEFSC. 2014. 58th Northeast Regional Stock Assessment Workshop (58th SAW) Assessment Report. US Department of Commerce NEFSC Reference Document 14-04. Available: https://www.nefsc.noaa.gov/publications/crd/crd1404/. (October 2017).

O'Donnell, T., and T. Darden. 2016. Preliminary Genetic Population Structure of Blueline Tilefish *Caulolatilus microps* along the East Coast of the United States. SEDAR50-DW06. SEDAR, North Charleston, SC. Available: http://sedarweb.org/docs/wpapers/S50_DW06_O%27Donnel%26Darden_BLTGenetics_7.12.2016_updated_0.pdf. (October 2017).

Pájaro, M., G. J. Macchi, E. Leonarduzzi, and J. Hansen. 2009. Spawning biomass of Argentine anchovy (*Engraulis anchoita*) from 1996 to 2004 using the Daily Egg Production Method. Journal of the Marine Biological Association of the United Kingdom 89(4):829-837.

Parker, R. O., Jr., and R. W. Mays. 1998. Southeastern US deepwater reef fish assemblages, habitat characteristics, catches, and life history summaries. NOAA Technical Report NMFS 138. Available: http://spo.nmfs.noaa.gov/sites/default/files/legacy-pdfs/tr138.pdf. (October 2017).

Parker, R. O., Jr., and S. W. Ross. 1986. Observing reef fishes from submersibles off North Carolina. Northeast Gulf Science 8:31-49.

Picquelle, S., and G. Stauffer. 1985. Parameter estimation for an egg production method of anchovy biomass assessment. Pages 7-16 *in* R. Lasker, editor. An Egg Production Method for Estimating Spawning Biomass of Pelagic Fish: Application to the Northern Anchovy, *Engraulis mordax*. NOAA Technical Report NMFS 36. Available: https://pdfs.semanticscholar.org/96e2/6a48c421a08ed67477473e2a057cf80f1065.pdf. (October 2017).

Quinn, T. J., and R. B. Deriso. 1999. Quantitative Fish Dynamics. Oxford University Press, New York, NY.

Quintero-Hunter, I., H. Grier, and M. Muscato. 1991. Enhancement of histological detail using metanil yellow as counterstain in periodic acid Schiff's hematoxylin staining of glycol methacrylate tissue sections. Biotechnic & Histochemistry 66:169-172.

Rodrigues, K.A., Macchi, G.J., Acha, E.M., Militelli, M.I. 2008. Reproduction of Jenyns´s sprat, *Ramnogaster arcuata*, a winter spawner in the temperate waters of the Río de la Plata estuary (Argentina – Uruguay). Journal of the Marine Biological Association of the United Kingdom 88 (2):423-428.

Ross, J. L., G. R. Huntsman. 1982. Age, growth, and mortality of blueline tilefish from North Carolina and South Carolina. Transactions of the American Fisheries Society 111:585-592.

Ross, J. L., V. Merriner. 1983. Reproductive biology of the blueline tilefish, *Caulolatilus microps*, off North Carolina and South Carolina. Fishery Bulletin 81:553-568.

Ruzecki, E. P., J. J. Norcross, and F. J. Wojcik. 1972. Oceanographic data collected in the Chesapeake Bight of the Virginian Sea from 1966 through 1969. Data Report (Virginia Institute of Marine Science) no. 8. Virginia Institute of Marine Science, College of William and Mary.

Schueller, A. M., E. H. Williams, and R. T. Cheshire. 2014. A proposed, tested, and applied adjustment to account for bias in growth parameter estimates due to selectivity. Fisheries Research 158:26-39.

Smith, D., A. Punt, N. Dowling, A. Smith, G. Tuck, and I. Knuckey. 2009. Reconciling approaches to the assessment and management of data-poor species and fisheries with Australia's harvest strategy policy. Marine and Coastal Fisheries 1:244-254.

Southeast Data Assessment and Review (SEDAR). 2011a. SEDAR 25 - South Atlantic Black Sea Bass Assessment Report. SEDAR, North Charleston, SC. Available: http://sedarweb.org/sedar-25. (October 2017).

SEDAR. 2011b. SEDAR 25 - South Atlantic Tilefish Assessment Report. SEDAR, North Charleston, SC. Available: http://sedarweb.org/sedar-25. (October 2017).

SEDAR. 2013. SEDAR 32 - South Atlantic Blueline Tilefish Assessment Report. SEDAR, North Charleston, SC. Available: http://sedarweb.org/sedar-32. (October 2017).

SEDAR. 2017. SEDAR 50 - South Atlantic Blueline Tilefish Assessment Report. SEDAR, North Charleston, SC. Available: http://sedarweb.org/sedar-50. (October 2017).

Stefánsson, U., L. P. Atkinson, and D. F. Bumpus. 1971. Hydrographic properties and circulation of the North Carolina shelf and slope waters. Deep Sea Research and Oceanographic Abstracts 18:383-420.

Stratoudakis, Y., M. Bernal, K. Ganias, and A. Uriarte. 2006. The daily egg production method: recent advances, current applications and future challenges. Fish and Fisheries 7:35–57.

Then, A., J. Hoenig, N. Hall, and D. Hewitt. 2014. Evaluating the predictive performance of empirical estimators of natural mortality rate using information on over 200 fish species. ICES Journal of Marine Science 72(1):82-92.

von Bertalanffy, L. 1938. A quantitative theory of organic growth (inquiries on growth laws. II). Human Biology 10:181-213.

Vetter, E. F. 1988. Estimation of natural mortality in fish stocks: a review. Fishery Bulletin 86(1):25-43.

White, D., D. Wyanski, and G. Sedberry. 1998. Age, growth, and reproductive biology of the blackbelly rosefish from the Carolinas, USA. Journal of Fish Biology 53:1274-1291.

Wilson, C. A., and D. L. Nieland. 1994. Reproductive biology of red drum, *Sciaenops ocellatus*, from the neritic waters of the northern Gulf of Mexico. Fishery Bulletin 92:841-850.

Wyanski, D. M., D. B. White, and C. A. Barans. 2000. Growth, population age structure, and aspects of the reproductive biology of snowy grouper, *Epinephelus niveatus*, off North Carolina and South Carolina. Fishery Bulletin 98:199-218.

Zar, J. 1996. Biostatistical Analysis, 3 ed. Prentice Hall, Upper Saddle River, NJ.

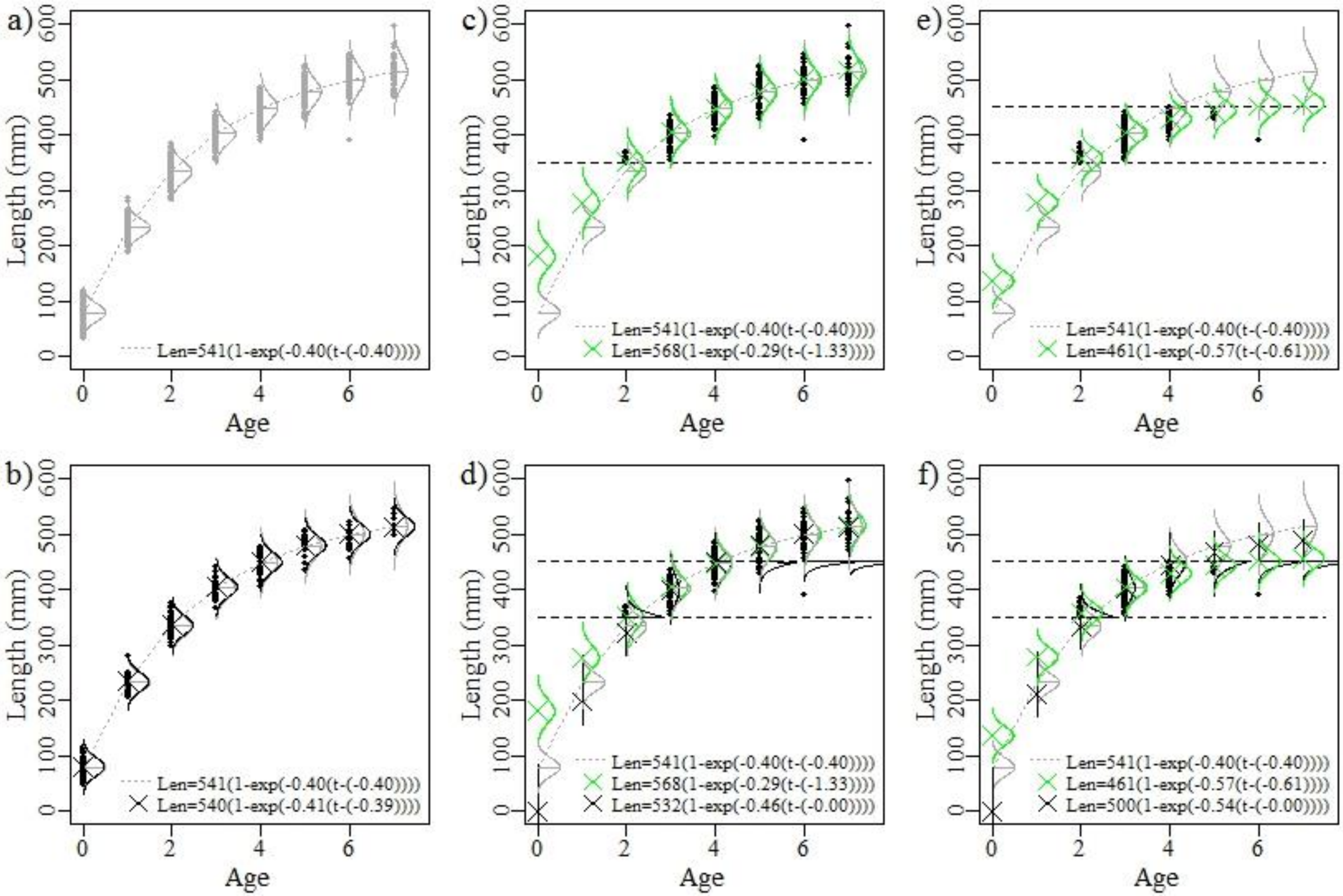
APPENDIX

Illustration of Bias Adjustment to von Bertalanffy Growth Model via Maximum Likelihood Estimation with Truncated Normal Errors

Consider a population of individuals with the lengths at age shown in Appendix Figure A1a. Solid bell-shaped lines show normal probability density functions (pdf) describing lengths at each age (with $Pr(y)=0$ at $x=t$ for ages $t=[0:7]$). All pdfs and error distributions are multiplied by 20 for illustrative purposes, unless otherwise stated. Solid horizontal lines highlight the maximum likelihood estimate (or, for the normal distribution, the mean) of lengths at each age. We attempt to model these underlying distributions using growth equations, such as the Von Bertalanffy (VB) growth equation. The VB equation that best estimates the population lengths at age is shown as a dashed line, with its equation listed in the legend.

Capture of all individuals within a population is often not feasible, so we attempt to model growth of individuals within a population by estimating growth for a sample from that population. In an optimal situation, a sample is of appropriate size and randomly selected from the entire population. For this exercise, we assume appropriate sample size throughout. In Appendix Figure A1b, length at age has been estimated for a randomly selected sample of our population, using maximum likelihood estimation of the VB equation with assumed normally distributed errors (equivalent to least squares regression). Maximum likelihood estimates from this model of mean lengths at age are shown as x's, with error distributions estimating the population pdfs of lengths at ages (using residual standard error to estimate standard deviations) shown in black. These model estimates portray the underlying growth pattern of our population quite nicely, evidenced by the great amount of overlap between black and grey distributions and close proximity of black x's to the horizontal grey lines.

Appendix Figure A1. Simulated population (a) and samples (b-f) of lengths at age. For the population (shown in grey for all graphs), points denote individual lengths at age (a; N=1600), bell-shaped lines denote normal probability density functions (pdf) of lengths at age, horizontal lines denote mean lengths at age, and dashed curves denote the Von Bertalanffy (VB) model fit using maximum likelihood estimation with normally distributed errors (normal MLE). Sampled lengths at age (all samples, n=320) are shown as black points (b-f). b) Sample was chosen randomly. VB model fit via normal MLE estimates means (black x's) and pdfs (black lines; using residual standard error to estimate standard deviations) of lengths at age. c-d) Sample was chosen randomly with a minimum length limit of 350 mm (dashed black line). VB model fit via normal MLE estimates means (green x's) and pdfs (green lines) of lengths at age. VB fit via maximum likelihood estimation with truncated normally distributed errors (truncated MLE) estimates means (black x's) and pdfs (solid black lines) of lengths at age. e-f) Sample was chosen randomly with a minimum length limit of 350 mm and a maximum length limit of 450 mm (dashed black lines). VB model fit via normal MLE estimates means (green x's) and pdfs (green lines) of lengths at age. VB fit via truncated MLE estimates means (black x's) and pdfs (solid black lines) of lengths at age.


a)
b)
c)
d)
e)
f)
Length (mm)
Age
0
100
200
300
400
500
600
0
2
4
6
Len=541(1-exp(-0.40(t-(-0.40))))
Len=540(1-exp(-0.41(t-(-0.39))))
Len=568(1-exp(-0.29(t-(-1.33))))
Len=532(1-exp(-0.46(t-(-0.00))))
Len=461(1-exp(-0.57(t-(-0.61))))
Len=500(1-exp(-0.54(t-(-0.00))))

Fishery scientists often collect data from fish captured by a fishery. These fish are typically not selected randomly and are thus representative of the subset of the population that is susceptible to fishing, rather than the entire population. This non-random selection can bias growth models if previous assumptions of random selection are not altered. The sample of our original population shown in Appendix Figures A1c and A1d has been selected from a fishery in which fish smaller than 350 mm are not captured. The VB model estimated with normally distributed errors for these data (green) estimates mean lengths at age well for mostly or fully selected ages (3-7). However, length estimates for partially or non-selected ages (0-2) are increasingly biased as length estimates move away from 350 mm and VB parameter estimates differ substantially from those describing our population.

The methods of McGarvey and Fowler (2002) and Schueller et al. (2014) reduce the bias caused by fishing selectivity by changing the assumption of normally distributed errors to one of errors having a truncated normal distribution. This reduces the probability of observing fish at lengths not selected by the fishery to zero and reapportions that probability to the rest of the distribution. This restructuring of the assumed error distribution (shown in Appendix Figure A1d in black; truncated distributions are multiplied by 10 for illustrative purposes) results in VB estimates of mean lengths at age (black x's) and parameter estimates that better approximate those describing our population.

The bias created by selectivity can be exaggerated even further in fisheries where minimum and maximum length limits, due to regulations or gears, restrict the size range captured by the fishery. The sample of our original population shown in Appendix Figures A1e and A1f is from a fishery in which fish smaller than 350 mm or larger than 450 mm are not captured. Estimation of VB parameters and lengths at age under the assumption of normally distributed errors (green) again reveals biases for partially or non-selected ages (0-2 and 4-7) that increase as length at age estimates move away from minimum and maximum length limits.

When we refit the VB model (Appendix Figure A1f) with a truncated normal error structure (black), truncating lengths less than 350 mm and greater than 450 mm, bias is reduced in partially or non-selected ages, and VB parameters better approximate those observed for the population. We do note that while this method reduces bias relative to least squares regression, there is a tendency to underestimate lengths at partially or non-selected ages, both above and below maximum and minimum length limits, respectively. For further detail on these methods, please see McGarvey and Fowler (2002) and Schueller et al. (2014).

VITA

Michael A. Schmidtke

Old Dominion University Department of Ocean, Earth & Atmospheric Sciences

406 Oceanography & Physical Sciences Building

Norfolk, VA 23529

Michael Schmidtke, originally from St. Louis, MO, graduated from North Carolina State University in 2011 with a B.S. in Zoology. In 2011, he joined the Center for Quantitative Fisheries Ecology at Old Dominion University as a Master's candidate and research assistant under the advisement of Dr. Cynthia M. Jones. In 2014, he graduated with an M.S. in Ocean and Earth Sciences and continued his education at Old Dominion as a doctoral candidate in the Department of Ocean, Earth, and Atmospheric Sciences. Michael is currently employed by the Atlantic States Marine Fisheries Commission as a Fishery Management Plan Coordinator, and continues to contribute to blueline tilefish management through work with the Mid-Atlantic Fishery Management Council.